

Length-Based Methods in Fisheries Research

Edited by
D. Pauly
G.R. Morgan

International Center for Living Aquatic Resources Management
Kuwait Institute for Scientific Research



in May '88
M

Length-Based Methods in Fisheries Research

Edited by

D. Pauly
G.R. Morgan

1987

International Center for Living Aquatic Resources Management
Kuwait Institute for Scientific Research

SH

207

CPL

#13

c. 1

MAY 10 1988

**Length-Based Methods
in Fisheries Research**

Proceedings of the International Conference
on the Theory and Application of Length-
Based Methods for Stock Assessment,
11-16 February 1985, Mazzara del
Vallo, Sicily, Italy

EDITED BY

D. PAULY

G.R. MORGAN

ICLARM Conference Proceedings 13.

1987

Printed in Manila, Philippines

Published by the International Center for Living
Aquatic Resources Management, MC P.O. Box 1501,
Makati, Metro Manila, Philippines, and Kuwait
Institute for Scientific Research, P.O. Box 24885,
Safat, Kuwait.

**Pauly, D. and G.R. Morgan, editors. 1987. Length-based
methods in fisheries research. ICLARM Conference
Proceedings 13, 468 p. International Center for Living
Aquatic Resources Management, Manila, Philippines,
and Kuwait Institute for Scientific Research, Safat,
Kuwait.**

**ISSN 0115-4389
ISBN 971-1022-28-1**

ICLARM Contribution No. 325.

CONTENTS

Preface • <i>The Editors</i>	vii
------------------------------------	-----

PART I: METHODS AND ANALYSES

Methods of Size-Frequency Analysis and Their Incorporation in Programs for Fish Stock Assessment in Developing Countries: FAO Interest in Receiving Advice • <i>J. Csirke, J.F. Caddy and S. Garcia</i>	1
A Review of the ELEFAN System for Analysis of Length- Frequency Data in Fish and Aquatic Invertebrates • <i>D. Pauly</i>	7
Estimating Growth and Mortality Parameters by Nonlinear Regression Using Average Size in Catches • <i>T.A. Ebert</i>	35
Some Modifications of Ebert's Method to Estimate Growth and Mortality Parameters from Average Lengths in a Population • <i>U. Damm</i>	45
Estimating Growth and Mortality in Steady-State Fish Stocks from Length-Frequency Data • <i>J.A. Wetherall, J.J. Polovina and S. Ralston</i>	53
A Method for the Estimation of Growth, Mortality and Gear Selection/ Recruitment Parameters from Length-Frequency Samples Weighted by Catch per Effort • <i>P. Sparre</i>	75
Two Methods for Simultaneously Estimating Growth, Mortality and Cohort Size Parameters from Time Series of Catch-at-Length Data from Research Vessel Surveys • <i>J.G. Pope</i>	103
A Weakly Parametric Method for Estimating Growth Parameters from Length Composition Data • <i>J.G. Shepherd</i>	113
Estimation of Growth and Mortality Parameters for Use in Length- Structured Stock Production Models • <i>J.M. Hoenig</i>	121
Estimating Mortalities from Length- or Age-Specific Sex Ratios • <i>M.J. Sanders</i>	129
Incorporating Age Data Into Length-Based Stock Assessment Methods • <i>G.R. Morgan</i>	137

Assessment of Short-Lived Stocks with Special Reference to Kuwait's Shrimp Fisheries: A Contrast of the Results Obtained from Traditional and Recent Size-Based Techniques • <i>C.P. Mathews, M. Al-Hossaini, A.R. Abdul Ghaffar and M. Al-Shoushani</i>	147
Towards a Method for Short-Term Forecasting of Catch Rates Based on Length Compositions • <i>J.G. Shepherd</i>	167
Phalanx Analysis: An Extension of Jones' Length Cohort Analysis to Multispecies Cohort Analysis • <i>J.G. Pope and Yang Jiming</i>	177
A Simulation Model for Generating Catch Length-Frequency Data • <i>J. Hampton and J. Majkowski</i>	193
An Examination of the Reliability of the ELEFAN Computer Programs for Length-Based Stock Assessment • <i>J. Hampton and J. Majkowski</i>	203
An Investigation of Length Composition Analysis Using Simulated Length Compositions • <i>R. Jones</i>	217
Analytical Investigations of Errors in Mortality Rates Estimated from Length Distributions of Catches • <i>A. Laurec and B. Mesnil</i>	239
Monte-Carlo Testing of Two Methods for Estimating Growth from Length-Frequency Data with General Conditions for Their Applicability • <i>A.A. Rosenberg and J.R. Beddington</i>	283
A Simple Method for Estimating the von Bertalanffy Growth Constants for Determining Length from Age and Age from Length • <i>M.J. Sanders</i>	299
Comments on Age-Length vs. Length-Age Relationships • <i>A.A. Rosenberg and J.G. Pope</i>	305
Use of ELEFAN I for Sampling Design • <i>D. Levi, M.G. Andreoli and L. Cannizzaro</i>	311
Fisheries Management in a Developing Country: the Most Appropriate Balance of Size- and Age-Related Methods for Practical Assessments • <i>C.P. Mathews</i>	321

PART II: IMPLEMENTATION OF METHODS

Length-Based Methods in Fisheries Research: from Theory to Application • <i>J.A. Gulland</i>	335
Data Acquisition for Length-Based Stock Assessment: Report of Working Group I • <i>J.M. Hoenig, J. Csirke, M.J. Sanders, A. Abella, M.G. Andreoli, D. Levi, S. Ragonese, M. Al-Shoushani and M.M. El-Musa</i>	343
Methods of Analysis and Assessment: Report of Working Group II • <i>J.G. Shepherd, G.R. Morgan, J.A. Gulland and C.P. Mathews</i>	353

Sensitivity of Length-Based Methods for Stock Assessment: Report of Working Group III • <i>J. Majkowski, J. Hampton, R. Jones, A. Laurec and A.A. Rosenberg</i>	363
---	-----

Analysis of Length-Frequency Data: Some Available Programs and User's Instructions • <i>G.R. Morgan and D. Pauly</i>	373
--	-----

Indexes*

Author Index	463
Geographic Index	466
Species Index	467

*By V.C. Sambilay, Jr., College of Fisheries, University of the Philippines in the Visayas, Diliman, Quezon City, Philippines.

Preface

The idea of a conference on length-based stock assessment methods evolved through the increased use and development of the methods by several groups in various parts of the world over the last few years. In many instances this renewed attention given to length-based methodology was a direct result of either the difficulty or inability to age fish satisfactorily (a particular problem in tropical areas) or the seeking of a more cost-efficient method of sampling fish populations for assessment purposes. In other cases, the interest was generated by the rediscovery of a whole new set of stock assessment tools which had been generally underutilized or ignored (apart from the pioneering work of Rodney Jones in Aberdeen working with crustacean stocks) for the past thirty years. The opportunities for further refinement and development of the methodology were, and still are, significant particularly with the proliferation of computers in fisheries laboratories around the world.

The theoretical model development and the problems of practical application to fish stocks came together initially through informal contacts and culminated in a conference on The Theory and Application of Length-Based Stock Assessment Methods which took place on 11-16 February 1985 at the Istituto di Tecnologia della Pesca e del Pescato (ITPP) in Mazara del Vallo, Sicily, Italy. These proceedings are a report of that conference.

Specific aims of the conference were:

- to identify the specific features of sampling schemes used to obtain length-frequency data for stock assessment;
- to review length-based methodologies for stock assessment of fish and aquatic invertebrates, with particular reference to their precision and accuracy;
- to review and test computer programs implementing length-based methods.

The conference was jointly organized by the International Center for Living Aquatic Resources Management (ICLARM) and the Kuwait Institute for Scientific Research (KISR). The Food and Agriculture Organization of the United Nations (FAO) supported the attendance of several participants, while the Italian research organization, Consiglio Nazionale delle Ricerche, supported the conference through ITPP, the host institution.

These proceedings include the twenty-two papers which were originally presented in addition to the reports of three of the four working groups formed during the conference. Also included are three contributions written immediately thereafter—the chairman's overview by J.A. Gulland, a review of programs discussed at the conference, compiled by the editors of this volume, and a note on some aspects of the problem of length-to-age *vs.* age-to-length conversion, proposed by two workshop participants.

This conference was the first ever to be devoted solely to length-based methods for stock assessment. In existence since C.G.J. Petersen's pioneering papers of 1891/92, length-based methods were masterly used by T.W. Fulton, F.I. Baranov and others. Then, a period followed where length-based methodology was inappropriately used and which led to its disappearance from major fisheries textbooks and to a gradual decline in its use, a trend which lasted until the mid-1950s, when R.J.H. Beverton and S.J. Holt proposed their now classic equation for estimation of total mortality from the mean length of fish in catch samples.

Further development of the methodology was slow and it was not until scientists working in various tropical areas of the world (where routine aging of fish presents significant problems) began re-examining those earlier methods that the momentum of development and use of these length-based techniques increased. The work of J.L. Munro in Jamaica may be mentioned in this context along with work done in India and by staff of the French agency ORSTOM in West Africa.

The present volume is based on and may be seen as representing the culmination of these earlier efforts and brings together the most recent and important theoretical model developments with examples of their applications. Listings and other details of computer programs are also given to allow the reader to use, adapt or further develop the models.

It was initially planned to present these proceedings in two volumes; the first was to include theoretical considerations, descriptions of methods and analyses, while the second was to consist of contributions related to the implementation of methods, program listings and user's instructions. As it turned out, not enough material was assembled for a second volume to be issued separately. The volume presented here has two parts, however, and they contain what we had originally earmarked for the two separate volumes.

We hope that in a few years, another group will assemble, and critically examine, as we have done in Sicily, the wide range of length-based methods being used by workers in the field, as well as new developments.

Finally, we would like to thank all those at ICLARM, KISR, FAO and ITPP, in addition to all the participants who made the conference the pleasant experience that it was.

D. PAULY
G.R. MORGAN

**Methods of Size-Frequency Analysis and Their Incorporation in
Programs for Fish Stock Assessment in Developing Countries:
FAO Interest in Receiving Advice**

J. CSIRKE
J.F. CADDY
S. GARCIA

*Fisheries Department
Food and Agriculture Organization of the United Nations
Via delle Terme di Caracalla
00100 Rome, Italy*

Csirke, J., J.F. Caddy and S. Garcia. 1987. Methods of size-frequency analysis and their incorporation in programs for fish stock assessment in developing countries: FAO interest in receiving advice, p. 1-6. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

FAO's task requires the adaptation of conventional methodologies developed in high latitudes for use in tropical situations and the development and further refinement of length-based methodologies. However, a number of problems need to be resolved before these methods can be applied on a grand scale. Questions for which FAO requires detailed answers relate among other things to the following items: "quick and dirty" vs. elaborate approaches, problems with cohort separation, bias due to selectivity and heterogeneous fishing effort, departures from steady state, sampling problems in multispecies, multigear fisheries, problems with very short-lived animals, appropriate trawl survey stratification for sampling length-frequency data, effects of sampling errors and biases on mortality and growth estimates from different methods, use of mean size statistics and translation of results from length-based models into management advice.

Introduction

Many questions have been raised over the last few years concerning the methodologies appropriate to stock assessment in developing countries, and in particular, the extent to which conventional methodologies developed in high latitudes may be adapted for use in the tropics. Along with other organizations, FAO has advocated the use of length-based methodologies (by publishing relevant manuals, see, e.g., Pauly 1983 or Jones 1984) as one subset of the range of methodologies available for fisheries research. We recognize that a precise evaluation of the strengths and weaknesses of the various techniques of length-frequency analysis continually arising in the literature is still lacking, and that this question must be a focal point for the present conference. One further class of questions should not be neglected pertaining to the ranking of the various methodologies in

terms of their content and use, and to the role they might play in an integrated program of stock assessment for developing countries.

A brief review of what, in our opinion, are the more important questions is given here with brief annotations. These are respectfully addressed to the participants of this conference for their expert consideration.

Short-Cut vs. Long-Term Advice

In some circumstances it may be useful to have available a methodology which, while not pretending to great accuracy, allows a "first cut" at estimating the parameters of, e.g., growth and mortality for the population in question.

Circumstances where short-cut methodologies might apply are, for example, in the preliminary stages of development of an assessment program, or for a resource on which investigation is just beginning.

Questions

Is there some validity in dividing length-frequency methods into "short-cut" or "quick and dirty" methods, and those more elaborate approaches appropriate to a long-term investigation?

What are the appropriate (length-based and other) methodologies which should be recommended to fisheries departments of developing countries for use in long-term programs of fish stock assessment in their national waters?

Length-Based vs. Age-Based Methods

COHORT SEPARATION

The use of age-based methods and the preparation of age-length keys allow estimation of mean length at age and of its variance for each age group separately.

These values can also be estimated by sequential modal separation for younger age groups provided mean length of different age groups are well separated and the variability in mean length is relatively small.

Questions

What sources of variability of the mean length-at-age may one expect in nature (e.g., for younger and older fish, or long- and short-lived species)? and how does variance change in age? (important for stochastic age-length simulation). What are the consequences for sequential modal separation?

BIAS DUE TO SELECTIVITY AND FISHING EFFORT

Growth studies of fish in the Northeast Atlantic were first based on length-frequency analysis, but this approach was abandoned when aging techniques using hard tissues (e.g., scales, otoliths) were developed and proved to be feasible and more precise (although more work intensive and costly). If proper sampling and subsequent age reading is possible, growth rate determination poses few practical problems. The main possible source of bias (as for length-based methods), is introduced by the selective capture of the larger sizes of younger age groups (due to gear selectivity) and the

removal of the older age groups from the larger sizes (due to fishing mortality), which will reduce the average age corresponding to a given length. These problems can be overcome by taking a random subsample for aging and construction of an age-length key. However, selectivity bias presents a serious problem for the use of length-based methods for estimating growth. Also, fishing gear commonly used in artisanal fisheries (hook and line, traps, gill nets) may not have logistic selection curves as occurs, e.g., in trawls.

Question

What are the effects of gear selectivity and increased fishing mortality on the parameters and apparent shape of the von Bertalanffy growth equation estimated through sequential modal analysis?

DETECTING DEPARTURES FROM STEADY STATE

If there are no crucial sampling problems, and steady-state conditions are satisfied (or are approximated by integrating over an appropriate period), most of the well-elaborated (theoretically precise) length-based methods should give results that are consistent with those that might be obtained through conventional age-based approaches. The problems start with departures from steady-state conditions. Age-based Virtual Population Analysis (VPA) and cohort analysis have demonstrated the interannual variability of year-class strength and of F for fish stocks of temperate (and some tropical) waters where aging has been possible. In fact, recent research has picked up recruitment variation and its causes as a key area of concern. For instance, there might be no problem in using Jones' length cohort analysis (for example, with shrimp) if only one cohort occurs in the data, but problems arise when two or more cohorts are represented (e.g., two recruitment pulses). Variations in recruitment strength will then seriously bias F vectors when these are expressed as a function of length.

Questions

How can changes in recruitment, availability and growth, for example, be detected from length-frequency samples alone? If these changes cannot be detected, what are the types and the magnitude of the errors that might be introduced into estimates of mortality obtained by length-converted catch curves or other length-based methods? Could a combination of length and, for example, a limited aging program be useful in coping with this problem? How intense should this combined sampling program be? What are the effects of changes in growth rate on the estimated mortality rates? Can these effects be detected?

Robustness of Length-Based Methods

Since all length-based methods assume some degree of stability or uniformity over time and ages, there are serious concerns about the precision of the results obtained when there are departures from equilibrium conditions or discontinuities in the effects on the population. In general, variability in recruitment, growth rate, differential fishing mortality of "fully recruited" fish (such as those caused by bell-shaped selection curves), year-to-year changes in fishing effort, etc. are all sources of bias that may introduce errors into the estimates of Z. Since almost every fishery tends to show a certain degree of variability, it would be worth testing the robustness of different length-based methods to different types of departures from the ideal equilibrium conditions before encouraging their extended application as routine methods of analysis.

Question

How robust are the length-based methods in use to departures from the ideal conditions of equilibrium with time and uniformity of certain parameters with age?

Sampling Problems

MULTISPECIES, MULTIGEAR FISHERIES

Length-based methods were offered as a replacement for catch-effort methods in multispecies, multigear fisheries because of the difficulty of defining effort in these cases. This benefit seems illusory since the same sampling problem occurs in both cases.

Questions

How does one correctly sample the catch or the population so as to obtain length-frequency distributions which represent either the size structure of the catch, or of the population? How is one to evaluate the degree of confidence to be given to the results of length-frequency analysis?

SHORT-LIVED ANIMALS

The use of length-based methods in short-lived animals (squid, shrimp) essentially involves a "dynamic nonequilibrium" situation. The annual catch size structure is "virtual", since the size composition changes. Also, there is a need for adequate sampling in both space and time, since biomass as well as numbers change greatly on a seasonal basis.

Question

What is the optimum weighting of the samples in time?

TRAWL SURVEY

Little use has been made of length-frequency data from trawl surveys, and their use should be encouraged. However, it should be taken into account that sizes are usually not distributed at random in the population. Also the fact that different species of a multispecies resource may have different distribution patterns should be considered.

Questions

What are the characteristics and uses of length-frequency data from trawl surveys?
How can the appropriate sampling schemes (e.g., systematic, nonrandom) that might be needed to account for nonrandom distribution of sizes, and the appropriate (post?) stratification scheme to account for multispecies resources, be accommodated while ensuring that the other important objective of trawl survey (i.e., estimating total biomass) is satisfied?

SAMPLING ERRORS AND BIAS

Length-frequency data are subject to sampling errors and bias due to the particular properties of the sampled population and due to the sampling procedures and the characteristics of the fishing (sampling) gear used.

Question

What are the validity and typical errors and biases in the estimates of mortality and growth parameters using various length-based methods?

Using Commercial Size Categories

One of the most widely available sets of data in tropical areas is the commercial size category: fuller use of this data source is clearly desirable. It has been noted that converting from, e.g., mean weight to mean length and then to mean age involves reversing the independent (age) and dependent (length) variables when fitting a conventional growth curve.

Question

What are the appropriate procedures for using commercial size categories and the likely sources of error?

Mean Size Statistics vs. Length-Frequency

Under certain circumstances it may be quicker and easier to obtain mean size data (e.g., as the reciprocal of the number of individuals per container holding a given weight) than a size-frequency distribution. In fact, commercial size categories may only allow computation of a mean size.

Question

Given an independent estimate of size-at-age, what are the relative advantages and drawbacks of estimates (of, e.g., mortality) from mean size statistics in the exploited population as opposed to those based on length-frequency analysis?

Integration With Other Approaches to Stock Assessment

Any method of stock assessment begins with data collection, leading through analysis, to conclusions on the state of the resource. Collecting length-frequency data on research or commercial vessels, in port, or in the market, needs to be integrated with other data collecting needs, methods of analysis and management conclusions. This is because in any science it is usual to check the validity of a hypothesis by using more than one independent set of data to arrive at two or more independent conclusions that can be compared. For developing countries, however, the collection of a large number of independent data sets (e.g., on catch, landings, catch rate, sizes, etc.) may be prohibitive in cost and manpower. What priorities should be given to length-based methods in these circumstances, and what other components will be needed for a stock assessment program?

Question

How should length-based methods be integrated with other approaches to stock assessment?

Management Advice

Independently of the method of analysis used in a stock assessment process, management advice will need to be given in terms that are understandable and usable by fisheries administrators. This may require the use of a yield model and the presentation of the final results in terms of an allowable catch (e.g., in terms of MSY, MEY, etc.), or its corresponding level of fishing effort (e.g., f_{MSY} , f_{MEY} , $f_{0.1}$, etc.).

Questions

How should parameter values resulting from length-frequency analysis be translated into management advice? Which models does it seem most appropriate to apply for length-based assessments? What possible errors, biases and qualifications should be taken into account in the application of these approaches?

Discussion

We believe that the answers to the questions posed above, both during the conference and in the follow-up work that will inevitably follow will greatly contribute to clarification of the validity and scope of length-based methodologies, possibly by defining more clearly the limits of their application. We hope that the answers obtained will confirm our conviction that a more general and appropriate use should be made of length-frequency data, especially but not only in the tropics. These methodologies, particularly the short-cut methodologies are very suitable for use in and by developing countries in the present situation, and are very useful in promoting more work and involvement in stock assessment and if necessary, the transition to methods more appropriate to long-term use.

High latitude fishery science has developed up to now a high level of sophistication, (physical inputs and trained manpower) that will not be available in the tropics for a few decades at least. In the meantime, scientists from developing countries, after obtaining their degree, are often not able to progress towards a global understanding of stock assessment and the provision of biological advice for national resource management since the methodologies proposed have until quite recently been largely nonapplicable, and sometimes irrelevant.

By reducing the demand for more and more detailed data, and therefore by reducing costs, the length-based theory offers them a possibility of starting work and learning. The price to pay in exchange for this is to accept some risk and to buffer the risk by cautious management advice. However, the potential gain in the long term is enormous, as this process might finally start the development of tropical expertise, "locally grounded", an essential condition for appropriate feedback and development of a more relevant tropical fisheries science.

References

- ✓Pauly, D. 1983. Some simple methods for the assessment of tropical fish stocks. FAO Fish. Tech. Pap. No. 234, Rome, 52 p. (Also available in French and Spanish)
- Jones, R. 1984. Assessing the effects of changes in exploitation pattern using length composition data. FAO Fish. Tech. Pap. No. 256, Rome, 118 p.

A Review of the ELEFAN System for Analysis of Length-Frequency Data in Fish and Aquatic Invertebrates*

D. PAULY

*International Center for Living Aquatic Resources Management
MC P.O. Box 1501, Makati, Metro Manila
Philippines*

Pauly, D. 1987. A review of the ELEFAN system for analysis of length-frequency data in fish and aquatic invertebrates, p. 7-34. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

Some aspects of the history and present role of length-based methods for the study of the growth, mortality and recruitment of fish and invertebrate stocks are reviewed. The recently-developed ELEFAN (Electronic LENGTH Frequency ANalysis) system of BASIC microcomputer programs is presented and some of its key features are discussed in detail.

Examples of the use of the ELEFAN system are provided; these examples are used to illustrate the need for further research on such length-based systems and to emphasize the need for adequate sampling.

Introduction

Fisheries research emerged at the turn of the century, mainly in countries bordering the North Atlantic (predominantly the North Sea area). This had the effect, among other things, that research on cod, herring and similar long-lived fish became paradigmatic for the field as a whole (Cushing 1983; Gulland 1983).

When fisheries research expanded after World War II to the developing countries of the Third World, many of the basic methods of "temperate" fisheries biology were not applicable or did not seem to be applicable to tropical fish (Menon 1953; Qasim 1973).

This is especially true of methods for aging individual fish, which, being based predominantly on "annuli" on scales, otoliths and other bony structures of fish, proved difficult to apply to tropical fish in which seasonal fluctuations of their environment are usually less marked than in temperate waters.

*ICLARM Contribution No. 232.

Fisheries biologists working on tropical fish prior to Pannella's (1971) discovery of daily rings as an ubiquitous feature of tropical fish had to rely predominantly on length-frequency analyses to draw inferences on the growth of fish. Methods using length-frequency data to infer the growth patterns of fish can be traced to the work of Petersen (1891), who developed two basic techniques.

The first of these, the "Petersen method" (*sensu stricto*), consists of the analysis of *one sample at a time*, with different (relative) ages being attributed to the fish comprising the prominent modes of a sample (Fig. 1a).

Graphical, semi-graphical (Harding 1949; Cassie 1954; Tanaka 1956; Bhattacharya 1967) and computer-based methods (Abramson 1971; Yong and Skillman 1971) for the separation of single samples into size groups have been widely used, often with mixed results (Mathews 1974). All of these methods assume the component of the distributions investigated to be normally distributed, an assumption which has been shown not to apply in a number of cases where it has been put to test (see, e.g., Morgan, Part I, this vol.). The separation of length groups (even when computer-aided) and the attribution of ages to these length groups are two distinct steps; the latter is purely arbitrary when no additional information on recruitment periodicity is available and can be very misleading when two recruitment pulses occur per year as is common in the tropics (Pauly and Navaluna 1983), while only one is assumed or *vice versa*.

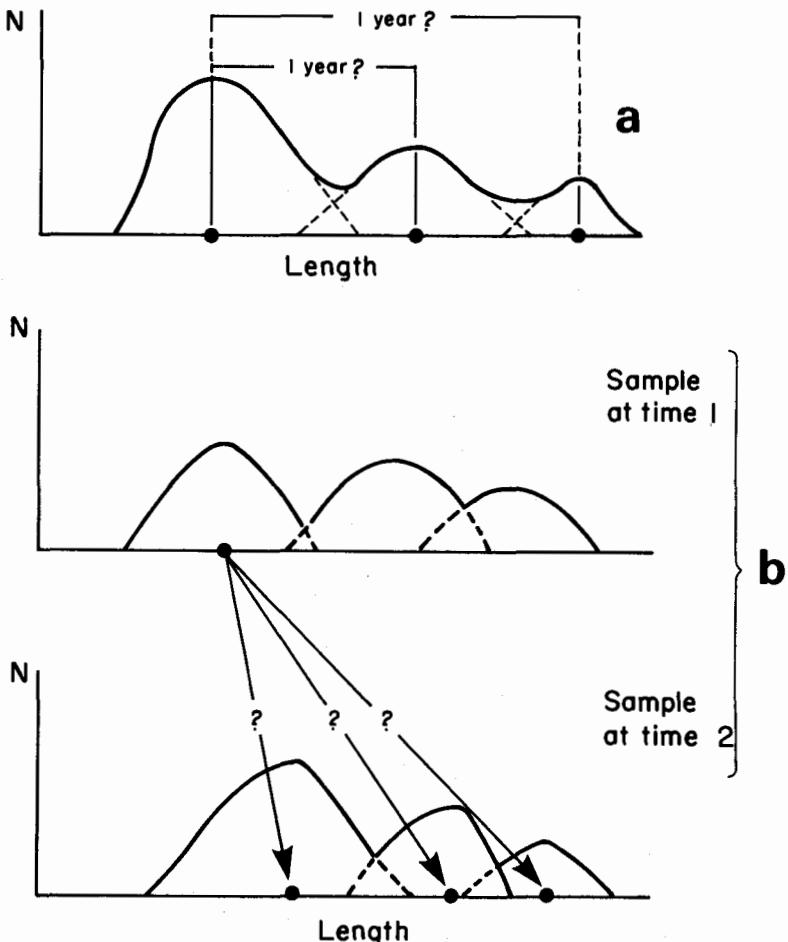


Fig. 1. Basic principles involved in traditional methods for length-frequency analysis (from Pauly et al. 1984). a) Application of the Petersen Method (*sensu stricto*) to a length-frequency sample. Note that the time separating peaks must be assumed, a difficult task in animals which may or may not spawn several times a year. b) Application of modal class progression analysis to a set of two samples obtained at known times. Note that the problem here is the proper identification of peaks to be interconnected, not that of time separating modes (as was the case in a).

The second of the techniques pioneered by Petersen (1891), now called "modal class progression analysis", involves the use of several samples collected at different times, whose modes or the means of their component distributions are linked in a fashion thought to represent growth (Fig. 1b).

Various approaches exist for combining the two techniques. One, called the "integrated method" was proposed by Pauly (1978, 1980a, 1983a and see Fig. 2).

The Petersen method *sensu stricto* and the modal class progression analysis were used rather extensively in the beginning of the century in the North Atlantic area, notably by Fulton (1904) and other authors cited in Thompson (1942). Indeed, the latter was so convinced of the superiority of length-frequency analyses in growth studies that he rejected the results of Lea (e.g., Lea 1911,

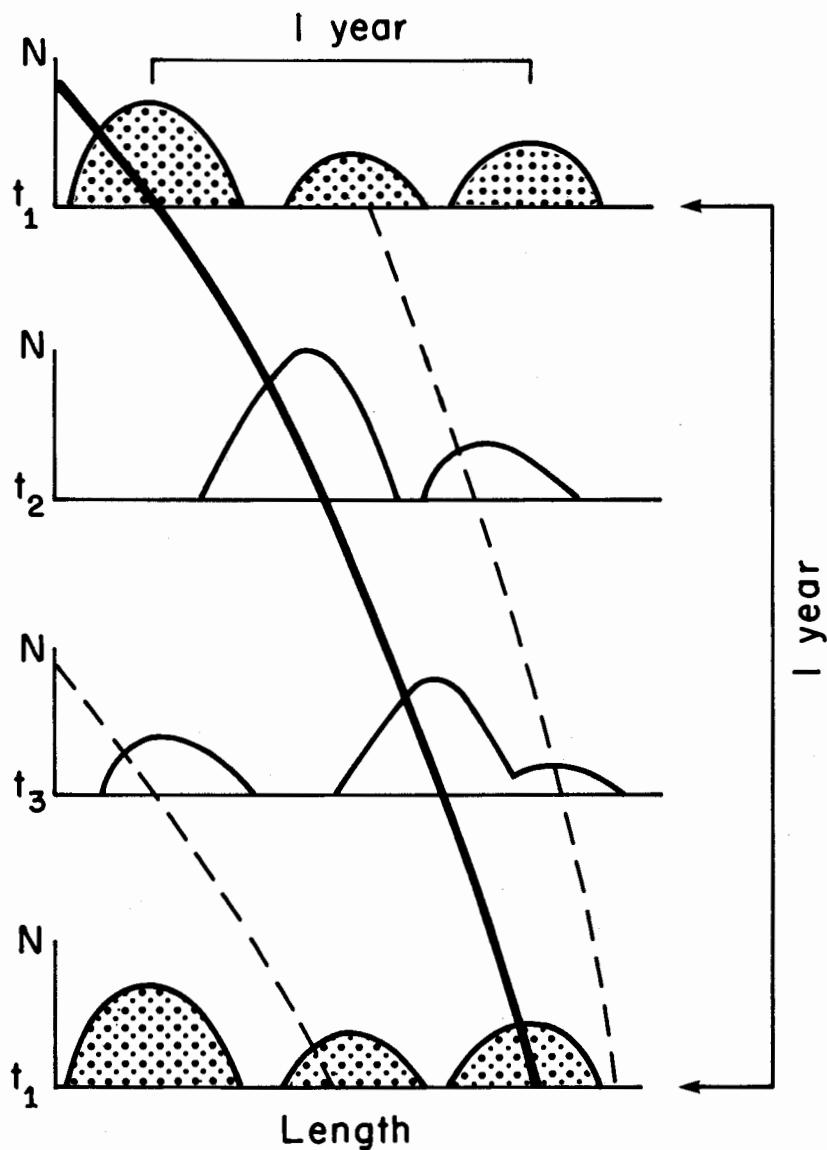


Fig. 2. An application of the 'integrated method' to a hypothetical set of length-frequency samples. Note that the attribution of a relative age to the third peak of sample t_1 is confirmed by the modal class progression, which suggests a growth curve passing through the major peaks of samples t_1 , t_2 , t_3 and through the third peak of sample t_1 repeated after one year (i.e., placed at the appropriate place on the time scale, after sample t_3). Thus a smooth growth curve can be traced which explains most of the peaks of a set of length-frequency samples, including those of earlier samples repeated once, twice or more along the time axis. A certain degree of reliability is achieved which could not be achieved by applying separately either of the two earlier methods (from Pauly et al. 1984).

1913) who, based on scale studies, was able to elucidate major aspects of the life history of North Sea herring. The conflict between D'Arcy Thompson (who was quite wrong) and Einer Lea delayed major advances in European fisheries research for at least a decade (Went 1972) and may be one reason why many contemporary fisheries biologists working in temperate waters are now reluctant to rely predominantly on length-frequency data when conducting growth studies.

The pendulum may have swung back too far, considering that most growth studies published in major journals now omit the *validation* of aging using independent methods or cross-validations (Beamish and McFarlane 1983).

While well-conducted growth studies based on otoliths, scales or other hard parts should generally lead to more precise inferences on growth than studies based on length frequencies only, length-based methods will remain extremely important, in the tropics at least, for the following reasons:

- (i) a number of important aquatic organisms, such as shrimp, cannot be aged individually, leaving one no choice but the use of length or of the more expensive tagging-recapture data;
- (ii) length-frequency sampling in fish markets is generally far cheaper than sampling say, for otoliths, because among other things, fish that are only measured need not be purchased; this is a factor limiting the work of fishery biologists in Southeast Asia; the phenomenon is also reported from Kuwait by Morgan (1983) and Mathews (Part I, this vol.);
- (iii) estimating growth from length-frequency data is far less work-intensive than using hard parts; when the resource under study consists of small, short-lived fish that can be aged only by counting daily rings, routine assessments based exclusively on age-structured models seem quite impractical, especially in a multispecies situation;
- (iv) large amounts of unanalyzed length-frequency data are available in most fisheries laboratories and often they are the only historical data available. Thus, for example, Ingles and Pauly (1982, 1984) were able to retrieve, from the files of various research institutions in the Philippines, length-frequency data covering more than one million measurements, from which inferences were drawn on the growth, mortality and recruitment of 56 species of Philippine fishes. Colleagues from various countries have had similar experiences and suggested that length-frequency data are probably the most underutilized type of information available on fish and other aquatic resources;
- (v) the "ages" used in many models are contrived because the process that is modelled is actually *size*-related. A number of biological and fishery-related processes involve sizes rather than ages such as predation (Ursin 1973), food conversion efficiency (Pauly 1981), gear (mesh) selection, target strength (in acoustic surveys), recruitment to fishing grounds and, quite importantly, marketability and price.

The ELEFAN System

INTRODUCTION TO THE SYSTEM

The ELEFAN system (*Electronic LENGTH Frequency ANalysis*) was developed at ICLARM in response to (1) the need for robust methods for analysis of length-frequency data; (2) the availability of cheap microcomputers. These two points provide some of the reasons why the system has found wide acceptance in developing countries, as was intended.

The system, as it now stands, consists of five programs, ELEFAN 0, I, II, III and IV.

ELEFAN 0 is used to create and modify length-frequency data files for use with the other four ELEFAN programs; the other ELEFAN programs have length-frequency data created by ELEFAN 0 as their main input (Fig. 3).

ELEFAN I is used to estimate the growth parameters of fish or invertebrates. The growth equation of which these parameters are estimated is a seasonally oscillating version of the von Bertalanffy Growth Formula (VBGF). ELEFAN I can thus be used to provide quantitative information on growth oscillations of fish and invertebrates, which can be correlated with oscillations of selected environmental parameters.

ELEFAN II performs a variety of computations, of which the following are the main ones:

- estimation of total mortality (Z) and derived quantities from the straight, descending arm of a length-converted catch curve;
- estimation of probabilities of capture by length and mean length at first capture (L_c) from the ascending, left arm of a length-converted catch curve;
- expression of the seasonal changes in recruitment intensity in the form of a graphical “recruitment pattern” (which can be further subdivided into normally distributed recruitment pulses, suggestive of the number of spawning and/or recruitment seasons per year).

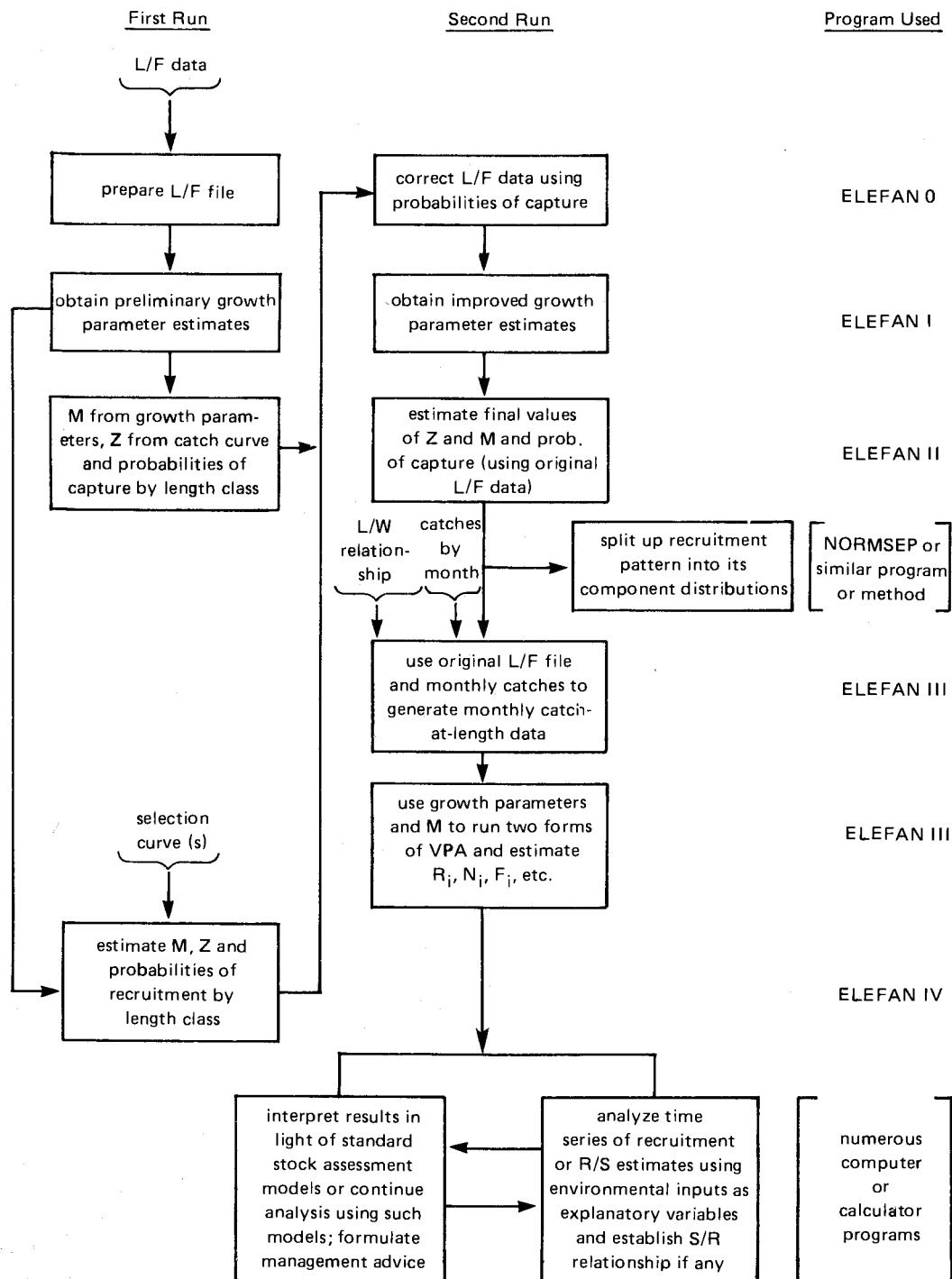


Fig. 3. Flowchart of the ELEFAN System, showing inputs, outputs and their subsequent analysis where appropriate. Definitions of parameters (Z , M , F , R , etc.) are given in the text, along with further details on system.

ELEFAN III incorporates three types of virtual population analysis (VPA), coded here VPA I, II and III.

VPA I estimates standing stock (in numbers) and fishing mortalities by time intervals (month, quarter, year, etc.). VPA II is used to estimate standing stock (in numbers) and fishing mortality by length class in a stock with stable age distribution, as can be simulated by combining data for several years. VPA III provides estimates of standing stock and fishing mortality by month and by length, which is achieved by "slicing" (pseudo-) cohorts through the catch-at-length data by means of a set of growth parameters. This approach assumes that little exchange occurs between the monthly "cohorts", which applies mainly in short-lived animals, such as anchovies or penaeid shrimps, for which the VPA III routine has been specifically designed.

ELEFAN IV is a program which, provided that gear selection is known (i.e., that probabilities of capture by length class are available), can be used to estimate M and probabilities of recruitment by length class from catch samples representative of an exploited population.

Table 1 gives a summary of the necessary inputs and of the output obtained from the ELEFAN programs. Further details are given below. As might be seen from Fig. 3, the overall system depends

Table 1. Inputs and outputs of the ELEFAN system for the analysis of length-frequency (L/F) and catch-at-length (C/L) data.

Input/output	ELEFAN program				IV
	0	I	II	III	
necessary inputs	L/F	L/F	L/F, growth parameters (excluding t_0)	L/F, monthly bulk catch, M, growth parameters (excluding t_0), length-weight relationship	L/F, selection curve, growth parameters (excluding t_0)
facultative inputs	selection curve	—	t_0 (a growth parameter)	C/L data computed externally can also be entered and used	—
output, given necessary inputs	files for use with ELEFAN I to IV	growth parameters (including seasonal growth)	Total, natural ^a and fishing mortalities, approximate selection curve, mean length at first capture, seasonal pattern of recruitment into the fishery.	Fishing mortality and population size by month and length group.	Total, natural ^a and fishing mortalities, recruitment curve.
output, given facultative inputs	L/F data corrected for selection effects (for use with ELEFAN I only)	—	seasonal pattern of recruitment with absolute time scale		

^aELEFAN II and IV use different methods to estimate natural mortality; with ELEFAN II, M is estimated from equation (6), which is built into the program; with ELEFAN IV, M is estimated from the left, ascending side of a length-converted catch curve and the selection curve following a method devised by Munro (1984, and see text).

critically on the availability of growth parameters (mainly L_∞ and K of the VBGF) or of length-frequency data from which these parameters can be estimated. The length-frequency type of data, on the other hand, while having to be representative of the stock from which they were sampled need not have been corrected for the effects of gear selection or incomplete recruitment. This is so because it is possible to first obtain preliminary growth parameter estimates from length-frequency data not corrected for selection effects; then the available L/F data are corrected for such effects using ELEFAN 0; then growth parameters are re-estimated from the corrected L/F data (see below for an application example).

The Conceptual Basis of ELEFAN I

INTRODUCTION TO ELEFAN I

The heart of the ELEFAN system is ELEFAN I, the program for the estimation of growth parameters from length-frequency data. The first version of ELEFAN I, developed in 1978 by the author with the assistance of H. Timmermann of Kiel University, consisted of two basic steps:

- (i) separation of a number of samples into their normally distributed components using the NORMSEP program of Abramson (1971),
- (ii) identification of the growth parameters generating the growth curve which minimized the sum of the squared deviations from the means of the component distributions.

The program "worked". However, it required as input the number of broods (or age groups) in each sample, in other words, it required as *input* the very information which was wanted as an *output* (but see Sparre, Part I, this vol.; Rosenberg and Beddington, Part I, this vol.). Methods do exist to get around this problem (e.g., using the ENORMSEP program of Yong and Skillman (1971) or the method of Bhattacharya (1967) as interpreted by Pauly and Caddy (1985) to provide preliminary estimates of the number of broods in each sample). However, ELEFAN I incorporates an algorithm which bypasses this step, and which fits the growth curve to peaks defined independently of any assumed underlying distribution.

PRINCIPLES OF FITTING GROWTH CURVES

The method used in ELEFAN I to identify peaks and troughs in length-frequency histograms is a simple highpass filter, i.e., a running average which leads to definition of peaks as those parts of a length-frequency distribution that are above the corresponding running average and conversely for the troughs separating peaks (Fig. 4). The running average used in ELEFAN I is taken over five classes; analyses of the effect of the number of classes included in the running average on the results of ELEFAN I have been conducted only once, by K. Sainsbury, CSIRO (pers. comm.), who reported that the results are sensitive, but not strongly so, to changes in the number of classes included in the running average.

Fitting of the growth curves is performed by ELEFAN I through the following steps:

- calculate the maximum sum of points "available" in a (set of) length-frequency sample(s) (see Fig. 4c) ["available points" refers here to points which can possibly be "accumulated" by one single growth curve; see below]. This sum is termed "available sum of peaks" (ASP),
- "trace" through the (set of) length-frequency sample(s) sequentially arranged in time, for any arbitrary "seed" input of L_∞ and K , a series of growth curves started from the base of each of the peaks, and projected backward and forward in time to meet all other samples of the sample set and/or the same sample repeated again and again,
- accumulate the "points" obtained by each growth curve when passing through peaks (positive points) or through the troughs separating peaks (negative points) (see Fig. 4b and 4c),
- select the curve which, by passing through most peaks and avoiding most troughs best "explains" the peaks in the (set of) sample(s) and therefore accumulates the largest number of points. This new sum is called "explained sum of peaks" (ESP).
- decrement or increment the "seeded" values of L_∞ and K until the ratio ESP/ASP reaches a maximum, and gives the growth parameters corresponding to this optimum ratio.

THE GROWTH MODEL IN ELEFAN I

Although ELEFAN I could in principle be used to fit any type of growth curve, e.g., those proposed by Gompertz (1825) or Krüger (1964), only one type of growth curve, albeit a very versatile one, was incorporated into ELEFAN I. It is a seasonally oscillating version of the generalized von Bertalanffy Growth Function (VBGF) (*sensu* Pauly 1981) of the form

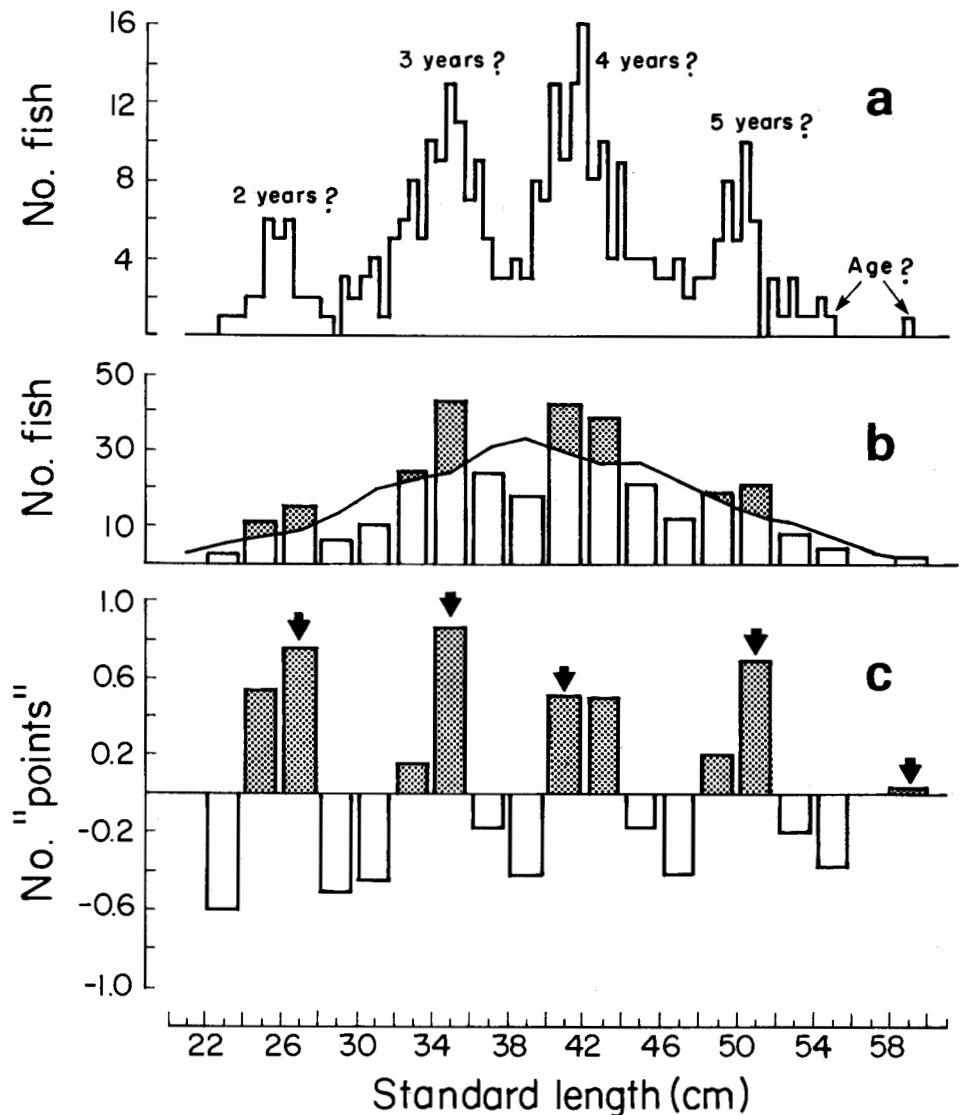


Fig. 4. Length-frequency data on coral trout (*Plectropomus leopardus*) caught near Heron Island (Great Barrier Reef, Australia) in October 1971. From Goeden (1978). a) Original data: the ages from Goeden, with question marks added; N = 319. Note inappropriately small class interval of 5 mm. b) Same data, replotted in 2-cm class intervals to smooth out small irregularities, showing running average frequencies (over 5 length classes) to emphasize peaks (shaded bars above running average) and intervening throughs. c) Same data as in b, after division of each frequency value by the corresponding running average frequency, subtraction of 1 from each of the resulting quotients and subsequent minor adjustments to remove potential sources of bias (see Pauly et al. 1980). Note that well-structured peaks have been allotted similar numbers of points, irrespective of the number of fish they represented. Arrows show the points used in the computation of ASP or "Available Sum of Peaks"; see also text (from Pauly and David 1981).

$$L_t = L_\infty (1 - e^{-KD(t - t_o)} + \frac{CKD}{2\pi} \sin 2\pi(t - t_s))^{1/D} \quad \dots 1)$$

where L_t is the predicted length at age t

L_∞ is the asymptotic length, or the mean length the fish of a given stock would reach if they were to grow forever;

K is a growth constant ("stress factor" in Pauly 1981);

D is another growth constant ("surface factor" in Pauly 1981);

- C is a factor which expresses the amplitude of the growth oscillations (Fig. 5);
- t_o is the "age" the fish would have had at zero length if they had always grown in the manner predicted by the equation;
- t_s sets the beginning of sinusoidal growth oscillation with respect to $t = 0$ (Pauly and Gaschütz 1979).

Within ELEFAN I, equation (1) is used with two of the original parameters replaced by others; thus t_s is replaced by the Winter Point (WP), which designates the period of the year (expressed as a fraction of a year) when growth is slowest. In the northern hemisphere, WP is often found in the neighborhood of 0.2, i.e., in February, while WP often has a value of 0.7 in the southern hemisphere. The Winter Point is related to t_s through

$$t_s + 0.5 = WP \quad \dots 2)$$

The other parameter of equation (1) which requires comment is t_o . The significance of this parameter is often misunderstood. It may be described simply as a factor used to adjust a growth curve to an absolute age scale. Length-frequency data, by themselves, *never* allow the estimation of t_o and ELEFAN I is no exception. Within ELEFAN I a parameter called "T0" is used internally which fulfills the usual role of t_o (that is positioning the growth curve along the time axis). In early versions of ELEFAN I this "T0" was output, such as to enable users to draw their growth curves (e.g., White 1982). Some users have misunderstood this output for a real t_o , however, which resulted in their growth equation generating "erroneous" length-at-age data. To prevent further misunderstandings, therefore, the output of "T0" is now suppressed, and a routine has been built instead

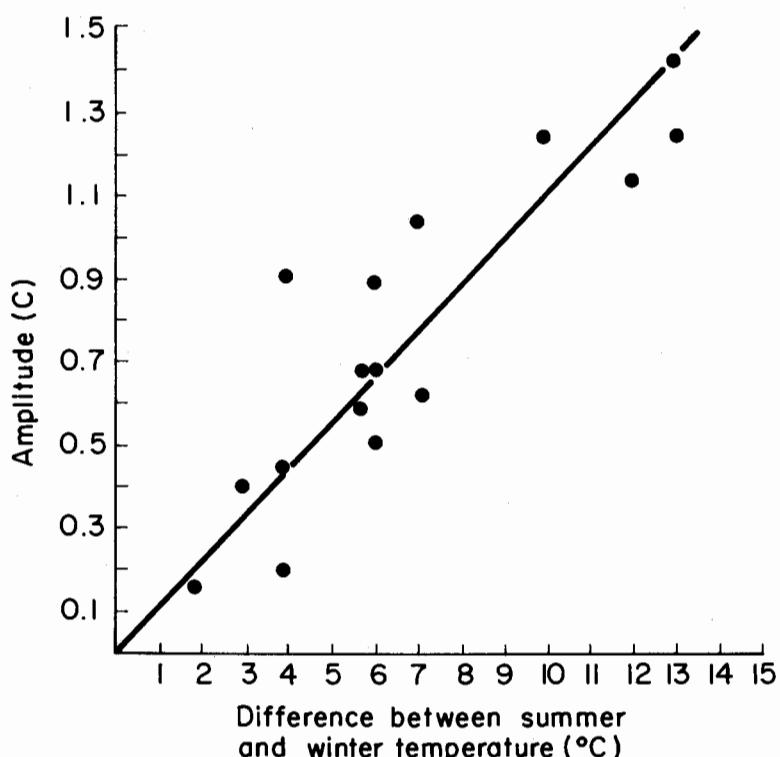


Fig. 5. Relationship between the amplitude (C) of seasonal growth oscillations in fish, penaeid shrimp and squid populations and the summer-winter temperature difference of their habitat (in °C). The data points used here are documented in Pauly and Ingles (1981), Pauly (1982; 1985a), Pauly and Gaschütz (1979) and Pauly et al. (1984). Note that values of $C > 1$ do not imply that the animals in question shrink in the winter time, but rather that they undergo a period of non-growth (see Pauly and Gaschütz 1979).

into ELEFAN I which computes and outputs length in given months, and thus allows the drawing of growth curves.

The parameter C in equation (1) is dimensionless; when it has a value of zero, equation (1) becomes

$$L_t = L_{\infty} (1 - e^{-KD(t - t_0)})^{1/D} \quad \dots 3)$$

that is, the equation takes the form of the generalized VBGF (Pauly 1981) which has only one parameter more (D) than the original equation (von Bertalanffy 1938).

Equation (1) is defined such that growth ceases (at WP) only briefly, once a year, when C = 1. When 0 < C < 1, growth oscillates seasonally, but is never completely halted. When C > 1, growth oscillates so strongly that the model predicts in winter a reduction of length. This latter case, it may be noted, should rarely happen in reality (as opposed to loss of weight). One paper referring to teleosts was thus found in which a significant decrease in length of 0.67 to 1.23% was reported after starvation (in captivity) for up to seven weeks (Nickelson and Larson 1974) which is unlikely to happen in nature. In most cases, therefore, values of C > 1 simply suggest sustained periods of nongrowth, such as may occur in cold freshwater habitats.

Extensive work with ELEFAN I and with other programs incorporating equation (1) has shown that C is generally correlated with the difference between mean monthly summer and winter water temperature, that is, with the range of temperature to which the fish of a given stock are exposed (Fig. 5). This rule may help in providing good "seed values" of C for use in ELEFAN I, although it must be restated here that C as well as WP and the parameter L_{∞} and K of the VBGF are outputs of and not inputs for ELEFAN I.

The parameter D of equation (1) on the other hand probably cannot be estimated from length-frequency data and ELEFAN I alone. When D = 1, equation (3) becomes

$$L_t = L_{\infty} (1 - e^{-K(t - t_0)}) \quad \dots 4)$$

which means that it reduces to the "special" VBGF, which is the normal VBGF used in fisheries research.

When D ≠ 1, equation (3), the generalized VBGF is actually a form of the curve proposed by Richards (1959), used to model the growth of fish and invertebrates by e.g., Mathisen and Olsen (1968) and Ebert (1980, 1981). While these authors estimated the fourth parameter (equivalent to D) of Richards' curve from size-at-age data, Pauly (1981, 1982) showed that D expresses the growth of fish gills as related to their body weight and also that D decreases with the maximum size that they can reach. Rough estimates of D can be obtained therefore, from

$$D = 3 \cdot [1 - (0.6742 + 0.03574 \cdot \log_{10} W_{\max})] \quad \dots 5)$$

where W_{\max} is the maximum weight (in grams) reached by the fish of a given stock. Equation (5) is built into ELEFAN I. However, it should be used, along with values of D ≠ 1 only when the arguments leading to the generalized VBGF are understood and agreed with (see Pauly 1981).

The earlier version of ELEFAN I required estimated asymptotic length to be larger than the largest animal in the set of samples analyzed, or $L_{\infty} > L_{\max}$. This requirement has been removed from more recent versions of the program; some of the consequences are listed in Table 2.

THE PROBLEM OF MULTIPLE COHORTS

One question which seems to have caused a great deal of confusion among users of ELEFAN I, and indeed among fishery biologists generally, is the allegedly "continuous" spawning and recruitment of tropical fish and invertebrates. In fact, several colleagues suspect that it is impossible for

Table 2. Available options for definition in ELEFAN I of asymptotic length (L_∞) in relation to maximum length in samples (L_{\max}), with a partial list of consequences (based on pers. obs. and various contributions in this vol.).

Options	Problem(s) solved	Problem(s) created
L_∞ must be $> L_{\max}$	"Age" and growth rate at L_{\max} are defined and hence VPA and catch curve can be used for all size groups; representation of growth curve still reasonable (because L_∞ and K are inversely related)	Some available empirical equations for estimation of M (e.g., equation 6) and t_o may produce biased estimates
$L_\infty \leq L_{\max}$ possible	L_∞ and K estimates may be less biased; estimates of M and t_o based on empirical equations (e.g., equation 6) may be improved	Another set of growth parameters must be used for length-structured VPA, and catch curve. (See also Jones, Part I, this vol., and Majkowski et al., Part II, this vol.)

any algorithm—including that built into ELEFAN I—to trace a single reliable growth curve through the jumble of modes generated by "continuous spawning". However, consider the following:

- (i) continuous spawning in the sense of spawning with constant intensity throughout the year has never been demonstrated for any fish or aquatic invertebrate, i.e., cases of "continuous" spawning are actually cases where *some* females are reported in ripe conditions throughout the year, or *some* eggs can be sampled throughout the year. The bulk of the females, however, do spawn within identifiable seasons, even in the tropics (Qasim 1973; Weber 1976; Johannes 1978; Longhurst and Pauly 1987);
- (ii) recruitment can oscillate seasonally even if spawning is continuous. The oscillating transfer function between these two processes is also known as a "recruitment window" which opens only during certain parts of the year (Sharp 1980; Bakun et al. 1981);
- (iii) if recruitment into a population of fish or invertebrates were continuous (i.e., did not oscillate seasonally), length-frequency samples would simply not contain peaks and troughs.

For these reasons, length-frequency data obtained from fish or invertebrate stocks (unless they are collected with a highly selective gear) usually contain modes pertaining to one or two major cohorts per year, even when they are sampled from populations in which some females spawn throughout the year, such as by following a lunar/tidal periodicity (Johannes 1978).

When two cohorts occur per year, ELEFAN I will, if left to itself, fit a growth curve to the more prominent of these two cohorts. The parameters of a second growth curve can be fitted by selecting a prominent mode belonging to the second cohort, then letting ELEFAN I identify the best set of growth parameters for a curve forced through the selected mode.

DEALING WITH GEAR SELECTION

In simulation experiments constructed with perfect data, it has been observed by users at ICLARM and by others (P. Sparre, pers. comm.) that ELEFAN I recovers the growth parameters used to generate the length-frequency data (see Hampton and Majkowski, Part I, this vol.; Rosenberg and Beddington, Part I, this vol.). Real length-frequency data, however, are obtained from populations with unknown growth parameters by means of gears that often select for certain sizes, a feature which is often aggravated by the migrations and/or schooling habits of the fish themselves.

There are three approaches by which gear selection problems can be addressed in the context of the ELEFAN system, and each of them pertains to a different combination of gear type and structure of the length-frequency data as follows:

- (i) length-frequency data (a) collected with gill nets with single mesh size (or hooks of a single gap size), whose selection curve is not known and (b) consisting essentially of the same size group of fish, with few or no shifts of modes through time.
Solution: growth cannot be estimated using ELEFAN I.
- (ii) length-frequency data (a) obtained from catches of a fleet of gill nets of different mesh size or from trawl or purse seines or other similar gears whose selection curve is not known and (b) showing a wide variability in the position of modes.
Solution: preliminary estimates of growth parameters are obtained using ELEFAN I; then, using ELEFAN II, approximate probabilities of capture are computed by length for the gear(s) used in sampling the available length-frequency data. These probabilities are used with ELEFAN 0 to correct the available length-frequency data for selection effect. ELEFAN I is then reapplied to the corrected data to obtain improved estimates of the growth parameters (see Fig. 3 and third application example below).
- (iii) length-frequency data and gear as in (ii), and selection curves of gears available, as obtained from a selection experiment.

Solution: correct data for selection using appropriate routine in ELEFAN 0, then apply ELEFAN I to corrected data. Also, estimate M, Z and F and recruitment curve using ELEFAN IV.

Certain behavioral and/or life-history features of fish and invertebrates can also prevent length-frequency data sets from being truly representative of the population sampled. Such features are emigration from fishing grounds or strong growth dimorphism, the former leading to overestimation of total mortality and the latter to biased growth parameter estimates, among other things. It is necessary, when interpreting results obtained through ELEFAN I (or any other program of the ELEFAN system) to consider all known aspects of a fishery and of the resources upon which it rests, and to realize that this all-important step cannot be replaced by the user's guide, however detailed, of a computer program.

APPLICATIONS OF ELEFAN I

The ELEFAN I program has been rather widely disseminated since 1980 and a relatively large number of papers and reports have been published which relied predominantly or at least partly on this program. Table 3 gives details on some of these applications. It will be noted that they cover a wide range of animals, from cold temperate to tropical, and from invertebrates to teleost fishes. With only one exception, these applications have not involved fatal misinterpretation of the program features and output, and have helped their authors extract more out of their data than if they had used classical paper-and-pencil methods.

This is illustrated in the brief review below of typical applications of the ELEFAN I program as used by this author.

The first example presented here pertains to the shrimp *Penaeus kerathurus* and is based on length-frequency data published by Rodriguez (1977). As might be seen (Fig. 6), the seasonally oscillating growth curve fitted by ELEFAN I to the data at hand provides what appears to be a good fit. In fact, it appears necessary to reemphasize here that such a curve is derived by ELEFAN I *without* external input as to the age structure in the available length-frequency data.

The second example presented here, pertaining to the squid *Loligo pealei*, features ELEFAN I as an alternative to subjective paper-and-pencil methods, illustrated in Fig. 7a. The figure demonstrates how such methods can (and often must!) lead to erroneous interpretations:

- (i) the peaks linked by growth increments were selected in subjective fashion (also they were not linked by their bases as should be done);
- (ii) seasonal growth oscillations were not considered;
- (iii) no test was performed to determine whether the resulting growth increments identified led to an acceptable growth curve.

Table 3. Some examples of applications of ELEFAN I as of early 1985.^a

Species investigated	Area	Computer system/language	Reference
Molluscs			
<i>Asaphis deflorata</i>	Bahamas	TRS-80 Model I/MBASIC	Berg (pers. comm.),
<i>Codakia orbicularis</i>	Kuwait	TRS-80 Model I/MBASIC	Berg and Atalo (1984)
"pearl oyster"		Hewlett-Packard 9845B/HPBASIC	Almatar et al. (1984)
Crustaceans			
<i>Panulirus spp.</i>	Galapagos Islands	n.a.	Reck (1983)
<i>Penaeus semisulcatus</i>	Kuwait	Hewlett-Packard 9845B/HPBASIC	Mathews et al. (Part I, this vol.)
<i>Metapenaeus dobsoni</i>	off Cochin, India	n.a.	Silas et al. (1984)
Teleosteans			
<i>Gadus morhua</i>	North Sea	PDP 10/FORTRAN V	Rohde (1982)
<i>Melanogrammus aeglefinus</i>			
<i>Merlangius merlangus</i>			
<i>Limanda limanda</i>			
<i>Clupea harengus</i>			
<i>Scomber japonicus</i>			
<i>Katsuwonus pelamis</i>			
<i>Katsuwonus pelamis</i>			
<i>Katsuwonus pelamis</i>			
<i>Thunnus albacares</i>			
<i>Thunnus albacares</i>	Peru	PDP10/FORTRAN V HP 85/87/HPBASIC n.a.	Mendo (1983) White (1982) Brouard et al. (1984) Tandog (1984) Brouard et al. (1984) White (1982) Cababan (1984a, 1984b) Dalzell (1983, 1984) Pagdilao (1982) Navaluna (1982) Gjøsaeter and Sousa (1983a, 1983b)
<i>Pterocoasio pisang</i>			
<i>Stolephorus spp.</i>			
<i>Leiognathus spp.</i>			
<i>Otolithes ruber</i>			
<i>Hilsa kelee</i>			
<i>Decapterus russelli</i>			
<i>Epinephelus tauvina</i>			
<i>Pampus argenteus</i>			
	Mozambique	Univac 1100/U-BASIC	
	Kuwait	Hewlett-Packard 9845B/HPBASIC	Morgan (1983, 1985)

^aThis table does not include applications of ELEFAN I by the author (see text); "n.a." refers to "not available".

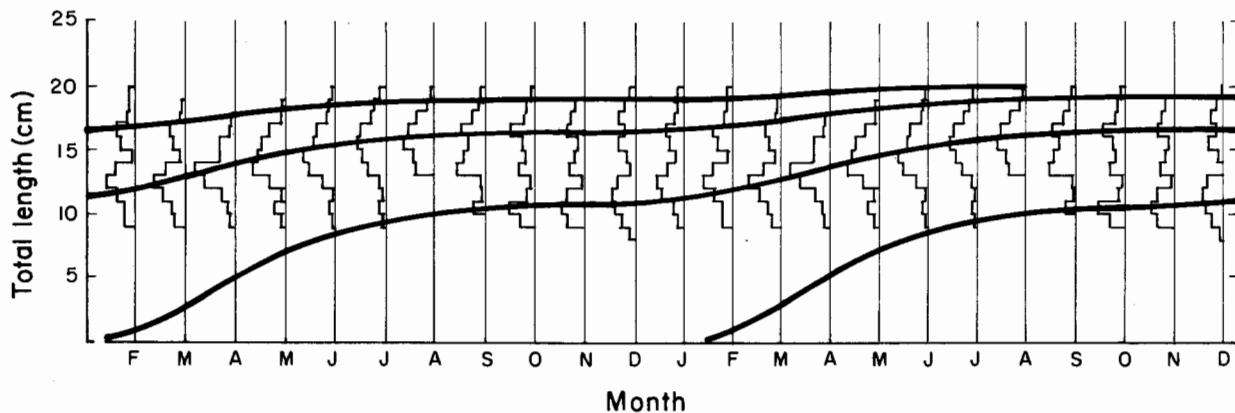


Fig. 6. Growth of females of the penaeid shrimp *Penaeus kerathurus* off Cadiz, Spain, based on length-frequency data in Rodriguez (1977) and the ELEFAN I program. The parameter estimates are: $L_{\infty} = 21.0$ (cm, total length), $K = 0.8$ (year $^{-1}$), $C = 0.9$. Winter Point = 0.8; the estimated ESP/ASP was 0.457. It is recalled that this growth curve was established by ELEFAN I without any external input as to the age-structure of the population or to the relative age(s) corresponding to any peak(s).

Figs. 7b and 7c show an alternative interpretation of the same data through ELEFAN I. The resulting growth curves have parameters consistent with those of other loliginid squids and the seasonal growth oscillations have an amplitude consistent with that of the oscillations of temperature in the Western Gulf of Mexico (see Pauly 1985a).

The third example pertains to the Peruvian anchoveta *Engraulis ringens* (Northern/Central stock) and demonstrates the impact of correcting length-frequency data for the effects of gear selection on growth parameter estimates, using an iterative approach first applied to *Rastrelliger brachysoma* in Burma (see Pauly and Sann Aung 1984), and illustrated in Fig. 8.

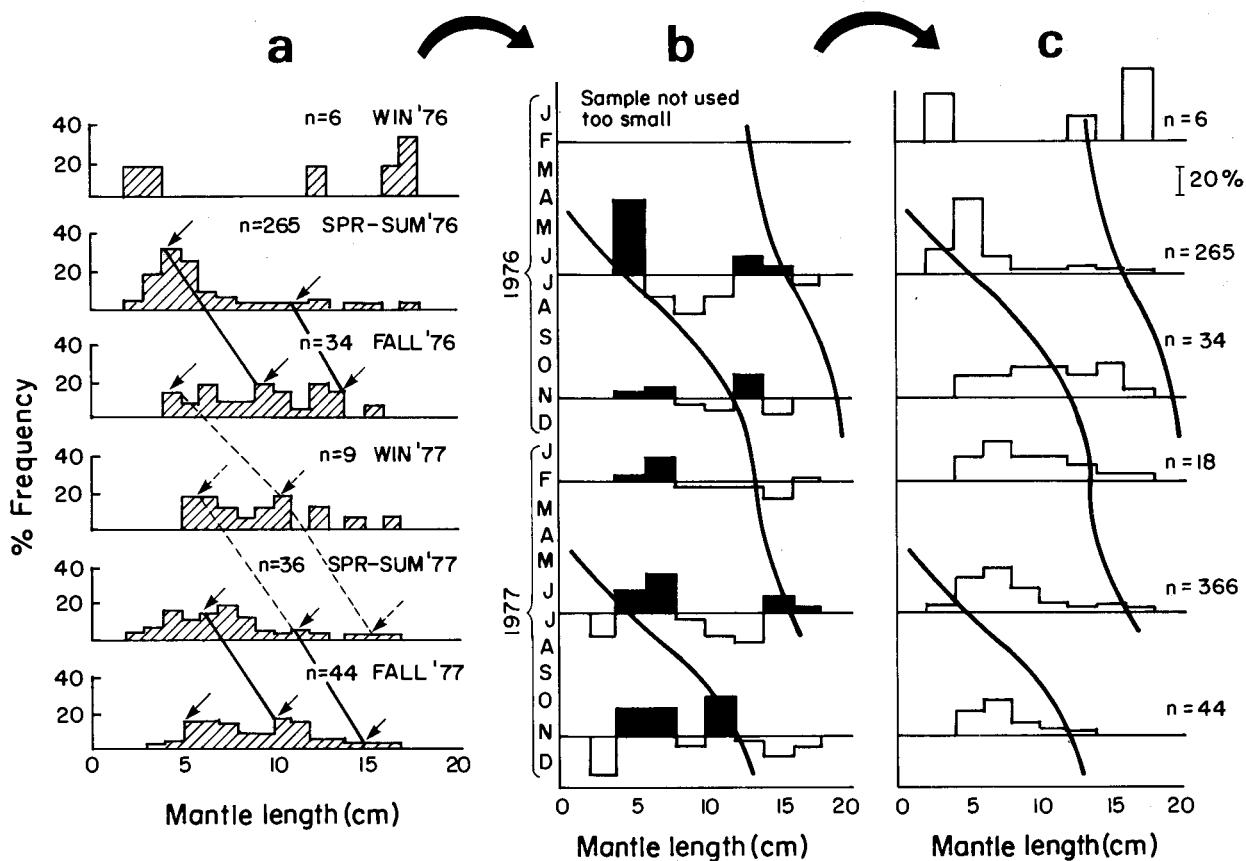


Fig. 7. Two interpretations of length-frequency data for *Loligo pealei* from the Western Gulf of Mexico (original length-frequency data from Hixon et al. 1981). a) Original interpretation (see text for comments). b) Restructured sample as created and used by ELEFAN I, with superimposed growth curves. (Note that original length-frequency data were regrouped into larger size classes for the ELEFAN I analysis). c) Length-frequency samples of Hixon et al. (1981), with superimposed growth curves as estimated using ELEFAN I (see text for comments).

In a first pass, initial estimates of L_∞ and K obtained from the analysis of several sets of length-frequency data were averaged and used, along with length-frequency data from the two months of each year from 1953 to 1982 with the smallest fishes (i.e., the month at which recruits enter the stock) to derive a length-converted catch curve (Fig. 8b). A selection curve was then derived (Fig. 8c) which provided the probabilities of capture by means of which the original length-frequency data (here October 1963 to September 1964 were corrected). Figs. 8a and 8d show the restructured samples for the uncorrected and corrected data, respectively. As might be seen from Fig. 8d, the peaks related to the smaller fish here shifted downward, suggesting a more rapid growth (i.e., higher K) than would have been inferred from Fig. 8a.

Other applications of the ELEFAN I program to penaeid shrimps and Philippine and Indonesian teleosts are presented in Pauly et al. (1984), Ingles and Pauly (1984) and Dwiponggo et al. (1987).

One important methodological result of these various applications is the identification of a flaw in the original version of ELEFAN I, which counted peaks everytime they were "hit" by a growth curve, although each peak, representing one age group, ought to be hit only once by a well-fitting growth curve.

This flaw has now been corrected (see Pauly 1985b) and the points in a given run of positive values are counted once only when they are "hit", which completely eliminates the "drifting" of seeded K values towards unrealistically low values discussed in Pauly et al. (1980).

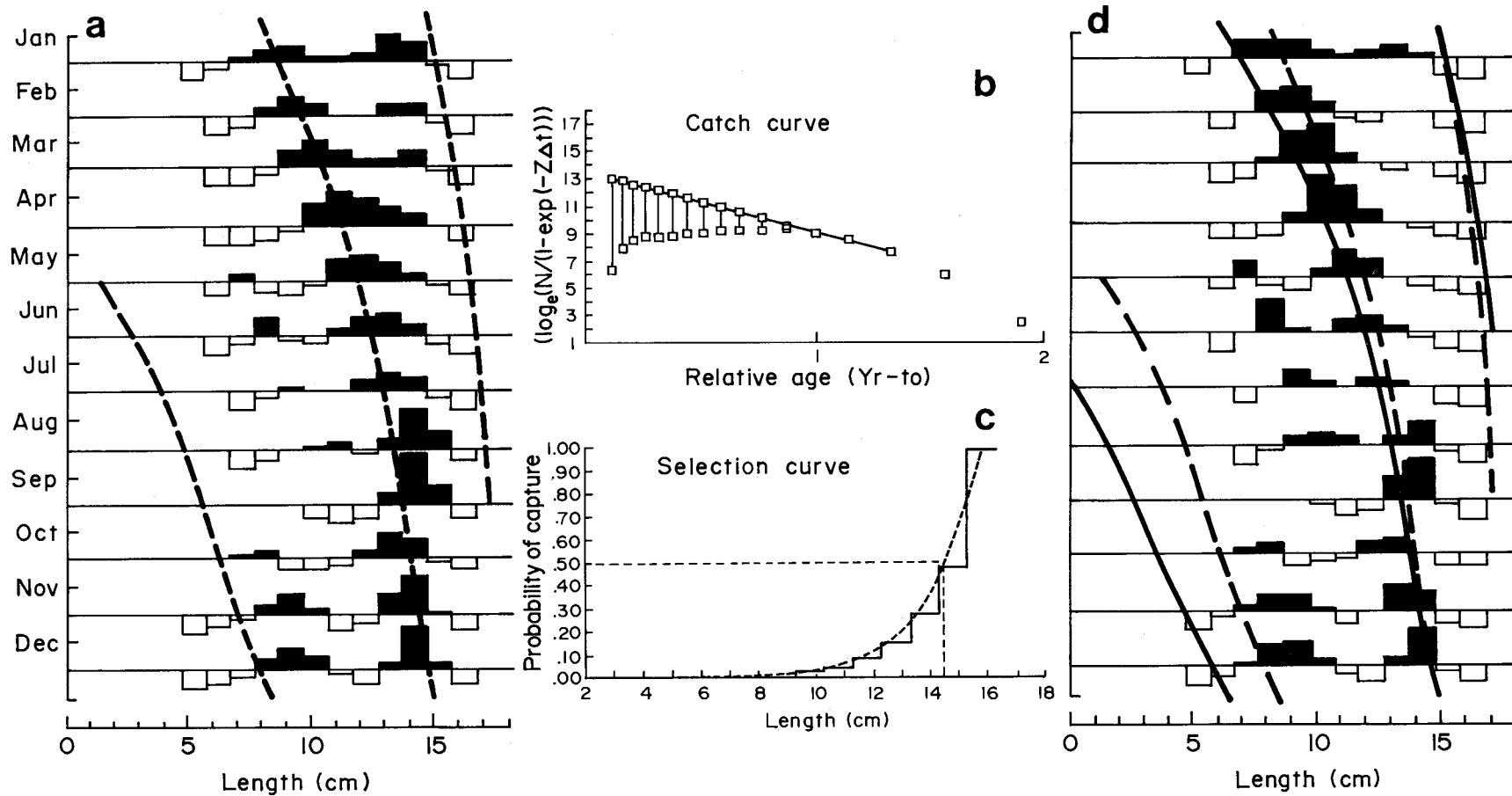


Fig. 8. Steps used to correct ELEFAN I estimates of growth parameters in Peruvian anchoveta, using the method of Pauly (1986). a) Estimation of a preliminary set of growth parameters for each cohort from 1954 to 1982 and averaging of the same. b) Derivation of a length-converted catch curve based on growth parameters estimated in (a) and an accumulated length-frequency file composed of data from the two months of each year containing the smallest fish (such as to obtain a correction for gear selection covering a size range as wide as possible; backward projection of catch curve to estimate number of fish that would have been caught, had it not been for gear selection and/or incomplete recruitment. c) Estimation of probabilities of capture from the ratio of fish caught to virtual fish, by length, and division of all original length-frequency data by the appropriate probabilities of capture, and d) Re-estimation of growth parameters. The correction leads to much improved estimates of the von Bertalanffy parameter K (Palomares et al. 1987 and see text).

ELEFAN II

INTRODUCTION TO ELEFAN II

ELEFAN II is a collection of routines which can be used, following the application of ELEFAN I to a given data set, to extract estimates of total, natural and fishing mortalities, information on the seasonality of recruitment and on probabilities of capture by length *without catch data being available or selection experiments having been conducted*. The program thus differs from ELEFAN III and IV in that it requires no inputs other than growth parameter estimates and length-frequency data.

LENGTH-CONVERTED CATCH CURVES

Total mortality in ELEFAN II is obtained via a length-converted catch curve, as described in Pauly (1980a, 1982, 1983a, 1983b, 1984a, 1984b). Important here is that, conceptually, catch curves based on length data are actually older than catch curves based on age composition data, with crude first estimates of mortality being obtained by Edser (1908), Heincke (1913) and Baranov (1918). Subsequent work has shown, however, that these early length-based catch curves were often biased and since then estimation of Z from catch curves has been based almost exclusively on age-structured catch curves (review in Ricker 1975).

Only recently have length-based catch curves been proposed which provide unbiased estimates of Z (Pauly 1980a; Gulland 1983). These "length-converted catch curves" have been shown to be far more versatile than age-structured catch curves, allowing for example, inferences to be drawn on the selection process of the gear used for sampling (Pauly 1984a, 1984b and see below) or, when used in conjunction with selection curves, the estimation of M from length samples (Munro 1984 and see below).

It is emphasized that the data used for catch curve estimates of Z must be representative of an equilibrium, or stable-age distribution. This problem, which was reviewed in great detail by Ricker (1975) in conjunction with age-structured catch curves, also occurs with length-converted catch curves, i.e., the samples used to construct the catch curve must reflect average conditions during which recruitment has varied little or randomly such that total mortality can be considered constant (see Hampton and Majkowski, Part I, this vol.).

One obvious way to simulate average conditions for construction of a length-converted catch curve is to pool length-frequency data from a longer period of time to smooth out recruitment pulses. Unfortunately, guidelines on the size of length-frequency samples suitable for the estimation of vital statistics are not presently available, although some rules of thumb have been suggested (see Munro 1982, Table 1 and Hoenig et al., Part II, this vol.). Research on this topic has been initiated by a number of authors and results should soon appear in the literature. For the time being, a routine has been incorporated into ELEFAN II which weights the samples by the square root of their size prior to combining them in a single sample for construction of the catch curve. This routine suggested by H. Lassen (pers. comm.) is based on (a) the need to use all the information incorporated in a given set of length-frequency samples (very small samples are not discarded as must be done with ELEFAN I) and (b) the fact that the standard error of means (i.e., of the mean lengths of fish in the various broods) is proportional to the square root of the number of fish used in the computation.

Obviously, schemes of sample weighting other than the dynamic range compression used here are feasible; two other approaches are incorporated in the ELEFAN system:

- (i) when catch-at-length data are available (rather than only length-frequency data), a single "sample" can be obtained by simply adding up the catches by length class. A routine provided in ELEFAN III (see below) can be used for such purpose;
- (ii) when sampling in certain periods is less intensive than in others, an adjustment can be conceived, as suggested by J.L. Munro (pers. comm.), such that those samples representing a longer period are given a larger weight than those obtained from periods represented

by many samples. Such an adjustment can be made using the appropriate routine in ELEFAN IV, where the "temporal weighting factors" are simply the distance (in time units) between a given sample and the nearest sample that precedes or succeeds it.

The ramifications of these various schemes to simulate equilibrium conditions have not been studied fully. Users of the ELEFAN system are invited to perform comparative studies of these and any other methods that might seem appropriate.

ESTIMATES OF M, F AND E

That the natural mortality (M) of fish and presumably also of invertebrates has a strong relationship with their growth parameters has been known at least since the landmark paper of Beverton and Holt (1959). Many fisheries biologists have used their empirical relationships to obtain preliminary estimates of M in cases where other approaches could not be used.

Pauly (1980b) extended the list of variables which can be used to predict M by showing that after accounting for the partial correlation between M and L_∞ on one hand, and M and K on the other, M was strongly correlated with mean environmental temperature. These interrelationships were then expressed in the form of a multiple regression which has the form

$$\log_{10} M = -0.0066 - 0.279 \log_{10} L_\infty + 0.6543 \log_{10} K + 0.4634 \log_{10} T \quad \dots 6)$$

and which can be used to obtain preliminary estimates of M , given L_∞ (total length, in cm), T = mean environmental temperature (in °C) and K (expressed, as is done throughout the ELEFAN system, on an annual basis). Equation (6) is built into ELEFAN II. Readers interested in details of the derivation and of the appropriate use of this equation should consult Pauly (1980b) and Gulland (1984), respectively.

Once Z has been estimated from a catch curve and M from equation (6), preliminary estimates of F can be estimated by subtraction, i.e., $F = Z - M$, while the exploitation rate (E) is estimated from $E = F/Z$. This information is sufficient, given an estimate of mean size at first capture (L_c , see below) to perform a yield-per-recruit analysis using the method of Beverton and Holt (1966).

Alternatively, if it is accepted that the optimum value of F in a given exploited stock (F_{opt}) is about equal to M (Gulland 1971), we also have

$$E_{opt} = 0.5 = F_{opt}/(F_{opt} + M) \quad \dots 7)$$

Thus, using ELEFAN II as a follow-up to ELEFAN I, it is *in principle possible to assess whether or not a stock is overfished, based solely on length-frequency data obtained from that stock*.

It must be emphasized that this is so only *in principle*. In reality, the estimates of Z will be biased one way or the other by the sampling gear and by the behavior of the animals sampled. Also, the estimated value of M will be an overestimate or underestimate of the real rate of natural mortality prevailing in the specific stock under investigation (see Gulland 1984). Finally, equation (7) may not apply (as suggested by Francis 1974; Caddy and Csirke 1983; Beddington and Cooke 1983). Still, the principle holds, and may serve as a rationale for attempts to improve the quality of a length-based system, including data collection.

ESTIMATION OF PROBABILITIES OF CAPTURE

The estimation of probabilities of capture from the ascending, left arm of length-converted catch curves goes back to an approach developed in 1981 (see Pauly et al. 1984) and to a paper by Munro (1984) in which this approach was put on a more rigorous footing. Essentially, the method (see Fig. 9) consists of extrapolating the right, descending left side of a catch curve such that fish that "ought" to have been caught (had it not been for the effect of incomplete selection and/or recruitment) are added to the curve, with the ratio of those "expected" numbers to those that are actually caught being used to estimate probabilities of capture (Fig. 10). This approach, which is related to work done by Jensen (1982) and Hoydal et al. (1982), has been found (Anon. 1982) to

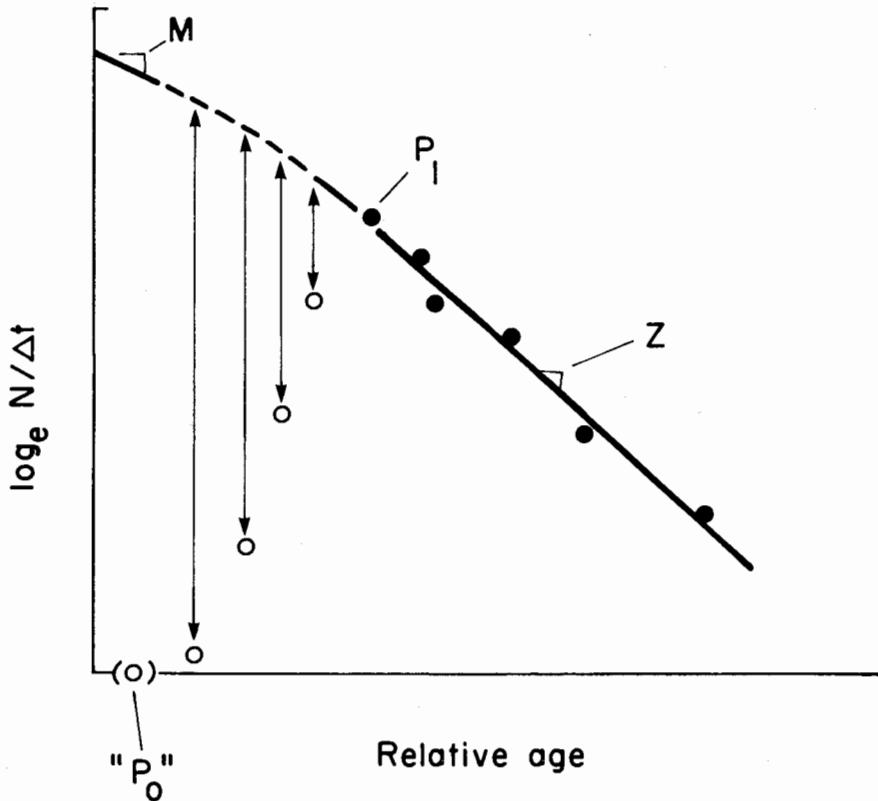


Fig. 9. Schematic representation of method to derive probabilities of capture from the left, ascending arm of a length-converted catch curve. P_1 identifies the first point from which on the probability of capture is 1, and hence the points to its right hand side which can be used for computation of Z . Point " P_0 " represents the first point (when going from right to left) where the probability of capture is 0; the mortality corresponding to this point is M , by definition. With M at P_0 , and Z at P_1 , intermediate mortalities can be interpolated, and population sizes reconstructed by backward extrapolation, starting from point P_1 . The probabilities of capture are then computed as the ratios between numbers caught and numbers expected. When $Z = M$, the method obviously requires no interpolation of mortalities and backward projection is done via the catch curve itself. The method also allows for mortality to be higher in small than in large animals; in this case, however, the value of M used is $> Z$.

provide reasonable estimates of mean size at first capture (L_c). Thus, it has the potential of making (some) mesh selection experiments superfluous, especially so when length-frequency data are available that include very small fish, i.e., covering the month(s) in which recruitment occur (M. Yahiaoui, pers. comm.).

RECRUITMENT PATTERNS

Since it is the pulsed nature of annual recruitment into a population which generates the peaks and troughs in length-frequency data, the converse also applies that, given a set of length-frequency data and growth parameters, one should be able to recover the pulsing of annual recruitment.

A routine which performs this task was incorporated into ELEFAN II; it produces what are here called "recruitment patterns", i.e., graphic descriptions of the recruitment process that generated the length-frequency data at hand.

The approach gives approximate results because recruitment patterns, while allowing statements on the number of recruitment pulses per year and on the relative importance of these pulses when compared to each other, are based on two assumptions which will hardly ever be met in reality:

- (i) all fish in a given data set grow according to the equation defined by the growth parameters used;
- (ii) one month out of twelve always has zero recruitment.

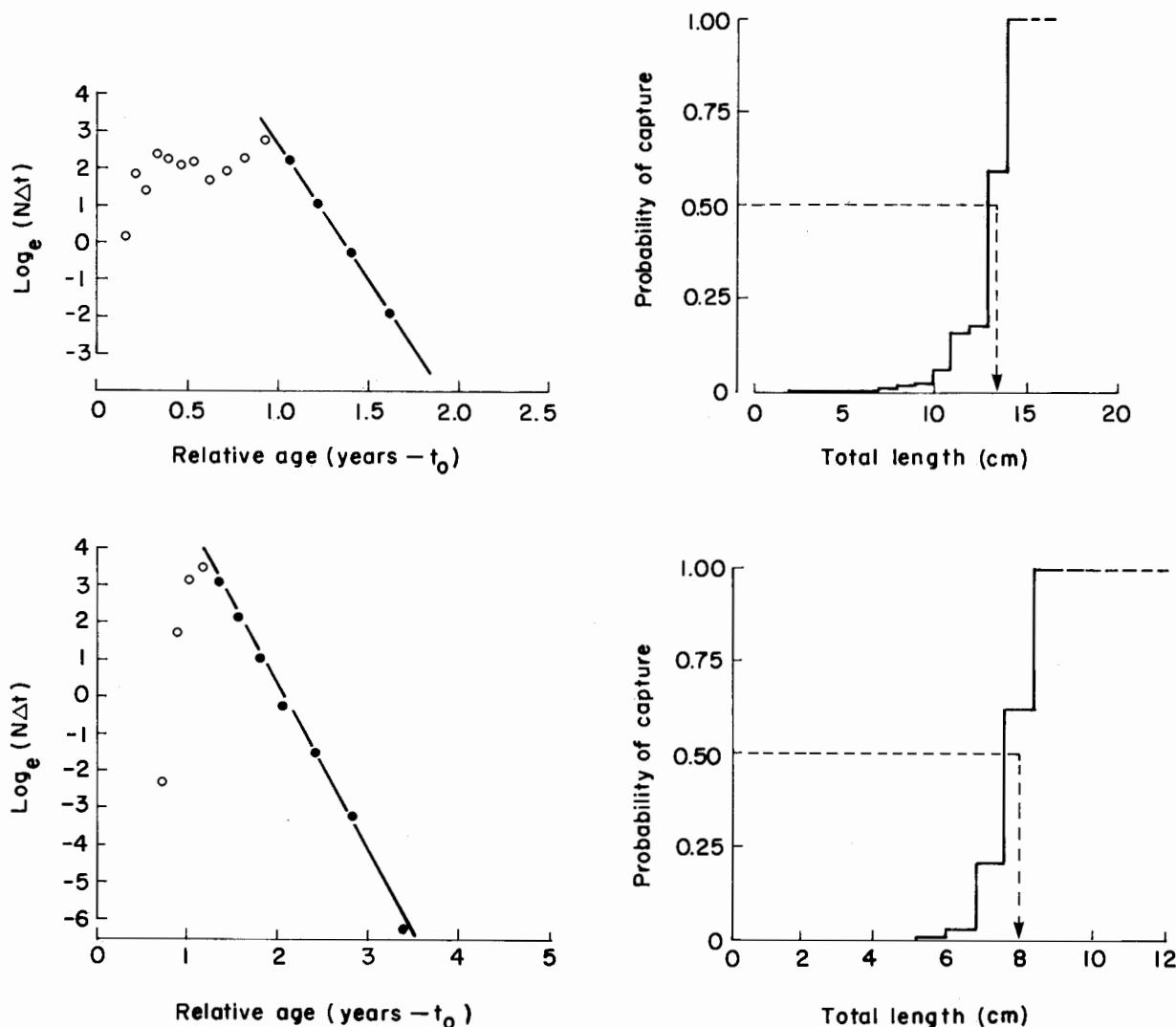


Fig. 10. Examples of length-converted catch curves and of the plots of probabilities of capture against length derived from them. Above: data on *Sardinella longiceps*, with estimated values of $Z = 5.77 (\text{year}^{-1})$ and $L_c = 13.5 \text{ cm}$; below: data on *Pomadasys argyreus* with estimated values of $Z = 2.93 (\text{year}^{-1})$ and $L_c = 8.1 \text{ cm}$ (from Ingles and Pauly 1984).

The first of these two assumptions is common to all routines and programs in the ELEFAN system; its validity and overall impacts are discussed in Hampton and Majkowski (Part I, this vol.) and Rosenberg and Beddington (Part I, this vol.). As far as recruitment patterns are concerned, it should have relatively little impact on the result because the actual computation of these patterns assigns a very small role to older, larger fish and a larger role to smaller fish, whose deviations from the growth curve typical of the whole population are much smaller than in larger, older fish.

The second assumption will probably never be strictly met, because (a) there may be more than one month per year in which no recruits enter a stock (especially in temperate fishes and invertebrates) or because (b) some recruits may be entering a stock every month (especially in tropical fishes and invertebrates).

Fig. 11 shows the match between a recruitment pattern as obtained by means of ELEFAN II and the seasonal pattern of gonad maturity in a stock of Philippine groupers. This and a number of application examples (see Pauly and Navaluna 1983 and Fig. 12) suggest that recruitment patterns obtained by ELEFAN II do contain useful information, from which legitimate inferences on the dynamics of fish and invertebrate stocks can be drawn.

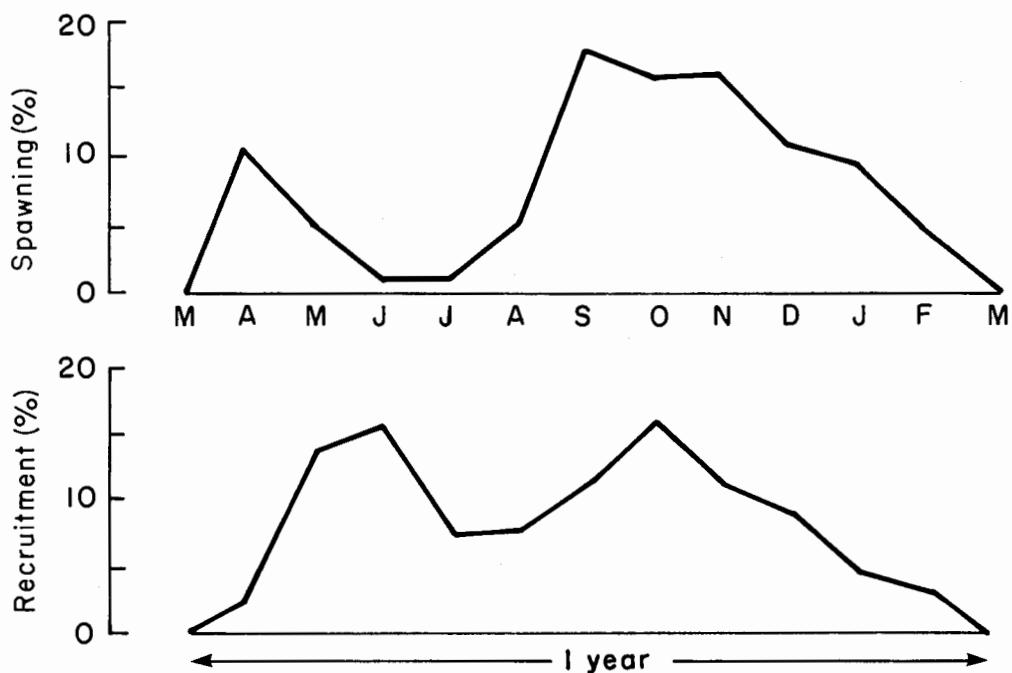


Fig. 11. Above: seasonal pattern of spawning condition in gonads for six-banded grouper (*Epinephelus sexfasciatus*) from the Visayan Sea, Philippines (data normalized by expressing as zero the month with lowest mean gonadal stage). Below: recruitment pattern in the same stock of fish. Note overall agreement of shape, suggestive of two spawning seasons and two recruitment pulses per year. Note however that absolute time scale of recruitment pattern is unknown. From Pauly and Ingles (1981).

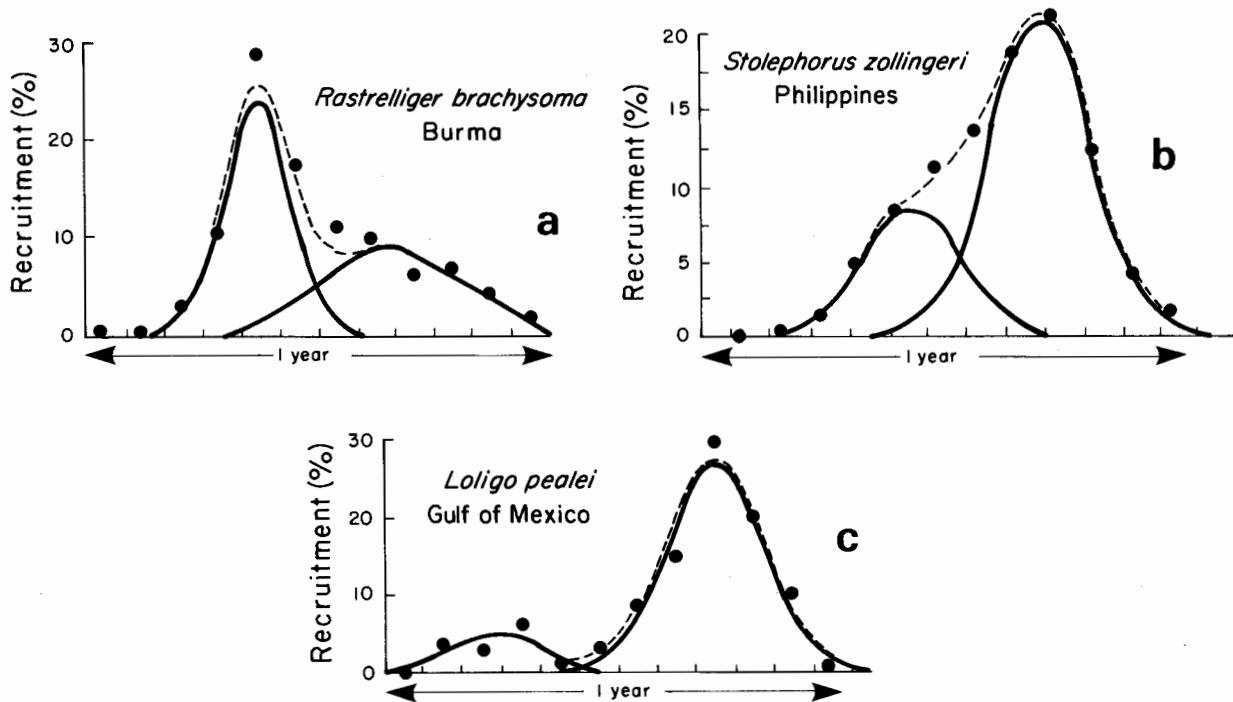


Fig. 12. Examples of recruitment patterns. a) *Rastrelliger brachysoma* (see also Fig. 9). b) *Stolephorus zollingeri* (from Pauly and Navaluna 1983). c) *Loligo pealei* (see also Fig. 8). Note in all three cases that annual recruitment consists of two pulses, one stronger than the other (see text).

ELEFAN III

INTRODUCTION TO ELEFAN III

ELEFAN III differs from ELEFAN I and II in that more data are needed to run the program than just length-frequency data. The added data which ELEFAN III requires are monthly catch data. Also, as is characteristic of most approaches based on Virtual Population Analysis (VPA), it is the *entire catch* from the whole stock which must be used, not only a part thereof.

ELEFAN III consists of four major routines:

- (i) the derivation of catch-at-length data from catch data and length-frequency samples;
- (ii) ordinary (age-structured) VPA, termed here VPA I;
- (iii) the VPA version of Jones' (1981) length cohort analysis, termed here VPA II;
- (iv) a routine to derive (pseudo-) cohorts from catch-at-length data, with subsequent application of age-structured VPA to these "cohorts"; this approach is termed VPA III.

The discussion below is limited to the principles behind these various routines and the reasons they were incorporated into ELEFAN III; further details on ELEFAN III are given in Morgan and Pauly, Part II, this vol.).

VIRTUAL POPULATION ANALYSIS (VPA I)

The principle behind VPA hinges around the concept of the "cohort", i.e., a group of fish or invertebrates born or hatched and recruited at more or less the same time and sharing throughout their lives a common pattern of exploitation by their predators and a fishery.

In most applications of VPA, the "exploitation" by predators is expressed by a single value of the rate of natural mortality M , which is assumed to apply throughout most of the life of the cohort (i.e., from the time the animals in the cohort are recruited to the time the cohort is decimated).

VPA, then, is a method to reconstruct a cohort (that is, to estimate the number of animals that were present in each age group or cohort) by using a value of M and the catch by the fishery in terms of numbers of each group in the cohort. The cohort is always reconstructed *backward* in time, starting with the last caught animals of a cohort (the "terminal catch"), which is used to obtain an approximation of the "terminal population" using a *guessed* value of the "terminal fishing mortality" exerted upon the "terminal population". Then, successive estimates of the population sizes are obtained for each age group by "adding" the catch of the fishery to a previous estimate of population size, to which the number of fish caught by predators has also been "added".

The method, which was formulated by Gulland (1965) has been reviewed by Pope (1972), Mesnil (1980) and Pauly (1984a). It is incorporated in ELEFAN III as VPA I to allow users of the ELEFAN system to acquaint themselves with VPA in its original form and to perform age-structured VPAs whenever the data for such analyses are available.

VPA II

Catch-at-length data differ from catch-at-age data (such as used in age-structured VPA or VPA I) in that they do not pertain—nor can readily be made to pertain—to a given cohort as defined above.

Still, versions of VPA can be run with catch-at-length data. Two very different versions of VPA are incorporated in ELEFAN III which use catch-at-length data. The first of these (VPA II) is conceived such that while it is not meant to help in reconstructing *any given cohort*, it can, given catch data covering the life span of several cohorts, reconstruct an *average cohort*. Thus, the results of VPA II are not structured in time; instead the population size and fishing mortalities which are typical outputs of VPAs are related to *sizes* (Table 4). These results, therefore cannot be used to manage a fishery in real time, nor to study temporal fluctuations of recruitment.

Table 4. Facsimile of printout of VPA II results obtained with the "Compleat ELEFAN" package. Data refer to West African hake *Merluccius merluccius* (see Morgan and Pauly, Part II, this vol.) for details on program used, and source of hake data.

VPA II results for MERSEN78

LENGTH CLASS (cm)	CATCHES	POPULATION	F. MORTALITY (1/year)
81.00- 87.00	46.00	92.00	0.2800
75.00- 81.00	16.00	145.79	0.1186
69.00- 75.00	96.00	304.23	0.4305
63.00- 69.00	181.00	599.59	0.4432
57.00- 63.00	228.00	1017.35	0.3364
51.00- 57.00	322.00	1626.46	0.3140
45.00- 51.00	653.00	2714.90	0.4199
39.00- 45.00	1871.00	5325.94	0.7079
33.00- 39.00	2959.00	9580.45	0.6395
27.00- 33.00	3889.00	15539.98	0.5259
21.00- 27.00	8134.00	26960.11	0.6931
15.00- 21.00	25227.00	58282.88	1.1588
9.00- 15.00	14463.00	82681.77	0.4076
3.00- 9.00	1823.00	96629.48	0.0421
Total catch :	59908	Natural mort. :	0.280 K : 0.100
Mean E :	0.620	Term. F. mort.:	0.280 Loo : 130 cm
Mean F :	0.457		

VPA III

VPA III was devised to combine the advantages of both VPA I and II, to allow catch-at-length data as input data and still to run an "age"-structured VPA, thus obtaining results that are structured both in size and in time (Fig. 13). This was achieved by assuming that all fish in the population under investigation have the same growth parameters, as is also assumed in the other ELEFAN programs. The effects of this assumption on the within-year estimates of recruitment variability have not been studied in detail. Exercises with the method suggest, however, that the VPA III routine of ELEFAN III generates monthly estimates of recruitment that are strongly autocorrelated (see also Mendelsohn and Mendo 1987) and which probably underestimate true within-year recruitment variability.

ELEFAN IV

THE ESTIMATION OF M

The ELEFAN IV program shares with ELEFAN III the feature of requiring more information than just length-frequency data. However, rather than requiring catch data, ELEFAN IV requires selection curves, that is, probabilities of capture by length with the gear used to obtain the available samples and of the commercial gear exploiting the stock under consideration (the two gears may be the same, in which case only one selection curve is required). Given these, and a set of length-frequency data covering a period of at least one year, ELEFAN IV can be used to estimate M and its standard error based on the method of Munro (1984) and also can be used to estimate probabilities of recruitment (by length) when recruitment overlaps with the selection range.

The basic idea behind this approach is illustrated in Fig. 14, which is based on constructed length data such as would be produced by a gill net fishery.

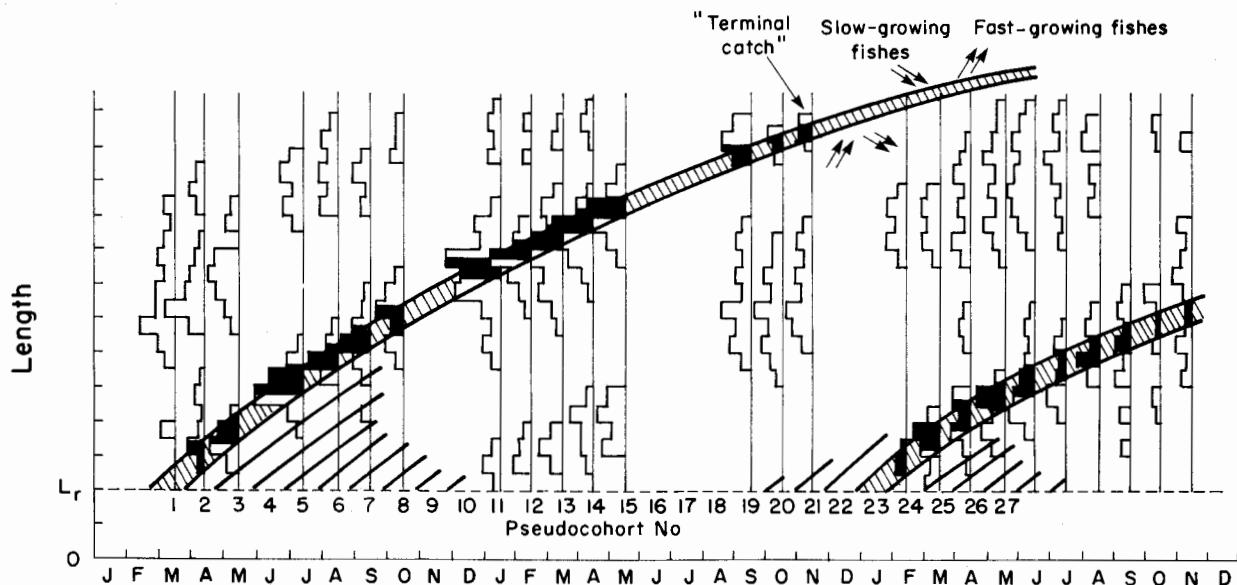


Fig. 13. Schematic representation of method to "slice" pseudo-cohorts from length-frequency data. As suggested by the arrows, the slow-growing fishes of a given pseudo-cohort will tend to be lost to preceding pseudo-cohorts and the fast-growing ones will tend to grow into the following pseudo-cohorts (from Pauly and Tsukayama 1983, redrawn).

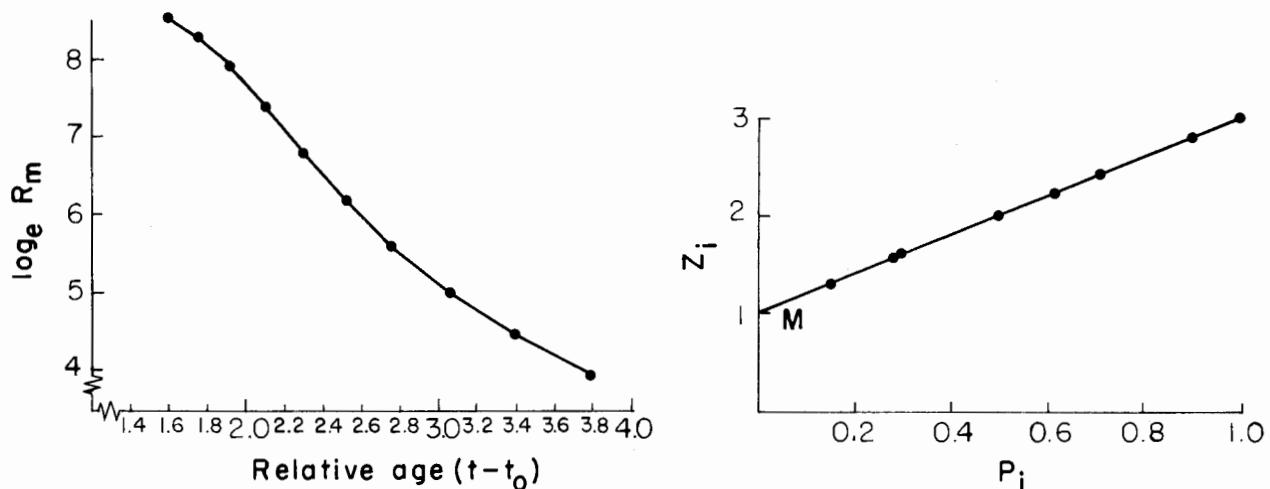


Fig. 14. Left: length-converted catch curve for fishes exploited by a single size of gill net in which the values of R_m (number caught/probability of capture) are plotted against relative age, and where the change of slope is due to mesh selectivity. Right: plot of Z_i (between successive length groups) against probability of capture P_i at the median between successive length groups. The intercept provides an estimate of M . (Both figures taken from Munro 1984).

The values of M estimated via ELEFAN IV differ from those obtained through ELEFAN II in that the former are estimated from the length-frequency data at hand, rather than from a built-in empirical equation. Therefore, it should be possible using ELEFAN IV to estimate values of M in different years, e.g., to follow the effects of predator removal on the natural mortality of a given stock of fish.

Users of ELEFAN IV will notice that this program differs also from ELEFAN II in being able to handle data that have been obtained using very selective gears, such as gill nets and hooks.

PRESENT PROBLEMS OF APPLICABILITY

Two problems must be mentioned which occur in conjunction with ELEFAN IV as it presently stands:

- (i) the method appears very sensitive to random variability in the length-frequency data used and reasonable estimates of M will generally be derived only from data that have been obtained through a rigorous, well-planned sampling design;
- (ii) because of (i) and because the idea behind ELEFAN IV has been presented only recently (in mid-1984), no data set has been found in the literature which could be used to illustrate the method, which therefore relies on a constructed example. It is therefore not certain that the approach implemented in ELEFAN IV will find wide applicability.

Discussion

The ELEFAN package of programs was developed to replace the suite of highly subjective "paper-and-pencil" methods which have been applied to length-frequency data since 1891.

Several, highly sophisticated alternatives to the ELEFAN package already exist (Sparre, Part I, this vol.; Pope and Yang, Part I, this vol.) or are presently being developed, and some may turn out to be more reliable, more robust and more rigorous than the ELEFAN programs (see Rosenberg and Beddington, Part I, this vol.).

Such programs, in order to become as useful and widespread as the ELEFAN package, will require, however, the following features:

- (i) they should run (e.g., in BASIC) on a widespread brand of microcomputer;
- (ii) their logic should be accessible to users without advanced degrees in mathematics and statistics;
- (iii) they should accommodate specific features of both temperate and tropical fishes and aquatic invertebrates, i.e.,
 - seasonally oscillating growth
 - one or two recruitment pulses per year;
- (iv) they should not require more than easy-to-obtain data such as length-frequency data with possibly some age or tagging data included (see below); particularly they should *not* require data to be weighted by C.P.U.E.;
- (v) they should accommodate data on incompletely selected and/or incompletely recruited animals.

One method which fulfills all these criteria and hence goes beyond the ELEFAN I program is that developed by Morgan (Part I, this vol.). There, the ELEFAN I approach is combined with an approach for the incorporation of age and/or tagging data such that a *single* set of growth parameters is estimated through simultaneous analysis of the combined length/age/tag returns data sets (or any paired combination thereof). It is probably approaches of this type, optimally using a variety of inputs, which will prevail in the future, since they would allow combining the wide availability of length-frequency data with the greater precision that can be obtained from age data.

Postscript: since this was written (in late 1984), numerous improvements of the ELEFAN approach—many of them based on inputs received from participants of this conference—were incorporated into ELEFAN I to IV. Only a few of these are discussed in this paper, which thus remains largely as originally presented; see Morgan and Pauly (Part II, this vol.) for a discussion of programs incorporating these changes and improvements.

References

- Abramson, N.J. 1971. Computer programs for fish stock assessments. FAO Fish. Tech. Paper 101, Rome. pag. var.
- Almatar, S., G.R. Morgan and S. Hakim. 1984. The pearl oyster fishery of Kuwait, p. 575-600. In C.P. Mathews (ed.) Proceedings of the Third Shrimp and Fin Fisheries Management Workshop, 9-11 October 1983. Vol. II. Kuwait Institute for Scientific Research.
- Anon. 1982. Comparative study of three methods for total mortality estimation, p. 91-125. In Report of the second technical consultation on stock assessment in the Balearic and Gulf of Lions statistical divisions, Casablanca, Morocco, 7-11 December 1981. FAO Fish. Tech. Rep. 263. 165 p.
- Bakun, A., J. Bayer, D. Pauly, J.G. Pope and G.V. Sharp. 1981. Ocean science in relation to living resource. Can. J. Fish. Aquat. Sci. 39(7): 1059-1070.
- Baranov, F.I. 1918. On the question of the biological basis of fisheries. Nauchn. Issled. Ikhtiol. Inst. Izv. 1: 81-128. (In Russian).
- Beamish, R.J. and G.A. McFarlane. 1983. The forgotten requirement for age validation in fisheries biology. Trans. Amer. Fish. Soc. 112(6): 735-743.
- Beddington, J.R. and J.G. Cooke. 1983. The potential yield of fish stocks. FAO Fish. Tech. Rep. 242, Rome, 47 p.
- Berg, C.J. Jr. and P. Atalo. 1984. Potential of chemosynthesis in molluscan aquaculture. Aquaculture 39: 165-179.
- Bertalanffy, L. von. 1938. A quantitative theory of organic growth (Inquiries on growth laws. II) Hum. Biol. 10(2): 181-213.
- Beverton, R.J.H. and S.J. Holt. 1959. A review of the lifespan and mortality rates of fish in nature and the relation to growth and other physiological characteristics, p. 142-177. In G.E.W. Wolsteholmen and M. O'Connor (eds.) Ciba Foundation colloquium on ageing: the lifespan of animals. Vol. 5, Churchill, London.
- Beverton, R.J.H. and S.J. Holt. 1966. Manual of methods for fish stock assessment. Part 2. Tables of yield functions. FAO Fish. Tech. Pap. 38. Rev. 1: 67.
- Bhattacharya, C.G. 1967. A simple method of resolution of a distribution into Gaussian components. Biometrics 23: 115-135.
- Brouard, F., R. Grandperrin and E. Cillaurren. 1984. Croissance des jeunes thous jaunes (*Thunnus albacares*) et des bonites (*Katsuwonus pelamis*) dans le Pacifique Tropical Occidental. Notes et Doc. d'Océanographie No. 10, Mission O.R.S.T.O.M., Port Vila, Vanuatu, 23 p.
- Cabanban, A.S. 1984a. Some aspects of the biology of *Pterocaesio pisang* (Bleeker, 1853) (Pisces, Caesionidae) in Central Visayas. University of the Philippines. MS Thesis.
- Cabanban, A.S. 1984b. Assessment of *Pterocaesio pisang* (Bleeker, 1853, Family Caesionidae) at Apo and Sumilon Island in Central Visayas, Philippines. Fishbyte 2(1): 20-21.
- Caddy, J.F. and J. Csirke. 1983. Approximation to sustainable yield for exploited and unexploited stocks. Océan. Trop. 18(1): 3-15.
- Cassie, R.M. 1954. Some use of probability paper in the analysis of frequency distributions. Aust. J. Mar. Freshwat. Res. 5: 513-522.
- Cushing, D.H., editor. 1983. Key papers on fish populations. IRL Press, Oxford-Washington. 405 p.
- Dalzell, P. 1983. Raw data and preliminary results for an analysis of the population dynamics of PNG bait-fish. Dept. Prim. Industry, Fish. Survey Branch, Rep. No. 83-04, Port Moresby, Papua New Guinea. 59 p.
- Dalzell, P. 1984. The population biology and management of bait-fish in Papua New Guinea waters. Dept. of Primary Industry. Fish. Res., Survey Branch. Report 84-05, Port Moresby, Papua New Guinea. 59 p.
- Ebert, T.A. 1980. Estimating parameter in a flexible growth equation, the Richards function. Can. J. Fish. Aquat. Sci. 37: 687-692.
- Ebert, T.A. 1981. Estimating mortality rates from size structure and growth parameters when recruitment is periodic. Limnol. Oceanogr. 26: 764-769.
- Edser, T. 1908. Notes on the number of plaice at each length, in certain samples from the southern part of the North Sea, 1906. J.R. Stat. Soc. 71: 686-690.
- Francis, C.R. 1974. Relationship of fishing mortality to natural mortality at the level of maximum sustainable yield under the logistic stock production model. J. Fish. Res. Board Can. 31: 1539-1542.
- Fulton, T.W. 1904. The rate of growth of fishes, p. 141-240. In 22nd Annual Report Fish. Board Scotland, Part III.
- Gjøsaeter, J. and M.I. Sousa. 1983a. Reproduction, age and growth of the kelee shad, *Hilsa kelee* (Cuvier 1829) (Pisces: Fam. Clupeidae) with information on its fishery in Maputo Bay, Mozambique. Rev. Invest. Pesqu. (Maputo) No. 8: 53-80.
- Gjøsaeter, J. and M.I. Sousa. 1983b. Reproduction, age and growth of the Russel's scad, *Decapterus russelli* (Rüppel, 1829) (Carangidae) from Sofala Bank, Mozambique. Rev. Invest. Pesqu. No. 8: 83-108.
- Goeden, G. 1978. A monograph of the coral trout. Res. Bull. No. 1 Queensland Fish. Serv. Brisbane, 42 p.
- Gompertz, B. 1825. On the nature of the function expressive of human mortality, and on a new mode of determining the value of life contingencies. Phil. Trans. (B) 115: 513-585.

- Gulland, J.A. 1965. Manual of sampling and statistical methods for fisheries biology. Part I. Sampling methods. FAO Man. Fish. Sci. 3. pag. var.
- Gulland, J.A. 1971. The fish resources of the oceans. FAO Fishing News Books, Ltd., Surrey, England.
- Gulland, J.A. 1983. Fish stock assessment. A manual of basic methods, John Wiley and Sons, New York.
- Gulland, J.A. 1984. Letter to the editor: on "best estimates". Fishbyte 2(3): 3-5.
- Harding, J.P. 1949. The use of probability paper for the graphical analysis of polymodal frequency distributions. J. Mar. Biol. Assoc. UK 28: 141-153.
- Heincke, F. 1913. Investigation on the plaice. General Report 1. Plaice fishery and protective regulations. Part I. Rapp. P.-V. Réun. Cons. Int. Explor. Mer 17A: 1-153.
- Hixon, R.F., R.T. Hanlon and W.H. Hulet. 1981. Growth and maximal size of the long-finned squid *Loligo pealei* in the Northwestern Gulf of Mexico. J. Shellfish Res. 1(2): 181-185.
- Hoydal, K., C.J. Rørvik and P. Sparre. 1982. Estimation of effective mesh size and their utilization in assessment. Dana 2: 69-95.
- Ingles, J. and D. Pauly. 1982. Raw data and intermediate results for an atlas on the growth, mortality and recruitment of Philippine Fishes. i-xv + 224 p., ICLARM (mimeo).
- Ingles, J. and D. Pauly. 1984. An atlas of the growth, mortality and recruitment of Philippine fishes. ICLARM Technical Reports 13. Manila, 127 p.
- Jensen, A.L. 1982. Adjusting fishery catch curves for trawl selection using the logistic distribution. J. Cons. Int. Explor. Mer 40(1): 17-20.
- Johannes, R.E. 1978. Reproductive strategies of coastal marine fishes in the tropics. Env. Biol. Fish 3: 65-84.
- Jones, R. 1981. The use of length composition data in fish stock assessment (with notes on VPA and Cohort Analysis) FAO Fish. Circ. 734, Rome, 55 p.
- Krüger, F. 1964. Neuere mathematische Formulierung der biologischen Temperaturfunktion und des Wachstums. Helgoländer Wiss. Meeresunters. 9: 108-124.
- Lea, E. 1911. Report on the international herring investigation during the year 1910. 3. A study of the growth of herrings. Publ. Circonst. Cons. Int. Explor. Mer (61): 35-57.
- Lea, E. 1913. Further studies concerning the methods of calculating the growth of herrings. Publ. Circonst. Cons. Int. Explor. Mer (66): 36 p.
- Longhurst, A. and D. Pauly. 1987. Ecology of tropical oceans. Academic Press, San Diego, Florida.
- Mathews, C.P. 1974. An account of some methods of overcoming errors in ageing tropical and subtropical fish populations when hard tissue growth marking are unreliable and the data sparse, p. 158-166. In T.B. Bagenal (ed.) Ageing of fish. Unwin Brothers, Old Woking, Surrey, England.
- Mathisen, O.A. and S. Olsen. 1968. Yield isopleths of the halibut, *Hippoglossus hippoglossus* in Northern Norway. Fish. Dir. Skr. Ser. Havunders. 14: 129-159.
- Mendo, J. 1983. Bestandsuntersuchung an Peruanischen Makrelen (*Scomber japonicus peruanus*). Kiel University, 73 p. MS Thesis.
- Mendelsohn, R. and J. Mendo. 1987. Exploratory analysis of anchoveta recruitment off Peru and related environmental series, p. 294-306. In D. Pauly and I. Tsukayama (eds.) The Peruvian anchoveta and its upwelling ecosystem: three decades of change. ICLARM Studies and Reviews 15, 351 p. Instituto del Mar del Peru (IMARPE), Callao, Peru, Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ), GmbH, Eschborn, Federal Republic of Germany and International Center for Living Aquatic Resources Management (ICLARM), Manila, Philippines.
- Menon, M.D. 1953. The determination of age and growth of fishes of tropical and subtropical waters. J. Bombay Nat. Hist. Soc. 51(3): 623-635.
- Mesnil, B. 1980. Théorie et pratique de l'analyse des cohortes. Rev. Trav. Inst. Pêches Marit. 44(2): 119-155.
- Morgan, G.R. 1983. Application of length-based stock assessments to Kuwait's fish stocks. ICLARM Newsletter 6(4): 3-4.
- Morgan, G.R. 1985. Stock assessment of the pomfret (*Pampus argenteus*) in Kuwait waters. J. Cons. Int. Explor. Mer 42: 3-10.
- Munro, J.L. 1982. Estimation of biological and fishery parameters in coral reef fisheries, p. 71-82. In D. Pauly and G.I. Murphy (eds.) Theory and management of tropical fisheries. ICLARM Conference Proceedings 9, Manila, 360 p.
- Munro, J.L. 1984. Estimation of natural mortality rates from selectivity and catch length-frequency data. Fishbyte 2(1): 11-14.
- Navaluna, N.A. 1982. Morphometrics, biology and population dynamics of the croaker fish, *Otolithes ruber*, p. 38-55. In D. Pauly and A.N. Mines (eds.) Small-scale fisheries of San Miguel Bay, Philippines: biology and stock assessment. ICLARM Technical Reports 7, Manila, 124 p.
- Nickelson, T.E. and G.L. Larson. 1974. Effect of weight loss on the decrease of length of coastal cutthroat trout. Progr. Fish-Cult. 36(2): 90-91.
- Pagdilao, C.R. 1982. Taxonomy, growth and mortality of slipmouths (Order Perciformes, Family Leiognathidae) in Manila Bay. University of the Philippines, 114 p. MS Thesis.

- Palomares, M.L., P. Muck, J. Mendo, E. Chuman, O. Gomez and D. Pauly. 1987. Growth of the Peruvian anchoveta (*Engraulis ringens*), 1953 to 1982, p. 117-141. In D. Pauly and I. Tsukayama (eds.) The Peruvian anchoveta and its upwelling ecosystem: three decades of change. ICLARM Studies and Reviews 15, 351 p. Instituto del Mar del Peru (IMARPE), Callao, Peru, Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ), GmbH, Eschborn, Federal Republic of Germany and International Center for Living Aquatic Resources Management (ICLARM), Manila, Philippines.
- Pannella, G. 1971. Fish otoliths: daily growth layers and periodical patterns. Science 137(4002): 1124-1127.
- Pauly, D. 1978. A preliminary compilation of fish length growth parameters. Ber. Inst. Meeresk. Univ. Kiel No. 55, 200 p.
- Pauly, D. 1980a. A selection of simple methods for the assessment of tropical fish stocks. FAO Fish. Circ. No. 729.
- Pauly, D. 1980b. On the interrelationships between natural mortality, growth parameters and mean environmental temperature in 175 fish stocks. J. Cons. Int. Explor. Mer 39(3): 175-192.
- Pauly, D. 1981. The relationship between gill surface area and growth performance in fish: a generalization of von Bertalanffy's theory of growth. Meeresforschung/Reports on Marine Research 28(4): 251-282.
- Pauly, D. 1982. Studying single-species dynamics in a multispecies context, p. 33-70. In D. Pauly and G.I. Murphy (eds.) Theory and management of tropical fisheries. ICLARM Conference Proceedings 9, Manila, 360 p.
- Pauly, D. 1983a. Some simple methods for the assessment of tropical fish stocks. FAO Fish. Tech. Pap. No. 234, 52 p.
- Pauly, D. 1983b. Length-converted catch curves: a powerful tool for fisheries research in the tropics (Part I). Fishbyte 1(2): 9-13.
- Pauly, D. 1984a. Fish population dynamics in tropical waters: a manual for use with programmable calculators. ICLARM Studies and Reviews 8, Manila, 325 p.
- Pauly, D. 1984b. Length-converted catch curves: a powerful tool for fisheries research in the tropics (Part II). Fishbyte 2(1): 17-19.
- Pauly, D. 1985a. The population dynamics of short-lived species, with emphasis on squids. NAFO Sci. Coun. Studies 9: 143-154.
- Pauly, D. 1985b. On improving operation and use of the ELEFAN programs. Part I: avoiding "drift" of K toward low values. Fishbyte 3(3): 13-14.
- Pauly, D. 1986. On improving operation and use of ELEFAN programs. Part III: correcting length-frequency data for effects of gear selection and/or incomplete recruitment. Fishbyte 4(9): 11-13.
- Pauly, D. and J. Caddy. 1985. A modification of Battacharya's method for the separation of normal distributions. FAO Fish. Circ. 781, Rome, 16 p. (+ errata sheet)
- Pauly, D. and N. David. 1981. A BASIC program for the objective extraction of growth parameters from length-frequency data. Meeresforsch./Rep. Mar. Res. 28(4): 205-211.
- Pauly, D. and G. Gaschütz. 1979. A simple method for fitting oscillating length growth data, with a program for pocket calculator. ICES CM 1979/G: 24. Demersal Fish Ctee., 26 p.
- Pauly, D. and J. Ingles. 1981. Aspects of the growth and mortality of exploited coral reef fishes, p. 89-98. In E. Gomez, C.E. Birkeland, R.W. Buddemeyer, R.E. Johannes, J.A. Marsch, Jr. and R.T. Tsuda (eds.) Proceedings of the Fourth International Coral Reef Symposium, Manila, Philippines, Vol. 1, 725 p.
- Pauly, D. and N.A. Navaluna. 1983. Monsoon-induced seasonality in the recruitment of Philippine fishes, p. 823-833. In G.D. Sharp and J. Csirke (eds.) Proceedings of the Expert Consultation to Examine Changes in Abundance and Species Composition of Neritic Fish Resources, San José, Costa Rica, 18-29 April 1983. FAO Fish. Rep. No. 291, Vol. 3.
- Pauly, D. and Sann Aung. 1984. Population dynamics of some marine fishes of Burma. BUR/77/003/FAO Field Document No. 7, FAO, Rome. 22 p.
- Pauly, D. and I. Tsukayama. 1983. On the seasonal growth, monthly recruitment and monthly biomass of the Peruvian anchoveta from 1961 to 1979, p. 987-1004. In G.D. Sharp and J. Csirke (eds.) Proceedings of the Expert Consultation to Examine Changes in Abundance and Species Composition of Neritic Fish Resources. San José, Costa Rica, 18-29 April 1981. FAO Fish. Rep. No. 291, Vol. 3.
- Pauly, D., N. David and J. Ingles. 1980. ELEFAN I: user's instructions and program listings. Mimeo, pag. var.
- Pauly, D., J. Ingles and R. Neal. 1984. Application to shrimp stocks of objective methods for the estimation of growth, mortality and recruitment-related parameters from length-frequency data (ELEFAN I and II), p. 220-234. In J.A. Gulland and B.J. Rothschild (eds.) Penaeid shrimps—their biology and management. Fishing News Books, Farnham, Surrey, England. 308 p.
- Petersen, C.G.J. 1891. Eine Methode zur Bestimmung des Alters und des Wuchses der Fische. Mitt. Dtsch. See-fischerei Ver. II: 226-235.
- Pope, J.G. 1972. An investigation of the accuracy of Virtual Population Analysis. Int. Comm. Northwest. Atl. Fish. Res. Bull. 9: 65-74.
- Qasim, S.Z. 1973. Some implication of the problem of age and growth in marine fisheries from the Indian waters. Indian J. Fish. 20(2): 351-371.

- Reck, G.K. 1983. The coastal fisheries in the Galapagos Island, Ecuador: description and consequences for management in the context of marine environmental protection and regional development. Universität Kiel, 231 p. Doctorate Thesis.
- Richards, F.J. 1959. A flexible growth function for empirical use. J. Exp. Bot. 10(29): 290-300.
- Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191. 382 p.
- Rodriguez, A. 1977. Contribution to the knowledge of the biology and fishing of the prawn, *Penaeus kerathurus* (Forskål 1775) from the Cadiz Gulf (Spanish South Atlantic Region). Invest. Pesq. 41(3): 603-637.
- Rohde, J. 1982. Wachstumsuntersuchungen an Nordseefischen. Kiel University, 133 p. MS Thesis.
- Sharp, G. 1980. Report of the Workshop on the effects of environmental variation on survival of larval pelagic fishes, p. 15-59. In G. Sharp (rapporteur). Workshop on the Effects of Environmental Variation on the Survival of Larval Pelagic Fishes. Intergovernmental Oceanographic Commission Workshop Rep. No. 28. 323 p.
- Silas, E.G., M.J. George and T. Jacob. 1984. A review of the shrimp fisheries of India: a scientific basis for the management of the resources, p. 83-103. In J.A. Gulland and B.J. Rothschild (eds.) Penaeid shrimps: their biology and management. Fishing News Books, Farnham, Surrey, England. 308 p.
- Tanaka, S. 1956. A method for analyzing the polymodal frequency distribution and its application to the length distribution of porgy, *Taius tumifrons* (T. & S.). Bull. Tokai Reg. Res. Lab. 14: 1-12.
- Tandog, D.D. 1984. State of exploitation and population dynamics of skipjack tuna (*Katsuwonus pelamis*) in waters off Misamis Oriental. University of the Philippines, College of Fisheries, Diliman, Quezon City. MS Thesis
- Thompson, D.W. 1942. Growth and form. Cambridge University Press.
- Ursin, E. 1973. On the prey size-preference of cod and dab. Meddr. Danm. Fisk. Havunders. N.S. 7: 85-98.
- Weber, W. 1976. The influence of hydrographical factors on the spawning time of tropical fish, p. 269-281. In K. Tiews (ed.) Fisheries resources and their management in South East Asia. West Berlin. 522 p.
- Went, A.E.J. 1972. Seventy years agrowing: a history of the International Council for the Exploration of the Sea, 1902-1972. Rapp. P.-V. Réun. Cons. Int. Explor. Mer 165. 252 p.
- White, T.F. 1982. The Philippine tuna fishery and aspects of the population dynamics of tuna in Philippine waters. South China Sea Fisheries Development and Coordinating Programme, SCS/82/WP/114, Manila, 64 p.
- Yong, M.Y. and R.A. Skillman. 1971. A computer program for analysis of polymodal frequency distributions (ENORMSEP), FORTRAN IV. US Fish. Bull. 73(3): 681.

Estimating Growth and Mortality Parameters by Nonlinear Regression Using Average Size in Catches

T.A. EBERT

*Department of Biology
San Diego State University
San Diego, CA 98182, USA*

Ebert, T.A. 1987. Estimating growth and mortality parameters by nonlinear regression using average size in catches, p. 35-44. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

When recruitment is periodic, average length in samples is smallest in the month of recruitment and largest in the month just before recruitment. The change in average length during a year is a function of growth and mortality parameters, which can be estimated by an iterative nonlinear regression technique, using average length as the dependent variable and time since recruitment as the independent variable. The model presented here uses a constant rate of mortality and the Richards function to model growth. Four parameters can be estimated: Z, the mortality coefficient; K, the growth constant; L_{∞} , asymptotic length; and n, the shape parameter.

Two applications of the method are presented. A simulation demonstrates that parameters used to generate size distributions can be recovered by the iterative technique. Data for flathead sole (*Hippoglossoides elassodon*) show both the sensitivity of the iterative technique to starting values for parameters and the effect of changing the shape parameter, n, on the improvement of the residual sums of squares.

Introduction

Size structure of a population is a record of its recent past history. A size-frequency distribution contains information on individual growth, mortality, recruitment, and within- and between-year variation of these attributes. Previously it was shown that if average length is determined twice a year then, given certain constraints, it is possible to estimate two unknown parameters such as K, the growth constant of the von Bertalanffy equation, and Z, the mortality coefficient (Ebert 1973). The purpose of the present paper is to extend this system of two equations and two unknowns to one in which there are more measurements of average size than unknown parameters to estimate. This is a least-squares problem and not only parameters but also standard deviations for the parameters can be estimated. This note also generalizes the growth equation so that the Richards function

(Richards 1959; Ebert 1980) can be used. The Richards function also is known as the generalized von Bertalanffy equation (Pienaar and Thompson 1973; Pauly 1981).

Methods

The average length of individuals in a population at time t after annual recruitment is

$$\bar{l}_t = \sum_{i=0}^{\omega} N_i l_{t+i} / \sum_{i=0}^{\omega} N_i \quad \dots 1)$$

in which N_i is the number and l_{t+i} is the length of individuals in the i th age class. The value of t is the time since recruitment and is expressed as a fraction of one year. Using t as a fraction of one year is a correction of the manner in which t was previously, and incorrectly, calculated; namely, by dividing months by 11 so $t = 1$ at the month just before recruitment (Ebert 1973).

Because relative numbers rather than absolute densities are used, the index "i" for numbers of individuals, N , does not have to be adjusted by time since annual recruitment.

Number, N_i , and length, l_{t+i} , can be replaced by suitable mortality and growth functions:

$$N_i = N_o e^{-Zi} \quad \dots 2)$$

$$l_{t+i} = L_\infty (1 - b e^{-K(t+i)})^{-n} \quad \dots 3)$$

where

$$b = (L_\infty^{-1/n} - L_R^{-1/n}) / L_\infty^{-1/n}$$

Size at recruitment is defined as L_R and is the mean of the first mode in the size distribution that was gathered on the date assigned $t=0$, which usually is the sample with the lowest mean. The mean of the first mode can be determined by eye or, more rigorously, the technique proposed by Macdonald and Pitcher (1979) or others.

Combining equations (2) and (3) leads to

$$\bar{l}_t = L_\infty (1 - e^{-Z}) \sum_{i=0}^{\omega} e^{-Zi} (1 - b e^{-K(t+i)})^{-n} \quad \dots 4)$$

When n , the shape parameter of the Richards function, is equal to -1 , that is, when the growth function is assumed to be of the von Bertalanffy type, equation (4) is further simplified by making use of the convergence of a series of partial sums with negative exponents:

$$\bar{l}_t = L_\infty (1 - b(e^Z - 1)) (e^{K(1-t)}) / (e^{K+Z} - 1) \quad \dots 5)$$

Frequently, average length of individuals is estimated several times during a year and sometimes several years of data are available. If annual recruitment is confined to a short period of time, such as one month, if growth can be described by the Richards function and mortality by a simple exponential, and if the population has a seasonally stationary age (and size) distribution (that is, $r=0$, neither growing nor declining), then the parameters Z , K , L_∞ , and n can be estimated by a least-squares technique.

Equation (4) is not reducible to linear form. Damm (Part I, this vol.) has linearized a form of equation (5) and has used an iterative technique to estimate K and L_∞ .

There is a standard nonlinear approach that can be used to estimate parameters in equations (4) and (5). Average size is selected as the dependent variable and time since annual recruitment is the independent variable, that is:

$$\bar{L}_t = f(t, Z, K, L_\infty, n) \quad \dots 6)$$

Pennington (1965), Jennrich (1981) and Ralston (1981) were followed in developing the program for iterative estimation of parameters and standard deviations.

Initial guesses must be made for each parameter. These are best estimates plus some unknown errors, E_j :

$$Z_i + E_1 = Z \quad \dots 7)$$

$$K_i + E_2 = K \quad \dots 8)$$

$$L_\infty i + E_3 = L_\infty \quad \dots 9)$$

$$n_i + E_4 = n \quad \dots 10)$$

equation (6) can then be rewritten:

$$\bar{L}_t = f(t, Z_i + E_1, K_i + E_2, L_\infty i + E_3, n_i + E_4) \quad \dots 11)$$

which can be expanded by Taylor's formula:

$$\bar{L}_t = f(t, Z_i, K_i, L_\infty i, n_i) + E_1 \frac{\partial f}{\partial Z} + E_2 \frac{\partial f}{\partial K} + E_3 \frac{\partial f}{\partial L_\infty} + E_4 \frac{\partial f}{\partial n} \quad \dots 12)$$

Using initial estimates of parameters, residuals ($\bar{L}_i - f(t_i, Z_i, K_i, L_\infty i, n_i)$) and partials are calculated for each data pair, (\bar{L}_j, t_j). A matrix, F , is constructed that contains the partials for each data value, t_j :

$$F = \begin{bmatrix} \frac{\partial f(t_1 \dots)}{\partial Z} & \frac{\partial f(t_1 \dots)}{\partial K} & \frac{\partial f(t_1 \dots)}{\partial L_\infty} & \frac{\partial f(t_1 \dots)}{\partial n} \\ \frac{\partial f(t_2 \dots)}{\partial Z} & \frac{\partial f(t_2 \dots)}{\partial K} & \frac{\partial f(t_2 \dots)}{\partial L_\infty} & \frac{\partial f(t_2 \dots)}{\partial n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f(t_j \dots)}{\partial Z} & \frac{\partial f(t_j \dots)}{\partial K} & \frac{\partial f(t_j \dots)}{\partial L_\infty} & \frac{\partial f(t_j \dots)}{\partial n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f(t_N \dots)}{\partial Z} & \frac{\partial f(t_N \dots)}{\partial K} & \frac{\partial f(t_N \dots)}{\partial L_\infty} & \frac{\partial f(t_N \dots)}{\partial n} \end{bmatrix} \quad \dots 13)$$

If \tilde{F} is the transpose of F , e is the vector with errors (E_j) and r is a vector with residuals, then:

$$\tilde{F}Fe = \tilde{Fr} \quad \dots 14)$$

Multiplying \mathbf{F} by its transpose, $\tilde{\mathbf{F}}$ creates a matrix, \mathbf{A} , which is the sum of cross products of the partials and is a matrix of the same order as there are unknown errors, E_i . With $\mathbf{Fr} = \mathbf{b}$, the matrix equation:

$$\mathbf{Ae} = \mathbf{b} \quad \dots 15)$$

provides a system of equations that can be solved for the E_i terms of the vector e . Matrix inversion of \mathbf{A} and multiplication by \mathbf{b} were used to obtain terms in the vector e , which are used to improve the estimates of Z_i , K_i , $L_{\infty i}$, and n_i :

$$Z_{i+1} = E_1 + Z_i \quad \dots 16)$$

$$K_{i+1} = E_2 + K_i \quad \dots 17)$$

$$L_{\infty i+1} = E_3 + L_{\infty i} \quad \dots 18)$$

$$n_{i+1} = E_4 + n_i \quad \dots 19)$$

The new values of Z_{i+1} , K_{i+1} , and n_{i+1} are used to calculate new partials and residuals and the procedure for calculating errors, E_1 , E_2 , E_3 , and E_4 , is repeated.

As the iterative process is continued, the error terms become smaller and smaller and the program can be terminated at some arbitrary point. The point where all error terms were less than or equal to 0.0001 was chosen here as a suitable level of convergence.

The error sums of squares, SS, is calculated by summing the squared residuals from the final iteration. The mean square, MS, has degrees of freedom equal to the number of data values, N, minus the number of parameters that are estimated, M:

$$MS = SS/(N - M) \quad \dots 20)$$

Asymptotic standard deviations of parameters are estimated from the MS and the inverted matrix of cross products, A' , which has diagonal elements $a_{11}, a_{22} \dots a_{mm}$:

$$sd_i = \sqrt{a_{ii} MS} \quad \dots 21)$$

The computer program to perform the nonlinear regression is written in BASIC for a 6502-based microcomputer (Apple II+). A listing and user's instructions for the program are included in Morgan and Pauly (Part II, this vol.); however, any nonlinear regression technique could be used. Other programs include the SIMPLEX method programmed for APPLE II microcomputers by Schnute (1982) or, on a mainframe computer, BMDPAR or BMDP3R (Dixon 1981).

Results

Two examples are used to illustrate how parameters are estimated by nonlinear regression with equations (4) and (5). The examples show the manner in which data are handled and also illustrate certain problems with the method.

Example 1.

A simulation to generate size distributions at various times following recruitment was used. The growth parameters were $n = -1$, $K = 0.32$, $L_{\infty} = 10.0$ and $L_R = 0.05$. The mortality coefficient, Z , was 0.8. A size-age relationship was generated using time since recruitment, $t = 0.2, 0.3, 0.5, 0.7, 0.9$ and 1.0, that is, the Richards function was used to calculate average length for each age class, l_{t+0}, l_{t+1} , etc. A standard deviation was selected for each age class l_{t+i} equal to $0.1(l_{t+i})$. Points on normal curves with means of l_{t+i} were calculated at intervals of 0.1 and areas estimated by the trapezoidal rule. Each area for an age class was reduced using $Z = 0.8$ and equation (2). One hundred

age classes were used in a simulation and all normal curves were added together and then adjusted to equal 100%. This is the model used previously (Ebert 1981) and two simulated size distributions are shown in that reference. Table 1 shows the average size of all individuals in the population at various times after recruitment.

Table 2 shows the results of estimating parameters using equations (4) and (5) and their associated partials, which are presented as an Appendix. In Table 2, section A, the final values are close

Table 1. Simulated average sizes (\bar{l}_t) of an entire population following recruitment at $t = 0$. Growth parameters for the Richards function are: $n = -1$, $K = 0.32$, $L_\infty = 10.0$, and $L_R = 0.05$; the mortality coefficient $Z = 0.80$; standard deviation for each age class, $i = 0.1 (\bar{l}_{t+i})$.

Time (t)	Ave. size (\bar{l}_t)
0.2	2.372
0.3	2.612
0.5	3.070
0.7	3.499
0.9	3.902
1.0	4.094

Table 2. Estimation of growth and survival parameters by nonlinear regression using data in Table 1. Section A shows estimation of all four parameters; section B shows estimation of just three parameters while holding the shape parameter constant at $n = -1.0$.

	A.		B.
	Solution by equation (4) df = 2	Solution by equation (4) df = 3	Solution by equation (5) df = 3
Z =	0.8002	0.7993	0.7994
K =	0.3136	0.3195	0.3195
L_∞ =	10.1031	10.0071	10.0070
n =	-0.9952	-1.0 (fixed)	-1.0 (fixed)
Residual SS =	6.944×10^{-8}	8.3706×10^{-8}	8.3707×10^{-8}
Mean square (MS) =	3.4722×10^{-8}	2.7902×10^{-8}	2.7902×10^{-8}
Standard deviations			
Z =	0.0013	0.0002	0.0002
K =	0.0093	0.0010	0.0010
L_∞ =	0.1545	0.0218	0.0218
n =	0.0074	n.a.	n.a.

to the values used in the simulation but not exactly the same. All estimates are within one standard deviation of the parameters used in the simulation and differences probably can be accounted for by roundoff errors in the simulation (only three places were saved for average sizes), errors associated with the trapezoidal rule for approximating an integral, roundoff errors associated with using a microcomputer with a 6502 microprocessor, and, possibly, errors accumulated by using summations in equation (4) and its associated partials with:

$$\omega = -\ln(.00001)/Z + .5 \quad \dots 21)$$

Table 2, section B, shows the results of estimating just three parameters while holding the shape parameter constant at $n = -1$. This provides a comparison of equations (4) and (5) and is a partial test of the effects of the use of summations in the function and its partials. It is clear that the results differ only at the fourth decimal place. It can be concluded from this that the problems of deviation of the estimates from the parametric values are due to roundoff errors in the simulation or in the computer rather than with errors introduced by using summations in the regression equation (equation 4) and its partials (see Appendix).

Example 2.

Data on flathead sole, *Hippoglossoides elassodon* (Miller and Wellings 1971) were previously analyzed (Ebert 1973) using just two average sizes to estimate two parameters, Z and K. Table 3 shows additional data extracted from their paper and used in the present analysis. Maximum average length was attained during the fall and recruitment took place between October and December. November was selected as the month of recruitment but the average size for this month was not included because it appeared that recruitment was not complete until sometime between November

Table 3. Average sizes (\bar{L}_t) of normal male flathead sole (*Hippoglossoides elassodon*) from East Sound, Orcas Island, Washington, 1962-1967 (data reconstructed from Appendix A in Miller and Wellings 1971). The average size of age-class 0 individuals in November was 93.1 mm, so a reasonable estimate of L_R is 90 mm with November taken as $t = 0$.

Month	Time (t)	Ave. size (\bar{L}_t) (mm)
Dec	0.083	118.2
Jan	0.167	128.3
Feb	0.250	143.5
Mar	0.333	143.8
Apr	0.417	182.4
May	0.500	157.6
Jun	0.583	181.2
Jul	0.667	181.7
Aug	0.750	202.6
Sep	0.833	178.8
Oct	0.917	202.4

and December. Consequently, Table 3 does not have an average size for November, when $t = 0$, and December is assigned a time of $t = 0.08$. Size of recruitment in November was 90 mm.

An attempt to estimate all four parameters failed when starting values were $K = Z = 1 \text{ yr}^{-1}$, $L_\infty = 290 \text{ mm}$ and $n = -1$. The change in the shape parameter n from one iteration to the next caused an increasing amplitude which failed to converge. Estimating all four parameters required selecting different initial values for the parameters. This was done by first fitting three parameters (Z , K , and L_∞) while holding n constant at -1 , observing the residual SS and then repeating the procedure while fixing n at -2 and using initial values for other parameters equal to the final values obtained from the trial using $n = -1$. The residual SS decreased and $n = -25$ was tried. Initial values for other parameters were the final values obtained with $n = -2$. The Richards function (equation (3)) approaches the Gompertz curve as the absolute value of n approaches infinity. Accordingly, the decrease in SS from $n = -2$ to $n = -25$ suggested that positive values of n might further decrease the residual sums of squares. The next trial was for $n = +25$ and then $n = +1$. At $n = +1$, it was again attempted to estimate all four parameters using initial values equal to those obtained as final values while fixing $n = +1$. With these initial values, the system converged and provided estimates for all four parameters (Table 4). The average lengths together with the fitted line are shown in Fig. 1.

It is clear that selection of initial estimates is very important in obtaining convergence as more parameters are estimated.

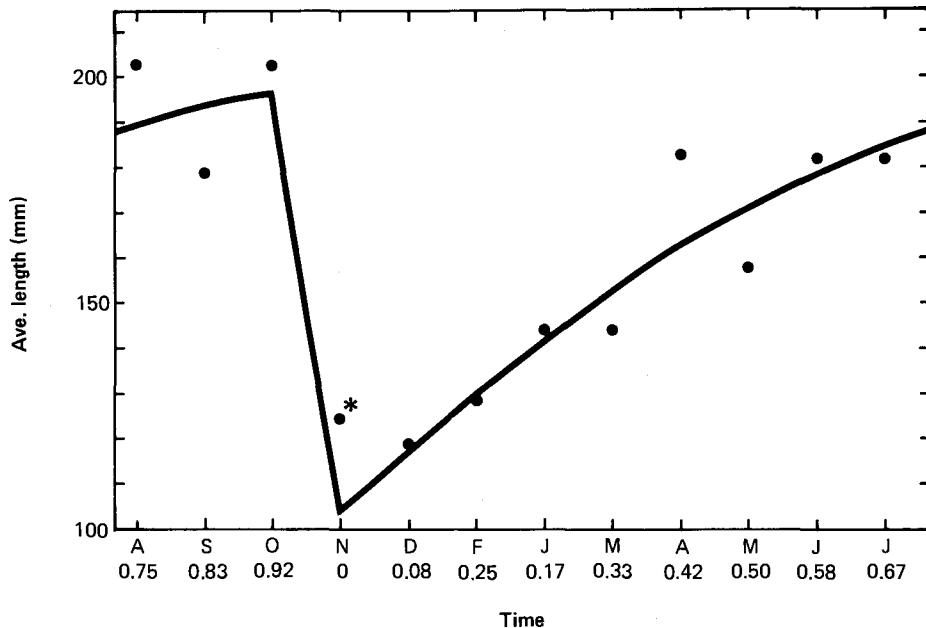


Fig. 1. Change in mean lengths in normal male flathead sole (*Hippoglossoides elassodon*) (data from Miller and Wellings 1971). Fitted line uses $K = 3.39$, $L = 206.6 \text{ mm}$, $Z = 2.01$ and $n = +0.97$; * data point excluded from analysis (see text).

Discussion

Using the models shown in equations (4) and (5), it is possible to obtain estimates of growth and survival parameters from a time progression of average sizes. The nonlinear technique used is sensitive to starting values and for some data sets it is necessary to try a number of initial values before the system stabilizes and converges on final estimates. It is clear that at least four parameters can be estimated and, in principle, the technique can be extended to include other parameters, such as a shape parameter for the mortality curve (Pinder et al. 1978; Caswell 1982) or seasonally varying

parameters (Pauly and David 1981). As shown by Damm (Part I, this vol.) the technique can also be modified to accommodate more than one recruitment episode per year.

Although many parameters can be estimated using the nonlinear technique presented here, it is worth noting the price that is paid for estimating many rather than few parameters with respect to the confidence that can be placed in the estimates. For example in Table 4, the residual SS obtained with $n = +0.974$ is 1,097.45 vs. 1,104.67 with $n = -1.0$ and 1,097.45 with $n = +1.0$. Between the guess of $n = -1.0$ and the estimated value of $n = 0.974$, the residual SS decreased by only 0.6%. Because of the additional loss of one degree of freedom, the estimated standard deviations all increased; sd for Z increased by 3.3%, while for K the increase was 593%(!) and for L_∞ the increase was 25%. Depending upon the nature of the questions being asked of the data, it may not always be worthwhile to try to estimate all four parameters, particularly when the data set is small.

The technique can be used to estimate only one parameter, such as Z , when other parameters have been determined by more conventional methods. The method also would find application in estimating both growth and mortality parameters in cases where lack of time or money prohibit the application of traditional techniques. Such circumstances may include species in benthic surveys or exploratory fisheries and species in very rich biotas, such as tropical fisheries, in which population information is needed for management but resources of time and money are limited.

Table 4. Estimates of growth and survival parameters for normal male flathead sole (*Hippoglossoides elassodon*) using data in Table 3 and taken from Miller and Wellings (1971); $L_R = 90$ mm.

Parameter estimates with different shape parameters					
	-1.0 (fixed)	-25 (fixed)	+25 (fixed)	+1.0 (fixed)	+0.974
Z =	2.441	2.199	2.183	2.019	2.015
K =	1.728	2.479	2.544	3.364	3.387
L_∞ =	223.0	213.3	212.6	206.7	206.6
df =	8	8	8	8	7
SS =	1,104.7	1,099.5	1,099.2	1,097.4	1,097.4
MS =	138.1	137.4	137.4	137.2	156.8
$sd\ n$ =	n.a.	n.a.	n.a.	n.a.	9.365
$sd\ Z$ =	1.705	1.184	1.155	0.892	1.761
$sd\ K$ =	1.260	1.243	1.242	1.238	8.737
$sd\ L_\infty$ =	41.16	26.72	25.92	18.94	51.41

References

- Caswell, H. 1982. Life history theory and the equilibrium status of populations. Am. Nat. 120: 317-339.
 Dixon, W.J., editor. 1981. BMDP statistical software. University of California Press, Berkeley, California.
 Ebert, T.A. 1973. Estimating growth and mortality rates from size data. Oecologia 11: 281-298.
 Ebert, T.A. 1980. Estimating parameters in a flexible growth equation, the Richards function. Can. J. Fish. Aquat. Sci. 37: 687-692.

- Ebert, T.A. 1981. Estimating mortality rate from size structure and growth parameters when recruitment is periodic. *Limnol. Oceanogr.* 26: 764-769.
- Jennrich, R. 1981. Nonlinear regression, p. 290-304. In W.J. Dixon (ed.) BMDP statistical software. University of California Press, Berkeley.
- Macdonald, P.D.M. and 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.
- Miller, B.S. and S.R. Wellings. 1971. Epizootiology of tumors on flathead sole (*Hippoglossoides elassodon*) in East Sound, Orcas Island, Washington. *Trans. Amer. Fish. Soc.* 100: 247-266.
- Pauly, D. 1981. The relationship between gill surface area and growth performance in fish: a generalization of von Bertalanffy's theory of growth. *Meeresforsch.* 28: 251-282.
- Pauly, D. and N. David. 1981. ELEFAN I, a BASIC program for the objective extraction of growth parameters from length-frequency data. *Meeresforsch.* 28: 205-211.
- Pennington, R.H. 1965. Introductory computer methods and numerical analysis. The Macmillan Co., New York.
- Pienaar, L.V. and J.A. Thompson. 1973. Three programs used in population dynamics: WVONB—ALOMA—BHYLD (FORTRAN 1130). *Fish. Res. Board Can. Tech. Rep.* 367. 33 p.
- Pinder, J.E., J.G. Wiener and M.H. Smith. 1978. The Weibull distribution: a new method of summarizing survivorship data. *Ecology* 59: 175-179.
- Ralston, M. 1981. Derivative-free nonlinear regression, p. 305-314. In W.J. Dixon (ed.) BMDP statistical software. University of California Press, Berkeley.
- Richards, F.J. 1959. A flexible growth function for empirical use. *J. Exp. Bot.* 10: 290-300.
- Schnute, J. 1982. A manual for easy nonlinear parameter estimating in fishery research with interactive microcomputer programs. *Can. Tech. Rep. Fish. Aquat. Sci.* No. 1140. 115 p.

Appendix: Partials of equations (4) and (5) with respect to parameters: A) partials for equation (4), which uses a generalized growth equation, the Richards function; B) partials for equation (5) which uses the von Bertalanffy growth function and is a special case of the Richards function where $n = -1$.

A

$$f(t, Z, K, L_{\infty}, n) = L_{\infty} (1 - e^{-Z}) \sum_{i=0}^{\infty} e^{-Zi} (1 - be^{-K(t+i)})^{-n}$$

$$\frac{\partial f}{\partial Z} = L_{\infty} e^{-Z} \left[\sum_{i=0}^{\infty} e^{-Zi} (1 - be^{-K(t+i)})^{-n} - (e^Z - 1) \sum_{i=0}^{\infty} e^{-Zi} (1 - be^{-K(t+i)})^{-n} \right]$$

$$\frac{\partial f}{\partial K} = -L_{\infty} nb(1 - e^{-Z}) \sum_{i=0}^{\infty} e^{-Zi} (1 - be^{-K(t+i)})^{-(n+1)} (t+i) e^{-K(t+i)}$$

$$\frac{\partial f}{\partial L_{\infty}} = \left[(1 - e^{-Z}) \sum_{i=0}^{\infty} e^{-Zi} (1 - be^{-K(t+i)})^{-n} \right] - \frac{L_{\infty}(1 - e^{-Z})}{L_R} \left(\frac{L_{\infty}}{L_R} \right)^{\frac{1}{n}-1} \sum_{i=0}^{\infty} e^{-Zi} (1 - be^{-K(t+i)})^{-(n+1)} e^{-K(t+i)}$$

$$\frac{\partial f}{\partial n} = L_{\infty} (1 - e^{-Z}) \sum_{i=0}^{\infty} e^{-Zi} (1 - be^{-K(t+i)})^{-n} \left[\frac{e^{-K(t+i)}}{n(1 - be^{-K(t+i)})} \ln \left(\frac{L_{\infty}}{L_R} \right) \left(\frac{L_{\infty}}{L_R} \right)^{\frac{1}{n}} - \ln(1 - be^{-K(t+i)}) \right]$$

B

$$f(t, Z, K, L_{\infty}) = L_{\infty} \left(1 - \frac{b(e^Z - 1)(e^{K(1-t)})}{e^{K+Z} - 1} \right)$$

$$\frac{\partial f}{\partial Z} = -L_{\infty} be^{K(1-t)} \left(\frac{e^Z(e^K - 1)}{(e^{K+Z} - 1)^2} \right)$$

$$\frac{\partial f}{\partial K} = \left(L_{\infty} - f(t, Z, K, L_{\infty}) \right) \left(\frac{t(e^{K+Z} - 1) + 1}{e^{K+Z} - 1} \right)$$

$$\frac{\partial f}{\partial L_{\infty}} = 1 - \frac{(e^Z - 1)(e^{K(1-t)})}{e^{K+Z} - 1}$$

Some Modifications of Ebert's Method to Estimate Growth and Mortality Parameters from Average Lengths in a Population

U. DAMM*

*Abteilung Fischereibiologie
Institut für Meereskunde
Düsternbrooker Weg 20
23 Kiel 1
Federal Republic of Germany*

Damm, U. 1987. Some modifications of Ebert's method to estimate growth and mortality parameters from average lengths in a population, p. 45-51. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

Ebert's method for the analysis of average sizes in a population is discussed and its similarity to other methods for fitting a von Bertalanffy growth curve is pointed out. A simple fitting method is proposed. It is shown how this method can be applied to the special case of two recruitments per year; the robustness of this approach is discussed. A BASIC program is presented which can fit the relevant function to a given set of data.

Introduction

Ebert (1973, 1981 and Part I, this vol.), extending an approach by Green (1970), suggested a method to estimate growth and mortality parameters from the increment of average individual size in a population. Saila and Lough (1981) presented equations which considerably simplified the computations needed to implement the method of Ebert (1973).

In this paper, it will be shown that if the regression of average size on time is used, a simple technique of fitting the von Bertalanffy growth curve can be applied. Further, it will be demonstrated that the method is in principle applicable to the special case with two recruitment events per year.

These considerations will be restricted to the case of the von Bertalanffy growth in length without seasonal fluctuations, although the method itself does not depend on a particular form for the growth curve.

*Present address: Institut für Küsten-und Binnenfischerei, Laboratorium Cuxhaven, Bei Der Alten Liebe 1, D-2190 Cuxhaven, Federal Republic of Germany.

The Basic Concept

When recruitment to a population is not constant but occurs in discrete groups (cohorts), one may imagine the average length of individuals in the population at a given time as being the average of the lengths of all cohorts which are present in the population, weighted by their absolute or relative abundance.

Length and abundance at time are determined by the growth and mortality functions, respectively. Mean size increases between recruitment events because of individual growth, and decreases more or less sharply when a new cohort is recruited. A plot of average size on time should therefore look like a sawtooth function. The "teeth" of the curve will be sharper, the shorter the recruitment period is.

The basic idea of the method is that when average sizes within one recruitment cycle are known, these can be used to determine the parameters of growth and mortality. It is a necessary assumption that growth and mortality parameters are constant with size-at-age and that the amount and timing of recruitment are constant between cohorts. In order to estimate mortality, size-at-age for cohorts (or individuals) must be known.

Some practical constraints are that the population studied should have a short recruitment period, and a correspondingly long phase of undisturbed increment of average size, and that growth between recruitment events should cover a relatively broad size range, which tends to give more precise regression estimates.

The Function

Let l_t , the length at time t of an individual or the average length of a cohort, follow a von Bertalanffy function

$$l_t = L_\infty (1 - e^{-K(t - t_0)}) \quad \dots .1)$$

and let N_t , the number of individuals of a cohort at time t follow an exponential decay function

$$N_t = N_0 e^{-Zt} \quad \dots .2)$$

then \bar{l}_t , the average length in a population at time t is given by

$$\bar{l}_t = L_\infty - (L_\infty - L_R) \frac{\sum_{x=0}^N e^{-(K+Z)x}}{\sum_{x=0}^N e^{-Zx}} \cdot e^{-Kt} \quad \dots .3)$$

which when rearranged, is equation (5) of Ebert (1973), where N is the oldest age class and L_∞ , K and Z are parameters of the growth and mortality functions, respectively. N_0 has dropped out, and t_0 is functionally replaced by L_R , "length at recruitment", which is the average length of the youngest cohort in the population at $t = 0$.

The parameter L_R is the same as L_0 in the original formulation of the growth function by von Bertalanffy (1934),

$$l_t = L_\infty - (L_\infty - L_0) e^{-Kt} \quad \dots .4)$$

If any information is available on length-at-age (preferably at low ages) of cohorts or individuals, this may be used to derive a value of L_R (see also Ebert, Part I, this vol.).

The value of t in equation (3) is the time which has passed since $t = 0$, which one may set at the time of recruitment, but this is arbitrary. It is important, however, that $t = 0$ be consistent with the L_R value and that length increases from then on only. The unit of time is arbitrary as well, but has to be consistent with equation (3) where a constant time interval of one between recruitments is assumed. With annual recruitment, t should be counted in fractions of a year.

When the summations in equation (3) can be done over an infinite number of cohorts (which in practice means that all cohorts are equally represented in the sample), one can use the relation

$$\sum_{i=0}^{\infty} e^{-Z_i} = \frac{e^Z}{e^Z - 1} \quad \dots 5)$$

(Ebert 1981). Then equation (3) can be written as

$$\bar{l}_t = L_\infty - (L_\infty - L_R) \frac{e^K + Z - e^K}{e^K + Z - 1} \cdot e^{-Kt} \quad \dots 6)$$

A Fitting Method

Note that equation (6), which is the regression of \bar{l}_t on e^{-Kt} , takes virtually the same form as the basic equation which Allen (1966) proposed to fit the von Bertalanffy function,

$$l_t = a + b'e^{-Kt} \quad \dots 7)$$

with $a = L_\infty$ and $b' = -(L_\infty - L_R)e^{Kt_0}$.

It can be seen from equation (4) that $b = L_\infty - L_0 = L_\infty - L_R$ is different from "b" as used by Ebert (1973, 1981) which is equal to $(L_\infty - L_r)/L_\infty$.

With an appropriate value of K , equation (6) is a straight-line relationship between \bar{l}_t and e^{-Kt} with intercept L_∞ and slope

$$b = -(L_\infty - L_R) \frac{e^K + Z - e^K}{e^K + Z - 1} \quad \dots 8)$$

This can be solved for Z , yielding

$$Z = \ln \left(\frac{b + e^K (L_\infty - L_R)}{b + L_\infty - L_R} \right) - K \quad \dots 9)$$

Consequently, when individuals grow according to a von Bertalanffy function, the average size in a population will also follow such a curve, the shape of which is completely determined by K and L_∞ , while the level of average sizes is determined by Z . Once K and L_∞ have been estimated, the estimate of Z depends only on L_R . If some value of the latter can be derived, Z can be estimated,

otherwise not. Knowing K and L_{∞} , any pair of values l_t and t can be used to compute L_R , using equation (4),

$$L_R = L_o = L_{\infty} - (L_{\infty} - l_t)e^{Kt} \quad \dots 10)$$

A fitting method can be used which searches for the value of K that gives the best straight-line relationship for equation (6) or (7) and a set of values of l and t, such as the Newton method which Allen (1966) uses in his algorithm. A simpler way might be to use a trial and error routine to find the best value of K, by minimizing the sum of squares of deviations. Once the optimum value of K is found, the intercept of the regression line gives L_{∞} and the slope yields Z for a given L_R . Such a routine could also be implemented on a pocket calculator. Where the program with Allen's (1966) method is at hand, it should be easy to adapt it to accept a value of L_R and give Z in place of t_o . Alternatively, Z could be calculated by hand, using equation (9) with $-L_{\infty} e^{Kt_o}$ instead of b, according to equation (7). It should be noted however, that calculating Z from a t_o value is only possible when the growth curve was fitted to average size in the whole population.

Allen's (1966) fitting method was tried on one of Ebert's (1973 and Part I, this vol.) examples, i.e., on the data for flathead sole ("normal males") which stem from Miller and Wellings (1971). The \bar{l}_t values had to be recalculated, because Ebert (1973) did not split up the juveniles by sexes. When weighting the length values by sample size, the following results were obtained:

$$\begin{array}{lll} L_{\infty} & = 266.3 & \text{var}(L_{\infty}) = 17656 \\ K & = 0.8501 & \text{var}(K) = 0.3747 \end{array}$$

$$Z = .6567 \text{ for } L_R = 90, t = 0 \text{ in November}$$

$$R^2 = 0.773$$

Fitting a von Bertalanffy curve to the aged sample gives $L_{\infty} = 269.7$, $K = 0.7281$, $t_o = -0.5141$. This looks like a good agreement, but the large variances indicate that this is partly coincidence. On another data set from the same source (flathead sole, normal females) the algorithm did not converge to positive values of K.

Note that in calculating these examples, values have been included from times of the year when recruitment was still (or already) going on, which, strictly speaking is not allowed.

The Special Case of Two Recruitments Per Year

Pauly and Navaluna (1983) identify a pattern of two recruitments per year, with different magnitudes of recruitment, as a typical case for fishes in Philippine waters. With some modifications, the method discussed here may be used in such a situation.

If it can be assumed that the relative strength of the two recruitment pulses remain consistent over the years, then one may imagine the total population to be composed of two subpopulations (say A and B), recruiting at different times. The average size in the total population would then be the weighted mean of the average sizes within subpopulations,

$$\bar{l}_t = p \cdot \bar{l}_{t_A} + (1 - p) \bar{l}_{t_B} \quad \dots 10)$$

where \bar{l}_{t_A} and \bar{l}_{t_B} is the average length in the respective subpopulation; p and $1 - p$ are weighting factors, indicating the relative amount of each recruitment, and thus also the average fraction of each group in the total population. The parameter \bar{l}_t , finally, is the resulting average length in the total population. Times t_A and t_B are counted from their respective zero origin (say, the times of recruitment), with a phase shift between t_A and t_B ; t is the time scale for the total population; for convenience, it will be set equal to t_A .

Average lengths within subpopulations are:

$$\bar{l}_{t_A} = L_\infty + b e^{-K t_A} \quad \dots 11)$$

$$\bar{l}_{t_B} = L_\infty + b e^{-K t_B} \quad \dots 12)$$

with b as specified in equation (8).

Equation (10) may be written as

$$\bar{l}_t = p \bar{l}_{t_A} + (1-p) \bar{l}_{t_B} = p (\bar{l}_{t_A} - \bar{l}_{t_B}) + \bar{l}_{t_B} \quad \dots 13)$$

$$\bar{l}_{t_A} - \bar{l}_{t_B} = L_\infty + b \cdot e^{-K t_A} - L_\infty - b e^{-K t_B} = b (e^{-K t_A} - e^{-K t_B}) \quad \dots 14)$$

thus

$$\bar{l}_t = L_\infty + pb (e^{-K t_A} - e^{-K t_B}) + b e^{-K t_B} \quad \dots 15)$$

This is obviously the form of a multiple linear regression, i.e.,

$$y = b_0 + b_1 x_1 + b_2 x_2 \quad \dots 16)$$

with

$$\begin{aligned} y &= \bar{l}_t \\ b_0 &= L_\infty \\ b_1 &= pb \\ x_1 &= e^{-K t_A} - e^{-K t_B} \\ b_2 &= b = \frac{(L_\infty - L_R) (e^{K+Z} - e^K)}{e^{K+Z} - 1} \end{aligned}$$

and

$$x_2 = e^{-K t_B}$$

Again the function may be fitted by searching for the value of K which gives the best linear relationship of equation (15) with a set of data. In the program "E2", documented in Morgan and Pauly (Part II, this vol.), this is performed by a simple direct search algorithm. Standard procedures of multiple linear regression are applied for calculating the sum of squares of deviations to be minimized, and once the minimum is found (which indicates the optimum value of K), the intercept gives L_∞ , while Z is estimated from b_2 , and $p = b_1/b_2$.

The program has the option of either fitting the total parameter set (K , L_∞ , Z and p) to the data, or only K , Z and p for a given L_∞ , or only Z and p for given K and L_∞ values. In the first case, the program searches for the K value that gives the best sum of squares, in the second case it searches for a K which produces an L_∞ value equal to the input value. In the third case, no iteration is necessary, since a solution can be obtained directly.

In order to find data to illustrate the method, Ingles and Pauly's (1982) compilation of length-frequency data was searched, since it had been the base of their study on recruitment patterns. It turned out, however, that no data set could be found which displayed the characteristic double sawtooth pattern discussed above.

Therefore, synthetic data were constructed with $L_\infty = 100$, $L_R = 10$, $p = 0.7$, and a time lag of 5 months between recruitments. A simulated error (C.V. of $\bar{l}_t = 10\%$) was added to one of the two sets (Table 1).

Data set I (see Table 1) was found to be very sensitive to error; the K and L_{∞} values with which it had been constructed could not be recovered when the length values were rounded without decimals. Values for Z and p, however, were very close to the expected values. The entire size range covered by this synthetic data set was only about 15% of the total possible range (0-100), so one should not expect good estimates of growth from such data.

Data set II, which spans a larger size range, was less sensitive, but adding an error term to the values prevented the program from converging onto a positive value of K. The amount of error—a standard deviation of 10% of the lengths—is arbitrary, but chosen to be realistic. Using the option of a fit with a given L_{∞} value gave the results in Table 2.

Table 1. Synthetic data sets used to illustrate the new method proposed here (see text).

	Data sets			
	I	II	III	
	K = 0.5	K = 1.0	K = 1.0	
	Z = 0.6	Z = 1.2	Z = 1.2	
t (months)		\bar{l}_t	\bar{l}_t^*	
0.5	44.91	41.15	45.85	1st recruitment
1.5	47.16	45.86	50.05	
2.5	49.31	50.19	44.29	
3.5	51.38	54.17	52.23	
4.5	53.37	57.83	60.02	2nd recruitment
5.5	48.23	48.34	42.35	
6.5	50.35	52.47	55.27	
7.5	52.37	56.27	43.97	
8.5	54.32	59.77	55.68	
9.5	56.18	62.98	61.88	
10.5	57.97	65.94	64.06	
11.5	59.68	68.67	67.33	

* $\bar{l}_t +$ random error with S.D. = $0.1 \cdot l_t$

Table 2. Results of analysis of data set II in Table 1.

L_{∞}	Input		
	K	Z	p
75	2.11	1.48	0.58
100	0.97	1.27	0.60
150	0.46	1.19	0.61

It was found that the parameter K fluctuated much with L_{∞} and that p remained biased, but that the values of Z that were recovered were reasonably close to the “true” value of 1.2. When the third option was selected—forcing K = 1 and $L_{\infty} = 100$, values of Z = 1.21 and p = 0.60 were recovered. Thus, when a set of good growth parameters is available, one might be able to obtain reasonable values of Z and of p when the growth parameters cannot be estimated from the data because of sampling error.

Discussion

The case of more than one recruitment event per year poses particular problems for Ebert's method. First, when recruitment can occur more than once per year, there is the chance that it might be somewhat unpredictable in time and strength, or that the recruiting seasons might not be as sharply marked as would be necessary.

Second, when there is little time between recruitment events, the increase in average size between these events will be rather small, hence resulting in regression estimates of little precision. The trials with synthetic data, though very few, gave some hope that under certain conditions (sufficiently high values of K and Z and well-marked recruitment seasons), reasonable estimates of Z could be obtained. It is disappointing, however, that real-life data to which this method could have been applied could not be found. This may partly be due to what I feel is a particular and very nasty source of sampling error: the clustering of sizes in length samples, such that entire size groups may be over- or underrepresented. This feature, unfortunately may not be uncommon, being caused, e.g., by unequal distribution of size groups with depth.

In such a case, only a very thoroughly planned sampling scheme, which covers the habitats of all size groups and allows the derivation of stratified estimates, might overcome the problems. Conversely, there is little hope that much can be concluded from sporadic or exploratory length-frequency samples, whatever the method of analysis.

References

- Allen, K.R. 1966. A method of fitting growth curves of the von Bertalanffy type to observed data. *J. Fish. Res. Board Can.* 23(2): 163-179.
- Bertalanffy, L.V. 1934. Untersuchungen über die Gesetzmäßigkeit des Wachstums I. *Roux' Arch.* 131: 613-652.
- Ebert, T.A. 1973. Estimating growth and mortality rates from size data. *Oecologia (Berl.)* 11: 281-298.
- Ebert, T.A. 1981. Estimating mortality from growth parameters and a size distribution when recruitment is periodic. *Limnol. Oceanogr.* 26(4): 764-769.
- Green, R.H. 1970. Graphical estimation of rates of mortality and growth. *J. Fish. Res. Board Can.* 27: 204-208.
- Ingles, J. and D. Pauly. 1982. Raw data and intermediate results for an atlas of the growth, mortality and recruitment of Philippine fishes, i-xv + 224 p. ICLARM (mimeo).
- Miller, B.S. and S.R. Wellings. 1971. Epizootiology of tumors on flathead sole (*Hippoglossoides elassodon*) in East Sound, Orcas Island, Washington. *Trans. Amer. Fish. Soc.* 100: 247-266.
- Pauly, D. and N.A. Navaluna. 1984. Monsoon-induced seasonality in the recruitment of Philippine fishes, p. 823-833. In G.D. Sharp and J. Csirke (eds.) *Proceedings of the Expert Consultation to Examine Changes in Abundance and Species Composition of Neritic Fish Resources*, San José, Costa Rica, 18-29 April 1983. FAO Fish. Rep. No. 291, Vol. 3.
- Saila, S.B. and R.G. Lough. 1981. Mortality and growth estimation from size data—an application to some Atlantic herring larvae. *Rapp. P.-V. Réun. Cons. Int. Explor. Mer* 178: 7-14.

Estimating Growth and Mortality in Steady-State Fish Stocks from Length-Frequency Data

J.A. WETHERALL
J.J. POLOVINA
S. RALSTON

*Southwest Fisheries Center, Honolulu Laboratory
National Marine Fisheries Service, NOAA
2570 Dole St., Honolulu, Hawaii 96822-2396
USA*

Wetherall, J.A., J.J. Polovina and S. Ralston. 1987. Estimating growth and mortality in steady-state fish stocks from length-frequency data, p. 53-74. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

Methods of using length-frequency statistics to estimate L_∞ and the ratio Z/K in steady-state fish stocks with von Bertalanffy growth and exponential mortality are studied. Several standard procedures applicable when mortality is constant are reviewed, and new methods are introduced which have superior performance. The various methods are evaluated using Monte Carlo techniques. Problems of systematic bias are discussed, and remedial measures are suggested. Finally, a method is developed (but not yet evaluated) to estimate length-specific mortality rates under the steady-state model.

Introduction

The current revival of interest in length-frequency based stock assessment methods has led to several new techniques which provide fisheries biologists with alternatives to well-established procedures or allow analysis in situations where traditional age-based methods are infeasible. Because age estimation is often difficult and usually expensive, length-frequency based methods may be the most efficient and reliable means for estimating some of the key parameters of fishery models (see Mathews, Part I, this vol.).

Most of the new length-based procedures, like the graphical and manual methods they are designed to replace, are applicable to situations where spawning is periodic and the identity of cohorts is adequately maintained in the population's length-frequency distribution. Under such circumstances, it is possible to estimate jointly parameters of growth, mortality and recruitment by fitting composite models to single length-frequency samples (e.g., Schnute and Fournier 1980), or to sequences of samples taken over time (e.g., Pauly and David 1981). Estimation and hypothesis testing are facilitated by adopting specific structural assumptions on the underlying processes and

then maximizing likelihood functions or fitting model expectations to observed length class frequencies by least squares or other criteria. Solutions are found by iterative search techniques which take advantage of, and indeed require, the numerical power of computers.

Less attention has been paid lately to the relatively simple methods applicable in situations where cohorts are not clearly distinguishable in catch length distributions, individual length-frequency samples are too small to be treated separately, first approximations are adequate, or sophisticated computer systems and optimization software are unavailable. Among these simple methods, probably the best known is the formula due to Beverton and Holt (1956):

$$Z = K \left(\frac{L_{\infty} - \bar{\ell}}{\bar{\ell} - L_c} \right) \quad \dots 1)$$

which estimates the total instantaneous mortality coefficient, Z , in a steady-state population with constant exponential mortality and von Bertalanffy growth, from the mean length, $\bar{\ell}$, in a random sample of fish above length L_c . In using (1), it is assumed that L_c and the von Bertalanffy parameters, K and L_{∞} , are given. If only L_c and L_{∞} are known (or estimated from other information), slight rearrangement of (1) yields an estimate of the ratio $\theta = Z/K$, a component of many standard yield models and stock assessment procedures (see Gulland 1983).

The Beverton-Holt formula represents a particular class of methods for computing analytical estimates of growth and mortality parameters on the basis of length-frequency data, procedures requiring rather strong assumptions but whose use is justified by their simplicity and their robustness under variable recruitment. This paper reviews three such methods reported in the literature, and introduces some new ones which require more computation but the same amount, or less, input information. First, the assumptions underlying the steady-state methods are given and a structure for length-frequency samples is described. Next, estimators based on the assumption of constant mortality rate are discussed, followed by a consideration of systematic biases. Next, the assumption of constant mortality is relaxed and methods for estimating length-specific mortality are developed. Finally, some Monte Carlo experiments on the constant-mortality methods are described; the concluding section offers advice on use of the various procedures.

Basic Assumptions

PROBABILITY DENSITY OF FISH LENGTHS

In all the methods treated here, a steady-state population is assumed. Growth is assumed to follow the deterministic von Bertalanffy curve with parameters K and L_{∞} . The curve's location parameter, t_0 , usually included in modelling length at age, is omitted since it does not influence length-frequency distributions. Unless otherwise stated, mortality of fish above a knife-edge selection size, L_c , is assumed to occur at a constant instantaneous rate, Z , and recruitment to this size to take place at a constant instantaneous rate, R .

Under these conditions, the probability density ($g(\ell)$) of fish length in the sampled stock is

$$g(\ell) = \frac{\theta (L_{\infty} - \ell)^{\theta - 1}}{(L_{\infty} - L_c)^{\theta}} \quad L_c \leq \ell < L_{\infty} \quad \dots 2)$$

On integrating (2), the cumulative length-distribution function ($G(\ell)$) is seen to be

$$G(\ell) = 1 - \left(\frac{L_\infty - \ell}{L_\infty - L_c} \right)^\theta \quad \dots 3)$$

$$L_c \leq \ell < L_\infty$$

and the mean and variance of length for fish above length L_c are

$$E(\ell) = L_\infty - \left(\frac{\theta}{\theta + 1} \right) (L_\infty - L_c) \quad \dots 4)$$

and

$$V(\ell) = \left(\frac{\theta}{\theta + 2} \right) \left(\frac{L_\infty - L_c}{\theta + 1} \right)^2 \quad \dots 5)$$

A convenient ancillary variable may be defined as $X = \ell/L_\infty$, the actual length of a fish in proportion to its maximum possible length. By changing variables in (2) it is seen that the density for X is

$$h(X) = \frac{\theta (1-X)^{\theta-1}}{(1-X_c)^\theta} \quad \dots 6)$$

$$X_c \leq X < 1$$

where $X_c = L_c/L_\infty$, and its distribution function is

$$H(X) = 1 - \left(\frac{1-X}{1-X_c} \right)^\theta \quad \dots 7)$$

SAMPLING SCHEME

It is assumed that a random sample of n fish is taken from $g(\ell)$ above the minimum catchable size, L_c , with lengths $\ell_1, \ell_2, \dots, \ell_n$. We assume the lengths are observed without error. In practice this assumption is usually violated, but the consequences may not be serious unless measurement errors are very large. We also assume, more critically, that there is no systematic measurement bias.

Measurement error aside, observed fish lengths are recorded with various degrees of "accuracy." As a result, sample length-frequency distributions range from the very sparse (multiple observations of a given length being uncommon in samples of moderate size) to the highly aggregated (the data

being grouped into relatively few length intervals, each containing relatively many observations). To describe each data set we adopt a general scheme wherein the $n.$ sample observations are grouped into $r > 2$ contiguous length intervals, each fish in a particular interval being assigned a length equal to the interval midpoint. Let \tilde{l}_i denote the midpoint of the i -th interval of width Δ_i ($i = 1, 2, \dots, r$). We define L_{\min} as the lower bound of the first length interval and L_{\max} as the upper bound of the last interval, i.e., $L_{\min} = \tilde{l}_1 - \Delta_1/2$, and $L_{\max} = \tilde{l}_r + \Delta_r/2$. A data set consists of a set of length interval midpoints $\{\tilde{l}_1, \tilde{l}_2, \dots, \tilde{l}_r\}$ and a corresponding set of frequencies $\{n_1, n_2, \dots, n_r\}$. (Note that the interval frequencies sum to the total sample size, $n.$)

Estimators

The procedures used to estimate growth and mortality parameters from length-frequency samples differ in their statistical properties, the assumptions they require, and the amount of information they produce. In this section, assuming constant mortality, we review three established methods and explore several new approaches. We first examine some methods applicable when both L_c and L_∞ are known and only θ is to be estimated. Then we consider the joint estimation of θ and L_∞ when only L_c is given.

BEVERTON-HOLT METHOD

The classical formula for $\hat{\theta}$, given by rearranging (1), is

$$\hat{\theta}_{BH} = \frac{L_\infty - \bar{l}}{\bar{l} - L_c} \quad \dots 8)$$

It was derived in a non-probabilistic way by Beverton and Holt (1956), from an expression for mean length in the catch. However, it is readily seen that $\hat{\theta}_{BH}$ is identical to the moment estimator based formally on the density at (2) and the resulting expected length (4). Further analysis with Taylor series shows that $\hat{\theta}_{BH}$ has a statistical bias approximately equal to

$$\text{Bias}(\hat{\theta}_{BH}) = \frac{\theta(\theta+1)}{n.(\theta+2)} \quad \dots 9)$$

and a large-sample variance of

$$\text{Var}(\hat{\theta}_{BH}) = \frac{\theta(\theta+1)^2}{n.(\theta+2)} \quad \dots 10)$$

In constructing $\hat{\theta}_{BH}$, L_∞ and L_c were assumed known. However, this is never true, so in applying $\hat{\theta}_{BH}$ independent estimates of these parameters, or simply educated guesses, must be substituted. The statistical properties of $\hat{\theta}_{BH}$ are consequently altered, so that (9) and (10) no longer apply. This problem of systematic bias will be discussed in a later section.

SSENTONGO-LARKIN METHOD

Ssentongo and Larkin (1973) also assumed L_c and L_∞ were known. They developed an estimator for θ by first deriving a moment estimator for Z , based on the mean age of fish in the sample, and then, given the assumed relationship between length and age, changing variables. The resulting expression, assuming fish lengths are measured and recorded exactly, is

$$\hat{\theta}_{SL} = (\bar{y} - y_c)^{-1} \quad \dots 11)$$

where

$$\bar{y} = -\frac{1}{n} \sum_{i=1}^n \ln \left(1 - \frac{l_i}{L_\infty} \right) = -\frac{1}{n} \sum_{i=1}^n \ln (1 - X_i)$$

and

$$y_c = -\ln (1 - X_c)$$

It may readily be shown that $\hat{\theta}_{SL}$ is also the maximum likelihood estimator based on the density at (6), and is a special case of a more general maximum likelihood procedure discussed later.

As Ssentongo and Larkin report, $\hat{\theta}_{SL}$ has a statistical bias of

$$\text{Bias}(\hat{\theta}_{SL}) = \frac{\theta}{n} \quad \dots 12)$$

and a large-sample variance of

$$\text{Var}(\hat{\theta}_{SL}) = \frac{\theta^2}{n} \quad \dots 13)$$

Thus the bias of $\hat{\theta}_{SL}$ is greater than that of $\hat{\theta}_{BH}$, but its asymptotic variance is smaller, as expected.

As in the Beverton-Holt estimator, the properties of $\hat{\theta}_{SL}$ are altered in practice, since L_∞ and L_c must invariably be estimated. Note further that the estimate of L_∞ must exceed L_{max} ; we return to this point later.

POWELL METHOD

Powell (1979) considered the problem of estimating both θ and L_∞ in a more general context, in which the asymptotic lengths of fish in a population are regarded as random variables, λ , with expectation L_∞ and variance σ_λ^2 . Of various estimation schemes suggested by Powell, the most

interesting is based on use of the first and second moments of the resulting probability distribution of ℓ . Manipulation of Powell's results (Equations (4) and (6) in his paper) reveals that

$$E(\ell) = L_\infty - \left(\frac{\theta}{\theta + 1} \right) (L_\infty - L_c) \quad \dots 14)$$

and

$$V(\ell) = \left(\frac{1}{\theta + 2} \right) \left\{ \theta \left(\frac{L_\infty - L_c}{\theta + 1} \right)^2 + 2\sigma_\lambda^2 \right\} \quad \dots 15)$$

We suppose L_c is given. Then, provided σ_λ^2 is known, or the range of λ is negligible compared with the expected maximum range of ℓ in the data (in which case set $\sigma_\lambda^2 = 0$), $E(\ell)$ and $V(\ell)$ can be equated to the corresponding sample statistics and solved for θ and L_∞ . Let $\bar{\ell}$ denote the sample mean length for fish above L_c , and S_ℓ^2 the sample variance. Then the resulting moment estimators of θ and L_∞ (when $\sigma_\lambda^2 = 0$) are

$$\hat{\theta}_P = \frac{2S_\ell^2}{(\bar{\ell} - L_c) - S_\ell^2} \quad \dots 16)$$

and

$$\hat{L}_{\infty P} = \bar{\ell} + \frac{2S_\ell^2(\bar{\ell} - L_c)}{(\bar{\ell} - L_c)^2 - S_\ell^2} \quad \dots 17)$$

$$= \bar{\ell} + \hat{\theta}_P (\bar{\ell} - L_c) \quad .$$

Since Powell's method makes use of additional information in the sample to estimate L_∞ , $\hat{\theta}_P$ is a considerable improvement over $\hat{\theta}_{BH}$ and $\hat{\theta}_{SL}$ when accurate, independent information on L_∞ is unavailable.

REGRESSION METHOD

Another method for estimating θ and L_∞ jointly may be developed by considering the equation at (4), and exploiting the fact that $E(\ell)$ is a linear function of L_c . Let $\{w_1, w_2, \dots, w_m\}$ be an increasing sequence of fish lengths in the interval (L_c, L_∞) . For a random sample of lengths in

this interval, let $\bar{\ell}_j$ denote the average length of the n_j fish in the sample whose length exceeds w_j . It follows from (4) that

$$E(\bar{\ell}_j) = L_\infty \left(\frac{1}{1 + \theta} \right) + w_j \left(\frac{\theta}{1 + \theta} \right) \quad \dots 18)$$

$$j = 1, 2, \dots, m$$

i.e., $E(\bar{\ell}_j)$ is linear in w_j . The idea of the method is to partition the length-frequency sample using a specified sequence $\{w\}$ and estimate α and β in the linear regression model

$$\bar{\ell}_j = \alpha + \beta w_j + \varepsilon_j \quad \dots 19)$$

where the ε_j are random errors with zero mean and covariance matrix A. Minimum variance unbiased estimates of α and β are found by weighted least squares, using the weight matrix A^{-1} . An estimate of A may be computed from the sample statistics, the (i, j) -th element being

$$A_{ij} = \text{Cov}(\bar{\ell}_i, \bar{\ell}_j) = \sigma_j^2 / n_i \quad \dots 20)$$

where σ_j^2 is the variance among the n_j lengths in the smaller, "included" sample, i.e., the variance among the common elements in the i -th and j -th partitions.

Consistent estimates of θ and L_∞ are then provided by

$$\hat{\theta}_R = \frac{\hat{\beta}}{1 - \hat{\beta}} \quad \dots 21)$$

and

$$\hat{L}_\infty R = \frac{\hat{\alpha}}{1 - \hat{\beta}} \quad \dots 22)$$

Although the choice of divisions in the length-frequency distribution is somewhat arbitrary, there is undoubtedly some optimal partitioning. To a degree, precision in parameter estimates is increased by creating more partitions (data points). One possibility is to let the observed length interval midpoints define a set of partitions, resulting in r data points for the regression, namely,

$$\omega_j = \tilde{\ell}_j - \frac{\Delta_j}{2} \quad \dots 23)$$

and

$$\bar{\ell}_j = \frac{\sum_{i=j}^r \tilde{\ell}_i n_i}{\sum_{i=j}^r n_i} \quad j = 1, 2, \dots, r.$$

The drawback to this approach is that a large weight matrix must be inverted. The task is greatly simplified when fewer data points are created. In particular, if the divisions are not close together the covariances among the partitions are reduced. Acceptable results can then be obtained by ignoring the off-diagonal elements of A , and weighting the mean lengths by the reciprocals of their variances, or simply by the partition sample sizes.

Asymptotic variances and covariances among the estimates of θ and L_∞ are approximately

$$\text{Var}(\hat{\theta}_R) = \frac{\sigma_{\hat{\beta}}^2}{(1-\beta)^4} \quad \dots 24)$$

$$\text{Var}(\hat{L}_\infty R) = \frac{1}{(1-\beta)^2} \left\{ \sigma_{\hat{\alpha}}^2 + \left(\frac{\alpha}{1-\beta} \right)^2 \sigma_{\hat{\beta}}^2 + \left(\frac{2\alpha}{1-\beta} \right) \sigma_{\hat{\alpha}}, \hat{\beta} \right\} \quad \dots 25)$$

and

$$\text{Cov}(\hat{\theta}_R, \hat{L}_\infty R) = \frac{\alpha}{(1-\beta)^4} \sigma_{\hat{\beta}}^2 + \frac{1}{(1-\beta)^3} \sigma_{\hat{\alpha}}, \hat{\beta} \quad \dots 26)$$

Another linear regression approach has been developed by Jones (1981). From (3), observe that the probability of a fish length exceeding w is

$$P(w) = 1 - G(w) \quad \dots 27)$$

$$= \left(\frac{L_\infty - w}{L_\infty - L_c} \right)^\theta \quad .$$

This may be estimated empirically by $p(w)$, the proportion of fish in the sample whose length exceeds w . Jones, assuming L_c and L_∞ are known, suggests plotting $\ln P(w_i)$ vs. $\ln(L_\infty - w_i)$, $i = 1, 2, \dots, m$ and estimating θ as the slope of a regression line fit to these points. We note that if this regression is fitted by least squares, an appropriate weighting scheme should be used to account for the correlated errors in the sequence of $\ln P(w)$. There are also unresolved questions related to choice of partitions.

Some other regression methods are discussed in a later section.

MAXIMUM LIKELIHOOD METHOD

If n_i fish are sampled from the density (2), and assigned to length interval midpoints as described above, the joint likelihood of the length-frequency sample is

$$\mathcal{L} = \prod_{i=1}^r f_i^{n_i} \quad \dots 28)$$

where

$$f_i = \frac{[L_\infty - (\tilde{\ell}_i - \frac{\Delta_i}{2})]^\theta - [L_\infty - (\tilde{\ell}_i + \frac{\Delta_i}{2})]^\theta}{(L_\infty - L_c)^\theta} \quad \dots 29)$$

is the probability of falling in the i -th length class interval. When the Δ_i are small, the exact interval probabilities may be approximated by

$$f'_i = \frac{\theta (L_\infty - \tilde{\ell}_i)^{\theta-1} \Delta_i}{(L_\infty - L_c)^\theta} \quad \dots 30)$$

For any values of θ and L_∞ , \mathcal{L} is maximized by setting L_c as large as possible. Therefore, under the sampling scheme adopted here, the maximum likelihood estimate of L_c is L_{\min} . Further, when $\theta = 1$, the maximum likelihood estimate of L_∞ is L_{\max} . Except in this singular case, the joint maximum likelihood estimates of θ and L_∞ ($\hat{\theta}_{ML}$ and $\hat{L}_{\infty ML}$) must be found numerically, applying iterative search methods to \mathcal{L} or $\ln \mathcal{L}$. If Newton-type methods are used, the asymptotic covariance matrix for the estimates is also easily computed. Note that in the search for $\hat{L}_{\infty ML}$, L_{\max} is a lower bound. Observe also that if both L_c and L_∞ are assumed known, the maximum likelihood estimator of θ is identical to the Ssentongo-Larkin estimator, $\hat{\theta}_{SL}$.

Another likelihood model may be constructed by considering the conditional probabilities of the length class frequencies given the total sample. Such frequencies are multinomial with probabilities

$$\mu_i = f_i / \sum_{i=1}^r f_i \quad \dots 31)$$

From (29), we have

$$\mu_i = \frac{[L_\infty - (\tilde{\ell}_i - \frac{\Delta_i}{2})]^\theta - [L_\infty - (\tilde{\ell}_i + \frac{\Delta_i}{2})]^\theta}{[L_\infty - L_{\min}]^\theta - [L_\infty - L_{\max}]^\theta} \quad \dots 32)$$

$$i = 1, 2, \dots, r$$

Given $n.$, L_{\min} , and L_{\max} , the likelihood of the sample is therefore

$$\mathcal{L} = \left(\frac{n.}{n_1 \ n_2 \ \dots \ n_r} \right)_{i=1}^r \prod_{i=1}^r \mu_i^{n_i} \quad \dots 33)$$

and maximum likelihood estimates of L_∞ and θ and their asymptotic standard errors may be found in the usual way by maximizing \mathcal{L} or $\ln \mathcal{L}$ directly, or, for example, by fitting the expected length class frequencies to corresponding observations using an iteratively reweighted Gauss-Newton algorithm. In the latter approach, one finds $\hat{\theta}$ and L_∞ which minimize

$$\mathcal{J} = \sum_{i=1}^r \omega_i (n_i - n. \mu_i)^2 \quad \dots 34)$$

with weights ω_i equal to the reciprocals of the expected frequencies. The resulting maximum likelihood estimates will also be minimum chi-square estimates, and the minimized \mathcal{J} will be asymptotically distributed as χ^2 with $r - 3$ degrees of freedom.

In fitting the multinomial model, a question arises concerning the optimal choice of the length classes into which the data are grouped. Intervals of equal width are customary in length-frequency analysis, but as a rule of thumb observations are often pooled where necessary to ensure an expected frequency of at least 5 in each interval. This is typically required in the right tail of the distribution. More specific guidance on grouping for a fixed number of intervals, r , may be offered if an optimality criterion is adopted. For example, one sensible choice would be to set the grouping intervals in such a way that the power of the usual multinomial chi-square goodness of fit test (or likelihood ratio test) for the null distribution is maximized with respect to some selected class of alternatives. Suppose we wish to test $g(\ell)$ against the open (unspecified) class of alternative length-frequency distributions. In this situation it has been shown (Cox and Hinkley 1974) that the optimal grouping is that which uniquely maximizes the expected "entropy" of samples drawn from the null distribution, namely,

$$I = - \sum_{i=1}^r \mu_i \log \mu_i, \quad \dots 35)$$

and that this is accomplished, in our context, by dividing the length range into intervals of equal probability under $g(\ell)$. Thus the r "optimal" length class intervals should have upper end points equal to the sequence of quantiles of the length distribution, of orders $1/r, 2/r, \dots, 1$. For our multinomial model these partitions are at

$$\ell_j^* = L_{\max} - (L_{\max} - L_{\min}) [1 - \frac{j-1}{r}]^\theta \quad \dots 36)$$

$j = 1, 2, \dots, r$

In practice the unknown parameters would be replaced by estimates. We have not evaluated this procedure.

Because we would expect grouping of data to result in loss of information and reduced efficiency of estimators, the actual number of intervals used should in general be as large as possible, subject to the rule that expected frequencies in each length class exceed 4. This recommendation applies provided the assumed steady-state model is correct. In cases where recruitment is actually cyclic but the steady-state model is being used to estimate θ , simulation studies (S. Ralston, unpublished data) suggest that coarser grouping of data will have the effect of smoothing out recruitment fluctuations and increasing the precision of $\hat{\theta}$ (in the same manner that pooling of several samples taken over the course of a spawning cycle will render the steady-state model applicable). In such cases, trial and error will indicate the optimal grouping.

Systematic Biases

We turn now to a consideration of systematic biases in the estimators, which arise when the underlying model assumptions are violated. In some situations serious biases result, in others the estimation procedures are robust and biases negligible.

SUBSTITUTION OF L_{\min} FOR L_c

Since L_c is a location parameter, and L_∞ a scale parameter, they have no bearing on the shape of the theoretical length distribution. This is determined solely by θ . Estimators of θ based on samples drawn randomly from (2), and assuming L_c and L_∞ are known, therefore have distributions unaffected by L_c and L_∞ , and dependent only on θ and n . However, in practice both L_c and L_∞ are unknown. In most of the procedures studied here an estimate of L_c , or of both L_c and L_∞ must be substituted, and properties of the estimators are affected.

We consider L_c first. In one of the likelihood models, the multinomial, problems related to L_c were circumvented by treating it as a nuisance parameter and conditioning the length interval probabilities on the sample end points, i.e., on L_{\min} and L_{\max} . In all the other methods specification of L_c is required. In these cases, L_c (or w_1 in the regression method) is typically set to L_{\min} , the lower bound of the sample length-frequency distribution. Indeed, in most applications the frequency distribution is formed by truncating a larger sample at a selected L_{\min} , setting aside the data below this length, and retaining for analysis only the upper "descending" portion. Setting L_c equal to L_{\min} in this manner is clearly reasonable. As pointed out in the section on Maximum Likelihood Method, doing so maximizes the likelihood of the sample under (2). Further, assuming continuous sampling (all $\Delta_i = 0$), analysis of the relevant order statistics shows that

$$E(L_{\min}) = L_c + \frac{(L_\infty - L_c)}{n\theta + 1}, \quad . . . 37)$$

so that for θ greater than about 2, L_{\min} is essentially unbiased as an estimator for L_c , even for fairly small samples. In virtually all cases, then, systematic biases due to estimating L_c by L_{\min} will be negligible.

SUBSTITUTION OF L_{\max} FOR L_∞ IN $\hat{\theta}_{BH}$ AND $\hat{\theta}_{SL}$

In the earlier sections, we noted that both $\hat{\theta}_{BH}$ and $\hat{\theta}_{SL}$ were susceptible to systematic bias resulting from the replacement of L_∞ by an independent estimate or guess. To avoid the bias, one

would need to apply one of the methods which estimate θ and L_∞ jointly. However, heretofore only Powell's method was available for this, and we are unaware of any situations in which it has been used. In the majority of cases, one of the simpler methods has been adopted instead. Except when reliable, independent estimates of L_∞ have been available from tagging or analysis of hard parts, the usual practice has been to set $L_\infty = L_{\max}$, the length of the longest fish measured, or the upper bound of the length distribution. When growth is deterministic with L_∞ the upper bound to length, as assumed in (2), this substitution biases $\hat{\theta}_{BH}$ and $\hat{\theta}_{SL}$ downward, the degree of negative bias depending on θ , X_c , and n .

If θ is large and X_c small, very large samples are required for this systematic bias to be negligible. Specifically, under (2) it may be shown that the sample size required to ensure that L_{\max} is within 100 δ % of L_∞ with probability ϕ is the smallest integer exceeding

$$n(\delta, \theta) = \frac{\ln(1 - \phi)}{\frac{\delta}{\ln[1 - (\frac{\theta}{1 - X_c})]}} \quad \dots 38$$

For example, if $X_c = 0.3$ and $\phi = 0.95$, we have the following results for $n(\delta, \theta)$:

δ	$\theta = 3.0$	$\theta = 4.0$
0.01	1.0×10^6	7.2×10^7
0.05	8.2×10^3	1.2×10^5
0.10	1.0×10^3	7.2×10^3
0.20	1.3×10^2	4.5×10^2

Thus for L_{\max} to be a reasonably accurate estimate of L_∞ , it must represent the longest fish in a sample of several thousand. This sample need not be the same one used to estimate θ . Indeed, if the deterministic von Bertalanffy model is correct and samples are drawn under identical conditions, L_{\max} should represent the longest specimen seen in all the samples taken.

To correct for this systematic bias, it is often recommended that L_{\max} be adjusted upward by some arbitrary amount, typically about 5% (e.g., Pauly 1983, who suggests dividing L_{\max} by 0.95). Depending on the true value of θ and the size of the sample from which L_{\max} is derived, a 5% adjustment may be a fortuitous choice, but clearly some better procedure is needed. The obvious choice, as suggested above, would be to make fuller use of the sample information to estimate θ and L_∞ jointly, i.e., use Powell's method, the regression method, or a maximum likelihood procedure. In the same vein, an interesting alternative approach which accounts for the systematic bias explicitly is based on the following pair of equations in θ and L_∞ , derived from (2):

$$E(\bar{L}) = L_\infty - (L_\infty - L_c) \left(\frac{\theta}{\theta + 1} \right) \quad \dots 39$$

$$E(L_{\max}) = L_\infty - (L_\infty - L_c) \prod_{i=1}^n \left(\frac{i\theta}{i\theta + 1} \right) \quad \dots 40$$

Substituting sample values of \bar{L} and L_{\max} for their expectations and solving by iteration, one can compute joint moment-type estimates of L_∞ and θ . Sampling variances and other properties of the parameter estimates can be derived empirically using bootstrap techniques.

Numerical evaluation of the bias term in (40) shows that a simple 5% adjustment of L_{\max} (or division by 0.95) is rarely appropriate (Table 1). In most fishery applications, where X_c is 0.5 or greater, the correct adjustment is apt to be anywhere between about 0.5% and 25%, depending on θ and n . The consequences of using an incorrect adjustment may be judged by various criteria, one being the so-called "D-measure" described by Majkowski (1982). For example, a D-measure analysis of $\hat{\theta}_{BH}$ shows that the estimator is quite sensitive to errors in the estimate or guess of L_∞ (Table 2). Since $\hat{\theta}_{BH}$ is linear in L_∞ , the relevant D-measure is directly proportional to the errors, and symmetrical. Combining the information in Tables 1 and 2, one can compute the expected bias in $\hat{\theta}_{BH}$ when a simple 5% upward adjustment is applied to L_{\max} (Table 3). Obviously, too large an adjustment simply replaces the negative bias in $\hat{\theta}_{BH}$ with an unknown positive bias.

Table 1. Percentage negative bias in L_{\max} as estimator of L_∞ , as a function of X_c , θ , and n .

X_c	n	θ		
		2.0	3.0	4.0
0.0	200	6.2	15.2	24.1
	500	4.0	11.2	19.2
	1,000	2.8	8.9	16.1
0.5	200	3.1	7.6	12.0
	500	2.0	5.6	9.6
	1,000	1.4	4.5	8.0
0.75	200	1.6	3.8	6.0
	500	1.0	2.8	4.7
	1,000	0.7	2.2	4.0

Table 2. Percentage carat, i.e., \hat{L}_∞ D-measure biases in $\hat{\theta}_{BH}$ due to systematic errors in \hat{L}_∞ of $\pm \delta\%$. (Biases in $\hat{\theta}_{BH}$ have same sign as errors in \hat{L}_∞ .)

X_c	δ	θ		
		2.0	3.0	4.0
0.0	1	1.5	1.3	1.3
	2	3.0	2.7	2.5
	4	6.0	5.3	5.0
	8	12.0	10.7	10.0
	16	24.0	21.3	20.0
	32	48.0	42.7	40.0
0.5	1	3.0	2.7	2.5
	2	6.0	5.3	5.0
	4	12.0	10.7	10.0
	8	24.0	21.3	20.0
	16	48.0	42.7	40.0
	32	96.0	85.3	80.0
0.75	1	6.0	5.3	5.0
	2	12.0	10.7	10.0
	4	24.0	21.3	20.0
	8	48.0	42.7	40.0
	16	96.0	85.3	80.0
	32	192.0	170.7	160.0

Table 3. Percentage bias in $\hat{\theta}_{BH}$ when L_∞ is set at $L_{\max}/0.95$, under various combinations of X_c , θ , and n .

X_c	n	θ		
		2.0	3.0	4.0
0.0	200	-1.9	-14.3	-25.1
	500	+1.6	-8.7	-18.7
	1,000	+3.5	-5.5	-14.6
	∞	+7.9	+7.0	+6.6
	200	+6.0	-7.3	-18.4
	500	+9.5	-1.7	-12.1
0.5	1,000	+11.4	+1.4	-7.9
	∞	+15.8	+14.0	+13.2
	200	+21.5	+6.7	-5.3
	500	+25.3	+12.3	+1.6
	1,000	+27.2	+15.7	+5.3
	∞	+31.6	+28.1	+26.3
0.75	200	+21.5	+6.7	-5.3
	500	+25.3	+12.3	+1.6
	1,000	+27.2	+15.7	+5.3
	∞	+31.6	+28.1	+26.3
	200	+21.5	+6.7	-5.3
	500	+25.3	+12.3	+1.6

VARIABILITY IN ASYMPTOTIC LENGTH

Another problem of systematic bias can arise if L_{\max} is substituted for L_∞ in the Beverton-Holt or Ssentongo-Larkin formulas but the deterministic von Bertalanffy model is incorrect. Specifically, if the growth trajectories of individual fish converge on different asymptotic lengths, i.e., if $\sigma_\lambda^2 > 0$, L_{\max} can easily exceed the expected asymptotic length, L_∞ , when derived from large samples. As Powell's model shows, the positive bias in L_{\max} as an estimator of L_∞ can be significant when σ_λ^2 is relatively large and θ is not. Subsequent estimates of θ will also be inflated.

Under such conditions, the maximum likelihood methods based on (2) will also give biased results, since L_{\max} is regarded as a lower bound on \hat{L}_∞^{ML} . Partial remedies might be to base estimation on the more realistic likelihood arising from the stochastic model of Powell, or to estimate L_∞ by some statistic reflecting the average size of the largest fish seen in the catches. Using Powell's likelihood would entail estimation of σ_λ^2 as well as θ and L_∞ , and would require numerical approximation of the theoretical length frequencies.

We note that because the regression method at (19) is based on fitting sample mean lengths to their expected values, which are unaffected by σ_λ^2 , the estimates of L_∞ it produces do not suffer from these types of bias. In contrast, Powell's estimators of θ and L_∞ are derived assuming $\sigma_\lambda^2 = 0$, so are presumably biased when this assumption is violated.

VARIABLE RECRUITMENT

Variability in recruitment violates a basic assumption of the methods, but is not particularly troublesome as long as the process is stationary, i.e., there has been no systematic trend in R . When spawning is periodic and reasonably regular, and survival of pre-recruits is stable, the effects of variable recruitment can be smoothed out by sampling the population on several occasions during a complete spawning cycle and pooling the individual length-frequency distributions. Distributions should be weighted by an index of population size, such as average catch per unit effort. When recruitment is trending up or down the smoothing procedure is ineffective, and serious bias in estimates of θ can result from use of any of the estimators, the bias being in the same direction as the trend in R .

VARIABLE MORTALITY

If the simple, constant mortality models are applied to length-frequency data when in fact mortality is variable, subsequent population assessments and yield computations based on the estimates of θ and L_∞ are likely to be biased. But since the robustness of the methods in such circumstances has not been studied, the specific ramifications are unclear.

Nevertheless, several forms of variability may be identified, and some are more easily handled than others. One kind is temporal variability in mortality, caused, for example, by a systematic change in fishing effort or stock vulnerability. Accommodating this kind of variability would be relatively difficult, because the simplifying steady state conditions no longer exist. A second type of variation is that due to size (or age) dependence of mortality. Assuming such size-specificity is time invariant, construction of steady-state population models and parameter estimation procedures is straightforward. We develop such procedures in the following section.

Estimating Length-Specific Fishing Mortality Rates

When aging of the catch is feasible, the analysis of cohort catch histories is one of the chief methods of estimating virtual populations and age-specific mortality rates. When it is not, but the

catch is known by length class interval, one may estimate population size and length-specific mortality using a length-based analog of the cohort methods (e.g., Jones 1981; Pauly 1984). These are applicable under more restrictive assumptions than cohort analysis. Recruitment must be approximately constant, or smoothed out by pooling catch-at-length data, a growth model must be assumed, and mortality rates must be constant in time.

The method developed by Jones and extended by Pauly is based on the deterministic von Bertalanffy growth model and a constant instantaneous natural mortality rate, M. Assuming M, the fishing mortality rate in the largest length class, F_r , and the von Bertalanffy parameters K and L_∞ are known, the length-based VPA proceeds recursively in the same manner as the popular age-based methods, beginning with the catch in the largest length class and working backward to estimate the sequence of length-specific population sizes and fishing mortality rates. Alternatively, if only L_∞ and the "exploitation rate" F_t/Z_t in the terminal length class are specified (where $Z_t = F_t + M$), estimates of M/K and the remaining F_i/Z_i may be found, there being in either case one parameter estimated for every data point, i.e., every length-interval catch.

In this section we adopt the same steady-state assumptions, and describe an alternative procedure for estimating length-specific mortality rates, growth parameters and population size jointly from a regression analysis of catch by length class. This is simply an extension of the age-based catch curve models developed by Ricker (1948, 1975) and frequently applied to cohort analysis.

Historically, in fact, the age-based catch curves were preceded by length-based methods due to Edser (1908), Heincke (1913), and Baranov (1918).

We assume a steady state, with constant recruitment, R, constant natural mortality rate M, and knife-edge selection at age t_c . Beyond t_c , the fishing mortality rate F is described as a step function of age t, namely

$$F(t) = F_i \quad \dots 41)$$

$$t_i < t < t_i + \delta_i$$

$$i = 1, 2, \dots, r$$

where $t_i = t_c + \sum_{j=1}^i \delta_{j-1}$ and the δ s are successive age increments ($\delta_0 = 0$). Under these conditions, the number of fish caught in the i-th age class, $(t_i, t_i + \delta_i)$, during a time period of duration τ , is

$$C(t_i, t_i + \delta_i) = R\tau \left(\frac{F_i}{F_i + M} \right) \left(1 - e^{-(F_i + M)\delta_i} \right) e^{-\sum_{j=1}^i (F_{j-1} + M)\delta_{j-1}} \quad \dots 42)$$

$$= C_i, \text{ say.}$$

Let the δ_i correspond, by virtue of a bijective growth model, to a specific sequence of length intervals into which the catch is conveniently grouped, say the sequence $(w_i, w_i + \Delta_i)$, $i = 1, 2, \dots, r$, where $w_i = L_c + \sum_{j=1}^i \Delta_{j-1} = \tilde{\lambda}_i - \Delta_i/2$ and $\Delta_0 = 0$. If the von Bertalanffy model is used, the correspondence is established by the relation

$$\delta_i = -\frac{1}{K} \ln \left\{ \frac{L_\infty - (w_i + \Delta_i)}{L_\infty - w_i} \right\} \quad \dots 43)$$

Making the substitution of length for age, we have

$$C_i = R\tau \left(\frac{F_i}{Z_i} \right) \left[1 - \left(\frac{L_\infty - (w_i + \Delta_i)}{L_\infty - w_i} \right)^{\theta_i} \right] \prod_{j=1}^{i-1} \left(\frac{L_\infty - (w_{j-1} + \Delta_{j-1})}{L_\infty - w_{j-1}} \right)^{\theta_{j-1}} \quad \dots 44)$$

where $Z_i = F_i + M$, $\theta_i = Z_i/K$, and $w_o = \Delta_o = Z_o = 0$.

If the fishing mortality rate is constant after recruitment, (44) reduces to

$$\begin{aligned} C_i &= R\tau \left(\frac{F}{Z} \right) \left[1 - \left(\frac{L_\infty - (w_i + \Delta_i)}{L_\infty - w_i} \right)^{\theta} \right] \left(\frac{L_\infty - w_i}{L_\infty - L_c} \right)^{\theta} \quad \dots 45) \\ &= C_* \left[\frac{(L_\infty - w_i)^{\theta} - (L_\infty - (w_i + \Delta_i))^{\theta}}{(L_\infty - L_c)^{\theta}} \right] \end{aligned}$$

where $C_*/\tau = R F/Z$ is the constant total catch in a unit time interval. Alternatively, (45) may be derived from the length density at (2), i.e.,

$$C_i = \tau F \int_{w_i}^{w_i + \Delta_i} N(\ell) d\ell \quad \dots 46)$$

$$= \tau F \frac{R}{Z} \int_{w_i}^{w_i + \Delta_i} g(\ell) d\ell$$

$$= C_* [G(w_i + \Delta_i) - G(w_i)]$$

where $N(\ell)$ is the number of fish of length ℓ in the steady-state population. Recalling the earlier discussion on maximum likelihood methods, C_i/C_* is seen to be the probability that a fish taken at random from the population will be in the i -th length class interval, a function of L_c , L_∞ , and Z/K , as in (29).

As formulated in (44), the length-based catch model is a set of r equations in $r + 4$ unknown parameters (R , M , K , L_∞ , and the F_i , $i = 1, 2, \dots, r$), so additional restrictions must be imposed to permit estimation. As usual, this may be done by making further structural assumptions about fishing mortality. One possibility is to assume that fishing mortality rates are constant and distinct

within $m < r - 4$ segments of the catchable length range. Another is to relate fishing mortality rate to length class by a particular parametric model, such as the quadratic $F(i) = a + bi + ci^2$. In the latter event the problem is reduced to estimating seven parameters from the r observed length class catches.

With proper structuring, so that adequate degrees of freedom are available, the q -model parameters can be estimated by iteratively reweighted least squares. When weights are set equal to the expected catches, this amounts to minimizing the familiar Pearson chi-square statistic. For large samples, this statistic will be distributed as χ^2 with $r - q - 1$ degrees of freedom. The resulting parameter estimates will also be maximum likelihood estimates, and if a Gauss-Newton algorithm is used in the optimization an estimate of their asymptotic covariance matrix will be readily available.

For some purposes individual estimates of M , K and the fishing mortality rates may not be necessary, and in this event the catch models can be parameterized in terms of R , L_∞ , and the ratios M/K and F_i/Z_i , $i = 1, 2, \dots, r$. (In the structured mortality model, let $F_i/Z_i = a' + b'i + c'i^2$). In (44), for example, this is accomplished by replacing θ_i by the product

$$\theta_i = \left(\frac{M}{K} \right) \left(1 - \frac{F_i}{Z_i} \right)^{-1} \quad \dots 47)$$

In the models just described, it is assumed that the entire catch is taken by a single gear, so that estimating the length-specific fishing mortality rates is equivalent to estimating the selectivity curve for that gear with respect to the species under study. If more than one gear is significantly involved, then the catch data from all of them must be combined (properly weighted), and the analysis applied to the aggregate information. As in age-based VPA methods, gear-specific fishing mortality rates can be computed by allocating the fishing mortality estimate for each length class in proportion to the associated catch by each gear.

We have not yet applied these procedures, nor investigated their robustness and other properties. Nevertheless, we would speculate that length-specific fishing mortality rates will not be identifiable, or estimable with much precision, unless they vary fairly dramatically with length. Further, separation of natural mortality and the fishing mortality rates and useful precision in the mortality rate estimates will likely require that a substantial fraction of the total mortality be due to fishing. Thus, the model will probably not estimate fishing mortality rates well when applied to lightly exploited stocks.

However, when adequate length-frequency data are available from virgin populations or very lightly exploited stocks, a regression model may be developed to estimate L_∞ , K and a set of length-specific natural mortality rates. Taking C_i to be a small sample proportional to τ and the steady state population size in the i -th length class, we will have, corresponding to (44), the model

$$C_i = \delta \left(\frac{R\tau}{M_i} \right) \left[1 - \left(\frac{L_\infty - (w_i + \Delta_i)}{L_\infty - w_i} \right)^{\frac{M_i}{K}} \right] \prod_{j=1}^i \left(\frac{L_\infty - (w_{j-1} + \Delta_{j-1})}{L_\infty - w_{j-1}} \right)^{\frac{M_{j-1}}{K}} \quad \dots 48)$$

where δ is the proportionality constant and M_i is the instantaneous natural mortality rate in the i -th length interval. Again, further restrictions on the parameters are necessary to allow estimation. Structural relationships among the M_i could be introduced, as above. If the segmented mortality model is adopted, such independent information as average size at maturity could be used to define points of change in the model. Note that the steady-state recruitment level, R , is no longer estimable, and difficulty in estimating the mortality parameters can be expected unless the length-specific changes in mortality rate are striking. When variation in natural mortality rates is negligible, only L_∞ and $\theta = M/K$ are estimable from the length-frequency information.

Evaluation of Methods

Four of the constant-mortality methods described above were compared by Monte Carlo procedures in situations where both θ and L_∞ were to be estimated from length-frequency data and L_c was assumed known. Specifically, performance of the following procedures was examined:

- (i) L_∞ was set equal to L_{\max} , and the classical Beverton-Holt estimator then applied to estimate θ .
- (ii) Powell's method of moments was applied to estimated L_∞ and θ , given the sample mean and variance of ℓ .
- (iii) The weighted regression method, based on the linear relationship between $\bar{\ell}$ and L_c , was used to estimate L_∞ and θ .
- (iv) The continuous maximum likelihood method was applied to estimate L_∞ and θ .

For specified values of X_c and θ samples of lengths were drawn from (6) by the inverse transformation method, a well-known Monte Carlo procedure. Let u be a random variate uniformly distributed on $(0, 1)$. Then a corresponding random variate from (6), x , is uniquely determined by

$$x = 1 - (1 - X_c)(1 - u)^{\frac{1}{\theta}} \quad \dots 49)$$

A uniform random number generator was used to produce sets of u_i and the associated random lengths, x_i , $i = 1, 2, \dots, n$. Except when the level of aggregation was being studied, experiments were conducted using the raw generated lengths. Thus a continuous distribution of the sample variates was assumed (all $\Delta_i = 0$). Recall that use of the auxiliary variable, X , is equivalent to scaling the length axis so that $L_\infty = 1.0$.

Performance of the estimators was measured by the coefficient of variation (CV), the relative bias (B), and the coefficient of error (CE). For a particular parameter π , we define these as

$$CV(\hat{\pi}) = \frac{\sqrt{\text{var}(\hat{\pi})}}{\pi}, \quad \dots 50)$$

$$B(\hat{\pi}) = \frac{\text{Bias}(\hat{\pi})}{\pi}, \quad \dots 51)$$

and

$$CE(\hat{\pi}) = \sqrt{CV(\hat{\pi})^2 + B(\hat{\pi})^2} = \sqrt{\frac{\text{MSE}(\hat{\pi})}{\pi}} \quad \dots 52)$$

where $\text{MSE}(\hat{\pi})$ is the mean squared error of $\hat{\pi}$. For each set of experimental circumstances, these measures were computed with respect to \hat{L}_∞ and $\hat{\theta}$ on the basis of 200 replicate samples. Trial simulations showed that this number of replicates was sufficient to describe the sampling distributions of the estimators accurately.

Experimental variables of primary interest were the level of θ and the sample size, n . Of lesser interest were $1 - X_c$, the proportion of the total possible length range selected in the sample, the level of aggregation in the data, and, in the regression method, the number of partitions and type of weighting used. Below we describe each particular simulation experiment, and the chief results.

Experiment 1

The performance of the four estimation procedures was compared at three levels of θ (2.0, 3.0, 4.0) and three levels of n , (200, 500, 1,000), with $X_c = 0$.

Results: The Beverton-Holt method produces estimates with the lowest CV, but these estimates have a substantial negative bias, especially with small n , and large θ (Tables 4 and 5). The Powell method, on the other hand, computes estimates with the highest CV of any method examined, but generally with the smallest B (positive) as well. Estimates derived by the regression

Table 4. Performance measures for various estimators of L_∞ , as a function of θ and n .

n	Method	CV	2.0			3.0			4.0		
			B	CE	CV	B	CE	CV	B	CE	
200	BH	3.3	-5.9	6.8	6.1	-15.3	16.5	7.1	-23.2	24.3	
	P	10.7	2.1	10.9	15.9	1.8	16.0	39.5	12.4	41.4	
	R	5.8	-1.0	5.9	10.3	-5.3	11.6	19.2	-7.2	20.5	
	ML	4.7	-1.7	5.0	10.5	-3.2	11.0	25.2	3.7	25.4	
500	BH	2.2	-3.7	4.3	4.0	-11.4	12.1	5.8	-19.2	20.1	
	P	6.6	0.8	6.6	11.2	1.7	11.3	13.0	1.3	13.0	
	R	3.4	-0.4	3.4	6.6	-1.9	6.9	11.1	-4.2	11.9	
	ML	2.6	-0.7	2.7	6.0	-1.1	6.1	9.6	-1.5	9.7	
1,000	BH	1.7	-2.7	3.2	3.5	-8.6	9.3	4.6	-16.7	17.3	
	P	4.6	-0.1	4.6	7.3	1.2	7.4	11.6	0.8	11.7	
	R	2.4	-0.4	2.4	5.0	-0.8	5.1	7.9	-3.8	8.8	
	ML	1.8	-0.6	1.9	4.5	-0.2	4.5	7.5	-1.5	7.7	

Table 5. Performance measures for various estimators of θ , as a function of θ and n .

n	Method	CV	2.0			3.0			4.0		
			B	CE	CV	B	CE	CV	B	CE	
200	BH	8.7	-9.0	12.5	8.8	-20.3	22.2	9.7	-28.5	30.1	
	P	20.0	3.5	20.3	24.0	3.0	24.2	54.8	17.5	57.5	
	R	12.2	-3.0	12.6	16.7	-8.0	18.5	28.6	-9.0	30.0	
	ML	12.2	-3.0	12.6	17.3	-4.3	17.9	36.3	5.8	36.8	
500	BH	5.0	-5.0	7.1	5.8	-15.0	16.1	7.5	-23.8	24.9	
	P	12.2	1.5	12.3	17.0	3.0	17.3	18.0	2.0	18.1	
	R	7.1	-0.5	7.1	10.5	-2.7	10.9	16.4	-4.8	17.1	
	ML	7.1	-1.0	7.1	10.5	-1.3	10.6	13.9	-1.5	14.0	
1,000	BH	5.0	-4.0	6.4	5.8	-10.7	12.1	6.1	-20.7	21.6	
	P	8.7	-0.5	8.7	11.1	2.3	11.3	15.4	1.3	15.5	
	R	5.0	-1.0	5.1	8.2	-0.3	8.2	10.6	-5.0	11.7	
	ML	5.0	-1.0	5.1	7.5	0.3	7.5	10.3	-2.0	10.5	

method, with optimal weighting and 10 partitions, appear to have slightly greater bias than the maximum likelihood estimates (negative in both cases) and the Powell method, but are far superior to the Beverton-Holt estimates in this regard. In terms of precision, the regression method is almost as good as the maximum likelihood method.

As expected, the maximum likelihood method, among the estimators examined, generally has the best overall performance in terms of mean squared error and CE, followed closely by the regression method. The maximum likelihood estimates of L_∞ almost always had the smallest CE. At $\theta = 2$, the Beverton-Holt procedure performed relatively well in terms of CE, virtually on a par with the maximum likelihood and regression methods, but it did comparatively poorly at higher values of θ . At $\theta = 2$, the Powell estimates generally had the largest CE of any method; otherwise, Beverton-Holt estimates tended to have the largest CE.

In all methods, the effect of increasing θ is to increase B, CV, and CE for both parameters estimated. As expected, increasing n , uniformly reduces B, CV, and CE.

Experiment 2

The effects on Beverton-Holt estimates of using a bias-corrected estimate of L_∞ were studied. In each generated sample, L_{\max} was increased by the theoretical bias term on the right side of (40), and this corrected estimate of L_∞ was used as a basis for estimating θ .

Results: Adjustment of L_{\max} for bias resulted in a significant reduction in B ($\hat{L}_{\infty BH}$), as anticipated, and also in B ($\hat{\theta}_{BH}$), and increases in CV for both parameter estimates. The CE's were changed very little by the adjustment. In practice, the bias correction procedure would have to be implemented by iterative solution of (39) and (40), i.e., the bias term itself would have to be estimated, and the properties of the estimators would be affected.

Experiment 3

The effects of X_c on the performance of all methods was studied, with n , fixed at 200, and θ at 3.0. Samples were drawn from (6) with X_c set at three levels (0.0, 0.5, 0.75) and estimators were applied with X_c assumed known.

Results: As expected, increasing the value of X_c (with n , unchanged) reduced B, CV, and CE for L_∞ in all methods, but had no effect on estimates of θ . With respect to estimation of L_∞ , increases in X_c have the same effect as increasing sample size with X_c fixed. On the other hand, as pointed out earlier, distributions of $\hat{\theta}$ are unaffected by the magnitude of X_c when X_c is known. The amount of information about θ conveyed by the sample depends only on θ and the sample size. (Note, for example, that in the regression model if L_c is known, θ is completely determined by the slope of the regression line, the expected value of which is independent of X_c .)

Experiment 4

The effects of weighting schemes and the number of partitions on estimates produced by the regression method were studied. Weights were determined in three ways: (1) the complete covariance matrix was inverted, (2) the diagonal elements only were inverted, and off diagonal elements neglected and (3) uniform weights were assigned. Under the diagonal weighting scheme, regression estimates were computed with both 10 and 20 partitions.

Results: Use of the full covariance weighting produced the most precise estimates, as expected, but these estimates had greater bias than those produced with either diagonal or uniform weighting. However, the fully weighted estimates had the lowest CEs. When 10 partitions are used, diagonal weighting, or even equal weighting, appears to be satisfactory.

Similarly, increasing the number of partitions (points) in the regression reduced the CVs at the expense of greater bias and resulting estimates had smaller CEs.

Experiment 5

Effects of aggregation were assessed by comparing estimates derived using continuous data (i.e., exact lengths) with those computed from data grouped by intervals of width $L_\infty/100$.

Results: In most cases, aggregation at this level slightly increased the CVs and reduced relative biases (except for increases in bias of Powell estimates). The CEs were increased slightly in all methods except the Beverton-Holt, in which case they were reduced. (Note that among the estimators studied, those of the Beverton-Holt procedure are distinguished by having their CEs composed mostly of bias. In the other methods, the variance dominates.)

Conclusions and Recommendations

Each of the methods outlined here has attributes which make it the method of choice in particular circumstances.

The classical estimator of θ due to Beverton and Holt, for example, is the only one applicable when the available information is limited to the mean length in a sample of fish above L_c and an independent estimate of L_∞ . The difficulty with the method, of course, is that considerable bias in $\hat{\theta}_{BH}$ will result unless L_∞ is known accurately. As discussed earlier, estimating L_∞ by L_{max} is advisable only if a very large sample has been measured. Alternatively, the estimate of L_∞ (and K) can be derived independently by fitting the von Bertalanffy model to tag recapture data or to readings of hard parts, and then substituted into the Beverton-Holt estimator. Here, too, L_∞ is rarely determined with sufficient accuracy unless fish are aged, or tagged and recovered, over the entire length range and an asymptote is clearly evident. These conditions are difficult to achieve when there is a low probability that a tagged fish will survive long enough to approach the asymptotic length and be recaptured, or when otoliths and other hard structures are difficult to read and interpret in old fish. The result is often serious bias in estimates of L_∞ and K. However, since these biases are inversely related, some compensation occurs when the pair of parameter estimates is inserted into the Beverton-Holt formula to compute Z.

If the variance among lengths in the sample is known in addition to the mean, a considerable improvement is afforded by Powell's method of moments, which provides joint estimates of θ and L_∞ without reliance on independent information. If tag recapture statistics or otolith data are available, they may then be used to estimate K with L_∞ fixed at $\hat{L}_\infty p$. An estimate of Z follows immediately.

The weighted regression method is attractive for several reasons, and has proven to be the most dependable of the constant-mortality methods in practice. In statistical terms, it is very efficient, giving parameter estimates with almost minimum mean squared error. It also is insensitive to assumptions on L_∞ , which may be regarded either as an absolute upper bound to length, or as the expected asymptotic length in a stochastic von Bertalanffy growth process. Among the simple methods studied, only the Beverton-Holt procedure shares this robustness with respect to L_∞ .

Another reason for recommending the regression method is that it provides a simple graphical means of judging the appropriateness of underlying model assumptions; serious violations can be detected directly by comparing observed mean lengths above the specified cut-off points with those predicted by the model.

Perhaps the only drawback to the regression method is the need to compute the covariance matrix A to achieve full efficiency. However, little is sacrificed by weighting simply by the partition sample sizes.

From a theoretical standpoint, the maximum likelihood procedure is highly attractive, yielding estimators with greatest asymptotic efficiency and other desirable properties. However, it is not robust to violations of the assumption on L_∞ , and this has apparently been a source of difficulty in applications. In a number of cases, searches for L_∞ have converged on L_{max} , suggesting overestimation of L_∞ and θ . Use of Powell's stochastic model (or a suitable variant) in a maximum likelihood context might be more appropriate, but we have not tried it.

In the methods outlined here, estimators are based on length-frequency data entirely, or on combinations of length-frequency data and other information, applied sequentially. Where consistent information on growth and mortality is available from several independent sources, e.g., from tagging, otolith analysis or length-frequency samples, it may also be used jointly to estimate the parameters of a common steady-state model. A maximum likelihood approach is most appropriate here, provided one can specify the forms of the various error distributions. An ancillary benefit of joint estimation is that the relative contribution of each type of information toward the mean squared errors of the parameter estimates can be determined, providing useful guidance for research planning and allocation of resources. Examples of joint estimation of growth parameters from length-increment data and age-length data have been reported by Laurs et al. (1983), who combined tag recapture results with counts of daily increments on sagittae of North Pacific albacore, by Kirkwood (1983), who pooled tag recapture data with information on "aged" length-frequency modes in southern bluefin tuna and by Morgan (Part I, this vol.). Extension of such techniques to joint estimation of growth, mortality and recruitment parameters in more general contexts is straightforward.

References

- Baranov, F.I. 1918. On the question of the biological basis of fisheries. Nauchn. Issled. Ikhtiol. Inst. Izv. 1: 81-128.
(In Russian.)
- Beverton, R.J.H. and S.J. Holt. 1956. A review of methods for estimating mortality rates in fish populations, with special reference to sources of bias in catch sampling. Rapp. P.-v. Réun. Cons. Int. Explor. Mer 140: 67-83.
- Cox, D.R. and D.V. Hinkley. 1974. Theoretical statistics. Chapman and Hall, London.
- Edser, T. 1908. Note on the number of plaice at each length, in certain samples from the southern part of the North Sea, 1906. J.R. Stat. Soc. 71: 686-690.
- Gulland, J.A. 1983. Fish stock assessment: a manual of basic methods. FAO/Wiley Series on Food and Agriculture.
- Heinke, F. 1913. Investigations on the plaice. General report. 1. Plaice fishery and protective regulations. Part I. Rapp. P.-v. Réun. Cons. Int. Explor. Mer 17A: 1-153.
- Jones, R. 1981. The use of length composition data in fish stock assessments (with notes on VPA and cohort analysis). FAO Fish. Circ. 734. 60 p.
- Kirkwood, G.P. 1983. Estimation of von Bertalanffy growth curve parameters using both length-increment and age-length data. Can. J. Fish. Aquat. Sci. 40: 1405-1411.
- Laurs, R.M., R. Nishimoto, J.H. Uchiyama and J.A. Wetherall. 1983. Age and growth of North Pacific albacore, *Thunnus alalunga*, determined from tag returns and daily increments on otoliths. Eighth North Pacific Albacore Workshop, Document NPALB/83/1. National Marine Fisheries Service, Southwest Fisheries Center Admin. Report LJ-84-04.
- Majkowski, J. 1982. Usefulness and applicability of sensitivity analysis in a multispecies approach to fisheries management, p. 149-165. In Pauly, D and G.I. Murphy (eds.) Theory and management of tropical fisheries. ICLARM Conference Proceedings 9, 360 p. International Center for Living Aquatic Resources Management, Manila, Philippines and Division of Fisheries Research, Commonwealth Scientific and Industrial Research Organisation, Cronulla, Australia.
- Pauly, D. 1983. Some simple methods for the assessment of tropical fish stocks. FAO Fish. Tech. Pap. 234, Rome, 52 p.
- Pauly, D. 1984. Fish population dynamics in tropical waters: a manual for use with programmable calculators. ICLARM Studies and Reviews 8, 325 p. International Center for Living Aquatic Resources Management, Manila, Philippines.
- Pauly, D. and N. David. 1981. ELEFAN I, a BASIC program for the objective extraction of growth parameters from length-frequency data. Meeresforsch. 28: 205-211.
- Powell, D.G. 1979. Estimation of mortality and growth parameters from the length frequency of a catch. Rapp. P.-v. Réun. Cons. Int. Explor. Mer 175: 167-169.
- Ricker, W.E. 1948. Methods of estimating vital statistics of fish populations. Indiana Univ. Publ. Sci., Ser. 15, 101 p.
- Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Fish. Res. Board Can. Bull., 191, 382 p.
- Schnute, J. and D. Fournier. 1980. A new approach to length-frequency analysis: growth structure. Can. J. Fish. Aquat. Sci. 37: 1337-1351.
- Ssentongo, G.W. and P.A. Larkin. 1973. Some simple methods of estimating mortality rates of exploited fish populations. J. Fish. Res. Board Can. 30: 698-698.

**A Method for the Estimation of Growth, Mortality
and Gear Selection/Recruitment Parameters from
Length-Frequency Samples Weighted by Catch per Effort**

P. SPARRE

*Danish Institute for Fisheries and Marine Research
Charlottenlund Slot
2920 Charlottenlund
Denmark*

Sparre, P. 1987. A method for the estimation of growth, mortality and gear selection/recruitment parameters from length-frequency samples weighted by catch per effort, p. 75-102. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

This paper describes a method and a FORTRAN program for the estimation of von Bertalanffy growth parameters, total mortality rates and gear selection/recruitment parameters from a time series of length-frequency samples weighted by catch per effort (e.g., numbers caught per hour by length class each quarter of the year). The length frequencies are separated into normally distributed components each of which represents a cohort. The mean values of the normally distributed components conform to a growth curve as suggested by various authors. Further, the solutions also conform to the exponential decay model and a gear selection/recruitment model. The chi-squared criterion as a measure of goodness of fit is used.

The method is constructed to allow for continuous, seasonally variable recruitment all year-round and is intended to be used especially for tropical and other fish stocks displaying an appropriate recruitment pattern.

The method is illustrated by means of a simulated example, i.e., an example for which the true parameters are known.

Introduction

The type of analysis to be presented in this paper deals with fish stocks for which it is difficult and expensive to obtain direct measurements of age compositions. This is the case for many tropical fish stocks where otoliths or scales do not show a clear annual ring structure. Further, unlike fish stocks in colder waters the spawning seasons of most tropical fish stocks are not restricted to a relatively short period once a year, which makes the definition of age composition complicated.

The present model assumes that only length-frequency samples over a period of time (e.g., a length-frequency sample per quarter of the year) are available. It assumes that recruitment occurs continuously all year-round, but is seasonally variable, with one or two peaks per year. In addition, the method requires some basic knowledge of the fish stock in question. The method works as an

iterative process and the user has to provide initial guesses of the results (mortality rates, growth parameters, etc.). One may say that the method attempts to improve the user's opinion.

Thus, this is not an entirely objective method. However, this author doubts that any objective method based entirely on length-frequency data will ever be developed. Of course, the higher the quantity and quality of data the less subjectivity is introduced in the analysis.

This contribution is one in a long series of works on analysis of length-frequency data, the start of which dates back to Petersen's work on identifying age groups from modes in length-frequency data of *Zoarces viviparus* (Petersen 1892). Pearson (1894) developed the first statistical theory on separation of overlapping component distributions.

For many years paper-and-pencil methods have been used in most practical applications. The "probability paper method" of Harding (1949) and Cassie (1954) is probably the most commonly used. The "parabola method" (Hald 1952; Tanaka 1953) is another example. The paper-and-pencil method preferred by the author is that developed by Bhattacharya (1967). All the paper-and-pencil methods are highly subjective and one often can get any result one wishes out of these exercises. Ideally one should test for each data set a large number of alternative solutions and then select the best one.

For practical reasons it is necessary to use computers. Hasselblad (1966) developed the first computerized method. A number of computerized versions of length-frequency analysis have appeared since then (e.g., Tomlinson 1971; McNew and Summerfelt 1978; Macdonald and Pitcher 1979; Schnute and Fournier 1980; Clark 1981).

A quite different approach for using computers in conjunction with length-frequency analysis is the ELEFAN I program of Pauly and David (1981). While all the other works mentioned so far are based on traditional statistical theory (as presented in Kendall and Stuart 1961), ELEFAN I uses a new statistical technique. This author was introduced to the field of length-frequency analysis by ELEFAN I. However, after some time with ELEFAN I, I found, although the basic ideas were good, that a lot of details could be considerably improved. The basic ideas behind the model of Schnute and Fournier (1980) are essentially similar to those behind ELEFAN I, but the estimation technique is considered superior to that of ELEFAN I. The basic ideas behind this paper are thus adopted from Schnute and Fournier (1981). They may be expressed as follows:

Often analysis of length-frequency data has been separated into a number of steps, each of which were carried out independently of the others.

Step 1 usually is the separation of the length-frequency data into normally distributed components.

Step 2 is the interpretation of the components (or of some of them) as cohorts and the assignment of an age to each cohort.

Step 3 is the derivation of growth parameters from age and mean length at age.

Step 4 is the derivation of estimates of mortality rates from relative differences in the cohort sizes.

Often this technique gave results that were ambiguous or made no biological sense. To avoid ambiguous solutions one can attempt to force a biological structure on the solutions. This requires that all steps are carried out simultaneously (e.g., only resolutions of the length-frequency samples into components whose mean values correspond to a von Bertalanffy growth curve are accepted). The work by Macdonald and Pitcher (1979) presents an excellent introduction to the statistical theory behind this approach to length-frequency analysis. However, the method they suggest may result in estimates of negative numbers in a cohort. Clark (1981) circumvents this problem by using the general technique of quadratic programming. In the present work the problem is solved by forcing the solutions to correspond to the exponential decay model, a solution which is along the lines suggested by Schnute and Fournier (1980) and was applied to a single sample problem by Fournier and Breen (1983). Also, underrepresentation of small fish due to gear selection and recruitment is accounted for by forcing the solutions to correspond to the selection/recruitment model of Hoydal et al. (1982).

The method presented here assumes that the available length-frequency data are expressed as catch per unit of effort (c/f) e.g., one sample each quarter of the year. From these data the method estimates:

1. Growth parameters of the seasonally oscillating version of the von Bertalanffy growth equation presented by Pauly and Gaschütz (1979).

2. Total mortality rates in each period between two consecutive samples.
3. Gear selection/recruitment parameters after Hoydal et al. (1982).

If the available length-frequency data are not expressed on a c/f basis, the method can still be used to estimate growth parameters and gear selection/recruitment parameters. If only one length-frequency sample is available, the method remains applicable; but in this case, of course, seasonal growth oscillations cannot be considered.

This paper differs somewhat from most other papers on the subject considered in that it concentrates more on the actual FORTRAN programming of the method. The symbols and notation used are a mixture of computer language and ordinary mathematical expressions. The intention is to make the translation from the mathematics to the computer program as easy as possible (see also Morgan and Pauly, Part II, this vol.).

INPUT DATA

The present model assumes that NOS research cruises have been carried out. A single date is assigned to each cruise (e.g., the date when half of a given cruise was completed).

Let S be the cruise index and let TIME(S) be the date associated with cruise no. S. If the date assigned to cruise no. S is in the format YEAR(S), MONTH(S) and DAY(S), then

$$\text{TIME}(S) = \text{YEAR}(S) + (\text{MO}(\text{MONTH}(S)) + \text{DAY}(S))/365 \\ S = 1, 2, \dots, \text{NOS}.$$

where $\text{MO}(1) = 0, \text{MO}(2) = 31, \text{MO}(3) = 89, \dots, \text{MO}(12) = 334$.

The time difference between two samples is defined by:

$$\text{DT}(S) = \text{TIME}(S) - \text{TIME}(S - 1). \quad S = 2, 3, \dots, \text{NOS}.$$

A cruise is considered a sample. Thus, a sample may be the pooled data from a number of e.g., trawl hauls. The basic data are numbers caught per unit of effort by length group.

Let L be the length group index. Length group L consists of the fish of lengths between

$$\text{Lower limit : FIRSTL} + \text{DL} * (L - 1)$$

and

$$\text{Upper limit : FIRSTL} + \text{DL} * L$$

where FIRSTL is the lower limit of the smallest length group and DL is the length interval size.

Let NOL be the number of length groups. Thus, the upper limit of the largest length group becomes FIRSTL+DL*NOL.

The basic observations are:

OBS (S, L) = The average number caught per unit of effort in sample S, belonging to length group L.

$$S = 1, 2, \dots, \text{NOS}. \quad L = 1, 2, \dots, \text{NOL}.$$

$$\text{OBSSUM}(S) = \sum_{L=1}^{\text{NOL}} \text{OBS}(S, L) = \text{total number of fish in sample no. } S$$

EXAMPLE:

To illustrate the formulae a hypothetical example has been constructed (how this was done is explained further below). This example which will be used throughout this paper consists of four samples in Table 1, collected at the dates given in Table 2.

The length interval size is $DL = 1.0$ cm. The lower limit of the smallest length group is $FIRSTL = 5.0$ cm. The number of length groups is $NOL = 50$. Thus the upper limit of the largest length group becomes 55 cm.

Table 1 shows the basic observations and Fig. 1 shows the bar diagrams corresponding to these data.

BASIC ASSUMPTIONS

It is assumed that each sample

$$OBS(S, L), \quad L = 1, 2, \dots, NOL; \quad S = 1, 2, \dots, NOS.$$

is a random sample of the fishable part of the stock. The fishable part of the stock is the part of the stock recruited to the area covered by the survey (the fishing grounds) and which consists of fish that would be retained if they entered the gear. Thus, every fish recruited to the fishing grounds which is too large to escape through the meshes of the gear in use should have the same probability of being sampled.

The method takes into account that some fish belonging to the stock have smaller probability of being sampled than other fish because they are not fully recruited or because the gear has reduced catchability for them. By applying an "adjustment factor" to the observations they are transformed into a random sample.

THE ESTIMATION PROBLEM

The approach taken here is the traditional one, the theory of which can be found in many textbooks of mathematical statistics (e.g., Cramer 1961, chapter 30; Kendall and Stuart, Vol. 2, 1961, chapter 30; Rao 1965, chapter 5; Draper and Smith 1981, chapter 10). This method has been applied to analysis of fish length frequencies by several workers, e.g., Macdonald and Pitcher (1979), Schnute and Fournier (1980) and Clark (1981).

Thus, there is nothing new in this section and the basic results of the general theory are presented here primarily to introduce the necessary notation for this special application:

If we consider $OBS(S, L)$ as a stochastic variable it may be written:

$$OBS(S, L) = CFREQ(S, L) + e \quad \dots 1)$$

where $CFREQ$ is the mean value of OBS and e is a stochastic term with mean value zero.

Suppose that we have a mathematical model of $CFREQ(S, L)$, i.e., some mathematical function of some dependent observations (e.g., the times of the sampling, $TIME(S)$) and a vector of $NPAR$ parameters:

$$THETA = (THETA(1), THETA(2), \dots, THETA(NPAR)) \quad \dots 2)$$

e.g.,

$$OBS(S, L) = CFREQ(S, L, TIME(S), THETA) \quad \dots 3)$$

Table 1. Basic observations (hypothetical length-frequency data, see also Figs. 1A-D).

L ^a	Lower limit of length class (cm)	15 Mar 75 OBS (1, L)	20 Jul 75 OBS (2, L)	12 Oct 75 OBS (3, L)	5 Jan 76 OBS (4, L)
1	5	0	0	0	0
2	6	0	0	0	0
3	7	0	0	0	0
4	8	1	0	0	0
5	9	11	5	0	0
6	10	54	21	0	0
7	11	138	65	0	0
8	12	156 (max)	91 (max)	0	0
9	13	84	52	2	0
10	14	23	19	6	0
11	15	3	5	15	0
12	16	0	8	36	0
13	17	0	20	53	1
14	18	0	39	69	3
15	19	1	77	66	7
16	20	3	85	45	17
17	21	5	90	31	23
18	22	11	69	32	40
19	23	20	61	43	43
20	24	32	32	54	45
21	25	54	17	55	38
22	26	64	8	72 (max)	48
23	27	68	10	57	49 (max)
24	28	91	18	43	40
25	29	88	26	32	47
26	30	99	33	25	46
27	31	95	43	27	43
28	32	111	53	32	40
29	33	86	55	38	35
30	34	96	77	41	31
31	35	68	69	57	34
32	36	68	77	49	39
33	37	63	61	58	43
34	38	70	69	62	35
35	39	60	51	52	41
36	40	55	45	63	35
37	41	65	38	52	41
38	42	60	24	36	35
39	43	51	27	38	26
40	44	42	18	27	23
41	45	33	13	20	18
42	46	29	9	15	14
43	47	19	8	12	8
44	48	15	6	8	5
45	49	10	3	6	4
46	50	5	2	3	2
47	51	4	1	3	1
48	52	2	1	2	1
49	53	1	0	1	0
50	54	1	0	1	0
OBSSUM (S)		2,115	1,601	1,439	1,001

^aL = length class index.

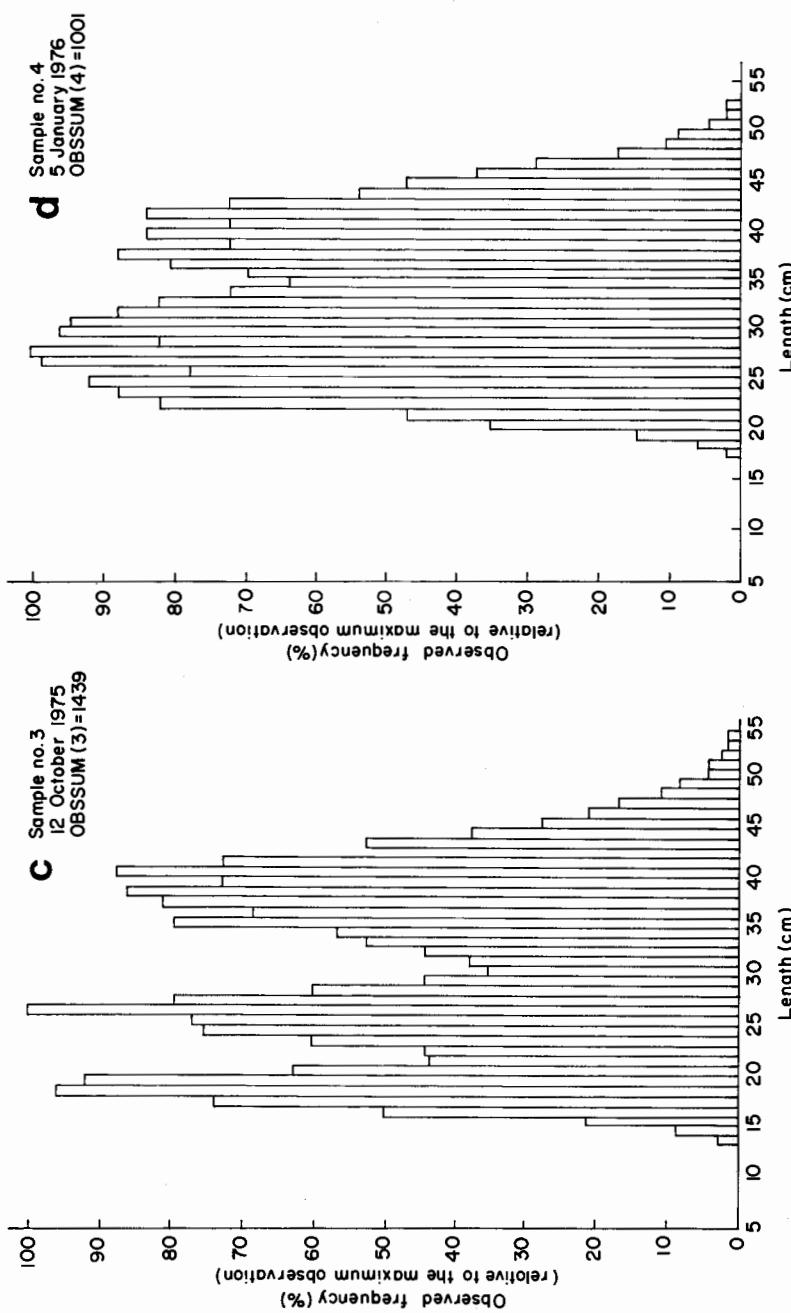
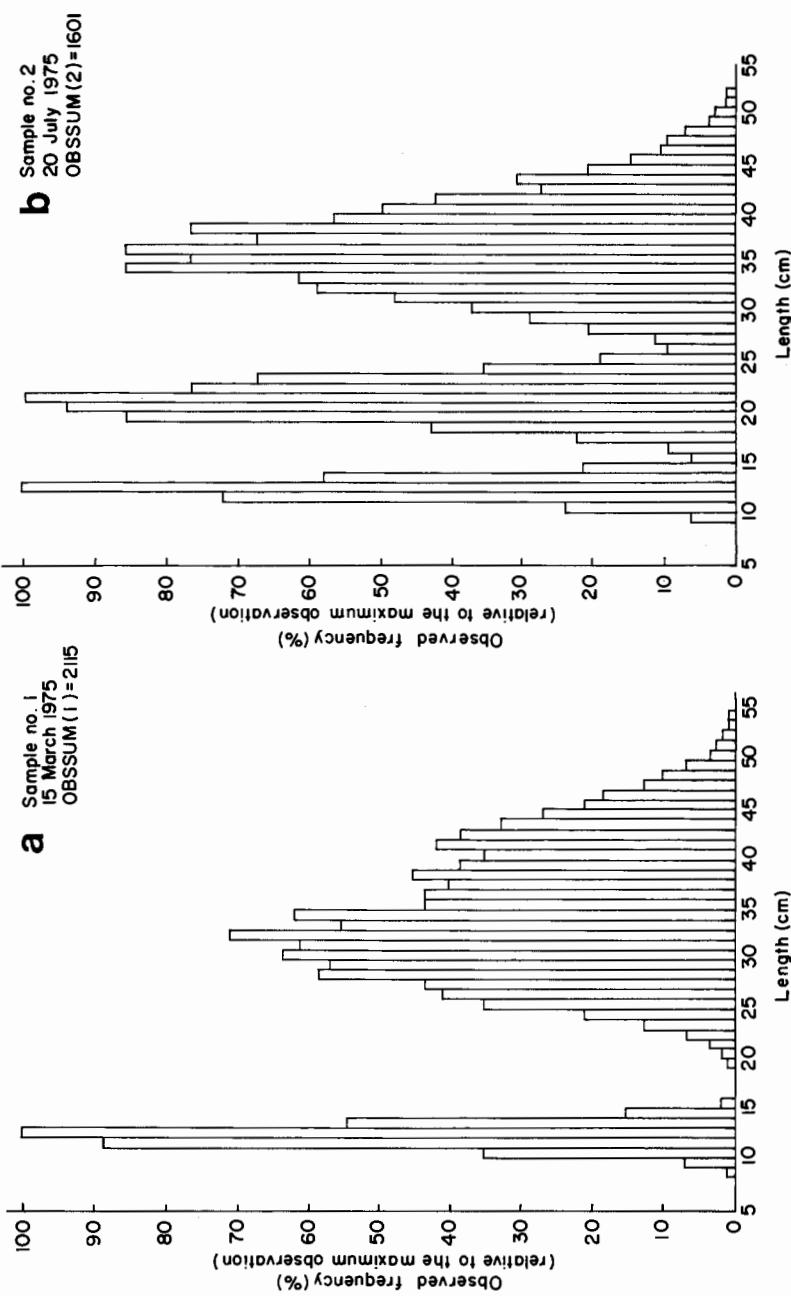


Fig. 1. Example of a set of length-frequency input data.

- a) OBS (1, L)/MAX (OBS (1, J))
J
- b) OBS (2, L)/MAX (OBS (2, J))
J
- c) OBS (3, L)/MAX (OBS (3, J))
J
- d) OBS (4, L)/MAX (OBS (4, J))
J

The basic estimation problem is to determine the value of THETA, which produces the best fit of the theoretical frequencies, CFREQ, to the observed ones, OBS.

As a measure of goodness of fit, the so-called modified chi-squared criterion is used.

CHI2 (THETA) =

$$\sum \sum \frac{(\text{OBS}(S, L) - \text{CFREQ}(S, L, \text{TIME}(S), \text{THETA}))^2}{\text{OBS}(S, L)} \quad \dots 4)$$

$\{(S, L) | \text{OBS}(S, L) > 0\}$

Table 2. Sample dates of illustrative samples (see also Table 1, Fig. 1).

S	Year(s)	Month(s)	Day(s)	Time(s)	DT(T)
1	1975	3	15	1975.2027	0.3479
2	1975	7	20	1975.5507	0.2301
3	1975	10	12	1975.7808	0.2329
4	1976	1	5	1976.0137	

Table 3. Structure of the four samples in Tables 1, 2, Fig. 1.

Sample no.	I	Birth yr STRUC (S, I, 1-2)	Spring/ fall 0/1	A (S, I)	MEANL (S, I)	SIGMA (S, I)	SURV (S, I)	Age dist.	X*	Y*	SEPINDE = X/Y
1	1	1975	0	0.203	11.99	1.20	700.0	0.30	15.59	1.98	7.8
	2	1974	1	1.203	27.59	2.76	400.0	0.17	5.34	3.03	1.8
	3	1974	0	1.203	32.92	3.29	600.0	0.26	7.01	3.64	1.9
	4	1973	1	2.203	39.93	3.99	200.0	0.09	2.40	4.11	0.6
	5	1973	0	2.203	42.33	4.23	400.0	0.17	—	—	—
2	1	1975	1	0.551	12.23	1.22	400.0	0.23	9.00	1.67	5.4
	2	1975	0	0.551	21.23	2.12	501.6	0.28	11.81	2.71	4.3
	3	1974	1	1.551	33.03	3.30	286.7	0.16	4.04	3.51	1.2
	4	1974	0	1.551	37.07	3.71	430.0	0.24	5.30	3.97	1.3
	5	1973	1	2.551	42.38	4.24	143.3	0.08	—	—	—
3	1	1975	1	0.781	18.59	1.86	329.9	0.23	7.48	2.23	3.3
	2	1975	0	0.781	26.07	2.61	413.8	0.28	9.82	3.10	3.2
	3	1974	1	1.781	35.89	3.59	236.4	0.16	3.36	3.76	0.9
	4	1974	0	1.781	39.25	3.93	354.7	0.24	4.41	4.15	1.1
	5	1973	1	2.781	43.66	4.37	118.2	0.08	—	—	—
4	1	1975	1	1.014	23.93	2.39	249.5	0.25	6.21	2.71	2.3
	2	1975	0	1.014	30.14	3.02	312.9	0.31	8.15	3.48	2.4
	3	1974	1	2.014	38.29	3.83	178.8	0.18	2.79	3.97	0.4
	4	1974	0	2.014	41.08	4.11	268.2	0.27	—	—	—

*X = MEANL (S, I + 1) - MEANL (S, I); Y = (SIGMA (S, I + 1) + SIGMA (S, I))/2.

$\{(S, L) | OBS(S, L) > 0\}$ is the set of all non-zero observations. Let NOBSP be the number of non-zero observations.

The value of THETA which minimizes CHI2 is (for large sample sizes) the maximum likelihood estimator of THETA, i.e., the value of THETA which assigns maximum probability to the observations. It can be shown that under certain assumptions CHI2 is asymptotically chi-squared distributed with NOBSP-NPAR-1 degrees of freedom (for large samples).

The method applied to minimize the weighted sum of squares of deviations (equation 4), is presented further below. An introduction to the general theory of estimation in non-linear models is given in Draper and Smith (1981).

Macdonald and Pitcher (1979) and Schnute and Fournier (1980) used another measure, namely:

$$2 * \sum_{\{(S, L) | OBS(S, L) > 0\}} \sum_{OBS(S, L)} \text{log}(OBS(S, L)/CFREQ(S, L, TIME(S), THETA))$$

However, as pointed out by the above-mentioned authors, this measure (and a few more alternatives) has asymptotically the same properties as equation (4). Thus, selection of a given function is merely a matter of taste and calculation convenience.

THE EXPONENTIAL DECAY MODEL

To define the model for CFREQ(S, L) in equation (1), it is convenient to start by considering

$$\text{TOTSUR}(S) = \sum_{L=1}^{\text{NOL}} \text{CFREQ}(S, L) \quad \dots 5)$$

and then gradually modify the model for this quantity until it is converted into a model for CFREQ. TOTSUR stands for total number of survivors. When the basic data are catch per unit of effort, TOTSUR(S) can be assumed proportional to the number of survivors from the stock at time TIME(S). Equation (5) expresses TOTSUR as the sum of length groups but for the present purpose another splitting is required, namely into broods (or cohorts):

$$\text{TOTSUR}(S) = \sum_I \text{SURV}(S, I) \quad \dots 6)$$

where I is index of cohort and SURV(S, I) is the number of survivors from cohort no. I at the time of sample S. (Actually, SURV is an index of number of survivors, which is tacitly assumed in the following). In the present model, one or two cohorts per year are assumed. As will be seen further below, the method is probably not capable of dealing with more than two cohorts per year. In the following two cohorts per year are assumed, as it is a simple thing to reduce the model to the case of one cohort per year.

A cohort (or brood) may be defined in the following way:

Spring cohort : All fish recruited from date TB0 to date TB1 (designated as "0-cohort" or "spring" cohort)

Autumn cohort : All fish recruited from date TB1 to date TB0 (designated as "1-cohort" or "fall" cohort)

A cohort is thus identified by the time the fish in it were hatched, i.e., (hatch year, season). Thus to each index, I, in equation (6) there is a corresponding pair:
 $\text{STRUC } (\text{S}, \text{I}, 2) = \text{hatch year}$

$$\text{STRUC } (\text{S}, \text{I}, 2) = \begin{cases} 0 & \text{if I is a spring cohort} \\ 1 & \text{if I is an autumn cohort} \end{cases} \dots 7)$$

The same cohort may be represented in several samples. The following connects the cohorts in the various samples:

$\text{STRUC } (\text{S}, \text{I}, 3) = \text{index of cohort strength (to be explained below)}$

$$\text{STRUC } (\text{S}, \text{I}, 4) = \begin{cases} \text{index of cohort I in the foregoing sample if the cohort did not recruit between samples S-1 and S.} \\ 0, \text{ if cohort I recruited between samples S-1 and S or if S = 1.} \end{cases} \dots 8)$$

Each cohort has a strength, i.e., the value of SURV when the cohort appears for the first time in a sample. These cohort strengths constitute a part of the parameter vector THETA, and the index $\text{STRUC } (\text{S}, \text{I}, 3)$ is used to identify this THETA-element, i.e.,

$$\text{THETA } (\text{STRUC}(\text{S}, \text{I}, 3)) = \text{SURV } (\text{S0}, \text{I0})$$

where

$$\begin{aligned} \text{STRUC } (\text{S}, \text{I}, 1) &= \text{STRUC } (\text{S0}, \text{I0}, 1) \\ \text{STRUC } (\text{S}, \text{I}, 2) &= \text{STRUC } (\text{S0}, \text{I0}, 2) \\ \text{STRUC } (\text{S0}, \text{I0}, 4) &= 0 \end{aligned} \dots 9)$$

(This is further explained in the example below).

The connection between

$$\text{SURV } (\text{S-1}, \text{STRUC } (\text{S}, \text{I}, 4)) \text{ and } \text{SURV } (\text{S}, \text{I})$$

in the case $\text{STRUC } (\text{S}, \text{I}, 4) \neq 0$ and $\text{S} > 1$ is the traditional exponential decay model:

$$\text{SURV } (\text{S}, \text{I}) = \text{SURV } (\text{S-1}, \text{STRUC } (\text{S}, \text{I}, 4)) * \text{EXP} (-\text{DT}(\text{S}) * \text{THETA}(9 + \text{S})) \dots 10)$$

Recall that $\text{DT}(\text{S}) = \text{TIME}(\text{S}) - \text{TIME}(\text{S-1})$, i.e., the time elapsed between samples S-1 and S.

$\text{THETA}(9 + \text{S})$ is the instantaneous rate of total mortality (usually denoted Z). Thus, in the traditional notation equation (10) reads:

$$N_{t + \Delta t} = N_t e^{-Z \Delta t}$$

The reason for defining the index $9 + \text{S}$ as it is, is that the first 10 elements of THETA are occupied by growth parameters and some other parameters which will be introduced further below.

Mortality does not have a cohort index, only a sample index. This is so because in this model, all cohorts are assumed to be exposed to the same total mortality, as soon as they are fully recruited and under full exploitation.

Fournier and Breen (1983) applied a similar model to analyze a single length-frequency sample. In this case, recruitment was assumed constant.

EXAMPLE:

The samples in the example above are composed of six cohorts with a structure as shown in Table 4. An alternative graphical description of the sample/cohort structure is shown in Fig. 2.

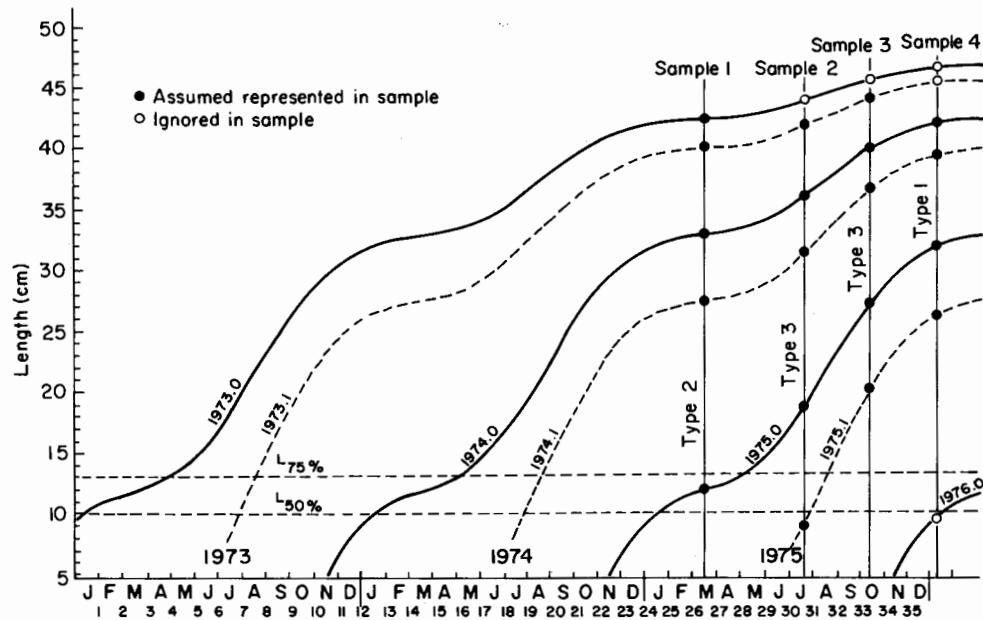


Fig. 2. Seasonally oscillating growth curves corresponding to data in Table 4 and illustrating the "Sample-Type" concept.

Table 4. Example of a sample/cohort structure (see also Fig. 2 and text).

Sample no.	1	2	3	4
I	STRUC (1, I, 1)	STRUC (2, I, 1)	STRUC (3, I, 1)	STRUC (4, I, 1)
	STRUC (1, I, 2)	STRUC (2, I, 2)	STRUC (3, I, 2)	STRUC (4, I, 2)
	STRUC (1, I, 3)	STRUC (2, I, 3)	STRUC (3, I, 3)	STRUC (4, I, 3)
	STRUC (1, I, 4)	STRUC (2, I, 4)	STRUC (3, I, 4)	STRUC (4, I, 4)
1	1975 0 14 0	1975 1 19 0	1975 1 19 1	1975 1 19 1
2	1974 1 15 0	1975 0 14 1	1975 0 14 2	1975 0 14 2
3	1974 0 16 0	1974 1 15 2	1974 1 15 3	1974 1 15 3
4	1973 1 17 0	1974 0 16 3	1974 0 16 4	1974 0 16 4
5	1973 0 18 0	1973 1 17 4	1973 1 17 5	— — — —

Table 5 shows an example of the parameter vector. This vector, of course, will not be known at the present stage of the analysis since THETA is the final result of the exercise. However, the input data given in Table 1 are constructed so that the parameters are those given in Table 4. For the hypothetical example used the true parameters are thus known. We may also consider Table 5 as the user's initial guess of the parameters. With the parameters given in Table 5, the numbers of survivors become those shown in Table 6.

Table 5. Example of (incomplete) parameter vector.

I	THETA (I)	Explanation
1	—	
2	—	various parameters (to be introduced later)
:	:	
:	:	
10	—	
11	1.0	Total mortality from TIME (1) to TIME (2)
12	0.8	Total mortality from TIME (2) to TIME (3)
13	1.2	Total mortality from TIME (3) to TIME (4)
14	700	Strength of cohort (1975, 0) (= SURV (1, 1))
15	400	Strength of cohort (1974, 1) (= SURV (1, 2))
16	600	Strength of cohort (1974, 0) (= SURV (1, 3))
17	200	Strength of cohort (1973, 1) (= SURV (1, 4))
18	400	Strength of cohort (1973, 0) (= SURV (1, 5))
19	400	Strength of cohort (1975, 1) (= SURV (2, 1))

Table 6. Example of calculation of the numbers of survivors, SURV. (THETA (S + 9) is the total mortality between samples S-1 and S and DT is the time period).

S	1	2	3	4
THETA (S + 9)	1.0	0.8	1.2	
DT (S)	0.3479	0.2301	0.2329	
I	SURV (1, I)	SURV (2, I)	SURV (3, I)	SURV (4, I)
1	700.0	400.0	329.9	249.5
2	400.0	501.6	413.8	312.9
3	600.0	286.7	236.4	178.8
4	200.0	430.0	354.7	268.2
5	400.0	143.0	118.2	—

NUMBER OF COHORTS IN A SAMPLE AND SAMPLE TYPE

The present method is not capable of estimating the number of cohorts represented in a sample. In case additional information is not available the user must guess the number of cohorts in each sample. If, e.g., subsamples of the length frequencies have been analyzed for age distribution by reading in hard parts (otoliths, scales or bones), this can be used to estimate how many cohorts are present in a sample.

If only length-frequency data are available, the user must provide a guess of the number of cohorts in each sample. Thus, this method is not a purely objective method, unless it is combined with additional information.

The array STRUC (S, I), I = 1, 2 in Table 4 must be given as input. However, the program does not take STRUC directly as input.

To secure a consistent sample/cohort structure, the program creates the array STRUC from the input parameters:

BM = maximum number of cohorts in any sample.

STYPE (S, 1) = type of sample S, S = 1, 2, . . . , NOS. . . 11)

The type of a sample can be 1, 2 or 3 and is defined:

Sample type 1 : The youngest cohort represented in the sample is the 1-cohort from the foregoing year. The number of cohorts is BM-1. (Sample no. 4 in Table 4 is a type 1 sample).

Sample type 2 : The youngest cohort is the 0-cohort of the sample year. The number of cohorts is BM. (Sample no. 1 of Table 4 is a type 2 sample).

Sample type 3 : The youngest cohort is the 1-cohort of the sample year. The number of cohorts is BM. (Samples 2 and 3 in Table 4 are type 3).

The second element in STYPE is

STYPE (S, 2) = number of cohorts in sample S . . . 12)

Table 7 illustrates the sample-type concept by the example of Table 4.

Table 7. Illustration of the sample type concept. (Compare Fig. 2 and Table 4).

Sample, S	1 1975.2027 type 2	2 1975.5507 type 3	3 1975.7808 type 3	4 1976.0137 type 1
I	STRUC (1, I, 1) STRUC (1, I, 2)	STRUC (2, I, 1) STRUC (2, I, 2)	STRUC (3, I, 1) STRUC (3, I, 2)	STRUC (4, I, 1) STRUC (4, I, 2)
1	1975 0	1975 1	1975 1	1975 (youngest 1 cohort)
2	1974 1	1975 0	1975 0	1975 0
3	1974 0	1974 1	1974 1	1974 1
4	1973 1	1974 0	1974 0	1974 0
5	1973 0	1973 1	1973 1	—
Number of cohorts STYPE (S, 2)	5 BM	5 BM	5 BM	4 BM-1

LENGTH DISTRIBUTION FOR GIVEN AGE

In order to introduce length groups into equation (6) it is written

$$\text{TOTSUR}(S) = \sum_I \sum_{L=1}^{\text{NOL}} \text{SURV}(S, I) * \text{CFREQ1}(S, I, L) \quad \dots 13$$

where

$\text{CFREQ1}(S, I, L)$ = the probability that a fish in sample S, cohort no. I belongs to length group L.

Changing the order of the summations gives:

$$\text{TOTSUR}(S) = \sum_{L=1}^{\text{NOL}} \sum_I \text{SURV}(S, I) * \text{CFREQ1}(S, I, L) \quad \dots 14$$

If $\text{CFREQ}(S, L)$ of equation (1) is defined by:

$$\text{CFREQ}(S, L) = \sum_I \text{SURV}(S, I) * \text{CFREQ1}(S, I, L), \quad \dots 15$$

then equation (5) has been converted into equation (1).

$\text{CFREQ1}(S, I, L)$ as a function of L is assumed to be a normal distribution. That is, for given age (given I and S), the probability that a fish belongs to length group L is:

$$\begin{aligned} \text{CFREQ1}(S, I, L) &= \\ &\text{FIRSTL} + DL*L \\ &\int \frac{1}{\text{SQRT}(2*\pi)*\text{SIGMA}(S, I)} \exp \left(-0.5 \left(\frac{X - \text{MEANL}(S, I)}{\text{SIGMA}(S, I)} \right)^2 \right) dX \quad \dots 16 \\ &\text{FIRSTL} + DL*(L - 1) \end{aligned}$$

We assume that there exists a linear relationship between MEANL and SIGMA (Schnute and Fournier 1980):

$$\text{SIGMA}(S, I) = \text{THETA}(5) + \text{THETA}(6)*\text{MEANL}(S, I) \quad \dots 17$$

The parameters $\text{THETA}(J)$, $J = 1, 2, 3, 4, 7$ and 8 are parameters of the seasonalized von Bertalanffy growth equation (von Bertalanffy 1934; Pauly and Gaschütz 1979).

MEANL is a function of these growth parameters, the age of the cohort and the time of the sample.

Let $\text{AGE}(S, I)$ be the age group of cohort I at $\text{TIME}(S)$.

$$\text{AGE}(S, I) = 0, 1, 2, 3, \dots$$

Let STIME(S) be the decimal part of TIME(S), i.e.,

$$\text{STIME}(S) = \text{TIME}(S) - \text{YEAR}(S) \quad \dots 18)$$

Let

$$A(S, I) = \text{AGE}(S, I) + \text{STIME}(S) \quad \dots 19)$$

Thus, A is the age of cohort I relative to the start of the hatching year of cohort I. MEANL is considered a function of A and not a function of absolute age. Usually, absolute age is used as the independent variable in the von Bertalanffy equation, but this is not possible in the present case because the absolute ages are not known.

MEANL is a function of

$A(S, I)$, $\text{STIME}(S)$, $\text{THETA}(J)$, $J = 1, 2, 3, 4, 7, 8$ and $\text{STRUC}(S, I, 2)$ ($= 0$ for spring cohorts and $= 1$ for fall cohorts)

Usually the von Bertalanffy growth equation in its simplest form is written:

$$L_t = L_\infty (1 - e^{-K(t - t_o)})$$

and in the seasonalized form (Pauly and Gaschütz 1979):

$$L_t = L_\infty (1 - e^{-k_o [t - t_o + (C/2\pi)*\sin(2\pi(t - t_w))]})$$

where t is the absolute age of the cohort.

In the Pauly and Gaschütz model K is assumed to be a function of time:

$$K(t) = k_o (1 - C \cos(2\pi(t - t_w)))$$

on the assumption that the cosine reflects seasonal oscillations of temperature, food availability and/or anything else influencing the growth rate. At time t_w (the winter point) K take its lowest value.

The six growth parameters used in the present model are closely related to those defined by Pauly and Gaschütz (1979):

THETA (1) : Asymptotic length (L_∞ , in the traditional notation)

THETA (2) : Curvature parameter (k_o)

THETA (3) : Initial condition parameter for the spring cohort with average birthday Tbo
 $((t_o + Tbo)*k_o)$

THETA (4) : Initial condition parameter for the fall cohort with average birthday Tb1
 $((t_o + Tb1)*k_o)$

THETA (7) : Amplitude of seasonal oscillation of K ($C/(k_o*2\pi)$)

THETA (8) : Winter point (t_w).

The absolute age is

$t = A(S, I) - Tbo$ for the spring cohort

$t = A(S, I) - Tb1$ for the fall cohort

The birthdays Tbo and Tb1 cannot be separated from t_o in the present estimation procedure, and consequently they are not used as parameters. Alternatively the parameters THETA (3) and THETA (4) which can be estimated are used in the growth equation.

With the growth parameters given above the growth equation reads:

MEANL (S, I) =

$$\begin{aligned} \text{THETA}(1)* & \left(1 - \exp \left\{ -\text{THETA}(2)*A(S, I) + \frac{\text{THETA}(3)}{\text{THETA}(4)} \right\} + \right. \\ & \left. + \text{THETA}(7)*\sin[2\pi*(\text{STIME}(S) - \text{THETA}(8))] \right\} \end{aligned} \quad \dots 20)$$

EXAMPLE:

We are now in a position to continue the specification of the parameter vector THETA initiated in Table 5. The extended parameter vector is shown in Table 8.

Fig. 2 shows the growth curves for the two cohorts corresponding to the parameters given in Table 8. Fig. 2 is also used to illustrate the sample-type concept.

Figs. 4A-D show examples of the separation of length-frequency samples into normally distributed components (see also Table 3).

Note that the youngest cohorts are underrepresented in Figs. 4A-D as compared to Table 6.

The reason for this discrepancy is that the number of survivors has been adjusted to account for low catchability due to gear selection and recruitment as discussed below.

Table 8. Example of extended (but still incomplete) parameter vector.

I	THETA (I)	Explanation
1	50.0	Asymptotic length, L_{∞}
2	0.8	Curvature parameter, Ko
3	-0.112	Initial condition parameter for spring cohort Ko (to + Tbo)
4	0.160	Initial condition parameter for autumn cohort Ko (to + Tb1)
5	0.0	Intercept in SIGMA
6	0.1	Slope in SIGMA
7	0.1019	Amplitude in seasonal osc. CK/2π
8	0.2	Winter point, tw
9	—	Gear selection parameter
10	—	(To be introduced later)
11	1.0	Total mortality from TIME (1) to TIME (2)
12	0.8	Total mortality from TIME (2) to TIME (3)
13	1.2	Total mortality from TIME (3) to TIME (4)
14	700	Strength of cohort (1975, 0) (= SURV (1, 1))
15	400	Strength of cohort (1974, 1) (= SURV (1, 2))
16	600	Strength of cohort (1974, 0) (= SURV (1, 3))
17	200	Strength of cohort (1973, 1) (= SURV (1, 4))
18	400	Strength of cohort (1973, 0) (= SURV (1, 5))
19	400	Strength of cohort (1975, 1) (= SURV (2, 1))

RECRUITMENT/GEAR SELECTION

Some fish may have reduced probability of being caught because they are either not on the fishing grounds (i.e., the area covered by the survey) or because of the selectivity of the gear (e.g., they may be so small that they are able to escape through the meshes of trawl's cod end).

A reduction factor expressed as a function of length of the fish (GCORR (L)) is thus defined:

$$\text{GCORR (L)} = \frac{\text{Probability of being caught in length group L}}{\max_L \{ \text{Probability of being caught} \}} \quad \dots 21)$$

Thus, $0 \leq \text{GCORR (L)} \leq 1.0$.

It is assumed that GCORR is dependent only on L (and not on S). This implies that the exploitation pattern of the stock is assumed to remain constant during the sampling period. The

migration pattern is assumed to be a function of the length of the fish only, and independent of the season of the year. The model for GCORR (L) presented by Hoydal et al. (1982) is used here:

$$\text{GCORR} (L) = \left(1 + \exp [\text{THETA} (9) + (\text{FIRSTL} + (L + 0.5) * \text{DL}) * \text{THETA} (10)] \right)^{-1} \dots 22)$$

The parameters can be interpreted in the following way:

Let L50% and L75% be the lengths at which the probability of being caught is 50% and 75% of the maximum probability, respectively. Then

$$\begin{aligned}\text{THETA} (9) &= \log(3) * \text{L50\%} / (\text{L75\%} - \text{L50\%}) \\ \text{and} \\ \text{THETA} (10) &= -\text{THETA} (9) / \text{L50\%}\end{aligned}$$

As defined by equation 15 with no regard to recruitment/gear selection we have:

$$\frac{\text{OBS} (S, L)}{\text{CFREQ} (S, L)} = \text{GCORR} (L)$$

Thus, if CFREQ is replaced by the adjusted value:

$$\text{CFREQ} (S, L) * \text{GCORR} (L) \dots 23)$$

the model is expected to give the best fit to the data.

EXAMPLE:

Table 9 shows the complete parameter vector, i.e., including the recruitment/gear selection parameters.

Fig. 3 shows the recruitment/gear selection ogives corresponding to the parameter values given in Table 9.

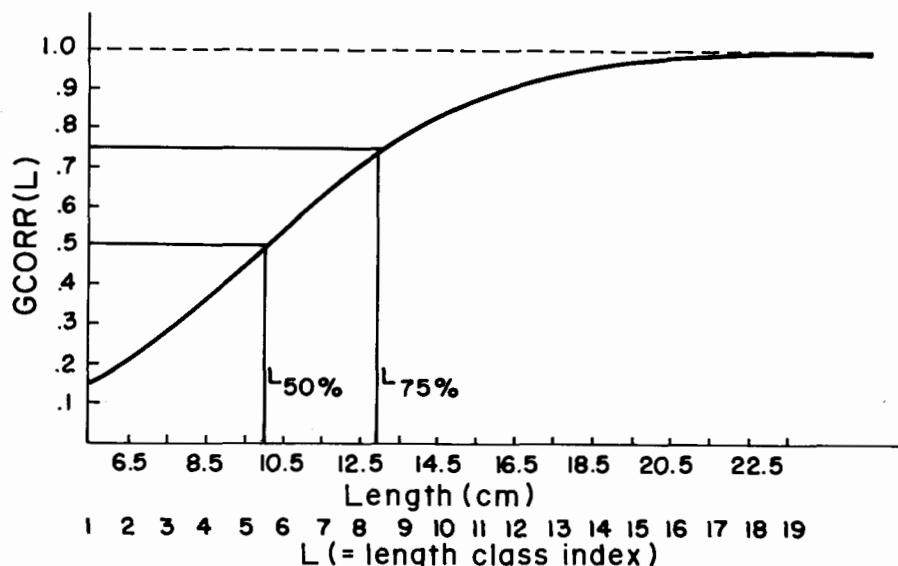


Fig. 3. Recruitment/selection ogive as used to correct length-frequency data for effects of incomplete recruitment and or selection (see text).

Table 9. Example of a complete parameter vector.

I	THETA (I)	Explanation
1	50.0	Asymptotic length, L_∞
2	0.8	Curvature parameter, K_o
3	-0.112	Initial condition parameter for spring cohort K_o ($t_0 + T_{bo}$)
4	0.160	Initial condition parameter for autumn cohort K_o ($t_0 + T_{bo1}$)
5	0.0	Intercept in SIGMA
6	0.1	Slope in SIGMA
7	0.1019	Amplitude in seasonal osc. $CK/2\pi$
8	0.2	Winter point, t_w
9	3.662	Gear selection parameter ($L_{50\%} = 10$ cm)
10	-0.3662	Gear selection parameter ($L_{75\%} = 13$ cm)
11	1.0	Total mortality from TIME (1) to TIME (2)
12	0.8	Total mortality from TIME (2) to TIME (3)
13	1.2	Total mortality from TIME (3) to TIME (4)
14	700	Strength of cohort (1975, 0) (= SURV (1, 1))
15	400	Strength of cohort (1974, 1) (= SURV (1, 2))
16	600	Strength of cohort (1974, 0) (= SURV (1, 3))
17	200	Strength of cohort (1973, 1) (= SURV (1, 4))
18	400	Strength of cohort (1973, 0) (= SURV (1, 5))
19	400	Strength of cohort (1975, 1) (= SURV (2, 1))

CONCISE MATHEMATICAL DEFINITION OF THE ESTIMATION PROBLEM

This section presents a stringent resumé of the mathematics introduced above. Special reference is given to the functional relationships, which may be difficult to see through in the mixture of words and formulas used above.

The estimation problem is defined as follows:

Determine the value of the parameter vector

$$\text{THETA} = (\text{THETA} (1), \text{THETA} (2), \dots, \text{THETA} (\text{NPAR}))$$

which for given values of BM and STYPE minimizes the object function:

$$\text{CHI2} (\text{OBS}, \text{BM}, \text{STYPE}, \text{THETA}) =$$

$$= \sum_{(S, L)} \sum_{\{ (S, L) | \text{OBS}(S, L) > 0 \}} \text{RES} (\text{OBS}(S, L), \text{TIME}(S), \text{BM}, \text{STYPE}, \text{THETA})^2 \quad \dots 24)$$

where

$$\text{RES} (\text{OBS}(S, L), \text{BM}, \text{STYPE}, \text{THETA}) = \frac{\text{OBS}(S, L) - \text{CFREQ}(L, S, \text{TIME}(S), \text{BM}, \text{STYPE}, \text{THETA})}{\text{SQRT}(\text{OBS}(S, L))} \quad \dots 25)$$

$\text{OBS}(S, L)$ = observations: The average number caught per unit of effort in sample S , belonging to length group L . $S = 1, 2, \dots, \text{NOS}$. $L = 1, 2, \dots, \text{NOL}$.

BM and STYPE may come from an independent investigation or be based on a qualified guess.

No general rule for selecting among competing solutions with different values of BM and STYPE is given. In that sense the method is subjective.

The theoretical frequency is defined by:

$$\text{CFREQ}(\text{S}, \text{L}, \text{TIME}(\text{S}), \text{BM}, \text{STYPE}, \text{THETA}) = \sum_{\text{I} = 1}^{\text{STYPE}(\text{S}, 2)} \text{SURV}(\text{S}, \text{I}, \text{STRUC}, \text{THETA}) * \text{CFREQ1}(\text{S}, \text{I}, \text{L}, \text{TIME}(\text{S}), \text{STRUC}, \text{THETA}) * \text{GCORR}(\text{L}, \text{THETA}) \quad \dots 26)$$

I is a cohort index and is related to the vector STYPE as follows:
 BM = maximum number of cohorts in any sample.

- STYPE ($\text{S}, 1$) = type of sample S .
- STYPE ($\text{S}, 2$) = number of cohorts in sample S . STYPE ($\text{S}, 2$) is determined by the value of BM and STYPE ($\text{S}, 1$):
- STYPE ($\text{S}, 1$) = 1: The youngest cohort is the autumn-cohort from foregoing year. STYPE STYPE ($\text{S}, 2$) = BM - 1
- STYPE ($\text{S}, 1$) = 2: The youngest cohort is the spring-cohort of the sample year. STYPE STYPE ($\text{S}, 2$) = BM
- STYPE ($\text{S}, 1$) = 3: The youngest cohort is the autumn-cohort of the sample year. STYPE STYPE ($\text{S}, 2$) = BM

The vector STRUC is determined by BM, STYPE and TIME:

$$\text{STRUC}(\text{S}, \text{I}, 2) = \begin{cases} 0 & \text{if I is a spring cohort} \\ 1 & \text{if I is an autumn cohort} \end{cases}$$

hatching year of cohort I

$$\text{STRUC}(\text{S}, \text{I}, 1) = \begin{cases} \text{THETA-index of cohort strength, i.e.,} \\ \text{SURV}(\text{SO}, 10) = \text{THETA}(\text{STRUC}(\text{S}, \text{I}, 3)) \text{ when} \\ \text{STRUC}(\text{S}, \text{I}, 1) = \text{STRUC}(\text{SO}, 10, 1), \\ \text{STRUC}(\text{S}, \text{I}, 2) = \text{STRUC}(\text{SO}, 10, 2) \text{ and} \\ \text{STRUC}(\text{SO}, 10, 4) = 0 \end{cases}$$

$$\text{STRUC}(\text{S}, \text{I}, 4) = \begin{cases} \text{index of cohort I in the foregoing sample if the cohort did not recruit} \\ \text{between samples S-1 and S.} \\ 0, \text{ if cohort I recruited between samples S-1 and S or if } \text{S} = 1. \end{cases}$$

The adjustment factor for gear selection/recruitment:

$$\text{GCORR}(\text{L}, \text{THETA}) = \left(1 + \exp \left[\text{THETA}(9) + (\text{FIRSTL} + (\text{L} + 0.5) * \text{DL}) * \text{THETA}(10) \right] \right)^{-1} \quad \dots 27)$$

The probability that a cohort I fish in sample S belongs to length group L [$\text{FIRSTL} + \text{DL} * (\text{L} - 1)$, $\text{FIRSTL} + \text{DL} * \text{L}$], is:

$$\text{CFREQ1}(\text{S}, \text{I}, \text{L}, \text{TIME}(\text{S}), \text{STRUC}, \text{THETA}) =$$

$$\text{FIRSTL} + \text{DL} * \text{L}$$

$$\int_{\text{FIRSTL} + \text{DL} * (\text{L} - 1)}^{\text{FIRSTL} + \text{DL} * \text{L}} \frac{1}{\text{SQRT}(2 * \pi) * \text{SIGMA}(-)} \exp \left(-0.5 \left(\frac{\text{X} - \text{MEANL}(-)}{\text{SIGMA}(-)} \right)^2 \right) d\text{X} \quad \dots 28)$$

$$\text{FIRSTL} + \text{DL} * (\text{L} - 1)$$

where “(-)” stands for “($\text{S}, \text{I}, \text{TIME}(\text{S}), \text{STRUC}, \text{THETA}$)”

The standard deviation of the length distribution is:

$$\text{SIGMA}(\text{S}, \text{I}, \text{TIME}(\text{S}), \text{STRUC}, \text{THETA}) = \text{THETA}(5) + \text{THETA}(6) * \text{MEANL}(\text{S}, \text{I}, \text{TIME}(\text{S}), \text{STRUC}, \text{THETA}) \quad \dots 29)$$

The mean length of cohort I in sample S is given by the seasonalized von Bertalanffy growth equation:

$$\text{MEANL (S, I, TIME(S), STRUC, THETA)} =$$

$$\begin{aligned} & \text{THETA (1)*(1 - exp (-THETA (2)*A (S, I) +}} \\ & + \text{THETA (3)*(1 - STRUC (S, I, 2))} + \text{THETA (4)*STRUC (S, I, 2) +} \\ & + \text{THETA (7)*(sin (2}\pi\text{*}(STIME(S) - THETA (8))))} \end{aligned}$$

where

... 30)

$$\begin{aligned} A (S, I) &= AGE (S, I) + STIME(S) \\ AGE (S, I) &= STRUC (S, I, 1) - YEAR(S) \\ STIME(S) &= TIME(S) - YEAR(S) \\ TIME(S) &= YEAR(S) + (MO(MONTH(S)) + DAY(S))/365. \end{aligned}$$

where MO(1) = 0, MO(2) = 31, MO(3) = 89, ..., MO(12) = 334.

The number of survivors in cohort I in sample S is:

$$= \begin{cases} \text{SURV (S-1, STRUC (S, I, 4), TIME(S), THETA)*} \\ * \exp (-DT(S)*THETA (9 + S)) & \text{if } STRUC (S, I, 4) \neq 0 \\ \text{THETA (STRUC (S, I, 3))} & \text{if } STRUC (S, I, 4) = 0 \end{cases} \quad \dots 31)$$

$$\text{where } DT(S) = \text{TIME}(S) - \text{TIME}(S-1). \quad S = 2, 3, \dots, \text{NOS}.$$

The complete parameter vector is defined in Table 10.

Table 10. Parameter vector.

I	THETA (I)	I	THETA (I)
1	Asymptotic length, L_∞	12	Total mortality from TIME (2) to TIME (3)
2	Curvature parameter, K_o	.	• • •
3	Initial condition parameter for spring cohort : K_o (to + T_{bo})	.	• • •
4	Initial condition parameter for autumn cohort : K_o (to + T_{b1})	.	• • •
5	Intercept in SIGMA	9 + NOS	Total mortality from TIME (NOS-1) to TIME (NOS)
6	Slope in SIGMA	10 + NOS	Strength of cohort (STRUC (1, 1, 1), STRUC (1, 1, 2))
7	Amplitude in seasonal oscillation: $CK_o/2\pi$	11 + NOS	Strength of cohort (STRUC (1, 2, 1), STRUC (1, 2, 2))
8	Winter point, t_w	12 + NOS	Strength of cohort (STRUC (1, 3, 1), STRUC (1, 3, 2))
9	Gear selection parameter: $\log(3)*L50%/(L75%-L50%)$.	• • •
10	Gear selection parameter: — THETA (9)/L50%	.	• • •
11	Total mortality from TIME (1) to TIME (2)	.	• • •
		NPAR ^a	Strength of cohort (STRUC (S1, I1, 1), STRUC (S1, I1, 2)) ^b

^aNPAR = 9 + NOS + number of cohorts.

^bS1 is the index for the last sample with a recruiting cohort. I1 is the index of the youngest recruit cohort in sample S1.

BASIC ALGORITHM WITH CONSTRAINED ESTIMATION

In the present context the algorithm which defines the residuals (equation 25) in the optimization problem (equation 24) is considered the "basic algorithm".

The minimization of the object function (equation 4 or 24) might also have been considered the basic algorithm, but since this is a standard technique which can be used to minimize any sum of squares, it is not considered worthwhile to go into detail about it in the present context.

However, it should be noted that the minimization of the object function (equation 24) is solved by the NAG-subroutine E04FCF (NAG-Library 1981). There exist a number of subroutine libraries containing subroutines similar to E04FCF, and any of these could have been used.

The routine E04FCF is of the Gauss-Newton type and is based on the algorithm by Gill and Murray (1978).

E04FCF does not require derivatives, but it assumes that continuous first and second derivatives do exist (this assumption is fulfilled in the present case).

E04FCF requires from the user the subroutine which calculates the residuals and also an initial guess on the parameters, the algorithm being an iterative process.

Before the basic algorithm for the calculation of residuals is presented, an outline of the main program is given:

- A: Read observations, OBS, max number of cohorts, BM and sample types STYPE.
- B: Read initial guess on the parameters, THETA.
- C: Calculate the residuals for the current THETA-value.
- D: Calculate the object function, CHI2 (THETA).
- If CHI2 (THETA) is minimum then go to F. (D is performed by E04FCF.)
- E: Change the value of THETA in a way which may lead to lower CHI2-value; go to C.
(E is performed by E04FCF.)
- F: Print results.

So far THETA has been considered a variable, NPAR-dimensional vector of unknown parameter values. However, the program allows for some of the THETA-elements to remain constant (i.e., known) or some elements to be proportional to other elements. For this purpose a parameter vector, PARAM, is introduced. Actually PARAM is the parameter vector used in equation (24) instead of THETA.

If we introduce the symbol CHI2P for the object function as a function of PARAM, then

$$\text{CHI2P}(\text{PARAM}) = \text{CHI2}(\text{THETA})$$

where

$$\begin{aligned} \text{THETA}(I) &= \text{PARAM}(\text{PPOINT}(I)) * \text{PFACTOR}(I) && \text{if } \text{PPOINT}(I) > 0 \\ \text{THETA}(I) &\text{ kept constant (= initial guess)} && \text{if } \text{PPOINT}(I) = 0 \end{aligned} \quad \dots 32$$

PPOINT and PFACTOR are input to be selected and entered by the user.

Suppose the total mortality remains constant for all periods, then

$$\begin{aligned} \text{PPOINT}(11) &= \text{PPOINT}(12) = \dots = \text{PPOINT}(9 + \text{NOS}) = 11 \text{ and} \\ \text{PFACTOR}(11) &= \text{PFACTOR}(12) = \dots = \text{PFACTOR}(9 + \text{NOS}) = 1.0 \end{aligned}$$

would produce an estimate of "common" total mortality, i.e., an estimation under the constraint:

$$\text{THETA}(11) = \text{THETA}(12) = \dots = \text{THETA}(9 + \text{NOS})$$

Entering, e.g., PPOINT(11) = 1.0 and PPOINT(12) = 2.0 would produce an estimation of total mortality under the constraint that the total mortality between first and second sample is double that between second and third sample.

Similar constraints can be imposed on the cohort strengths.

The algorithm for calculation of residuals works as follows:

- A: Select subset PARAM of THETA (equation 32)
- B: S := 1 (index of sample)
- C: I := 1 (index of cohort)
- D: Calculate relative age of cohort I (equation 19)
Calculate mean length of cohort I (equation 20 or 30)
Calculate standard deviation of cohort I (equation 17 or 29)
- E: Calculate theoretical frequencies (equation 26)
- G: Correct for gear selection/recruitment (equation 22 or 27)
- H: If I < STYPE (S, 2) then I := I + 1; go to D.
- I: If S < NOS then S := S + 1; go to C.
- J: Calculate residuals (for all samples) (equation 25).

Morgan and Pauly (Part II, this vol.) present the FORTRAN 77 program corresponding to this algorithm along with practical hints on how to run the program.

SELECTING AMONG COMPETING SOLUTIONS

The basic estimation problem was defined above as the minimization of CHI2 with respect of THETA (equation 4).

However, the implication is not that in general the parameter set which minimizes CHI2 is necessarily the best one. In case the number of cohorts (BM) and the sample types, STYPE, are kept constant (and thus also the number of parameters, NPAR, kept constant), the minimum value of CHI2 can be used to select the best estimate of THETA.

But CHI2 for two competing solutions with different number of cohorts and, thus, different NPAR-values should not be compared. Usually (but not necessarily) a better fit is achieved the more parameters the model contains. To use the THETA value which minimizes CHI2 for all values of NPAR may lead to a very poor choice (Schnute and Fournier 1980).

I am not in a position to suggest a solution to the problem as far as statistical testing is concerned. If I were, the analysis would not have the subjective component mentioned above. In fact, I do not believe that any statistical test exists, which would make a judgment based on biological considerations superfluous.

Nevertheless, some guidance can be extracted from the CHI2-values for different NPAR-values. It can be shown that CHI2 is asymptotically chi-squared distributed with NOBS-NPAR-1 degrees of freedom. However, the distribution of CHI2 is usually unknown (Cramer 1961). Schnute and Fournier (1980) recommend the chi-squared test for choosing among competing solutions. They do not consider it a final criterion, but a useful guide.

Schnute and Fournier (1980) did not apply expression (24) as the chi-squared statistics. The approximation of equation (24) to the chi-squared distribution becomes poor when some intervals contain few observations (say, less than three). They suggested, rather, the use of what they called the Br-statistics:

$$Br = \sum_{\{(S, L) | OBS(S, L) \geq R\}} \frac{(OBS(S, L) - CFREQ(S, L))^2}{CFREQ(S, L)} + \frac{(SMALLOBS - SMALLC)^2}{SMALLC}$$

where

$$SMALLOBS = \sum_{\{(S, L) | OBS(S, L) < R\}} OBS(S, L) \text{ and } SMALLC = \sum_{\{(S, L) | OBS(S, L) < R\}} CFERQ(S, L)$$

and R is some small number, say 2, 3 or 4.

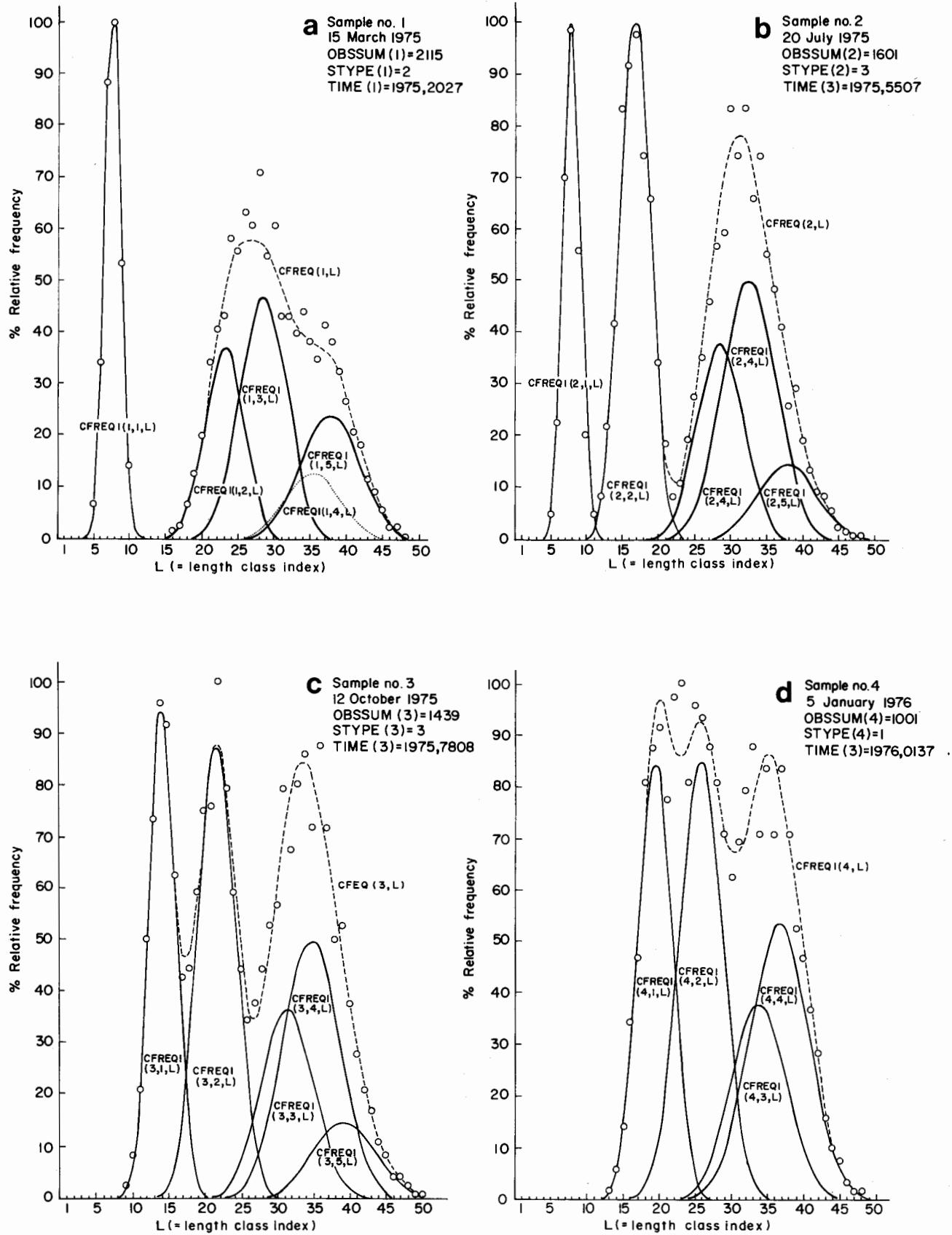


Fig. 4. Separation of length-frequency samples into their normally distributed components (see text).

Let N1 be the number of observations greater than or equal to R. Schnute and Fournier (1980) assume Br to be approximately chi-squared distributed with N1-NPAR degrees of freedom.

However, rather than trying to find a solution in the theory of hypothesis testing, it is as important to search for solutions which make biological sense.

If we were so lucky that some results from direct estimation of age composition (otoliths, scales or tagging experiments) were available from an earlier period or from another stock of the species in question and some estimate of the total mortality, we would have an estimate of BM and the problem would vanish. If this is not the case, we are left with common sense as the only tool. The maximum number of cohorts represented in the samples is not only determined by the biological maximum longevity of the species but also the mortality.

Related to the comparison of competing solutions are the variances and covariances of the estimate of THETA. Large covariances indicate that ranges of THETA-estimates may give nearly equal goodness of fit between observations and model.

An exact expression for the variance-covariance matrix VAR (estim. THETA) does not exist in the general case. It can be shown asymptotically that (Kendall and Stuart 1961):

$$\text{VAR} (\text{estim. THETA}) = (D^T * D)^{-1} \quad \text{where } D \text{ is the Jacobian:}$$

$$D = \left\{ \frac{1}{\text{SQRT}(\text{OBS}(S, L))} * \frac{\partial \text{CFREQ}(S, L)}{\partial \text{THETA}(J)} \right\}$$

D is a matrix with NPAR columns and NOBS (= number of non-zero observations) rows.

The Jacobian is calculated by E04FCF as a part of the optimization process, so technically it is easy to obtain estimates of approximate variances and covariances.

Discussion

The question to be discussed is not so much the validity of the model presented here. It does not contain any new or controversial approaches in fish stock assessment or statistical analysis. Rather, the question is whether the estimation techniques will work in practice. Even though the model is the best possible one, we may not be able to achieve the quality and quantity of data necessary for a dependable estimate of the parameters.

The problem may be structured so that there exists a range of solutions all of which give nearly equal goodness of fit to the observations, i.e., parameter estimates may be highly correlated. This obviously has something to do with how pronounced the structures representing the cohorts are.

In case a cohort appears as a peak in the length distribution, there is no big problem (e.g., cohort no. 1 in sample 1, Fig. 4A). But when they overlap considerably (as, e.g., the four oldest cohorts in sample 1, Fig. 4A), very large unbiased samples are required. In case only one sample, e.g., sample 1, was available it would simply not be possible to arrive at any safe estimate of the cohort components. For all four samples in the example only dependable estimates for the two youngest cohorts can be expected.

In sample 4 (Fig. 4D) the two youngest cohorts overlap considerably, but because samples 1-3 provided good estimates of them, they are not likely to create problems in sample 4. This example underlines the importance of collecting samples covering the entire length range. If samples contained only fish from, say, 20 cm and onwards, the data would be of little use.

The interesting quantity in this connection is the ratio between the standard deviation, SIGMA, and the difference between mean-lengths, MEANL, of consecutive cohorts, i.e.,

$$\text{SEPINP} = \frac{\text{MEANL}(S, I+1) - \text{MEANL}(S, I)}{(\text{SIGMA}(S, I+1) + \text{SIGMA}(S, I))/2} \quad \dots 33)$$

SEPIIND stands for separation index — degree of overlapping.

Hasselblad (1966) and McNew and Summerfelt (1978) have observed that if this quantity is less than two, the estimation on the normally distributed components becomes poor. Clark (1981) using simulation techniques arrives at a similar result.

The values of SEPIIND (equation 33) are given in Table 3. Notice that in this example SEPIIND-values greater than two produce peaks for the youngest cohorts.

Fig. 5 shows the effect of a doubling and a halving of SEPIIND. In this case THETA (5) = 0 (see Table 9) so that:

$$\text{SIGMA} (S, I) = \text{THETA} (6) * \text{MEANL} (S, I)$$

Thus, a doubling of THETA (6) produces a halving of SEPIIND and vice versa.

The data simulated for THETA (6) = 0.20 and 0.05 are not subject to random noise as was the case for the data in Table 1 (and Fig. 1). For comparison the graphs of Fig. 1 is also shown in Fig. 5.

Obviously the values of SEPIIND have a great influence on how precisely the cohort characteristics can be estimated. This is illustrated in Fig. 5 by the true growth curves (those corresponding to the parameters from which the data were generated). McNew and Summerfelt (1978) investigated

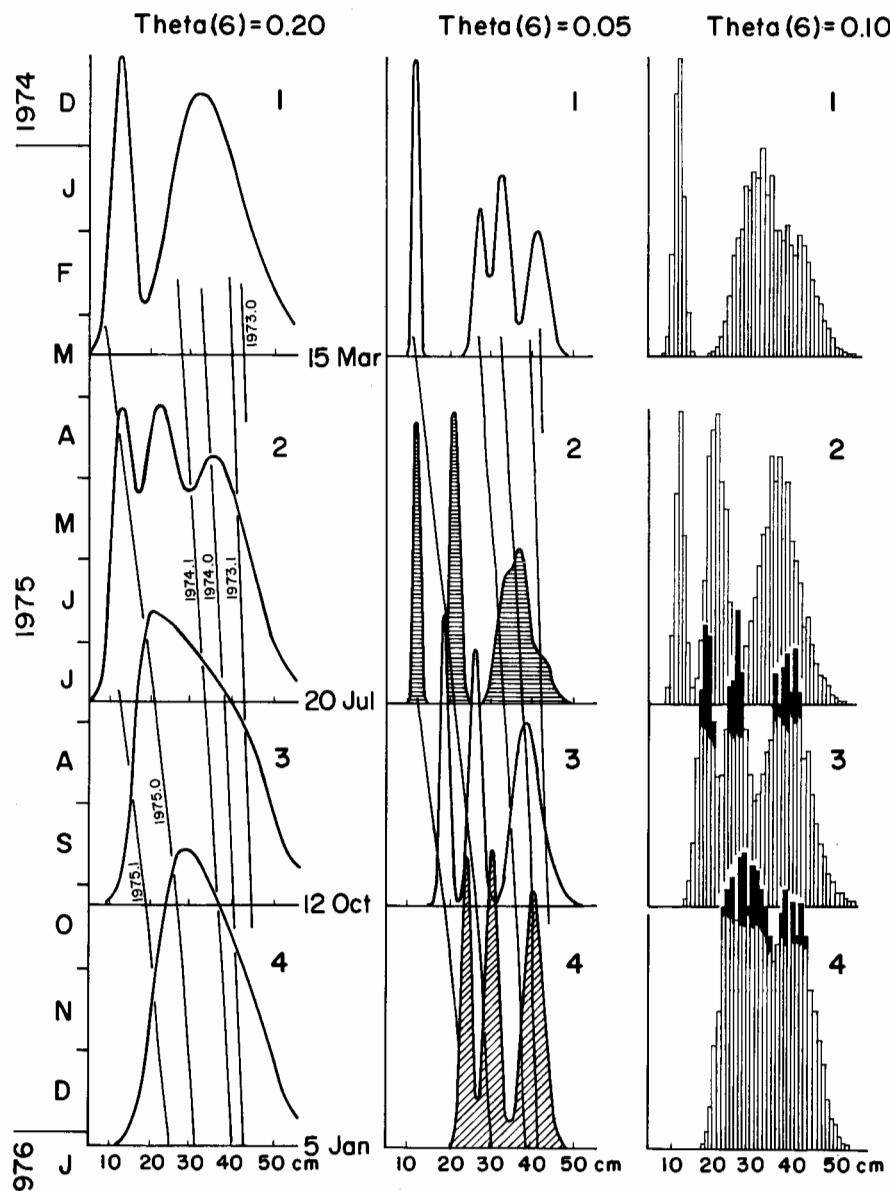


Fig. 5. Illustration of the effect of doubling ($\text{THETA} (6) = 0.20$) and halving ($\text{THETA} (6) = 0.05$) the separation index SEPIIND (see also Fig. 4A-D).

the problems of separating length frequencies of *Micropterus salmoides*. Because ages were directly measured from scale reading the true age compositions were known allowing for an evaluation of the results of the length-frequency analysis. It would be very interesting to carry out a similar study on a tropical fish stock (e.g., based on reading of daily rings in otoliths).

Such investigations may give useful guidelines for planning of sampling schemes. Especially sample sizes and sample frequencies necessary for a reasonable safe estimation is of utmost importance for the applicability of the method.

However, as pointed out by Foucher et al. (1984), the age readings may be so questionable that one should rather use the length-frequency analysis to test the age readings. Foucher et al. (1984) compared results from age readings and length-frequency analysis (the Schnute and Fournier method) on Pacific cod and found a considerable discrepancy between results of the two methods. They concluded that the results from length-frequency analysis in that particular case were the most reasonable ones.

In case the data are not C.P.U.E. data, but length frequencies not associated with effort, the algorithm can still be applied, but in that case the estimates of the mortality rates and cohort strengths have no meaning. However, even if the values of the estimates are of no use, their use will ensure that the relatively large and the relatively small cohorts occur in the appropriate proportions in all samples. Even negative mortality rates may be useful tools to force a biological structure on the solution.

In case of only one length-frequency sample, the model can still work. If the sample is entered twice—the first time with actual sampling date and next time with a date one year later—constant recruitment can be assumed, and the growth parameters THETA (J), J = 1, 2, 3, 4 can be estimated. Estimation of the seasonality parameters THETA (J), J = 7, 8 requires at least four samples per year. Under the assumption of constant recruitment, this application of the algorithm provides also estimates of total mortality rates (as an annual average).

In case only one annual recruitment period is assumed, the model is adapted to this assumption simply by assigning the fixed value 0 to every second cohort strength.

Obviously, the reliability of parameter estimates decreases with the number of samples and appropriate sampling schemes should include far more samples than are absolutely necessary.

References

- Bertalanffy, L. von. 1934. Untersuchungen über die Gesetzmäßigkeiten des Wachstums. I. Roux' Archiv Entwicklungsmech. 131: 613-652.
- Bhattacharya, C.G. 1967. A simple method of resolution of a distribution into Gaussian components. Biometrics 23: 115-135.
- Cassie, R.M. 1954. Some uses of probability paper in the analysis of size frequency distributions. 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.
- Cramer, H. 1961. Mathematical methods of statistics. Princeton University Press.
- Fournier, D.A. and P.A. Breen. 1983. Estimation of abalone mortality rates with growth analysis. Trans. Amer. Fish. Soc. 112: 403-411.
- Foucher, R.P., R.G. Bakkala and D. Fournier. 1984. Comparison of age frequency derived by length-frequency analysis and scale reading for Pacific cod in the Northern Pacific Ocean. Int. North Pac. Fish. Comm. Bull. 42: 232-242.
- Draper, N. and H. Smith. 1981. Applied regression analysis. (2nd ed.) John Wiley and Sons, New York.
- Gill, P.E. and W. Murray. 1978. Algorithms for the solution of the non-linear least squares problem. SIAM J. Numer. Anal. 15: 977-992.
- Hald, A. 1952. Statistical theory with engineering applications. Wiley, New York.
- Harding, J.P. 1949. The use of probability paper for the graphical analysis of polymodal frequency distributions. J. Mar. Biol. Assoc. (U.K.) 28: 141-153.
- Hasselblad, V. 1966. Estimation of parameters for a mixture of normal distributions. Technometrics 8: 431-444.

- Hoydal, K., C.J. Rorvik and P. Sparre. 1982. Estimation of effective mesh sizes and their utilization in assessment. Dana 2: 69-95.
- Kendall, M.G. and A. Stuart. 1961. The advanced theory of statistics. Vol. 2. Charles Griffin & Com. Ltd.
- Macdonald, P.D.M. and T.J. Pitcher. 1979. Age group from size-frequency data: a versatile and efficient method of analyzing distribution mixtures. J. Fish. Res. Board Can. 36: 987-1001.
- McNew, R.W. and R.C. Summerfelt. 1978. Evaluation of maximum likelihood estimators for analysis of length-frequency distributions. Trans. Amer. Fish. Soc. 197: 730-736.
- NAG-Library. 1981. EO4FCF: a comprehensive FORTRAN algorithm for finding an unconstrained minimum of a sum of squares of M non-linear functions in N variables ($M \geq N$). No derivations required. NAGLIB: 1652/0: MK7:Jan 81.
- Pauly, D. and N. David. 1981. ELEFAN I, a BASIC program for the objective extraction of growth parameters from length-frequency data. Meeresforsch. 28: 205-211.
- Pauly, D. and G. Gaschütz. 1979. A simple method for fitting oscillating growth data, with a program for pocket calculators. ICES C.M. 1979/G:24 (mimeo.)
- Pearson, K. 1984. Contribution to the mathematical theory of evolution. Philos. Trans. R. Soc. London, Ser. A. 185: 71-110.
- Petersen, C.G.J. 1892. Fiskeribiologiske forhold i Holbaek Fjord, 1890-91. Beret. Danm. Biol. St. 1890(91) 1: 121-183.
- Rao, C.R. 1965. Linear statistical interference and its application. 2nd ed. John Wiley and Sons, New York.
- Schnute, J. and D. Fournier. 1980. A new approach to length frequency analysis: growth structure. Can. J. Fish. Aquat. Sci. 37: 1337-1351.
- Tanaka, S. 1953. Precision of age-composition of fish estimated by double sampling method using the length for stratification. Bull. Japan. Soc. Sci. Fish. 19: 657-670.
- Tomlinson, P.K. 1971. NORMSEP: normal distribution separation, p. 11.(1).2.1-11.(1).2.3. In N.J. Abramson (ed.) Computer programs for fish stock assessment. FAO Fish. Tech. Pap. 101 (FIRD/T101), FAO, Rome.

Appendix: List of Symbols

A (S, I)	: age of cohort I in sample S relative to the birth year of cohort S
AGE (S, I)	: age group of cohort I in sample
BM	: maximum number of cohorts in any sample
Br	: statistics for comparison on competing solutions
C	: amplitude in seasonal growth oscillation
CFREQ (S, L)	: theoretical frequency of length group L in sample no. S
CFREQ1 (S, I, L)	: theoretical frequency of length group L in sample no. S belonging to cohort no. I.
CHI2	: chi-squared criterion
CHI2P	: chi-squared criterion for constrained estimation
<u>D</u>	: Jacoby matrix for CFREQ with respect of THETA
DAY (S)	: day no. of sample S
DL	: length group interval size
DT (S)	: length of time period between samples
e	: stochastic term
E04FCF	: NAG-library minimization routine
FIRSTL	: lower limit of first length group
GCORR (L)	: adjustment factor for recruitment/gear selection of length group L
I	: index of cohort. I = 1, 2, . . . , STYPE (S, 2). Also used as index for parameter
K, Ko	: curvature parameters in growth equation
L	: length class index. L = 1, 2, . . . , NOL
L_∞	: asymptotic body length
L50%, L75%	: selection/recruitment parameters
MEANL (S, I)	: mean value of length distribution of cohort I in sample S
MO (J)	: day no. of the first day of each month
MONTH (S)	: month no. of sample S
NOBSP	: total number of non-zero observations
NOL	: number of length groups
NOS	: number of samples (or cruises)
NPAR	: number of parameters
N1	: sum of all observations $\geq R$
OBS (S, L)	: number of fish in length group L in sample S
OBSSUM (S)	: total number of fish in sample S
PARAM	: subset of the parameter vector THETA
PFACTOR (I)	: factor of proportionality between parameter no. I and its basic parameter
PI	: 3.1415926
PPOINT (I)	: pointer from PARAM to THETA

R	: lowest allowable number of observations in a length group allowing for the use of Br
RES (S, L)	: residual
S	: index of sample (or cruise), S = 1, 2, . . . , NOS
SEPIND	: separation index
SIGMA (S, I)	: standard deviation in length distribution of cohort I in sample S
SMALLOB	: sum of all OBS < R
SMALLC	: sum of all CFREQ for which OBS < R
STIME (S)	: decimal part of TIME(S)
STRUC (S, I, J)	: sample/cohort structure J = 1 : birth year of cohort I in sample S J = 2 : 0/1-cohort (spring/fall) J = 3 : THETA index for cohort strength J = 4 : index in foregoing sample (0 in the case of recruitment)
STYPE (S, 1)	: sample type (1, 2 or 3)
STYPE (S, 2)	: number of cohorts in sample no. S
SURV (S, I)	: number of survivors of cohort I sample S
Tbo, Tb1	: birthday
t _o	: von Bertalanffy growth parameter
THETA	: parameter vector
TIME(S)	: time of sample S (unit: year)
TOTSUR(S)	: total number of survivors at TIME(S)
t _w	: winter point
YEAR(S)	: year of sample no. S

**Two Methods for Simultaneously Estimating Growth, Mortality
and Cohort Size Parameters from Time Series of Catch-at-Length Data
from Research Vessel Surveys**

J.G. POPE

*Ministry of Agriculture, Fisheries and Food
Fisheries Laboratory
Lowestoft, Suffolk, NR33 0HT
United Kingdom*

Pope, J.G. 1987. Two methods for simultaneously estimating growth, mortality and cohort size parameters from time series of catch-at-length data from research vessel surveys, p. 103-111. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

Two methods for estimating growth, mortality and cohort size parameters from time series of groundfish survey catch-at-length data are described. Examples of their use are given based on North Sea cod (*Gadus morhua*) data from English groundfish surveys.

Introduction

Total catch-at-length data and catch-at-length per unit of effort data are generally much easier and cheaper to collect than the corresponding catch-at-age based data sets, but they are more difficult to interpret in terms of growth rates, mortality rates and cohort size. Perhaps as a result many of the methods of interpreting them have been kept as simple as possible. For example, it is common to estimate growth rates and only then (if at all) to estimate mortality rates and cohort size. Very frequently such estimates are based only on one year's result or on a composite distribution achieved from averaging together several years' data. Such approaches do have several drawbacks. These are notably that (i) mortality rate estimates will almost certainly be correlated with growth rate estimates and (ii) where a single length distribution is used, assumptions about the stationarity of recruitment and mortality have to be made which may not be justified.

This working paper investigates the possibility of developing some methods where growth rates, mortality rates and cohort sizes (relative sizes) are estimated from a time series of catch-at-length data.

A Simple Generalized Linear Model (GLIM) Model of the Problem

Dr. J. Shepherd (pers. comm.) has shown that catch-at-age data from groundfish surveys and total international catch-at-age data can be simply explained as an incomplete two-way analysis of variance in two multiplicative factors (year-class and age). A rather similar approach can be used for catch-at-length data as follows:

Estimates of K , L_∞ and t_o for a von Bertalanffy growth curve are adopted and each length group given a putative age. Length groups are then assigned to cohorts and an analysis of variance carried out on the log numbers at length. As a refinement the cosine and sine of 2π (age) might be included as covariates to account for recruitment having a seasonal distribution. The approach is simple and the factor estimates indicate the relative year-class strength and progressive total mortality on different ages of fish. Reruns of the ANOVA procedure with different K , L_∞ and t_o give a sum of squares surface for these factors which can be examined for a minimum. In practice discrimination of modes is best achieved for the smaller length groups.

Since

$$t - t_o = \frac{-1}{K} \ln(1 - L/L_\infty) \quad \dots 1)$$

it follows that in this range

$$t - t_o \approx \frac{1}{K} \cdot \frac{L}{L_\infty} \quad \dots 2)$$

Thus, we should expect the sum of squares to show a minimum for some particular value of K , L_∞ .

A simple example of this technique can be constructed using data from English groundfish surveys which are carried out in August-September each year in the North Sea. Table 1 shows cod catch rates from one of the subdivisions of the North Sea from 1980 to 1984. Examination of Table 1 suggests that catches of 5 and 10 cm fish are somewhat variable and a decision was made to remove all values less than 0.5 due to the extreme values they would receive under a logarithmic transformation. The remaining 43 points were transformed as follows: (i) the numbers at length to log number at length, (ii) the length to a putative age, (iii) the putative age and year to a putative year-class. For each K , L_∞ , t_o shown in Table 2, a two-way ANOVA of log numbers was performed, the factors being putative age and putative year-class. The sum of squares and degrees of freedom of these analyses are shown in Table 2. The results indicate a weak minimum in the sum of squares function for

$$\begin{array}{ll} L_\infty = 115; K = .20; t_o = .6, \\ \text{for } L_\infty = 130; K = .15; t_o = .3, \\ \text{and for } L_\infty = 145; K = .15; t_o = .6. \end{array}$$

Although at best the fit only explains 60% of the total variation of age and year-class, effects are generally significant as can be judged from the ANOVA at $K = 0.2$, $L_\infty = 130$, $t_o = 0$ shown on Table 3.

A second example was constructed for North Sea cod (*Gadus morhua*) using the catch rate in the same survey series for the entire North Sea. This data set is more extensive and is less variable than the previous one. However, because of the averaging of substocks' recruitment and growth rates it does not show the same clear year-class structure.

Table 4 shows the catch rate (actually the total catch) of each survey from 1977 to 1981. A similar ANOVA procedure was adopted as for the first example and results are shown in Table 5. (The results in the cells show the minimum sum of squares obtained for any t_o value and the degrees of freedom. Residuals of the fit for $K = 0.2$, $L_\infty = 130$, $t_o = 0$ are shown in Fig. 1. Table 5 shows that the sum of squares (and the degrees of freedom) both fell with reduced K and L_∞ and the lowest sum of squares result obtained is at a frankly unbelievable level of $K = 0.05$. These results are presumably due to the lack of clear year-class structure in the data and to the greater numbers of year-class and age factors fitted at lower K and L_∞ values. It is clear, therefore, that the results

Table 1. Catch rate of North Sea cod (ICES Roundfish Area 4) from English groundfish surveys, 1980-1984.

Length (cm)	Year				
	1980	1981	1982	1983	1984
0	0.00	130.11	.67	11.00	.17
10	.10	63.78	0.00	.83	0.00
15	6.50	11.33	1.44	.17	10.09
20	31.10	1.56	27.11	4.50	43.27
25	14.40	8.56	32.33	5.17	24.30
30	6.40	9.11	11.33	2.00	6.17
35	2.50	6.33	6.33	6.83	2.50
40	.70	4.22	2.56	15.00	1.50
45	1.40	.89	2.67	15.83	1.67
50	3.20	3.22	3.44	11.84	7.65

Table 2. Sum of squares of ANOVA fits to cod catch-at-length data for various values of K and L_{∞} (t_o adjusted for minimum sum of squares) (ICES Roundfish Area 4, 1980-1984).

K (year ⁻¹)	L _∞ (cm)					
	115	130	145			
	Sum of squares	d.f.	Sum of squares	d.f.	Sum of squares	d.f.
.10	40.4	29	33.7	30	36.2	32
.15	31.2	32	28.6	32	28.6	32
.20	28.6	32	33.7	33	32.9	35

Table 3. ANOVA for $K = 0.1 \text{ year}^{-1}$, $L_{\infty} = 130 \text{ cm}$, $t_o = 0$ for cod from Groundfish Area 4 (1980-1984).

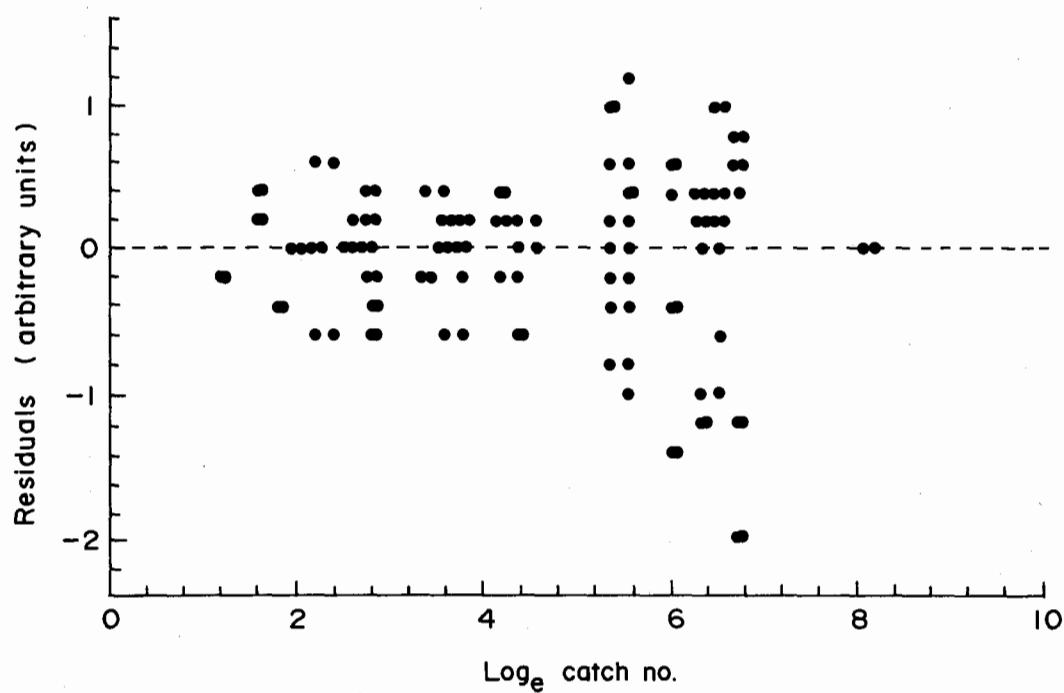
Cause	d.f.	Sum of squares	Mean squares	F	P
Age	3	13.53	4.510	4.41	< .05
Year-class	6	22.34	3.723	3.64	< .01
Residual	33	33.72	1.022		
Total	42	69.59			

obtained from the sum of squares are essentially meaningless in this case and that the method breaks down when there are no prominent year-classes to observe.

Table 6 shows the estimates obtained with $L_{\infty} = 130 \text{ cm}$, $K = 0.2 \text{ year}^{-1}$ and $t_o = 0.0 \text{ year}$ and indicates that year-class effects are never significant and that age effects become increasingly negative up to age 8. Note also that the differences in effect from age to age become reduced on the older ages suggesting that Z is for some reason being underestimated on these ages. This suggests that perhaps a more restrictive hypothesis about Z might be helpful and accordingly an analysis of

Table 4. Catch rate of North Sea cod from English groundfish surveys (1977-1981) (ICES Roundfish Area 4).

Length (cm)	Year				
	1977	1978	1979	1980	1981
5.	1,096.00	912.00	2,045.00	715.00	4,093.00
10.	630.00	406.00	125.00	608.00	3,626.00
15.	258.00	178.00	224.00	271.00	106.00
20.	1,569.00	700.00	777.00	2,110.00	242.00
25.	1,924.00	967.00	942.00	2,215.00	539.00
30.	1,411.00	720.00	701.00	1,140.00	764.00
35.	937.00	331.00	363.00	532.00	417.00
40.	559.00	285.00	208.00	280.00	321.00
45.	227.00	253.00	143.00	183.00	241.00
50.	111.00	192.00	96.00	127.00	121.00
55.	105.00	105.00	89.00	97.00	78.00
60.	88.00	68.00	47.00	76.00	49.00
65.	51.00	57.00	53.00	59.00	44.00
70.	47.00	52.00	41.00	38.00	22.00
75.	33.00	25.00	21.00	36.00	23.00
80.	21.00	23.00	19.00	23.00	14.00
85.	7.00	14.00	12.00	8.00	16.00
90.	9.00	12.00	21.00	11.00	17.00
95.	3.00	7.00	7.00	5.00	16.00
100.	7.00	6.00	4.00	10.00	17.00

Fig. 1. Plot of residuals of ANOVA (using $L_\infty = 130$ cm and $K = 0.2 \text{ year}^{-1}$) for data in Table 4.

covariance with age as the covariate and year-class as the factor was made. The estimate of the age effect is thus an estimator of the average Z on all years and ages.

Table 7 shows the Z's for a number of combinations of K and L_{∞} values. Although this analysis indicates clearly the correlation of Z on growth rate estimates the resulting sums of squares produced much the same results as the two-way ANOVA and $K = 0.05 \text{ year}^{-1}$ again gave the lowest sum of squares. This result is almost certainly due to the very small year-class effect detected in the ANOVA where the mean square of year-class effects was very similar to the residual variation. It is thus interesting that this second data set, although having only half the variability of the first, fails to produce the evidence for growth rate that is found in the first. It also suggests that a test for a significant year-class effect should be an essential part of the approach.

Table 5. Sum of squares of ANOVA fits to cod catch-at-length data for various values of K and L_{∞} (t_0 adjusted for minimum sum of squares) (all areas, 1977-1981).

K (year $^{-1}$)	L $_{\infty}$ (cm)		
	115	130	145
.05			10.8 46
.10	17.1 59	18.3 67	20.3 71
.15	21.3 69	25.2 75	24.4 79
.20	23.0 74	26.1 79	26.4 83

Table 6. GLIM estimates for North Sea cod (all areas, 1977-1981) (length = 5 cm; groups 5-100 cm; for K = 0.2 (year $^{-1}$) and $L_{\infty} = 130$ cm).

#	Estimate	Standard error	Parameter
1	6.867	.7717	% GM
2	-3.675	.7310	YC (2)
3	-8544E-01	.6924	YC (3)
4	.4630	.6828	YC (4)
5	.5200	.6804	YC (5)
6	.4428	.6995	YC (6)
7	.6229	.7043	YC (7)
8	.3396	.7123	YC (8)
9	.7198E-01	.7197	YC (9)
10	.7172E-02	.7270	YC (10)
11	.5302	.7389	YC (11)
12	-6758E-01	.7574	YC (12)
13	1.450	.9625	YC (13)
14	-5.896	.3121	Age (2)
15	-1.436	.3360	Age (3)
16	-2.653	.3480	Age (4)
17	-3.461	.3782	Age (5)
18	-4.207	.3973	Age (6)
19	-4.772	.4184	Age (7)
20	-5.187	.4801	Age (8)
21	-4.921	.5145	Age (9)

Scale parameter taken as .3309

Table 7. Z-values from analysis of covariance fits to cod catch-at-length data for various values of K and L_{∞} .

K (year $^{-1}$)	L_{∞} (cm)		
	115	130	145
.05		.21	
.10	.28	.41	.49
.15	.43	.59	.71
.20	.57	.75	.93

A Least Squares Approach to Estimating Growth Rates, Mortality Rates and Cohort Size

In the GLIM model of the previous section each length group was assigned a putative age. An alternative approach is to assume a particular cohort structure, mortality rates and growth rates and then to use this to predict the catch in each length group. Estimates of cohort structure, mortality rates and growth rates can then be progressively altered to minimize the sum of squares of the difference between actual and predicted number per length group or some suitable transformation of them.

A prototype method of this type was developed in order to identify the problems. This was designed to interpret groundfish survey type data and has the following assumptions:

- *Cohorts.* Annual cohorts of variable size are assumed, the size being specified for the cohort at its age in year 1 or at the youngest age group for cohorts which do not extend that far back. Each cohort has the same growth function (von Bertalanffy with the same K , L_{∞} and t_0). It was further assumed that the distribution of length at a particular age (a) is given by a Gaussian curve $N(L_a, \sigma_a)$ where σ_a might take various forms. For example, if all fish grew to the same L_{∞} then this would give $\sigma_{a+1} = \sigma_a e^{-K}$ and thus $\sigma_{\infty} = 0$. However, since some fish do grow larger than L_{∞} it seemed better to assume that $\sigma_a = \text{constant}$ so that only one value had to be estimated. Since it is not possible to consider an infinite number of cohorts each year a fixed number (g) was considered and the oldest cohort treated as a "plus" group.
- *Growth.* Average growth in the model follows a simple von Bertalanffy curve for all cohorts but since observations fall only once a year at the same time this is compatible with seasonally varying growth.
- *Mortality.* Mortality rates are assumed to be constant for all ages but can vary from year to year.
- *Catchability.* Catchability in the survey is assumed to be equal to 1.0 for all lengths greater than some specified length so that catchability effects have to be estimated for the smallest sizes only.
- *Variability.* Stochastic variation in the model and sampling error are considered jointly as error. Length distributions are split into b length groups specified by $b+1$ size delimiters $D_1, D_2 \dots D_{b+1}$. Within each such length group the total numbers are assumed to be distributed such that some simple transformation will cause the residuals to follow a normal distribution. For simplicity we will consider a square root transformation since this reduces the emphasis on the biggest numbers without overemphasizing the smallest numbers. The choice of $D_1 \dots D_{b+1}$ could of course be made partly with the intention of revealing cohort structure and partly to equalize variances of residuals. It would obviously make sense for the biggest length group D_b, D_{b+1} to be large and to span the size range of the oldest (plus group) cohort. The sum of the squared residuals, of course, forms the objective function to be minimized.

A computer program was developed to try out this approach to estimation. The procedure is as follows: estimates of population at age a in year y ($P(ajy)$) are made by first assuming $P(a, 1)$ and

$P(1, y)$ (where age and year 1 indicate the youngest age and first year, respectively). Assuming total mortality $Z(y)$ for each year then allows $P(a, y)$ to be estimated progressively from the recurrence relationships

$$P(a, y) = P(a - 1, y - 1) \exp [-Z(y - 1)]$$

for $a < g$,

and

$$P(g, y) = (P(g - 1, y - 1) + P(g, y - 1)) \exp [-Z(y - 1)].$$

Assuming values of K , L_∞ , t_0 and σ , we specify the distribution of length at each age $N(L_\infty, \sigma_a)$. We can now calculate the contribution of each cohort to population numbers at each size group i in each year as

$$\frac{1}{2} P(a, y) \left[\operatorname{erf} \frac{(D(i) - L_a)}{\sqrt{2} \sigma} - \operatorname{erf} \frac{(D(i+1) - L_a)}{\sqrt{2} \sigma} \right] .$$

where erf is the error function.

Summing these contributions from all ages gives the population at each length group which can be converted to estimated survey catch numbers $C_{\text{est}}(i, y)$ at length and year by multiplying by the catchability $q(i)$ of the survey for fish of the i th length group.

The transformation (square root) of these estimates can then be compared to the transformed observed numbers at length to give a residual. In order to find estimates that minimize the sum of squares of these residuals a standard minimization routine (National Algorithm Library (NAG) EO4FDF) has been used. This is a robust algorithm which requires a subroutine which only estimates the values of residuals for each length and year and is therefore convenient for getting a prototype running. It is, however, very restrictive in the number of parameters that can be considered due to the large overhead storage requirement of the routine. In practice, solutions were restricted to six cohorts per year for five years in which 10 length groups are specified and catchability is only estimated for the two smallest length groups.

With considerable correlation between the estimates of L_∞ and K and between the estimates of Z and L_∞ , K it seemed sensible to optimize other estimates for specified levels of K and L_∞ and thus to build up a sum of squares surface for these two parameters from successive runs with different levels.

Even with this very restrictive model (18 parameters from 50 data points) the algorithm was slow and expensive in storage. Each run for a fixed K , L_∞ took about 30-60 minutes and a day was needed to produce a reasonably comprehensive sum of squares surface on K and L_∞ . The method will have to be speeded up and made portable before it can be regarded as a very helpful technique but the preliminary runs do indicate the feasibility of this approach to estimating the various parameters needed to explain groundfish survey catch-at-length data. (One should also perhaps remember that 14 years ago it took the author a week to produce a virtual population analysis of North Sea plaice, so speed is not everything, even though it is obviously highly desirable.) The method has been tried out on exact test data for which it successfully recovered the initial input values and for the North Sea cod groundfish survey data shown on Table 1.

Results from this method applied to the data of Table 1 are as follows. Table 8 shows the sum of squares estimated for each combination of K and L_∞ shown. This indicates the trend of the minimum from low K high L_∞ to higher K lower L_∞ with little to discriminate between the various possibilities. Table 9 shows the estimated catches and the residuals for $K = 0.2$, $L_\infty = 130$.

Table 10 shows the estimates of the various parameters at this position. The results for mortality seem particularly wild and a more restrictive hypothesis should be preferable on this data set, for example, a constant mortality for all years and ages. The fact that the algorithm has usually converged on this very rough data set is, however, encouraging.

Apart from obvious improvements to the algorithm it would seem possible to extend the idea to a least squares separable analysis for total catch-at-length data. In practice this would not seem very likely to succeed in estimating all parameters unless catch-at-length data could be produced for finer subdivisions of time than a year in order to emphasize the growth of year-classes. In the absence of such data the method could still be used to estimate fishing mortality by year, partial recruitment at age and cohort sizes, given estimates of K, L_∞ and M.

Table 8. Sum of squares of least squares fits to cod catch-at-length data for various values of K and L_∞ .^a

K (year ⁻¹)	L_∞ (cm)			
	100	115	130	145
	Sum of squares	Sum of squares	Sum of squares	Sum of squares
.15	68.0	59.4	44.8	36.4
.20	48.3	35.4	28.4	29.2
.20	32.7*	27.2**	32.9***	104.0***

^aRuns marked * failed to converge completely according to the criteria of the algorithm. Extra *'s indicate progressively less confidence in the value of the sum of squares function.

Table 9. Catch estimates and residuals using restricted least square model (see text).

Length (cm)	Catch estimates in year				
	1980	1981	1982	1983	1984
0.00	1.26	132.12	.44	8.07	.08
10.00	.67	57.93	.36	3.56	.21
15.00	5.53	13.39	7.67	1.86	8.06
20.00	19.13	5.52	26.96	4.02	28.45
25.00	21.29	5.95	30.00	4.49	31.66
30.00	7.52	2.32	10.54	2.29	11.16
35.00	1.27	2.33	1.18	6.64	1.58
40.00	1.30	5.83	.12	17.90	1.06
45.00	1.12	5.10	.07	15.63	1.01
50.00	3.37	3.21	.23	4.92	10.45

Length (cm)	Residuals in year				
	1980	1981	1982	1983	1984
0.00	-1.12	-.09	.16	.48	.12
10.00	-.50	.38	-.60	-.98	-.46
15.00	.20	-.29	-1.57	-.95	.34
20.00	1.20	-1.10	.01	.12	1.24
25.00	-.82	.49	.21	.16	-.70
30.00	-.21	1.49	.12	-.10	-.86
35.00	.45	.99	1.43	.04	.32
40.00	-.30	-.36	1.26	-.36	.19
45.00	.12	-1.31	1.37	.02	.29
50.00	-.05	.00	1.37	1.22	-.47

Table 10. Estimates of parameters corresponding to $L_{\infty} = 130$ cm, $K = 0.2$ (year $^{-1}$), Sigma = 4.42 and $t_0 = 0.91$ (year), at best fit point of least squares analysis applied to North Sea cod (data from ICES Roundfish Area 4, 1980-1984).

Z (I) =	1.307	4.338	.528	1.488	1.000
S (K) =	.03-25	1.00-1.00	1.00-1.00	1.00-1.00	1.00-1.00
R (I) =	56.32	5,894.75	19.43	359.86	3.70
P (J) =	56.32	54.63	3.22	.93	2.03

Discussion

Both methods presented here appear capable of estimating growth, mortality and relative cohort strength from time series of groundfish survey length data. Neither, however, is likely to be able to estimate growth parameters with any great confidence due to the correlated nature of these parameters unless distinguishable cohorts can be detected in the data up to sizes approaching L_{∞} . In the absence of such information in the data the methods will only be able to estimate other parameters conditional upon growth rate estimates which might be obtained externally (e.g., from aging a few fish). This would, however, itself be useful as the methods make fewer assumptions about the nature of cohort size and mortality rate than do some simpler methods. Development and tests of these two methods would, thus, seem well worthwhile.

A Weakly Parametric Method for Estimating Growth Parameters from Length Composition Data

J.G. SHEPHERD

*Ministry of Agriculture, Fisheries and Food
Directorate of Fisheries Research
Fisheries Laboratory
Lowestoft, Suffolk NR33 0HT
United Kingdom*

Shepherd, J.G. 1987. A weakly parametric method for estimating growth parameters from length composition data, p. 113-119. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

A new robust method for the analysis of length compositions in terms of the parameters of a specified growth curve (e.g., von Bertalanffy) has been developed. It does not require the number or width of modes to be specified and is insensitive to sampling noise of the larger length groups. A goodness-of-fit measure is mapped as a function of growth parameters, which points out problems caused by multiple maxima and enables the size and shape of approximate confidence regions to be estimated. Once growth parameters have been selected, these may be used to 'slice' the length composition into age groups or as the basis for more sophisticated methods if required. However, the decomposition into age groups is highly dependent on the choice of L_∞ within the confidence region and it is probable that further analysis of length data directly is preferable to methods relying on such a decomposition.

Introduction

A variety of methods has been proposed for the analysis of length compositions into age groups and particularly for the determination of the parameters of specified growth curves such as that of von Bertalanffy. The state of the art is probably best represented by the statistical poly-modal analysis developed by MacDonald (1980) and the method known as ELEFAN I (Pauly and David 1981). Neither of these methods is entirely satisfactory, however. The former normally requires some estimate to be provided of the number of modes and their widths (which is often difficult to do and subjective) and does not constrain the mode locations to conform to any growth function (which may be an advantage or not, depending on the application), nor assigns ages to the modes directly. ELEFAN I, on the other hand, is based on the goodness-of-fit of the data to a parametric representation of a growth curve (usually von Bertalanffy), but the basis of the goodness-of-fit criterion and the preliminary data treatment (compression and filtering) is arguably not

ideal and the method sometimes converges to strange results. Also, the method gives little indication of the confidence limits of the results.

This paper describes a method which is conceptually very similar to ELEFAN I in that it is based on the goodness-of-fit of the modes (if any) in one or more length compositions to the locations expected from a specified growth curve. This places a strong constraint on the possible locations of modes and their relationship to one another, eliminating many infeasible interpretations of the data and yielding directly values of growth parameters, and an interpretation in terms of ages, without further analysis.

The goodness-of-fit criterion used is akin to a correlation coefficient between the data and a test function constructed from a growth curve. The idea is similar to that used in the time series analysis technique known as complex demodulation (see, e.g., Bloomfield 1976). The criterion is designed to take account of data in proportion to its quantity and likely reliability without pre-processing, and not to be unduly sensitive to possibly spurious modes created by sampling noise.

These modifications are intended to result in a more robust method and this is further assisted by not automatically seeking the parameters yielding the maximum goodness-of-fit (since there may be several maxima), but simply mapping its value over a feasible range of parameter values and choosing the most appropriate values in the light of this information.

Method

We attempt to interpret one or more length compositions (arising from data collected in different years or seasons, perhaps) in terms of a single growth curve. For simplicity this is assumed to be that of von Bertalanffy, without seasonal modulation of growth rate, though more complicated formulations could be used if required.

Given current estimates of the parameters L_∞ , K and t_o of the growth curve, the lengths where modes are to be expected in a season which is a fraction t_s through the annual (or other) cycle are given by

$$L = L_\infty [1 - \exp \{ -K(t + t_s - t_o)\}] \quad \dots .1)$$

for $t = 0, 1, 2, \dots$ etc.

Thus observations occurring at or near these lengths may be considered as evidence in favor of the current parameter choices. Conversely, observations occurring at the intermodal lengths (given by $t = 0.5, 1.5, 2.5$ etc.) constitute evidence against them.

For any season one may construct a test function which is positive near the expected modal lengths and negative near the intermodal lengths. The sum of products of the observations with this test function then constitutes a criterion of goodness-of-fit: if it is large and positive the data are consistent with the current parameter choices, whilst if it is large and negative they are not consistent. This sum of products may be computed for each length composition available (using a test function appropriate to the season) and an overall goodness-of-fit measure (or 'score') may be constructed by summing these totals, thereby using all available data efficiently.

Two refinements of this idea are used in practice. First, because one may expect the numbers at length in any length interval to be roughly Poisson distributed, the square root of the numbers observed is used in the summation. This provides an appropriately modest degree of dynamic range compression so that numerous observations still carry more weight than rare ones, but not overwhelmingly so.

Second, since modal lengths do not generally fall in the center of length intervals, and for length groups approaching L_∞ there may be several modes in a length interval, it is necessary to average the test function over the length intervals in use. A suitable basic test function is simply

$\cos 2\pi(t - t_s)$ where t is just estimated age-at-length (measured from 1 January or whatever other origin is routinely used), i.e.,

$$t = \frac{1}{K} \ln \left[\frac{L_\infty - L_o}{L_\infty - L} \right] \quad \dots 2)$$

$$= t_o - \frac{1}{K} \ln (1 - L/L_\infty) \quad \dots 2a)$$

$$\text{where } L_o = L_\infty \{ 1 - \exp(Kt_o) \} \quad \dots 3)$$

For a length interval whose upper and lower bounds correspond to ages-at-length of t_{\max} and t_{\min} , it is easily shown that the average of this test function over the interval is just

$$T(\ell) = \frac{\sin \pi (t_{\max} - t_{\min})}{\pi (t_{\max} - t_{\min})} \cos 2\pi (\bar{t} - t_s) \quad \dots 4)$$

where \bar{t} is the average of t_{\max} and t_{\min} . The first term of this expression is the well-known Fraunhofer diffraction function, which in this context allows for the number and placement of modes in the interval. It becomes small if there is more than one mode in the interval, thus correctly allowing for the obvious fact that observations in such an interval yield little evidence for modal positions and should be heavily discounted.

The final form of the goodness-of-fit (or 'score') function is then

$$S = \sum_{\ell} \sum_i T(\ell, i) N^{1/2}(\ell, i) \quad \dots 5)$$

where ℓ indexes the length groups and i indexes the various distributions available.

Noting that, because of equation (2a), t_{\max} and t_{\min} and, therefore, also \bar{t} and $T(\ell)$ are periodic in t_o , it is clear that S is similarly periodic (with a period of unity). It is therefore possible to determine very easily the value of t_o leading to a maximum of S for any choice of values for K and L_∞ . If we denote by A the value of S obtained with t_o set to zero, and B that obtained with t_o set to 0.25, it can be determined that

$$S_{\max} = (A^2 + B^2)^{1/2} \quad \dots 6)$$

and

$$t_o = \frac{1}{2\pi} \arctan(B/A) \quad \dots 7)$$

Thus, the full three dimensional mapping of S as a function of K , L_∞ and t_o may be reduced to a two-dimensional mapping as a function of K and L_∞ only, conditional on the choice of t_o given by (7) above. Since such a two dimensional mapping is easily displayed and interpreted (e.g., by contouring) this constitutes a further technical advantage of this method.

A concise algorithmic description of the full procedure follows:

1. Read data.
2. Select range of growth parameters (K , L_∞) to be used.
3. For each value of L_∞ . . .
4. For each value of K . . .
5. For $t_o = 0$ and 0.25 . . .
6. Initialize score (to zero).
7. For each length group . . .

8. Calculate ages corresponding to length group boundaries (t_{\min} and t_{\max}) and $\Delta t = \frac{t_{\max} - t_{\min}}{n}$.
9. Calculate diffraction function $D = \frac{\sin \pi \Delta t}{\pi \Delta t}$.
10. For each length distribution ...
11. Calculate $C = \cos 2\pi (\bar{t} - t_s)$.
12. Calculate $\Delta S = D * C * \text{SQRT}(N(L))$ and add to score.
13. Next length distribution.
14. Next length group.
15. Next value of t_o .
16. Set $A = \text{Score for } t_o = 0$.
Set $B = \text{Score for } t_o = 0.25$.
17. Calculate maximum score for current values of L_∞ and K , $S_{\max} = \text{SQRT}(A^2 + B^2)$.
18. Calculate t_o to give max score as $t_o = \arctan(B/A)/2\pi$.
19. Next K .
20. Next L_∞ .
21. Tabulate results.

A listing of the source code (in the Microsoft dialect of Fortran 77) is given in Morgan and Pauly (Part II, this vol.). This should be suitable for running on many microcomputers, and may easily be translated into BASIC or any other suitable language if required.

Results

Preliminary tests of the method have been carried out. A test data set for *Raja clavata* in the Irish Sea (courtesy of K. Brander) is given in Table 1. These data were obtained during routine groundfish surveys carried out three times per year and the numbers given are the total numbers measured in each season over several years. The data could equally well (perhaps better) have been analyzed as the individual distributions obtained from each survey.

The results from the analysis are given in Table 2. It is apparent that the score function is not entirely smooth, having several local maxima within a broad band of high values and also a maximum on the boundary of the region of parameter space selected for investigation. Both these features could have caused problems if automatic maximization had been attempted.

It is also clear that (as expected) there is a wide range of almost equally good pairs of correlated values of K and L_∞ , lying between (0.14, 150) and (0.26, 100) and presumably beyond. The parameter values (0.21, 120) selected by Brander and Palmer (1985) on the basis of subjective modal analysis are very close to the center of this region of good parameter values. The optimum value for the (fractional part of) t_o is also strongly dependent on the choice of K and L_∞ .

The statistical properties of the score function are not known so it is not possible to determine proper confidence regions for the parameter values. However, it seems plausible that contours of constant score would coincide with the shape of the confidence region and these may easily be sketched or computed. As a rough approximation, one may guess that the score function relative to its maximum is probably analogous to a variance ratio and that the contour corresponding to half the maximum is associated with something like a 95% confidence level (since they are likely to be observations in several tens of length groups contributing to the total).

A more detailed analysis of the basis of the solution corresponding to $(K, L_\infty) = (0.20, 120)$ is given in Table 3. It should be noted that because of the periodicity of the score function in t_o , the values for this and the actual ages assigned to particular length groups, are indeterminate to the addition or subtraction of *any* whole number of years. Assignment of *true* ages is unlikely to be possible unless small mesh data (e.g., from research vessel surveys) are available, which permit the unambiguous identification of 0 or 1-group fish. It is also important to observe that almost all the positive contributions to the score function have been accumulated by length groups of less than 50 cm. Thus, almost all the useful information about modes and ages is carried by the smaller length

Table 1. Length-frequency data on *Raja clavata*, from the Irish Sea, used for testing method proposed here.

Lower limit of length class (cm)	Numbers at length Season (years)			Lower limit of length class (cm)	Numbers at length Season (years)		
	0.2	0.45	0.8		0.2	0.45	0.8
9	10	0	2	60	29	31	24
12	180	4	62	63	25	31	37
15	35	15	2	66	17	36	29
18	4	2	0	69	32	30	40
21	2	0	11	72	29	26	35
24	23	2	56	75	24	45	35
27	84	3	176	78	29	37	33
30	124	18	108	81	45	27	34
33	63	34	43	84	30	25	33
36	39	21	29	87	17	19	27
39	39	16	58	90	4	15	19
42	52	14	65	93	9	15	10
45	47	31	56	96	1	5	6
48	46	18	44	99	0	1	0
51	28	27	29	102	0	1	1
54	26	38	35	105	0	1	0
57	32	38	28				

Table 2. Results of length composition analysis for *Raja clavata* (Irish Sea).

L _∞ (cm)	Value of K (1/year)										
	.10	.12	.14	.16	.18	.20	.22	.24	.26	.28	.30
100.00	7.1	9.3	8.9	5.2	16.5	24.0	27.5	31.4	32.6	30.5	29.3
110.00	10.4	7.4	4.6	16.2	23.9	30.7	32.2	31.7	31.2	22.7	21.4
120.00	6.6	4.2	14.8	23.3	29.9	34.5	30.5	26.7	21.7	23.4	19.8
130.00	9.4	9.1	21.9	27.5	32.9	31.4	24.7	25.8	19.7	14.0	11.0
140.00	4.1	15.4	27.4	35.3	27.6	27.7	24.7	17.6	13.9	10.4	8.1
150.00	9.2	22.4	34.4	29.2	29.1	24.1	18.3	13.6	9.8	8.1	8.5

B. Table of t_o values

L _∞ (cm)	Value of K (1/year)										
	.10	.12	.14	.16	.18	.20	.22	.24	.26	.28	.30
100.00	.5	.9	.1	.2	.2	.3	.4	.5	.6	.7	.8
110.00	.8	.0	.1	.2	.3	.5	.6	.7	.8	.8	.8
120.00	.9	.2	.2	.3	.5	.6	.7	.8	.8	.9	1.0
130.00	.1	.1	.2	.4	.6	.7	.7	.8	.9	1.0	1.0
140.00	.1	.1	.3	.5	.7	.7	.8	.9	.9	1.0	1.0
150.00	.1	.3	.4	.6	.7	.8	.9	.9	1.0	1.0	.9

groups, as would be expected, because for older fish the modes become broader and closer together. For this reason commercial length compositions (which are usually truncated by minimum mesh and landing sizes) are much less useful for modal analysis than those from research vessel surveys. Indeed, for this same stock the commercial length compositions contain very few fish less than 50 cm in length, and are virtually useless for this type of analysis (i.e., the determination of growth parameters).

Finally, the numbers at age corresponding to the same choice of parameters are also given in Table 4. This has been obtained by simply 'slicing' the length composition at the intermodal values, with appropriate distribution between older and younger ages where necessary, and truncation to give conventional integer age groups referred to 1 January. Experimentation with the results obtained with almost equally good choices of parameter values shows (as would be expected) that the numbers in the youngest age groups are quite well determined, but those in the older age groups are strongly dependent on the choice of L_∞ . The utility of decomposition into age groups is therefore in doubt, especially for any purpose which would subsequently involve any resummation over age groups. Since this includes many common assessment procedures (e.g., catch forecast and yield-

Table 3. Length composition analysis for *Raja clavata* from the Irish Sea. (Evaluation of solution for $K = 0.2$, $L_\infty = 120.0$, $t_o = .57$).

Length (cm) ^a	Age (years)	Cumulative score
9.0	.96	1.63
12.0	1.10	8.83
15.0	1.24	14.25
18.0	1.38	15.53
21.0	1.53	15.53
24.0	1.69	17.78
27.0	1.84	23.74
30.0	2.01	26.88
33.0	2.18	29.93
36.0	2.35	31.33
39.0	2.54	31.07
42.0	2.72	31.25
45.0	2.92	30.60
48.0	3.12	31.95
51.0	3.34	33.83
54.0	3.56	33.92
57.0	3.79	30.95
60.0	4.04	31.65
63.0	4.29	33.40
66.0	4.56	33.52
69.0	4.85	32.23
72.0	5.15	33.49
75.0	5.47	34.27
78.0	5.82	32.46
81.0	6.19	34.26
84.0	6.59	33.07
87.0	7.02	33.35
90.0	7.50	33.88
93.0	8.03	34.78
96.0	8.62	34.45
99.0	9.28	34.50
102.0	10.06	34.56
105.0	10.97	34.47

^aLower limit of length class.

Table 4. "Age composition" of data in Table 1, as obtained through "slicing", using $L_\infty = 120$ cm and $K = 0.2$ (see also text).

Age	Numbers at age		
0	.0	.0	64.9
1	233.1	24.9	382.1
2	365.0	108.6	257.4
3	193.7	134.1	124.7
4	104.4	121.4	121.0
5	85.5	104.0	96.3
6	89.9	71.6	73.2
7	40.9	36.1	35.0
8	11.6	19.7	11.4
9	.9	3.4	.3
10	.0	1.1	.7
11	.0	.9	.0
12	.0	.1	.0
13	.0	.0	.0
14	.0	.0	.0
15	.0	.0	.0
16	.0	.0	.0
17	.0	.0	.0
18	.0	.0	.0
19	.0	.0	.0
20	.0	.0	.0

per-recruit calculations), it seems likely that it would be preferable to seek equivalent computational procedures which utilize length data directly, rather than relying on an uncertain division into age groups and subsequent recombination. For this reason the obvious possibility of improving the separation into age groups using a formal deconvolution procedure has not been pursued.

Other tests of the method have also been carried out for stocks which can be aged and the results have been found to be in excellent agreement with those obtained by conventional methods. Further and more exhaustive tests are required, particularly in order to determine whether this method is indeed more robust and informative than alternative methods when used on difficult data sets as it is intended to be.

Discussion

The method which has been developed for the analysis of length compositions avoids the need to specify parameters which may be in doubt (e.g., especially the number and width of assumed modes), and provides a direct indication of suitable values of K , L_∞ and t_0 , together with an indication of their likely precision. Preliminary tests indicate that it works well on suitable data and the results are in good agreement with those from more conventional subjective methods and those based on age determination where this has been tested. The ability to determine growth parameters depends strongly on the suitability of the data set, which must (of course) exhibit some modal structure. Data from research work using small mesh nets (especially from series of surveys using research vessels) are therefore usually much more suitable than that from commercial landings. It should be noted however that growth parameters determined from one sort of data (e.g., from RV surveys) *cannot* be used immediately to interpret data from other sources (e.g., commercial data), because of the differences of sampling bias (selection).

It is possible to use growth parameters to 'slice' length compositions into age compositions. The results are however strongly dependent on the choice of L_∞ within the likely confidence region, especially for older age groups. It is not therefore clear that it is useful to do this in order to use conventional age-based calculations and it seems likely that it would be preferable to develop alternative procedures based on the use of length compositions directly.

References

- Bloomfield, P. 1976. Fourier analysis of time series: an introduction. J. Wiley, New York.
- Brander, K. and D. Palmer. 1985. Growth rate of *Raja clavata* in the northeast Irish Sea. J. Cons., Cons. Int. Explor. Mer 42(2): 125-128.
- MacDonald, P.D.M. 1980. A Fortran program for analyzing distribution mixtures. McMaster University, Dept. of Mathematical Sciences (Stats), Tech. Rep. No. 80-ST-1.
- ✓ Pauly, D. and N. David. 1981. ELEFAN I: A BASIC program for the objective extraction of growth parameters from length-frequency data. Meeresforsch. 28(4): 205-211.

Estimation of Growth and Mortality Parameters for Use in Length-Structured Stock Production Models

J.M. HOENIG*

*Minnesota Department of Natural Resources
500 Cedar Avenue, St. Paul
Minnesota 55146, USA*

Hoenig, J.M. 1987. Estimation of growth and mortality parameters for use in length-structured stock production models, p. 121-128. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

The suggestion of some authors that linear methods of fitting von Bertalanffy growth curves generate biases may be premature, at least in one case discussed here, where the direction of the observed bias is explainable by the use of a predictive instead of a functional regression. When growth parameters are needed for length-converted catch curves or length-based cohort analysis, an inverse regression (age on length) may be appropriate. This can be computed through ordinary linear regression. An estimator for the total mortality, based on mean length, is derived for use when recruitment occurs periodically rather than continuously. Surplus-production models employing total mortality as the independent variable may become tilted to the left if natural mortality is compensatory. When growth parameters are unknown, surplus production can be modelled as a function of Z/K. When available data are limited to a small portion of the surplus production curve, fitting curves by regression may give unreasonable results. In this case, one can constrain a parameter to the value of an independent estimate so that the other estimates become more reasonable.

Introduction

Surplus production modelling as a function of total instantaneous mortality Z, or as a function of Z/K, was proposed by Csirke and Caddy (1983) as an alternative to collecting and calibrating data on fishing effort. Total mortality data can be obtained in a variety of ways, the simplest of which utilize data on mean length in a sample.

*Present address: RSMAS/CIMAS, University of Miami, 4600 Rickenbacker Causeway, Miami, Florida 33149, USA.

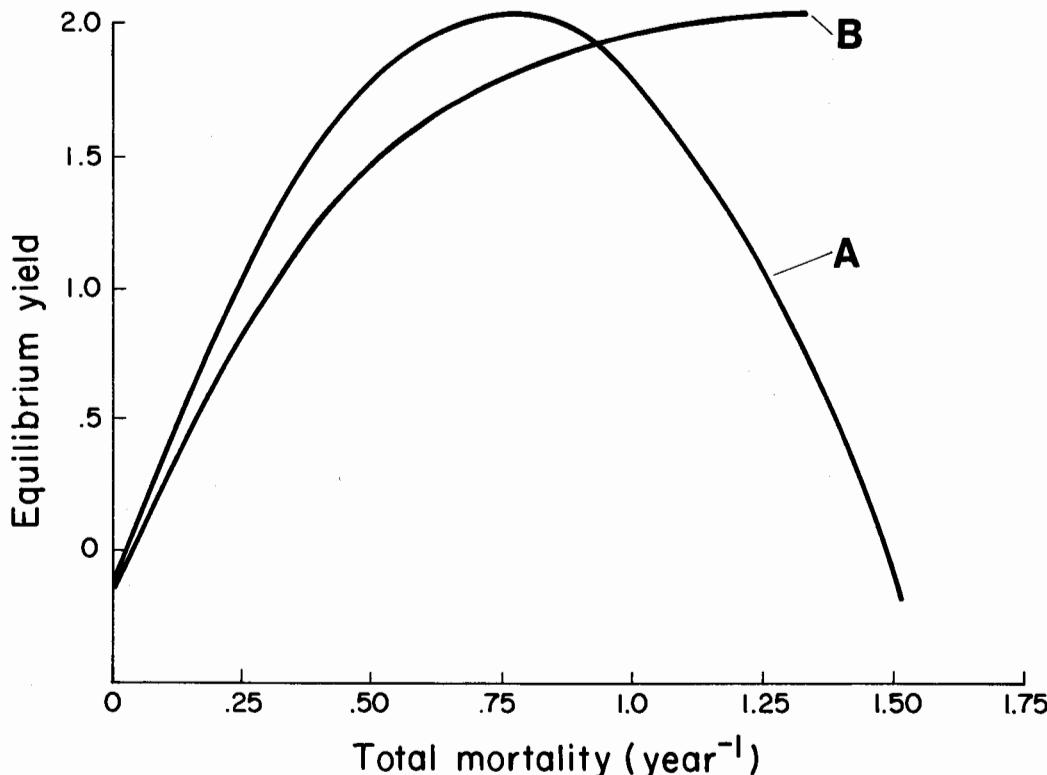


Fig. 1. Production curves generated using two models for estimating mortality. Curve A represents the model $Y_{EQ} = 4Z^2 + 6Z - .5$. Assuming that $L_\infty = 40$, $K = 0.3$ and $L' = 10$, Z values were converted to mean lengths using equation 4. Curve B was generated by converting the mean lengths back to Z values using equation (2) and represents effect of assuming recruitment to be continuous and using estimator (2) when the estimator for discrete recruitment (4) would be appropriate.

when estimating Z_{opt}/K : a 10% increase in the value of L_∞ caused an 18% increase in $(Z/K)_{opt}$ (Fig. 2).

Optimum length can be obtained from Z_{opt}/K using the relationship (Pauly 1984):

$$\frac{Z_{opt}}{K} = \frac{L_\infty - \bar{L}_{opt}}{\bar{L}_{opt} - L'}$$

Thus,

$$\bar{L}_{opt} = \frac{L_\infty + \left(\frac{Z_{opt}}{K} \cdot L' \right)}{\frac{Z_{opt}}{K} + 1}$$

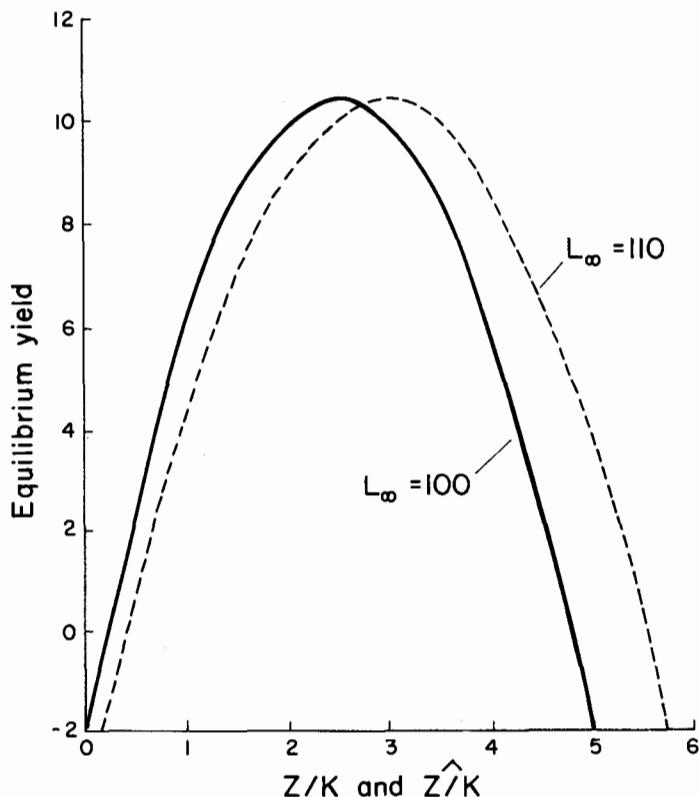


Fig. 2. Sensitivity analysis for the value of L_{∞} in the relationship $Z/K = (L_{\infty} - \bar{L})/(\bar{L} - L')$. One curve uses a value of L_{∞} which is 10% greater than the other. MSY is little affected while $(Z/K)_{opt}$ increases by about 18%.

Replacing Z_{opt} by any value of Z allows one to establish the relationship between Z (or Z/K) and the mean length of animals above L' in the catch.

When stock production data are limited or overly noisy, it may be necessary to incorporate an independently derived estimate of one parameter in the estimation procedure in order to obtain reasonable estimates of the other parameters. For example, the production curve can be constrained to have height equal to an estimate of MSY obtained independently from a comparative study. Or, the production curve can be forced to have a left X-axis intercept equal to a given value of M or M/K .

Let the production parabola be described by

$$Y = b_1 X^2 + b_2 X + b_3$$

where X denotes either Z or Z/K . The parabola has a maximum height of

$$-\frac{b_2^2}{4b_1} + b_3$$

which will be constrained to a value of d using the method of Lagrange multipliers.

The task is to minimize

$$\Sigma (Y - b_1 X^2 - b_2 X - b_3)^2 + \lambda (-b_2^2 + 4b_1 b_3 - 4b_1 d)$$

where λ is a Lagrange multiplier, by setting the partial derivatives with respect to each parameter equal to zero. The resulting system of four equations can be solved simultaneously using Newton's method (Hoenig and Hoenig 1986).

Discussion

Length-based surplus production modelling is an important new tool for fisheries biologists. However, careful attention to technical details is required to avoid insidious systematic errors, which will remain undetected if a simple goodness-of-fit criterion is used to assess the validity of a model.

If an inappropriate mortality model or a poor estimate of L_∞ is used, the stock production model may be distorted in shape or shifted in location. Even so, the MSY can be estimated. If the methodology used does not change, then the fishery can be managed on the basis of the relationship between equilibrium yield and the X variable since it will remain a simple function of the true mortality rate. Difficulties arise, however, when estimates derived from such a production curve, such as Z_{opt} and M or Z_{opt}/K and M/K , are taken out of this context and used for other purposes.

Acknowledgements

Partial support for this study was provided by the Minnesota Department of Natural Resources. Additional support was provided through the Cooperative Institute of Marine and Atmospheric Studies by National Oceanic and Atmospheric Administration Cooperative Agreement No. NA 85-WCH-06134. The U.S. Government is authorized to produce and distribute reprints for governmental purposes notwithstanding any copyright notation that may appear hereon.

References

- Beverton, R.J.H. and S.J. Holt. 1956. A review of methods for estimating mortality rates in fish populations, with special reference to sources of bias in catch sampling. *Rapp. P.v. Réun. Cons. Int. Explor. Mer* 140: 67-83.
- Csirke, J. and J.F. Caddy. 1983. Production modelling using mortality estimates. *Can. J. Fish. Aquat. Sci.* 40: 43-51 (+ Errata sheet in 40: 255-256).
- Hoenig, J. 1979. The vertebral centra of sharks and their use in age determination. Univ. Rhode Island, Kingston, 142 p. MS Thesis.
- Hoenig, J.M. and N.A. Hoenig. 1986. Fitting stock production models subject to a constraint. Univ. Rhode Island Dept.-Computer Sci. and Statistics Tech. Rep. TR86-186, 10 p. Kingston.
- Hoenig, J.M., W.D. Lawing and N.A. Hoenig. 1983. Using mean age, mean length and median length data to estimate the total mortality rate. *ICES CM* 1983/D: 23. 11 p.
- Pauly, D. 1983. Length-converted catch curves: a powerful tool for fisheries research in the tropics (Part 1). *Fishbyte* 1(2): 9-13.
- Pauly, D. 1984. Fish population dynamics in tropical waters: a manual for use with programmable calculators. ICLARM Studies and Reviews 8. International Center for Living Aquatic Resources Management, Manila, Philippines. 325 p.
- Pauly, D. and G. Gaschütz. 1979. Simple method for fitting oscillating length-growth data, with a program for pocket calculators. *ICES CM* 1979/G: 24. 26 p.
- Vaughan, D.S. and P. Kanciruk. 1982. An empirical comparison of estimation procedures for the von Bertalanffy growth equation. *J. Cons. Int. Explor. Mer* 40: 211-219.

Estimating Mortalities from Length- or Age-Specific Sex Ratios

M.J. SANDERS

*Marine Resources Service
Fisheries Resources and Environment Division
Fisheries Department
Food and Agriculture Organization of the United Nations
Via delle Terme di Caracalla
00100 Rome, Italy*

Sanders, M.J. 1987. Estimating mortalities from length- or age-specific sex ratios, p. 129-136. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

Methods are presented for estimating F and M from the proportions by length or age of each sex in a fish catch, and information on the total mortality coefficient and of the age at first capture for each sex. When dealing with seasonal fisheries, the additional data required are the ages at the end of the first fishing season and the duration of the fishing season. The methods are supported by programs for HP 67/97 calculators.

Introduction

Sanders (1977 and unpublished data) presented methods for estimating the mortality coefficients from data for which the sexes differ in mortalities and/or age at first capture. The data requirements include the overall proportions of each sex in the catch. A development of these methods is presented which makes fuller use of the data when the proportions of each sex are known for each length or age class.

The additional data requirements are the total mortality coefficient and age at first capture for each sex. When dealing with seasonal fisheries, the ages at the end of the first fishing season and the fishing season duration are also required. Programs based on these methods and suitable for use with HP 67/97 calculators are listed (see Morgan and Pauly, Part II, this vol.), along with some worked examples based on data simulated from known parameter values.

Theory

The methods as programmed are associated with the basic assumption that the sexes share the same values of the natural mortality coefficient and/or the fishing mortality coefficient, that is,

$F_m = F_f$ or $M_m = M_f$. Other assumptions are that the mortality coefficients remain constant for each time period under consideration and the stock numbers at zero age for each sex and cohort are the same.

Nonseasonal Fishery Method

The multiple-cohort catch number, $C_{t_d \rightarrow t_m}$, of individuals whose ages are equal and greater than t_d , can be described by the equation

$$C_{t_d \rightarrow t_m} = F/Z \cdot N_o \exp(-Mt_c - Z(t_d - t_c)) \quad \dots 1)$$

where F , M and Z are the fishing, natural and total mortality coefficients, N_o is the number of pre-recruits at zero age, t_c is the age at first capture and t_m is the maximum age (Fig. 1).

Subscripting separately for males (m) and females (f), the catch number ratio is given by

$$\frac{\hat{C}_{t_d \rightarrow t_m}^f}{\hat{C}_{t_d \rightarrow t_m}^m} = \frac{(Z_f - M_f)/Z_f \cdot N_o^f \exp(-M_f t_c^f - Z_f(t_d^f - t_c^f))}{(Z_m - M_m)/Z_m \cdot N_o^m \exp(-M_m t_c^m - Z_m(t_d^m - t_c^m))} \quad \dots 2)$$

for which it is considered reasonable to assume that $N_o^m = N_o^f$. This equation can then be used for the estimation of catch number ratios for each observed value of both t_d^m and t_d^f from predetermined values of Z_m , Z_f , t_c^m and t_c^f , and trial values of M when $M_m = M_f = M$ are assumed, or trial values of F when $F_m = F_f = F$ are assumed.

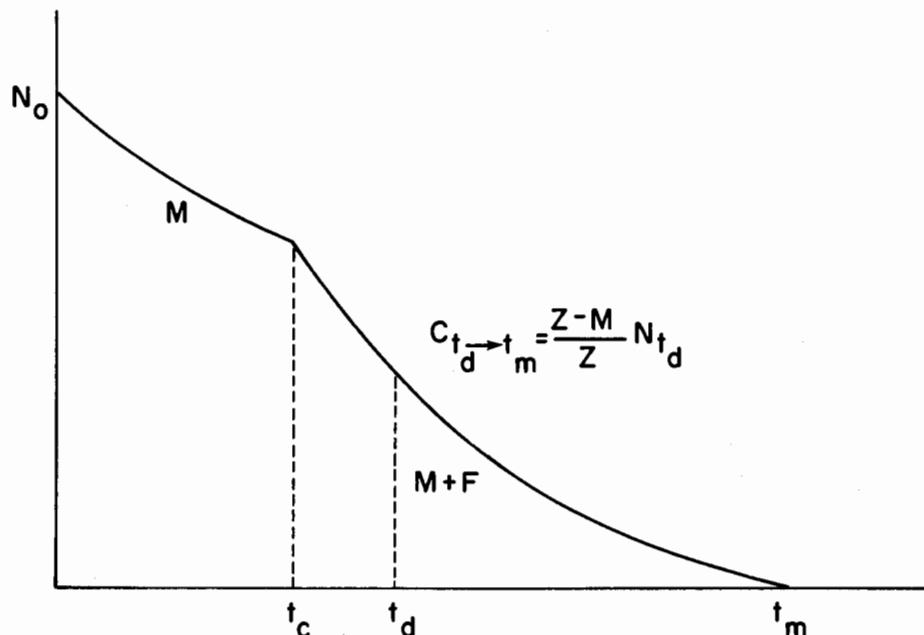


Fig. 1. Diagrammatic representation of the changes in stock numbers with age for a multiple of cohorts where the N_o for each cohort is constant.

The extent of proportional "correlation" between the estimated and observed catch number ratios is used to determine which of the trial values of M or F is the best choice. The parameters estimated for this purpose are the coefficient of determination (r^2), the y-intercept (u) and slope (v) as obtained from a linear (geometric mean) regression analysis, and the sum of the squared differences between the estimated and observed catch number ratios (ΣD^2). Ideally, the best choice is presumed to occur when the r^2 is maximized, u is zero, v is unity, and the ΣD^2 is minimized.

Seasonal Fishery Method

For seasonal fisheries it is convenient to consider three separate time periods: from the age at first capture to the end of the first fishing season immediately following reading the age at first capture; from the end of this season to the end of the next; and from then to maximum age.

The equations enabling estimation of the catch numbers for each period are as follows:

$$C_{t_d \rightarrow t_s} = (Z-M)/(Z-M(1-\Delta t)) \cdot N_o \exp(-Mt_c - (Z-M(1-\Delta t))(t_d - t_c)) \cdot \frac{(1 - \exp(-(Z-M(1-\Delta t))(t_s - t_d)))}{\Delta t} \quad \dots 3)$$

in which $t_d = t_s$ when $t_d > t_s$;

$$C_{t_s \rightarrow t_{s+1}} = (Z-M)/(Z-M(1-\Delta t)) \cdot N_o \exp(-Mt_c - (Z-M(1-\Delta t))(t_d - t_c) - M(1-\Delta t)) \cdot \frac{(\exp(-(Z-M(1-\Delta t))(t_d - (t_s + (1-\Delta t)))) - \exp(-(Z-M(1-\Delta t))))}{\Delta t} \quad \dots 4)$$

in which $t_d = t_s + (1 - \Delta t)$ when $t_d < t_s + (1 - \Delta t)$; and $t_d - (t_s + (1 - \Delta t)) = \Delta t$ when

$$t_d > t_{s+1};$$

$$C_{t_{s+1} \rightarrow t_m} = (Z-M)/Z \cdot N_o \exp(-Mt_c - (Z-M(1-\Delta t))(t_s - t_c) - \frac{3}{2}M(1-\Delta t) - (Z-M(1-\Delta t)) - Z(t_d - (t_s + 1 + \frac{(1-\Delta t)}{2}))) \quad \dots 5)$$

in which $t_d = t_s + 1 + \frac{(1-\Delta t)}{2}$ when $t_d < t_s + 1 + \frac{(1-\Delta t)}{2}$,

t_s is the age at the end of the first fishing season, and Δt is the season duration (Fig. 2).

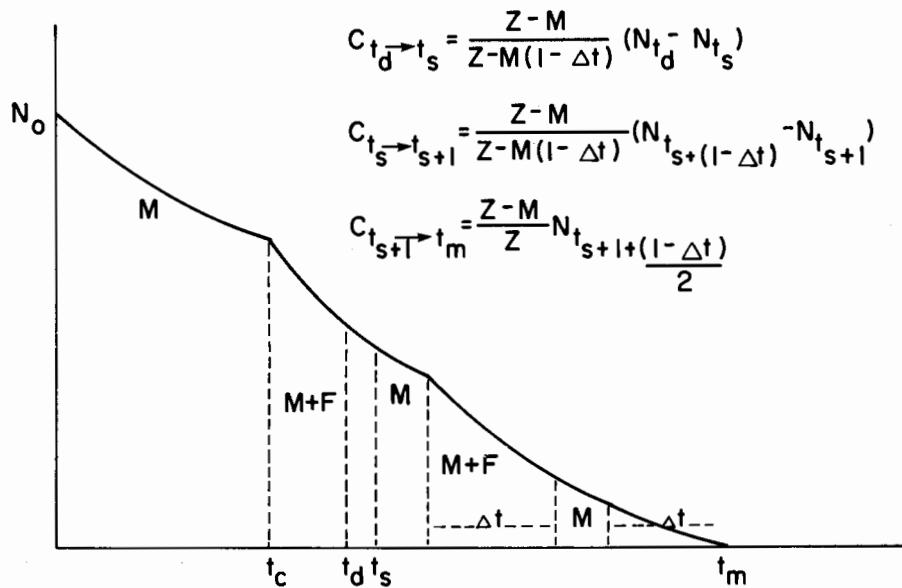


Fig. 2. Diagrammatic representation of the changes in stock numbers with age for a multiple of cohorts where the N_0 for each cohort is constant.

Subscripting separately for the males and females, and combining the catches for the three periods gives

$$\frac{C_{t_d \rightarrow t_m}^f}{C_{t_d \rightarrow t_m}^m} = \frac{C_{t_d \rightarrow t_s}^f + C_{t_s \rightarrow t_{s+1}}^f + C_{t_{s+1} \rightarrow t_m}^f}{C_{t_d \rightarrow t_m}^m + C_{t_d \rightarrow t_s}^m + C_{t_s \rightarrow t_{s+1}}^m + C_{t_{s+1} \rightarrow t_m}^m} \quad \dots 6)$$

This equation enables estimation of the catch number ratios for each observed combination of t_d^m and t_d^f from predetermined values of Z_m , Z_f , t_c^m , t_c^f , t_s^m , t_s^f and Δt (taken as the same for each sex), and trial values of M when $M_m = M_f = M$ are assumed, or trial values of F when $F_m = F_f = F$ are assumed.

Again the extent of proportional "correlation" between the estimated and observed catch number ratios is used to determine which of the trial values of M or F are the best choice values. The ideal values are presumed to occur when the r^2 is maximized, u is zero, v is unity, and the ΣD^2 is minimized.

Application

Each of the methods was applied to data simulated from known parameter values. The model used for the simulations was such as to enable the life history for each sex to be considered as the sum of many small time intervals, each of which could be associated with different parameter values.

The simulations firstly involved estimating the stock numbers at the end of each time interval as for a single cohort, from

$$N_2 = N_1 \exp(-(F+M) \Delta t) \quad \dots 7)$$

where N_1 and N_2 are the stock numbers at the beginning and end of the chosen interval of duration Δt . The catch numbers during each interval were then estimated from

$$C = F/Z (N_1 - N_2) \quad \dots 8)$$

The number (N_o) at the beginning of the first time interval was taken as 10,000.

These estimated catch numbers for each sex, plus known parameter values (Z_m , Z_f , t_c^m , t_c^f , etc) were used with the relevant program to provide estimates for the natural and fishing mortality coefficients. The validity of the methods was then judged on the basis of a comparison of the estimated and known values for F and M .

Results

The known parameter values used with respect to each data simulation are given in Tables 1 and 2, along with the estimates obtained for the mortality coefficients. In Examples 1, 3, 5 and 7 where the $M_m = M_f$ assumption is valid, close agreement was established between the known and estimated M values. Similarly, in Examples 1, 2, 5, 6 and 8 where the $F_m = F_f$ assumption is valid, close agreement was established between the known and estimated F values.

In all cases where the assumptions were invalid, such as Examples 2, 4, 6 and 8 in the $M_m = M_f$ assumption, and Examples 3 and 7 in the $F_m = F_f$ assumption, obviously unrealistic estimates were obtained for the mortality coefficients. No solutions were possible in the case of Example 4 where $F_m = F_f$ is assumed.

The "correlation" parameters which proved most useful for identifying the best choice estimates for the F and M values were v and ΣD^2 . The single exception was for Example 8 where $F_m = F_f$ is assumed, and for which the best choice was decided entirely on the value of v . The parameters r^2 and u were found not useful in isolation from the other parameters.

Discussion

On the basis of these results, it appears that both the non-seasonal and seasonal methods are useful in providing estimates of the mortality coefficients, provided the underlying assumptions are adequately met. The only situation in which it was not possible to obtain solutions occurred when $t_c^m = t_c^f$ with the $F_m = F_f$ assumption in Example 4.

The latter can be readily understood from the simplification of equation (2). When $F_m = F_f$ and $t_c^m = t_c^f$, this equation becomes

$$C_{t_c \rightarrow t_m}^f / C_{t_c \rightarrow t_m}^m = Z_m / Z_f \cdot N_o^f / N_o^m \cdot \exp (Z_m t_d^m - Z_f t_d^f) \quad \dots 9)$$

for which the right hand side is constant for all values of F .

The above limitation does not apply to the method applied to a seasonal fishery. In Example 8, $t_c^m = t_c^f$ with the $F_m = F_f$ assumption, yet a solution could be obtained for F . It is suspected, however, that in most cases where $t_c^m = t_c^f$, problems will occur when identifying the best choice F and M values. As mentioned earlier, the parameter ΣD^2 was not a good indicator for Example 8, whereas it had been for all the others.

A positive aspect of the method is that the results themselves provide a basis for deciding whether the assumption $F_m = F_f$ or $M_m = M_f$ is valid. In all the examples where one of the assumptions was known to be invalid, obviously unrealistic best choice estimates were obtained in association with the invalid assumption, and correct estimates in association with the valid assumption.

In Examples 4 and 8 where each of $F_m = F_f$ and $M_m = M_f$ were valid, correct best choice estimates for the mortality coefficients were obtained in both assumptions. Although this was not tested, it is suspected that when neither assumption is valid, this may not always be indicated by the output of obviously unrealistic estimates of the mortality coefficients.

In such circumstances it will not be possible to determine a single best choice solution for the mortality parameters. It is possible to obtain an array of solutions, many of them being clearly unrealistic. The approach would be to apply the method (based on equation 2 or 6 as appropriate) with a range of realistic values for, e.g., M_m , and by trial and error determine the associated values of M_f (and hence F_m and F_f) for which there is agreement between the estimated and observed length- or age-frequency distributions.

Concerning the other assumptions, it is important that the N_o for each sex and cohort be reasonably constant; a somewhat similar method by Caddy (1984) is less restrictive in this sense.

Table 1. Known parameter values^a and estimated mortality coefficients associated with applying the non-seasonal fishery method.

General description	Known parameter values		Estimated mortality coefficients when			
	Males	Females	$M_m = M_f$ assumed Males	Females	$F_m = F_f$ assumed Males	Females
Example 1						
$F_m = F_f$	$F=1.8$ $M=0.4$	$F=1.8$ $M=0.4$	$F=1.8$ $M=0.4$	$F=1.8$ $M=0.4$	$F=1.8$ $M=0.4$	$F=1.8$ $M=0.4$
$M_m = M_f$	$Z=2.2$	$Z=2.2$	$Z=2.2$	$Z=2.2$	$Z=2.2$	$Z=2.2$
$t_c^m \neq t_c^f$	$t_c=1.333$	$t_c=1.666$				
Example 2						
$F_m = F_f$	$F=1.8$ $M=0.4$	$F=1.8$ $M=0.6$	$F=3.53$ $M=-1.33$	$F=3.73$ $M=-1.33$	$F=1.8$ $M=0.4$	$F=1.8$ $M=0.6$
$M_m \neq M_f$	$Z=2.2$	$Z=2.4$	$Z=2.2$	$Z=2.4$	$Z=2.2$	$Z=2.4$
$t_c^m \neq t_c^f$	$t_c=1.333$	$t_c=1.666$		solutions unrealistic		
Example 3						
$F_m \neq F_f$	$F=2.25$ $M=0.6$	$F=1.8$ $M=0.6$	$F=2.25$ $M=0.6$	$F=1.8$ $M=0.6$	$F=3.7$ $M=-0.85$	$F=3.7$ $M=-1.3$
$M_m = M_f$	$Z=2.85$	$Z=2.4$	$Z=2.85$	$Z=2.4$	$Z=2.85$	$Z=2.4$
$t_c^m \neq t_c^f$	$t_c=1.666$	$t_c=1.166$			solutions unrealistic	
Example 4						
$F_m = F_f$	$F=1.8$ $M=0.9$	$F=1.8$ $M=0.6$	$F=a$ $M=a$	$F=a$ $M=a$	$F=a$ $M=a$	$F=a$ $M=a$
$M_m \neq M_f$	$Z=2.7$	$Z=2.4$	$Z=a$	$Z=a$	$Z=a$	$Z=a$
$t_c^m = t_c^f$	$t_c=1.166$	$t_c=1.166$		^a solutions unrealistic		^a no solutions

^aIn all examples (see Morgan and Pauly, Part II, this vol.), where the catch numbers are grouped by length intervals, the values used for the von Bertalanffy growth constants were: $L_\infty = 25.0$, $K = 0.4$ and $t_o = -0.5$.

Table 2. Known parameter values^a and estimated mortality coefficients associated with applying the seasonal fishery method.

General description	Known parameter values		Estimated mortality coefficients when			
	Males	Females	$M_m = M_f$ assumed Males	M_f assumed Females	$F_m = F_f$ assumed Males	F_f assumed Females
Example 5						
$F_m = F_f$	$F=0.3$ $M=0.3$	$F=0.3$ $M=0.3$	$F=0.3$ $M=0.3$	$F=0.3$ $M=0.3$	$F=0.3$ $M=0.3$	$F=0.3$ $M=0.3$
$M_m = M_f$ $t_c^m \neq t_c^f$	$Z=0.6$ $t_c^m = 1.416$ $t_s^m = 1.916$ $\Delta t = 0.666$	$Z=0.6$ $t_c^f = 1.25$ $t_s^f = 1.916$	$Z=0.6$	$Z=0.6$	$Z=0.6$	$Z=0.6$
Example 6						
$F_m = F_f$	$F=0.3$ $M=0.6$	$F=0.3$ $M=0.3$	$F=-0.4$ $M=1.3$	$F=-0.7$ $M=1.3$	$F=0.3$ $M=0.6$	$F=0.3$ $M=0.3$
$M_m \neq M_f$	$Z=0.9$	$Z=0.6$	$Z=0.9$	$Z=0.6$	$Z=0.9$	$Z=0.6$
$t_c^m \neq t_c^f$	$t_c^m = 1.25$ $t_s^m = 1.916$ $\Delta t = 0.666$	$t_c^f = 1.416$ $t_s^f = 1.916$	solutions unrealistic			
Example 7						
$F_m \neq F_f$	$F=0.3$ $M=0.6$ $Z=0.9$	$F=0.6$ $M=0.6$ $Z=1.2$	$F=0.3$ $M=0.6$ $Z=0.9$	$F=0.6$ $M=0.6$ $Z=1.2$	$F=4.8$ $M=-3.9$ $Z=0.9$	$F=4.8$ $M=-3.6$ $Z=1.2$
$M_m = M_f$	$t_c^m = 1.25$	$t_c^f = 1.416$				
$t_c^m \neq t_c^f$	$t_s^m = 1.916$ $\Delta t = 0.666$	$t_s^f = 1.916$	solutions unrealistic			
Example 8						
$F_m = F_f$	$F=0.3$ $M=0.6$ $Z=0.9$	$F=0.3$ $M=0.3$ $Z=0.6$	$F=-0.8$ $M=1.7$ $Z=0.9$	$F=-1.1$ $M=1.7$ $Z=0.6$	$F=0.3$ $M=0.6$ $Z=0.9$	$F=0.3$ $M=0.3$ $Z=0.6$
$M_m \neq M_f$	$t_c^m = 1.25$	$t_c^f = 1.25$				
$t_c^m = t_c^f$	$t_s^m = 1.916$ $\Delta t = 0.666$	$t_s^f = 1.916$	solutions unrealistic			

^aIn all examples (see Morgan and Pauly, Part II, this vol.), where the catch numbers are grouped by length intervals, the values used for the von Bertalanffy growth constants were: $L_\infty = 25.0$, $K = 0.4$ and $t_0 = -0.5$.

Equal numbers of males and females at zero age would seem to be reasonable in respect to a single cohort. Where it is not reasonable between cohorts, an attempt should be made to minimize the likely error by applying the methods to the catch number distributions for a combination of years. This combining of data should be for years in which F , M , t_c , etc. have remained reasonably constant.

Problems with the N_0 not being the same will be most acute for fisheries based on few cohorts. Where the number of cohorts represented in the catches is large, differences in the N_0 is likely to have a negligible effect on the estimates for the mortality coefficients.

Another safeguard from using catch numbers relevant to a relatively large time interval (at least a year) is to ensure that the observed proportions by sex in the catches accurately reflect the proportions in the sea. This would be necessary when the sexes exhibit short-term differences in their vulnerabilities to capture.

References

- Caddy, J.F. 1984. Method of solving for natural mortality rate for stock components with different schedules of growth and mortality. *Can. J. Fish. Aquat. Sci.* 41: 1226-1230.
Sanders, M.J. 1977. Estimation of mortality coefficients for exploited fish populations in which the sexes exhibit divergent growth rates. *J. Cons. Int. Explor. Mer* 37(2): 192-193.

Incorporating Age Data Into Length-Based Stock Assessment Methods

G.R. MORGAN*

Kuwait Institute for Scientific Research
P.O. Box 24885, Safat
Kuwait

Morgan, G.R. 1987. Incorporating age data into length-based stock assessment methods, p. 137-146. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

A technique is described which enables data on length-at-age to be incorporated into the ELEFAN I method of length-composition analysis. It is shown that by utilizing this additional information in a single assessment method, several important advantages result. Firstly, the resolution within the ESP/ASP response surface of ELEFAN I is improved, resulting in the ability to identify parameter values even when the response surface has extensive "plateaus" or multiple peaks. Secondly, a test of consistency between the length-at-age and the length composition is possible and one such test is described.

As an example of the application of the new technique, data on hamra (*Lutjanus coccineus*) taken in Kuwait waters in 1983 has been used and the optimal sampling strategy for this species considered.

Introduction

Length-based stock assessment methods are being increasingly used in fish stock assessment, not only in instances where age composition is unavailable (Morgan 1985) but also where age composition either exists or is easily collected. The rationale behind this latter use of length-based methods is that it is generally assumed or has been shown (Morgan 1984) that assessment based on easily collected length composition samples is more cost-effective than those based on the more tedious and time-consuming methods of fish aging, although as J.A. Gulland (pers. comm.) points out, comparisons between the costs of assessment based on length composition data and those based on age composition data should take the relative precision of the two approaches into account.

The division of assessment methods into those based on length composition data and those based on age composition data is essentially an artificial one since, if data on both parameters

*Present address: c/o W.A. Marine Research Laboratories, P.O. Box 20, North Beach 6020, W.A., Australia.

exist, transformation of the various equations from length-based to age-based is, mathematically, a trivial conversion. Such conversion, in practice, is carried out via some type of growth equation (usually von Bertalanffy) relating age and length. Kirkwood (1983) has used joint maximum likelihood estimation procedures to estimate growth curve parameters from length increment and age-length data.

The continued separation of assessment methods into these two largely artificial classes also has important practical disadvantages.

Firstly, and most importantly, by treating age and length-based assessments as independent entities, a loss of information and consequent loss of precision in the assessment results when compared with an assessment which uses all the information available on length and age. If both age and length data are available, separation of the two assessment procedures, therefore, results in two less precise assessments (from which the fisheries biologist usually has to choose the most "realistic") rather than a single more precise assessment.

Secondly, separation of the assessment techniques does not allow a direct comparison of the length and age composition data for consistency. If the two data sets are very inconsistent, the separate assessments based on these data will probably diverge significantly, again leaving the choice of the most realistic assessment to the judgment of the fisheries biologist.

Thirdly, separation of the assessment techniques does not allow an evaluation of the most appropriate mix of data to be collected for stock monitoring and assessment purposes. For example, estimates of the resultant increase in precision of the various growth and mortality parameters used for assessment can be made if the sampling intensity of age composition data is increased. Likewise, some estimate of increased parameter precision can be made if sampling intensity of length composition data is increased (although this is generally more difficult). However, the separate assessment approaches do not allow an examination of the effects on parameter estimation of simultaneous changes in collection intensity of age and length composition data. The optimum sampling strategy for assessment of a particular fish stock cannot, therefore, be easily determined.

The purpose of this contribution is, therefore, to examine ways by which age composition data can be incorporated into the various length-based assessment methods to produce estimates of the growth parameters (and, therefore, the mortality rates) of a fish stock. As an example of the application of this integrative method, length and age composition data of red snapper or hamra (*Lutjanus coccineus* = *L. malabaricus*?) from Kuwait has been used to derive estimates of the growth parameters necessary for stock assessment purposes. In addition, the question of optimal sampling strategy for stock assessment purposes is examined.

Methods

Two of the more common length-based stock assessment methods are the length-based cohort analysis and virtual population analysis (VPA) techniques (Jones 1979 and 1981; Pauly 1984) and the ELEFAN I program of Pauly and David (1981). The first of these methods uses catch at length data from one instant in time to produce estimates of population size and mortality rates while the ELEFAN programs utilize a restructured time series of length-composition data to estimate both growth parameters and mortality rates (see Morgan and Pauly, Part II, this vol. and other contributions in this volume for other aspects of the ELEFAN package).

Pauly (1984) has shown that VPA can be easily modified to accommodate catch-at-length data and, more importantly, points out why such length-related VPA should be used in preference to the approximate (but, computationally simpler) length-related cohort analyses technique of Jones (1979). It is likely, however, that the conclusions reached with respect to length-related VPA will also apply to its approximation, cohort analysis.

If two data sets are available (one catch-at-age with x age intervals and one catch-at-length with y length intervals), these may be combined with length converted to age using the appropriate growth equation. This, however, results in a combined data set of $x+y+1$ intervals with $x+y$ values of catch. Values of catch for each interval in this combined data set cannot, therefore, be calculated from the information on catch in each age/length interval contained in the two data sets.

However, if the two sets cover *different* ranges of the age/length composition of the catch (so that the total catch for the range considered in each data set is different), the totals of the two data sets may be used to produce $x+y+2$ values of catch for $x+y+1$ intervals. This set of simultaneous equations may, therefore, be solved to produce a catch for each interval of the combined data set.

Data on age at length can also be usefully incorporated into the ELEFAN technique in several different ways.

First, the data can be used to calculate an absolute value of t_0 rather than a relative value. This may be done by using the inverse of the seasonally oscillating von Bertalanffy growth equation although it should be recognized that, if there is an extensive distribution of ages-at-length, such a technique may give a quite different (and inappropriate) value of t_0 than that calculated from length-at-age data. Jones (1981 and Part I, this vol.), Laurec and Mesnil (Part I, this vol.) and Majkowski et al. (Part II, this vol.) have provided useful discussions on the errors in growth parameter estimation generated by the two-dimensional distribution of ages and lengths. If there is more than one point estimate of age and length available, perhaps the most appropriate method of estimating t_0 is to utilize the other (known) von Bertalanffy growth parameters K, L_∞ , C, WP and D (see Pauly and David (1981) and Pauly (Part I, this vol.) for parameter definitions) to determine, using an iterative technique, the value of t_0 at which the sums of squares of the deviations of the length-at-age data is minimized. This technique has been used in the example following.

Second, the point estimates of age and length can also be used in an analogous way to the series of length composition data to estimate, for any set of growth parameters, the proportion of the total age/length observations through which the growth curve passes. This is equivalent, in Pauly and David's (1981) terminology of estimating an ESP/ASP ratio for the age/length data.

The growth curve which provides the best fit to *both* the length composition and the age/length data sets will, therefore, be that curve which maximizes not the length-related ESP/ASP ratio but that which maximizes the expression (ESP/ASP + Pa) where Pa is the proportion of the total age/length data through which the curve passes.

Table 1. Age and length data for hamra (*Lutjanus coccineus*) taken in Kuwait 1983. Tests of normality were carried out by using a chi squared distribution to test the differences in observed and expected (on the basis of a normal distribution) frequencies. NT = test not possible; * = significantly different at 5% level; ** = significantly different at 1% level; NS = not significantly different.

Length group (cm TL)	Age at length			Age group	Length at age			Normality test
	Mean age	Standard deviation	Normality test		Mean length (cm)	Standard deviation		
10-14	0.51	0.02	NT	0	13.5	3.61		NT
15-19	—	—	—	1	27.4	5.64		**
20-24	1.46	0.11	NT	2	38.1	8.66		**
25-29	1.63	0.24	NT	3	44.8	8.44		**
30-34	1.86	0.85	NS	4	52.9	6.18		*
35-39	2.37	0.61	NS	5	57.0	6.15		**
40-44	2.73	0.68	NS	6	60.2	5.08		NS
45-49	3.66	0.81	NS	7	63.3	3.35		**
50-54	4.23	2.77	NS	8	67.0	3.01		NT
55-59	4.88	2.87	NS	9	68.7	3.03		NT
60-64	8.28	7.01	NS	10	66.6	2.61		NT
65-69	16.49	8.30	NS	11	70.2	1.94		NT
70-74	18.69	10.61	NS	12	67.8	2.48		NT
75-79	20.20	13.84		13	66.2	2.32		NT
				14	66.4	2.46		NT
				15	68.1	1.99		NT
				> 15	69.3	—		NT

This technique avoids the statistical complications of comparing stochastic age-at-length with length-at-age data although the distribution of, particularly, ages-at-length needs to be considered in providing an adequate sample of the age/length data. This will be further considered below.

Thirdly, a test for consistency between age/length data and the set of length composition data can be devised. Since the two data sets should be related through a growth curve, it follows that any given growth curve should, if the data sets are consistent, fit both data sets equally well (or badly). A simple test for consistency is then the correlation between the goodness of fit of a series of growth curves to the length composition data (from the ESP/ASP ratios) and to the age/length data (from the values of Pa). Such a test can be carried out in the course of the iterations preceding the identification of the maximum value of (ESP/ASP + Pa).

The age/length data can also be used to test some of the assumptions inherent in the ELEFAN and other techniques. For example, recent methods of size composition analyses (Schnute and Fournier 1980; Fournier and Breen 1983) involving simultaneous estimation of growth and mortality parameters assume a normal distribution of length-at-age. For *Lutjanus coccineus* taken in Kuwait in 1983, this was shown to be an unrealistic assumption although age-at-length data were normally distributed (Table 1).

Similarly the ELEFAN I program assumes that the peaks in a series of restructured length composition data represent age groups (usually year classes). Such an assumption can be tested if age/length data are available and more importantly, if shown to be invalid, the age/length data can be utilized to dissect a peak into its component age groups.

This dissection can be done in several ways but perhaps the simplest is to first use the age/length data to estimate the distribution of ages in the length sample. If this is normally distributed (Table 1 indicates that such age-at-length data are normally distributed in catches of *Lutjanus coccineus* taken in Kuwait in 1983), then the expected cumulative number of age groups, A_i , in length interval i will be a function of the deviation D from the mean age, the sample standard deviation, S and the sample size, n , according to the t distribution (or the normal deviate if n is large). Since $t_{p,n} = D/S$, it follows that $1-p$ of the age groups within a length class will fall within $2St_{p,n}$ of the mean. Weighting of the raw, restructured frequencies by $1/St_{p,n}$ (or $1/S$ if sample sizes are sufficiently large), therefore, takes into account the possibility that the frequency of any length class comprises contributions from a number of different age groups.

Other tests of assumption according to specific requirements may also be devised.

Application of the Modified ELEFAN Technique to Data from Kuwait's Snapper Fishery

Both monthly length composition data and age/length data are available for hamra (*Lutjanus coccineus*), a red snapper which is taken commercially by fish traps in Kuwait and for which landings over the last few years have averaged around 1,000 tonnes, representing some 18% of Kuwait's total fish landings. Hamra is characterized by being a very long-lived species with fish up to 46 years old (as determined from otoliths) having been recorded. The growth rate is such that at the larger sizes a 5-cm length interval may consist of up to 25 separate year classes, although a 5-cm length interval at the smaller sizes after recruitment usually consists of a single year class (see Table 1).

As a result, modes in the length composition samples are not clearly defined and the additional information provided by the age/length data might, therefore, be expected to reduce the uncertainty inherent in the application of a purely length-related approach to assessment of such a stock.

A modified version of the ELEFAN I program was used in a joint analysis of the length composition and age/length data by incorporating into the ELEFAN I program a routine for maximizing the value of $(d_1 \cdot \text{ESP/ASP} + d_2 \cdot \text{Pa})$ where Pa is the proportion of the total age/length data points through which a given growth curve passes, ESP and ASP are as defined (Pauly and David 1981 and Pauly, Part I, this vol.) and d_1 and d_2 are weighting factors which determine the relative influence of the length composition and age/length data. The value of d_1 is set either at 0 (which results in length composition data being ignored) or 1.

If $d_1 = 0$ then $d_2 = 1$ and if $d_1 = 1$ then $d_2 \geq 0$. A value of $d_2 = 0$ and $d_1 = 1$ results in length data being ignored in the analyses while $d_1 = d_2 = 1$ results in equal weighting being given to the two data sets. The values of d_1 and d_2 are required inputs into the maximization routine and, in the present example, d_1 was set at 1 and d_2 values of 0 to 2 in step sizes of 0.1 were tested.

Measures of the degree of consistency between the length composition and the age/length data sets were calculated from the linear correlation coefficient between d_1 ESP/ASP and d_2 Pa for a series of growth curves as described earlier.

VPA analyses were not used with the hamra data sets since the extension of length-related VPA to include age/length data appears trivial.

Computer programs in HP BASIC, designed to run on an HP 9845 microcomputer, are available for the joint analyses of length composition and age/length data using this modification of ELEFAN I.^a

Results

LENGTH COMPOSITION AND AGE/LENGTH DATA

Age and length data were available for 519 fish taken during 1983 covering a size range of 15-80 cm total length (TL). Table 1 presents the essential characteristics of this data set. Ages were adjusted to take into account the sampling month relative to a nominal birthday of 1 January, so that, for example, a 2+ fish sampled in February was assigned an age of 2.125 years. From Table 1, it is apparent that in the majority of cases, age data were normally distributed within a length class whereas length data were generally not distributed normally within year classes.

Length composition data were available for each month of 1983 with monthly sample sizes ranging from 115 to 656 fish. The data were grouped into 5-cm length intervals for further analysis.

APPLICATION OF THE ELEFAN TECHNIQUE TO SIZE COMPOSITION DATA

The value of d_2 was set at 0 and d_1 at 1 for this analysis. The modified ELEFAN I program, when applied to the 1983 monthly set of length composition samples of hamra, resulted in an optimum value of the ESP/ASP ratio of 0.711 with the following parameters:

$$\begin{aligned} L_{\infty} &= 102 \text{ cm TL} \\ K &= 0.17 \\ WP &= 0.3 \\ C &= 0.7 \\ "t_o" &= -0.185 \end{aligned}$$

These parameters differ somewhat from those previously calculated using data for 1982 (Morgan 1984) with the differences probably reflecting sampling errors in the collection of the length data.

^aEditor's note: Brey and Pauly (1986) have developed versions of this routine for Apple II (CP/M), or IBM PC and their compatibles (see Program Number 5 in Morgan and Pauly, Part II, this vol.).

CALCULATION OF GROWTH PARAMETERS
USING LENGTH-AT-AGE DATA

The value of d_2 was set at 1 and d_1 at 0 for this analysis. This resulted in a growth curve with the following parameters:

$$\begin{aligned} L_\infty &= 87 \text{ cm TL} \\ K &= 0.23 \\ WP &= 0 \\ C &= 1.0 \\ "t_o" &= 0.41 \end{aligned}$$

[WP and C refer to seasonal growth oscillations; see Gaschütz et al. (1980) and Pauly, Part I, this vol.]

At this point, the value of Pa was 0.353.

JOINT ANALYSIS OF LENGTH AND AGE DATA

With d_1 set at 1.0, values of d_2 between 0.1 and 2.0 were tested and the parameter combination which maximized the value of $(d_1 \cdot ESP/ASP + d_2 \cdot Pa)$ identified for each d_2 value.

Table 2 presents these parameter combinations together with those for the length composition and age/length data. It is apparent from Table 2 that even small weightings of age/length data (i.e.,

Table 2. Parameter combinations which maximized the value of $(d_1 \cdot ESP/ASP + d_2 \cdot Pa)$ for various values of d_1 and d_2 . See text for symbol explanations.

d_1	d_2	K	L_∞ (cm)	WP	C	t_o	Comments
1	0.0	0.17	102	0.3	0.7	—	Only length data
1	0.1	0.18	101	0.2	0.9	0.09	
1	0.2	0.16	100	-0.3	1.0	-0.38	
1	0.3	0.24	87	0.2	1.0	-0.04	
1	0.4	0.24	88	0.1	1.0	-0.08	
1	0.5	0.23	87	0	1.0	0.38	
1	0.6	0.23	88	0	1.0	0.26	
1	0.7	0.24	88	0.1	0.9	0.21	
1	0.8	0.24	87	1.0	1.0	0.36	
1	0.9	0.22	89	0	0.9	0.43	
1	1.0 ^a	0.24	88	0.1	0.9	0.41	Gradually increasing weight of age data
1	1.5	0.23	87	0	1.0	0.40	
1	2.0	0.23	87	0	1.0	0.41	
0	1.0	0.23	87	0	1.0	0.41	

^aEqual weighting of age and length data.

values of $d_2 > 0.3$) result in the optimum parameter combination moving away from that identified on the basis of length data alone, to rapidly reach values more in accordance with those based on the age/length data set. This, therefore, indicates that even a small amount of age/length data would, in combination with the length composition data, be sufficient for improving the resolution within the ESP/ASP response surface.

The most appropriate weighting of length composition and age/length data has, in this instance, been taken as $d_1 = 1$ and $d_2 = 1$ (i.e., equal weighting of the two data sets). The parameter combination which maximized the value of $(d_1 \cdot \text{ESP/ASP} + d_2 \cdot \text{Pa})$ at these values of d_1 and d_2 was, from Table 2:

$$\begin{aligned} L_\infty &= 88 \text{ cm TL} \\ K &= 0.24 \\ WP &= 0.1 \\ C &= 0.9 \\ "t_o" &= 0.41 \end{aligned}$$

At this point, the value of the ESP/ASP ratio for the length composition data was 0.694 (compared with the optimum of 0.711 when the length data were considered alone) and the value of Pa for the age composition data was 0.267 (compared with 0.353 for the parameter combination at the point of best fit of the length-at-age data). This low value of Pa no doubt reflects the high variability of length-at-age data, particularly for ages 2 to 6 (Table 1).

As a test to determine whether the value of Pa for any given growth curve reflects the goodness of fit of that curve to the length-at-age data, a number of growth curves were fitted and both the value of Pa and the proportion of variation accounted for (R^2) by the curve calculated. The relationship between these two quantities is shown in Fig. 1 from which it is apparent that the value of Pa is closely related to the variation accounted for ($R^2 = 0.907$, $df = 20$, $P < 0.001$) and, hence, provides an adequate measure of the goodness of fit of a growth curve to the age/length data.

Using the correlation between 15 sets of ESP/ASP ratios and values of Pa for a range of growth parameters, a test of consistency between the two data sets was carried out as previously described. This resulted in a value of R, the correlation coefficient, of 0.798, indicating that the two data sets were consistent with each other at the 5% level.

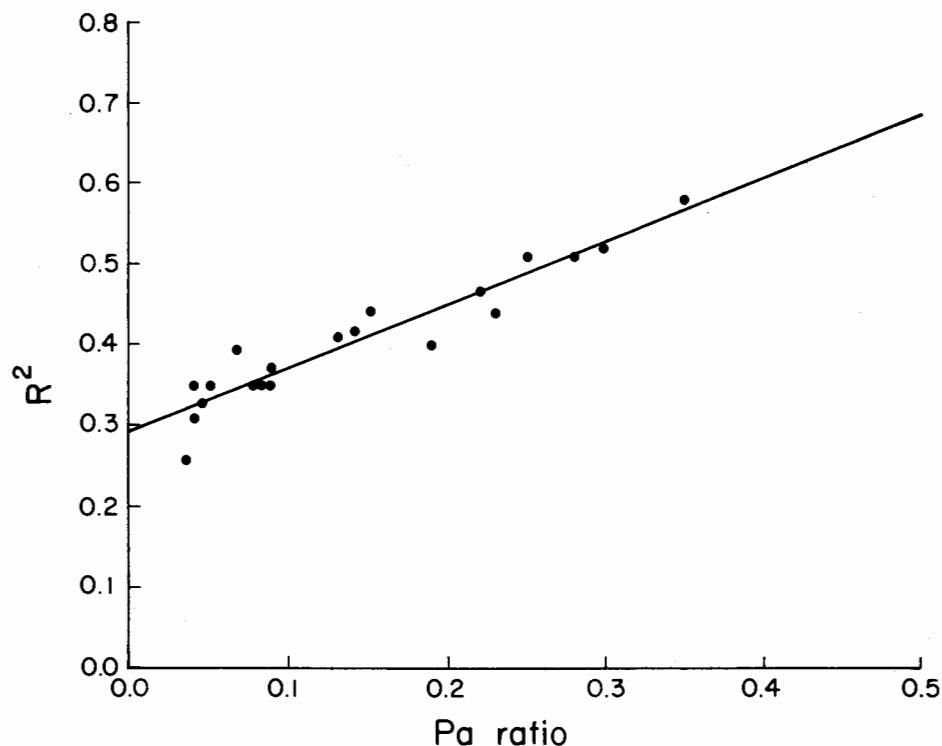


Fig. 1. The relationship between Pa (the number of length-at-age data points accumulated by a single growth curve expressed as a proportion of the total number of available data points) and R^2 (the amount of variation accounted for) for various growth curves of hamra *Lutjanus coccineus* taken in Kuwait in 1982. The significant relationship indicates that Pa is an adequate measure of the goodness of fit of a growth curve to the length-at-age data.

The importance of carrying out this test over a suitably wide range of growth parameter values (particularly K and L_{∞}) was highlighted during preliminary running of the program when restricting the parameter values to narrow ranges sometimes resulted in the two data sets being indicated as significantly inconsistent. This was a result of either or both data sets being constrained to areas where local optima existed, hence giving spurious results. It, therefore, appears important to conduct the test of consistency over as wide a range of realistic growth parameter values as possible.

Fig. 2 shows the growth curve estimated from the two data sets together with the set of size composition data and the observed mean lengths-at-age for the dominant year classes.

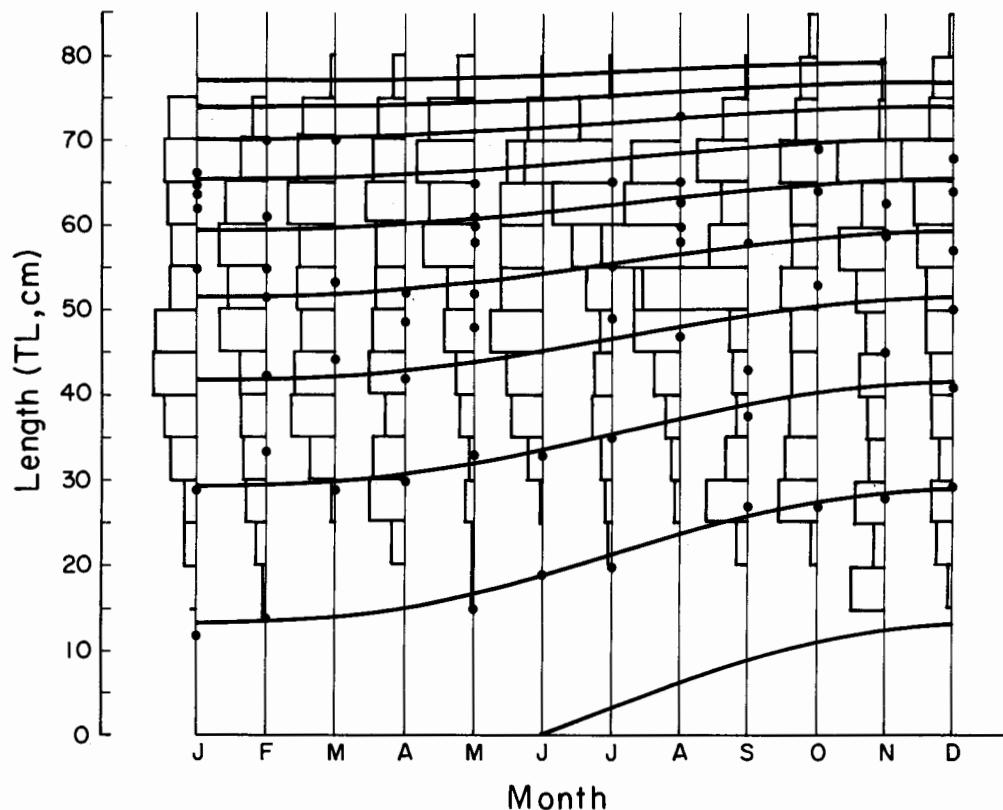


Fig. 2. Seasonally oscillating growth curve fitted to the combined length composition and length-at-age data for hamra *Lutjanus coccineus* taken in Kuwait in 1983. The monthly size composition data are shown along the vertical axis while the mean lengths-at-age (based on a varying number of observations) each month are indicated by a dot.

Discussion

The estimation of growth parameters is critical in the application of any length-based stock assessment procedure since it is this growth curve which provides the link between length data and the time base. Its use in converting lengths to either relative or absolute ages enables subsequent estimation of time-related mortality rates, essential inputs for fisheries stock assessment purposes.

The estimation of these critical growth parameters from length composition data for hamra (*Lutjanus coccineus*) in Kuwait provided values which were not in accordance with observed data on length-at-age. The reason for this discrepancy is readily apparent when the values of the ESP/ASP ratios derived from the ELEFAN technique for various growth parameter inputs are examined (Table 3). The response surface, of which Table 3 shows a part, has a broad "plateau" over a range of K and L_{∞} values with a series of local optima. The maximum value of the ESP/ASP ratio selected

Table 3. Values of the ESP/ASP (x 1,000) ratio derived using the ELEFAN I program of Pauly and David (1981) for various values of growth parameters, K and L_{∞} of hamra (*Lutjanus coccineus*) taken in Kuwait in 1983. Values of the other growth parameters are C = 1.0, WP = 0.1. Dotted line indicates the extensive "plateau" in the ESP/ASP response surface.

		L_{∞} (cm TL)						
		86	88	90	95	100	105	110
K	.14	380	304	324	523	340	104	246
	.15	324	211	676	555	562	481	211
	.16	222	611	672	679	580	430	580
	.17	446	550	684	584	621	702	685
	.18	320	684	522	582	691	640	520
	.19	514	656	426	638	620	626	404
	.20	621	580	154	646	636	673	121
	.25	674	690	499	422	146	422	85
	.30	61	25	101	24	99	311	143

by the ELEFAN I program was one of these local optima which was marginally higher than some of the other local optima (see also Rosenberg and Beddington, Part I, this vol.). Unfortunately, the point selected did not reside in the region of the correct growth parameter values. The value of incorporating the additional information provided by the age/length data into the analysis was that these data could be used to identify which of the local optima of the ESP/ASP response surface more closely corresponded to the correct growth parameter values.

The identification of the "correct" local optimum on the ESP/ASP response surface could not have been achieved for hamra without some length-at-age data since the large number of year classes and the variability of length within a year class (Table 1) tended to obscure definitive peaks in the length composition data and so lead to "plateaus" in the ESP/ASP response surface. Such a problem might not be so important in fish where either the variability in length within a year class is small and/or there are few year classes. For example, the ELEFAN I program identified almost identical growth parameters as those calculated from length-at-age data for the short-lived, fast-growing newaby (*Otolithes argenteus*) in Kuwait (Morgan 1984).

The question of how much and what type of length-at-age data is required to support an analysis based on ELEFAN I of length composition data (or more broadly, what mix of length and age data is desirable to evaluate growth parameters) requires further examination.

Firstly, if only a single point of length-at-age data is available or if multiple observations of length at essentially the same age exist, little can be achieved apart from calculating a value of t_o . However, if observations of length at various ages are available, these data can be incorporated into the ELEFAN I program as described earlier.

The degree to which such length-at-age data is useful in describing the growth curve will depend not only on the number and range of observations available but also on the variability of length at any age. At ages where length variability is great, a greater number of observations of length at that age will be needed to adequately define the distribution of length while, conversely, a smaller number of observations should be required to define the length distribution at ages where length variability is small (usually the younger ages).

Therefore, the first steps in designing a sampling program to gather growth information might usefully be a large-scale, short-duration collection of length-at-age data designed to measure the

variability of length-at-age as well as an examination of the ESP/ASP response surface derived by the ELEFAN I program from the length composition data. If such a response surface has extensive plateaus, it may reasonably be assumed that definition of the growth parameters will rely more heavily on the length-at-age data than it would if the surface showed an obvious peak. Data on length variability at age and the number of age groups in the population can then be used to determine the number of samples needed to define the growth curve to the required level of precision.

It should be noted that, unless the length composition and length-at-age data are inconsistent (which indicates a sampling problem), the ESP/ASP response surface should include the "correct" growth parameter combination in a plateau or peak. The role of length-at-age data might, therefore, be viewed as locating, within that plateau or peak, the most appropriate set of growth parameters and, hence, improving the resolution of those parameters. Sampling intensity of length-at-age data, therefore, need only be sufficient to identify the growth parameter combination within the range encompassed by the ESP/ASP plateau and, if, as is usual, variability in length at the younger ages is small, collection of a small amount of length-at-age data for the younger age groups may well provide sufficient information to achieve this.

The final choice of an adequate sampling system will obviously be species-specific and should take into consideration not only the sampling intensity for length composition and length-at-age data (after taking into account the variability of length-at-age) but also the costs involved in collecting such data. However, by combining length and age composition data into one analysis, the sampling intensity required to reach any given level of precision will be less than that required in carrying out separate analyses on the two data sets.

The incorporation of length-at-age data into length-based assessment methods (principally the ELEFAN I program) has, therefore, allowed a more realistic identification of growth parameters for hamra in Kuwait waters than was possible using length data alone. The technique appears efficient in that it makes use of a large proportion of the information content in the length and age data and, hence, provides significant advantages (such as data consistency tests and possible reduction in sampling intensity) over treating the two data sets as independent entities. Further work is needed both to extend this technique to mortality estimation procedures and to examine the use of joint maximum likelihood methods in analyzing combined sets since the general principle of data integration and joint analyses appears worthy of further investigation.

References

- Brey, T. and D. Pauly. 1986. Electronic length-frequency analysis: a revised and expanded user's guide to ELEFAN 0, 1 and 2. Berichte des Instituts für Meereskunde an der Christian-Albrecht Universität Kiel. No. 149, Kiel, 76 p.
- Fournier, D. and P. Breen. 1983. Estimation of abalone mortality rates with growth analysis. Trans. Amer. Fish. Soc. 112: 403-411.
- Gaschütz, G., D. Pauly and N. David. 1980. A versatile BASIC program for fitting weight and seasonally oscillating length growth data. J. Cons., Cons. Int. Explor. Mer Doc C.M. 1980/D: 6. 23 p. Mimeo.
- Jones, R. 1979. An analysis of a *Nephrops* stock using length composition data. Rapp. P.-V. Réun. CIEM 175: 259-269.
- Jones, R. 1981. The use of length composition data in fish stock assessment (with notes on VPA and cohort analysis). FAO Fish. Circ. 734. 60 p.
- Kirkwood, G.P. 1983. Estimation of von Bertalanffy growth curve parameters using both length increment and age-length data. Can. J. Fish. Aquat. Sci. 40: 1405-1411.
- Morgan, G.R. 1984. Application of length-based stock assessments to Kuwait's fish stocks. ICLARM Newsl. 6(4): 3-4.
- Morgan, G.R. 1985. Stock assessment of pomfret (*Pampus argenteus*) in Kuwait waters. J. Cons., Cons. Int. Explor. Mer 42: 3-10.
- Pauly, D. 1984. Fish population dynamics in tropical waters: a manual for use with programmable calculators. International Center for Living Aquatic Resources Management, Manila, Philippines.
- Pauly, D. and N. David. 1981. ELEFAN I, a BASIC program for the objective extraction of growth parameters from length frequency data. Meeresforsch. 28: 205-211.
- Schnute, J. and D. Fournier. 1980. A new approach to length frequency analysis: growth structure. Can. J. Fish. Aquat. Sci. 37: 1337-1351.

**Assessment of Short-Lived Stocks with Special Reference to Kuwait's
Shrimp Fisheries: A Contrast of the Results Obtained from Traditional and
Recent Size-Based Techniques**

C.P. MATHEWS
M. AL-HOSSAINI
A.R. ABDUL GHAFFAR
M. AL-SHOUSHANI

*Mariculture and Fisheries Department
Kuwait Institute for Scientific Research
P.O. Box 24885, Safat
Kuwait*

Mathews, C.P., M. Al-Hossaini, A.R. Abdul Ghaffar and M. Al-Shoushani. 1987. Assessment of short-lived stocks with special reference to Kuwait's shrimp fisheries: a contrast of the results obtained from traditional and recent size-based techniques, p. 147-166. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

A review of the biology and stock assessments of three species of shrimps important in Kuwait (*Penaeus semisulcatus*, *Metapenaeus affinis* and *Parapenaeopsis stylifera*) is presented, with emphasis on their growth parameters and mortality rates, and on the use of ELEFAN I and II for estimation of these. The major result obtained is the evidence for a relative stability of growth parameters, and for a strong dependence of natural mortality rate on predator stock sizes.

Of all methods used to date in Kuwait, the ELEFAN programs were the most flexible and provided the most reliable estimates of growth and mortality. However, successful application of the programs presupposes and must rest upon a good understanding of the biology and life-cycle of the species studied. When such previous knowledge does not exist, preliminary inferences based on application of ELEFAN I and II must be confirmed by subsequent biological studies.

Introduction

Assessment of stocks of short-lived aquatic organisms, such as squid (Pauly 1985) and shrimp (Mathews and Abdul-Ghaffar 1986) presents special problems. These are particularly acute when, as in the case of shrimp, it is impossible to carry out aging by examining the harder parts of the body. In spite of these difficulties, it is often necessary to provide timely and accurate advice on the management of such stocks, many of which are very valuable. This is particularly true when there is strong evidence of a long-term decline in recruitment, as has been shown for Kuwait's shrimp stocks (Morgan and Garcia 1982).

Mathews and Abdul-Ghaffar (1986) describe the results of shrimp stock assessments carried out in Kuwait. A two-pronged approach was adopted: catch-and-effort data obtained from interviews with fishermen and from fishing companies' records were used to construct a surplus yield model. At the same time, biological data were used to provide estimates of growth and mortality parameters and a dynamic pool model was applied to the shrimp populations. Results showed that management advice obtained independently by application of these two models was similar even though the databases and analytical techniques employed were entirely different. This analysis pinpointed the levels at which biomass yield and economic yield could be optimized (~6,500 and ~3,000 standard days fishing per year, respectively, which is much less than the present level of ~10,000 to 12,000 days per year). The dynamic pool analysis used was based on estimates of growth and mortality rates obtained by a combination of techniques (Mathews and Al-Hossaini 1984), most of which have been used in other fisheries with varying degrees of success. In 1982, it was decided that ELEFAN-based techniques (Pauly and David 1980; Pauly 1982) should be applied to Kuwait shrimp, together with the techniques used traditionally. The object of this paper is to report on some of the results of the work conducted since.

Methods

TRADITIONAL METHODOLOGY

Growth rates were determined using traditional modal progression analysis (Mathews and Al-Hossaini 1984). Von Bertalanffy growth curves were fitted to size-at-age data so obtained. Maturation of female shrimp was quantified using gonadal index (Al-Hossaini 1981; Shoushani 1985). These data were combined with data obtained from sea surveys of R/V Oloum I, and with data on catch, effort and catch per unit effort obtained through monthly interviews from 1978 to 1984. The life cycles of *Penaeus semisulcatus*, *Metapenaeopsis affinis* and *Parapenaeopsis stylifera* were studied by Mohammed et al. (1981b), Mathews (1982b, 1982c), the Food and Agriculture Organization (FAO 1982), and Mathews and Al-Hossaini (1984). Using the data presented in these reports, it was possible to determine the number of shrimp of each sex and species in each cohort for each month from 1978 to 1981 after catch rates in the artisanal fishery were adjusted for the effort expended in each month. Time at recruitment was fixed for each cohort, usually in April or May for the "spring" spawning cohort. The resulting plots were then used for estimating total mortality by fitting a semi-logarithmic regression. However, Mathews and Al-Hossaini (1984), who carried out this work, expressed dissatisfaction with the results obtained and were concerned with possible biases in estimates of growth parameters (particularly K) and the effects that these might have on estimates of M obtained by Pauly's (1980) technique. Attempts to estimate M by a plot of Z on total effort failed because the data produced an apparently random scatter. All commercial size categories referred to below were converted to carapace length using morphometric data in Farmer (1980).

RECENT SIZE-BASED TECHNIQUES OTHER THAN ELEFAN

The need for size-based assessments in fisheries research was noted by Mathews and Mead (1971). Mathews (1974) tried various techniques, including analysis by the methods of Cassie (1954) and the NORMSEP computer program (Abramson 1971). Mathews found these techniques useful but less satisfactory than aging by established techniques. The need for more reliable size-based assessment techniques led to the development of a size-based cohort analysis (Jones 1981) which was applied, following a suggestion by A.S.D. Farmer, to Kuwaiti shrimp data by Jones and van Zalinge (1981).

The technique of Jones and van Zalinge (1981) is based on the availability of L_{∞} and K estimates. The appropriate von Bertalanffy parameters derived for *P. semisulcatus* and *M. affinis* by modal progression analysis were used in each case. This procedure was applied to the dhow and industrial fisheries separately and combined, so the effects on the resulting estimates of including

the dhow landings could be determined. The authors had excluded the dhow landings from their analysis. Because of this, their technique provided mortality estimates that theoretically apply only to a very narrow size range (~ 5 mm CL). It was hoped to widen the size range and to obtain mortality estimates applicable to the whole of Kuwait's fisheries and shrimp species, not just the industrial landings and *P. semisulcatus*.

ELEFAN-BASED ASSESSMENTS

The size-based stock assessment techniques called ELEFAN 0, I and II have been described by Pauly and David (1980) and Pauly (1982b), while Pauly (1980) should be consulted for his method for estimating M from values of L_{∞} and K (the von Bertalanffy growth parameters) and the mean water temperature (T). The ELEFAN package of methods has been applied by Morgan (1985), Samuel and Morgan (1984) and Baddar and Morgan (1984) to Kuwaiti fish populations; Mathews (Part I, this vol.) discusses the results of stock assessment of long-lived organisms (5-45 years) separately. Estimates of natural mortality obtained by Pauly's equation are here referred to as M_p while the symbol M is reserved for natural mortality measured using the traditional techniques (e.g., through a plot of Z on effort, or using $Z - F$).

Results

GROWTH RATES AND LIFE CYCLES OF KUWAITI SHRIMP

Table 1 presents growth parameters obtained using model class progression analysis for *P. semisulcatus*, the dominant species in Kuwait's landings, and for *M. affinis* and *P. stylifera*, the two other commercially important species.

Fig. 1 depicts the growth of male and female *P. semisulcatus* in Kuwait. Two spawning seasons and two cohorts occur annually; spring recruits enter the fishery in May and June and spawn in

Table 1. Von Bertalanffy growth parameters for Kuwait's three main commercial species of shrimp for the 1978-1979 to 1981-1982 growth seasons, based on modal progression analysis. Adapted from El-Musa (1984) and Mathews (1982a).

	1978-1979		1979-1980		1980-1981		1981-1982		Mean values	
	L_{∞} (CL, mm)	K (year $^{-1}$)								
<i>M. affinis</i>										
Spring										
♂	—	—	—	—	29.9	0.4214	31.3	0.3022	28.6	0.34
♀	30.9	0.2792	34.4	0.2295	34.0	0.5312	27.3	0.3680	32.8	0.38
Summer-winter										
♂	24.2	0.9774	—	—	27.5	0.3861	—	—	25.8	0.68
♀	30.0	0.4197	—	—	25.2	0.9774	—	—	27.7	0.70
<i>P. stylifera</i>										
Spring										
♂	—	—	—	—	25.0	0.1493	29.0	—	—	—
♀	—	—	—	—	30.1	0.1709	29.9	0.1485	30.0	0.16
Summer-winter										
♂	33.3	0.2941	34.4	0.1670	39.8	0.2185	37.6	0.1685	36.3	0.20
♀	54.5	0.1180	58.1	0.0932	—	—	72.7	0.0892	56.3	0.10
<i>P. semisulcatus</i>										
Spring										
♂	31.2	0.2300	33.4	0.9054	—	—	—	—	32.6	0.57
♀	39.7	0.2282	—	—	—	—	—	—	39.7	0.23

September to October, whereas autumn recruits enter the fishery in September, October or November and spawn the following April or May. This life cycle is discussed fully in Mathews (1982b) and FAO (1982). Data from 1981 to 1984 confirm these results.

Fig. 2 shows a schematic life cycle for *M. affinis* based on available data on the life, growth and spawning of this species. *M. affinis* females mature sexually at sizes greater than ~ 20 mm carapace length (CL) and recruit from May to August at 18-20 mm. Selection experiments using a cod end cover of full mesh size ~ 5 mm provided large samples of small shrimp during April, May and June in 1980 and 1981; large numbers of *M. affinis* were taken on the main fishing grounds at Rixa, but all were at or above the 18-20 mm (CL) size range. Other species, e.g., *P. stylifera*, were taken at sizes as small as ~ 5 mm (CL) in these nets and the absence of *M. affinis* below 18 mm CL

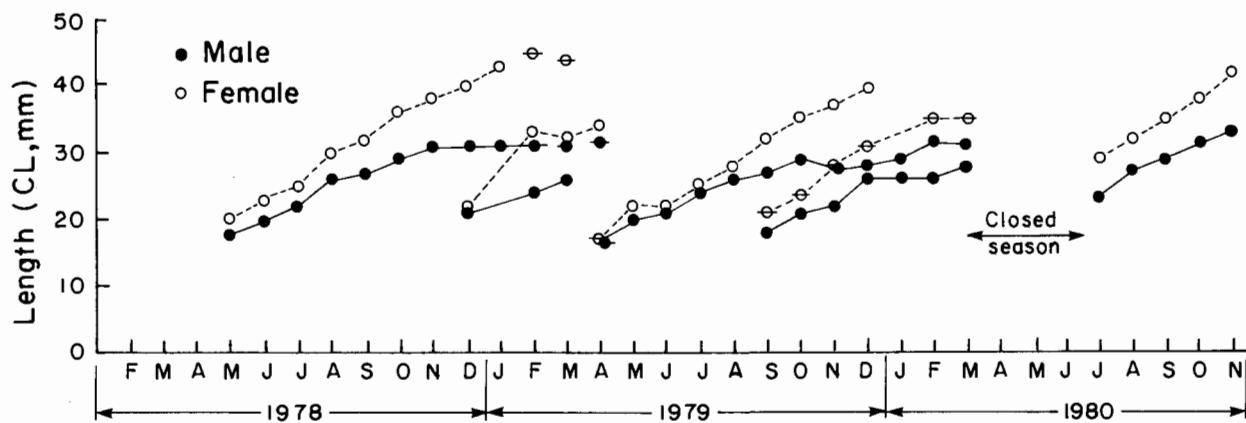


Fig. 1. Different cohorts for male and female *P. semisulcatus* in Kuwait's fisheries (data from the artisanal and industrial fisheries combined).

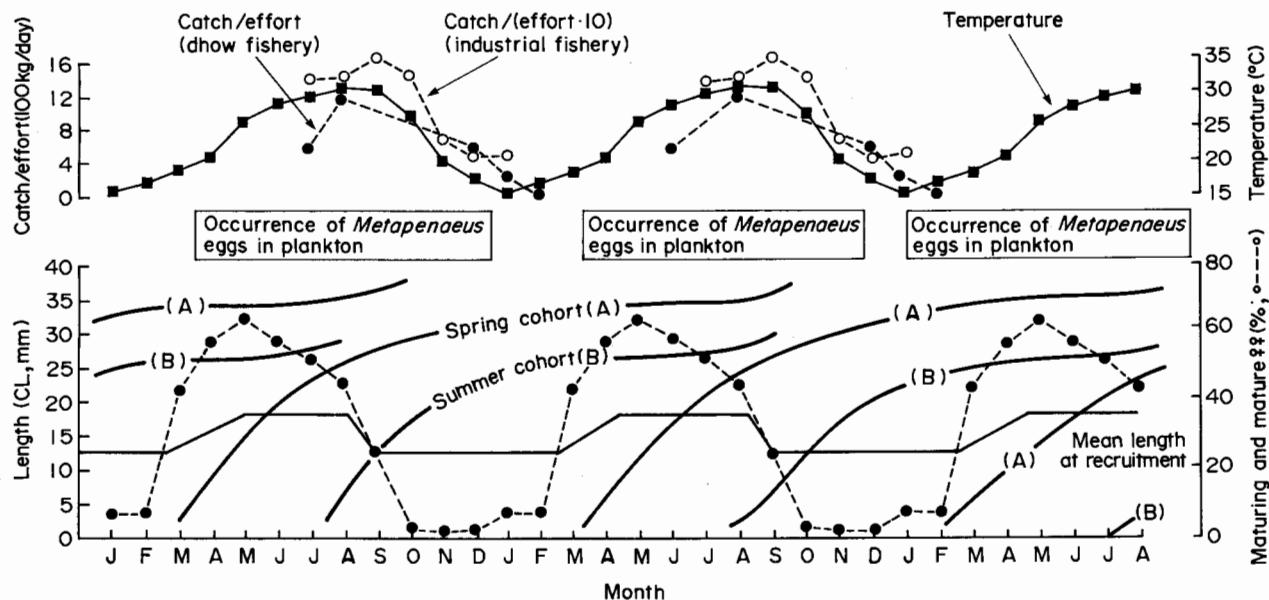


Fig. 2. Summary of information (growth, spawning, recruitment) on the life cycle of *Metapenaeus affinis* in Kuwait waters (see text).

suggests that it recruited to the fishing grounds in Rixa at or above this size. Much smaller *M. affinis* have been taken in trawls in the winter months, especially in the discards and landings of the artisanal boats. This suggests that small shrimp occur on the Rixa fishing grounds from around October to February.

The spring cohort recruits to the Rixa fishing grounds around May and June at a size of ~18 mm CL and grows steadily until the shrimp leave the fishery during the following May or June; they stay on the fishing grounds for ~12 months, during which they are fished continuously. Even during the closed season, fishing mortality occurs because shrimp taken while trawling for fin fish are discarded at sea. There has also been a limited fishery on *M. affinis* during the closed season since 1983.

A second cohort is recruited to the population from August to October, but at a much smaller size (~13-15 mm CL); the autumn recruiting cohort stays in the fishery until the following June or July, and then leaves the fishing grounds (or is fished out of existence at the same time as the spring recruiting cohort). Abdul-Ghaffar (1984) showed that *M. affinis* recruitment was very low in July while in August recruitment to the smaller size groups occurred. In September and October, very heavy recruitment of small *M. affinis* occurred, but the larger *M. affinis* were found only in small numbers. *M. affinis* spawned from April to October, but showed major peaks in May-June and September-October according to Al-Attar (1982, 1984b), who studied the incidence of eggs of *Metapenaeus* spp. in the plankton. Although eggs of *M. affinis* cannot be distinguished from those of *M. stebbingi*, the only other species of the genus taken in commercial landings, it is likely that *M. affinis* (the major species found in the landings belonging to this genus) dominates the *Metapenaeus* spp. egg plankton, which may, therefore, be taken as an index of spawning in *M. affinis*. Peak spawning periods coincide with the two recruitment periods, and probably also with the presence of sexually mature cohorts in the population, especially during the spring recruitment period. Peak spawning periods also coincide with temperature levels at ~22-28°C; the onset of spawning in spring and the end of spawning in autumn occur at ~20°C.

Entry to the fishery occurs over the size range 11-14 mm CL, because 11 mm is the 50% retention length for dhow shrimp nets (full mesh size 32-35 mm) and ~14 mm is that for industrial mesh (44-46 mm, full mesh size; El-Musa 1982). Al-Hossaini et al. (1984) suggested that recruitment occurred over a wide range of sizes.

Because *M. affinis* of 18-20 mm CL recruit to the fishing grounds in spring (May to June) at the same time as older and larger *M. affinis* spawn, spring recruits are almost certainly the product of autumn spawning. It is also probable that autumn recruits are the product of spring spawning, although the possibility cannot be excluded that an early peak in autumn spawning (perhaps during years with cooler summers) might provide some late autumn recruits. Entry of spring recruits in large numbers to the fishing grounds appears to occur from outside Kuwait waters, probably to the north of the Rixa and Bubiyan areas. Mathews (1982b) noted a significant inverse correlation between numbers and hours of fishing of *M. affinis* in Rixa and Bubiyan and in Kuwait Bay. He suggested that small shrimp are recruited from Bubiyan to Rixa, but *M. affinis* tended to avoid Kuwait Bay. These data are in agreement with the migration of small *M. affinis* towards Kuwait waters, perhaps from the Shatt Al-Arab area and the northeastern part of the Gulf. Analysis of a larger data series than is presently available is necessary to confirm this observation.

The description of the life cycle of *M. affinis* given here is similar to that given by FAO (1982), based largely on data from Kuwait and Saudi Arabia. Perez et al. (1984a, 1984b) studied the life cycle and population dynamics of Cuban pink and white shrimp (*Penaeus notialis* and *P. schmitti*). They reported the same type of life cycle as found in Kuwait's shrimp, although recruitment and spawning tended to be more continuous than in Kuwait. They also noted that for both Cuban species, spring recruitment was stronger than autumn recruitment; the same observation was made for Kuwaiti shrimp by Mohammed et al. (1981a), Mathews (1982c) and FAO (1982). Sivasubramaniam and Ibrahim (1982) reported essentially the same life cycle for *P. semisulcatus* in Qatar as was recorded for Kuwait.

It is likely that the type of life cycle found in Kuwaiti shrimp is general in penaeids: two distinct cohorts, one with spring recruitment and autumn spawning (the stronger cohort), the other with autumn recruitment and spring spawning. More work is needed to further elucidate this life cycle, which has the paradoxical feature that the major spawning and recruitment episodes occur in spring, while the major (spring) recruitment results from the minor (autumn) spawning (FAO 1982).

MORTALITY

Estimation by the Technique of Jones and van Zalinge (1981)

The method assumes that the fishery is based on a single unit stock. For application to Kuwait's shrimp stocks, Jones and van Zalinge (1981) assumed that over 90% of the industrial landings were *P. semisulcatus* during the period of the fishery (principally from July to January). Dhow landings were excluded from their analyses because these were known to contain substantial quantities of other species. The data used by Jones and van Zalinge were obtained from the industrial landings in each commercial size grade (tail counts), which were converted to numbers landed by length groups.

Van Zalinge et al. (1981) subsequently extended this method to cover the fishery from 1972 to 1979 and obtained estimates of Z for each year and each sex. In July 1981, a new sampling system was introduced (Bedford 1982) allowing a full breakdown by species, size, sex and gonadal phase of the Shuaiba industrial fishery. Using these data, Abdul-Ghaffar (1984) showed that approximately 50% of the industrial landings were *M. affinis* in 1981 and that the unit stock assumption of van Zalinge et al. (1981) was violated. In subsequent years, a much smaller portion of the industrial landings were composed of *M. affinis* because of a change in fishing strategy. The opening of a separate fishery from April to June on the *M. affinis* stock caused the industrial fishery to confine effort to the *P. semisulcatus* stock during the remainder of the year. Nevertheless, it is clear that *M. affinis* was an important component of Kuwaiti landings in the 1970s, as it was in 1980 and 1981.

Because of these difficulties, the method of Jones and van Zalinge (1981) was applied to the dhow and industrial fisheries separately. Mean values of L_{∞} and K used for three species are shown in Table 1. Abdul-Ghaffar (1984) provides the necessary data on the number of each species and sex in the industrial fishery. These data were used to produce estimates of Z/K from the slope of logarithmic plots of numbers landed per size grade (ordinate) and $L_{\infty} - L_t$ (abscissa) for each sex and species in each fishery, and for fisheries and species combined, according to the technique of Jones and van Zalinge (1981). Table 2 shows typical results, while full results for all years and cohorts are given by Mathews and Al-Hossaini (1984). This procedure allowed the following:

- Estimating mortality for each fishery separately to test whether mortality rates differed between the fisheries and whether bias was likely if data for only one fishery were used (e.g., the industrial fishery only, van Zalinge et al. 1981). This also allowed mortality rates to be determined for each sex separately.
- Expanding the size range used in mortality estimates. Jones and van Zalinge (1981) found that their estimates of Z were obtained for a very narrow size range, i.e., 28-23 mm carapace length.
- Testing the assumption that all shrimp landed belonging to one species tends to cause bias in the estimates of Z .

Fig. 3 above shows a typical set of curves for *P. semisulcatus* from the dhow fishery with a linear middle section from which the slope, a direct estimate of Z/K , may be obtained for values of $L_{\infty} - L_t$ ranging from 26.5 to 18.4 mm, corresponding to a carapace length of 29.8-37.9 mm (see Table 2, female *P. semisulcatus*). The upper parts of the curves have a declining slope because small shrimp are underrepresented in the landings due to selectivity or because recruitment of small shrimp to the fishery is incomplete in these size classes. However, recruitment to the fishery is probably completed within the 20-21 mm carapace length size range.

The same figure shows the data for *P. semisulcatus* for the industrial fishery and for both fisheries combined; the rectilinear curves generated are similar (see also Table 3). The curves are not smooth but a rectilinear section needed to estimate Z/K can be distinguished in each case.

Data for *M. affinis* in the two fisheries are shown in Fig. 3 (central part). Only three or four points are available for females and two for males (not shown) because the values of L_{∞} are much lower, and relatively coarse size grades are used. The technique of Jones and van Zalinge (1981), depending on the use of commercial size grades, is, therefore, not as useful for *M. affinis* and other small shrimp such as *P. stylifera*, as it is for *P. semisulcatus*.

Results for all species combined, i.e., combining total landings of *P. semisulcatus*, *M. affinis* and *P. stylifera*, were obtained by assuming that the growth parameters of *P. semisulcatus* could

Table 2. Estimates of Z for *P. semisulcatus* from the dhow fishery for 1981-1982 using the method of Jones and van Zalinge (1981).

Males ($L_{\infty} = 36.3$ mm; $K = 2.4$ per year)				
Market category	Carapace length (mm)	Numbers landed	Cumulative landings (no.)	$L_{\infty} - L$
15	43.3	12,762	12,762	—
16/20	37.9	39,307	52,069	—
21/25	34.4	58,335	110,404	1.86
26/30	31.8	400,624	511,028	4.46
31/35	29.8	896,127	1,407,155	6.46
36/40	28.2	692,629	2,997,784	8.06
41/50	26.3	1,511,043	3,610,827	9.96
51/60	24.3	1,167,588	4,778,415	11.96
61/70	22.8	703,075	5,481,490	13.46
71/90	21.1	648,954	6,130,444	15.16

Females ($L_{\infty} = 56.3$ mm; $K = 1.2$ per year)				
Market category	Carapace length (mm)	Numbers landed	Cumulative landings (no.)	$L_{\infty} - L$
15	43.3	183,833	183,833	12.99
16/20	37.9	563,675	747,508	18.39
21/25	34.4	838,058	1,585,566	21.89
26/30	31.8	1,231,593	2,817,159	24.49
31/35	29.8	1,878,212	4,695,371	26.49
36/40	28.2	439,274	5,134,645	28.09
41/50	26.3	775,889	5,910,534	29.99
51/60	24.3	521,697	6,432,231	31.99
61/70	22.8	344,878	6,777,109	33.49
71/90	21.1	589,139	7,356,248	35.19

be applied to such a mixed population. The curves are similar to those for *P. semisulcatus* but are dissimilar to those for *M. affinis* (see Fig. 3).

For *P. semisulcatus* estimated values of Z are close for both sexes regardless of the fishery from which they were obtained or whether the data from the two fisheries were combined (see Table 3). However, values for females are much lower than for males. Values for *M. affinis* are generally much higher for both sexes. For this species, there were no strong differences between the sexes, and estimates from the different fisheries (single or combined) were similar.

Estimates for the different fisheries using all species combined were more variable but were generally comparable to those for *P. semisulcatus* when *P. semisulcatus* growth parameters were used. Combining fisheries did not seem to affect the estimates of Z.

This contrasts with the results obtained when the data for all species combined were analyzed assuming *M. affinis* growth parameters. This gave estimates of Z of 1.9 for females and 57.9 for males, i.e., very much larger than when the same data were analyzed using *P. semisulcatus* growth parameters. It thus seems that the method of Jones and van Zalinge (1981) is particularly sensitive to variations in the values of L_{∞} and K. It follows that estimates of Z based on application of growth parameter estimates from one species to a fishery based on two or more species with different growth parameter estimates may be erroneous. The estimates of van Zalinge et al. (1981) and of van Zalinge (1984) were almost certainly based on mixed landings of *P. semisulcatus*, *M. affinis* and *P. stylifera*,

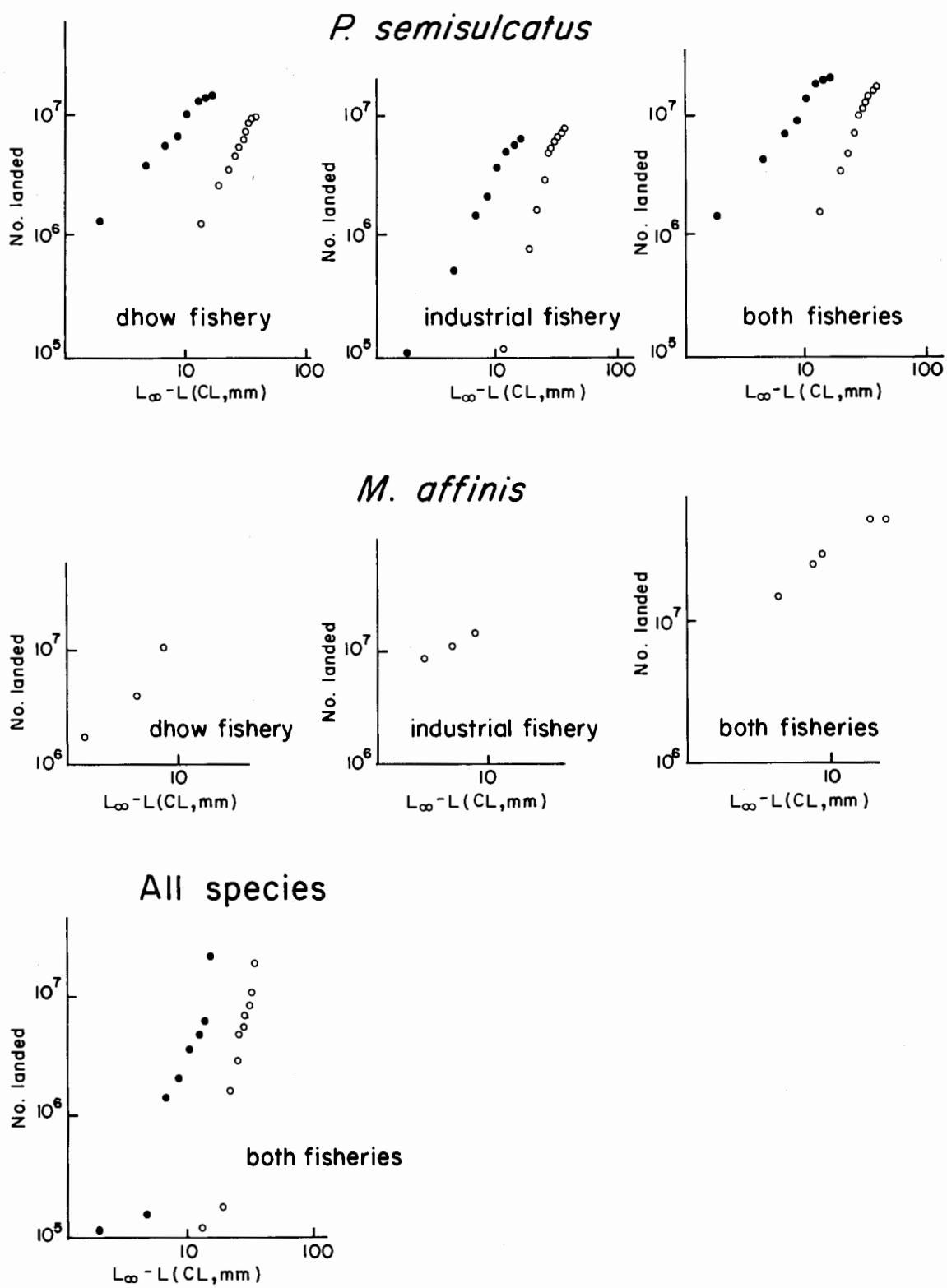


Fig. 3. Plots obtained through application of the method of Jones and Van Zalinge (1981) for various species of Kuwait shrimp, and two different fisheries (the plot for all species used growth parameters derived from data on *P. semisulcatus*); ● = ♂; ○ = ♀.

Table 3. Mortality estimates obtained by applying the method of Jones and van Zalinge (1981) to Kuwait's shrimp population, for the 1981-1982 fishing season (see also Fig. 3).

Fishery	Species	Sex	Size used range (mm, CL)	L_{∞} (CL, mm)	K (year $^{-1}$)	Z (year $^{-1}$) ^c
Dhow	<i>P. semisulcatus</i>	♀	24.3 — 34.4	56.3	1.2	8.71
Industrial	<i>P. semisulcatus</i>	♀	24.3 — 34.4	56.3	1.2	8.01
Combined	<i>P. semisulcatus</i>	♀	34.4 — 43.3	56.3	1.2	8.19
Dhow	<i>P. semisulcatus</i>	♂	24.3 — 29.8	36.3	2.4	21.22
Industrial	<i>P. semisulcatus</i>	♂	24.3 — 29.8	36.3	2.4	22.25
Combined	<i>P. semisulcatus</i>	♂	28.2 — 34.4	36.3	2.4	20.79
Dhow	<i>M. affinis</i>	♀	25.1 — 31.3	32.8	4.56	65.21
Industrial	<i>M. affinis</i>	♀	25.1 — 30.0	32.8	4.56	60.38
Combined	<i>M. affinis</i>	♀	25.1 — 28.4	32.8	4.56	61.82
Dhow	<i>M. affinis</i>	♂	25.1 — 28.0	28.5	4.08	54.51
Industrial	<i>M. affinis</i>	♂	25.1 — 28.0	28.5	4.08	60.38
Combined	<i>M. affinis</i>	♂	25.1 — 28.0	28.5	4.08	56.23
Dhow	All species combined ^a	♀	24.3 — 29.8	56.3	1.2	10.58
Dhow	All species combined ^a	♂	22.8 — 29.8	36.3	2.4	16.78
Industrial	All species combined ^a	♀	29.2 — 34.4	56.3	1.2	10.20
Industrial	All species combined ^a	♂	21.1 — 28.2	36.3	2.4	20.47
Dhow + industrial	All species combined ^a	♀	28.2 — 34.4	56.3	1.2	10.56
Dhow + industrial	All species combined ^a	♂	22.8 — 28.2	36.3	2.4	23.64
Dhow + industrial	All species combined ^b	♂	25.1 — 31.3	32.8	4.56	41.91
Dhow + industrial	All species combined ^b	♂	25.1 — 28.4	28.6	4.08	57.94

^aUsing growth parameters for *P. semisulcatus*.

^bUsing growth parameters for *M. affinis*.

^cNote that these values all appear to largely overestimate total mortality (see text).

with the growth parameters from *P. semisulcatus* being applied to a fishery frequently composed of less than 50% of this species. Their results were thus probably erroneous. While the method of Jones and van Zalinge (1981) itself is not in question, it is evident that it can only be applied to a single species population, or to a population consisting of species with similar growth parameters. This is not the case for Kuwait's shrimp.

The very high mortality rates obtained by means of this technique, even for data pertaining to one species only (see Table 3), are also suspiciously high and suggest that some of the basic assumptions may be violated.

Estimates of Mortality from Survival Curves

Survival curves were obtained for male and female *P. semisulcatus* from spring cohorts and for those autumn cohorts for which sufficient data are available (Fig. 4). Good fits were obtained for the spring cohorts although very low catch-per-effort values were found for some cohorts in the winter months. Such data points were excluded from the analysis because, at very low temperatures, very low catch rates were recorded by Mathews and Samuel (1982), who noted a significant relationship between catch per effort and temperature. For the autumn cohorts, only male shrimp in 1979-1980 give a useful survival curve.

For *M. affinis*, acceptable survival curves could only be obtained for the 1978-1979 cohorts; other cohorts showed a clear pattern of increasing catch per effort of both sexes with time, suggesting strong and increasing recruitment from July to at least November. A similar, although less marked, pattern was observed for the two cohorts of *P. stylifera* for which data could be obtained (Fig. 5). Table 4 presents the estimates of total mortality based on Figs. 4 and 5.

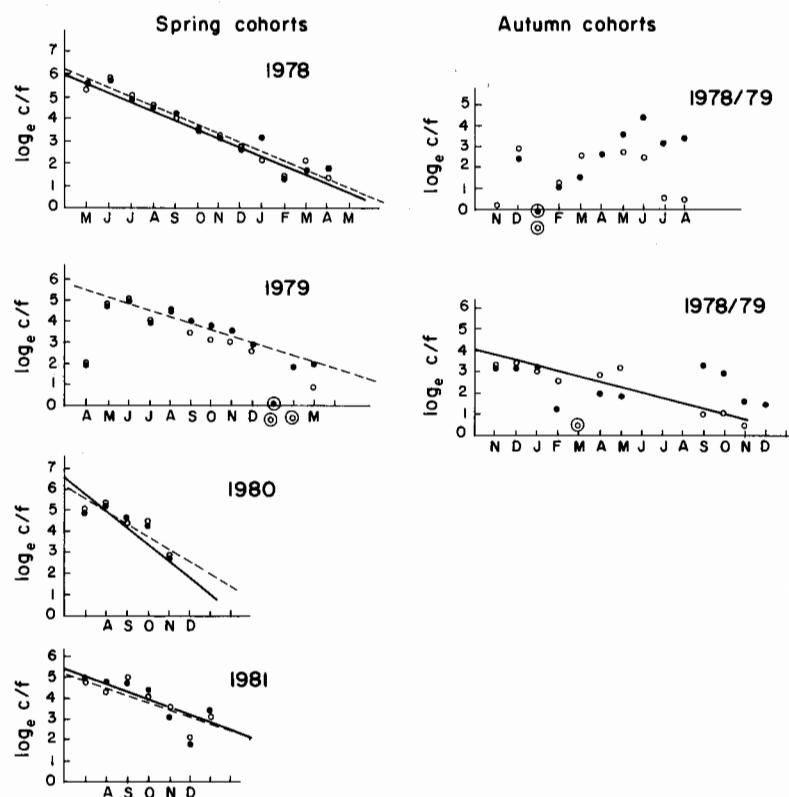


Fig. 4. Temporal change in relative abundance of *P. semisulcatus*, as used to estimate their total mortality (see text and Table 4).

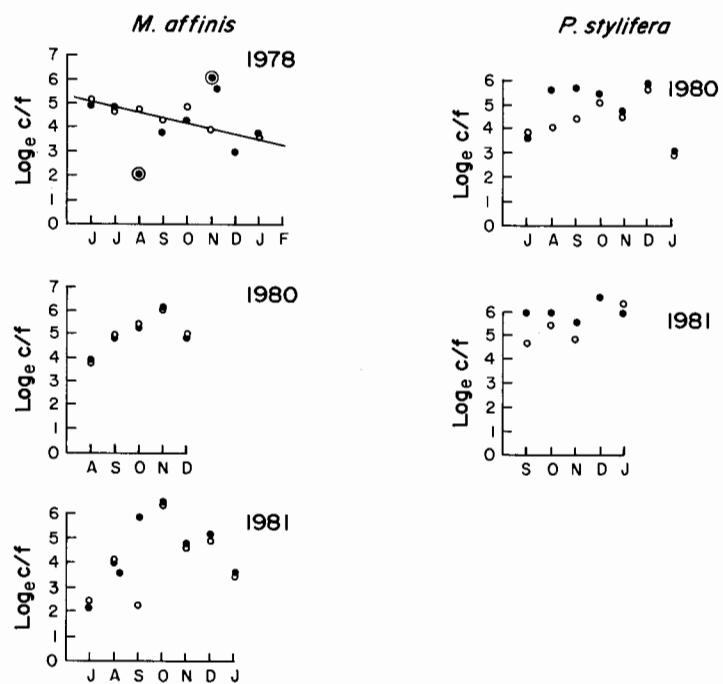


Fig. 5. Time series of catch/effort for male (●) and female (○) *Metapenaeus affinis* and *Parapenaeopsis stylifera*. Only one of these plots (*M. affinis* 1978) could be interpreted as survival curve (circled points excluded from regression).

Table 4. Total mortality estimates in two species of Kuwaiti shrimp, based on the survival curves in Figs 4 and 5.

	Z (year ⁻¹)	df	Significance level
<i>P. semisulcatus</i>			
Spring 1978			
♂	5.14	10	P < 0.001
♀	4.84	10	P < 0.001
Spring 1979			
♂	4.76	7	P < 0.001
♀	3.56	8	P < 0.00
Spring 1980			
♂	6.53	3	0.10 > P > 0.05
♀	6.33	3	0.10 > P > 0.05
Spring 1981			
♂	4.61	5	0.05 > P > 0.02
♀	4.92	5	0.05 > P > 0.02
Autumn 1980			
♂	2.88	7	P < 0.001
<i>M. affinis</i>			
Spring 1978			
♂	2.94	6	0.01 > P > 0.001
♀	2.71	4	0.05 > P > 0.01

Estimates of Natural Mortality (M_p) Based on Pauly's Empirical Equation

Estimates of M_p for the Kuwait *P. semisulcatus* stocks were provided by FAO (1982), which suggested that $M = 3$ for both sexes of *P. semisulcatus*, based on the equation

$$\log_{10} M_p = -0.0066 - 0.279 \log_{10} L_\infty + 0.6543 \log_{10} K + 0.4634 \log_{10} T \quad \dots 1$$

(Pauly 1980) where, M_p = natural mortality estimate, L_∞ (total length, cm) and K are parameters of the von Bertalanffy growth equation, and T is the mean water temperature (in °C). These estimates, however, were based on van Zalinge et al. (1981) who used carapace length instead of total length as the units of measurement for L_∞ . Values of M_p for all Kuwaiti species and cohorts for which growth parameter estimates based on modal progression analyses were available, were computed. Values obtained varied from 5.0-16.0 for *P. semisulcatus*, 13.0-20.0 for *M. affinis* and 7.0-8.0 for *P. stylifera*, for different sexes and cohorts (1978-1981). These were comparable to, or in excess of, the values of total mortality obtained from cohort analysis or survival curves for *P. semisulcatus*. The estimates of M_p were also higher than any values of M reported in the literature for shrimp populations (Mathews 1981). This can probably be attributed to the use of parameters which did not account for the strong growth oscillations depicted by *P. semisulcatus* in Kuwait, and which lead to overestimates of K , i.e., to values of K much higher than would have been obtained, had seasonal growth oscillations been considered explicitly. Table 5 summarizes mortality (M_p) estimates based on the growth parameters using equation (1). Estimates of M_p are substantially lower than estimates of Z (see below), and are much lower than estimates of M obtained

using dates based on modal progression analysis (i.e., growth estimation which exclude temperature and other environmental effects on growth). This emphasizes the need to estimate M_p only when reliable growth parameter estimates are available (see text below and Table 6).

Table 5. Mean growth parameter estimates for the three species of Kuwait shrimps, as estimated using ELEFAN I; the table also presents estimates of natural mortality (M_p) based on these parameters, a temperature of 23.2°C and the empirical equation of Pauly (1980) (see text).

Parameter estimates (95% conf. interval)	<i>P. semisulcatus</i>		<i>M. affinis</i>		<i>P. stylifera</i>	
	♂	♀	♂	♀	♂	♀
L_∞ (mm)	48.2 (± 2.04)	53.2 (± 1.83)	37.7 (± 1.03)	48.0 (± 1.10)	31.5 (± 3.39)	35.0 (± 1.79)
K (year^{-1})	0.96 (± 0.05)	1.07 (± 0.11)	1.09 (± 0.21)	1.22 (± 0.25)	1.23 (± 0.10)	1.23 (± 0.36)
WP (year)	0.95 (± 0.10)	0.92 (± 0.13)	0.96 (± 0.19)	1.02 (± 0.10)	0.87 (± 0.12)	1.04 (± 0.10)
C	0.83 (± 0.20)	0.92 (± 0.20)	1.00 (± 0.22)	0.99 (± 0.27)	0.98 (± 0.08)	0.98 (± 0.12)
M_p	1.8	1.8	2.1	2.0	2.3	2.7

Table 6. Comparison of growth parameters and natural mortality estimates in uninfested and infested male *P. semisulcatus* in Kuwait (based on data for 1979-1984), as obtained from using ELEFAN I and the empirical equation of Pauly (1980), respectively.^a

	CL_∞ (mm) (95% conf. interval)	K (year^{-1}) (95% conf. interval)	M_p (year^{-1}) (95% conf. interval)
Uninfested males only ^b	40.0 (± 1.3)	0.99 (± 0.04)	1.8 (± 0.05)
Infested + uninfested males	48.2 (± 2.0)	0.96 (± 0.05)	1.8 (± 0.07)

^aWith $T = 23.2^\circ\text{C}$.

^bObtained by deletion of infested males and of a few extremely large males thought to have lost their parasites and have moulted (see text).

GROWTH PARAMETER ESTIMATES BASED ON ELEFAN I

The following data were analyzed:

- Size-frequency distributions from monthly samples for male and female *P. semisulcatus*, *M. affinis* and *P. stylifera* for the dhow fishery from 1979 to 1984.
- Size-frequency distributions obtained from monthly cruises on R/V Oloum I in 1979; these samples were considered to be similar to industrial landings because R/V Oloum I is similar to the industrial vessels fishing in Kuwait and covered the same fishing grounds, at least for *P. semisulcatus*.

- Size-frequency distributions for industrial landings obtained by sampling at the Shuaiba shrimp processing plant. Weight frequencies and species compositions were obtained for several size packs in each commercial size grade. Weight of the monthly landings was provided by the fishing company for each size grade. Combining these data with those on shrimp morphometrics provided by Farmer (1980), it was possible to derive monthly length frequencies of the shrimp landed in the industrial fishery. These data, obtained each month, were also analyzed for each species and sex from July 1981 to 1983.

A total of 45 data sets (covering 6 years for 3 species and 2 sexes) were analyzed. For male *P. semisulcatus*, extra analyses were conducted because growth was probably affected by the presence of a parasite (see below).

In these analyses, the values of ESP/ASP varied from 0.167 to 0.960, but were often over 0.4. Although growth parameter estimates varied between data sets for the same species and sex, the results show that the estimates of L_{∞} , K, WP and C vary by a relatively small amount. For L_{∞} , the 95% confidence limits vary from ~2-5% of the mean (Table 5). The mean values of the various growth parameters (for each species and sex) obtained over the six-year study period provide good fits to all the data and can be used to describe growth for any year with acceptable accuracy. This implies that the growth parameters of the shrimp populations investigated here may vary little from year to year.

Table 5 presents mean growth parameters for each sex and species studied, on the assumption that annual values for each parameter vary randomly from year to year, although the randomness may be apparent only and may indeed conceal the effects of environmentally driven variations. The estimates of K and C were more variable from year to year than those of other parameters. It is likely that K and C vary from year to year with key environmental variables, probably temperature.

INFLUENCE OF PARASITES ON GROWTH ESTIMATES OBTAINED BY ELEFAN I

P. semisulcatus individuals of both sexes are often infested by *Epipenaeon elegans* (Chopra). This bopyrid parasite occurs in the gill chamber of the shrimp and appears to prevent females from spawning successfully. Infested females do not reach gonad stages III to IV and tend to reach much larger sizes than uninfested males. It is possible that infested males are physiologically castrated and grow as if they were females (Rheinhard 1950). Mohammed et al. (1981b) reported that infested male *P. semisulcatus* in Kuwait reached a maximum carapace length of 44 mm. Branford (1980) reported that male *P. semisulcatus* infested by *E. elegans* also grew larger than uninfested males: 38.3 mm for infested males and ~30 mm for uninfested males. El-Musa et al. (1981) described the effects of *E. elegans* on *P. semisulcatus* in Kuwait.

Analysis of size-frequency data by sex and degree of infestation showed that infested males are more common in large size groups (15-100% shrimp > 30 mm in 1979) than in small sizes (1.1-4.2% for shrimp from 13-30 mm). Large infested male *P. semisulcatus* were sometimes accompanied in catches on R/V Oloum I by occasional large males or females that had shed their parasite, but which had the typically distended and swollen gill chamber that accompanies infestation by a parasite. It is likely that the very large uninfested male *P. semisulcatus* were infested previously but had lost or shed their parasite, perhaps because of completion of its life cycle. When such shrimp moult, all signs of infestation may disappear, except for the fact that their previous growth history was influenced by the parasite.

Size-frequency data for male *P. semisulcatus* were analyzed from 1979 to 1984. Data sets including all male *P. semisulcatus*, i.e., whether carrying a parasite or not, and definitely uninfested males were contrasted. Table 6 presents the results; as might be seen, the mean values for K and M_p were not very different but the mean values of L_{∞} were.

The versions of ELEFAN I and II used in the present analysis were based on the assumption that $L_{\infty} > L_{\max}$, the largest individual in the sample. It seems likely that adjustment of these programs such that L_{\max} may be less than L_{∞} would eliminate the types of bias discussed here (see Hampton and Majkowski, Part I, this vol. and Pauly, Part I, this vol.).

MORTALITY ESTIMATES

Table 7 presents estimates of Z , F ($= Z - M_p$) and M_p values for spring cohorts of the 3 species of shrimps investigated here, based on the length-converted catch curve routines of the empirical equation of Pauly (1980), both of which are part of ELEFAN II. These values complement the values of M_p in Table 5. As might be seen, these mortality estimates appear throughout far more realistic than those obtained using alternative techniques, particularly tagging which overestimated M by one order of magnitude.

Table 7. Selected estimates of mortality for Kuwaiti shrimp (data for spring cohorts only; all mortalities are on a per year basis).

		Tagging	Size-based cohort analysis	Survival curves (with range)	ELEFAN II (with 95% confidence limits)
<i>P. semisulcatus</i>					
Z	♂	28.5	12.5	5.26 (4.61 — 5.63)	4.80 ± 1.18
	♀	31.5	4.8	4.91 (3.55 — 6.33)	2.90 ± 0.42
F	♂	2.42	—	—	3.14 ($= Z - M_p$)
	♀	3.49	—	—	1.13 ($= Z - M_p$)
M or M_p	♂	26.1 ($= Z - F$)	—	—	1.66 ± 0.08 ($= M_p$)
	♀	27.6 ($= Z - F$)	—	—	1.77 ± 0.12 ($= M_p$)
<i>M. affinis</i>					
Z	♂	—	—	—	5.42 ± 1.11
	♀	—	—	—	4.36 ± 0.98
F	♂	—	—	—	3.36 ($= Z - M_p$)
	♀	—	—	—	2.32 ($= Z - M_p$)
M_p	♂	—	—	—	2.06 ± 0.30
	♀	—	—	—	2.04 ± 0.26
<i>P. stylifera</i>					
Z	♂	—	—	—	7.04 ± 2.69
	♀	—	—	—	4.75 ± 1.44
F	♂	—	—	—	4.72 ($= Z - M_p$)
	♀	—	—	—	2.04 ($= Z - M_p$)
M_p	♂	—	—	—	2.32 ± 0.21
	♀	—	—	—	2.71 ± 1.46

However, changes in effort during the study period were considerable (Table 8). There is, for *P. semisulcatus*, a significant negative correlation between Z and F , and an increase of total effort. Data for the other two species are too scattered to show a trend but are consistent with a constant value of Z .

These results are incompatible with the assumptions of (a) proportionality of effort and fishing mortality and/or (b) constant natural mortality, and in fact, suggest that natural mortality of Kuwait shrimp stocks has drastically declined during the period covered here (see Discussion).

Table 8. Estimates of total mortality (Z) for males and females combined, based on the length-converted catch curve routine of ELEFAN II, and total effort in Kuwait's shrimp fisheries.

Year	<i>P. semisulcatus</i> ^a Z (year ⁻¹)	<i>M. affinis</i> Z (year ⁻¹)	<i>P. stylifera</i> Z (year ⁻¹)	Effort (standard days fishing)
1979	3.87	4.30	5.39	6,600
1980	4.12	4.92	7.08	6,500
1981	4.47	5.88	5.14	3,200
1982	3.79	6.11	7.04	8,900
1983	3.64	6.26	6.46	8,700
1984	3.32	3.91	4.23	11,600

^aThe correlation between Z and total effort is significant ($P < 0.001$) for *P. semisulcatus*, but the correlation has a negative sign (see text).

REPRODUCTION AND RECRUITMENT TO THE FISHERY

Fig. 6 summarizes the available data on seasonal changes of mean gonadal stages for the three major shrimp species in Kuwait.

Recruitment of *P. semisulcatus* to the fishery as estimated by changes in catch/effort probably peaks in July and August and decreases rapidly to low levels in October and November.

In *P. semisulcatus*, the proportion of mature female shrimp is generally very high from about November to February, and starts to decline in March. Assuming that the shrimp recover from spawning fairly quickly (i.e., in a few days), the rather high proportion of mature shrimp from

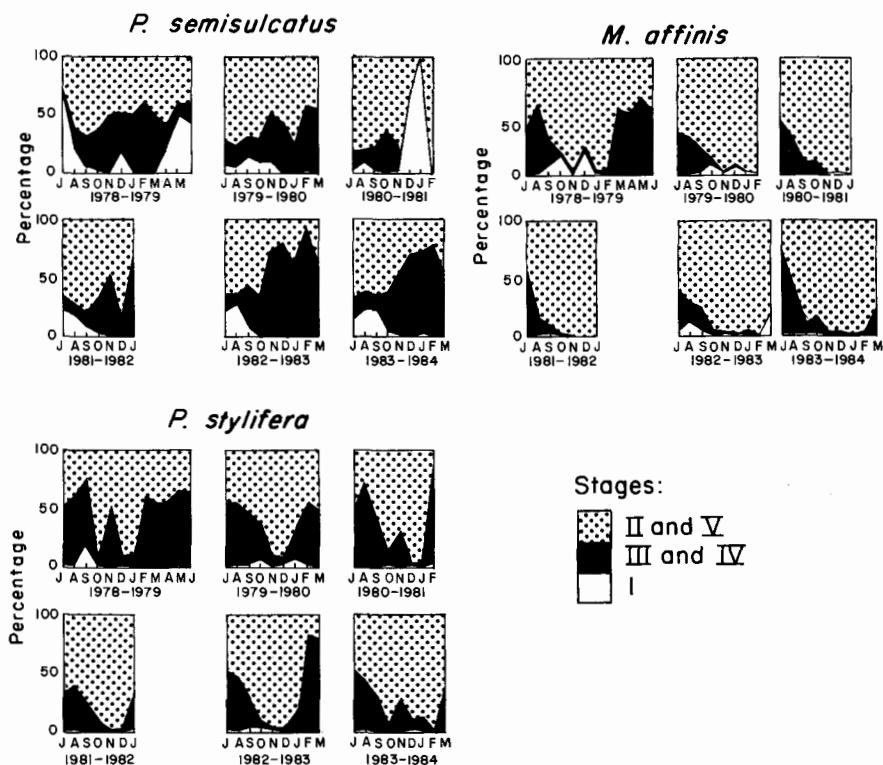


Fig. 6. Seasonal progression of gonadal maturation stages in the females of the three species of shrimp reported on here.

November to February is consistent with a peak spawning in December and January. However, Al-Attar (1982, 1984a) and FAO (1982) noted that there were two peaks of spawning in *P. semisulcatus*, one in spring (April to June) and one in autumn (September to November). It may be argued that mature shrimp are ready to spawn but do not do so until March, when increasing temperatures trigger spring spawning (Al-Attar and Ikenoue 1974). Thus, the high proportion of mature females in December would not indicate spawning. Spawning may increase in autumn, perhaps because summer temperatures are too high. Similar consideration should also apply to *M. affinis* and *P. stylifera*.

Fig. 7 shows the pattern of recruitment to the fishery as derived from catch-and-effort data and "recruitment patterns" obtained through ELEFAN II and which it must be emphasized pertain to "recruitment" at size zero (and, hence, should be related to the plots in Fig. 6).

Overall, the impression emerging from Figs. 6 and 7 is that use of ELEFAN II to estimate recruitment patterns may give a less reliable picture of recruitment than gonad maturation and/or planktonic studies (Al-Attar 1982, 1984b; FAO 1982) which showed clear peaks in numbers of benthic post larvae (~ 4 mm CL) in February, March and April of 1979, 1980 and 1981, with peak numbers of *P. semisulcatus* generally occurring in February. Winter growth rates are likely to be very low and these shrimp could be the results of autumn (September or October) spawning. Over-wintering of young shrimp could occur, with growth suspended from about December to February when temperatures are low ($\sim 14^{\circ}\text{C}$) compared to summer temperatures ($28\text{-}33^{\circ}\text{C}$). Shrimp recruiting to the fishery at different times could be of similar sizes and different ages, and this might confuse recruitment patterns estimated by means of ELEFAN II. Research designed to obtain direct estimates of recruitment at earlier and smaller stages than those studied here is under way (J. Bishop, KISR, pers. comm.) and is aimed at further clarification of recruitment patterns of

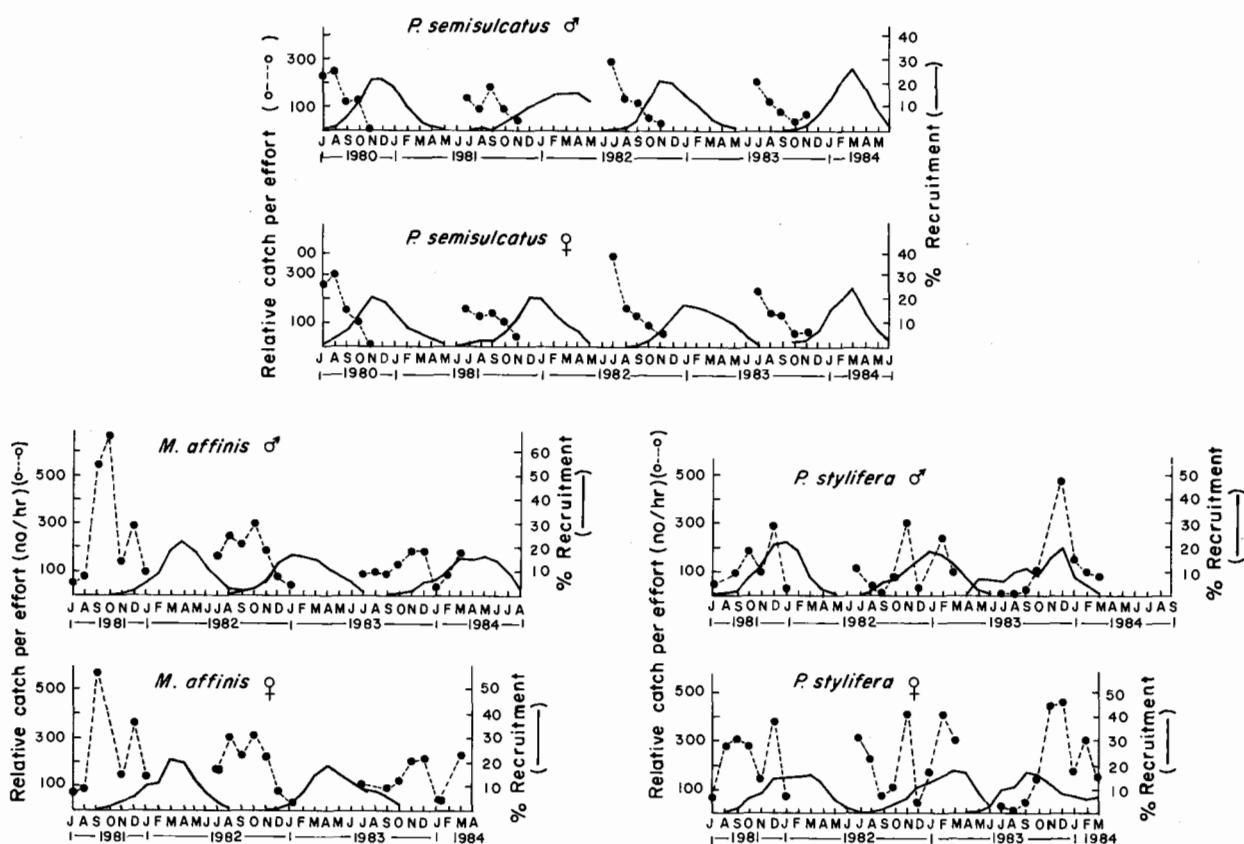


Fig. 7. Patterns of recruitment of penaeid shrimp in Kuwait, as inferred from relative abundance data and from ELEFAN II (see also text and Fig. 6).

shrimp in Kuwait. The estimates of size at entry into the dhow fishery (24-27 mm CL) for male and female *P. semisulcatus*; (about 20 mm for female *M. affinis* and about 16 mm for female *P. stylifera*) obtained using ELEFAN II through analysis of the left side of length-converted catch curves were, on the other hand, similar to those obtained through monthly biological sampling.

Discussion

The ELEFAN I and II programs produced the most reliable estimates of growth and mortality parameters for the three species that are important components of Kuwait shrimp stocks. The programs indeed produced the only reliable estimates for two of these species. No estimate for the comparative cost effectiveness of the ELEFAN approach was made but it was certainly no more expensive and was much more flexible than any of the other techniques applied.

While ELEFAN I appears to be a much more flexible and robust analytical procedure than any hitherto applied to the study of growth in shrimp populations, it is not always objective. The general growth patterns that emerged from this analysis were detected in spite of some anomalies. For instance, in several cases, the growth parameters giving the highest value of ESP/ASP in a given year were quite different from those obtained for other years. In such cases, a set of parameters closer to the others for that species and sex was used, even if the corresponding value of ESP/ASP was lower. On one occasion, a high value of ESP/ASP (~ 0.5) was obtained when $C = -2.25$; this implies marked negative growth or, in view of the sinusoid nature of the growth equation in ELEFAN I, a five-month period of zero growth. A much worse fit ($ESP/ASP \approx 0.20$) was, therefore, accepted because the population clearly did not display such a long growth stop. Occasionally, values of WP were anomalous compared with size-frequency distributions. In these instances also, a more biologically realistic value of WP was chosen. Thus, successful application of ELEFAN analysis for growth and mortality estimates presupposes and must rest upon a good understanding of the biology and life cycle of the species studied. On the other hand, if no such knowledge is available, ELEFAN I and II could be used to determine the growth and life cycle of the species studied on a provisional basis. Conclusions obtained would then need to be confirmed by biological observations.

It appears from our work that the growth parameters of shrimp populations remain similar between years and do not change as the level of exploitation increases. The growth pattern shown in the population may be viewed as the result of interactions between genetically fixed growth rates of individual shrimp and environmental factors. It appears to be largely independent of biomass, as influenced either by a fishery or by natural predators. However, infestations of many shrimp parasites may bias growth rate estimates, and care should be taken to exclude infested individuals from the analysis. This observation may be true for other short-lived animals. Although growth rates of longer-lived animals may change, the biases may be smaller if the parasite's life cycle is short.

Mathews (1981 and unpublished data) and Pauly (1982a, 1982b, 1984) showed that M probably decreases as fishing mortality increases in penaeid shrimp fisheries in the Gulf of California (*P. vannamei*) and in the Gulf of Thailand, respectively. The possibility that $M \propto 1/F$ in Kuwait also exists; Z could be approximately constant and F could increase at the same time as M decreases. This would explain why Z remained approximately constant where effort (and presumably F) increased by a factor of 4 (see also Munro 1980, Fig. 2). The equation of Pauly (1980) being tied to growth parameters which do not seem to vary much, tends to generate values of M_p that are fairly constant for a given stock—independently of any compensatory changes in its real value of M. For this reason, the usefulness of this equation will be limited in cases—such as in Kuwait—where massive changes of natural mortality may be expected.

The use of ELEFAN II to estimate recruitment at size zero (i.e., spawning) was compared with field data on the seasonal patterns of entry to the fishery, with maturation patterns and with published estimates of the numbers of individuals in various planktonic stages. It was found that minor errors in growth parameter estimates could lead to a major shift in the peak of the recruitment pattern along the time axis. Biological work (Al-Attar 1982, 1984a, 1984b) (Fig. 3) suggests two spawning seasons of equal strength for both *M. affinis* and *P. semisulcatus*, with a major spawning period in spring. The recruitment patterns obtained using ELEFAN II program often displayed only the major recruitment peak.

It is also possible, however, that recruitment of the autumn cohort occurs in a different area than that of the spring cohort (e.g., in deep waters of the central Gulf for *P. semisulcatus*), and this could explain the generally weaker recruitment in autumn. More work on the biological aspects of recruitment and spawning is needed before the use of recruitment patterns obtained from ELEFAN II can be fully appraised. For the time being, recruitment studies based on ELEFAN II should be used whenever possible, but results should be complementary to and not a substitute for biological studies of shrimps.

ELEFAN analysis, finally, may be used to control the quality of sampling systems. In Kuwait, shrimp are sampled in two ways:

1. Dhow boat landings are sampled regularly for size frequency, species composition and gonadal maturation (van Zalinge et al. 1981; Bedford 1982).
2. The shrimp populations at sea are sampled from the R/V Bahith.

A comparison of the results by ELEFAN II analysis shows that the dhow sampling system consistently provides samples covering a wide size range and probably provides generally more consistent results.

Acknowledgements

A preliminary version of this paper, which emphasized the biology and life cycles of penaeid shrimp in Kuwait, but did not include the ELEFAN-based analyses, was presented at the Symposium on the Scientific Basis for the Management of Penaeid Shrimp held at Key West, Florida, USA, 18-24 November 1981. The earlier paper was not submitted for publication because the authors felt that the questions addressed required a more satisfactory answer than could then be obtained. We wish to thank the hosts of the Key West Symposium, Drs. J. Gulland and B. Rothschild, for the opportunity to present our work at that meeting.

References

- Abdul-Ghaffar, A.R. 1984. The industrial fishery based at Shuaiba July 1981 to January 1982, p. 1-222. In C.P. Mathews (ed.) Proceedings of the Shrimp and Fin Fisheries Management Workshop, 19 June 1982, Vol. 1. Kuwait Institute for Scientific Research, Kuwait.
- Abramson, N.J. 1971. Computer programmes for fish stock assessment. FAO Fish. Tech. Pap. 101.
- Al-Attar, M.A. 1982. Shrimp spawning and nursery areas in Kuwait Bay, p. 1-342. In C.P. Mathews (ed.) Proceedings of the Shrimp Fisheries Management Workshop, 17 January 1981. Kuwait Institute for Scientific Research, Kuwait.
- Al-Attar, M.A. 1984a. Kuwait Bay, a nursery ground for Kuwait shrimp, p. 85-105. In C.P. Mathews (ed.) Proceedings of the Shrimp and Fin Fisheries Management Workshop, 19 June 1982, Vol. 1. Kuwait Institute for Scientific Research, Kuwait.
- Al-Attar, M.A. 1984b. Kuwait Bay, a nursery area for penaeid shrimp: II. *Metapenaeus affinis*, p. 207-222. In C.P. Mathews (ed.) Proceedings of the Shrimp and Fin Fisheries Management Workshop, 9-11 October 1983, Vol. 1. KISR 1366/MB-48. Kuwait Institute for Scientific Research, Kuwait.
- Al-Attar, M.A. and H. Ikenoue. 1974. Spawning season of shrimp *Penaeus semisulcatus* in the sea along the coast of Kuwait. Kuwait Institute for Scientific Research, Kuwait.
- Al-Hossaini, M. 1981. Some observation on recruitment and growth in Kuwait's shrimp fishery, p. 342. In C.P. Mathews (ed.) Proceedings of the Shrimp Fisheries Management Workshop, 17 January 1981. Kuwait Institute for Scientific Research, Kuwait.
- Al-Hossaini, M. 1982. Some observation on recruitment and growth in Kuwait's shrimp fishery, p. 92-105. In C.P. Mathews (ed.) Proceedings of the Shrimp Fisheries Management Workshop, 17 January 1981, Vol. 1. Kuwait Institute for Scientific Research, Kuwait.
- Al-Hossaini, M. 1984. The industrial fishery operating from the Sief area, p. 11-18. In C.P. Mathews (ed.) Proceedings of the Third Shrimp and Fin Fisheries Management Workshop, 4-5 December, Vol. I. Kuwait Institute for Scientific Research, Kuwait.

- Al-Hossaini, M., A.R. Abdul-Ghaffar and M. El-Musa. 1984. A preliminary study of gear selection in Kuwait's shrimp fishery and its effects on *P. semisulcatus* recruitment, p. 398-422. In C.P. Mathews (ed.) Proceedings of the 1983 Shrimp and Fin Fisheries Management Workshop, 9-11 October 1983, Vol. 2. KISR 1366/MB-48. Kuwait Institute for Scientific Research, Kuwait.
- Baddar, M. and G.R. Morgan. 1984. Stock assessment of hamoor, *Epinephelus tauvina*, in Kuwait waters, p. 499-522. In C.P. Mathews (ed.) Proceedings of the Shrimp and Fin Fisheries Management Workshop, 9-11 October 1983, Vol. 2. KISR 1366/MB-48. Kuwait Institute for Scientific Research, Kuwait.
- Bedford, B. 1982. The industrial fishery for shrimp. Tech. Rep. KISR 648. Kuwait Institute for Scientific Research, Kuwait.
- Branford, J.R. 1980. Bopyrid parasites of penaeid shrimp from the Tokur Delta (Red Sea), p. 275-280. In C.P. Mathews (ed.) Proceedings of Symposium on the Coastal and Marine Environment of the Red Sea, Gulf of Aden and Tropical Western Indian Ocean, 9-14 January 1980, Khartoum, Vol. II.
- Cassie, R.M. 1954. Some uses of probability paper in the analysis of size frequency distribution. Aust. J. Mar. Freshwat. Res. 5: 513-522.
- El-Musa, M., M. Al-Hossaini and A.R. Abdul-Ghaffar. 1981. Infestation of the shrimp *Penaeus semisulcatus* (de Haan) by the bopyrid *Epipenaeon elegans* (Chopra) in Kuwait. Kuwait Institute for Scientific Research, Kuwait.
- El-Musa, M. 1984. Growth estimates for Kuwaiti shrimp populations with special reference to *M. affinis* and *P. styliifera*. In C.P. Mathews (ed.) Proceedings of the Third Shrimp and Fin Fisheries Management Workshop, 19 June 1982, Vol. I. Kuwait Institute for Scientific Research, Kuwait.
- FAO. 1982. Assessment of the shrimp stocks of the west coast of the Gulf between Iran and Arabian Peninsula. FAO-UN Rep. No. FI:DP/RAB/015. Rome.
- Farmer, A.S.D. 1980. Morphometric relationships of commercially important species of penaeid shrimp from the Arabian Gulf. Mariculture and Fisheries Department, Kuwait Institute for Scientific Research, Kuwait.
- Jones, R. 1981. The use of length composition data in fish stock assessment (with notes on VPA and cohort analysis). FAO Fish. Circ. 734. 55 p.
- Jones, R. and N.P. van Zalinge. 1981. Estimates of mortality and population size for shrimp in Kuwaiti water. Kuwait Bull. Mar. Sci. 2: 273-388.
- Mathews, C.P. 1974. An account of some methods of overcoming errors in ageing tropical and subtropical fish populations when hard tissue growth markings are unreliable and the data sparse, p. 158-160. In T.B. Bagena (ed.) Ageing of fish: proceedings of an international symposium. Unwin Brothers Ltd., Surrey.
- Mathews, C.P. 1981. A review of the North American penaeid fisheries, with particular reference to Mexico. Kuwait Bull. Mar. Sci. 2: 325-409.
- Mathews, C.P., editor. 1982a. Proceedings of the shrimp fisheries management workshop. Kuwait Institute for Scientific Research, Kuwait.
- Mathews, C.P. 1982b. The history of Kuwait's fisheries, p. 6-43. In C.P. Mathews (ed.) Proceedings of the Shrimp Fisheries Management Workshop, 17 January 1981. Kuwait Institute for Scientific Research, Kuwait.
- Mathews, C.P. 1982c. Mortality, growth and management of Kuwait's shrimp fishery, p. 287-321. In C.P. Mathews (ed.) Proceedings of the Shrimp Fisheries Management Workshop, 17 January 1981. Kuwait Institute for Scientific Research, Kuwait.
- Mathews, C.P. and A.R. Abdul-Ghaffar. 1986. A review of the present status of Kuwait's shrimp fisheries with special reference to the need for effort limitation, p. 100-126. In A.M. Landry and E.F. Klima (eds.) Proceedings of the Shrimp Yield Workshop, Galveston, Texas, November, 1985.
- Mathews, C.P. and M. Al-Hossaini. 1984. Stock assessment of Kuwait's shrimp population: growth, mortality and life cycles of *Penaeus semisulcatus*, *Metapenaeus affinis* and *Parapenaeopsis styliifera*, and the management of Kuwait's shrimp stocks, p. 112-170. In C.P. Mathews (ed.) Proceedings of the Third Shrimp and Fisheries Management Workshop, 19 June 1982, Vol. 1. Kuwait Institute for Scientific Research, Kuwait.
- Mathews, C.P. and M. Samuel. 1982. The oceanographic environment and its influence on Kuwait's shrimp stocks, p. 142-153. In C.P. Mathews (ed.) Proceedings of the Shrimp Fisheries Management Workshop, 17 January 1981. Kuwait Institute for Scientific Research, Kuwait.
- Mathews, C.P. and R. Mead. 1971. Estimation of ecological production where age of the population is unknown. Nature 230(5293): 403-405.
- Mohammed, K.H., N.P. van Zalinge, R. Jones, M. El-Musa, M. Al-Hossaini and A.R. Abdul-Ghaffar. 1981a. Mark-recapture experiments on the Gulf shrimp *Penaeus semisulcatus* De Haan in Kuwait waters. KISR 401/MB-29. Kuwait Institute for Scientific Research, Kuwait.
- Mohammed, K.H., M. El-Musa and A.R. Al-Ghaffar. 1981b. Observations on the biology of an exploited species of shrimps, *Penaeus semisulcatus* De Haan, p. 33-53. In A.S.D. Farmer (ed.) Proceedings of the International Shrimp Releasing, Marking and Recruitment Workshop, 25-29 November 1978, Salmiyah, State of Kuwait. Kuwait Bull. Mar. Sci. 2.
- Morgan, G.R. 1985. Stock assessment of the Pomfret (*Pampus argenteus*) in Kuwaiti waters. J. Cons., Cons. Int. Explor. Mer 42: 3-10.

- Morgan, G.R. and S. Garcia. 1982. The relationship between stocks of shrimp in Kuwait and Saudi Arabia. *Océanogr. Trop.* 17(2): 133-137.
- Munro, J.L. 1980. Stock assessment models: applicability and utility in tropical small-scale fisheries, p. 35-47. In P.M. Roedel and S.B. Saila (eds.) *Stock assessment for tropical small-scale fisheries*. University of Rhode Island, International Center for Marine Resources Development, Kingston, Rhode Island.
- Pauly, D. 1980. On the interrelationship between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. *J. Cons., Cons. Int. Explor. Mer* 39: 175-192.
- Pauly, D. 1982a. A method to estimate the stock-recruitment relationship of shrimps. *Trans. Amer. Fish. Soc.* 111: 13-29.
- Pauly, D. 1982b. Studying single species dynamics in a tropical multispecies context, p. 33-70. In D. Pauly and G.I. Murphy (eds.) *Theory and management of tropical fisheries*. International Center for Living Aquatic Resources Management, Manila, Philippines and Commonwealth Scientific and Industrial Research Organisation, Cronulla, Australia.
- Pauly, D. 1984. Reply to comments on prerecruit mortality in Gulf of Thailand shrimp stocks. *Trans. Amer. Fish. Soc.* 113: 404-406.
- Pauly, D. 1985. The population dynamics of short-lived species, with emphasis on squids. *N. Atl. Fish. Coun. Sci. Stud.* 9: 143-154.
- Pauly, D. and N. David. 1980. ELEFAN I, a BASIC program for the objective extraction of growth parameters from length frequency data. *Ber. Dtsch. Wiss. Komm. Meeresforsch.* 28: 205-211.
- Perez, A., R. Puga and J. Rodrigues. 1984a. The stock assessment and management of Cuban shrimp stocks, p. 48-119. In C.P. Mathews (ed.) *Proceedings of the Shrimp and Fin Fisheries Management Workshop*, 9-11 October 1983, Vol. 1. KISR 1366/MB-48. Kuwait Institute for Scientific Research, Kuwait.
- Perez, A., R. Puga and G. Venta. 1984b. Studies of recruitment and stock recruitment relations in the Cuban shrimp populations, p. 163-206. In C.P. Mathews (ed.) *Proceedings of the Shrimp and Fin Fisheries Management Workshop*, 9-11 October 1983, Vol. 1. KISR 1366/MB-48. Kuwait Institute for Scientific Research, Kuwait.
- Rheinhard, E.G. 1956. Parasitic castration of Crustacea. *Exp. Parasitol.* 5: 79-107.
- Samuel, M. and G.R. Morgan. 1984. A comparison of length-related and age-related stock assessment of Newaiby, *Otolithes argenteus*, in Kuwait waters, p. 472-494. In C.P. Mathews (ed.) *Proceedings of the Shrimp and Fin Fisheries Management Workshop*, 9-11 October 1983, Vol. 2. KISR 1366/MB-48. Kuwait Institute for Scientific Research, Kuwait.
- Shoushani, M. 1985. The sexual ripening of Kuwaiti shrimp, p. 162-166. In C.P. Mathews (ed.) *Proceedings of the 1984 Shrimp and Other Fisheries Management Workshop*. KISR 1856/MB-51. Kuwait Institute for Scientific Research, Kuwait.
- Sivasubramaniam, K. and M.A. Ibrahim. 1982. Bionomics of the shrimp fishery in the inshore waters of Qatar. *Univ. Sci. Bull.* 3(1): 265-354.
- van Zalinge, N.P. 1984. The shrimp fisheries in the Gulf between Iran and the Arabian Peninsula, p. 71-83. In J. Gulland and B. Rothschild (eds.) *Penaeid shrimps—their biology and management*. Fishing News Books Ltd., Surrey, England.
- van Zalinge, N.P., M. El-Musa, M. Al-Hossaini and A.R. Abdul-Ghaffar. 1981. The Kuwait shrimp fisheries and the shrimp resources in Kuwait waters. KISR 395 and KISR/FAO 7. Kuwait Institute for Scientific Research, Kuwait.

Towards a Method for Short-Term Forecasting of Catch Rates Based on Length Compositions

J.G. SHEPHERD

*Ministry of Agriculture
Fisheries and Food Fisheries Laboratory
Lowestoft, Suffolk NR33 0HT
United Kingdom*

Shepherd, J.G. 1987. Towards a method for short-term forecasting of catch rates based on length compositions, p. 167-176. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

A method has been developed for combining a time series of length compositions and estimates of growth parameters to produce estimates of future length compositions in the absence of mortality. This is the kernel of the problem of making short-term forecasts of catch rates (and hence catches). Various methods for estimating and allowing for mortality are possible, but the most suitable has yet to be determined.

Introduction

Methods of assessment based on length compositions have hitherto tended to be of essentially two types. Firstly, there are methods (e.g., MacDonald 1980; Pauly and David 1981; Shepherd, Part I, this vol.) for use in the estimation of growth parameters which attack the interpretative rather than the assessment end of the whole spectrum of methods. Secondly, there are methods such as those of Jones (1981) which do attack the assessment problem but only at the expense of averaging over time to produce results appropriate to an approximate steady state. It is feasible to base certain other standard assessment calculations on length-based data and some estimate of growth parameters. Woolner and Pope (unpublished data) have for example developed such a method for the computation of yield and biomass per recruit. The ELEFAN III program (see Morgan and Pauly, Part II, this vol.) performs several standard assessment procedures (including cohort analysis), but only by using the growth parameters to "slice" the length compositions into age compositions, a process which is rather parameter dependent and probably unnecessary for many purposes (Shepherd, Part I, this vol.).

A central problem of fishery management is the forecasting of catch rates (and hence catches) one or two years ahead. Recent work on simple methods for catch forecasts (Shepherd 1984; Anon. 1984) has shown that very simple methods which distinguish only between adults and recruits

can perform quite well for this purpose. It is, therefore, most likely that a method with even better performance can be devised using length compositions without explicitly reinterpreting these as age compositions. Such a method could in many ways be similar to that of Jones (1981), but should avoid the need for averaging and the restriction to steady-state conditions, as it would be intended to take account of large and small year-classes (modes) in the length composition, and estimate their propagation rather than averaging over them.

Method

It is here assumed that one has available a time series (not necessarily a long one) of length compositions which may be taken to be indicative of population abundance together with estimates of the parameters of a suitable growth equation for the stock. The length compositions should therefore be of catch per unit effort or from a research vessel survey; for convenience it is assumed that they all relate to the same season or annual average. For simplicity it is assumed in what follows that the growth relationship is that of von Bertalanffy, the parameters of which may be estimated in various ways such as that of Shepherd (Part I, this vol.).

Given the growth relationship the future length of any individual fish can be estimated at any time; in particular, of course, it can be estimated at the time for which a forecast is required. For simplicity we assume that this is for one year following the last data available. The method may be repeated as often as required if necessary. This growth information is the essence of what is required for the forecasting process. It is very easy to estimate the forward propagation of individual lengths or length group boundaries, using the Ford-Walford relationship

$$L(t+1) = L_{\infty} (1 - \rho) + \rho L(t) \quad \dots 1)$$

$$\text{where } \rho = \exp(-K\Delta t) \quad \dots 2)$$

However, it is desirable that the future estimated length composition shall have the same form as the historic ones, i.e., that it should be based on the same length intervals. It is obvious that growth moves the boundaries of length groups and that the new boundaries will not in general coincide with the old ones. Furthermore, it is clear from equation (1) that the size of length intervals is reduced by growth. The length groups are thus also compressed. If $L(\ell, t)$ is the lower length limit of the ℓ^{th} length group at time t ,

$$L(\ell, t+1) = L_{\infty} (1 - \rho) + \rho L(\ell, t)$$

$$L(m, t+1) = L_{\infty} (1 - \rho) + \rho L(m, t)$$

and therefore

$$\begin{aligned} \Delta L(t+1) &= L(\ell, t+1) - L(m, t+1) \\ &= \rho \{ L(\ell, t) - L(m, t) \} \\ &= \rho \Delta L(t) \end{aligned} \quad \dots 3)$$

Thus (coincidentally), for the von Bertalanffy growth relationship all length intervals are compressed by the same factor (ρ).

The most convenient way to represent the transformation of one length composition into another is by means of a transition matrix so that

$$N(\ell, t + 1) = \sum_k A(\ell, k)N(k, t) \quad \dots 4)$$

Here the transition matrix A must depend on the effects of both growth (as outlined above) and mortality. Although these processes occur concurrently, we assume that for practical purposes they can be separated as the product of two terms, i.e.,

$$A(\ell, k) = G(\ell, k)S(k) \quad \dots 5)$$

Here the matrix G represents the effects of growth in the absence of mortality, S represents the effects of mortality (survival) and also of selection to the sampling gear.

The matrix G can easily be estimated and the procedure proposed is to use this to estimate from each length composition what the next one would be expected to be in the absence of mortality and selection and then to compare observed and expected to estimate these effects.

The elements of the matrix G are simply the proportions of each initial length group k which contribute to each final length group ℓ . Referring back to equation (1), the length $\tilde{L}(\ell)$ at time t corresponding to the lower length boundary $L(\ell)$ of the ℓ^{th} length group is

$$\tilde{L}(\ell) = \{ L(\ell) - \lambda \} / \rho \quad \dots 6)$$

where $\lambda = L_\infty (1 - \rho)$.

The matrix G is then calculated as

$$\begin{aligned} G(\ell, k) &= 0 && \text{if } L(k + 1) < \tilde{L}(\ell) \\ G(\ell, k) &= \{ L(k + 1) - \tilde{L}(\ell) \} / \Delta L && \text{if } L(k) < \tilde{L}(\ell) < L(k + 1) \\ G(\ell, k) &= 1.0 && \text{if } \tilde{L}(\ell) < L(k) \\ &&& \text{and } L(k + 1) < \tilde{L}(\ell + 1) \\ G(\ell, k) &= \{ \tilde{L}(\ell + 1) - L(k) \} / \Delta L && \text{if } L(k) < \tilde{L}(\ell + 1) < L(k + 1) \\ G(\ell, k) &= 0 && \text{if } \tilde{L}(\ell + 1) < L(k) \end{aligned} \quad \dots 7)$$

where $\Delta L = L(k + 1) - L(k)$ is assumed to be constant for all k. The first, third and fifth cases are obvious, and the second and fourth have the effect of distributing numbers in accordance with linear interpolation on length.

To allow for the (usual) case in which the final length group is, in fact, a plus group one should also set

$$G(\ell, k) = 1.0 \quad \text{if } \ell = \ell_{\max} \quad \dots 7a$$

and $\tilde{L}(\ell) < L(k)$

An example of this matrix is shown in Table 1. It has elements which are mostly zero but otherwise less than or equal to one. The non-zero elements lie below the diagonal in a fairly narrow band not parallel to the diagonal. This matrix is just the counterpart of the usual age composition transition matrix which is so simple that it is hardly ever written down as such, i.e., a subdiagonal matrix with unit elements on the subdiagonal

$$\begin{matrix} 0 & 0 & 0 & 0 & . & . & . & . \\ 1 & 0 & 0 & 0 & . & . & . & . \\ 0 & 1 & 0 & 0 & . & . & . & . \\ 0 & 0 & 1 & 0 & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & . & . & 1 & 1 \end{matrix}$$

It may also be regarded as a matrix representation of the Ford-Walford equation (the slope of the non-zero subdiagonal band is just ρ). It should be noted that the elements sum to 1 down the columns (because they are proportions) and sum to $1/\rho$ along the rows (because of the compression effect).

Once the G matrix has been estimated, it is possible to compute for each length composition $N(\ell, t)$ available the next composition $\tilde{N}(\ell, t + 1)$ in the series in the absence of mortality and partial selection. The ratios (more generally, the relationships) between these "predicted" compositions and those actually observed may then be used to estimate the combined effects, since for each t , $N(\ell, t)/\tilde{N}(\ell, t)$ provides an estimate of $S(\ell)$.

The best way to make use of these estimates is not yet clear. They clearly may be interpreted to provide estimates of mortality and this is a possibility which needs to be explored.

Table 1. Length-based catch forecast for North Sea Groundfish Survey cod RF Area 4 for $L_\infty = 110$ cm and $K = 0.2$ per year.

Transition matrix

(cm)	5.	10.	15.	20.	25.	30.	35.	40.	45.	50.	55.	60.	65.	70.	75.	80.	85.
5.	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
10.	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
15.	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
20.	.2	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
25.	.8	.5	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
30.	.0	.5	.7	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
35.	.0	.0	.3	.9	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
40.	.0	.0	.0	.1	1.0	.1	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
45.	.0	.0	.0	.0	.0	.9	.3	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
50.	.0	.0	.0	.0	.0	.0	.7	.6	.0	.0	.0	.0	.0	.0	.0	.0	.0
55.	.0	.0	.0	.0	.0	.0	.0	.4	.8	.0	.0	.0	.0	.0	.0	.0	.0
60.	.0	.0	.0	.0	.0	.0	.0	.0	.2	1.0	.0	.0	.0	.0	.0	.0	.0
65.	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	1.0	.2	.0	.0	.0	.0	.0
70.	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.8	.5	.0	.0	.0	.0
75.	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.5	.7	.0	.0	.0
80.	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.3	.9	.0	.0
85.	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.1	1.0	1.0

For the present, however, the operational task of simply using the numbers available to make some sort of forecast by a simple *ad hoc* procedure will be discussed. The methods are presented below in the context of an example.

Results

The method has been applied to data from the English Groundfish Survey of the North Sea. Annual mean catch per hour by 5-cm length groups was extracted for each of the eight years available for Roundfish Area 4 (D. Harding and J. Casey, pers. comm.). The data are tabulated in Table 2, and were analyzed by a new method (Shepherd, Part I, this vol.) to provide estimates of K and L_∞ . The transition matrix (G) is given in Table 1 and has been discussed above. The raw "predictions" based on applying the G-matrix to the data are given in Table 3. The predictions are zero

Table 2. Actual catches of North Sea cod (see also Table 1).

Length class (cm)	Age group (year)							
	1	2	3	4	5	6	7	8
5.0	.00	.50	47.67	.00	130.11	.67	11.00	.17
10.0	.08	.00	.50	.10	63.78	.00	.83	.00
15.0	.67	6.00	.17	6.50	11.33	1.44	.17	10.09
20.0	7.92	21.00	1.25	31.10	1.56	27.11	4.50	43.27
25.0	9.92	20.50	3.58	14.40	8.56	32.33	5.17	24.30
30.0	3.83	14.90	6.50	6.40	9.11	11.33	2.00	6.17
35.0	1.83	4.30	4.83	2.50	6.33	6.33	6.83	2.50
40.0	1.33	3.00	2.33	.70	4.22	2.56	15.00	1.50
45.0	.42	1.50	1.83	1.40	.89	2.67	15.83	1.67
50.0	.92	2.50	1.83	1.00	1.56	.44	7.67	2.16
55.0	.25	1.20	1.25	.60	.44	.56	2.83	3.15
60.0	.75	.40	.33	.70	.56	1.00	.67	.84
65.0	.67	.50	.25	.40	.11	.78	.17	1.00
70.0	.33	.10	.33	.40	.44	.33	.17	.50
75.0	.17	.10	.00	.00	.11	.11	.33	.00
80.0	.33	.00	.08	.00	.00	.11	.00	.00
85.0	.08	.00	.00	.10	.00	.11	.00	.00
Biomass	2.7E+06	1.8E+06	2.0E+06	2.0E+06	3.2E+06	5.5E+06	3.3E+06	

for the first three length groups because the growth parameters indicate that these relate to fish appearing in the length composition for the first time. They are, therefore, new recruits in this context and prediction of their abundance must be based on information other than previous length compositions. Results for the first non-zero length group (the fourth in this case) must also contain a contribution of recruits (the row sum of the G-matrix is less than the usual value of e^K) and must be regarded as suspect, especially in view of the clear indication in the data that mode locations are variable to the extent of at least one length group. The ratios (see Table 4) for the remaining length groups show some consistency with time tending to be large for smaller (and presumably partially selected) length groups and generally less than one for longer length groups where mortality is dominant. There is, however, much variation about the mean ratios.

For this reason it is not clear that simply using the mean ratios (or even weighted mean ratios) would provide a useful basis for prediction. It seems possible that the ratio may depend (*inter alia*) on the size of year-class involved, and for practical purposes it would also be desirable to strive for a more accurate prediction of the larger year-classes. A regression of actual catches on the catches expected in the absence of mortality would allow for both these factors and this has been tried out. The program used incorporated a regression subroutine which permits (amongst other things) the regression to be forced through the origin if required, thereby approaching the mean ratio method and also permits a variety of weighting schemes to be used if desired.

Table 3. Expected catches of North Sea cod in absence of mortality (see also Table 1).

Length class (cm)	Age group (year)							
	1	2	3	4	5	6	7	8
5.0	.00	.00	.00	.00	.00	.00	.00	.00
10.0	.00	.00	.00	.00	.00	.00	.00	.00
15.0	.00	.00	.00	.00	.00	.00	.00	.00
20.0	.00	.00	.12	11.26	.00	30.73	.16	2.60
25.0	.00	.04	.38	36.64	.05	128.57	.51	8.78
30.0	.00	.50	4.07	.39	4.47	42.29	.98	.57
35.0	.00	7.35	20.83	1.18	30.09	5.04	24.87	4.11
40.0	.00	11.18	24.41	4.50	18.28	9.82	36.41	5.86
45.0	.00	3.99	14.56	7.37	6.48	10.17	12.12	4.10
50.0	.00	1.95	4.52	4.49	2.04	6.54	5.60	12.95
55.0	.00	.91	2.49	2.45	1.41	2.54	3.21	18.97
60.0	.00	1.01	2.83	2.23	1.30	1.75	1.02	11.08
65.0	.00	.42	1.28	1.32	.76	.56	.78	2.96
70.0	.00	.88	.53	.37	.72	.48	1.12	.59
75.0	.00	.59	.34	.36	.49	.36	.65	.21
80.0	.00	.26	.12	.11	.13	.24	.21	.35
85.0	.00	.43	.01	.08	.10	.01	.23	.04
Biomass	3.6E+06	7.3E+06	4.6E+06	5.6E+06	9.1E+06	8.4E+06	1.1E+07	

Table 4. Catch ratios of North Sea cod (see also Table 1).

Length class (cm)	Age group (year)							
	2	3	4	5	6	7	8	
20.0	.00	10.59	2.76	.00	.88	28.44	16.66	
25.0	560.04	9.37	.39	187.08	.25	10.10	2.77	
30.0	29.90	1.60	16.55	2.04	.27	2.05	10.91	
35.0	.59	.23	2.12	.21	1.26	.27	.61	
40.0	.27	.10	.16	.23	.26	.41	.26	
45.0	.38	.13	.19	.14	.26	1.31	.41	
50.0	1.28	.41	.22	.77	.07	1.37	.17	
55.0	1.32	.50	.24	.31	.22	.88	.17	
60.0	.40	.12	.31	.43	.57	.66	.08	
65.0	1.19	.19	.30	.15	1.38	.22	.34	
70.0	.11	.62	1.09	.61	.69	.15	.84	
75.0	.17	.00	.00	.23	.31	.51	.00	
80.0	.00	.65	.00	.00	.45	.00	.00	
85.0	.00	.00	1.25	.00	9.34	.00	.00	

So far the method has been tested using (a) a normal unweighted predictive regression (Table 5); (b) an unweighted predictive regression forced through the origin (Table 6); (c) a regression through the origin whose slope is the ratio of mean numbers at length (Table 7). The first two cases implicitly assume that the variance of numbers is constant, which (since in practice it is probably an increasing function of numbers) effectively gives undue weight to the largest numbers which may be desirable for a forecasting procedure.

These results may be summarized in a more intelligible way by constructing an index of survey catch rate, by simply summing the numbers at length multiplied by $(\text{length})^3$, a surrogate for weight at length. This is labelled "biomass" in the tables, and the results for this index are illustrated in Fig. 1. They may perhaps be described as encouraging but not impressive. All the predictions increase between 1978 and 1979 and none correctly forecasts the major feature, the peak in 1983.

Table 5. Regression analysis (normal predictive regression, see text).

L·GRP	INTCPT	GRADIENT	RSQUD	XBAR	YBAR	RESSQ	SXX	SXY	SYY	CV
20.0	14.7681	.5880	.1671	6.4	18.5	1,360.2	112.8	66.3	233.3	.70
25.0	11.5308	.1607	.5130	25.0	15.5	333.1	1,939.5	311.8	97.7	.41
30.0	7.3495	.0932	.1186	7.6	8.1	91.7	203.1	18.9	14.9	.42
35.0	3.3352	.1099	.4970	13.4	4.8	9.9	115.4	12.7	2.8	.23
40.0	-1.5810	.3656	.7328	15.8	4.2	38.4	112.7	41.2	20.5	.52
45.0	-1.3057	.5941	.1983	8.4	3.7	139.4	13.9	8.3	24.8	1.13
50.0	2.3686	.0152	.0006	5.4	2.5	34.6	11.9	.2	5.0	.85
55.0	.8611	.1252	.5181	4.6	1.4	3.6	35.1	4.4	1.1	.47
60.0	.5759	.0221	.1141	3.0	.6	.3	11.2	.2	.0	.30
65.0	.2097	.2155	.3211	1.2	.5	.4	.6	.1	.1	.51
70.0	.5730	-3.707	.4460	.7	.3	.1	.1	.0	.0	.29
75.0	-1.219	.5323	.6092	.4	.1	.0	.0	.0	.0	.54
80.0	.0524	-0.187	.0025	.2	.0	.0	.0	.0	.0	.59
85.0	.0631	-0.0908	.1426	.1	.1	.0	.0	.0	.0	.57

Predicted catches of cod from regression.

Length class (cm)	Age group (year)							
	1	2	3	4	5	6	7	8
20.0	.00	14.79	14.84	21.39	14.79	32.84	14.86	16.30
25.0	.00	11.54	11.59	17.42	11.54	32.20	11.61	12.94
30.0	.00	7.40	7.73	7.39	7.77	11.29	7.44	7.40
35.0	.00	4.14	5.63	3.46	6.64	3.89	6.07	3.79
40.0	.00	2.50	7.34	.06	5.10	2.01	11.73	.56
45.0	.00	1.07	7.35	3.07	2.54	4.74	5.90	1.13
50.0	.00	2.40	2.44	2.44	2.40	2.47	2.45	2.57
55.0	.00	.97	1.17	1.17	1.04	1.18	1.26	3.24
60.0	.00	.60	.64	.63	.60	.61	.60	.82
65.0	.00	.30	.49	.49	.37	.33	.38	.85
70.0	.00	.25	.38	.44	.31	.39	.16	.35
75.0	.00	.19	.06	.07	.14	.07	.22	-0.01
80.0	.00	.05	.05	.05	.05	.05	.05	.05
85.0	.00	.02	.06	.06	.05	.06	.04	.06
Biomass	2.2E+06	3.4E+06	2.5E+06	2.7E+06	3.4E+06	3.6E+06	2.7E+06	

It is arguable whether any of these forecasts would be of any use in practice or offer any significant advantage over the use of the average as a predictor.

Discussion

The forecasts presented above were made using the most obvious *ad hoc* procedures and it is doubtless possible to construct superior estimators. Close examination of all the results does, however, suggest that the major discrepancies arise because of the variability of the data and could not

Table 6. Predictive regression forced through origin (see text).

Length class (cm)	Gradient	RSQUD	XBAR	YBAR	RESSQ	SXX	SXY	SYY	CV
20.0	1.2036	.3864	6.4	18.5	2,478.6	153.9	185.3	577.1	.95
25.0	.2731	.5636	25.0	15.5	1,037.0	2,564.2	700.4	339.4	.73
30.0	.3075	.3091	7.6	8.1	385.9	260.9	80.2	79.8	.86
35.0	.2616	.7766	13.4	4.8	40.5	293.6	76.8	25.9	.47
40.0	.2966	.8354	15.8	4.2	43.9	361.6	107.3	38.1	.56
45.0	.4643	.4743	8.4	3.7	141.3	84.5	39.2	38.4	1.14
50.0	.3258	.4020	5.4	2.5	45.9	41.5	13.5	11.0	.98
55.0	.1955	.6865	4.6	1.4	6.8	56.0	10.9	3.1	.65
60.0	.1078	.5136	3.0	.6	1.6	20.4	2.2	.5	.69
65.0	.3380	.7450	1.2	.5	.5	2.0	.7	.3	.57
70.0	.3864	.6171	.7	.3	.3	.5	.2	.1	.63
75.0	.2762	.7461	.4	.1	.0	.2	.1	.0	.64
80.0	.2029	.6014	.2	.0	.0	.0	.0	.0	.70
85.0	.1313	.1652	.1	.1	.0	.0	.0	.0	1.02

Predicted catches from regression

Length class (cm)	Age group (year)							
	1	2	3	4	5	6	7	8
20.0	.00	.04	.14	13.55	.04	36.98	.19	3.13
25.0	.00	.01	.10	10.01	.01	35.12	.14	2.40
30.0	.00	.15	1.25	.12	1.37	13.00	.30	.17
35.0	.00	1.92	5.45	.31	7.87	1.32	6.51	1.07
40.0	.00	3.31	7.24	1.33	5.42	2.91	10.80	1.74
45.0	.00	1.85	6.76	3.42	3.01	4.72	5.63	1.90
50.0	.00	.64	1.47	1.46	.66	2.13	1.83	4.22
55.0	.00	.18	.49	.48	.27	.50	.63	3.71
60.0	.00	.11	.31	.24	.14	.19	.11	1.19
65.0	.00	.14	.43	.44	.26	.19	.27	1.00
70.0	.00	.34	.21	.14	.28	.19	.43	.23
75.0	.00	.16	.09	.10	.13	.10	.18	.06
80.0	.00	.05	.02	.02	.03	.05	.04	.07
85.0	.00	.06	.00	.01	.01	.00	.03	.00
Biomass	1.0E+06	2.3E+06	1.5E+06	1.7E+06	3.1E+06	2.6E+06	2.6E+06	

be much reduced by fine-tuning the methodology. Any such forecasting procedure is necessarily limited by the quality of the data available and it is entirely possible that this data set (derived from single annual surveys of a small part of the North Sea) is simply not adequate for the job. Further tests on other data sets are obviously necessary. Nevertheless, it is clear that length-based methods of forecasting are feasible and one must now seek for practical methods and establish the quality of data necessary on which to base them.

Table 7. Use of weighted mean ratios (see text).

Length class (cm)	Gradient	RSQUD	XBAR	YBAR	RESSQ	SXX	SXY	SYY	CV
20.0	2.8895	.3978	.9	3.9	196.4	21.7	24.9	72.0	1.28
25.0	.6221	.4628	9.1	9.0	126.3	779.3	224.3	139.6	.44
30.0	1.0592	.3082	4.5	5.7	126.6	131.9	43.0	45.6	.70
35.0	.3597	.7482	10.2	4.1	11.3	204.1	54.9	19.8	.29
40.0	.2654	.7509	9.9	2.0	9.7	156.1	31.1	8.3	.56
45.0	.4387	.4955	7.4	1.7	26.3	66.2	14.4	6.3	1.06
50.0	.4505	.3943	5.6	1.3	26.4	39.3	7.0	3.1	1.41
55.0	.3137	.6816	2.7	.9	4.7	18.6	4.0	1.2	.88
60.0	.2120	.5333	2.7	.6	3.9	15.1	1.7	.4	1.25
65.0	.3970	.6961	.9	.3	1.4	1.1	.3	.1	1.54
70.0	.4832	.5133	.8	.2	2.2	.7	.2	.1	2.13
75.0	.2472	.6835	.3	.1	.3	.1	.0	.0	3.88
80.0	.2391	.6284	.2	.0	.2	.1	.0	.0	4.26
85.0	.4010	.2682	.1	.0	1.0	.0	.0	.0	9.30

Predicted catches from regression

Length class (cm)	Age group (year)							
	1	2	3	4	5	6	7	8
20.0	.00	.09	.34	32.53	.09	88.78	.46	7.51
25.0	.00	.02	.24	22.79	.03	79.98	.32	5.46
30.0	.00	.53	4.31	.41	4.73	44.79	1.04	.60
35.0	.00	2.64	7.49	.42	10.82	1.81	8.95	1.48
40.0	.00	2.97	6.48	1.19	4.85	2.61	9.66	1.56
45.0	.00	1.75	6.39	3.23	2.84	4.46	5.32	1.80
50.0	.00	.88	2.04	2.02	.92	2.95	2.52	5.84
55.0	.00	.29	.78	.77	.44	.80	1.01	5.95
60.0	.00	.21	.60	.47	.28	.37	.22	2.35
65.0	.00	.17	.51	.52	.30	.22	.31	1.18
70.0	.00	.43	.26	.18	.35	.23	.54	.29
75.0	.00	.15	.08	.09	.12	.09	.16	.05
80.0	.00	.06	.03	.03	.03	.06	.05	.08
85.0	.00	.17	.00	.03	.04	.00	.09	.01
Biomass		1.2E+06	2.6E+06	2.2E+06	2.0E+06	5.9E+06	2.9E+06	3.7E+06

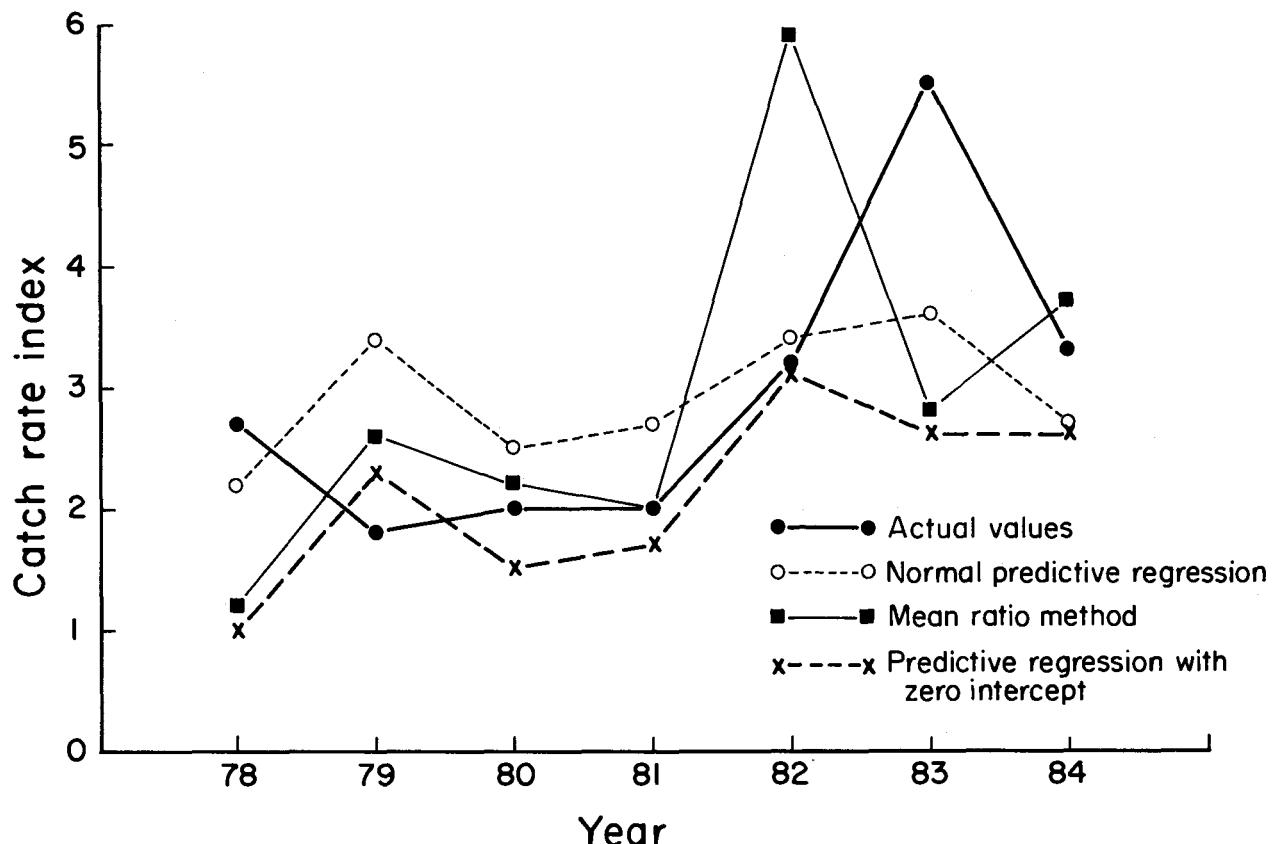


Fig. 1. Comparison of actual catch-rate indices and forecasts by various methods.

References

- Anon. 1984. Report of the ICES working group on the methods of fish stock assessment, ICES CM 1984. International Council for the Exploration of the Sea, Copenhagen. (Mimeo)
- Jones, R. 1981. The use of length composition data in fish stock assessment (with notes on VPA and cohort analysis). FAO. Fish. Circ. No. 734.
- MacDonald, P.D.M. 1980. A Fortran program for analyzing distribution mixtures. McMaster University, Dept. Math. Sci. (Stats.) Tech. Rep. 80-ST-1.
- Pauly, D. and N. David. 1981. ELEFAN I: a BASIC program for the objective extraction of growth parameters from length frequency data. Meeresforsch. 28(4): 205-211.
- Shepherd, J.G. 1984. Status quo catch estimation and its use in fisheries management, ICES CM 1984/G: 5. International Council for the Exploration of the Sea, Copenhagen. (Mimeo)

Phalanx Analysis: An Extension of Jones' Length Cohort Analysis to Multispecies Cohort Analysis

J.G. POPE

*Ministry of Agriculture, Fisheries and Food
Fisheries Laboratory
Lowestoft, Suffolk, NR33 0HT
United Kingdom*

and

YANG JIMING

*Institute of Oceanology, Academia Sinica
Qingdao, China*

Pope, J.G. and Yang Jiming. 1987. Phalanx analysis: an extension of Jones' length-cohort analysis to multispecies cohort analysis, p. 177-191. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

Phalanx analysis (multispecies length cohort analysis) is described. This method calculates fishing mortality and population numbers at length, given information on the growth and feeding characteristics and the catch length distributions from a multispecies fishery. The approach can also be adapted to provide a method for calculating the steady-state yields which would result from a change in fishing patterns. It is thus a multispecies extension of Jones' length cohort analysis.

An example of the use of phalanx analysis is shown based on constructed data which represent the North Sea roundfish. This illustrates the results obtained and also indicates that the method might be used in consideration both of changes in overall size compositions resulting from changes in fishing patterns and of size dependent mortality rates.

Introduction

In recent years the need for fisheries management to consider multispecies effects has been more clearly perceived and the state of the art of multispecies management has been reviewed, notably in Mercer (1982).

A common problem is that multispecies fisheries models are apt to require large data sets if they are to be used successfully. Models based on catch-at-age data are particularly demanding in

this respect. This makes them difficult to use even for developed fisheries such as those in the North Sea, and it makes them virtually impossible to use for areas where catch-at-age data have not been routinely collected for the most numerous species for a number of years. For the fisheries of many areas a more realistic alternative to such a data set would be average catch composition by length. It follows that developing extensions of Jones' length cohort analysis to a multispecies fishery would be highly desirable. Jones (1974, 1976) shows how cohort analysis can be adapted to assess numbers-at-length data and how to reverse this model to assess the effects of changes in mesh size and of fishing effort. The same approach can be used to convert Pope's (1979) multispecies age cohort analysis to a multispecies length cohort analysis. One method of extending length cohort analysis called phalanx analysis is described in this paper.

Development of Phalanx Analysis

Pope (1972) developed a method of retrospective analysis of catch numbers-at-age data usually called "Pope's cohort analysis". The results of this method closely approximate the results obtained from Virtual Population Analysis (VPA) (Gulland 1965) and from the equivalent cohort analysis of Murphy (1965) provided the natural mortality M per time period is less than about 0.3 and that fishing mortality F per time period is less than about 1.2. The basic formula of this method is

$$N_t = C_t \exp(M/2) + N_{t+1} \exp(M) \quad \dots 1)$$

where N_t is the number of fish from a cohort alive at age t and C_t is the catch number of fish of age between t and $t+1$.

Jones' length cohort analysis (Jones 1974) is based on the extension of Pope's cohort analysis formula to variable time intervals (Δt) which modifies equation (1) and gives the form

$$N_1 = C_{1/2} \exp(M\Delta t_{1/2}/2) + N_2 \exp(M\Delta t_{1/2}) \quad \dots 2)$$

Then assuming von Bertalanffy growth, Jones modifies this to the form

$$N_1 = N_2 \left(\frac{L_\infty - L_1}{L_\infty - L_2} \right)^{M/K} + C_{1/2} \left(\frac{L_\infty - L_1}{L_\infty - L_2} \right)^{M/2K} \quad \dots 3)$$

where N_1 and N_2 are the numbers in the sea at length L_1 and L_2 , respectively; $C_{1/2}$ is the catch number in the interval $L_1 - L_2$ and K and L_∞ are the parameters of the von Bertalanffy equation.

Jones' length cohort analysis enables catch-at-length data to be used instead of catch-at-age data in a cohort analysis. The drawback is that the method requires the data to come from a group of cohorts where fishing mortality and recruitment have been in a steady state, since although equation (3) applies strictly only to cohorts these cannot be easily identified in the length distribution.

In practice steady-state conditions are approximated by averaging the catch-at-length data for a number of years and only average results can be obtained. Nevertheless, the method is useful as a quick summary of events which can be based on fewer years' data than an annual (VPA) or cohort analysis.

A practical criticism is that when the interval $L_1 - L_2$ lies close to L_∞ then $(\Delta t_{1/2})$ may become large. In these circumstances $M \Delta t_{1/2}$ may become sufficiently large (> 0.3) to render the cohort analysis approximation to VPA inaccurate. In practice therefore a VPA formulation of the problem will be preferable to equation (2) in a computer algorithm. Equation (2) is, however, far simpler to understand and provides a better basis for developing the multispecies extension of Jones' method. This development may, of course, easily be converted to the more accurate but less comprehensible VPA formulation. (Pope 1972 shows details of the VPA equivalent to equation (1) which may be

easily adapted to the variable time form of equation (3), to the multispecies form of equation (4) and to the variable time multispecies form of equation (5) shown below.)

Given catch-at-age data for several species it is possible to develop multispecies extensions to VPA and cohort analysis which take some account of predation as a source of natural mortality. Such methods were developed by Forney (1977), Pope (1979), Helgason and Gislason (1979) and Majkowski (1981) and discussed by Ursin (1982). They are often described as "legion analyses". Pope (1979) gives as the basic formula of his method

$$N_t = (C_t + D_t) \exp(M_1/2) + N_{t+1} \exp(M_1) \quad \dots 4)$$

where D_t is the number of fish of the cohort devoured by cohorts of fish species for which catch-at-age data are available and which are included in the analysis. D_t may be thought of as the catch taken by fish included in the analysis, while M_1 is natural mortality from causes other than this predation.

Formula (3) may be converted into a multispecies length cohort form which we will call the phalanx analysis formula

$$N_1 = (C_{1/2} + D_{1/2}) \left(\frac{L_\infty - L_1}{L_\infty - L_2} \right)^{M_1/2K} + N_2 \left(\frac{L_\infty - L_1}{L_\infty - L_2} \right)^{M_1/K} \quad \dots 5)$$

where $D_{1/2}$ is the number of fish eaten by other fish included in the analysis between length L_1 and L_2 . This can be straightforwardly calculated as in Pope (1979) using estimates of predator/prey size ratio and food preference developed in the next section. This approach to feeding which considers other food to be a constant proportion of diet was considered by Ursin (1982) to be less realistic than other approaches but also described as "charmingly simple". As such, it is most appropriate for the first explanation of a phalanx analysis. It could, however, be modified if this were considered desirable.

Calculating the Numbers of Fish Eaten

$D_{1/2}$ may be more conveniently written as $D(i, l)$, the number of fish eaten on average of the l th length group of the i th prey species (hereafter called prey (i, l)). Clearly

$$D(i, l) = \sum_{\text{all } I} \sum_{\text{all } L} d(i, l, I, L) \quad \dots 6)$$

where $d(i, l, I, L)$ is the average number of fish of prey (i, l) eaten by the L th length group of the I th predator (hereafter called predator (I, L)). In turn we may write

$$d(i, l, I, L) = B(i, l, I, L) R(I, L) \Delta t(I, L) \bar{N}(I, L) / Wt(i, l) \quad \dots 7)$$

where $Wt(i, l)$ is the average weight of an individual prey (i, l) , where $\bar{N}(I, L)$ is the average number of predators (I, L) , $\Delta t(I, L)$ is the time an average individual predator takes to grow through the L th length interval, $R(I, L)$ is the ration (in weight) per unit time that it eats from all possible prey (i, l) included in the analysis and $B(i, l, I, L)$ is the proportion (in weight) of this ration that it takes from prey (i, l) . Formula (6), thus, states that the biomass of prey (i, l) eaten by predator (I, L) equals the proportion of prey (i, l) in the diet of predator (I, L) multiplied by the aggregate consumption of all prey in the analysis and divided by the weight of prey (i, l) to convert the biomass to numbers.

WEIGHT AT LENGTH AND TIME IN EACH LENGTH INTERVAL

In the formula all the terms except $\bar{N}(I, L)$ and $B(i, l, I, L)$ which is partly a function of $\bar{N}(i, l)$ may be written as relatively simple functions of length. Clearly $Wt(i, l)$ may be simply expressed as a length-weight relationship

$$Wt(i, l) = a(i) Lt(i, l)^b(i) \quad \dots 8)$$

where $Lt(i, 1)$ is the average length of prey ($i, 1$). A similar formula holds for $Wt(I, L)$. Similarly $\Delta t(I, L)$ may be given, following Jones (1974) as

$$\Delta t(I, L) = \frac{1}{K(I)} \ln \left(\frac{L_\infty(I) - L_1(I, L)}{L_\infty(I) - L_2(I, L)} \right) \quad \dots 9)$$

where $K(I)$ and $L_\infty(I)$ are the von Bertalanffy parameters for stock I and where $L_1(I, L)$ and $L_2(I, L)$ are the initial and final lengths of the L th length group for predator (I, L) . A similar formula holds for $\Delta t(I, L)$.

INDIVIDUAL RATION

The ration ($R(I, L)$) may conveniently be taken in two parts. The first part is a total ration given by a power function of the mid-length of the interval (Lt)

$$f(I) Lt(I, L)^g(I) \quad \dots 10)$$

where $f(I)$ and $g(I)$ are parameters which characterize predator stock (I) 's feeding requirements per unit time. An example of such a function is given in Daan (1973) who suggests the formula

$$\phi_L = \frac{2 \times \text{average food in stomach}}{\text{stomach depletion time}} \quad \dots 11a)$$

which he takes as

$$\phi_L = 0.00533 Lt^2 \quad \dots 11b)$$

for North Sea cod, where Lt is in cm and ϕ_L is daily ration (g).

The second part of $R(I, L)$ is the proportion of the total diet satisfied by all the possible prey ($i, 1$) in the analysis. This may be simply expressed as a logistic ogive

$$r(I)/(1.0 + \exp \{-p(I)[Lt(I, L) - q(I)]\}) \quad \dots 12)$$

where $p(I)$, $q(I)$ and $r(I)$ characterize the rate of rise, 50% point and maximum height of the ogive, respectively. Sparre (1979) shows an ogive of feeding proportion for North Sea cod.

RATION PROPORTION TAKEN FROM PREY (i, l)

The factor $B(i, l, I, L)$ in equation (6), the ration proportion taken from prey (i, j) is perhaps the most difficult factor to understand. It is in fact the ratio that preference for prey (i, j) multiplied by the available biomass of prey (i, j) has to the sum of this product for all prey. Hence

$$B(i, l, I, L) = \frac{\bar{N}(i, l) \Delta t(i, l) Wt(i, l) A(i, l, I, L)}{\sum_{\text{all } i} \sum_{\text{all } l} \bar{N}(i, l) \Delta t(i, l) Wt(i, l) A(i, l, I, L)} \quad \dots 13)$$

where $A(i, l, I, L)$ is the preference of predator (I, L) for prey (i, l). This is expressed as the product of a predator species/prey species factor $\alpha(i, I)$ and a log normal prey size preference as suggested by Ursin (1973). Hence

$$A(i, l, I, L) = \frac{\alpha(i, I)}{\sigma(I)\sqrt{2\pi}} \exp - \frac{1}{2} \left(\frac{\ln \{Wt(i, l)/Wt(I, L)\} - \mu(I)}{\sigma(I)} \right)^2 \quad \dots 14)$$

where $\mu(I)$ and $\sigma(I)$ are the mean and standard deviation of the log normal prey size preference function of predator I .

ESTIMATING MEAN NUMBERS AT LENGTH

We see from the above subsections that the numbers of prey (i, l) eaten on average by predator (I, L) are a function of length and of the average population numbers of predator (I, L) and of all the prey (i, l) in the analysis.

These $\bar{N}(I, L)$, $\bar{N}(i, l)$ may be estimated by successive approximations in a way analogous to that shown in Pope (1979). The method consists of initially setting $D(i, l) = 0$ and estimating the average numbers in the highest length interval l_{max} as $C(i, l_{max})/(F/Z(i))$. The numbers at the beginning of each lower length interval $N(i, l)$ are then estimated using the equivalent to equation (4) in (i, l) notation, i.e.,

$$N(i, l) = \{C(i, l) + D(i, l)\} \left(\frac{L_\infty(i) - L_1(i, l)}{L_\infty(i) - L_2(i, l)} \right)^{M_1(i)/2K(i)} + N(i, l+1) \left(\frac{L_\infty(i) - L_1(i, l)}{L_\infty(i) - L_2(i, l)} \right)^{M_1(i)/K(i)} \quad \dots 15)$$

Thus, first estimates of $N(i, l)$ for all prey (i, l) included in the analysis are obtained. Since all length groups of all fish species are considered as potential predators and all length groups except l_{max} of all fish species are considered as potential prey these first population estimates can be used to obtain first estimates of $\bar{N}(i, l)$ and $\bar{N}(I, L)$. These are calculated by first estimating $Z(i, l)$, the instantaneous rate of total mortality using

$$Z(i, l) = \frac{1}{\Delta t(i, l)} \ln \left(\frac{N(i, l)}{N(i, l+1)} \right) \quad \dots 16)$$

Then

$$\bar{N}(i, l) = (N(i, l) - N(i, l+1))/Z(i, l)*\Delta t(i, l) \quad \dots 17)$$

Equations (5) through (14) are then used to estimate all the $D(i, l)$ and new values of the $N(i, l)$ computed using equation (15). This process of iteration is continued until no changes occur in any of the $Z(i, l)$ estimated. The values of $N(i, l)$, $D(i, l)$ are then adopted as the final estimates for phalanx analysis. The mathematical appendix shows a flow chart of the operations carried out to make these calculations (Appendix Fig. A1).

Using an Inverse of Phalanx Analysis to Consider the Effects of Mesh Changes and Effort Changes

One very useful feature of Jones' length cohort analysis is that the results may be used to consider the effects of mesh and/or effort changes.

Jones (1976) considers the effect of a mesh change or effort change on catch-at-length data using an approach based on Jones (1961). This approach is not possible with the phalanx analysis because a change in the coefficient of fishing mortality will lead to a change in the coefficient of natural mortality. The inversion problem can, however, be solved as follows. The numbers in the smallest length group of each species are held constant. The numbers at each successively greater length are then calculated using preliminary estimates of total mortality based on the new level of fishing mortality and the natural mortality obtained in the phalanx analysis. These population sizes are used to calculate predation levels which in turn are used to calculate better estimates of natural mortality due to predation. It is then possible to recalculate the population numbers at length using the improved estimates of total mortality which can be used to improve the predation estimates. This cycle of calculations is repeated until changes in the total mortality estimates are less than some specified level. Details of calculations are shown in a flow chart in the mathematical appendix (Fig. A2).

An advantage of this method is that it would be simple to extend this approach by including a stock/recruitment relationship in these calculations. This would modify the numbers in the smallest length group of each species in line with the biomass of fish of spawning size. This might be simply included in the iterations indicated above.

Worked Example of Phalanx Analysis and Its Inverse

The absence of published tables of international catch-at-length data made it difficult in the short term to present a detailed analysis of the North Sea fisheries using phalanx analysis. Moreover such a detailed example would be inappropriate to a first description of the method. An illustrative example has, therefore, been constructed based on the North Sea roundfish. Thus, species A, B and C in the results are somewhat like cod, haddock and whiting. Table 1 shows the input values used for the various coefficients of growth, mortality and feeding. Tables 2, 3 and 4 show the results of the phalanx analysis for species A, B and C, respectively. These results may be used as inputs to the inverse form. As an example of this, reducing the fishing mortality by 50% on all species results in changes in the predation levels. As an example, the effects of this change are shown for species A in Table 5. By making a series of such changes yield curves can be constructed (see Fig. 1) for the three species. In this case these are somewhat artificial because two of the species obtain food at the expense of smaller fish sizes of the same three species, there being no smaller food species in the example (e.g., sandeels). Thus, the yield curves tend to increase sharply with increasing fishing. The lack of a stock/recruitment relationship is also a clear defect in the model in its current form.

An interesting feature of phalanx analysis and its inverse is that they provide a method for investigating overall size distributions and some size-related processes. It has been suggested that regular particle size distributions may be a consistent feature of marine ecosystems (Sheldon et al. 1977). In the North Sea some regularity of overall size distributions has been observed (Pope and Knights 1982). Thus, an ability to investigate overall size distributions may prove to be the most valuable feature of the phalanx model. As an example of this feature, the length distribution obtained from the phalanx analysis is compared with the aggregate length distributions for the North Sea obtained from English groundfish surveys (Fig. 2) (see Pope and Knights 1982). The change in numbers with length can be seen to be rather similar in these two series, and, while this result may be coincidental, it certainly does indicate that phalanx analysis can be used to describe aggregate length distributions. Similarly, the inverse of phalanx analysis could be used to model how overall length distributions might change with changes in exploitation patterns.

Table 1. Stock specific coefficients.

Property	Species A	Species B	Species C	Symbol in appendix
1 L_∞	130.00	85.00	55.00	L_∞ (i)
2 K	0.10	.125	.125	K (i)
3 Natural M1	0.10	0.10	0.10	M1 (i)
4 F/Z for L max	0.70	0.70	0.70	F/Z (i)
5 Wt/length A	0.000010	0.000009	0.000008	a (i)
6 Wt/length B	3.0	3.0	3.0	b (i)
7 Av. prey/pred	-4.60	-4.60	-4.60	μ (i)
8 Sd. prey/pred	1.00	1.00	1.00	σ (i)
9 Ration coeff.	0.000020	0.000000	0.000016	f (i)
10 Ration power	3.00	3.00	3.00	g (i)
11 Food ogive p	0.50	0.50	0.50	p (i)
12 Food ogive q	30.00	30.00	30.00	q (i)
13 Food ogive r	0.50	0.50	0.50	r (i)
14 Largest len.	120	81	50	
15 No. of len. S	23	24	40	
16 Length inc.	5	3	1	
17 α (a, I)	1.0	1.0	1.0	
18 α (b, I)	1.0	1.0	1.0	
19 α (c, I)	1.0	1.0	1.0	

Table 2. Phalanx analysis results for species A.

Length (cm)	Population	Catch	Predation	Z	F	M2	ΔT
10	845,392	0	434,097	1.850	0.000	1.750	.426
15	384,659	0	87,145	.693	0.000	.593	.445
20	282,692	0	35,989	.400	0.000	.300	.465
25	234,672	2,167	16,371	.273	.020	.153	.488
30	205,402	21,177	7,022	.396	.222	.074	.513
35	167,646	30,775	2,692	.525	.390	.034	.541
40	126,248	33,633	924	.679	.563	.015	.572
45	85,650	25,392	293	.710	.603	.007	.606
50	55,694	13,574	94	.553	.450	.003	.645
55	38,979	8,395	32	.467	.366	.001	.690
60	28,239	5,539	12	.408	.307	.001	.741
65	20,873	3,709	4	.356	.256	.000	.800
70	15,700	2,538	2	.313	.213	.000	.870
75	11,960	2,076	1	.311	.211	.000	.953
80	8,893	1,624	0	.303	.203	.000	1.054
85	6,463	1,292	0	.302	.202	.000	1.178
90	4,526	999	0	.302	.202	.000	1.335
95	3,026	1,071	0	.413	.312	.000	1.542
100	1,602	858	0	.585	.485	.000	1.823
105	552	219	0	.363	.263	.000	2.231
110	246	141	0	.478	.378	.000	2.877
115	62	49	0	.931	.831	.000	4.055

Table 3. Phalanx analysis results for species B.

Length (cm)	Population	Catch	Predation	Z	F	M2	ΔT
12	4,103,954	0	1,594,507	1.597	0.000	1.497	.336
15	2,400,465	0	558,830	.871	0.000	.771	.350
18	1,768,770	0	284,449	.588	0.000	.488	.366
21	1,425,793	0	167,446	.432	0.000	.332	.384
24	1,207,767	6,917	103,191	.342	.015	.227	.403
27	1,052,076	97,168	61,440	.494	.241	.152	.425
30	853,025	171,156	32,824	.725	.524	.101	.449
33	616,134	200,169	15,087	1.031	.866	.065	.475
36	377,327	142,053	6,081	1.119	.978	.042	.505
39	214,292	99,433	2,183	1.338	1.211	.027	.540
42	104,125	40,236	766	.998	.882	.017	.579
45	58,438	21,510	296	.879	.769	.011	.624
48	33,769	12,935	115	.855	.748	.007	.676
51	18,944	8,836	41	1.001	.897	.004	.739
54	9,040	5,217	12	1.233	1.130	.003	.814
57	3,312	2,007	3	1.212	1.110	.002	.907
60	1,104	465	1	.674	.573	.001	1.023
63	554	158	0	.408	.307	.001	1.173
66	344	104	0	.386	.285	.000	1.375
69	202	31	0	.210	.110	.000	1.661
72	143	19	0	.176	.076	.000	2.099
75	99	26	0	.227	.127	.000	2.853
78	52	9	0	.155	.055	.000	4.477

Table 4. Phalanx analysis results for species C.

Length (cm)	Population	Catch	Predation	Z	F	M2	ΔT
12	25,431,960	0	13,270,758	2.093	0.000	1.993	.381
14	11,457,542	0	4,242,511	1.286	0.000	1.186	.400
16	6,850,546	0	1,906,117	.893	0.000	.793	.421
18	4,702,830	0	1,046,044	.680	0.000	.580	.445
20	3,475,737	0	642,260	.545	0.000	.445	.471
22	2,688,665	0	419,584	.449	0.000	.349	.500
24	2,148,225	1,076	284,006	.375	.001	.274	.534
26	1,758,976	324,729	176,644	.707	.393	.214	.572
28	1,173,956	404,362	88,825	1.022	.756	.166	.616
30	625,603	256,697	37,671	1.099	.871	.128	.667
32	300,514	144,012	14,248	1.185	.988	.098	.728
34	126,813	63,565	4,925	1.131	.957	.074	.801
36	51,253	21,820	1,788	.838	.682	.056	.890
38	24,309	10,906	700	.796	.654	.042	1.001
40	10,953	5,741	250	.856	.725	.032	1.145
42	4,111	1,884	86	.637	.514	.024	1.336
44	1,754	873	31	.610	.492	.018	1.605
46	659	295	11	.459	.345	.013	2.011
48	262	187	3	.756	.646	.010	2.692

Table 5. Results for species A, where fishing mortality is reduced by 50% on all length groups of all species.

Length (cm)	Population	Catch	Predation	Z	F	M2	ΔT
10	845,392	0	363,492	1.454	0.000	1.354	.426
15	455,391	0	162,684	1.119	0.000	1.020	.445
20	276,892	0	82,953	.885	0.000	.785	.465
25	183,443	782	39,055	.615	.010	.505	.488
30	135,901	6,840	17,556	.496	.111	.285	.513
35	105,363	9,886	7,539	.444	.195	.149	.541
40	82,878	11,746	3,094	.456	.282	.074	.572
45	63,870	10,254	1,226	.438	.301	.036	.606
50	48,989	6,380	491	.342	.225	.017	.645
55	39,281	4,490	203	.291	.183	.008	.690
60	32,132	3,330	86	.258	.154	.004	.741
65	26,549	2,480	37	.230	.128	.002	.800
70	22,090	1,869	16	.207	.106	.001	.870
75	18,446	1,683	7	.206	.105	.000	.953
80	15,159	1,460	3	.202	.101	.000	1.054
85	12,257	1,300	1	.201	.101	.000	1.178
90	9,670	1,141	1	.201	.101	.000	1.335
95	7,396	1,472	0	.256	.156	.000	1.542
100	4,982	1,637	0	.342	.242	.000	1.823
105	2,669	611	0	.231	.131	.000	2.231
110	1,593	588	0	.289	.189	.000	2.877
115	694	490	0	.515	.415	.000	4.055

Change in F = .50

Catch weight = 157,331

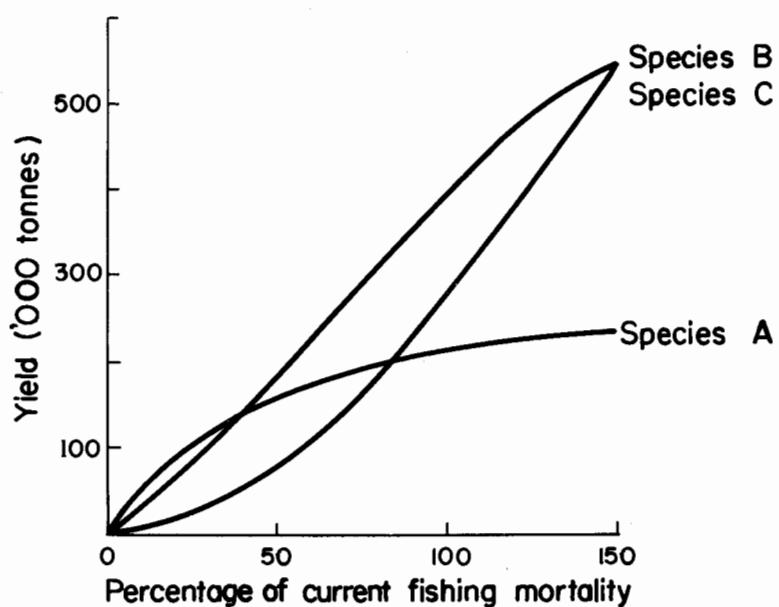


Fig. 1. Yield curves for species A, B and C when all fishing mortalities are changed in the same proportion.

In the example the only preferences predators have for different species are based on weight differences. Consequently, the predation mortality at length is much the same for the three species considered, as can be seen from Fig. 3. This also indicates that for this data set the predation mortality estimates declined exponentially with length in a very regular fashion and thus that phalanx analysis provides a means of investigating this feature of a fish ecosystem besides having obvious uses for fisheries assessment.

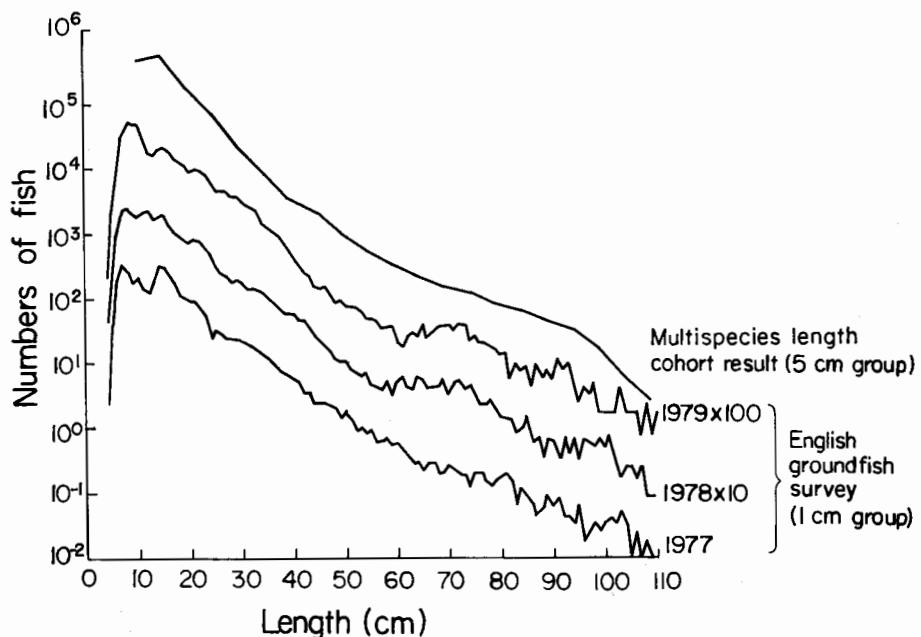


Fig. 2. The total length distribution from phalanx analysis compared to those observed from groundfish surveys of the North Sea.

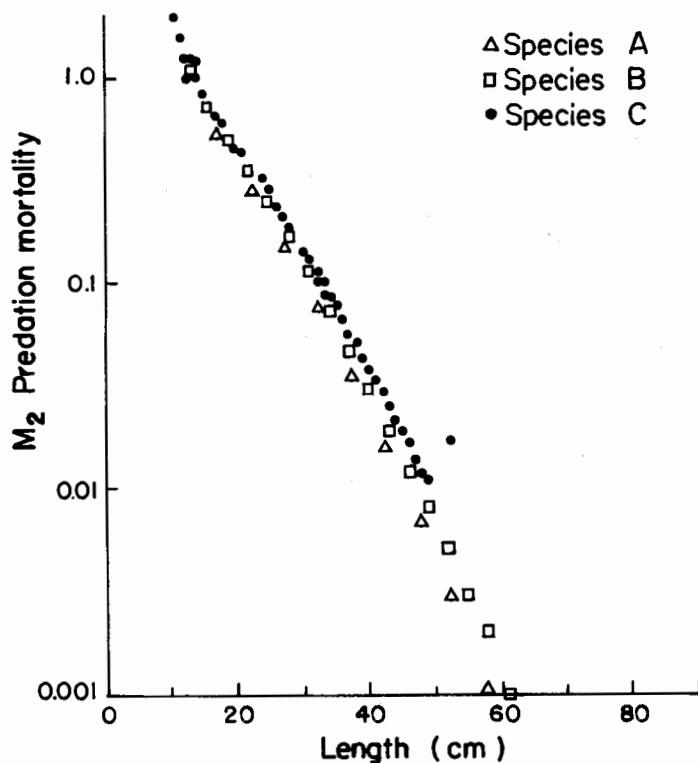


Fig. 3. Phalanx analysis: predation mortality in the example as a function at length.

Discussion

The model presented of a multispecies length cohort analysis has essentially the same virtues and drawbacks as Jones' (1974) length cohort analysis. It is less demanding of data than the equivalent multispecies age cohort models but it may be sensitive to nonstationary data. There would seem some prospect of using such a model on a little documented tropical fishery where average catch length and growth parameters might be inferred but where annual catch-at-age data are not practical to estimate. It will also be worth developing the model for the North Sea fisheries where it can be contrasted with age models and where its simple structure should help the development of insight in multispecies model development. Where necessary, extra realism might be introduced into the version presented here. More realistic feeding models (see Ursin 1982) could certainly be developed. Similarly, stock recruitment effects and possibly density dependent growth effects might be introduced into the inverse of phalanx analysis.

Since the method gives mortality components and population sizes by length it should be useful in interpreting the structures of biomass size distributions in the North Sea. The method should also be of value in obtaining insights into the relationship between growth parameters and natural mortality.

Phalanx analysis may, thus, be seen as both a possible assessment tool and a possible bridge between age-structured, functional approaches to the analysis of complex fisheries systems and system overviews based on an empirical approach. In our present state of knowledge this latter aspect of phalanx analysis may prove to be of greater utility than its parallel use of providing another bookkeeping model of multispecies fisheries. The next stage is, therefore, to develop a full phalanx analysis to the North Sea which should provide a stimulus for further implications in our models of the North Sea fisheries.

Other fisheries biologists may, however, see the method as a way to integrate their knowledge of their fisheries, using the data they have accumulated from standard studies of fish biology such as growth and feeding.

References

- Daan, N. 1973. A quantitative analysis of the food intake of North Sea cod, *Gadus morhua*. Neth. J. Sea Res. 6: 479-517.
- Forney, J.L. 1977. Reconstruction of yellow perch (*Perca flavescens*) cohorts from examination of walleye (*Stizostedion vitreum vitreum*) stomachs. J. Fish. Res. Board Can. 34: 925-932.
- Gulland, J.A. 1965. Estimation of mortality rate. Annex to Report of the Arctic Fisheries Working Group (meeting in Hamburg, January 1965). ICES C.M. 1965, Doc No. 3, 9 p. (mimeo).
- Helgason, T and H. Gislason. 1979. VPA-analysis with species interaction due to predation. ICES CM 1979/G: 52, 10 p. (mimeo).
- Jones, R. 1961. The assessment of the long-term effects of changes in gear selectivity and fishing effort. Mar. Res. 2. 19 p.
- Jones, R. 1974. Assessing the long-term effects of changes in fishing effort and mesh size from length composition data. ICES CM 1974/F: 33, 14 p. (mimeo).
- Jones, R. 1976. A preliminary assessment of the Firth of Forth stock of *Nephrops*. ICES Special Meeting of Population Assessments of Shellfish Stocks 24, 27 p. (mimeo).
- Majkowski, J. 1981. Application of a multispecies approach for assessing the population abundance and the age structure of fish stocks. Can. J. Fish. Aquat. Sci. 38: 424-431.
- Mercer, M.C., editor. 1982. Multispecies approaches to fisheries management advice. Can. Spec. Publ. Fish. Aquat. Sci. 59. 169 p.
- Murphy, G.I. 1965. A solution of the catch equation. J. Fish. Res. Board Can. 22: 191-202.
- Pope, J.G. 1972. An investigation of the accuracy of virtual population analysis. Res. Bull. Int. Comm. NW Atlant. Fish. 9: 65-74.
- Pope, J.G. 1979. A modified cohort analysis in which constant natural mortality is replaced by estimates of predation levels. ICES CM 1979/H: 16, 5 p. (mimeo).
- Pope, J.G. and B.J. Knights. 1982. Comparison of length distributions of combined catches of all demersal fishes in surveys in the North Sea and at Faroe Bank. Can. Spec. Publ. Fish. Aquat. Sci. 59: 116-118.

- Sheldon, R.W., J.R. Sutcliffe and M.A. Paranjape. 1977. Structure of pelagic food chain and relationship between plankton and fish production. *J. Fish. Res. Board Can.* 34: 2344-2353.
- Sparre, P. 1979. Some remarks on the application of yield/recruit curves in estimation of maximum sustainable yield. *ICES CM 1979/G: 41*, 21 p. (mimeo).
- Ursin, E. 1973. On the prey size preferences of cod and dab. *Meddr. Danm. Fisk.-og Havunders., (N.S.) 3*: 85-98.
- Ursin, E. 1982. Multispecies fish stock and yield assessment in ICES. *Can. Spec. Publ. Fish. Aquat. Sci.* 59: 39-47.

Mathematical Appendix

The sequence of calculations performed in making a multispecies length cohort analysis is shown in Appendix Fig. A1. The (supposed) consequences to species yield caused by altering fishing mortality rates result from the calculations shown in Appendix Fig. A2. This figure should be regarded as an extension of Fig. A1 where many of the basic variables are calculated.

In these charts the variables shown in Tables A1, A2 and A3 are used. In general, these are expressed in terms of the following: fish stock is designated by the index i or, in the case of a predator, by the index I ; length interval is designated by the index l or, in the case of a predator, by the index L ; all lengths of all species are regarded as potential predators; all but the greatest length of each species are regarded as potential prey; at present the fish of the greatest length group of each species are regarded as exempt from predation for the sake of simplicity; the largest lengths of each species are designated by the index l_{\max} and the smallest by l_{\min} .

Predation in the model is considered to consist only of predation by those length groups (L) of species (I) for which catch data are input. This source of mortality creates the $M_2(i, l)$ component of natural mortality of the prey species i of length L . The $M_1(i)$ component is that part of the natural mortality of stock i caused by predation by species external to the model or from other causes. At present it is set at an arbitrary low value of 0.1 for all lengths of each species.

Table A1. Stock variables: The following stock variables are used for all lengths of each species (i) of predator (I). Together with catch-at-length data by species they form the basic inputs to the multispecies length cohort analysis.

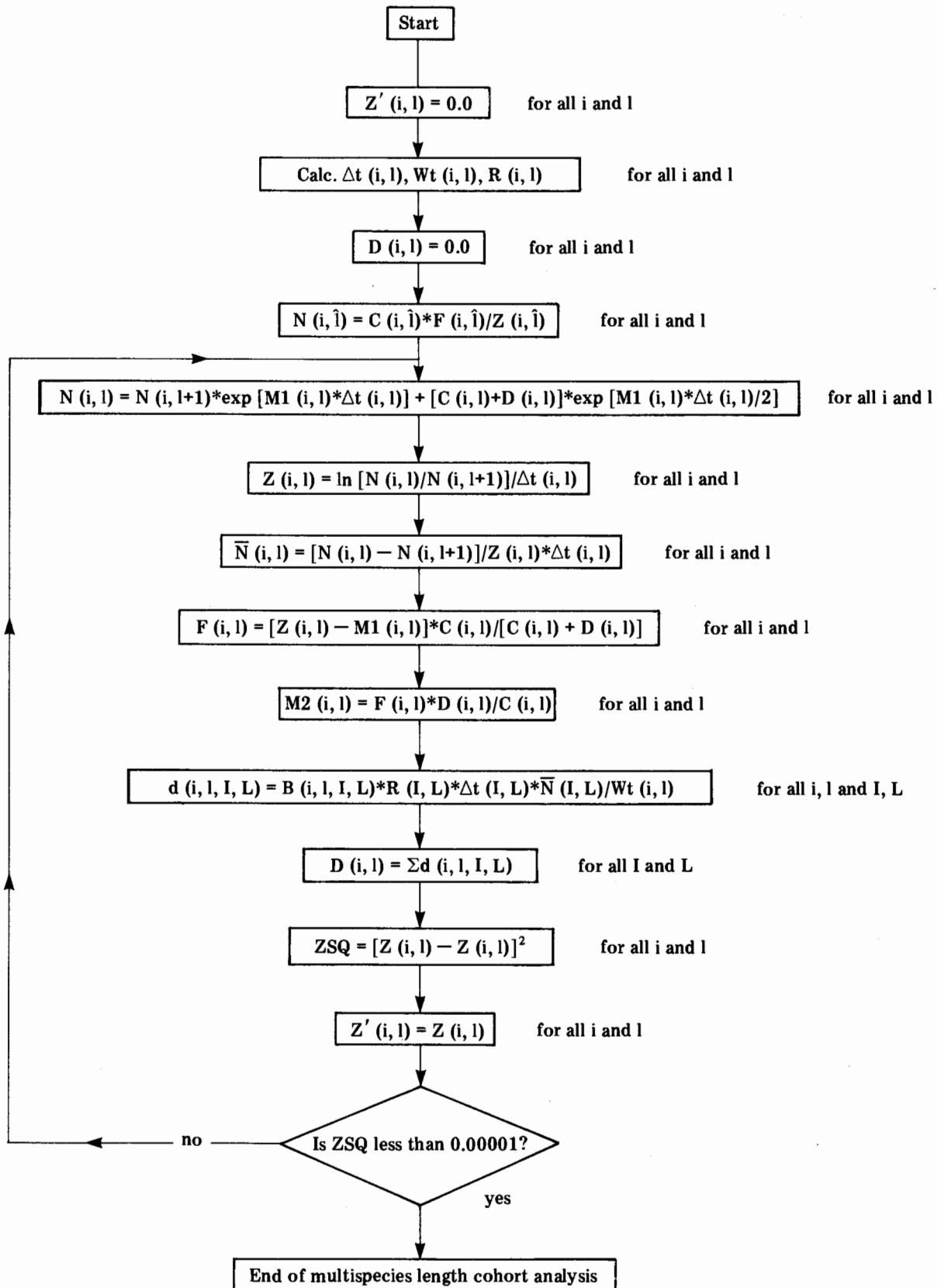
Variable	Use
$L_{\infty}(i)$	Von Bertalanffy coefficients for stock i used
$K(i)$	to calculate $\Delta t(i, l)$
$M_1(i)$	Non-predation mortality of stock i
$F/Z(i)$	F/Z ratio for largest length group of stock i
$a(i)$	Weight-length relationship coefficients for
$b(i)$	stock i used to calculate $Wt(i, l)$
$f(I)$	Ration requirement coefficients for
$g(I)$	stock I
$p(I)$	Proportion of ration taken as predation.
$q(I)$	Coefficients of ogive for stock I used
$r(i)$	to calculate proportion
$\mu(I)$	Food size preference coefficients for
$\sigma(I)$	stock I
$\alpha(i, I)$	General preference matrix of stock I for stock i
	used to calculate $R(I, L)$
	used to calculate $A(i, l, I, L)$

Table A2. Stock length variables: the following variables apply to the length interval l of stock i or in the case of predators to the length interval L of stock I .

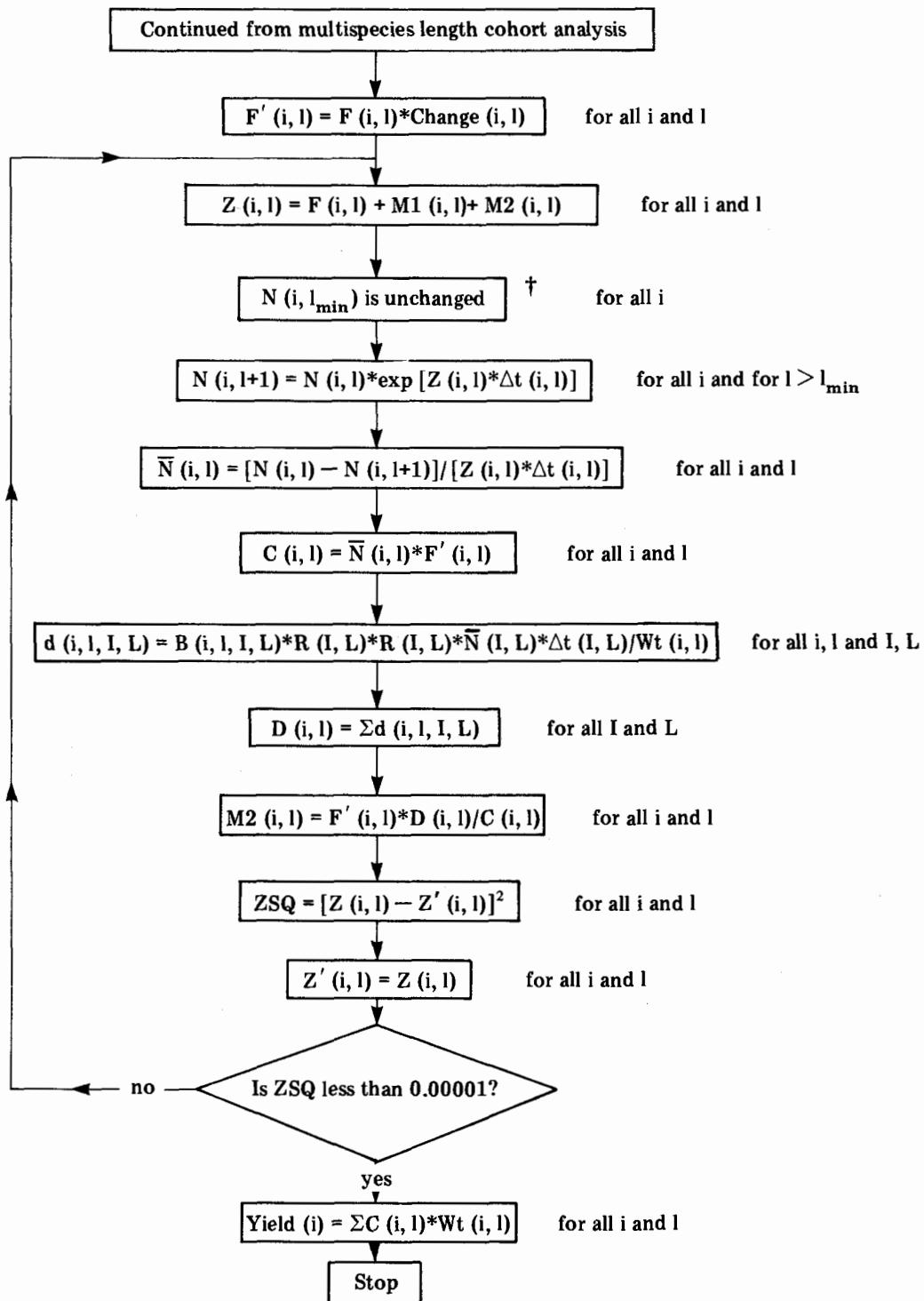
Variable	Use
$N(i, l)$	Population numbers at beginning of length interval
$\bar{N}(i, l)$	Average population numbers in the interval
$C(i, l)$	Catch numbers in the interval ($C_{1/2}$ in equation 3)
$D(i, l)$	Predation numbers in the interval ($D_{1/2}$ in equation 5)
$\Delta t(i, l)$	Time spent in the interval (see Jones 1974)
$F(i, l)$	Fishing mortality rate in the interval
$F'(i, l)$	New level of fishing mortality rate for yield assessment
$M2(i, l)$	Predation mortality in the interval
$Z(i, l)$	Current total mortality in the interval
$Z'(i, l)$	Total mortality in previous iteration
$Lt(i, l)$	Average length in the interval
$Wt(i, l)$	Average weight in the interval = $a(i) Lt(i, l)^b(i)$
$R(I, L)$	Predators' ration from fish in analysis taken here as $f(I) Lt(I, L)^g(I) \times r(I)/(1.0 + \exp\{-p(I)(Lt(I, L) - q(I))\})$

Table A3. Predator/prey length variables: the following variables link predators of length interval L of stock I to their prey of length l of stock i .

Variable	Use
$A(i, l, I, L)$	Preference of predators (I, L) for prey (i, l) . This could be a matrix (see Pope 1979) but here it is taken as
	$\frac{\alpha(i, I)}{\sqrt{2\pi} \sigma(I)} \exp - \left[\frac{1}{2} \left(\frac{\ln \{ Wt(i, l)/Wt(I, L) \} - \mu(I)}{\sigma(I)} \right)^2 \right]$
$B(i, l, I, L)$	is diet proportion of predator (I, L) coming from prey (i, l) which is taken as $\bar{N}(i, l) * \Delta t(i, l) * Wt(i, l) * A(i, l, I, L)$
$d(i, l, I, L)$	is number of prey (i, l) devoured by predator (I, L)



Appendix Fig. A1. Multispecies length cohort analysis.



Appendix Fig. A2. Multispecies yield changes.

A Simulation Model for Generating Catch Length-Frequency Data

J. HAMPTON
J. MAJKOWSKI

*Division of Fisheries Research, Marine Laboratories
Commonwealth Scientific and Industrial Research Organisation
GPO Box 1538, Hobart, Tasmania 7001, Australia*

Hampton, J. and J. Majkowski, 1987. A simulation model for generating catch length-frequency data, p. 193-202. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

This paper describes a simulation model for generating catch length-frequency data for a hypothetical population and associated fishery. Such a model can be used to test the reliability of commonly used mathematical procedures for estimating, from catch length-frequency data, population parameters important to fisheries management. Data generated by the model, which correspond to a population and fishery of known characteristics, can be used as input information to these procedures. The reliability of the estimation procedures can then be assessed by comparing parameter estimates derived by these procedures with their "true" values (i.e., those used for the simulation model). The model is illustrated by an application to a hypothetical tropical population and fishery.

Introduction

Catch length-frequency data from a fish population often provide the basis for estimating important parameters, such as rates of growth and mortality, age composition and abundance of that population. Several mathematical procedures (e.g., Pauly 1980; Schnute and Fournier 1980; Jones 1981; Pauly and David 1981; Pope, Part I, this vol.; Shepherd, Part I, this vol.) are commonly used for estimating such parameters. However, the reliability of these procedures remains largely untested (see review in Gulland 1983), despite the importance of the estimates in making fisheries management recommendations.

The estimation procedures can be tested with the aid of a simulation model for generating catch length-frequency data (for examples, see Jones, Part I, this vol.; Rosenberg and Beddington, Part I, this vol.). By selecting values for input parameters of the model, length-frequency data can be generated for a hypothetical population and fishery of known (i.e., assumed) characteristics. These data can then be used as input information for the procedures for estimating the parameters

of that hypothetical population. Subsequently, the derived estimates can be compared with their "true" values (i.e., those used for the simulation model). By varying the input parameters of the simulation model, an indication of the reliability of the estimation procedures can be obtained for a variety of exploited fish populations.

In this paper, a simulation model for generating catch length-frequency data is described. Features of the model include an allowance for a relatively long spawning period with one or more* peaks of activity (common in tropical populations) seasonally oscillating growth with variation amongst individual fish, a variable number of fish recruited to the fishable stock each year (referred to throughout this paper as "cohort strength"), time-, age- and size-independent instantaneous rates of natural mortality and encounter with the fishing gear and size-selective exploitation resulting from a size-dependent rate of escape from the fishing gear. Although these features of the model were incorporated specifically to examine the reliability of the ELEFAN package (see Hampton and Majkowski, Part I, this vol.), other features could easily be included as required for specific purposes.

General Description of the Model

In the model, each cohort (i.e., all fish hatched during a particular year) is considered separately. Each fish in a cohort is treated individually in terms of time of hatching, growth and death from fishing or natural causes. By sampling from probability distributions which describe these processes, the life of each individual fish is followed until (i) death occurs or (ii) the age of the fish becomes such that this fish is no longer subject to fishing (justification for this condition is given below). When an individual is identified as caught, the time of capture and its length and age at capture are recorded for later compilation.

Processes Simulated

RECRUITMENT

It is assumed that no fish are caught before all fish from the cohort considered are hatched. This condition is introduced for the convenience of computations and model presentation. The time (T_o) of recruitment of a given cohort to the fishable stock is defined as the earliest moment from which any fish of that cohort has a probability of capture greater than zero. The cohort strength (N_r) is assumed to be a random normal variate characterized by a mean and a standard deviation denoted by N_r and δ_{N_r} , respectively.

The distribution of times of hatching (T_h) is represented by n normal distributions. The age (t_r) of a fish at the time of recruitment is calculated as the difference between T_h and the time of recruitment (T_o). The probability density function (P_h) associated with fish hatching at the moment T_h is, therefore, equal to the weighted sum of the probability density functions of the n normal distributions having means and standard deviations denoted by x_i and σ_i , respectively ($i = 1, \dots, n$). Each component distribution density function is multiplied by p_i (a weighting factor specifying the relative contribution of each component distribution to the cohort). Consequently,

$$P_h(T_h) = \sum_{i=1}^n p_i \left\{ [1/(\sigma_i \sqrt{2\pi})] e^{-(T_h - x_i)^2 / 2\sigma_i^2} \right\} \quad \dots 1)$$

*Two such peaks can often be identified in tropical fish populations (Editors' note).

and

$$\sum_{i=1}^n p_i = 1.$$

To determine T_h for a given fish, a pseudo-random number (s) from the range [0,1] is generated by a computer. The fish is assumed to be associated with distribution k if,

$$\sum_{i=1}^{k-1} p_i < s < \sum_{i=1}^k p_i \quad \dots 2)$$

The age (t_r) at recruitment is then determined by generating a random normal variate from distribution i .

GROWTH

Growth of an individual is assumed to follow a von Bertalanffy growth equation modified for seasonal oscillations (Pauly 1982):

$$l_t = L_\infty \{ 1 - e^{-[K(t - t_o) + (CK/2\pi) \sin 2\pi(t - t_s)]} \} \quad \dots 3)$$

where l_t is the length at age t and L_∞ , K , t_o , C and t_s are parameters. L_∞ , K and t_o are the standard von Bertalanffy growth parameters. C determines the magnitude of seasonal growth oscillations and t_s refers to the age at which conditions are most favorable for growth during the first year of life (usually assumed to correspond to mid-summer). In all model implementations, T_o is assumed to correspond to the middle of winter. Hence

$$t_s = t_r - 0.5 \quad \dots 4)$$

if age is expressed in years. As t_r varies among fish, t_s is also variable. L_∞ and K are assumed to vary randomly among individual fish and their values are sampled from the bivariate normal distribution defined by mean values of these parameters and a variance-covariance matrix. The parameters t_o and C are assumed to take the same values for all fish.

DEATH WHEN FISHING IS NOT SIZE SELECTIVE

Initially we shall assume the instantaneous rate of fishing mortality (F) to be independent of fish size (as well as time and age). When tracing individual fish in the population, the time of death for each fish is a random variable conforming to the constraints imposed by the values of F and the

instantaneous rate of natural mortality (M). This is accomplished using a technique described by Kleijnen (1974), who showed that a value \underline{x} can be sampled from its cumulative probability distribution $f(x)$ by using the inverse of the function $f(f^{-1})$ and a pseudo-random number \underline{r} from the range $[0,1]$ so that

$$\underline{x} = f^{-1}(\underline{r}). \quad \dots 5)$$

The cumulative probability distribution of time of death, \underline{t} , of recruits is denoted by $f(t)$. It can be calculated by the equation (Hearn et al., in press).

$$f(t) = \int_{T_o}^{\underline{t}} (F+M)e^{-(F+M)t'} dt' = 1 - e^{-(F+M)(\underline{t} - T_o)} \quad \dots 6)$$

By applying equation (5), we can put $f(t)$ equal to the pseudo-random number \underline{r} and solve for \underline{t} , i.e.,

$$\underline{r} = 1 - e^{-(F+M)(\underline{t} - T_o)} \quad \dots 7)$$

and

$$\underline{t} = T_o - \ln(1 - \underline{r})/(F+M) \quad \dots 8)$$

Equation (8) can be used to generate a time of death for each fish in the cohort. Note that equation (8) makes no distinction between death from fishing and death from natural causes. The probability (P_f) that a given death is due to fishing is determined by the equation

$$P_f = F/(F+M) \quad \dots 9)$$

The cause of death can be identified by generating a pseudo-random number \underline{d} from the range $[0,1]$. If $\underline{d} < P_f$, the fish died from fishing; if $\underline{d} > P_f$, the fish died from natural causes (Hearn et al., 1987).).

From equation (8) it can be seen that \underline{t} can take values ranging from T_o to ∞ . However, for computational reasons (computer time and storage capacity), it is necessary to specify an upper limit (T_{max}) on \underline{t} which would still encompass the majority (say, 95%) of generated capture times. Then, only fish caught during the period $[T_o, T_{max}]$ are considered in terms of their contribution to catches. As long as a sensible selection of T_{max} is made, this is a very minor condition which could only have a small effect on the extreme upper end of the catch length-frequency distribution.

When all fish of a cohort are treated in this way, a number (C') of them will have been deemed to have been caught. C' is approximately equal to C as given in the equation

$$C = [F/(F+M)]N_r [1 - e^{-(F+M)(T_o - T_{max})}] \quad \dots 10)$$

where N_r is the number of fish alive at time T_o . Differences between C' and C are due only to sampling errors involved in the determination of C' .

DEATH WHEN FISHING IS SIZE SELECTIVE

In the foregoing description of determination of capture time, it was assumed that there was no size selectivity of the fishing gear. However, for many fisheries there is a size-dependent probability of capture following encounter with the fishing gear (Gulland 1983). For example, small fish are more likely than larger fish to escape through the meshes of a trawl. To incorporate size selectivity into the simulation model, some modifications are necessary to the above method of determining t 's. The simulated distribution of t 's should reflect the possibility of (i) fish escaping after encountering the gear and (ii) subsequent encounter(s) with the gear.

Let us redefine F in equation (8) as the instantaneous rate of encounter with the fishing gear. Then, the time (t_1) of first encounter or death from natural causes can be calculated on the basis of the following equation

$$t_1 = T_o - \ln(1 - r_1)/(F+M) \quad \dots 11)$$

The encounter/death-from-natural-causes decision can be reached as described previously. For those cases where t_1 represents a time of encounter, an additional decision must then be made as to whether or not the fish is retained by the gear. The logistic curve is used to model size selection (see Pope 1966). The probability (P_r) of retention, therefore, can be calculated from the equation

$$P_r = 1/[1 + e^{-(a+b l_t)}] \quad \dots 12)$$

where l_t is the length at the time of first encounter (calculated from equation (1)), and a and b are parameters of the logistic curve. The magnitude of P_r in relation to a generated pseudo-random number p from the range [0,1] determines whether or not the fish is retained (retention if $p < P_r$; escapement if $p > P_r$). If the fish is retained, information on its age, length and time at capture is stored as before. In the case of escapement, allowance must be made for the possibility of future encounter. Therefore, the time of next encounter or death due to natural causes is calculated for this individual from the equation

$$t_2 = t_1 + T_o - \ln(1 - r_2)/(F+M) \quad \dots 13)$$

and decisions are again made regarding encounter/death-from-natural-causes and capture/escape. This procedure is repeated until (i) the fish dies from natural causes; (ii) the value of t_j (the time of the j -th encounter/death-from-natural-causes) is greater than T_{max} ; or (iii) the fish is caught.

It is clear that if l_t is such that P_r is smaller than one, the real instantaneous rate of fishing mortality, F' ($F' = FP_r$), of fish of this size will be somewhat smaller than the instantaneous rate of encounter, F . However, as l_t increases, P_r approaches 1 and F' approaches F . This results in the catch number (C') being less than that predicted by equation (10), which is valid only if there is no size selection. Whilst it may be possible to develop modifications of equations (8) and (10) that allow for size selection, their solution would require the use of complex numerical methods that would probably negate any advantages of the analytical formulation. Therefore, this approach is not dealt with in this paper.

Implementation of the Model

The sequence of steps during a run of the simulation model is as follows.

1. Determine the cohort strength (the number of fish at time T_o), N_r .
2. For each of the N_r fish:
 - (i) Determine a time of encounter/natural death, \underline{t} , from equation (8).
 - (ii) If $\underline{t} > T_{max}$, disregard this fish and return to step 2(i) to consider the next fish; otherwise, continue.
 - (iii) Decide whether this individual died from natural causes or encountered the fishing gear. If it died from natural causes, return to step 2(i) to consider the next fish; otherwise, continue.
 - (iv) Determine the age (t_r) at time T_o .
 - (v) Determine the age at encounter ($a = t_r + \underline{t}$).
 - (vi) Determine the growth parameters L_∞ , K and t_s .
 - (vii) Determine the length (l_t) at encounter from equation (3).
 - (viii) Determine the probability (P_r) of retention by the gear from equation (12). (P_r is equal to 1 if no size selectivity is assumed).
 - (ix) Decide if the fish is caught or it it escapes. If it is caught, store a , l_t and \underline{t} and return to step 2(i) to consider the next fish; otherwise, continue.
 - (x) Determine the time of next encounter/natural death from equation (11). Repeat steps 2(ii), (iii), (v) and (vii) to (x) until the individual lives beyond time T_{max} , is caught, or dies from natural causes.
3. Assign caught fish to specified time intervals of capture and construct catch length-frequencies for each time interval for that cohort.
4. Repeat steps 1 to 3 for as many cohorts as required and aggregate individual cohort length-frequencies into a total length frequency.

Potential Extension of the Model

Various extensions of the model described are possible. The mathematical description of recruitment could be extended to include a stock-recruitment relationship. This would require the calculation of the biomass of mature females at the time of spawning. Also, the parameters x_i , σ_i , and p_i 's could be made random variables, changing from cohort to cohort. Density-dependent mechanisms, e.g., in relation to growth parameters, could also be introduced. Changes in M and F with time and a dependence of M upon fish length, age or population abundance and age structure could also possibly be incorporated. Whilst these extensions may add to the realism of the model, they would also add substantially to computer time and memory requirements and, therefore, would only be added for specific simulation requirements.

Example

The following example demonstrates results obtainable using the simulation model. A hypothetical fish population having the following characteristics is considered.

SPAWNING ACTIVITY

Spawning takes place in two activity peaks represented by overlapping normal distributions. The more intense of the two peaks is centered at midsummer and has a standard deviation of one month. Therefore, the mean age at time T_o (midwinter) of fish contributed by this spawning peak (75% of all recruits) is six months. The other peak is centered two months before the first* and hence the mean age at time T_o of fish contributed by this peak (the remaining 25% of all recruits) is eight months. It also has a standard deviation of one month.

COHORT STRENGTH

Ten cohorts are simulated, with the cohort strength (N_r) at time T_o being sampled from a normal distribution with a mean of 10,000 and a standard deviation of 2,000.

GROWTH

The non-variable parameters (t_o and C) of the seasonally oscillating growth curve take values of 0 year and 0.4, respectively. Mean values of L_∞ and K are 50.0 cm and 0.5 year⁻¹, respectively. Their variances are 8.0 cm² and 0.0002 year⁻², respectively, and they have a covariance of -0.01. The parameter t_s depends on t_r , the age at time T_o (see equation (4)).

MORTALITY

The instantaneous rates of encounter and natural mortality are both 0.3 year⁻¹. The fishery is assumed to operate on individual cohorts for five years (i.e., $T_{max} - T_o = 5$ years).

SIZE SELECTIVITY

The parameters a and b in equation (12) take values of -10 and 0.67 cm⁻¹, respectively. This means that few fish less than 10 cm long are caught, while almost all of those longer than 20 cm are caught if they encounter the fishing gear.

Information on fish length, age and time of capture for ten cohorts is generated and aggregated into monthly periods.

Cohort strengths for the ten cohorts are shown in Table 1. Note the relative weakness of cohort 4.

Length-frequency histograms of fish caught during 13 successive one-month periods are presented in Fig. 1. Each of the first five cohorts simulated has contributed to catches in the first 12 one-month periods presented in Fig. 1, i.e., cohorts 1 to 5 provide age-classes 4 (fish aged 4 to 5 years) to 0, respectively. In the thirteenth period, new recruits have been introduced into the fishable population from cohort 6, fish from cohorts 2 to 5 have graduated to the next age-class and cohort 1 has left the fishable population. The effects of size selectivity, seasonally oscillating growth and the weakness of cohort 4 (represented by age-class 1 in periods 1 to 12 and age-class 2 in period 13) are very apparent. There is also some evidence of the bimodal spawning activity assumed in the simulations. Although age-class 0 has a well-formed mode and age-class 1 generally can be discerned, it is clearly impossible to separate by eye the remainder of the distribution into age-classes 2, 3 and 4.

*This two-month period separating the two spawning peaks is not intended to simulate a realistic situation, in which the period between two peaks of activity would usually be between five and seven months (Editors' note).

Table 1. Cohort strength for ten simulated cohorts.

Cohort number	Cohort strength
1	11,177
2	11,478
3	9,764
4	5,647
5	11,414
6	12,368
7	9,254
8	10,976
9	11,541
10	10,606

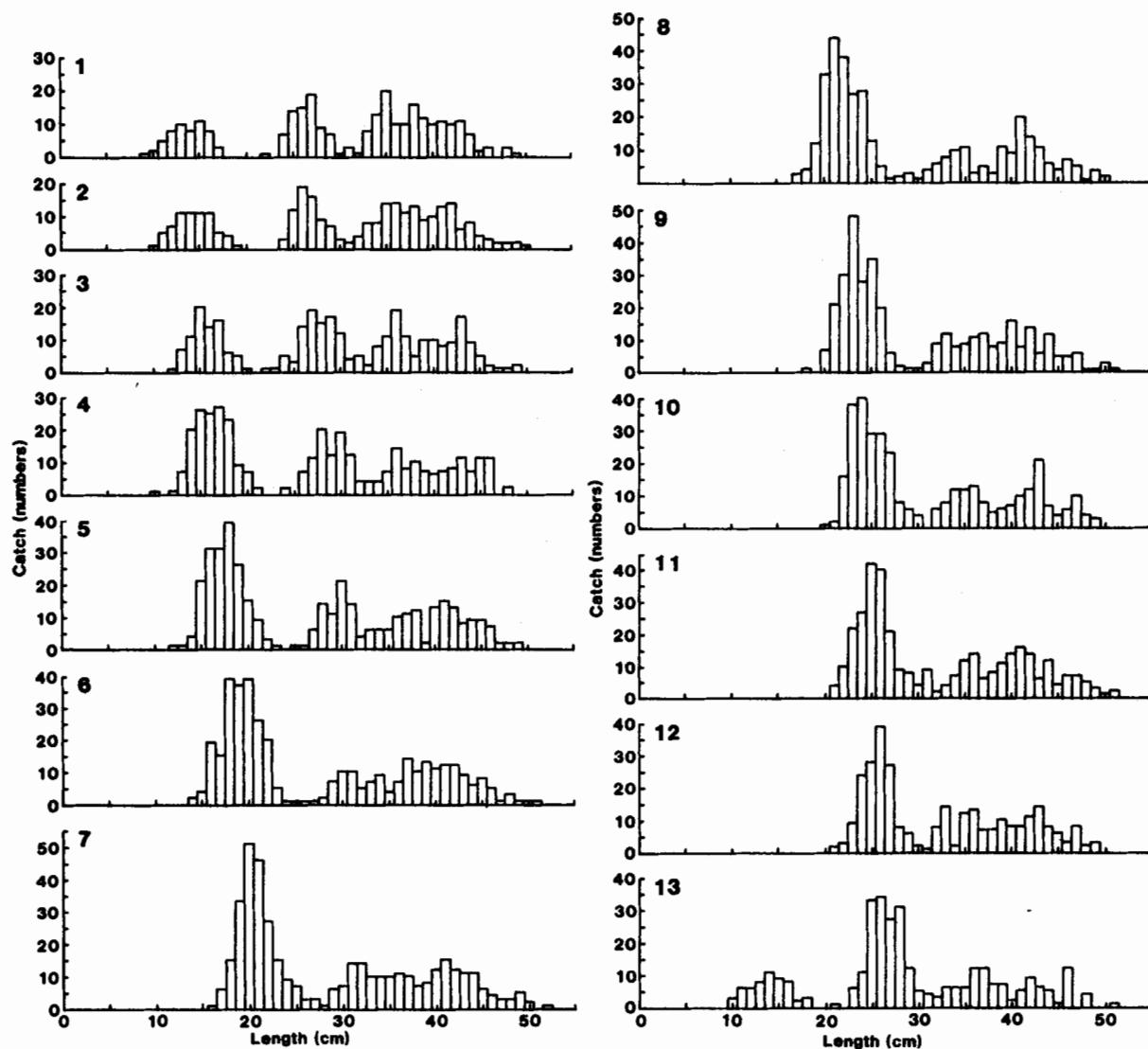


Fig. 1. Example of monthly length-frequency "samples", generated as described in the text.

Because of the nature of the simulation model presented, it is necessary to consider sufficiently large numbers of (i) fish in each cohort at the time of recruitment to the fishable stock and (ii) fish caught during each period for which length-frequencies are constructed. This is to ensure that errors associated with sampling during the implementation of the model do not change the assumed characteristics of the hypothetical fish population and associated fishery. To verify this condition, the characteristics of the simulated population and fishery have to be assessed.

Fig. 2 shows a plot of the natural logarithm of catch number against period of capture for cohort 1. Full recruitment to the fishery occurs after about the seventh month. The application of ordinary least squares regression to the data for months 8 to 60 gives the instantaneous rate of total mortality equal to 0.593 year^{-1} , a value very close to that assumed. In a similar way, other characteristics of the simulated population and associated fishery can be examined. Spawning activity, individual variation in growth and size selectivity of the fishing gear all show characteristics in the simulated population and associated fishery consistent with the assumptions of the simulation model.

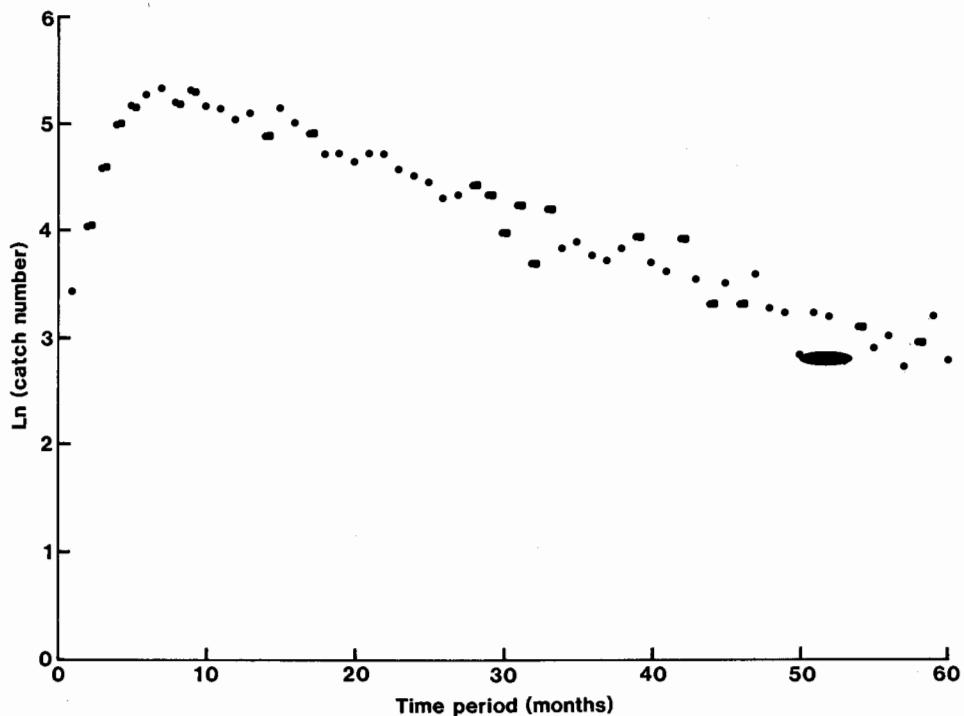


Fig. 2. Natural logarithm of catch numbers against period of capture for a simulated cohort of fish (see text).

Acknowledgements

Drs. John Caddy (Food and Agriculture Organization, United Nations, Rome, Italy) and John Gulland (Imperial College, London, England) suggested the development of the model described in this paper. Partial financial support for the project was provided by the Food and Agriculture Organization of the United Nations. Mr. William Hearn (CSIRO Division of Fisheries Research, Hobart, Tasmania, Australia) and Dr. Ron Sandland (CSIRO Division of Mathematics and Statistics, Lindfield, New South Wales, Australia) reviewed the mathematical procedures in the model and made helpful suggestions.

References

- Gulland, J.A. 1983. Fish stock assessment: a manual of basic methods. FAO/Wiley Series on Food and Agriculture.
- Hearn, W.S., R.L. Sandland and J. Hampton. 1987. Robust estimation of natural mortality in a completed tagging experiment with variable fishing intensity. *J. Cons., Cons. Int. Explor. Mer* 43(2): 107-117.
- Jones, R. 1981. The use of length composition data in fish stock assessments (with notes on VPA and cohort analysis). FAO Fish. Cric. No. 734. 55 p.
- Kleijnen, J.P.C. 1974. Statistical techniques in simulation. Marcel Dekker, New York.
- Pauly, D. 1980. A selection of simple methods for the assessment of tropical fish stocks. FAO Fish. Circ. No. 729. 54 p.
- Pauly, D. 1982. Studying single-species dynamics in a tropical multi-species context, p. 33-70. In D. Pauly and G.I. Murphy (eds.) Theory and management of tropical fisheries. ICLARM Conference Proceedings 9, 360 p.
- International Center for Living Aquatic Resources Management, Manila, Philippines and Division of Fisheries Research, Commonwealth Scientific and Industrial Research Organisation, Cronulla, Australia.
- Pauly, D. and N. David. 1981. ELEFAN I, a BASIC program for the objective extraction of growth parameters from length-frequency data. *Meeresforsch.* 28(4): 205-211.
- Pope, J.A. 1966. Manual of methods for fish stock assessment. Part III. Selectivity of fishing gear. FAO Fish. Tech. Pap. No. 41. 41 p.
- Schnute, J. and D. Fournier. 1980. A new approach to length-frequency analysis: growth structure. *Can. J. Fish. Aquat. Sci.* 37: 1337-1351.

An Examination of the Reliability of the ELEFAN Computer Programs for Length-Based Stock Assessment

J. HAMPTON
J. MAJKOWSKI

*Division of Fisheries Research, Marine Laboratories
Commonwealth Scientific and Industrial Research Organisation
GPO Box 1538, Hobart, Tasmania 7001, Australia*

Hampton, J. and J. Majkowski. 1987. An examination of the reliability of the ELEFAN computer programs for length-based stock assessment, p. 203-216. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

Catch length-frequency data generated by a computer simulation model are analyzed by the ELEFAN programs and the resulting parameter estimates compared with their true values (i.e., those assumed in the simulation model). Four experiments were devised and ten replicate data sets (with differences caused by random sampling only) produced for each. In experiment 1, the data are characteristic of a fishery in which exploitation is size selective and there is relatively high variation in the von Bertalanffy growth parameters, L_∞ and K. In experiment 2, the data from experiment 1 are truncated such that the largest length-class is one length-class interval less than the true value of L_∞ . In experiment 3, fishing is size selective, but the individual variation in growth parameters is relatively low. In experiment 4, data characteristic of a fishery with low individual variation in growth parameters and size-independent fishing are analyzed. In all experiments, recruitment is assumed to occur in two activity peaks per year which are approximated by normal distributions. The number of recruits per year is assumed to be a normal random variable.

ELEFAN I overestimates L_∞ by 11-23%, and underestimates K by 16-36% in experiments where fishing is size selective (1 to 3). In experiment 4, where fishing is size-independent and growth parameter variation is low, these estimates are almost exactly correct. In experiments 2 and 3, where no samples included fish larger than the true mean value of L_∞ , the biases are reduced by employing a procedure that corrects the data for size-selection. This procedure is ineffective in experiment 1 because of the presence in the samples of fish larger than the true mean value of L_∞ .

The two estimates of the instantaneous rate of total mortality (Z) given by ELEFAN II are both positively biased in experiments 1 and 2 (by about 20% on average) but are generally quite accurate in experiments 3 and 4. Variation in year-class strength is shown to be a serious potential source of error in estimates of Z. Patterns of recruitment derived by ELEFAN II slightly overestimate the temporal spread of recruitment and patterns of selection derived by the same program may underestimate the true probabilities of retention by length-class.

ELEFAN III gives reasonable estimates of average recruitment, with the observed bias (about 10% on average for experiments 1 and 2) being mainly due to the bias in the previously derived Z estimates. However, the instantaneous rate of fishing mortality (averaged over the fully recruited length classes) is overestimated by as much as 100% in some cases and appears to be quite sensitive to the selection of the terminal length-class.

Introduction

The management of fisheries in many countries is handicapped by a lack of information necessary for applying standard stock-assessment models, inadequate manpower for data collection and analysis and limited access to modern, high-speed computers. In an attempt to circumvent these problems, a package of computer programs named ELEFAN (*Electronic LEngh Frequency ANalysis*) was developed (Pauly, Part I, this vol.). These programs are written in BASIC for operation on small, low-cost microcomputers. The only input data required are relatively easy-to-collect catch length frequencies, preferably in a time series (e.g., samples for each month of a year).

The version of the ELEFAN package tested by the authors consists of four main programs. ELEFAN 0 (David et al., unpublished data) is a program for creating and editing data files for input to the ELEFAN I, II and III programs. It is not discussed in this paper.

ELEFAN I (Pauly and David 1981; Pauly, Part I, this vol.) is a program that enables von Bertalanffy growth parameters to be extracted from a set of length-frequency samples arranged sequentially in time. Under certain assumptions, the asymptotic length (L_∞), the growth constant (K), the coefficient reflecting the magnitude of seasonal growth oscillations (C) and the parameter indicating the time of year at which growth is slowest (the so-called winter point, WP) may be estimated.

Given estimates of L_∞ , K, C and WP and compliance with various assumptions, ELEFAN II (Pauly 1982 and Part I, this vol.) evaluates the instantaneous rate of total mortality (Z). Two estimates of Z are provided by this program, one based on length-converted catch curves (see Pauly, Part I, this vol.) which are similar to the catch-curve method of Robson and Chapman (1961) and the other based on the mean length of fully recruited size-classes (Beverton and Holt 1956). Given an independent estimate of the instantaneous rate of natural mortality (M), ELEFAN II can also be used to derive fishing gear selection patterns by backwards projection of the descending, right limb of the catch curve. Recruitment patterns may be estimated using ELEFAN II by backwards projection of the length-frequency data on the time axis. This procedure, when used with an independent estimate of the parameter t_0 , provides a graphic representation of the monthly fluctuation into the stock under investigation; when reliable estimates of t_0 are not available, the method can still be used to show seasonal changes in recruitment. However, no information is then provided as to *when* recruitment occurs.

ELEFAN III (Pope et al., unpublished data; Pauly and Tsukayama 1983) consists of three routines, each being a form of virtual population analysis (VPA). VPA 1 is a conventional age-structured VPA (Pope 1972). VPA 2 is a version of Jones' (1979, 1981) length-cohort analysis (Pauly 1984, Chapter 7). Each of these routines requires independent estimates of M and the instantaneous rate of fishing mortality (F) upon the last age or length group represented in the samples. This latter quantity is usually referred to as "terminal F" and has the symbol F_t .

The availability of the ELEFAN programs for use in stock assessment has attracted the interest of many workers in this field (Pauly, Part I, this vol.). However, to establish whether the programs can be used with confidence, their accuracy must be assessed and possible biases identified. For this purpose, a computer simulation model for generating catch length-frequency data for a hypothetical population and associated fishery was developed (Hampton and Majkowski, Part I, this vol.). By selecting values for input parameters to this model, catch length-frequency data can be generated for a population and fishery of known characteristics. These data can then be used as input information for the ELEFAN programs and the derived parameter estimates compared with their true values (i.e., those used in the simulation model). By changing the seed values for the various number generating routines (for random or normal deviates), a number of independent data sets associated with identical population and fishery characteristics can be generated and the effects of sampling errors evaluated.

In the descriptions of the simulation experiments that follow, it is assumed that the reader has a working knowledge of the ELEFAN programs and is familiar with the terms used in their documentation. It is also assumed that the description of the catch length-frequency simulation model (Hampton and Majkowski, Part I, this vol.) has been read.

In this paper, the reliability of a number of the more important procedures which are parts of the ELEFAN system are examined. There is no attempt to investigate the efficiency (from a programming point of view) of the computer programs themselves; concern is solely with the models and the assumptions underlying them.

General Characteristics of the Simulation Model

For full details of the catch length-frequency simulation model, Hampton and Majkowski (Part I, this vol.) should be consulted; only a brief outline of the model is presented here.

Cohorts are treated individually. The life of each recruit to the fishable stock is traced from the time of recruitment to the time of death (from fishing or natural causes) or until the fish is assumed to no longer be a potential contributor to the catch. The processes simulated are: recruitment (which is influenced by a relatively long spawning period with two peaks of activity and variable cohort strength); growth (which is seasonal and variable amongst individuals); time-, age- and size-independent instantaneous rates of natural mortality and encounter with the fishing gear; and size-selective exploitation resulting from a size-dependent rate of escape from the fishing gear. The time of death of an individual is a random variable subject to the constraints imposed by the instantaneous rates of encounter and natural mortality and the process of size selection.

Methods

THE SIMULATED POPULATION AND FISHERY

For the simulation experiments described below, a standard population and fishery was considered, as follows:

Spawning takes place in two activity peaks represented by overlapping normal distributions. The more intense of the two peaks (75% of all recruits) is centered at midsummer and has a standard deviation of one month. Recruitment of an entire year-class to the fishery (i.e., the moment when the probability of capture for each fish from that year-class is greater than zero) is assumed to be instantaneous at midwinter; therefore, the mean age at recruitment of fish contributed by this spawning peak is six months. A minor peak is centered two months earlier than the major peak; hence, the mean age at recruitment of fish contributed by this peak (the remaining 25% of recruits) is eight months. It also has a standard deviation of one month. Cohort strength (total number of recruits) is sampled from a normal distribution with a mean of 10,000 and a standard deviation of 2,000.

A seasonally oscillating von Bertalanffy growth curve (Pauly 1982) is assumed to describe the growth, in length, of individual fish. All fish have individually determined values of L_{∞} and K. These are sampled from a bivariate normal distribution having means of 50 cm and 0.50 year⁻¹, respectively. Two levels of variation in L_{∞} and K are tested; they correspond to coefficients of variation (C.V.s) of 10% and 2% for both parameters. L_{∞} and K are assumed to be correlated ($r = -0.8$).

The instantaneous rates of encounter and natural mortality are both set at 0.5 year⁻¹. In simulations where size selection is assumed not to occur, the instantaneous rate of encounter is equal to F. Where size selection is assumed, it is described by a logistic curve with the a and b parameters equal to -10 and 0.67 cm⁻¹, respectively (see equation 12 in Hampton and Majkowski, Part I, this vol.). This effectively results in few fish smaller than 10 cm being caught, while causing almost all of those longer than 20 cm to be caught if they encounter the fishing gear. These processes are used to define a probability density function from which the time of death by fishing or natural causes for each fish may be determined by random sampling. For computational simplicity, the maximum time from recruitment that fish are assumed to contribute to the fishery is set at five years. Less

than 1% of the initial recruits would remain alive after five years with the above mortality levels. This, however, does slightly reduce the probability of very large fish being caught. The possible effect of this is discussed in later sections.

DATA GENERATED BY THE SIMULATION MODEL

The length-frequency data generated by the simulation model for analysis by the ELEFAN programs have the following features:

1. Fish contribute to the fishery for a maximum of five years.
2. Caught fish are grouped into 2 cm length-classes. This results in the total number of length-classes being 20 to 25.
3. Six year-classes are generated and combined such that there are two complete years of catch data, each comprising five age-classes.
4. Fish are aggregated into 24 monthly samples according to their dates of capture. This is the maximum number of samples that can be analyzed by the version of ELEFAN used in this study.

ELEFAN TESTING PROCEDURES

Two aspects of fish population dynamics and a data treatment procedure that could affect the performance of the ELEFAN programs were considered.

1. The magnitude of growth parameter variation for tropical fish stocks has never, to these authors' knowledge, been precisely quantified. However, Sainsbury (1980) has shown that the variation in L_{∞} may be approximated by the variation in length-at-age of very old fish. As can be seen from data in Ingles and Pauly (1984), this variation may be considerable for many tropical species. Simulated length-frequency data with two arbitrary levels of growth parameter (L_{∞} and K) variation were tested to assess the effect of this variation on ELEFAN performance.
2. Simulated length-frequency data were routinely generated on the assumption that fishing is size-selective, i.e., F is size-dependent. To see whether or not this affects the performance of the ELEFAN programs, simulated length-frequency data generated assuming size-independent fishing were also analyzed.
3. Because the higher level of growth parameter variation results in fish with lengths greater than the mean L_{∞} in the length-frequency samples, the effect of truncating the samples at one length-class below the mean L_{∞} was investigated. In data so treated, the midpoint of the largest length-class becomes 47.5 cm. This is done to remove large fish as a source of bias in the estimation of L_{∞} . Although this procedure is artificial and would be difficult to implement in reality, these authors wanted to see whether or not it is possible for ELEFAN I to arrive at the correct estimate of L_{∞} in the absence of this biasing effect.

Using different seed values for the number-generating routines in the length-frequency simulator, ten data sets were produced for each of the four experiments. The characteristics of the data for these experiments are summarized in Table 1.

The ELEFAN package has facilities to provide a variety of population parameter estimates. In this paper, we have restricted ourselves to examining the reliability of the following procedures:

ELEFAN I—Estimation of L_{∞} and K (ten replicates per experiment);

ELEFAN II—Estimation of Z (ten replicates per experiment), the pattern of recruitment (one replicate for experiments 1, 3 and 4) and the pattern of size selection (one replicate for experiments 1 and 3);

ELEFAN III—Estimation of steady-state recruitment and average F using VPA 2 (ten replicates per experiment).

Table 1. Characteristics of the simulated data produced for the four experiments.

Characteristic	Experiment			
	1	2	3	4
C.V.s of L_{∞} and K	10%	10%	2%	2%
Samples truncated	No	Yes	No	No
Size-selective fishing	Yes	Yes	Yes	No

EXECUTING THE ELEFAN PROGRAMS

The ELEFAN programs were installed and executed on a Digital Rainbow 100 microcomputer. The procedures were as follows:

ELEFAN I—Data sets consisting of 24 length-frequency samples were analyzed. For all replicates of each experiment, the true values of C (0.4) and WP (0) were used. In order to identify the L_{∞} and K combination resulting in the maximum ESP/ASP ratio for a particular set of data, a systematic searching technique was adopted. Values of L_{∞} ranging from the lowest possible value to 65 cm were searched with a step size of 0.5 cm. For each value of L_{∞} searched, the best (i.e., that resulting in the maximum ESP/ASP ratio for that value of L_{∞}) K value was determined. K values were searched in the range 0.30 to 0.60 year⁻¹ with a step size of 0.01 year⁻¹.

ELEFAN II—Using the optimum values of L_{∞} and K as derived above, estimates of Z based on the mean length of fully recruited caught fish (Z_m) and a length-converted catch curve (Z_r) were obtained for the 10 replicates of each experiment. All 24 length-frequency samples were used. For Z_r estimates, the first fully recruited length-class is assumed (as recommended by Pauly (1983 and 1984, Chapter 5) to be that immediately to the right (on the catch curve) of the length-class containing the highest catch number. Because of the variability of catch numbers in the largest length-classes and its pronounced effect on the Z_r estimate, the largest length-class included in the regression was 44.5-46.5 cm. All fully recruited length-classes were included for the calculation of the Z_m estimate. The pattern of recruitment was determined for the first replicate of experiments 1, 3 and 4, assuming a value for t_0 of zero (its true value). The pattern of size-selection was determined for the first replicate of experiments 1 and 3 assuming a value of M equal to half the derived value of Z_r .

ELEFAN III—The VPA 2 routine is employed to obtain estimates of steady-state population numbers and F by length-class. The derived estimates of L_{∞} , K and Z_r and the 24 length-frequency samples (assumed to represent the entire catch) were used as input to VPA 2. A direct comparison of actual and estimated steady-state recruitment is not possible here because recruitment in the simulation model is a discrete event with respect to time. The known number of recruits is not readily comparable with the estimates of population number by length-class given by VPA 2. As an approximation, the population number of the length-class which includes the actual mean length of recruits (length class 10.5-12.5 cm; actual mean length of recruits was 11.87 cm) was compared with the known number of recruits given by the simulation model. We assumed that both M and F_t were equal to half the calculated value of Z_r . The true F was compared with the estimated average F for fully recruited length-classes (the same length-classes as used for the calculation of Z_r were included in the average).

Results and Discussion

ELEFAN I

The estimates of L_{∞} and K derived using ELEFAN I are given in Table 2. In experiment 1, L_{∞} is substantially overestimated (by 14% on average), while K is substantially underestimated (by 22% on average).

A critical assumption made in the ELEFAN I program is that "all fishes in the samples have the same length at the same age, and therefore, differences in length can be attributed to differences in age" (Pauly and David 1981, p. 209). However, as Pauly and David (1981) recognized, fishes of identical age do vary in length. The original derivation of the von Bertalanffy growth model considered only the growth of an individual (von Bertalanffy 1938), although it is commonly applied to describe the mean growth of a group of animals (Sainsbury 1980). When this is done, the growth parameters are usually assumed to show variation between individual animals, so the estimates, whether derived from length-at-age, length-increment or length-frequency data, represent the *mean* values of these parameters for the group of animals being considered.

Table 2. Estimates of L_{∞} (cm) and K (year^{-1}) derived using ELEFAN I. Their real values (input to the simulation model) are 50.0 cm and 0.50 year^{-1} , respectively.

Replicate	Experiment							
	1		2		3		4	
	L_{∞}	K	L_{∞}	K	L_{∞}	K	L_{∞}	K
1	56.0	0.40	56.0	0.40	55.0	0.42	50.5	0.51
2	57.5	0.36	56.0	0.40	54.5	0.41	51.5	0.47
3	55.5	0.39	56.5	0.41	54.5	0.40	50.5	0.51
4	55.5	0.40	56.0	0.39	55.5	0.39	50.0	0.51
5	57.5	0.41	55.0	0.39	55.5	0.39	50.5	0.51
6	57.5	0.39	56.0	0.41	56.0	0.39	50.0	0.51
7	55.5	0.40	55.5	0.40	56.0	0.39	50.5	0.51
8	55.5	0.41	55.5	0.41	56.0	0.39	50.5	0.50
9	61.5	0.32	57.5	0.40	55.5	0.38	51.5	0.47
10	57.5	0.37	56.0	0.40	55.5	0.39	50.0	0.52
Mean	57.0	0.39	56.0	0.40	55.5	0.40	50.5	0.50

However, in ELEFAN I, L_{∞} , K, C and WP are assumed not to vary between individuals. Depending on sample size, the amount of individual variation in L_{∞} and the levels of F and M, any representative set of length-frequency samples is likely to contain some fish that are larger than the mean L_{∞} of the population. Because individual variation in growth is a feature of the simulation model, the length-frequency samples generated in experiment 1 (where 10% C.V.s for L_{∞} and K are assumed) include fish 3-10 cm longer than the true mean L_{∞} of 50 cm. However, these authors' version of ELEFAN I requires that the computed L_{∞} be at least one length-class interval above the midpoint of the length-class containing the largest fish in the set of samples. This requirement alone would result in the biased L_{∞} estimates in experiment 1. (It is possible, in fact, that this bias is slightly underestimated in experiment 1 because only fish up to five years after recruitment were considered.)

To overcome this problem (if only artificially), experiment 2 was devised to truncate the length frequencies from experiment 1 such that the midpoint of the largest length-class considered was at least one length interval smaller than the true mean L_{∞} . This produced little improvement in the L_{∞} and K estimates (Table 2). The reduction of the C.V.s of L_{∞} and K from 10% to 2% (Table

2, experiment 3) also had little effect. In contrast, the estimates of L_{∞} and K in experiment 4 (Table 2) were almost perfect. The only difference between experiments 3 and 4 is the operation of size-selective fishing in experiment 3 and size-independent fishing in experiment 4. As explained by Pauly and Sann Aung (1984) and Pauly (Part I, this vol.), size selectivity favoring the capture of larger fish should cause a negative bias in estimated K and a positive bias in L_{∞} (because the gear initially selects the faster growing fish from the new recruits). This means that the midpoint of the first peak is somewhat larger (in length) than it would have been without size selection. In succeeding time periods, the effect of gear selection becomes less and less; hence, the slow growers will now also be represented in the catch. Therefore, the change in length over time of the midpoint of the first peak will be less when size selection operates. Clearly, this will cause K to be underestimated, and, because they are negatively correlated, L_{∞} to be overestimated.

Pauly and Sann Aung (1984) and Pauly (Part I, this vol.) describe a method for correction of the effects of size selection. The L_{∞} and K estimates obtained in an initial run on the data are used to derive a length-converted catch curve from which the probabilities of capture by length-class are determined using ELEFAN II. These probabilities are used to correct the original length-frequency data and a new set of growth parameters is derived. In the limited number of trials undertaken, we found that this procedure can, in some circumstances, result in more accurate estimates of L_{∞} and K. The results of applying the procedure to replicate 1 of experiments 1, 2 and 3 are given in Table 3. For experiment 1, only a very slight improvement in the estimates was obtained because a lower limit (55.5 cm) was imposed upon L_{∞} by the largest length class in the samples (midpoint 53.5 cm). This not only restricted the estimate of L_{∞} , but also, because of its high negative correlation with K, reduced the estimate of that parameter. However, in experiment 2 the length-frequency data used in experiment 1 were truncated (midpoint of the largest length-class considered was 47.5 cm), consequently allowing more improvement in the L_{∞} and K estimates. Similarly in experiment 3, where growth parameter variation is low (the largest length-class had a midpoint of 47.5 cm), more accurate estimates of these parameters were obtained.

These results suggest that the application of the correction procedure described by Pauly and Sann Aung (1984) and Pauly (Part I, this vol.) may result in more reliable estimates of L_{∞} and K so long as the selection pattern can be accurately estimated (see section below) and the length-frequency samples exclude (either naturally or by truncation) fish of lengths greater than the true value of mean L_{∞} . The difficulty in a real situation is, of course, that one does not know if these conditions are satisfied. The selection of an appropriate length for truncation is a problem for which there is no easy solution. A useful procedure may be to exclude successive length-classes from the analysis and see whether the L_{∞} and K estimates stabilize. This procedure was undertaken for the data, corrected for size selection, in replicate 1 of experiment 1 (Table 4). In this case, there was little change in the estimates after the exclusion of the largest length-class (midpoint 53.5 cm). This indicated that, for this data set: (1) truncation of the data did not adversely affect the growth parameter estimates and (2) growth parameters can be obtained that, almost certainly, are not biased by the presence of large fish in the samples.

Table 3. A comparison of L_{∞} (cm) and K (year^{-1}) estimates for replicate 1 of experiments 1, 2 and 3 before and after correction of length-frequency data for size selection.

Experiment	Before correction		After correction	
	L_{∞}	K	L_{∞}	K
1	56.0	0.40	55.5	0.43
2	56.0	0.40	53.5	0.46
3	55.0	0.42	52.0	0.47

Table 4. Estimates of L_{∞} (cm) and K (year^{-1}) derived using ELEFAN I for data truncated at different points. The data, for replicate 1 of experiment 1, have been corrected for the effects of size selection.

Midpoint of the largest length-class	L_{∞}	K
53.5 (all data)	55.5	0.43
51.5	53.5	0.46
49.5	53.5	0.46
47.5	53.5	0.46
45.5	55.5	0.46
43.5	55.5	0.46
41.5	53.5	0.48
39.5	53.0	0.47

An aspect of ELEFAN I that has not been dealt with specifically in this paper, but which deserves mention, is the certainty with which ELEFAN I might arrive at the maximum ESP/ASP ratio for a particular set of data. It was clear from our use of the program that the maximum ESP/ASP could not always be identified from a single ELEFAN I run. The L_∞ and K values giving the maximum ESP/ASP can only be verified by examining the ESP/ASP response surface over a suitable range of L_∞ and K. Fig. 1 shows the response of the ESP/ASP ratio in replicate 1 of experiment 1 to different L_∞ and K. For clarity, only three values of L_∞ are considered in Fig. 1; a wider range of values is recommended in practice. There are clear maxima in the ESP/ASP ratio for specific values of L_∞ , but the overall maximum (i.e., when different values of L_∞ are considered) is less distinct. It is strongly recommended that users of ELEFAN I examine response curves, such as those shown in Fig. 1, over a suitable range of both K and L_∞ .

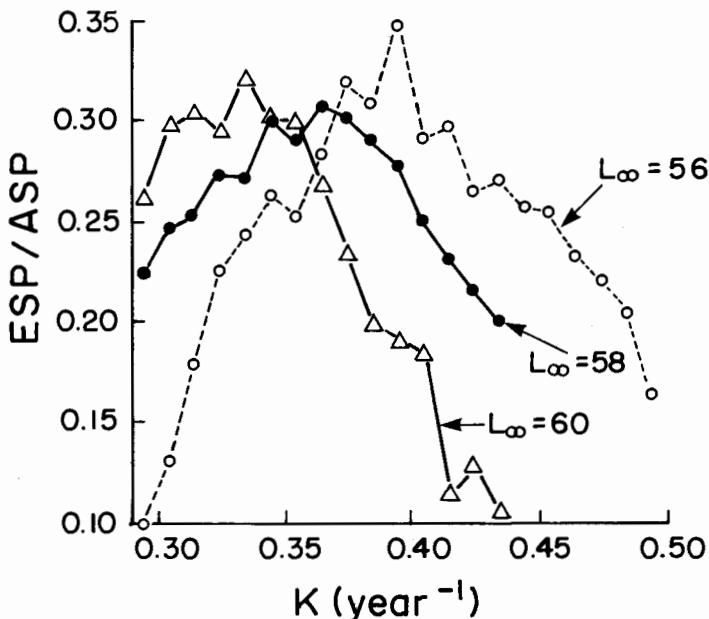


Fig. 1. ESP/ASP response to different values of K and L_∞ in replicate 1 of experiment 1.

ELEFAN II

The estimates of Z, derived using ELEFAN II, for the ten replicates of experiments 1 to 4 are given in Table 5. The estimates based on a regression of the fully recruited segment of the length-converted catch curve (Z_r) showed a strong positive bias in experiments 1 and 2 (+19% and +20% on average, respectively). The Z_r estimates for experiments 3 and 4, on the other hand, were generally quite accurate. The estimates derived on the basis of the mean length of fully recruited length-classes (Z_m) performed similarly to Z_r in each experiment.

The much better performance of both estimators in experiments 3 and 4 suggests that the higher growth parameter variation in experiments 1 and 2 was the cause of the large positive bias in those experiments. There is no obvious reason why this should be so; further work is required to resolve this question.

An important assumption of the ELEFAN II procedure and indeed of all catch curve or mean length-based estimation of Z is that cohort strength is constant both before and during the period over which the fish are sampled, or varies only randomly. In applying ELEFAN II, or any similar method, to estimate Z, one needs to be wary of non-compliance with this assumption. Even if cohort strength varies only randomly, problems may still arise, particularly if few cohorts are represented in the samples. As an example, consider replicate 7 of experiment 2. As for all the replicates, 5 age-classes are present in any given sample. Since the data analyzed consist of 2 years (24 monthly samples) of length-frequencies, the total number of year-classes represented in the data

Table 5. Estimates Z (year^{-1}) using the regression (Z_r) and mean length (Z_m) routines of ELEFAN II. The corresponding estimates of L_∞ and K derived by ELEFAN I (Table 2) are used in the ELEFAN II procedures. The real value of Z input to the simulation model is 1.0 year^{-1} .

Replicate	Experiment							
	1		2		3		4	
	Z_r	Z_m	Z_r	Z_m	Z_r	Z_m	Z_r	Z_m
1	1.15	1.14	1.15	1.17	1.03	1.14	1.02	1.02
2	1.16	1.17	1.19	1.23	1.09	1.18	0.93	0.89
3	1.07	1.08	1.16	1.21	1.02	1.06	1.19	1.11
4	1.12	1.12	1.11	1.16	0.99	1.16	1.00	0.93
5	1.34	1.30	1.16	1.13	1.00	1.10	1.06	1.00
6	1.26	1.24	1.25	1.26	1.10	1.18	0.83	0.83
7	1.30	1.31	1.30	1.35	0.91	1.07	0.97	1.00
8	1.19	1.20	1.19	1.24	1.06	1.16	1.05	0.97
9	1.24	1.27	1.34	1.40	1.03	1.10	0.99	0.95
10	1.11	1.11	1.14	1.16	0.98	1.10	1.07	1.03
Mean	1.19	1.19	1.20	1.23	1.02	1.13	1.01	0.97

is six. The number of recruits for each of these year-classes and the age-classes they represent in the 24 samples is shown in Table 6. By chance, 2 of the year-classes (1 and 3) from which the older fish in the samples are derived are very weak. In contrast, those that make up the younger fish in the samples (year-classes 4, 5 and 6) are strong. This would have the effect of positively biasing both estimators of Z . In Table 5, we see that this is indeed the case, with estimates of 1.30 and 1.35 year^{-1} for Z_r and Z_m , respectively.

To confirm that this bias has been at least partially caused by the chance sequence of year-class strengths, we reversed the order of the year-classes, i.e., the weak year-classes now contribute the younger fish in the samples and the strong year-classes contribute the older fish. The analysis was repeated (using the same L_∞ and K as before) and produced Z_r and Z_m estimates of 0.98 and 1.09 year^{-1} , respectively. This indicates that the error introduced by variable recruitment may be considerable. However, recruitment variability alone cannot fully account for the observed bias in the original total mortality estimates (otherwise, the above re-estimates should have been considerably less than 1.00 year^{-1}).

Table 6. Details of the year-classes occurring in the samples associated with replicate 7 of experiment 2. The age-classes represented in the samples are denoted by the number of completed years since recruitment.

Year-class no.	No. of recruits	Age-classes represented	
		Samples 1-12	Samples 13-24
1	7,650	4	—
2	11,523	3	4
3	5,877	2	3
4	11,443	1	2
5	13,200	0	1
6	12,407	—	0

Given an independent estimate of t_o , ELEFAN II may also be used to estimate temporal recruitment patterns by projecting the length-frequency data backwards in time using the growth parameters derived by ELEFAN I. This procedure was undertaken for replicate 1 of experiments 1, 3 and 4, the results of which, along with the actual pattern as derived from the simulation model, are

presented in Fig. 2 as frequency histograms of time (month) at age zero. Note that the patterns determined by ELEFAN II overestimate the spread of the recruitment period as determined in the simulation model. This occurs because ELEFAN II attributes all variation in length to variation in age, whereas in the simulated data, some of the variation in length is due to variation in growth (as would be the case in reality). Therefore, the variation in age at a particular time, or vice versa, is overestimated by ELEFAN II. Also, any error in the assumed value of t_0 will cause identical errors in the times of recruitment of all fish.

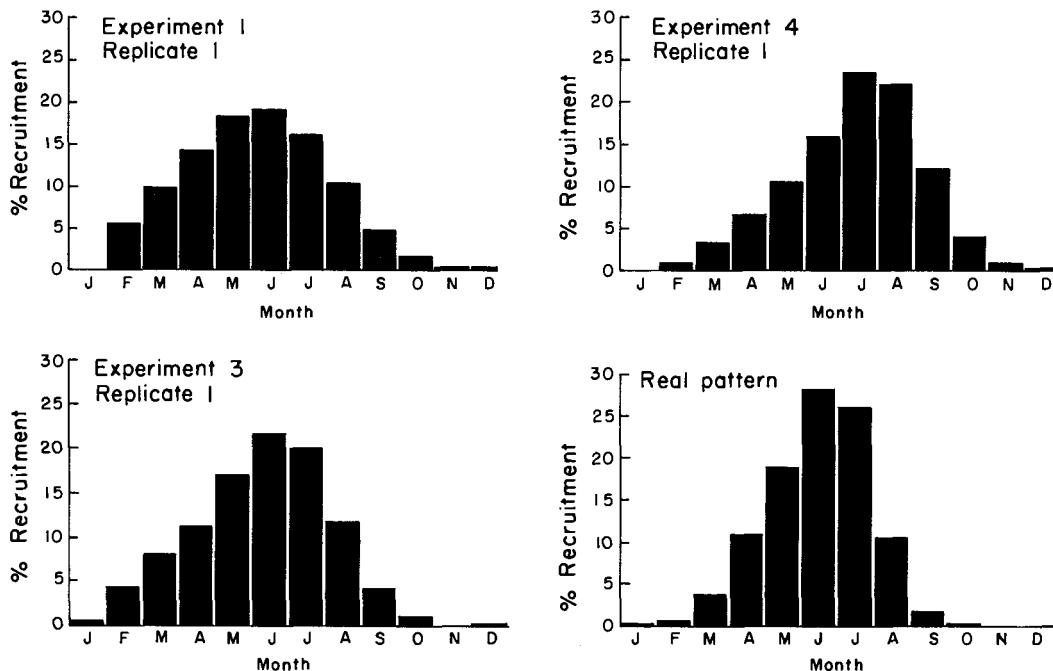


Fig. 2. Pattern of recruitment estimated by ELEFAN II for replicate 1 of experiments 1, 3 and 4. The pattern derived from the simulation model is given for comparison.

ELEFAN II is also able to provide estimates of the pattern of size selection by the fishing gear (i.e., approximate probabilities of retention, by length-class) by projecting backwards the descending limb of the catch curve and calculating the difference between observed and expected catches (see Pauly, Part I, this vol.). This procedure is subject to two main sources of possible error. The first arises from the assumption regarding the mode of recruitment to the fishery. The ELEFAN system assumes that fish are recruited by length. This means that, if there is no size selection, the smallest fish in the catch would be the most numerous. The alternative to recruitment by length is recruitment at a specific time. Here, recruitment could be triggered by a change in some environmental factor or the lifting of a management regulation such as a closed season. In such a case, the lengths of the new recruits would be variable because of differences in times of hatching, and therefore, the length-frequency of the recruits would appear as a peak, with the abundance of the very smallest fish in the catch tailing off to zero. This latter scheme was assumed in the simulation model (Hampton and Majkowski, Part I, this vol.).

This type of recruitment makes the interpretation of the selection curve derived by ELEFAN II (and also the application of the length-frequency correction procedure) difficult because the tailing off of the catch of the very smallest fish would be assumed by ELEFAN II to be the result of size-selection by the fishing gear. This would lead to underestimates of the probabilities of retention for these fish. Secondly, any error in the estimate of the slope of the catch-curve regression line and in the estimate of M (which cannot be derived from length-frequency data alone and is, in many cases, highly uncertain) will, of course, result in errors in the estimates of probabilities of retention. Mainly because of the first source of error, we might expect that ELEFAN II would tend to underestimate the probabilities of retention, and hence, overestimate the effect of size selection.

This seems to be the case in experiments 1 and 3, where the pattern of size selection derived by ELEFAN II is compared with a typical pattern generated directly from the simulation model (Table 7).

ELEFAN III

The VPA 2 routine of ELEFAN III appears to provide, with few exceptions, quite reasonable estimates of average recruitment. The observed errors (Table 8) seem to be directly related to errors in the estimates of Z derived using ELEFAN II.

The values of F averaged for the fully recruited length-classes show strong positive bias (Table 9). This is due to very high values of F being generated for the three or four largest length-classes. The terminal length-class used for these calculations was 44.5-46.5 cm. The bias is even stronger than that indicated in Table 9 if larger length-classes are assigned as the terminal length-class. This

Table 7. Selection patterns for replicate 1 of experiments 1 and 3 as determined by ELEFAN II, and the real selection pattern as defined in the simulation model. The corresponding estimates of L_∞ and K derived by ELEFAN I (Table 2) are used in the ELEFAN II procedure.

Midlength (cm)	Probability of retention		Actual
	Experiment 1	Experiment 3	
9.5	0.00	0.00	0.03
11.5	0.03	0.03	0.09
13.5	0.09	0.08	0.28
15.5	0.22	0.19	0.60
17.5	0.32	0.29	0.85
19.5	0.38	0.34	0.96
21.5	0.47	0.39	0.99
23.5	0.62	0.55	1.00
25.5	0.81	0.82	1.00
27.5	1.00	1.00	1.00

Table 8. Comparisons between the mean recruitment levels determined by the simulation model (R_s) and by the VPA 2 routine of ELEFAN III (R_e). The percentage difference is shown in parentheses. The corresponding estimates of L_∞ and K derived by ELEFAN I (Table 2) and Z_r derived by ELEFAN II (Table 4) are used in the VPA 2 procedure. We assume $M = F_t = Z_r/2$.

Replicate	Experiment							
	R_s	R_e	R_s	R_e	R_s	R_e	R_s	R_e
1	8,560	9,056 (+6)	8,560	9,053 (+6)	9,541	9,221 (-3)	10,203	10,662 (+4)
2	11,907	12,983 (+9)	11,907	12,923 (+9)	10,282	11,335 (+10)	10,209	9,664 (-5)
3	9,700	9,918 (+2)	9,700	9,910 (+2)	9,180	9,409 (+2)	10,567	12,896 (+22)
4	11,132	11,081 (-0)	11,132	11,085 (-0)	11,137	10,710 (-4)	10,496	9,934 (-5)
5	8,049	9,265 (+15)	8,049	9,211 (+14)	9,493	9,300 (-2)	8,979	8,997 (+0)
6	9,336	10,354 (+11)	9,336	10,314 (+10)	10,975	12,669 (+15)	9,751	7,643 (-22)
7	10,350	13,230 (+28)	10,350	13,230 (+28)	8,647	7,182 (-17)	9,137	8,244 (-10)
8	10,356	11,272 (+9)	10,356	11,270 (+9)	9,317	9,784 (+5)	10,022	10,428 (+4)
9	9,649	11,740 (+22)	9,649	11,490 (+19)	10,338	11,208 (+8)	9,740	9,576 (-2)
10	11,199	10,916 (-3)	11,199	10,905 (-3)	10,193	9,192 (-10)	8,948	8,445 (-6)
Mean	10,027	10,982 (+10)	10,027	10,939 (+9)	9,910	10,001 (+1)	9,805	9,649 (-2)

Table 9. Estimates of average F derived using the VPA 2 routine of ELEFAN III. The corresponding estimates of L_∞ and K derived by ELEFAN I (Table 2) and Z_r derived by ELEFAN II (Table 4) are used in the VPA 2 procedure. We assume $M = F_t = Z_r/2$. The real value of F for fully recruited length-classes input to the simulation model is 0.5 year⁻¹.

Replicate	Experiment			
	1	2	3	4
1	0.87	0.87	0.79	0.65
2	0.87	0.86	0.76	0.64
3	0.78	0.88	0.79	0.68
4	0.80	0.80	0.72	0.59
5	1.03	0.84	0.79	0.66
6	0.93	0.90	0.84	0.56
7	0.85	0.85	0.70	0.60
8	0.85	0.85	0.80	0.66
9	0.94	0.96	0.75	0.61
10	0.87	0.86	0.73	0.61
Mean	0.88	0.87	0.77	0.63

may be due, at least in part, to the artificial termination of the fishery at five years after recruitment. This could have led to the larger length-classes being slightly underrepresented, particularly in experiments 1 and 2 where growth variation is high. If the terminal length-class is underrepresented in this way, the preceding length-classes with larger catches could have F values that are much larger than the terminal F used. This indicates the need to ensure that: (1) fish from the terminal length-class are properly sampled and (2) an optimum length-class interval is chosen so as to minimize the variation within length-classes while retaining most of the between-length-class information.

The remarks made here probably also apply as well to Jones' length-cohort analysis (Jones 1979, 1981), which has properties similar to the VPA 2 routine of ELEFAN III (Pauly 1984, Chapter 5 and Jones, Part I, this vol.).

Concluding Remarks

In the analysis of the simulated data, we occasionally had to make assumptions that might be much more difficult to make in practice (e.g., those regarding the values of t_o and M). In addition, some of the assumptions of the simulation model itself were, for the sake of simplicity, somewhat optimistic (e.g., 100% sampling of the catch and no variation in M or the instantaneous rate of encounter). Therefore, we might expect that the analysis of real data would include additional complications relating to these assumptions. Nevertheless, the results of this study have highlighted a number of points regarding the ELEFAN system. They are briefly summarized as follows:

1. Unless one is quite confident that the largest length-class included in the analysis is smaller than the true mean L_∞ , ELEFAN I should not be used to estimate this parameter. If this is not the case, the best (or least biased) estimate of L_∞ that ELEFAN I can provide is one length-class interval larger than the midpoint of the largest length-class in the samples.
2. Truncating the data might help to overcome the problem described in (1), although where the point of truncation should be is debatable. The deletion of successive length-classes from the samples in the hope that the L_∞ and K estimates will stabilize may be a sensible course of action, but this procedure requires further validation.
3. Size-selection by the fishing gear also causes ELEFAN I to underestimate K and overestimate L_∞ . This may be at least partially compensated for if accurate estimates of the probabilities of retention by length-class can be used to correct the length-frequency data

for this effect. While ELEFAN II can provide such estimates, there may be problems with this procedure if fish are not recruited strictly by length and/or there are substantial errors in the estimates of Z and M.

4. ELEFAN II tends to overestimate Z in experiments where the growth parameters are highly variable. Also, normal (or random) variation in year-class strength can cause substantial errors in Z estimates.
5. The temporal spread of recruitment is overestimated by ELEFAN II; however, its peak can be identified with reasonable accuracy if the correct value of t_o is provided.
6. ELEFAN III provided reasonable estimates of mean recruitment (as measured in these experiments). As with all other forms of VPA, accurate estimates of M and terminal F are necessary.
7. The estimates of average F provided by ELEFAN III are positively biased, probably due to a combination of factors, including fish being subject to exploitation for five years only, individual variation in growth and the difficulty in identifying a correct value of F_t .

The ELEFAN system incorporates empirical equations to obtain values for parameters which usually cannot be estimated from length-frequency data alone (t_o and M). Their robust estimation requires supplementary data (e.g., age-at-length, tag-recapture and/or fishing effort data). We did not make an assessment of the methods incorporated in ELEFAN II for estimating these parameters, but, as recognized by Pauly (Part I, this vol.), they should only be used where superior methods are not available.

We strongly recommend that thorough sensitivity analyses (see reviews in Majkowski 1982, 1983; Majkowski and Hampton 1983) be undertaken when using the ELEFAN programs. The ranges for the various parameters to be used in sensitivity analysis will vary for different data sets, but our experiences in working with simulated data would suggest the following minimum ranges: $L_\infty \pm 20\%$, $K \pm 30\%$ (for input to ELEFAN II and III) and $Z \pm 40\%$ (for input to ELEFAN III). Also, the sensitivity of results to uncertainties in t_o and M should be addressed. This approach would allow management advice that takes into account the shortcomings of the methods used.

This paper has identified several areas that warrant further investigation. These include the effects of truncating the length-frequency data to remove bias; the effects of larger variations in growth parameters and recruitment than used in this paper upon estimates of L_∞ , K and Z; and the reliability of the procedure for correcting for size-specific selection. Information on these topics is required before the ELEFAN system can be confidently recommended for routine use.

Acknowledgements

Drs. John Caddy (Food and Agriculture Organization, United Nations, Rome, Italy) and John Gulland, F.R.S. (Imperial College, London, England) suggested that this study be undertaken. Partial financial support for the project was provided by the Food and Agriculture Organization. We are grateful to Dr. Daniel Pauly (ICLARM, Manila, Philippines) who provided copies of the ELEFAN software and documentation; to Mr. Kevin McLoughlin (CSIRO Division of Fisheries Research, Hobart, Tasmania, Australia) for his assistance in installing and testing the ELEFAN programs on our microcomputer; and to Dr. Keith Sainsbury (CSIRO Division of Fisheries Research, Hobart, Tasmania, Australia) for his helpful suggestions.

References

- Bertalanffy, L. von. 1938. A quantitative theory of organic growth (Inquiries in growth laws II). *Hum. Biol.* 10: 181-213.
- Beverton, R.J.H. and S.J. Holt. 1956. A review of methods for estimating mortality rates in fish populations, with special references to sources of bias in catch sampling. *Rapp. P.-V. Réun. Cons., Cons. Int. Explor. Mer* 140: 67-83.
- Ingles, J. and D. Pauly. 1984. An atlas of the growth, mortality and recruitment of Philippine fishes. International Center for Living Aquatic Resources Management, Manila, Philippines.
- Jones, R. 1979. An analysis of *Nephrops* stocks using length composition data. *Rapp. P.-V. Réun. Cons., Cons. Int. Explor. Mer* 175: 259-269.
- Jones, R. 1981. The use of length composition data in fish stock assessment (with notes on VPA and cohort analysis). *FAO Fish. Circ.* 734. 55 p.
- Majkowski, J. 1982. Usefulness and applicability of sensitivity analysis in a multi-species approach to fisheries management, p. 149-165. In D. Pauly and G.I. Murphy (eds.) *Theory and management of tropical fisheries*. International Center for Living Aquatic Resources Management, Manila, Philippines and Division of Fisheries Research, Commonwealth Scientific and Industrial Research Organisation, Cronulla, Australia.
- Majkowski, J. 1983. Usefulness and applicability of sensitivity analysis in resource management, p. 247-261. In L.C. Wadhwa, M.K. James, A.B. Pomeroy and K.P. Stark (eds.) *Resource management issues, techniques, applications*. Proceedings of the Workshop on Resource Management, James Cook University, 14-17 July 1981. James Cook University of North Queensland, Townsville, Australia.
- Majkowski, J. and J. Hampton. 1983. The effect of parameter uncertainties in an age-length relationship upon estimating the age composition of catches. *Can. J. Fish. Aquat. Sci.* 40: 272-280.
- Pauly, D. 1982. Studying single-species dynamics in a tropical multi-species context, p. 33-70. In D. Pauly and G.I. Murphy (eds.) *Theory and management of tropical fisheries*. International Center for Living Aquatic Resources Management, Manila, Philippines and Division of Fisheries, Commonwealth Scientific and Industrial Research Organisation, Cronulla, Australia.
- Pauly, D. 1983. Length-converted catch curves: a powerful tool for fisheries research in the tropics (Part I). *Fishbyte* 1(2): 9-13.
- Pauly, D. 1984. Fish population dynamics in tropical waters: a manual for use with programmable calculators. International Center for Living Aquatic Resources Management, Manila, Philippines.
- Pauly, D. and N. David. 1981. ELEFAN I, a BASIC program for the objective extraction of growth parameters from length-frequency data. *Meeresforschung* 28: 205-211.
- Pauly, D. and Sann Aung. 1984. Population dynamics of some fishes of Burma based on length-frequency data. BUR/77/007/FAO Field Doc. No. 7. 22 p.
- Pauly, D. and I. Tsukayama. 1983. On the seasonal growth, monthly recruitment and monthly biomass of Peruvian anchoveta (*Engraulis ringens*) from 1961 to 1979, p. 987-1004. In G.D. Sharp and J. Csirke (eds.) *Proceedings of the Expert Consultation to Examine Changes in Abundance and Species Composition of Neritic Fish Resources*, San José, Costa Rica, 18-19 April 1983. FAO Fish. Rep. 291, Vol. 3.
- Pope, J.G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. *ICNAF Res. Bull.* 9: 65-74.
- Robson, D.S. and D.G. Chapman. 1961. Catch curves and mortality rates. *Trans. Amer. Fish. Soc.* 90: 181-189.
- Sainsbury, K.J. 1980. Effect of individual variability on the von Bertalanffy growth equation. *Can. J. Fish. Aquat. Sci.* 37: 241-247.

An Investigation of Length Composition Analysis Using Simulated Length Compositions

R. JONES

Marine Laboratories

*Department of Agriculture and Fisheries for Scotland
PO Box 101, Aberdeen ABG 8DB, Scotland, United Kingdom*

Jones, R. 1987. An investigation of length composition analysis using simulated length compositions, p. 217-238. In D. Pauly and G.R. Morgan (eds.) *Length-based methods in fisheries research*. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

This paper summarizes the results of applying two kinds of length composition analysis to simulated length compositions. The kinds of analyses investigated are length cohort analysis for determining the long-term effects of changes in exploitation pattern and methods of converting the right hand limb of a length composition to a straight line for estimating Z/K or Z - K.

Length compositions were simulated by combining length compositions derived by following five cohorts with different growth parameters throughout their lives, and combining sets of normally distributed length compositions-at-age for a number of age groups. For both approaches, the effect of selectivity on the relative proportions of small individuals was allowed for using a logistic selection curve.

Introduction

This paper presents an investigation of the effects of changes in growth and mortality parameters on the results of length cohort analysis and Z or Z/K estimates for length composition data based on simulated catch-at-length data. The aim of this investigation is to provide practical guidelines for the application of methods such as those presented in Jones (1984) and Pauly (1984).

SIMULATION OF CATCH-AT-LENGTH DATA

Simulation of a single cohort

A single cohort can be simulated quite simply, either by following a year class throughout its life or by simulating a steady-state length composition given constant recruitment throughout

the year. For a cohort followed through its life, the number removed from any length interval $L_1 - L_2$ will be equivalent to:

$$F/Z N_o (e^{-Zt_1} - e^{-Zt_2}) \quad \dots 1)$$

where $t_1 = (-1/K) \ln (1 - L_1/L_\infty)$ and $t_2 = (-1/K) \ln (1 - L_2/L_\infty)$. Note that t_1 and t_2 refer simply to relative ages (i.e., " t_o " can be ignored).

Alternatively a steady state length composition can be simulated by assuming, as in Fig. 1, that during some time interval, all individuals previously in an interval $L_1 - L_3$ will grow to lie within an interval $L_2 - L_4$. L_2 can be calculated from L_1 , and L_4 from L_3 using the same von Bertalanffy relationship. Simulations show that both methods lead to results which are identical within the limits of computational accuracy.

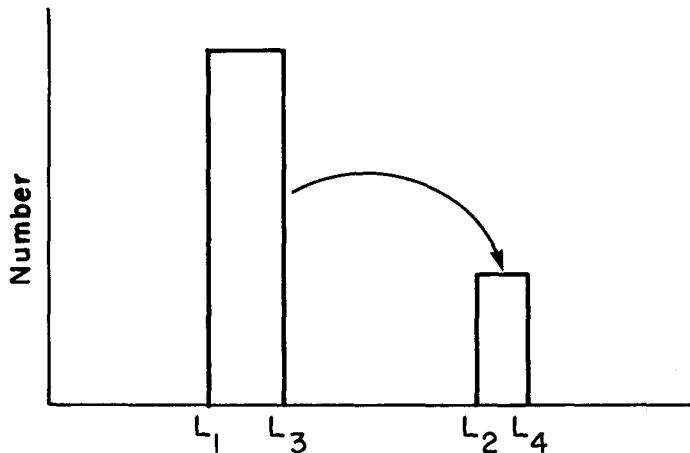


Fig. 1. Basic assumption in a simulated set of catch-at-length data: all individuals in an interval $L_1 - L_3$ will grow to lie if they survive within an interval $L_2 - L_4$.

Simulation of growth curves—whether to vary L_∞ or K

In practice, the variance of length-at-age tends to increase with increasing age, and the indication is that fish grow according to different growth curves throughout their lives. Variation in K, but not in L_∞ , would lead to length compositions for which the variance of the lengths increased with age up to a point, and then gradually decreased to zero (Fig. 2). It is, however, more realistic to assume that variation is largely due to variations in L_∞ , as indicated in Fig. 3.

Two approaches to length composition simulation

From the considerations above, two methods have been adopted for length composition simulation. One, as indicated in Fig. 4, was to assume a normal distribution of lengths-at-age. Mean lengths-and numbers-at-age were calculated using von Bertalanffy and exponential relationships in the usual manner. This method is suitable for generating length composition representative of the stock.

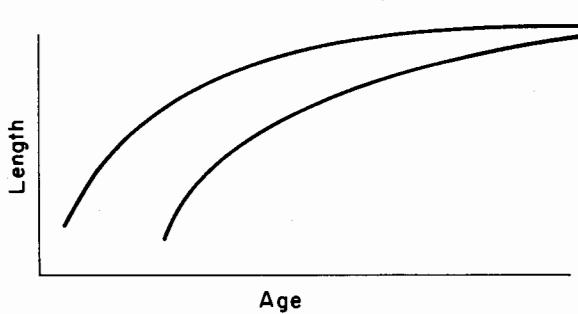


Fig. 2. Growth curves that would result if variations of K were strong, while variation of L_∞ were zero (see text).

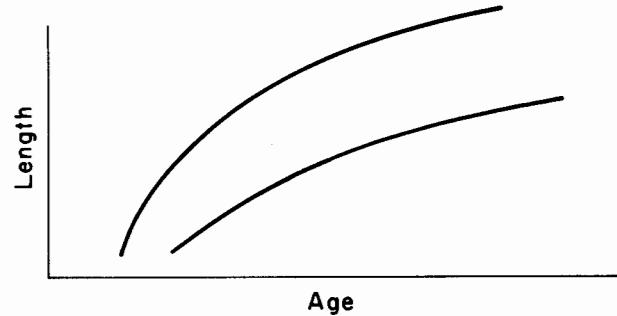


Fig. 3. Basic growth model used for simulation: all fish grow according to different growth curves. Their individual growth parameters, however, remain constant throughout life. Variance of length-at-age is assumed to increase with age.

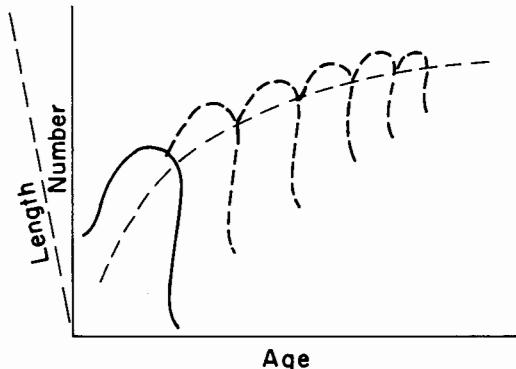


Fig. 4. Schematic representation of method used for constructing simulated population length-at-age data (see text).

The second method was to use equation (1) to determine length compositions for five cohorts and then to combine these by simple addition or by weighting. This method (numerical examples are given below) is suitable for generating length compositions representative of catch.

Table 1 shows the mean and variances of North Sea haddock for ages 1-7. Also shown are the values of each mean \pm the standard deviation and the values of each mean \pm twice the standard deviation. If von Bertalanffy curves are fitted to the mean lengths at age it is found that the L_∞ for the mean lengths is 58 and the value of K is 0.42. For the values of mean \pm once and twice the standard deviation the corresponding von Bertalanffy curves have different values of L_∞ but similar values of K .

To allow for variations in growth rate two approaches have, therefore, been used:

- i. computing the normal distribution of lengths for each age group using the means and variances in Table 1 and combining these into a single length composition (Table 2);
- ii. simulating several cohorts, each based on equation (1), but with different values of L_∞ . Several examples are given in Tables 3-5.

Tables 3-5 allow for various combinations of possibilities. In Tables 3 and 4 the values of K and M are relatively high (0.42 and 0.63, respectively). In Tables 5 and 6 the values of K and M are relatively low (0.15 and 0.225, respectively). For all four tables, however, the value of $M/K = 1.5$.

To allow for differences in the relative proportions of fish with different values of L_∞ , the individual cohorts have been combined in two different ways. In Tables 3 and 5, the numbers per

Table 1. Input data for simulating length compositions using variances-at-length for North Sea haddock.

Age	Mean length	Variance	SD	Mean - SD	Mean + SD	Mean + 2 SD	Mean - 2 SD
1	15	3	1.73	13.3	16.7	18.5	11.5
2	30	8	2.83	27.2	32.8	35.7	24.3
3	40	14	3.74	36.3	43.7	47.5	32.5
4	46	19	4.36	41.6	50.4	54.7	37.3
5	50	25	5.00	45.0	55.0	60.0	40.0
6	52.7	25*	5.00	47.7	57.7	62.7	42.7
7	54.5	25*	5.00	49.5	59.5	64.5	44.5
L_{∞}	58	—	—	52.0	65.0	70.0	45.0
K	0.42	—	—	0.44	0.41	0.47	0.41

* = Assumed values.

SD = Standard Deviation.

Table 2. Simulated length compositions based on summations of normal distributions at each age. Mean lengths-at-age and variances are as given in Table 1 for $L_{\infty} = 58$ cm and K = 0.42 for ages from 1 to 16 years, with variance-at-age = 25 for ages > = 5 years.

Length group (cm) ^a	Z					
	0.1	0.5	1.0	1.4	1.6	2.0
< 10	19	19	19	19	19	19
10	397	397	397	397	397	397
12	2,402	2,402	2,402	2,402	2,402	2,402
14	4,363	4,363	4,363	4,363	4,363	4,363
16	2,402	2,402	2,402	2,402	2,402	2,402
18	399	398	398	397	397	397
20	39	32	27	24	23	22
22	132	89	54	36	30	20
24	558	374	227	152	124	83
26	1,461	978	592	397	324	218
28	2,379	1,589	959	642	524	352
30	2,460	1,624	972	648	528	354
32	1,789	1,122	644	420	339	224
34	1,352	718	350	206	160	99
36	1,604	712	274	132	93	48
38	2,198	897	310	135	90	41
40	2,704	1,014	329	137	90	39
42	3,049	994	291	114	74	31
44	3,424	907	227	81	50	19
46	4,012	823	168	52	31	10
48	4,842	751	124	32	18	5
50	5,798	673	90	20	10	2
52	6,636	580	62	11	6	1
54	7,026	473	40	6	3	
56	6,707	359	24	3	1	
58	5,670	249	13	1	1	
60	4,197	155	7	1		
62	2,700	86	3			
64	1,503	42	1			
66	721	18				
68	298	7				
> 70	149	3				

^aLower class limit.

cohort have been adjusted so that the initial numbers at the smallest lengths are in the proportions 6:24:40:24:6.

In Tables 4 and 6, the numbers per cohort have been adjusted so that the weight of fish above a certain length (here taken as 28 cm) is the same for all five cohorts. This simulates a situation in which the potential reproductive output is the same for all five cohorts. Without some condition such as this, there would presumably be strong selective pressure for that combination of growth parameters that consistently favored the maximum reproductive output.

Table 3. Simulated catch-at-length data based on equation (1).

	A	B	C	D	E	A+B+C+D+E
Input parameters:						
L _∞	45	52	58	65	70	
K	0.42	0.42	0.42	0.42	0.42	
F	0.8	0.8	0.8	0.8	0.8	
M	0.63	0.63	0.63	0.63	0.63	
Numbers in sea at a length of 6 cm						
	6,000	24,000	40,000	24,000	6,000	100,000
Simulated length compositions						
6	1	4	6	3	1	15
8	3	9	14	7	2	35
10	6	22	33	18	4	83
12	15	51	77	41	9	193
14	33	116	174	93	22	438
16	71	248	373	201	47	940
18	132	471	714	387	90	1,794
20	205	745	1,144	627	147	2,868
22	251	945	1,478	822	195	3,691
24	251	989	1,587	900	216	3,943
26	218	913	1,512	879	213	3,735
28	172	783	1,348	807	199	3,309
30	128	642	1,158	717	179	2,824
32	89	509	970	624	159	2,351
34	58	391	797	536	140	1,922
36	34	290	641	455	122	1,542
38	17	206	504	381	105	1,213
40	7	138	386	314	89	934
42	1	85	286	255	75	702
44		47	203	203	62	515
46		21	136	158	51	366
48		6	84	119	41	250
50			46	86	32	164
52			21	60	25	106
54			6	38	18	62
56				23	13	36
58				11	9	20
60				4	5	9
62					3	3
64					1	1
Total	1,692	7,631	13,698	8,769	2,274	34,064

The effect of selectivity on the smaller individuals has been allowed for by using a logistic selection curve with a selection factor of 3.0 and a slope at the 50% length of 0.1.

Investigations based on both principles have been carried out and preliminary results are presented later.

Table 4. Simulated catch-at-length data (as in Table 3 but initial numbers adjusted to equalize biomass of fish > 28 cm).

	A	B	C	D	E	A+B+C+D+E
Input parameters:						
L _∞	45	52	58	65	70	
K	0.42	0.42	0.42	0.42	0.42	
F	0.8	0.8	0.8	0.8	0.8	
M	0.63	0.63	0.63	0.63	0.63	
Numbers in sea at a length of 6 cm						
	27,910	15,396	10,576	7,430	5,991	67,303
Simulated length compositions						
6	5	3	2	1	1	12
8	13	6	4	2	2	27
10	30	14	9	5	4	62
12	69	33	20	13	9	144
14	156	75	46	29	22	328
16	328	159	99	62	47	695
18	616	302	189	120	90	1,317
20	953	478	303	194	147	2,075
22	1,169	606	391	254	194	2,614
24	1,167	635	420	279	215	2,716
26	1,012	586	400	272	213	2,483
28	802	502	356	250	198	2,108
30	595	412	306	222	179	1,714
32	415	327	256	193	159	1,350
34	269	251	211	166	139	1,036
36	158	186	170	141	121	776
38	80	132	133	118	104	567
40	31	89	102	97	89	408
42	6	55	76	79	75	291
44		30	54	63	62	209
46		14	36	49	51	150
48		4	22	37	41	104
50			12	27	32	71
52			5	18	25	48
54			2	12	18	32
56				7	13	20
58				4	9	13
60				1	5	6
62					3	3
64					1	1
Total	7,874	4,899	3,624	2,715	2,268	21,380

Table 5. Simulated catch-at-length data using equation (1) and based on growth parameters for Faroe Haddock.

	A	B	C	D	E	A+B+C+D+E
Input parameters:						
L _∞	55	69	83	94	105	
K	0.15	0.15	0.15	0.15	0.15	
F	0.8	0.8	0.8	0.8	0.8	
M	0.225	0.225	0.225	0.225	0.225	
Numbers in sea at a length of 5 cm						
	6,000	24,000	40,000	24,000	6,000	100,000
Simulated length compositions						
5	8	25	34	18	4	89
10	66	212	294	156	35	763
15	450	1,487	2,101	1,125	254	5,417
20	1,154	4,247	6,380	3,526	814	16,121
25	757	3,559	6,078	3,583	867	14,844
30	260	1,814	3,744	2,418	624	8,860
35	63	780	2,059	1,487	415	4,804
40	10	288	1,049	868	265	2,480
45	1	87	489	480	163	1,220
50		19	203	248	96	566
55		3	72	118	54	247
60			21	50	28	99
65			4	19	14	37
70				6	6	12
75				1	2	3
80					1	1
Total	2,769	12,521	22,528	14,103	3,642	55,563

INVESTIGATION OF LENGTH COHORT ANALYSIS

Cohort analysis (Pope 1972 and see Fig. 5) is a simple approximation of the virtual population technique in which an exponential survival curve is replaced by a "step function", by assuming that (a) the whole of the catch for an age group is taken at exactly the middle of that age interval and (b) only natural losses occur continuously on an exponential basis. Pope's cohort formula can then be derived quite simply by proceeding backwards in time from the oldest to the youngest ages.

Thus, let N_{t+1} = number in sea at sea age $t + 1$ (point D in Fig. 5) and let M = the instantaneous loss rate due to natural causes.

Then by proceeding backwards in time, the number in the sea at the middle of the age interval, just after the catch has been taken, is given by:

$$N_{t+1} e^{M/2} \text{ (point C in Fig. 5)}$$

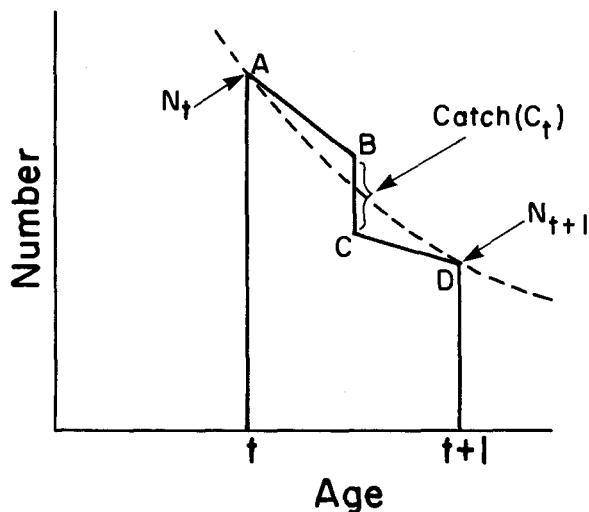


Fig. 5. Showing how, in cohort analysis, a negative exponential decay curve is replaced by a step function.

Table 6. Simulated catch-at-length data based on growth parameters for Faroe Haddock (as for Table 5 but initial numbers adjusted to equalize biomass of fish > 28 cm).

	A	B	C	D	E	A+B+C+D+E
Input parameters:						
L _∞	55	69	83	94	105	
K	0.15	0.15	0.15	0.15	0.15	
F	0.8	0.8	0.8	0.8	0.8	
M	0.225	0.225	0.225	0.225	0.225	
Numbers in sea at a length of 5 cm						
	36,681	13,852	7,608	5,345	3,986	67,472
Simulated length compositions						
5	48	14	6	4	3	75
10	407	122	56	35	23	643
15	2,753	858	400	251	169	4,431
20	7,052	2,451	1,214	785	541	12,043
25	4,626	2,054	1,156	798	576	9,210
30	1,587	1,047	712	539	415	4,300
35	384	450	392	331	276	1,833
40	59	166	199	193	176	793
45	4	50	93	107	108	362
50		11	39	55	64	169
55		2	14	26	36	78
60			4	11	19	34
65				4	9	13
70				1	4	5
75					2	2
80						
Total	16,920	7,225	4,285	3,140	2,421	33,991

Also, at the middle of the age interval, but just before the catch is removed, the number in the sea is given by:

$$N_{t+1} e^{M/2} + C_t \text{ (point B in Fig. 5)}$$

where C_t = catch during age interval.

Next by proceeding back in time one more step to the beginning of the interval, the number in the sea at age t is given by:

$$N_t = (N_{t+1} e^{M/2} + C_t) e^{M/2} \text{ (point A in Fig. 5)} \quad \dots 2)$$

where N_t = number in sea at age t .

Numerical analysis proceeds by first determining the number in the sea for the oldest age, and then successively applying this formula for determining the numbers in the sea at successively younger ages.

Further details and numerical examples are given by Jones (1984).

Fig. 6 shows the relationship between length and real time for a number of cohorts each derived from eggs that hatch on different days. Cohort number 1, for example, may grow from L_1 to L_2 over one time period. Cohort number 4 grows through the same length interval considerably later in real time. The time intervals between the different curves depends partly on the different hatching times and partly on different growth rates prior to the attainment of the reference length L_1 cm.

Fig. 7 shows the relationship between length and real age for the same cohorts as in Fig. 2. It is assumed that each cohort is assigned a zero age at the time of hatching. There will be differences in growth rate prior to the attainment of the length L_1 and consequently each cohort will grow from L_1 to L_2 cm over a different age range. The differences in hatching time will be eliminated, however, and one would expect these curves to be much closer together than those in Fig. 2 differing only due to differences in early growth rate.

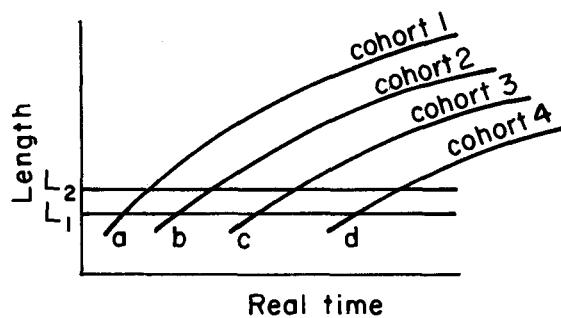


Fig. 6. Relationship between length and time for cohorts born at different times.

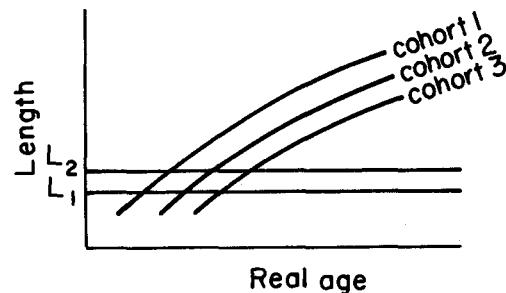


Fig. 7. Relationship between length and age for cohorts in which length L_1 is reached at the same age.

Fig. 8 shows the relationship between length and relative age for the same four cohorts. By "relative" age is meant age relative to the age at the time of attainment of a particular length, such as L_1 . Over a small length interval, it is to be expected that much of the variability would be removed, and that the differences between the various curves would be relatively small. The cohort length analysis is based on this assumption, i.e., it is assumed that the variation in time needed to grow from L_1 to L_2 cm is small, compared to the variations in the ages or real times involved.

For small and medium fish it may be reasonable to assume that there is not a very large variation in the time required to grow through a particular length group. For very large fish, however, this assumption will be less valid.

Fig. 9 shows the relationship between length and relative age for a length grouping of large fish. In this situation, some individuals may continue to grow throughout the length interval (cohort A) whereas others may actually stop growing and never reach a length of L_2 (cohort B). In this situation the individual variability is clearly very large, and the adoption of an average growth rate to apply to all individuals can be only a poor approximation.

To some extent this should be compensated for by the fact that large individuals usually make up a relatively small proportion of the total catch. Errors, due to the greater variability of growth of the very large individuals, should, therefore, only apply to a relatively small part of the total population.

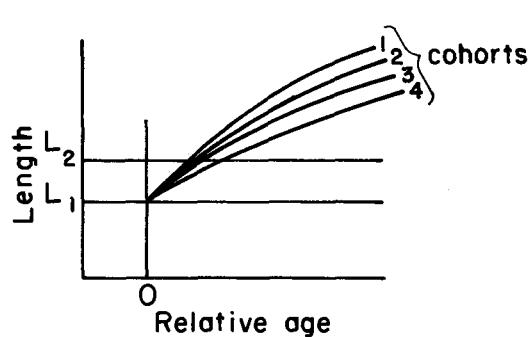


Fig. 8. Relationship between length and relative age for cohorts with the same relative age (here 0) at some specified length (here L_1).

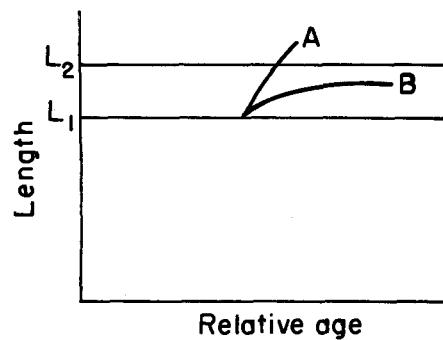


Fig. 9. Relationship between length and relative age for large individuals. Within length group L_1 - L_2 , individual A is still growing, while individual B has stopped growing.

For length composition analysis it is important to note that the values of L_∞ and K obtained from the best fitting von Bertalanffy growth curve, may not necessarily be the best values to use. This is because in reality, there are two underlying relationships (Fig. 10). One is the relationship between length and age (i.e., corresponding to each age there is a mean length and the relationship between these constitutes a conventional growth curve—curve (a)). With this relationship the value of L_∞ may be smaller than the largest individuals in the length compositions (as shown in Fig. 11).

The other is the relationship between age and length (i.e., corresponding to each length there is a mean age). This relationship provides a different growth curve with different values of L_∞ and K (curve (b) in Fig. 10). With this relationship, the value of L_∞ will not be smaller than the largest length.

Investigation of mean ages per length group

As part of the preparation of the length compositions shown in Table 2, it has been possible to calculate mean ages for each length group and hence to prepare curves of mean ages on length. This leads to a whole set of curves, each depending on the value of Z adopted. The results are summarized in Fig. 12. In this example, the curves of age on length tend to an asymptotic age rather than an asymptotic length. This appears to be a consequence of including the right-hand limb of the oldest age groups in each curve in the figure. However, the slope of the right-hand limb of any age group using this method of simulation is not a function of growth and mortality but a property of a normal distribution.

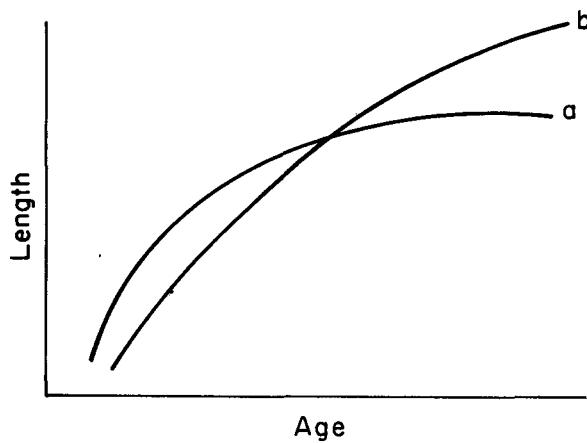


Fig. 10. Schematic representation of difference between the curves resulting from a plot of length on age (a) and age on length (b) (see text for details).

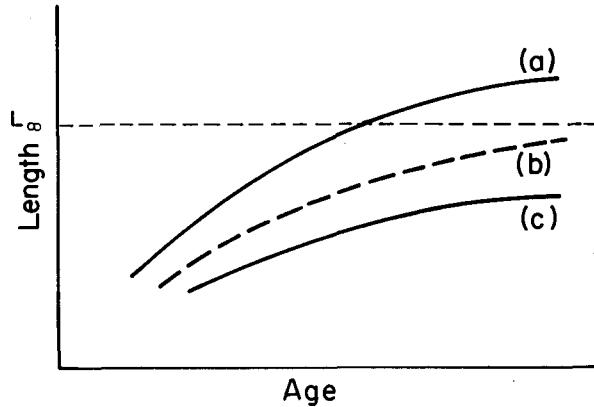


Fig. 11. Growth curves for fast growing (a), and slow growing (b) fish, and the relation between average length and age (dotted). The figure shows how some fish can be larger than the L_∞ of the population.

If the tails of the distribution in Table 2 are neglected, the curves of age on length appear to be quite close to the underlying curve of length on age but not identical.

This investigation shows that curves of age on length do not conform to any one growth model and are not likely to be readily predictable in practice. Because of this, some other way of choosing values of L_∞ and K is required. One procedure is to first choose a value of L_∞ that is greater than the largest length group in the sample and then to adopt a value of K that is compatible with it.

Determination of compatible values of L_∞ and K

It is important, when choosing a value for K , to make sure that this is compatible with the value of L_∞ adopted. One way of doing this is to relate values of L_∞ and K to a particular growth increment by means of a Ford-Walford plot. Fig. 13 shows an example for the situation where animals grow from 42 to 46 cm in one year. This determines point C in Fig. 13. A second point (A) is then determined by arbitrarily choosing a value for L_∞ greater than the largest length group in the samples. Here a value of 70 cm has been adopted. A straight line can be drawn through point C and through point A from which it is possible to calculate values of L_∞ and K that are compatible with the observed growth increment. Thus the slope of this line is equivalent to AB/BC, i.e.,

$$= (70 - 46)/(70 - 42) = e^{-K} = 0.86$$

The value of K is therefore $-\ln 0.86 = 0.15$.

Other lines can be drawn for other values of L_∞ . In this way values of L_∞ and K can be obtained that are compatible with the observed growth increment.

If various lines are drawn it is clear that the larger the value of L_∞ the smaller the slope and hence the smaller the value of K .

For length composition analysis, therefore, it is appropriate first to choose a value for L_∞ and then to determine a value for K that is compatible with this and with one growth increment. The choice of growth increment is necessarily arbitrary, but a useful procedure is to adopt a growth increment that happens to coincide with growth in the region of the modal length of the population.

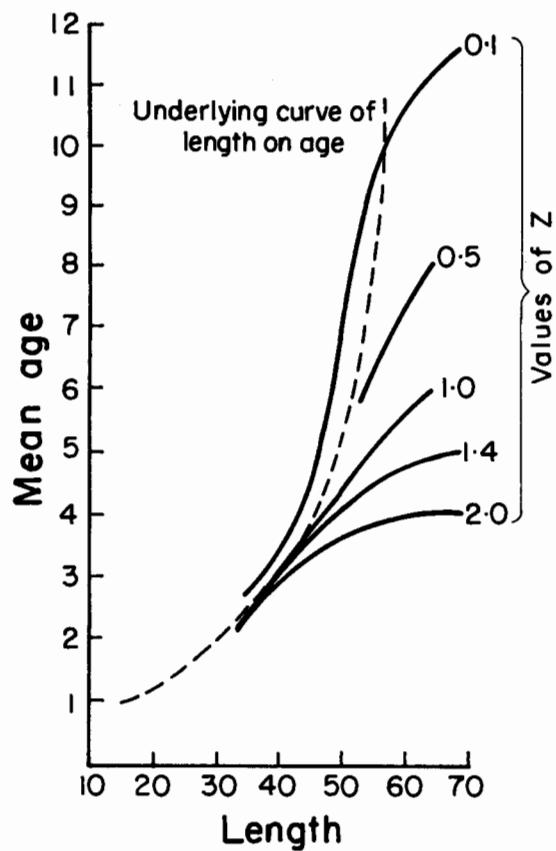
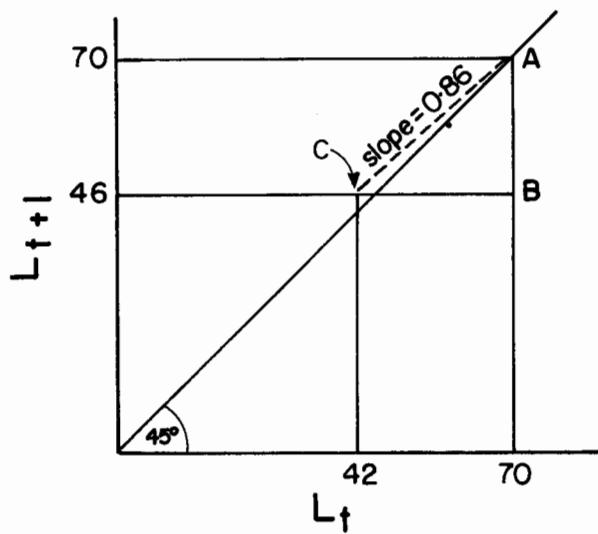


Fig. 12. Effect of total mortality (Z) on the growth curve resulting from a plot of age on length (see text).



$$\text{slope} = e^{-K} = (L_\infty - L_{t+1}) / (L_\infty - L_t)$$

$$\text{hence } K = \ln [(L_\infty - L_t) / (L_\infty - L_{t+1})]$$

Fig. 13. Method for estimation of a value of K compatible with a given growth increment and a pre-selected value of L_∞ (see text for details).

Principle of determination of percentage change in steady-state catch

Fig. 14 shows the relationship between the steady-state numbers-at-length in the catch and the steady-state numbers-at-length in the sea for a given pattern of fishing.

Length cohort analysis can be used to determine the percentage change in steady-state catch for some other pattern of fishing.

One way of doing this would be by converting length groups to relative age groups and then employing conventional age-related techniques. Alternatively, the various computational steps can be combined into a single sequence that makes use of the values of $F \Delta t$ obtained from cohort analysis. This permits an assessment of the long-term effects of a change in exploitation pattern to be made without any further input data being required. The results are numerically almost identical to those obtained by converting lengths to ages and applying conventional age-related techniques, but fewer input assumptions have to be made. In fact, the only difference is that numbers caught are calculated from $F \Delta t N_L$ instead of from

$$F N_B (1 - e^{-Z \Delta t}) / Z$$

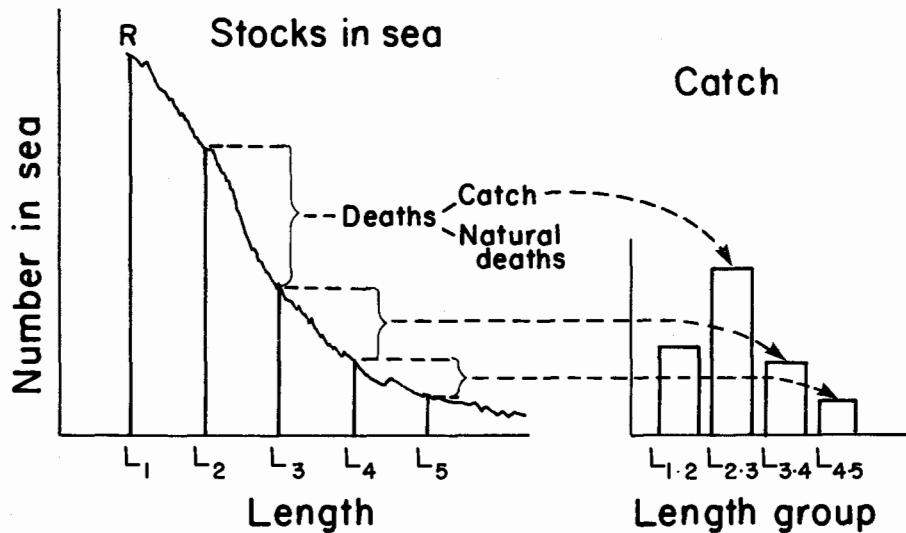


Fig. 14. Relationships between steady-state number-at-length in the sea and the catch resulting from a given pattern of fishery.

where

N_L is the number at a length corresponding to the middle of a length group

N_B is the number at a length corresponding to the beginning of a length group and
 Δt is the time needed to grow through the length interval

The choice of either formulation is not likely to affect the results significantly.

Results and Discussions

INVESTIGATION OF THE EFFECT OF CHANGES IN EXPLOITATION PATTERN (BASED ON COHORT ANALYSIS)

Estimates of the long-term effect of changes in exploitation pattern have been made using the simulated length composition data in Tables 3-6. In each instance, cohort analysis has been applied to sets A...E combined and the results compared with the results obtained for sets A...E separately.

LONG-TERM EFFECTS OF A 40% REDUCTION IN EFFORT

Table 7a shows the long-term effect of a 40% reduction in effort on sets A...E separately from Tables 3 and 4. $M/K = 1.5$, final $F/Z = 0.56$.

Two weighted means are given. One of -5% was obtained by using the numbers at the foot of Table 3 as weighting factors. The other of -8% was obtained by using the corresponding numbers at the foot of Table 4.

The object of this study, therefore, is to determine the effect of the 40% reduction in effort on the combined length composition (A...E) from Tables 3 and 4) and to compare these with the values -5% and -8%, respectively.

For comparison with the mean value of -5%, Table 7b below shows the percentage changes in weight for a 40% reduction in effort for A...E combined, from Table 3 for three options.

Comparison with the correct value of -5% shows that with a mean value of $L_{\infty} = 58$ cm all three options underestimate the long-term benefit. For values of $L_{\infty} = 70$ and 90 cm, options 1 and 2 overestimate the long-term benefit but option 3 underestimates it.

Table 7c shows the result of similar calculations using the data in Table 4 for comparison with the value of -8%.

The results of this comparison are the same as those above, i.e., compared with the correct value of -8%, all three options underestimate the long-term benefit for $L_{\infty} = 58$ cm. For $L_{\infty} = 70$ and 90 cm, options 1 and 2 overestimate the long-term benefit but option 3 underestimates it.

SECOND INVESTIGATION OF A 40% REDUCTION IN EFFORT

Tables 8a, 8b and 8c show the results of a similar investigation using the simulated length compositions in Tables 5 and 6.

For comparison with the mean value of +25% in Table 8a, Table 8b below shows the percentage changes in weight for a 40% reduction in effort for sets A...E combined from Table 5 for three options.

Comparison with the correct value of +25% shows that with a mean value of $L_{\infty} = 83$ cm all three options underestimate it but not by very much. For larger values of L_{∞} , the tendency is for options 1 and 2 to overestimate the correct value but for option 3 to underestimate it.

For comparison with the correct value of +18% in Table 8a, Table 8c shows that for all three options, the tendency is to overestimate the long-term benefit.

LONG-TERM EFFECTS OF AN INCREASE IN MESH SIZE TO 90 MM

Tables 9 and 10 show the effect of an increase in mesh size to 90 mm using data from Tables 3-6.

For comparison with the value of +9% in Table 9a, Table 9b shows the results of calculations for sets A...E combined from Table 3 for three options.

Compared with the correct value of +9%, the mean value for $L_{\infty} = 58$ cm underestimates the long-term benefit for all three options. For larger values of L_{∞} options 1 and 2 overestimate and option 3 underestimates the long-term benefit.

Compared with the correct value of +6% in Table 9a, similar results were obtained in Table 9c as for Table 9b.

SECOND INVESTIGATION OF THE EFFECTS OF AN INCREASE IN MESH SIZE

Tables 10a, 10b and 10c show the results of a similar investigation using the data from Tables 5 and 6.

For comparison with the mean value of +30% in Table 10a, Table 10b shows the results for sets A...E combined for three options.

Comparison with the correct values of +30% shows that with a mean value of $L_{\infty} = 83$ cm all three options agree very closely with it (i.e., 31% compared with 30%). For larger values of L_{∞} , options 1 and 2 overestimate and option 3 tends to underestimate the correct long-term benefit.

For comparison with the correct value of +28% in Table 10a, Table 10c shows that for all values of L_{∞} and for three options, the long-term benefits are overestimated.

The results of this very limited study appear to point to the following conclusions, as far as estimates of the long-term effect of changes in exploitation pattern are concerned:

- 1) If the correct or mean value of L_{∞} is used along with the correct value of M/K, long-term effects may be still under- or overestimated.
- 2) If larger values of L_{∞} are used along with the correct values of M/K, long-term effects are liable to be underestimated.
- 3) If large-values of L_{∞} are used along with values of M/K based on compatible values of L_{∞} and K, long-term effects may or may not be overestimated. In general, an overestimate appears to be likely, if K is calculated from L_{∞} plus a growth increment for small fish. If the growth increment is based on growth over a mid-sized range of individuals, the likelihood of overestimating the long-term benefits appears to be considerably reduced.

Table 7. Investigation of the long-term effects of a 40% reduction in effort, based on simulated data in Tables 3 and 4.

a) Analysis based on data sets A-E (separately), from Tables 3 and 4; M/K = 1.5 and F/Z = 0.56.

Data set	L_{∞}	% change in weight
A	45	-11
B	52	-8
C	58	-5
D	65	-3
E	70	-1

weighted mean for Table 3

-5

weighted mean for Table 4

-8

b) Same as (a), but based on data sets A-E (combined) from Table 3 and three combinations of growth parameters.

L_{∞}	Option 1	Option 2	Option 3
58	-17	-17	-17
70	+11	+1	-14
90	+29	+9	-12

Option 1 : L_{∞} varied, M/K = 1.5, F/Z = 0.56.

Option 2 : L_{∞} varied and K selected to correspond to L_{∞} and an annual increment of 15 to 30 cm, M = 0.63, F/Z = 0.56.

Option 3 : L_{∞} varied and K selected to correspond to L_{∞} and an annual increment of 40 to 46 cm, M = 0.63, F/Z = 0.56.

c) Same as (b), but based on data in Table 4.

L_{∞}	Option 1	Option 2	Option 3
58	-14	-14	-14
70	+12	+1	-14
90	+31	+11	-12

Table 8. Investigation of the long-term effects of a 40% reduction in effort, based on simulated data in Tables 5 and 6.

a) Analysis based on data sets A-E (separately), from Tables 5 and 6; M/K = 1.5 and F/Z = 0.78.

Data set	L_{∞}	% change in weight
A	55	+12
B	69	+19
C	83	+26
D	94	+31
E	105	+35

weighted mean for Table 5

+25

weighted mean for Table 6

+18

b) Same as (a), but based on data sets A-E (combined) from Table 5 and three combinations of growth parameters.

L_{∞}	Option 1	Option 2	Option 3
83	+22	+22	+22
90	+30	+25	+21
100	+38	+28	+21
110	+44	+30	+20

Option 1 : L_{∞} varied, M/K = 1.5, F/Z = 0.78.

Option 2 : L_{∞} varied and K selected to correspond to L_{∞} and an annual increment of 25 to 33 cm, M = 0.225, F/Z = 0.78.

Option 3 : L_{∞} varied and K selected to correspond to L_{∞} and an annual increment of 50 to 54.5 cm, M = 0.225, F/Z = 0.78.

c) Same as (b), but based on data in Table 6.

L_{∞}	Option 1	Option 2	Option 3
83	+26	+26	+26
90	+34	+29	+25
100	+42	+32	+25
110	+47	+34	+24

Table 9. Investigation of the long-term effects of an increase in mesh size from 70 to 90 mm, based on simulated data in Tables 3 and 4.

a) Analysis based on data sets A-E (separately), from Tables 3 and 4; M/K = 1.5 and F/Z = 0.56.

Data set	L_∞	% change in weight
A	45	+1
B	52	+7
C	58	+10
D	65	+12
E	70	+13
weighted mean for Table 3		+9
weighted mean for Table 4		+6

b) Same as in (a), but based on data sets A-E (combined) from Table 3 and three combinations of growth parameters.

L_∞	Option 1	Option 2	Option 3
58	+5	+5	+5
70	+21	+14	+3
90	+29	+17	+1

Option 1 : L_∞ varied, M/K = 1.5, F/Z = 0.56.

Option 2 : L_∞ varied and K selected to correspond to L_∞ and an annual increment of 15 to 30 cm, M = 0.63, F/Z = 0.56.

Option 3 : L_∞ varied and K selected to correspond to L_∞ and an annual increment of 40 to 46 cm, M = 0.63, F/Z = 0.56.

c) Same as in (b), but based on data in Table 4.

L_∞	Option 1	Option 2	Option 3
58	+5	+5	+5
70	+25	+16	+3
90	+35	+20	+2

Table 10. Investigation of the long-term effects of an increase in mesh size from 70 to 90 mm, based on simulated data in Tables 5 and 6.

a) Analysis based on data sets A-E (separately), from Tables 5 and 6; M/K = 1.5 and F/Z = 0.78.

Data set	L_∞	% change in weight
A	55	+26
B	69	+30
C	83	+31
D	94	+30
E	105	+28
weighted mean for Table 5		+30
weighted mean for Table 6		+28

b) Same as in (a), but based on data sets A-E (combined) from Table 6 and three combinations of growth parameters.

L_∞	Option 1	Option 2	Option 3
83	+31	+31	+31
90	+35	+32	+30
100	+38	+32	+28
110	+40	+33	+27

Option 1 : L_∞ varied, M/K = 1.5, F/Z = 0.78.

Option 2 : L_∞ varied and K selected to correspond to L_∞ and an annual increment of 25 to 33 cm, M = 0.225, F/Z = 0.78.

Option 3 : L_∞ varied and K selected to correspond to L_∞ and an annual increment of 50 to 54.5 cm, M = 0.225, F/Z = 0.78.

c) Same as in (b), but based on data in Table 6.

L_∞	Option 1	Option 2	Option 3
83	+42	+42	+42
90	+46	+42	+40
100	+50	+43	+38
110	+53	+43	+36

OTHER APPLICATIONS OF LENGTH COHORT ANALYSIS

Use of yield-isopleth diagram

Length cohort analysis can be used without having separate values of M and K to determine the current value of F/Z. This can be applied to a yield isopleth diagram directly, if this is prepared using the tables of yield functions provided by Beverton and Holt (1964). They showed that their yield equation can be written with the principal part as a function simply of the three variables E, c and M/K, where E equals F/Z, and c equals L_c/L_∞ .

Investigations have been carried out using the data sets in Table 5. These extend over a reasonable length range and F/Z and Z are not influenced by selection for small fish. For sets A...E separately, cohort analysis correctly estimates $F/Z = 0.78$ for all length groups above the selection range when the correct input parameters are used (Table 11). This is as expected. Investigation also showed that if L_{∞} is overestimated, F/Z is overestimated and if L_{∞} is underestimated, F/Z is underestimated.

An L_{∞} of 83 cm corresponds to the mean of the values of L_{∞} used for sets A...E separately in Table 5. Using this value for L_{∞} the values of F/Z for fully exploited length groups obtained by cohort analysis above tended to be less than the correct value of 0.78 although not by very much (Table 11). For larger values of L_{∞} , larger values of F/Z were obtained for option 1. Option 2 gave values of F/Z quite close to 0.78 and option 3 gave values that were a little smaller.

Table 11. Cohort analysis based on estimates of F/Z for two different values of L_{∞} (83 and 150 cm) and three options (see Table 8b) based on catch-at-length data in Table 5.

L_{∞} lower length class limit (cm)	Option 1 ^a		Option 2		Option 3	
	83	150	83	150	83	150
5	0.01	0.02	0.01	0.01	0.01	0.004
10	0.08	0.17	0.08	0.07	0.08	0.04
15	0.39	0.61	0.39	0.36	0.39	0.25
20	0.70	0.86	0.70	0.69	0.70	0.57
25	0.76	0.90	0.76	0.76	0.76	0.67
30	0.76	0.90	0.76	0.78	0.76	0.68
35	0.75	0.90	0.75	0.78	0.75	0.69
40	0.74	0.90	0.74	0.79	0.74	0.70
45	0.72	0.91	0.72	0.79	0.72	0.70
50	0.70	0.91	0.67	0.80	0.67	0.72
55	0.67	0.91	0.67	0.80	0.67	0.72
60	0.64	0.91	0.64	0.80	0.63	0.72
65	0.62	0.92	0.61	0.81	0.61	0.74
70	—	0.92	—	0.82	—	0.75
75	—	0.91	—	0.81	—	0.74

^aSee footnote to Table 8b.

Methods based on the determination of Z/K or $(Z - K)$

Cohort analysis provides estimates of F/Z and $F\Delta t$ for each length group and is a suitable method of analysis when mortality is not constant with age. A different kind of approach is to assume that mortality is constant with age or length.

This can be done by a modification of the classical "catch curve" approach. When dealing with species that can be aged, a plot of the natural logarithms of the numbers-at-age against age gives a relationship for which the right hand part has a slope of $-Z$.

For species that cannot be aged, the equivalent plot of \log (numbers-at-length) against length is not likely to be linear, but forms a curve for which the slope at any point is a function of mortality rate and growth rate (Fig. 15).

There are various methods for analyzing curves such as those in Fig. 15. The simplest and the ones considered below are all based on an assumption of a deterministic relationship between length and age from which it is possible to derive estimates of the ratio Z/K or of $(Z - K)$.

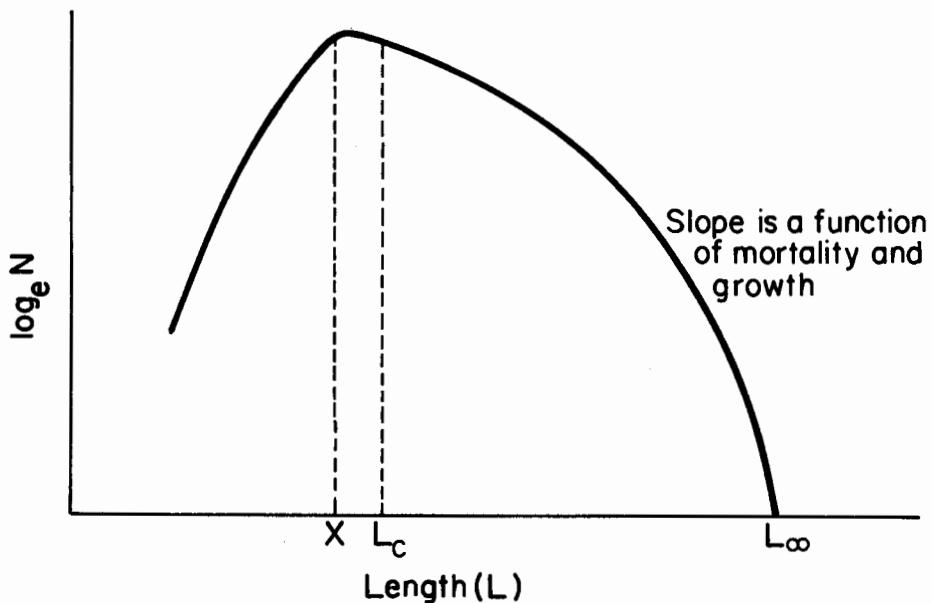


Fig. 15. Relationship between catch numbers and length under steady-state conditions.

Investigations using simulated data-transformation to a straight line

For this part of the study, investigations have been carried out using data from Tables 2-6. In each instance, to linearize the right hand part of each length composition, log numbers-at-length were plotted against the logs of $L_{\infty} - L'$, where L' refers to the center point of each length group. In theory, this should provide straight lines with slopes of $(Z/K) - 1$.

When real data are used, a plot of log numbers against the logarithms of $L_{\infty} - L'$ should be linear over at least the central part of the range. The slope of this portion should be equal to $(Z/K) - 1$. The curve is likely to depart from linearity due to the non-representative sampling of small fish, and due to the fact that the growth characteristics of large fish are likely to be atypical.

Fig. 16 shows an alternative method, not further discussed here, for estimated Z/K .

Estimates of Z/K (from slope of right-hand part of each length composition in Tables 2-6)

Estimates of Z/K were obtained from the plots of log numbers against $(L_{\infty} - L')$ by adopting the maximum slopes at the inflexions of the curves in those instances where the curves were sigmoid. Table 12 shows results for different values of L_{∞} using data from Table 2.

For the correct value of $L_{\infty} = 58$ cm, the estimated value of Z/K was close to the expected value. For larger values of L_{∞} the estimates of Z/K were close to the expected values provided K was calculated from L_{∞} and a growth increment near to the mid-part of the length range (i.e., option 2 in Table 12).

Table 13 shows the results of similar calculations using the combined length compositions A+B+C+D+E from Tables 3-6.

In each instance, the mean value of L_{∞} was adopted (i.e., 58 cm for Tables 3 and 4 and 83 cm for Tables 5 and 6). Using these values of L_{∞} , the estimated values of L_{∞} were close to the expected values for Tables 3 and 5, but larger than the expected values for Tables 4 and 6.

Table 12. Estimates of Z/K^a from the maximum slopes of transformed data (see Fig. 1b and text).

Input values L_∞	K^b	K^c	Z	Expected values		Estimated values of Z/K	Size range included in plot (cm)
				Z/K^b	Z/K^c		
58	0.42	0.42	0.1	0.2	0.2		
			0.5	1.2	1.2	1.6	42-48
			1.0	2.3	2.3	3.1	44-48
			1.4	3.3	3.3	3.6	44-48
			1.6	3.8	3.8	3.9	44-48
			2.0	4.8	4.8	4.6	42-52
65	0.36	0.27	0.5	1.4	1.8	1.9	42-48
			1.0	2.8	3.7	3.9	44-48
			1.4	3.9	5.2	5.2	44-48
			1.6	4.4	5.9	5.6	46-50
			2.0	5.6	7.4	7.1	46-52
100	0.19	0.11	0.5	2.6	4.5	3.6	42-48
			1.0	5.3	9.1	9.1	44-48
			1.4	7.4	12.8	13.6	46-50
			1.6	8.4	14.6	15.7	48-52
			2.0	10.5	18.2	19.1	42-52

^aBased on simulated catch-at-length data in Table 2.^bValues of K selected as consistent with L_∞ and an annual increment of 15 to 30 cm.^cValues of K selected as consistent with an annual increment of 40 to 46 cm.Table 13. Comparison of expected values with estimates of Z/K from the maximum slope of transformed data (see Fig. 1b and text).

Data source (Table #)	L_∞ (cm)	Expected values of Z/K	Estimated values of Z/K	Size range included in plot (cm)
3	58	3.4	3.5	30-40
4	58	3.4	4.3	30-38
5	83	7.0	7.0	30-40
6	83	7.0	9.5	25-40

Method of Beverton and Holt (1956)

Beverton and Holt (1956) show that for the right hand part of Fig. 15, an estimate of Z/K is given by:

$$Z/K = (L_\infty - \bar{L}) / (\bar{L} - L_c) \quad \dots 3)$$

where \bar{L} is the mean length of individual in the range $L_c \dots L_\infty$ and is calculated from

$$\bar{L} = f\bar{L}' / \sum f$$

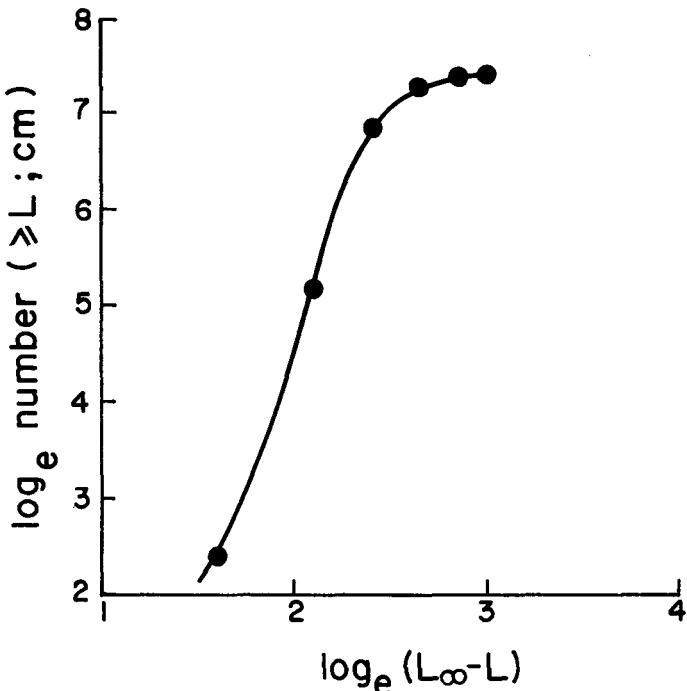


Fig. 16. An alternative approach for estimation of Z/K , which is equal to the slope of the central, linear section of a plot of cumulative numbers $\geq L$ vs. $\log_e L_\infty - L$ (from Jones 1984).

L' being the mid-point of each length group and f the number of observations in each length group. L_c is any length on the right hand part of a curve such as that in Fig. 15. L_c is often referred to as the length at first capture. This is likely to be misleading, however, since it may cause one to adopt a length on the left hand part of the curve (i.e., to the left of X in Fig. 15). To avoid bias, L_c should be chosen somewhere on the right hand part of the curve. In practice, as much as possible of the right hand part of the curve should be included in the calculations, so that L_c should be as close to X as possible as seems reasonable for the data concerned.

Estimates of Z/K (using Beverton and Holt formula)

Estimates of Z/K were obtained from the data in Table 2 using the Beverton and Holt formula (equation 3). The results are shown in Fig. 14. Table 15 shows similar results using data from Tables 3-6.

For all data sets, estimated values of Z/K were reasonably close to the expected values when mean values of L_∞ were used (i.e., 58 cm for Tables 3 and 4 and 83 cm for Tables 5 and 6). For larger values of L_∞ , the estimated values of Z/K can also be close to the expected values. The results show, however, that this result is critically dependent on the part of the length range from which a growth increment is selected for determining K .

Overall, results for estimates of Z/K lead to similar conclusions to those reached in the previous section. Reasonable estimates of Z/K can be obtained. However the derivation of Z from Z/K is critically dependent on the choice of value of K for a given input value of L_∞ . If K is estimated from L_∞ along with one annual increment, this increment should relate to growth over the mid-part of the length range.

Table 14. Comparison of expected values with estimates^a of Z/K based on the Beverton and Holt mean length formula (equation 3).

Input values		K ^c	Z	Expected values		Estimated values of Z/K
L _∞	K ^b			Z/K ^b	Z/K ^c	
58	0.42	0.42	0.1	0.2	0.2	0.33
			0.5	1.2	1.2	1.1
			1.0	2.3	2.3	2.1
			1.4	3.3	3.3	2.9
			1.6	3.8	3.8	3.2
			2.0	4.8	4.8	4.4
65	0.36	0.27	1.4	3.9	5.2	4.5
			1.6	4.4	5.9	4.9
			2.0	5.6	7.6	6.7
75	0.29	0.19	0.1	0.35	0.53	1.7
			0.5	1.7	2.6	3.1
			1.0	3.4	5.3	5.1
			1.4	4.8	7.4	7.0
			1.6	5.5	8.4	7.6
			2.0	6.9	10.5	10.2
100	0.19	0.11	0.1	0.53	0.91	3.8
			0.5	2.6	4.5	6.3
			1.0	5.3	9.1	9.9
			1.4	7.4	12.8	13.2
			1.6	8.4	14.6	14.4
			2.0	10.5	18.2	19.0

^aBased on catch-at-length data in Table 2; all fish considered ≥ 42 cm.

^bValue of K selected as consistent with L_∞ and an annual increment of 15 to 30 cm.

^cValue of K selected as consistent with L_∞ and an annual increment of 40 to 46 cm.

Table 15. Comparisons of expected values with estimates of Z/K from maximum slope of transformed data^a (see Fig. 1b and text).

Source of data (Table #)	L _∞	Expected values of Z/K	Estimated values of Z/K
3	58 ^b	3.4	3.1
3	70 ^b	5.0	4.9
4	58 ^b	3.4	3.5
4	70 ^b	5.0	5.5
5	83	7.0	6.2
5	110 ^c	12.5	9.8
6	83	7.0	7.3
6	110 ^c	12.5	11.5

^aBased on catch-at-length data in Tables 3-6, and on fish > 30 cm.

^bValues of K selected as consistent with L_∞ and annual growth increment from 30 to 40 cm.

^cValues of K selected as consistent with L_∞ and annual growth increment from 35 to 41 cm.

References

- Beverton, R.J.H. and S.J. Holt. 1956. A review of methods for estimating mortality rates in exploited fish populations with special references to sources of bias in catch sampling. *Rapp. P.-V. Réun. Cons. Int. Explor. Mer* 140: 67-83.
- Jones, R. 1984. Assessing the effect of changes in exploitation patterns using length composition data (with notes on VPA and cohort analysis). *FAO Fish. Tech. Pap.* 256. 118 p.
- Pauly, D. 1984. Fish population dynamics in tropical waters: a manual for use with programmable calculators. *ICLARM Studies and Reviews* 8. 325 p.
- Pope, J.G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. *Res. Bull. Int. Comm. North. Atl. Fish.* 9: 65-74.

Analytical Investigations of Errors in Mortality Rates Estimated from Length Distributions of Catches

A. LAUREC

B. MESNIL

Centre IFREMER de Nantes

B.P. 1049 44037

Nantes, Cédex, France

Laurec, A. and B. Mesnil. 1987. Analytical investigations of errors in mortality rates estimated from length distributions of catches, p. 239-281. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

Taylor's expansions offer an efficient tool to investigate some of the sensitivity problems encountered in the estimation of mortalities from length distribution when a von Bertalanffy growth model is assumed.

It was found that large class intervals had a very negative impact on the results and should be replaced by smoothing wherever possible.

Risk of bias associated with sampling and estimation error of K , L_{∞} and of catch-at-length data is usually moderate, except near L_{∞} . The same is true for individual variabilities in growth parameters.

This and related problems should be investigated when conducting a length-based analysis, for each specific case; the method presented by the authors allows such investigation and can be generalized to applications involving yield-per-recruit calculations.

Introduction

Sensitivity analyses are now generally conducted by stochastic simulations. The power of such tools has been widely demonstrated. Approximations through Taylor's expansion, when feasible, can also be very useful. Of course they do not give exact results, but they can offer an analytical understanding of the results, which are only observed by simulations.

This paper is devoted to an extensive use of Taylor's expansions to address a number of sensitivity problems arising in the estimation of fishing mortalities from length-frequency distribution of catches. Von Bertalanffy growth curves will be assumed.

The first part considers the simple situation where mortalities at length are constant. The second and larger part deals with Jones' length-cohort analysis. Simple microcomputer programs have been developed to provide numerical illustrations which, in the present case, are based on a widely published data set (Jones 1979; 1981).

Several equations have been considered important enough to be given a number, while other equations which are "locally" important are associated with capital letters (A, B, ...) and are only referred to within the section where they appear.

THE SIMPLE CASE: BEVERTON AND HOLT'S FORMULA

Beverton and Holt's formula is based on the relationships:

$$Z = K \cdot \frac{(L_{\infty} - \bar{l})}{\bar{l} - l_c} \quad \dots 1)$$

or its equivalent

$$\bar{l} - l_c = \frac{K}{Z + K} \cdot (L_{\infty} - l_c) \quad \dots 2)$$

l_c being the length at first capture, and \bar{l} the mean length in the catches.

These equations are valid in general provided that:

- \bar{l} is accurately known, free of sampling error, and without any discrepancy due to the approximation of a continuous length distribution by a histogram with finite step.
- Growth of all fish larger than l_c is adequately described by the von Bertalanffy formula (the suitability of which is not discussed here), in which the parameters K and L_{∞} are supposed to be perfectly known, no individual variability being allowed for (see below).

Impact of Discretisation¹

A histogram being given with discrete bounds l_i , $i = 1 \dots I + 1$, the use of the mean lengths within each interval (l_i, l_{i+1}) , as estimated by

$$\bar{l}_i = \frac{l_i + l_{i+1}}{2}$$

can only yield an approximation $\hat{\bar{l}}$ to the true \bar{l} , and the wider the grouping intervals, the larger the difference can be.

It is possible however to define the true relationship between $\hat{\bar{l}}$ and Z , and thus to deduce Z from $\hat{\bar{l}}$.

The number $N(l_i)$ of individuals attaining the length l_i is given by

$$N(l_i) = N_c \cdot \left(\frac{L_{\infty} - l_c}{L_{\infty} - l_i} \right)^{-Z/K} = N_c \left(\frac{L_{\infty} - l_i}{L_{\infty} - l_c} \right)^{Z/K}$$

¹Our attention was drawn on the importance of this effect by our colleague, Y. Morizur (pers. comm.) (Centre IFREMER de Brest, B.P. 337 29273, Brest Cédex).

The number of individuals the size of which falls within the interval $[l_i, l_{i+1}]$ is equal to:

$$\frac{N_c}{(L_\infty - l_c)^{Z/K}} \cdot \left[(L_\infty - l_i)^{Z/K} - (L_\infty - l_{i+1})^{Z/K} \right] \text{ for } i = 1, \dots, I.$$

As $l_{I+1} \geq L_\infty$ on the last interval the number is:

$$N(l_I) = \frac{N_c}{(L_\infty - l_c)^{Z/K}} \cdot (L_\infty - l_I)^{Z/K}$$

Those individuals are given the average length $\frac{l_I + L_\infty}{2}$

Thus:

$$\tilde{\bar{l}} = \frac{1}{(L_\infty - l_c)^{Z/K}} \sum_{i=1}^{I-1} \frac{(l_i + l_{i+1})}{2} \left[(L_\infty - l_i)^{Z/K} - (L_\infty - l_{i+1})^{Z/K} \right] + (L_\infty - l_I)^{Z/K} \frac{(l_I + L_\infty)}{2}$$

which can be viewed as a function $g(Z)$.

Approximations can be used but K and L_∞ being known, it remains possible to solve numerically the equation $g(Z) = \bar{l}$ for an observed \bar{l} .

When the width of the intervals (l_i, l_{i+1}) decreases to 0, $\tilde{\bar{l}}$ approaches $\bar{l} = l_c + \frac{K}{Z + K}$ ($L_\infty - l_c$) and we are back with Beverton and Holt's formula.

With $K = 0.5$ and $L_\infty = 70$ (see *Nephrops*-like data set, Table 1), Fig. 1 was built to show the effect upon \bar{l} of varying the length-class width. $\bar{l}(Z)$ (solid line) is plotted together with $\tilde{\bar{l}}$ calculated for a width of 1 (dashes and dots) and of 5 (dashes).

When Beverton and Holt's formula is used without taking account the length grouping step, the difference between the curves is merely neglected although it can be quite considerable for large values of Z .

Sampling Errors on \bar{l}

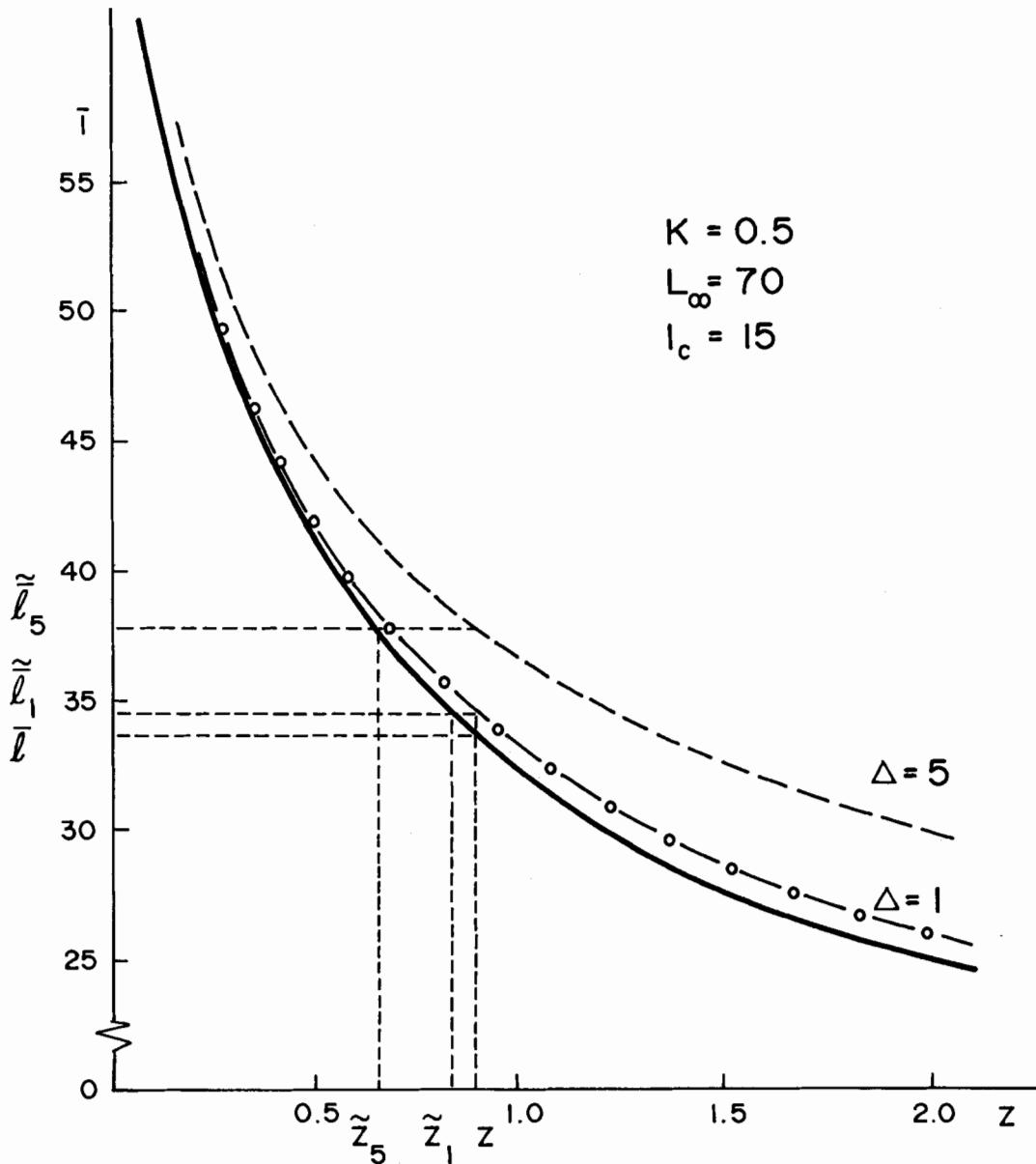
Let \hat{l} be a sample estimate for \bar{l} , and \hat{Z} the estimate of total mortality calculated on this basis. Formula (1) can be rewritten as:

$$\hat{Z} = K \left(\frac{L_\infty - l_c}{\hat{l} - l_c} - 1 \right) = -K + K \left(\frac{L_\infty - l_c}{\bar{l} - l_c} \right) \left(\frac{1}{1 + \hat{\Delta}} \right) = -K + (Z + K) \left(\frac{1}{1 + \hat{\Delta}} \right)$$

where $\hat{\Delta} = \frac{\hat{l} - \bar{l}}{\bar{l} - l_c}$

$\frac{1}{1 + \hat{\Delta}}$ can be expanded as $1 - \hat{\Delta} + \hat{\Delta}^2$ so that $\hat{Z} \approx -K + (Z + K)(1 - \hat{\Delta} + \hat{\Delta}^2)$

or $\hat{Z} \approx Z + (Z + K)(- \hat{\Delta} + \hat{\Delta}^2)$



Z = True mortality.

$\tilde{\ell}_1$ = Apparent mean length using a length-class width of 1.

\tilde{Z}_1 = Apparent mortality corresponding to $\tilde{\ell}_1$.

$\tilde{\ell}_5$ = Apparent mean length using a length-class width of 5.

\tilde{Z}_5 = Apparent mortality corresponding to $\tilde{\ell}_5$.

Fig. 1. Plot of ℓ vs. Z . Comparison of actual values from Beverton and Holt's equation and results observed for two different length grouping intervals.

Table 1. The standard data sets used in the examples.

L	Step : 1		Step : 2		Step : 5			
	Catch	Δt	L	Catch	Δt	L	Catch	Δt
14	1	0.036	14	19	0.073	10	1	0.174
15	18	0.037	16	64	0.075	15	163	0.191
16	25	0.037	18	81	0.078	20	1,390	0.211
17	39	0.038	20	163	0.082	25	4,120	0.236
18	46	0.039	22	702	0.085	30	4,730	0.267
19	35	0.040	24	1,340	0.089	35	3,040	0.308
20	68	0.040	26	1,738	0.093	40	1,650	0.365
21	95	0.041	28	1,567	0.098	45	827	0.446
22	342	0.042	30	1,777	0.103	50	312	0.575
23	360	0.043	32	2,053	0.108	55	94	0.811
24	525	0.044	34	1,720	0.114	60	10	0.386
25	815	0.045	36	1,345	0.121	65+	3	0.386
26	1,012	0.046	38	875	0.129			
27	726	0.047	40	668	0.138		16,340	
28	760	0.048	42	685	0.148			
29	807	0.049	44	506	0.160			
30	850	0.051	46	352	0.174			
31	927	0.052	48	266	0.191			
32	1,023	0.053	50	161	0.211			
33	1,030	0.055	52	115	0.236			
34	900	0.056	54	64	0.267			
35	820	0.058	56	42	0.308			
36	715	0.060	58	24	0.365			
37	630	0.062	60	6	0.446			
38	505	0.063	62	3	0.575			
39	370	0.066	64+	4				
40	320	0.068						
41	348	0.070						
42	365	0.073						
43	320	0.075						
44	297	0.078						
45	209	0.082						
46	185	0.085						
47	167	0.089						
48	149	0.093						
49	117	0.098						
50	85	0.103						
51	76	0.108						
52	61	0.114						
53	54	0.121						
54	36	0.129						
55	28	0.138						
56	24	0.148						
57	18	0.160						
58	15	0.174						
59	9	0.191						
60	4	0.211						
61	2	0.236						
62	2	0.267						
63	1	0.308						
64	1	0.365						
65+	3							

— Catches in number at length reconstructed from Jones (1979, 1981).

— Δt : time required to grow from lower to upper limit of each group as calculated with the parameters: $L_\infty = 70$; $K = 0.5$.— $M = 0.2$.

from which

$$E(\hat{Z}) \approx -K + (Z + K)(1 + \text{Var}(\hat{\Delta})) = Z + (Z + K)\text{Var}(\hat{\Delta})$$

or

$$\frac{1}{Z}(E(\hat{Z}) - Z) \approx \frac{Z + K}{Z}\text{Var}(\hat{\Delta})$$

which shows the existence of a (generally small) bias. In the same way:

$$\text{Var}(\hat{Z}) \approx (Z + K)^2 \text{Var}(\hat{\Delta})$$

$\text{Var}(\hat{\Delta})$ is thus the critical parameter. It can be calculated from $\text{Var}(\hat{l})$ since

$$\text{Var}(\hat{\Delta}) = \frac{1}{(\bar{l} - l_c)^2} \text{Var}(\hat{l})$$

If p_i is the true frequency in the interval (l_i, l_{i+1}) and \hat{p}_i the estimated one, then:

$$\hat{l} = \sum_i \hat{p}_i \bar{l}_i \quad \text{and} \quad \text{Var}(\hat{l}) = \sum_{i,j} \bar{l}_i \bar{l}_j \text{Cov}(\hat{p}_i, \hat{p}_j)$$

As an example, one can take the case where \bar{l} is estimated from a simple random sample of size n .

$$\text{In this scheme : } \text{Var}(\hat{p}_i) = \frac{1}{n} \cdot p_i \cdot (1 - p_i)$$

$$\text{Cov}(\hat{p}_i, \hat{p}_j) = -\frac{1}{n} \cdot p_i \cdot p_j \quad \text{for } i \neq j$$

$$\text{therefore} \quad \text{Var}(\hat{l}) = \frac{1}{n} \left[\sum_i p_i \cdot \bar{l}_i^2 - (\sum_i p_i \cdot \bar{l}_i)^2 \right] = \frac{1}{n} \left[\sum_i p_i \bar{l}_i^2 - \bar{l}^2 \right]$$

This variance can be approximated in a simplified approach whereby the effect of discretisation is neglected. This leads to the formulae given by Hoenig et al. (1983), Wetherall et al. (Part I, this vol.) and Hoenig (Part I, this vol.):

$$\text{var}(\hat{\Delta}) = \frac{\text{Var}(\hat{l} - l_c)}{(\bar{l} - l_c)^2} \approx \frac{1}{n} \cdot \frac{Z}{Z + 2K}$$

$$\text{Thus } \frac{1}{Z}(E(\hat{Z}) - Z) = \frac{1}{n} \cdot \frac{Z + K}{Z + 2K} \quad (\text{A})$$

and

$$\text{Var}(\hat{Z}) \approx \frac{(Z+K)^2}{n} - \frac{Z}{Z+2K}$$

or

$$: \quad \frac{\text{Var}(\hat{Z})}{Z^2} \approx \frac{1}{n} - \frac{(Z+K)^2}{Z(Z+2K)} \quad (\text{B})$$

Coming back to the example given in I.1 with $K = .5$, $Z = .8$

$$\frac{1}{Z} (E(\hat{Z}) - Z) \approx \frac{.72}{n} \quad \text{and} \quad \frac{\text{Var}(\hat{Z})}{Z^2} \approx \frac{1.25}{n}$$

With $n = 100$ the bias is very small (.0072) and the coefficient of variation of \hat{Z} close to 11%.

When l is too close to l_c a first order Taylor's expansion may lead to poor approximations (Majkowski 1982).

Higher order expansions may, thus, be necessary, for instance:

$$\hat{Z} \approx -K + (Z+K)(1 - \hat{\Delta} + \hat{\Delta}^2 + \hat{\Delta}^3 - \hat{\Delta}^4) = Z + (Z+K)(-\hat{\Delta} + \hat{\Delta}^2 - \hat{\Delta}^3 + \hat{\Delta}^4)$$

Since \hat{l} is an arithmetic mean, normality can be used as an approximation, so that:

$$E(\hat{\Delta}^3) \approx 0 \text{ and } E(\hat{\Delta}^4) \approx 3(E(\hat{\Delta}^2))^2$$

$$E(\hat{Z}) \approx Z + (Z+K)(\text{Var}(\hat{\Delta}) + 3\text{Var}(\hat{\Delta})^2)$$

In the case of simple random sampling where: $\text{Var}(\hat{\Delta}) \approx \frac{1}{n} \frac{Z}{Z+2K}$ this leads to:

$$E(\hat{Z}) \approx Z + (Z+K) \left(\frac{1}{n} \frac{Z}{Z+2K} + \frac{3}{n^2} \frac{Z^2}{(Z+2K)^2} \right)$$

$$\frac{1}{Z} (E(\hat{Z}) - Z) \approx \frac{(Z+K)}{n(Z+2K)} + \frac{3}{n^2} \cdot \frac{Z(Z+K)}{(Z+2K)^2} \quad (\text{C})$$

The term $\frac{3}{n^2} \cdot \frac{Z(Z+K)}{(Z+2K)^2}$ is negligible when sample size increases. This suggests that formula (A) is a robust approximation. In the previous example $\frac{3}{n^2} \cdot \frac{Z(Z+K)}{(Z+2K)^2} = \frac{.96}{(100)^2}$ which is negligible. Similar derivations suggest the robustness of formula (B).

Uncertainty Regarding K and L_∞

Let \hat{K} and \hat{L}_∞ be estimates of K and L_∞ , and \hat{Z} the estimate of total mortality rate Z given by equation (1).

Formulas for expected value of a product imply that:

$$E(\hat{Z}) = Z + \frac{1}{(1-l_c)} \cdot \text{Cov}(\hat{K}, \hat{L}_\infty)$$

where the bias is equal to $\frac{1}{(\bar{l} - l_c)} \text{Cov}(\hat{K}, \hat{L}_\infty)$

appears. Further developments show the relationship:

$$\frac{E(\hat{Z}) - Z}{Z} = \frac{\text{Cov.}(\hat{K}, \hat{L}_\infty)}{K \cdot (L_\infty - \bar{l})}$$

A very simple Taylor expansion yields:

$$\text{Var}(\hat{Z}) \approx \frac{1}{(\bar{l} - l_c)^2} \left[K^2 \cdot \text{Var}(\hat{L}_\infty) + 2(L_\infty - \bar{l}) \cdot K \cdot \text{Cov}(\hat{K}, \hat{L}_\infty) + (L_\infty - \bar{l})^2 \cdot \text{Var}(\hat{K}) \right]$$

or a squared coefficient of variation:

$$\frac{\text{Var}(\hat{Z})}{Z^2} = \left[\frac{\text{Var}(\hat{K})}{K^2} + 2 \cdot \frac{\text{Cov}(\hat{K}, \hat{L}_\infty)}{K \cdot (L_\infty - \bar{l})} + \frac{\text{Var}(\hat{L}_\infty)}{(L_\infty - \bar{l})^2} \right]$$

VARIABILITY IN THE GROWTH PATTERNS

Individual Variability

We consider that each individual within the population has its own pair of values K and L_∞ , and that the distribution of K and L_∞ follows a density function $\omega(K, L_\infty)$. The mean length corresponding to given K and L_∞ , \bar{l} , is also a function of K and L_∞ . The overall mean in the population is written $\bar{l} = E(\bar{l})$ and \bar{K} and \bar{L}_∞ stand for the expected values of each parameter, respectively. Z is assumed to be constant. Calculating the expectations of both sides of equation (2) yields

$$(\bar{l} - l_c) = E \left[\frac{\bar{L}_\infty - l_c}{1 + Z/\bar{K}} \right] \quad (\text{A})$$

For any given distribution function $\Omega(K, L_\infty)$ it is possible, at least numerically, to calculate the right-hand side. Generally it will not be equal to

$$\left[\frac{\bar{L}_\infty - l_c}{1 + Z/\bar{K}} \right]$$

as the nonlinearity introduces a bias.

In other words, the problem of individual variability is generally not solved by directly applying Beverton and Holt's formula using \bar{K} and \bar{L}_∞ .

It should be noticed that the problem does not arise directly from L_∞ . If the variability in K cancels out, equation (A) turns to:

$$(\bar{l} - l_c) = \frac{1}{1 + \frac{Z}{K}} \cdot E(L_\infty - l_c) = \frac{K}{Z + K} (\bar{L}_\infty - l_c)$$

In this case the variability in L_∞ affects in similar fashion both ($\bar{l} - l_c$) and $L_\infty - l_c$.

Coming back to the general case, we write:

$$f(K, L_\infty) = \frac{K}{Z + K} (L_\infty - l_c)$$

By Taylor's expansion, we get:

$$\bar{l} - l_c \approx f(\bar{K}, \bar{L}_\infty) + \frac{1}{2} \frac{\partial^2 f}{\partial K^2} \text{Var}(K) + \frac{\partial^2 f}{\partial K \cdot \partial L_\infty} \text{Cov}(K, L_\infty) + \frac{1}{2} \frac{\partial^2 f}{\partial L_\infty^2} \text{Var}(L_\infty)$$

and it can be shown that:

$$\frac{\partial^2 f}{\partial K^2} = -2 \cdot \frac{Z(L_\infty - l_c)}{(Z + K)^3}$$

$$\frac{\partial^2 f}{\partial K \cdot \partial L_\infty} = \frac{Z}{(Z + K)^2}$$

and

$$\frac{\partial^2 f}{\partial L_\infty^2} = 0$$

therefore:

$$\bar{l} - l_c \approx \frac{\bar{K}(\bar{L}_\infty - l_c)}{Z + \bar{K}} + \frac{Z}{(Z + \bar{K})^2} \text{Cov}(K, L_\infty) - \frac{Z}{(Z + \bar{K})^3} (\bar{L}_\infty - l_c) \cdot \text{Var}(K).$$

or, with the notation:

$$\Delta = \frac{Z}{Z + \bar{K}} \cdot \frac{\text{Cov}(K, L_\infty)}{\bar{K} \cdot (\bar{L}_\infty - l_c)} - \frac{Z}{\bar{K} \cdot (Z + \bar{K})^2} \text{Var}(K)$$

$$\bar{l} - l_c = \frac{\bar{K}(\bar{L}_\infty - l_c)}{Z + \bar{K}} \cdot (1 + \Delta) \quad (B)$$

A simple application of equation (1) with a given set \bar{l} , \bar{K} and \bar{L}_∞ would yield an apparent total mortality rate:

$$\tilde{Z} = \bar{K} \cdot \frac{(\bar{L}_\infty - \bar{l})}{(\bar{l} - l_c)}$$

also given by

$$\tilde{Z} + \bar{K} = \bar{K} \left[\frac{\bar{L}_\infty - l_c}{\bar{l} - l_c} \right]$$

Using formula (B) and another approximation:

$$\frac{1}{1 + \Delta} \approx 1 - \Delta$$

we get:

$$\frac{\tilde{Z} - Z}{Z} \approx \frac{\bar{K}}{Z + \bar{K}} \cdot \frac{\text{Var}(K)}{\bar{K}^2} - \frac{\text{Cov}(K, L_\infty)}{\bar{K}(\bar{L}_\infty - l_c)} \quad (C)$$

If K and L_∞ were independent, (C) would be simplified into:

$$\frac{\tilde{Z} - Z}{Z} \approx \frac{\bar{K}}{Z + \bar{K}} \cdot \frac{\text{Var}(K)}{\bar{K}^2}$$

This shows first that the bias is positive, and secondly that it is moderate. Even with a coefficient of variation of 20% for K and thus

$$\frac{\text{Var}(K)}{\bar{K}^2} = .04, \quad \frac{\tilde{Z} - Z}{Z} \approx \frac{.04}{1 + \frac{Z}{\bar{K}}} < .04$$

If for example $\bar{K} = 0.5$ and $Z = 0.8$, $\frac{\tilde{Z} - Z}{Z}$ comes to about 1.5%.

Variability Between Sexes or Subsets

In cases where males and females have different growth curves but are affected by the same mortality rate Z , the mean lengths will differ between sexes. Write K^1, L^1_∞ and \bar{l}^1 (resp. K^2, L^2_∞ and \bar{l}^2) for the values pertaining to the males (resp. to the females).

Equation (2) leads to two relationships:

$$(Z + K^1) \cdot (\bar{l}^1 - l_c) = K^1 \cdot (L^1_\infty - l_c) \quad (A)$$

$$(Z + K^2) \cdot (\bar{l}^2 - l_c) = K^2 \cdot (L^2_\infty - l_c) \quad (B)$$

The proportions p_1 and p_2 of males and females in the catches being known, we have: $\bar{l} = p_1 \cdot \bar{l}^1 + p_2 \cdot \bar{l}^2$ and the relationships (A) and (B) can be combined to yield:

$$Z \cdot (\bar{l} - l_c) + p_1 K^1 (\bar{l}^1 - l_c) + p_2 K^2 (\bar{l}^2 - l_c) = p_1 K^1 (L^1_\infty - l_c) + p_2 K^2 (L^2_\infty - l_c)$$

If we put $K^1 = K^2 = K$, this leads to:

$$(Z + K) \cdot (\bar{l} - l_c) = K \left((p_1 \cdot L^1_\infty + p_2 \cdot L^2_\infty) - l_c \right)$$

Equally, an analogy to the basic formula is found when $\bar{L}_\infty = p_1 \cdot L_{\infty}^1 + p_2 \cdot L_{\infty}^2$, a result already observed when individual variability only affects L_∞ .

However, when $K^1 \neq K^2$ there is no simple means to find an analogy with equation (1) that would be free of error.

In any case, the difference should be moderate. To illustrate this, use an example with $Z = 0.8$, $l_c = 14$, $L_\infty = 70$, $K^1 = 0.5$, $K^2 = 0.3$ and $p_1 = p_2 = 0.5$.

One can verify that $l_1 = 35.54$, and $l_2 = 29.27$, and thus $\bar{l} = 32.4$. Applying formula (1) and taking an average $\bar{K} = 0.4$ would result in an apparent total mortality $Z = 0.817$ instead of 0.8. The bias is small. It can also be noticed that formula (C) above is applied, the variability between sexes being considered as one case of individual variability with $\text{Var}(K) = .01$, formula (C) would predict $Z = 0.816$. This formula, thus, seems to provide a good approximation. Overall, it appears that Beverton and Holt's formula seems quite robust with respect to growth variability.

Cohort Analysis

THE BASIC EQUATION

Equation Using Time

Although it was certainly not the authors' intention, Jones' length-cohort analysis is sometimes thought of as a derivation of Pope's (1972) cohort analysis. As such, it would seem to be dependent on the latter's approximation. In order to clarify the issue, restart from the basic equations.

Let N_c be the size in numbers of a cohort at age t_c . The catches from age t_c to t will be noted $C(t)$ (bold letter) as opposed to the instantaneous rate $c(t)$ where $c(t) = dC(t)/dt$

Likewise the instantaneous rates of mortality $F(t)$ and $M(t)$ will contrast with $F(t)$ and $M(t)$ integrated over ages t_c to t . Beyond age T , $F(t)$ and $M(t)$, then $Z(t)$ considered as constant, will be noted F^+ , M^+ , Z^+ . Lastly, the cumulated catches beyond time T will be noted C^+ . $E = F^+/Z^+$ is the terminal exploitation rate, assumed to be constant, or more generally the ratio between $N(T)$ and C^+ .

$$E = C^+ / N(T)$$

In this context, the basic equation can be written:

$$N_c = \int_{t_c}^T c(t) \cdot e^{M(t)} \cdot dt + e^{M(T)} \cdot \frac{C^+}{E} \quad \dots 1)$$

To prove this:

$$c(t) = F(t) \cdot N(t)$$

and

$$N(t) = N_c \cdot e^{-(F(t) + M(t))}$$

Thus

$$\begin{aligned} \int_{t_c}^T c(t) \cdot e^{M(t)} \cdot dt &= N_c \int_{t_c}^T F(t) \cdot e^{- (F(t) + M(t))} \cdot e^{M(t)} \cdot dt \\ &= N_c \int_{t_c}^T F(t) \cdot e^{- F(t)} \cdot dt \end{aligned}$$

Since, by definition, $F(t) = dF(t)/dt$, the term under the integral is simply the derivative of $-e^{-F(t)}$ and thus:

$$N_c \int_{t_c}^T F(t) \cdot e^{- F(t)} \cdot dt = -N_c \left(e^{-F(t)} \right)_{t_c}^T = N_c \cdot \left(1 - e^{-F(T)} \right)$$

$$\text{Similarly, } C^+ = E \cdot N(T) = E \cdot N_c \cdot e^{-(F(T) + M(T))}$$

Thus, the second term in the right hand side of equation (1) is equal to:

$$N_c \cdot e^{-F(T)}$$

$$\text{yielding: } N_c = N_c \cdot 1 - e^{-F(T)} + N_c \cdot e^{-F(T)} = N_c$$

and this establishes the proof.

If M is a constant, then $M(t) = M \cdot (t - t_c)$ and equation (1) reduces to

$$N_c = \int_{t_c}^T c(t) \cdot e^{M(t-t_c)} \cdot dt + \frac{C^+}{E} \cdot e^{M(T-t_c)} \quad \dots 2)$$

If the integration is carried out on discrete steps instead of the continuous function, expressions similar to Pope's simplified equation can be obtained.

Moreover, equations (1) and (2) can be used to develop convergence formulas with respect to variations in E as was done by Pope.

Consider variations $\Delta(E)$ or, equally $\Delta(1/E)$. Since only the last term is subject to such variations, one gets:

$$\begin{aligned} \Delta(N_c) &= \Delta\left(\frac{1}{E}\right) \cdot C^+ \cdot e^{M(T)} \\ &= \Delta\left(\frac{1}{E}\right) \cdot E \cdot N(T) \cdot e^{M(T)} = \Delta\left(\frac{1}{E}\right) \cdot E \cdot N_c \cdot e^{-F(T)} \end{aligned}$$

or, in relative terms:

$$\frac{\Delta(N_c)}{N_c} = e^{-F(T)} \cdot \frac{\Delta\left(\frac{1}{E}\right)}{\left(\frac{1}{E}\right)} \approx -e^{-F(T)} \cdot \frac{\Delta(E)}{E}$$

where the convergence is again demonstrated to depend on the term in cumulative fishing mortality $e^{-F(T)}$.

Since $F(t) = \frac{c(t)}{N(t)}$ a similar expression can be obtained for the fishing mortalities:

$$\frac{\Delta(F(t_c))}{F(t_c)} = -\frac{\Delta(N_c)}{N_c} = e^{-F(T)} \cdot \frac{\Delta(E)}{E}$$

In both formulas, t_c can be replaced by any age t and the general relationship for $F(t)$ written as:

$$\frac{\Delta(F(t))}{F(t)} \approx e^{-F(t, T)} \cdot \frac{\Delta(E)}{E}$$

where $F(t, T)$ is the cumulative fishing mortality from age t to T .

Equation Using Lengths

By simply changing variables, equation (1) can be expressed in terms of length.

If l_c is the length of fish at some reference age t_c , we obtain from the von Bertalanffy growth model

$$t - t_c = -\frac{1}{K} \cdot \log \left(\frac{L_\infty - l}{L_\infty - l_c} \right)$$

Equation (2) (constant M) then becomes:

$$N_c = \int_{l_c}^L c'(l) \cdot \left(\frac{L_\infty - l_c}{L_\infty - l} \right)^{M/K} \cdot dl + \frac{1}{E} \cdot c + \left(\frac{L_\infty - l_c}{L_\infty - L} \right)^{M/K} \quad \dots 3)$$

where $c'(l)$ is defined as. $\frac{dC(l)}{dl} = \frac{dC(l(t))}{dt} \cdot \frac{dt}{dl} = c(t) \cdot \frac{dt}{dl}$

Stock numbers at length l and $c'(l)$ being given, one can derive the corresponding instantaneous rate of fishing mortality:

$$F(l) = K \cdot (L_\infty - l) \cdot \frac{c'(l)}{N(l)}$$

The proof of this can be developed starting with the relationship:

$$F(t) = \frac{c(t)}{N(t)} = \frac{c'(l(t))}{N(l(t))} \cdot \frac{dl}{dt}$$

Here again convergence formulas can be established:

$$\frac{\Delta(F(l))}{F(l)} \approx \frac{\Delta(E)}{E} \cdot e^{-F(l, L)}$$

where $F(l, L) = F(L) - F(l)$ is the cumulated fishing mortality from length l to L .

L being the length corresponding to age T .

IMPACT OF DISCRETISATION

Length intervals (l_i, l_{i+1}) will be considered, with $l_1 = l_c$ and $l_I = L$.

Generalities

In general, the density $c'(l)$ is unknown and can only be approached through the histogram $\{C_i\}$ fitted along length intervals (l_i, l_{i+1}) , $i = 1, \dots, I$. C_i designates the catches over the interval (l_i, l_{i+1}) and is written in bold letters in order to avoid confusions with density. The quantity

$$\frac{c_i}{l_{i+1} - l_i}$$

may be seen as an apparent mean density but the true distribution $c'(l)$ within the interval (l_i, l_{i+1}) remains unknown. In our opinion, this makes the difference between Virtual Population Analysis based on equations (2) or (3) and cohort analysis based on catch equations.

The catch equations can then be solved sequentially without Pope's approximation used by Jones by:

$$c_i = \frac{\overset{\circ}{F}_i}{\overset{\circ}{F}_i + M} \cdot N_{i+1} \left(e^{(\overset{\circ}{F}_i + M) \cdot (t_{i+1} - t_i)} - 1 \right)$$

where the fishing mortality $\overset{\circ}{F}_i$ is assumed to be constant over the time interval (t_i, t_{i+1}) corresponding to the length interval (l_i, l_{i+1}) . For simplification, the natural mortality M is assumed to be constant at all ages. The assumption of constancy of F within each interval can also compensate for our ignorance of "within intervals" fluctuations of $c'(l)$. The above equation can also be written in terms of length as:

$$c_i = \frac{\overset{\circ}{F}_i}{\overset{\circ}{F}_i + M} \cdot N_{i+1} \left(\left(\frac{L_\infty - l_i}{L_\infty - l_{i+1}} \right)^{\frac{\overset{\circ}{F}_i + M}{K}} - 1 \right) \quad \dots 4)$$

which can be solved for $\overset{\circ}{F}_i$ by any iterative method as is done for backwards computation in ordinary VPA.

Formulas for demonstrating the convergence are no longer appropriate for this equation in which relationships between $\Delta(E)$ and $\Delta(F)$ cannot be explicitly developed. However, simulations carried out using the formulas stated previously prove that these remain valid.

Generally speaking, the effects of discretisation are more important when the steps (l_i, l_{i+1}) are larger. In a trial to bracket the maximum error, consider the catches to occur at either end of the interval (l_i, l_{i+1}) .

The corresponding virtual populations estimated at some initial length l_c are given by:

$$a_i^1 = \left(\frac{L_\infty - l_c}{L_\infty - l_i} \right)^{M/K} \cdot c_i \quad \text{and} \quad a_i^2 = \left(\frac{L_\infty - l_c}{L_\infty - l_{i+1}} \right)^{M/K} \cdot c_i$$

Hereafter, adopt the average definition:

$$a_i = \left(\frac{L_\infty - l_c}{L_\infty - \bar{l}_i} \right)^{M/K} \cdot c_i \quad \text{with} \quad \bar{l}_i = \frac{l_i + l_{i+1}}{2}$$

Through a Taylor's expansion, one can easily obtain an estimate of $a_i^1 - a_i$ and $a_i^2 - a_i$

$$\frac{a_i^1 - a_i}{a_i} \approx \frac{a_i - a_i^2}{a_i} \approx \frac{M}{2 \cdot K} \left(\frac{l_{i+1} - l_i}{L_\infty - \bar{l}_i} \right)$$

with a constant step $\Delta l = l_{i+1} - l_i$, the error can be larger as l approaches L_∞ . This is related to the increase of time intervals Δt as growth in length slows down. When the lower bound of the last interval ($l_I = L$), is far enough from L_∞ , which is preferable for a number of reasons (Pereiro 1984), serious errors are unlikely.

Furthermore, a massive aggregation of the catch lengths right on the interval bounds is quite unlikely in practice. The actual density distribution $c'(l)$ should be regular since length distributions used for analysis often are averaged over the year and, thus, bound to be smoothed as a consequence of individual variability in growth.²

Actually, except in the vicinity of L_∞ , the size of the length intervals used correspond to time lapses which are quite short when compared to the time units (generally the year, quarters sometimes) currently used for age-structured VPA. Thus, it is natural that the effects demonstrated here are not as big as those evidenced by Sims (1984).

The Basic Approximation

Consider the scale of sizes $(l_i, l_{i+1}), i = 1 \dots I$, the ultimate size class having no upper bound. C_i designates the catches in number over the interval (l_i, l_{i+1}) . In default of the true density $c'(l)$ within the interval, approach it by a mean density $C_i/l_{i+1} - l_i$. Replaced in equation (3), this becomes:

$$N_C = (L_\infty - l_C) M/K \left(\sum_{i=1}^I \frac{C_i}{l_{i+1} - l_i} \int_{l_i}^{l_{i+1}} (L_\infty - l)^{-M/K} + \frac{C_i}{E} (L_\infty - L)^{-M/K} \right)$$

Integrating this formula makes no problem and gives explicit results. However, for consistency with traditional VPA, computations of stock sizes at the beginning of the classes, and of fishing mortalities within intervals, were carried out by solving iteratively the generalized form (equation 4) of the catch equation.

However, for sensitivity analysis and assessment of the effects of individual variations, the equation established in this paper proved easier to use with an approximation to the middle of the length classes $\bar{l}_i = (l_i + l_{i+1})/2$ as was done by Jones (1974) and Pope (1972).

In this context, equation (3) provides an approximation \tilde{N}_1 of N_1 :

$$(A) \quad \tilde{N}_1 \approx \sum_{i=1}^I v_i \cdot C_i$$

where

$$v_i = \left(\frac{L_\infty - \bar{l}_i}{L_\infty - l_i} \right)^{M/K}, \quad i = 1, \dots, I-1$$

and

$$v_I = \frac{1}{E} \left(\frac{L_\infty - \bar{l}_I}{L_\infty - l_I} \right)^{M/K}$$

The last coefficient is different from the others. From $C^+ = C_I$ one directly derives a number of survivors C_I/E at length \bar{l}_I , which corresponds to a "virtual" number of survivors at length \bar{l}_1 equal to:

$$\left(\frac{C_I}{E} \right) \left(\frac{L_\infty - \bar{l}_1}{L_\infty - l_I} \right)^{M/K}$$

From equation (A) an estimate of the fishing mortality is given by:

$$(B) \quad \tilde{F}_1 = K (L_\infty - \bar{l}_1) \frac{C_1}{\tilde{N}_1} \cdot \frac{1}{l_2 - l_1}$$

²When the length-frequency histograms are fitted on narrow steps and appear to be locally irregular, it is preferable to smooth them instead of grouping over larger steps. A narrow step reduces the side-effects of discretisation while smoothing is known to reduce sampling variances, although it introduces a bias which can be larger as the true histograms in the population demonstrate marked inflexions. For the reasons above, such local accidents should be uncommon and the risk of bias sufficiently low.

Table 2. Results of length VPA on the standard data sets (solutions of the generalized catch equation).

L	F	Numbers	L	F	Numbers	L	F	Numbers
14	.001	19,658.3	14	.013	19,656.7	10	.000	20,180.9
15	.025	19,516.2	16	.044	19,354.0	15	.045	19,489.6
16	.035	19,355.6	18	.055	19,000.5	20	.377	18,600.7
17	.054	19,186.5	20	.109	18,624.4	25	1.254	16,472.6
18	.063	19,002.1	22	.467	18,161.1	30	2.003	11,695.4
19	.047	18,809.2	24	.923	17,158.6	35	2.128	6,493.2
20	.091	18,626.0	26	1.288	15,528.2	40	2.115	3,167.6
21	.126	18,408.4	28	1.275	13,520.4	45	2.230	1,361.6
22	.453	18,162.6	30	1.622	11,707.7	50	2.147	460.4
23	.481	17,669.8	32	2.222	9,711.6	55	2.216	119.3
24	.710	17,160.0	34	2.317	7,473.9	60	.787	16.8
25	1.133	16,487.1	36	2.293	5,605.4	65+	.467	4.3
26	1.473	15,528.2	38	1.864	4,143.1			
27	1.106	14,378.8	40	1.738	3,174.2			
28	1.206	13,521.6	42	2.272	2,429.3			
29	1.343	12,635.6	44	2.273	1,684.0			
30	1.496	11,708.4	46	2.179	1,133.5			
31	1.746	10,744.8	48	2.352	749.2			
32	2.099	9,711.7	50	2.091	460.6			
33	2.344	8,591.2	52	2.264	284.2			
34	2.291	7,473.3	54	1.990	159.0			
35	2.342	6,494.7	56	2.165	88.6			
36	2.300	5,604.7	58	2.374	42.7			
37	2.287	4,827.5	60	1.049	16.7			
38	2.061	4,142.4	62	.697	9.6			
39	1.671	3,588.4	64+	.467	5.7			
40	1.578	3,174.2						
41	1.895	2,813.6						
42	2.256	2,428.9						
43	2.289	2,031.5						
44	2.495	1,683.6						
45	2.057	1,362.8						
46	2.113	1,133.4						
47	2.245	930.9						
48	2.408	749.1						
49	2.299	587.7						
50	2.011	460.5						
51	2.171	367.0						
52	2.141	284.0						
53	2.387	217.3						
54	2.020	158.8						
55	1.969	119.3						
56	2.172	88.4						
57	2.173	62.2						
58	2.549	42.5						
59	2.238	26.4						
60	1.342	16.6						
61	.795	12.0						
62	.914	9.5						
63	.514	7.0						
64	.555	5.6						
65+	.467	4.3						

Terminal E = 0.7

 L_∞ = 70.

M = 0.2

K = 0.5

Formulas similar to formulas (A) and (B) can be used to define \tilde{N}_i and \tilde{F}_i at any length interval (l_i, l_{i+1}) .

$$\tilde{N}_i = \sum_{j \geq i} b_j^i \cdot c_j \quad \dots 5)$$

$$\tilde{F}_i = K (L_\infty - \bar{l}_i) \cdot \frac{c_i}{N_i} \cdot \frac{1}{l_{i+1} - l_i} \quad \dots 6)$$

with $b_j^i = \left(\frac{L_\infty - \bar{l}_i}{L_\infty - \bar{l}_j} \right)^{M/K}$ while $j \leq i-1 ; i = 1, \dots, I-1$

and $b_I^i = \left(\frac{L_\infty - \bar{l}_i}{L_\infty - l_I} \right)^{M/K} \cdot \frac{1}{E} \quad i = 1, \dots, I-1$

The approximations (5) and (6) are essential for the subsequent sections. They give explicit expressions for \tilde{N}_i and \tilde{F}_i , from which it will be possible to get explicit expressions for their partial derivatives. Equation (5) is linear in c_j . Both equations are simple enough not to make the calculations too tedious. Usually, they give results close to those obtained by solving the catch equations. They are not suggested as an alternative to the solution of the exact catch equations, however. Rather, the sensitivity, the bias or the variance of the fishing mortality, noted \hat{F}_i , obtained through the catch equations will be approximated by the corresponding sensitivity bias or variance of \tilde{F}_i . This in fact is the basic idea for the following sections.

It must be pointed out that the \tilde{N}_i cannot be compared directly to the stock numbers calculated at the beginning of any length interval. They refer to the middle of such intervals. However they are used mainly as intermediate steps for the calculations of the \tilde{F}_i . If one is directly interested in stock numbers at the beginning of length intervals, a better approximation at length l_i would be given by

$$\sum_{I > j \geq i} \left(\frac{L_\infty - l_i}{L_\infty - l_j} \right)^{M/K} c_j + \frac{1}{E} \left(\frac{L_\infty - l_i}{L_\infty - l_I} \right)^{M/K} \cdot c_I$$

Sequential Calculations

Applying directly equation (5) would make it necessary for each length \bar{l}_i to build a loop for $j \geq i$. Making use of the V_i previously defined this can be avoided. It can be verified that: $b_j^i = \frac{\sqrt{i}}{\sqrt{j}}$. Thus, with the notation $U_i = \sum_{j \geq i} V_j \cdot c_j$,

$$\tilde{N}_i = \frac{U_i}{V_i} , i = 1, \dots, J-1 \quad \dots 7)$$

This way, a backward calculation will provide U_i for each interval, using the recurrence formula $U_i = U_{i+1} + V_i C_i$ and $U_I = V_I C_I$. From U_i , \tilde{N}_i can be calculated through equation (7), and \tilde{F}_i through equation (6). If one is interested in stock numbers at the beginning of length intervals, one can use expressions U_i/W_i , where:

$$W_i = \left(\frac{L_\infty - l_1}{L_\infty - l_i} \right)^{M/K}$$

EFFECTS OF UNCERTAINTIES IN C_i

The General Case

In order to distinguish the true values from the sample estimates, we shall use the symbol $\hat{\cdot}$. Thus, C_i is estimated by \hat{C}_i ; N_i is estimated by \tilde{N}_i and F_i by \hat{F}_i . In the same way \bar{F}_i would be estimated by $\hat{\bar{F}}_i$. As previously mentioned, no direct calculation will be conducted on the expectation and the variance of \hat{F}_i . Any bias or variance will be approximated from those calculated on \hat{F}_i .

From the previous sections, we take:

$$\hat{N} = \sum_{j \geq i} \frac{v_j}{v_i} \cdot \hat{C}_j = \sum_{j \geq i} b^i_j \cdot \hat{C}_j$$

Notice by the way that if the \hat{C}_j are unbiased the \hat{N}_j also are. So, only the variances will be considered.

We assume that the matrix of variances-covariances $\text{Cov}(\hat{C}_i, \hat{C}_j) = \theta_{i,j}$ is also known.
With these notations:

$$\text{Var}(\hat{N}_i) = \sum_{k,m \geq i} b^i_k \cdot b^i_m \cdot \theta_{k,m} = \frac{1}{v_i^2} \sum_{k,m > i} v_k \cdot v_m \cdot \theta_{k,m} \quad \dots 8)$$

During a sequential computation, the quantities are collected:

$$w_i = \sum_{k,m \geq i} v_k \cdot v_m \cdot \theta_{k,m}$$

using the recurrence:

$$w_i = w_{i+1} + v_i^2 \theta_{i,i} + 2 \cdot v_i \sum_{k \geq i} v_k \cdot \theta_{i,k}$$

For the calculation of the moments of \hat{F}_i , we need one more approximation. \hat{F}_i is given by:

$$\hat{F}_i = \frac{K \cdot (L_\infty - \bar{l}_i)}{\Delta l_i} \cdot \frac{C_i}{\hat{N}_i} \text{ with } \Delta l_i = l_{i+1} - l_i$$

and

$$E(\hat{F}_i) = K \cdot \frac{(L_\infty - \bar{l}_i)}{\Delta l_i} \cdot E\left(\frac{\hat{C}_i}{\hat{N}_i}\right)$$

by Taylor's expansion,³ we get:

$$\frac{E\left(\frac{\hat{C}_i}{\hat{N}_i}\right) - \frac{C_i}{\tilde{N}_i}}{\frac{C_i}{\tilde{N}_i}} \approx \frac{\text{Var}(\hat{N}_i)}{\tilde{N}_i^2} - \frac{\text{Cov}(\hat{C}_i, \hat{N}_i)}{C_i \cdot \tilde{N}_i}$$

and by a similar process:

$$\text{Var}(\hat{F}_i) = \left(\frac{K \cdot (L_\infty - \bar{l}_i)}{\Delta l_i} \right)^2 \cdot \text{Var}\left(\frac{\hat{C}_i}{\hat{N}_i}\right)$$

³ It must be noticed that if the variability of \hat{N}_i is too high, this expansion will lead to poor approximations.

Making use of the usual approximation for the variance of a ratio yields:

$$\text{Var} \left(\frac{\hat{C}_i}{\tilde{N}} \right) \approx \frac{\text{Var}(\hat{C}_i)}{(\tilde{N}_i)^2} + \frac{C_i^2}{(\tilde{N}_i)^4} \cdot \text{Var}(\tilde{N}_i) - 2 \frac{C_i}{(\tilde{N}_i)^3} \cdot \text{Cov}(\hat{C}_i, \tilde{N}_i)$$

and consequently:

$$\frac{\text{Var}(\hat{F}_i)}{(\tilde{F}_i)^2} = \frac{\text{Var}(\hat{C}_i)}{C_i^2} + \frac{\text{Var}(\tilde{N}_i)}{\tilde{N}_i^2} - 2 \frac{\text{Cov}(\hat{C}_i, \tilde{N}_i)}{C_i \cdot \tilde{N}_i} \quad \dots 9)$$

$\text{Var}(\tilde{N}_i)$ is given by equation (8) and $\text{Var}(\hat{C}_i) = \theta_{i,j}$, so we still have to solve $\text{Cov}(\hat{C}_i, \tilde{N}_i)$:

$$\text{Cov}(\hat{C}_i, \tilde{N}_i) = \text{Cov}(\hat{C}_i, \sum_{j \geq i} b_j^i \cdot \hat{C}_j) = \sum_{j \geq i} b_j^i \cdot \theta_{i,j}$$

thus

$$\text{Cov}(\hat{C}_i, \tilde{N}_i) = \frac{1}{V_i} \sum_{j \geq i} v_j \cdot \theta_{i,j}$$

As compared to the expression for $\frac{\text{Var}(\tilde{N}_i)}{(N_i)^2}$, equation (9) for $\frac{\text{Var}(\tilde{F}_i)}{(\tilde{F}_i)^2}$

introduces an additional term:

$$\frac{\text{Var}(\hat{C}_i)}{(C_i)^2} - 2 \frac{\text{Cov}(\hat{C}_i, \tilde{N}_i)}{C_i \cdot \tilde{N}_i} \frac{\text{Var}(\hat{C}_i)}{(C_i)^2}$$

which, in most cases, should play a major part due to the dominance of $\frac{(C_i)^2}{(C_i)^2}$. As a matter of fact, the variability of \tilde{F}_i will generally be larger than is the case of \tilde{N}_i , the latter taking advantage of a buffering effect upon the uncertainties in \hat{C}_j when $j \geq i$. This is not true for the \hat{C}_i . As will be seen later, the narrower the interval width, the larger $\frac{\text{Var}(\hat{C}_i)}{(C_i)^2}$

Each individual \hat{F}_i can be affected by a large amount of uncertainty. This does not imply that the overall pattern and level of fishing mortality is similarly affected. If for example, the \hat{F}_i array is smoothed, the variance of the smoothed estimate \hat{F} should be lower, and an increased stability may be observed:

$$\text{Let } \hat{F} = \sum_i g_i \cdot \hat{F}_i \text{ and } \sum_i g_i = 1$$

$$\text{Var}(\hat{F}) = \sum_i g_i^2 \cdot \text{Var}(F_i) + \sum_{i \neq j} g_i \cdot g_j \cdot \text{Cov}(\hat{F}_i, \hat{F}_j)$$

At this point, one needs to know the covariances in addition to the variances developed previously, and again the same type of approximation is used to derive:

$$\frac{1}{\tilde{F}_i \cdot \tilde{F}_j} \text{Cov}(\hat{F}_i, \hat{F}_j) \approx \frac{\text{Cov}(\hat{C}_i, \hat{C}_j)}{C_i \cdot C_j} + \frac{\text{Cov}(\tilde{N}_i, \tilde{N}_j)}{\tilde{N}_i \cdot \tilde{N}_j} - \frac{\text{Cov}(\hat{C}_i, \tilde{N}_j)}{C_i \cdot \tilde{N}_j} - \frac{\text{Cov}(\hat{C}_j, \tilde{N}_i)}{C_j \cdot \tilde{N}_i}$$

$$\text{and } \frac{\text{Cov}(\hat{C}_i, \hat{C}_j)}{C_i \cdot C_j} = \frac{\theta_{i,j}}{C_i \cdot C_j}$$

The other terms are simple to obtain. When $i < j$:

$$\text{Cov}(\hat{N}_i, \hat{N}_j) = \frac{1}{V_i V_j} \left[\sum_{k=i}^I v_k v_m \theta_{k,m} \right]$$

$$\text{Cov}(\hat{C}_i, \hat{N}_j) = \frac{1}{V_j} \left[\sum_{k=j}^I v_k \theta_{i,k} \right]$$

$$\text{Cov}(\hat{C}_j, \hat{N}_i) = \frac{1}{V_i} \left[\sum_{i=i}^I v_k \theta_{j,k} \right]$$

where estimates of the variances $\theta_{i,j}$ are used whenever the true values are unknown.

Simple Random Sampling

Obviously, circumstances in which the catches \hat{C}_i are sampled through a simple random design will rarely occur in practice. However, to give a simple illustration, it is supposed that this case occurred here. It is also assumed that the total catch in number C is known without error.

For a sample of size n :

$$\text{Var}\left(\frac{\hat{C}_i}{C}\right) = \frac{1}{n} \cdot \frac{C_i}{C} \left(1 - \frac{C_i}{C}\right) \text{ or } \theta_{i,i} = \frac{1}{n} \cdot C_i (C - C_i)$$

$$\text{Cov}\left(\frac{\hat{C}_i}{C}, \frac{\hat{C}_j}{C}\right) = -\frac{1}{n} \cdot \frac{C_i}{C} \cdot \frac{C_j}{C} \text{ or } \theta_{i,j} = -\frac{1}{n} \cdot \frac{C_i \cdot C_j}{C^2} \text{ for } i \neq j$$

In this context, formulae of the previous section are simplified into:

$$\frac{\text{Var}(\hat{N}_i)}{\tilde{N}_i^2} = \frac{1}{n} \left(C \cdot \frac{\sum_{j=i}^I v_j^2 \cdot C_j}{\sum_{j=i}^I (v_j \cdot C_j)^2} - 1 \right)$$

and, since

$$\frac{\text{Cov}(\hat{C}_i, \hat{N}_i)}{C_i \cdot \tilde{N}_i} = \frac{1}{n} \cdot \left[\frac{C}{\tilde{N}_i} - 1 \right]$$

$$\frac{\text{Var}(\hat{F}_i)}{\tilde{F}_i^2} = \frac{C}{n} \cdot \left(\frac{1}{C_i} - \frac{2}{\tilde{N}_i} \right) + \frac{1}{n} + \frac{\text{Var}(\hat{N}_i)}{\tilde{N}_i^2}$$

This also yields an estimation of the bias in \hat{F}_i when replaced in the equation:

$$\frac{E(\hat{F}_i) - F_i}{\tilde{F}_i} = \frac{\text{Var}(\hat{N}_i)}{\tilde{N}_i^2} - \frac{\text{Cov}(\hat{C}_i, \hat{N}_i)}{C_i \cdot N_i}$$

(estimates of \tilde{N}_i were proved to be unbiased)

Example

Results of calculations based on the variance formulae given above are presented in Table 3. Most dramatic is the decrease of the (squared) coefficient of variation of N_i as the computation proceeds backwards from the largest to the smallest length groups. The low variability in stock numbers of small individuals is due to the concentration of the bulk of the catches over a limited range of classes, mainly from 25 to 40, and also depends on the assumption that the total catch is perfectly known.

Table 3 suggests that large biases in the fishing mortalities are not likely.

The larger variability of the fishing mortality estimates is obvious as well. It can be observed that the relative variance decreases first with decreasing lengths, then increases over the smallest classes, in spite of the decreasing contribution of $\text{Var}(\hat{N}_i)/\tilde{N}_i^2$. This pattern demonstrates the importance of the relative uncertainty in the catches when these are low. Thus, even in the length groups with the lowest variances, large samples are required. For the class 32, for example, with a sample size of 1,000, the coefficient of variation is $\sqrt{0.015} = 12\%$.

However one is generally more concerned with the overall pattern of the F array than with the precision of individual values, so again the problem of calculating variances for \bar{F} is addressed. Simple random sampling allows for tractable expressions.

One may verify, for $j > i$:

$$\begin{aligned} \text{Cov}(\hat{F}_i, \hat{F}_j) &= \frac{1}{n} \tilde{F}_i \cdot \tilde{F}_j \left(\frac{\tilde{N}_j \cdot v_j}{\tilde{N}_i \cdot v_i} \right) \cdot \left(c \frac{\sum_{k \geq j} v_k^2 \cdot c_k}{(N_j \cdot v_j)^2} - \frac{c}{\tilde{N}_j} \right) \\ &= \frac{1}{n} \cdot \tilde{F}_i \cdot \tilde{F}_j \left(\frac{\tilde{N}_j \cdot v_j}{\tilde{N}_i \cdot v_i} \right) \cdot \left(\left(\frac{(n \cdot \text{Var}(\hat{N}_j))}{\tilde{N}_j^2} + 1 \right) - \frac{c}{\tilde{N}_j} \right) \end{aligned}$$

This gives, for the average $\frac{1}{2}(\hat{F}_{32} + \hat{F}_{33})$ for example, a coefficient of variation reduced to 8.7% when the sample size is 1,000.

EFFECTS OF UNCERTAINTIES IN K, L_∞ OR M

Generalities

Sensitivity of stock numbers at length

Starting again with equation (3):

$$N_c = \int_{l_c}^L c \cdot (1) \cdot \left(\frac{L_\infty - l_c}{L_\infty - L} \right)^{M/K} \cdot dl + \frac{1}{E} c^+ \cdot \left(\frac{L_\infty - l_c}{L_\infty - L} \right)^{M/K}$$

Table 3. Variance components when catches at length are drawn by simple random sampling.

L	U	V	\tilde{N}	\tilde{F}	$n*cv^2(\hat{N})$	$n*cv^2(\hat{F})$
14	19,590.0	1.00	19,590.0	0.00	0.01	16,339.3
15	19,589.0	1.01	19,447.1	0.02	0.01	907.1
16	19,570.9	1.01	19,285.7	0.03	0.01	652.9
17	19,545.5	1.02	19,115.9	0.05	0.01	418.3
18	19,505.7	1.03	18,930.7	0.06	0.01	354.5
19	19,458.3	1.04	18,737.1	0.05	0.02	466.1
20	19,421.9	1.05	18,553.1	0.09	0.02	239.5
21	19,350.7	1.05	18,334.8	1.13	0.02	171.2
22	19,250.5	1.06	18,088.5	0.45	0.03	47.0
23	18,886.5	1.07	17,596.1	0.48	0.05	44.6
24	18,500.1	1.08	17,086.8	0.70	0.08	30.3
25	17,931.7	1.09	16,415.3	1.10	0.11	19.2
26	17,041.4	1.10	15,459.1	1.42	0.18	15.2
27	15,925.8	1.11	14,313.3	1.08	0.27	21.5
28	15,118.0	1.12	13,458.5	1.17	0.35	20.4
29	14,264.3	1.13	12,575.2	1.30	0.44	19.1
30	13,348.9	1.15	11,651.1	1.44	0.54	18.0
31	12,375.0	1.16	10,690.9	1.67	0.68	16.2
32	11,302.0	1.17	9,661.6	1.98	0.85	14.4
33	10,105.3	1.18	8,545.7	2.20	1.09	14.1
34	8,887.3	1.20	7,432.7	2.15	1.41	16.2
35	7,811.2	1.21	6,458.5	2.19	1.77	17.6
36	6,819.5	1.22	5,572.5	2.15	2.22	20.2
37	5,944.5	1.24	4,799.0	2.13	2.74	22.9
38	5,164.1	1.25	4,117.2	1.93	3.36	28.8
39	4,530.7	1.27	3,565.9	1.58	4.04	40.0
40	4,060.6	1.29	3,153.5	1.50	4.68	46.4
41	3,648.5	1.31	2,794.7	1.77	5.38	41.6
42	3,194.2	1.32	2,412.0	2.08	6.38	38.6
43	2,710.8	1.34	2,016.9	2.10	7.82	43.7
44	2,280.7	1.36	1,671.0	2.27	9.65	46.1
45	1,875.4	1.39	1,352.2	1.89	12.18	67.2
46	1,585.5	1.41	1,124.3	1.93	14.83	75.1
47	1,324.6	1.43	923.1	2.03	18.25	81.7
48	1,085.0	1.46	742.5	2.16	22.92	89.6
49	867.2	1.49	582.3	2.06	29.53	114.1
50	693.0	1.52	456.0	1.82	37.98	159.6
51	563.8	1.55	363.3	1.93	47.80	173.8
52	445.9	1.59	281.0	1.90	61.99	214.6
53	349.1	1.62	214.9	2.07	81.17	232.7
54	261.4	1.67	156.9	1.78	111.49	358.1
55	201.4	1.71	117.7	1.72	148.21	455.2
56	153.5	1.76	87.2	1.86	199.04	506.1
57	111.2	1.81	61.3	1.84	281.64	657.1
58	78.6	1.88	41.9	2.06	408.60	718.2
59	50.4	1.95	25.9	1.82	654.09	1,208.9
60	32.9	2.03	16.2	1.17	1,022.71	3,096.1
61	24.8	2.12	11.7	0.73	1,364.88	6,744.0
62	20.6	2.23	9.2	0.81	1,638.87	6,629.8
63	16.1	2.36	6.8	0.48	2,046.46	13,602.0
64	13.7	2.52	5.5	0.50	2,328.32	12,675.5
65	11.2	2.62	4.3	0.47	2,667.87	

designate

$$\phi(l, l_c, M, K, L_\infty) = \left(\frac{L_\infty - l_c}{L_\infty - l} \right)^{M/K}$$

This corresponds to the V_i defined earlier.⁴

N_c is a function of M, K and L_∞ which can be developed using such derivatives as:

$$\frac{\partial N_c}{\partial K} = \int_{l_c}^L \frac{\partial \phi}{\partial K} \cdot c'(l) \cdot dl + \frac{1}{E} \cdot c + \frac{\partial \phi(l, l_c, M, K, L_\infty)}{\partial K}$$

(Possible problems which may occur when deriving under the integral sign were not examined.)

In the same manner, get the first derivatives in M and L_∞ , and the second derivatives based on the corresponding first derivatives. These derivations raise no particular difficulty. Logarithmic derivatives, and their equivalent for second derivatives, were actually used since they provide simpler expressions, especially for computation.

It can be shown in this context that:⁵

$$(A) \cdot \frac{1}{\phi} \cdot \frac{\partial \phi}{\partial K} = \phi_K = - \frac{M}{K^2} \cdot \log \left(\frac{L_\infty - l_c}{L_\infty - l} \right)$$

$$(B) \cdot \frac{1}{\phi} \cdot \frac{\partial \phi}{\partial L_\infty} = \phi_L = \frac{M}{K} \left(\frac{1}{L_\infty - l_c} - \frac{1}{L_\infty - l} \right)$$

$$(C) \cdot \frac{1}{\phi} \cdot \frac{\partial \phi}{\partial M} = \phi_M = \frac{1}{K} \cdot \log \left(\frac{L_\infty - l_c}{L_\infty - l} \right)$$

$$(D) \cdot \frac{1}{\phi} \cdot \frac{\partial^2 \phi}{\partial K^2} = \phi_{KK} = \frac{2 \cdot M}{K^3} \cdot \log \left(\frac{L_\infty - l_c}{L_\infty - l} \right) + (\phi_K)^2$$

$$(E) \cdot \frac{1}{\phi} \cdot \frac{\partial^2 \phi}{\partial K \cdot \partial L_\infty} = \phi_{KL} = \frac{-M}{K^2} \left(\frac{1}{L_\infty - l_c} - \frac{1}{L_\infty - l} \right) + \phi_K \cdot \phi_L$$

$$(F) \cdot \frac{1}{\phi} \cdot \frac{\partial^2 \phi}{\partial L_\infty^2} = \phi_{LL} = \frac{M}{K} \left(\frac{1}{(L_\infty - l)^2} - \frac{1}{(L_\infty - l_c)^2} \right) + (\phi_L)^2$$

$$(G) \cdot \frac{1}{\phi} \cdot \frac{\partial^2 \phi}{\partial M^2} = \phi_{MM} = (\phi_M)^2$$

⁴The following sections use functions such as $\Phi(l, \bar{l}_1, M, K, L_\infty)$, with \bar{l}_1 instead of l_c . Equations (A) to (G) of this section can be directly adapted, l_c being here replaced by \bar{l}_1 .

⁵Such computations can take benefit of the easily demonstrated relations:

$$\Phi_{KK} = \frac{\partial \Phi_K}{\partial K} + (\Phi_K)^2 ; \quad \Phi_{LL} = \frac{\partial \Phi_L}{\partial L_\infty} + (\Phi_L)^2$$

$$\Phi_{KL} = \frac{\partial \Phi_K}{\partial L_\infty} + \Phi_K \cdot \Phi_L$$

$$\frac{\partial^2 \phi}{\partial K \partial M} \quad \text{and} \quad \frac{\partial^2 \phi}{\partial M \cdot \partial L_\infty}$$

are equally possible to obtain but are of little interest.

The ability to calculate the derivatives of N_c with respect to the parameters provides an opportunity to make use of Taylor's expansions. In the first place, these allow to investigate directly the sensitivity of N_c with respect to K , L_∞ and M by means of such quantities as:

$$\Delta N_c \approx \frac{\partial N_c}{\partial K} \cdot \Delta K$$

or, preferably:

$$\Delta N_c \approx \frac{\partial N_c}{\partial K} \cdot \Delta K + \frac{1}{2} \cdot \frac{\partial^2 N_c}{\partial K^2} \cdot (\Delta K)^2$$

These are still easier to use when expressed in relative terms, such as:

$$\frac{\Delta N_c}{N_c} \approx \left(\frac{1}{N_c} \cdot \frac{\partial N_c}{\partial K} \cdot K \right) \frac{\Delta K}{K} + \left(\frac{1}{2 \cdot N_c} \cdot \frac{\partial^2 N_c}{\partial K^2} \cdot K^2 \right) \left(\frac{\Delta K}{K} \right)^2$$

and the equivalents in M and L_∞

A coefficient such as $A_K = \left(\frac{1}{N_c} \cdot \frac{\partial N_c}{\partial K} \cdot K \right)$ will be called a sensitivity coefficient of first order.

$A_{KK} = \left(\frac{1}{2 \cdot N_c} \cdot \frac{\partial^2 N_c}{\partial K^2} \right) K^2$ is a coefficient of second order.

A_L and A_{LL} are defined in a similar way for L_∞ .

A_{KL} could be called a coefficient of "cosensitivity". It is defined by:

$$A_{KL} = \left(\frac{1}{N_c} \cdot \frac{\partial^2 N_c}{\partial K \partial L_\infty} \right) K \cdot L_\infty$$

As for the sensitivity with respect to M , one may wish to take into account the additional effect of M upon E for a given terminal F^+ . Since $E = F^+ / (F^+ + M)$, this consists of adding to the expressions given above a term:

$$\frac{\partial N_c}{\partial E} \cdot \frac{\partial E}{\partial M}$$

Taylor's expansions can also be used to derive estimates of the expectation and variance of \hat{N}_c based on the variances of the parameters K , L_∞ and M . Consider for example that field estimates \hat{K} and \hat{L}_∞ are available from which \hat{N}_c is estimated. Then

$$\begin{aligned} E(\hat{N}_c) &\approx N_c + \frac{\partial N_c}{\partial K} (E(\hat{K}) - K) + \frac{\partial N_c}{\partial L_\infty} (E(\hat{L}_\infty) - L_\infty) + \frac{1}{2} \cdot \frac{\partial^2 N_c}{\partial K^2} \text{Var}(\hat{K}) \\ &+ \frac{\partial^2 N_c}{\partial K \cdot \partial L_\infty} \text{Cov}(\hat{K}, \hat{L}_\infty) + \frac{1}{2} \cdot \frac{\partial^2 N_c}{\partial L_\infty^2} \text{Var}(\hat{L}_\infty) \end{aligned}$$

Here again, the deviations are expressed in relative terms:

$$\begin{aligned} \frac{E(\hat{N}_c) - N_c}{N_c} &\approx \left(\frac{1}{N_c} \cdot \frac{\partial N_c}{\partial K} \cdot K \right) \left(\frac{E(\hat{K}) - K}{K} \right) + \left(\frac{1}{N_c} \cdot \frac{\partial N_c}{\partial L_\infty} \cdot L_\infty \right) \left(\frac{E(\hat{L}_\infty) - L_\infty}{L_\infty} \right) \\ &+ \left(\frac{1}{2 \cdot N_c} \cdot \frac{\partial^2 N_c}{\partial K^2} \cdot K^2 \right) \frac{\text{Var}(\hat{K})}{K^2} + \left(\frac{1}{N_c} \cdot \frac{\partial^2 N_c}{\partial K \cdot \partial L_\infty} \cdot K \cdot L_\infty \right) \frac{\text{Cov}(\hat{K}, \hat{L}_\infty)}{K \cdot L_\infty} \\ &+ \left(\frac{1}{2 \cdot N_c} \cdot \frac{\partial^2 N_c}{\partial L_\infty^2} \cdot L_\infty^2 \right) \frac{\text{Var}(\hat{L}_\infty)}{L_\infty^2} \end{aligned}$$

When unbiased estimators are available for K and L_∞ , the only bias will be due to the second order coefficients. This explains why a second order expansion has been performed. On the other hand if one is interested in the variance of \hat{N}_c , a first order expansion will provide useful terms, while those associated with higher orders will generally be much smaller and difficult to calculate. For the basic approximations only the first order coefficients will be used; for instance:

$$\begin{aligned} \frac{\text{var}(\hat{N}_c)}{N_c^2} &\approx \left(\frac{1}{N_c} \cdot \frac{\partial N_c}{\partial K} K \right)^2 \frac{\text{Var}(K)}{K^2} \\ &+ \left(\frac{1}{N_c} \cdot \frac{\partial N_c}{\partial K} K \right) \left(\frac{1}{N_c} \cdot \frac{\partial N_c}{\partial L_\infty} L_\infty \right) 2 \frac{\text{Cov}(K, \hat{L}_\infty)}{K L_\infty} \\ &+ \left(\frac{1}{N_c} \cdot \frac{\partial N_c}{\partial L_\infty} L_\infty \right)^2 \frac{\text{Var}(\hat{L}_\infty)}{L_\infty} \end{aligned}$$

The formulas given in this section for the initial length l_c are straightforward to generalize for any intermediate length.

Sensitivity of mortality rates

$N(l)$ being the stock number, and $c'(l)$ the catch density at length l :

$$F(l) = \frac{1}{N(l)} \frac{dC(t)}{dt} = \frac{1}{N(l)} \cdot K \cdot (L_\infty - l) \cdot c'(l)$$

which is a function of K , L_∞ and M through $N(l)$ in the one hand, and $K(L_\infty - l)$ in the other. First and second derivatives can thus be developed easily from those of $N(l)$, giving

- $\frac{1}{F} \frac{\partial F}{\partial K} = \frac{1}{K} - \frac{1}{N} \cdot \frac{\partial N}{\partial K} = \eta_K$
- $\frac{1}{F} \frac{\partial F}{\partial L_\infty} = \frac{1}{L_\infty - l} - \frac{1}{N} \cdot \frac{\partial N}{\partial L_\infty} = \eta_L$
- $\frac{1}{F} \frac{\partial^2 F}{\partial K^2} = \frac{-1}{K^2} - \frac{1}{N} \cdot \frac{\partial^2 N}{\partial K^2} + \left(\frac{1}{N} \cdot \frac{\partial N}{\partial K} \right)^2 + \eta_K^2$
- $\frac{1}{F} \frac{\partial^2 F}{\partial K \cdot \partial L_\infty} = - \frac{1}{N} \cdot \frac{\partial^2 N}{\partial K \cdot \partial L_\infty} + \left(\frac{1}{N} \cdot \frac{\partial N}{\partial K} \right) \left(\frac{1}{N} \cdot \frac{\partial N}{\partial L_\infty} \right) + \eta_K \cdot \eta_L$
- $\frac{1}{F} \frac{\partial^2 F}{\partial L^2} = - \frac{1}{(L_\infty - l)^2} - \frac{1}{N} \cdot \frac{\partial^2 N}{\partial L_\infty^2} + \left(\frac{1}{N} \cdot \frac{\partial N}{\partial L_\infty} \right)^2 + \eta_L^2$

and:

- $\frac{1}{F} \cdot \frac{\partial F}{\partial M} = - \frac{1}{N} \cdot \frac{\partial N}{\partial M}$
- $\frac{1}{F} \cdot \frac{\partial^2 F}{\partial M^2} = - \frac{1}{N} \cdot \frac{\partial^2 N}{\partial M^2} + 2 \left(\frac{1}{N} \cdot \frac{\partial N}{\partial M} \right)^2$

Discretisation

As mentioned previously, when considering discrete intervals, use the following relationships:

$$\tilde{N}_1 = \sum_{j=1}^I V_j \cdot C_j \quad F_1 = K \cdot (L_\infty - \bar{l}_1) \cdot \frac{C_1}{\tilde{N}_1} \cdot \frac{1}{(l_2 - l_1)}$$

$$\tilde{N}_i = \sum_{j \geq i}^I \frac{V_j}{V_i} \cdot C_j = \frac{U_i}{V_i} \quad F_i = K \cdot (L_\infty - \bar{l}_i) \cdot \frac{C_i}{\tilde{N}_i} \cdot \frac{1}{(l_{i+1} - l_i)}$$

$$i = 1, \dots, I-1$$

where:

$$V_i = \phi(\bar{l}_1, \bar{l}_i, K, L_\infty, M)$$

$$\text{Then: } \frac{\partial \tilde{N}_1}{\partial K} = \sum_{j=1}^I \frac{\partial V_j}{\partial K} \cdot C_j$$

The same holds for the derivatives in the other parameters and the second derivatives, which all follow from formulas given above, and the same process could have been used for the derivatives of the N_i and F_i , with help of the relationship:

$$\frac{V_j}{V_i} = \phi(l_i, l_j, K, L_\infty, M)$$

The formulas given above could be directly applied. It would require for each length interval a whole set of calculations. In order to simplify the computer program, another approach was preferred. It is based on the relationship: $\tilde{N}_i = (U_i / V_i)$

As a matter of fact, backwards sequential computation proceeds with calculations of the first and second derivatives of the V_j for each interval, so that the derivatives of the U_i are obtained by simple sums. For example:

$$(A) \frac{\partial U_i}{\partial K} = \frac{\partial U_{i+1}}{\partial K} + C_i \frac{\partial V_i}{\partial K}$$

It corresponds to the recurrence formula:

$$U_i = U_{i+1} + C_i V_i, \text{ previously given.}$$

Here again take the logarithmic derivatives which simplify the expressions:

$$(B) \frac{1}{\tilde{N}_i} \cdot \frac{\partial \tilde{N}_i}{\partial K} = \frac{1}{U_i} \cdot \frac{\partial U_i}{\partial K} - \frac{1}{V_i} \cdot \frac{\partial V_i}{\partial K}$$

In the ultimate class $\frac{\partial U_i}{\partial K}$ is calculated by adapting equation (A) and the definition:

$$U_I = \frac{1}{E} \left(\frac{L_\infty - \bar{l}_1}{L_\infty - L} \right)^{M/K} = \frac{1}{E} \phi(L, \bar{l}_1, M, K, L_\infty)$$

Then, for each length, $\frac{\partial V_i}{\partial K}$ is calculated using the same ad hoc formula. Equation (A) from this section will give $\frac{\partial U_I}{\partial K}$. Equation (B) will give

$$\frac{1}{\tilde{N}_i} \cdot \frac{\partial \tilde{N}_i}{\partial K}$$

It can be noticed that this scheme will make it possible to obtain all the $(1/\tilde{N}_i)(\partial \tilde{N}_i / \partial K)$ within a single backward sequential calculation.

The same calculations can be applied to the other first derivatives.

The calculations of the second derivatives will also proceed from a similar scheme.

Equation (C) $\frac{\partial^2 U_i}{\partial K^2} = \frac{\partial^2 U_{i+1}}{\partial K^2} + C_i \frac{\partial^2 V_i}{\partial K^2}$ gives a recurrence formula.

$\frac{\partial^2 U_i}{\partial K^2}$ is obtained by an adaptation of formula (D) above. At each step $\frac{\partial^2 V_i}{\partial K^2}$ is obtained through this same process. Then $\frac{1}{\tilde{N}_i} \cdot \frac{\partial^2 \tilde{N}_i}{\partial K^2}$ is given by:

$$\frac{1}{\tilde{N}_i} \cdot \frac{\partial^2 \tilde{N}_i}{\partial K^2} = \frac{1}{U_i} \cdot \frac{\partial^2 U_i}{\partial K^2} - \frac{1}{V_i} \cdot \frac{\partial^2 V_i}{\partial K^2} + 2 \left(\frac{1}{V_i} \cdot \frac{\partial V_i}{\partial K} \right)^2 - 2 \left(\frac{1}{U_i} \cdot \frac{\partial U_i}{\partial K} \right) \left(\frac{1}{V_i} \cdot \frac{\partial V_i}{\partial K} \right)$$

By the same means, we get $\frac{1}{\tilde{N}_i} \cdot \frac{\partial^2 \tilde{N}_i}{\partial L_\infty^2}$ and $\frac{1}{\tilde{N}_i} \cdot \frac{\partial^2 \tilde{N}_i}{\partial M^2}$, and

$$\begin{aligned} \frac{1}{\tilde{N}_i} \cdot \frac{\partial^2 \tilde{N}_i}{\partial K \cdot \partial L_\infty} &= \frac{1}{U_i} \cdot \frac{\partial^2 U_i}{\partial K \cdot \partial L_\infty} - \frac{1}{V_i} \cdot \frac{\partial^2 V_i}{\partial K \cdot \partial L_\infty} + 2 \left(\frac{1}{V_i} \cdot \frac{\partial V_i}{\partial K} \right) \left(\frac{1}{V_i} \cdot \frac{\partial V_i}{\partial L_\infty} \right) \\ &\quad - \left(\frac{1}{U_i} \cdot \frac{\partial U_i}{\partial K} \right) \left(\frac{1}{V_i} \cdot \frac{\partial V_i}{\partial L_\infty} \right) - \left(\frac{1}{U_i} \cdot \frac{\partial U_i}{\partial L_\infty} \right) \left(\frac{1}{V_i} \cdot \frac{\partial V_i}{\partial K} \right) \end{aligned}$$

The derivatives related to the fishing mortalities F_i are derived from those calculated on the N_i . This can be done through the formulas given earlier.

Example

The results are illustrated in Tables 4, 5 and 6, based on the standard data set given in Table 1. Table 4 makes it possible, for length 40 chosen as an example, to check the quality of the approximations derived from the second order Taylor's expansion. These expansions offer direct approximations of $\Delta \tilde{F}/\tilde{F}$ and $\Delta \tilde{N}/\tilde{N}$, used in fact as approximations for $\Delta \hat{F}/\hat{F}$ and $\Delta N/N$. For instance, a change $\Delta K/K = 10\%$ from the reference value $K = .5$ will lead to a value $\hat{F} = 1.751$ instead of the reference value 1.75, while the approximations suggest $\hat{F} = 1.7502$. An overall view of Table 4 shows the quality of the approximations.

Table 5 gives the sensitivity coefficients of the estimated stock numbers, \tilde{N} , defined previously. Table 6 is the equivalent table for fishing mortalities.

Considering again length 40 in Table 6, a change $\Delta K/K = 10\%$ will create a change $\Delta \tilde{F}/\tilde{F}$ approximated at first order by $1.089 \times .1$, at second order by $1.089 \times .10 - 0.002 \times (.10)^2$.

Consider now a situation where the pair (K, L_∞) is estimated by $(\hat{K}, \hat{L}_\infty)$. The uncertainty in the growth parameters will affect \hat{F}_{40} . From Table 6:

$$\begin{aligned} \frac{E(\hat{F}_{40}) - \hat{F}_{40}}{\hat{F}_{40}} &\approx \frac{E(\hat{F}_{40}) - \tilde{F}_{40}}{\tilde{F}_{40}} \approx 1.089 \left(\frac{E(\hat{K}) - K}{K} \right) + 2.649 \left(\frac{E(\hat{L}_\infty) - L_\infty}{L_\infty} \right) \\ &\quad - 0.002 \cdot \frac{\text{Var}(\hat{K})}{K^2} + 5.507 \cdot \frac{\text{Cov}(\hat{K}, \hat{L}_\infty)}{K \cdot L_\infty} - 3.53 \cdot \frac{\text{Var}(\hat{L}_\infty)}{L_\infty^2} \end{aligned}$$

If unbiased estimators are available for K and L_∞ , only the second order coefficients will appear.

The variance of \hat{F}_{40} , and then \tilde{F}_{40} can also be approximated. Only the first order coefficients will be used, so that:

$$\frac{\text{Var}(\hat{F}_{40})}{F_{40}} \approx \frac{\text{Var}(\tilde{F}_{40})}{\tilde{F}_{40}} \approx (1.089)^2 \cdot \frac{\text{Var}(\hat{K})}{K^2}$$

$$+ (1.089)(2.649) \cdot \frac{2 \text{Cov}(\hat{K}, \hat{L}_\infty)}{K^2} + (2.649)^2 \cdot \frac{\text{Var}(\hat{L}_\infty)}{L_\infty^2}$$

When only the variance in \hat{K} is considered, the coefficient of variation of \hat{F}_{40} will be simply equal to 1.089 cv(\hat{K}), cv(\hat{K}) being the coefficient of variation of \hat{K} . However, due to the generally high correlations between \hat{K} and \hat{L}_∞ it is wiser to refer to the joint uncertainty of \hat{K} and \hat{L}_∞ .

It can be noticed in Table 5 that $A_K = -A_M$. This corresponds to the fact that in the formulas only K/M appears. It can also be noticed that the coefficients A_L remain moderate for small and medium lengths, but increase very rapidly when the length comes close to L_∞ .

INDIVIDUAL VARIABILITY IN THE GROWTH PATTERN

General Background

Following are comments on the general approach and successive steps used.

It is assumed that the rate of natural mortality is constant, and that the fishing mortality $F(l)$ equally affects all the individuals of length l , whatever their age at that size. This assumption deserves some discussion. In particular, it implies that the behavior and distribution of all the animals of a given size is independent of their age.

Table 4. A comparison between 'true' \hat{F} , solution of the catch equation for class 40, step = 1, \tilde{F} , and \hat{F}^* calculated by expansions about the reference value of \hat{F} , for various relative errors in the parameters.

	K	L_∞	M	\hat{F}	\tilde{F}	\hat{F}^*
Reference	0.50	70.0	0.20	1.578	1.497	
K - 30%	0.35	70.0	0.20	1.061	1.008	1.062
K - 20%	0.40	70.0	0.20	1.233	1.171	1.234
K - 10%	0.45	70.0	0.20	1.406	1.344	1.406
K + 10%	0.55	70.0	0.20	1.751	1.660	1.750
K + 20%	0.60	70.0	0.20	1.924	1.823	1.922
K + 30%	0.65	70.0	0.20	2.096	1.986	2.094
L_∞ - 5%	0.50	66.5	0.20	1.366	1.296	1.368
L_∞ + 5%	0.50	73.5	0.20	1.788	1.694	1.786
L_∞ + 10%	0.50	77.0	0.20	1.995	1.890	1.991
L_∞ + 20%	0.50	84.0	0.20	2.409	2.280	2.392
L_∞ + 30%	0.50	91.0	0.20	2.821	2.669	2.783
M - 50%	0.50	70.0	0.10	1.652	1.563	1.648
M - 25%	0.50	70.0	0.15	1.615	1.530	1.613
M - 10%	0.50	70.0	0.18	1.593	1.510	1.592
M + 10%	0.50	70.0	0.22	1.564	1.483	1.564
M + 25%	0.50	70.0	0.25	1.541	1.463	1.543
M + 50%	0.50	70.0	0.30	1.504	1.430	1.508
M + 75%	0.50	70.0	0.35	1.467	1.397	1.473
M + 100%	0.50	70.0	0.40	1.431	1.363	1.437

Table 5. Effects upon estimated stock numbers of uncertainties in the parameters.

I_i	A_K	A_L	A_{KK}	A_{KL}	A_{LL}	A_M	A_{MM}	$A_{1/E}$
14	-.186	-.329	.209	-6.073	.747	.186	.022	.0006
15	-.179	-.320	.200	-5.824	.733	.179	.021	.0006
16	-.172	-.311	.192	-5.575	.718	.172	.020	.0006
17	-.164	-.301	.183	-5.324	.703	.164	.018	.0006
18	-.157	-.291	.175	-5.073	.687	.157	.017	.0006
19	-.149	-.281	.166	-4.819	.671	.149	.016	.0006
20	-.142	-.271	.157	-4.557	.653	.142	.015	.0006
21	-.134	-.260	.148	-4.298	.636	.134	.014	.0006
22	-.126	-.249	.140	-4.040	.619	.126	.013	.0006
23	-.120	-.241	.133	-3.833	.608	.120	.012	.0006
24	-.114	-.233	.126	-3.622	.598	.114	.011	.0006
25	-.109	-.227	.120	-3.442	.592	.109	.011	.0007
26	-.106	-.224	.116	-3.321	.597	.106	.010	.0007
27	-.104	-.225	.114	-3.247	.610	.104	.010	.0007
28	-.100	-.221	.110	-3.108	.613	.100	.009	.0008
29	-.096	-.217	.105	-2.976	.617	.096	.009	.0008
30	-.092	-.215	.102	-2.855	.625	.092	.009	.0009
31	-.089	-.213	.098	-2.748	.637	.089	.008	.0009
32	-.087	-.214	.096	-2.671	.657	.087	.008	.0010
33	-.087	-.219	.096	-2.642	.692	.087	.008	.0012
34	-.088	-.227	.097	-2.651	.739	.088	.008	.0013
35	-.088	-.236	.098	-2.656	.790	.088	.008	.0015
36	-.089	-.246	.099	-2.674	.849	.089	.008	.0017
37	-.090	-.256	.100	-2.690	.912	.090	.009	.0020
38	-.092	-.267	.102	-2.710	.983	.092	.009	.0023
39	-.092	-.276	.102	-2.693	1.046	.092	.009	.0026
40	-.089	-.276	.099	-2.598	1.085	.089	.008	.0029
41	-.085	-.274	.095	-2.474	1.117	.085	.008	.0032
42	-.083	-.278	.093	-2.396	1.176	.083	.008	.0036
43	-.083	-.289	.093	-2.380	1.275	.083	.008	.0043
44	-.083	-.302	.094	-2.373	1.390	.083	.008	.0051
45	-.085	-.322	.097	-2.416	1.550	.085	.008	.0062
46	-.084	-.332	.096	-2.371	1.674	.084	.008	.0073
47	-.083	-.345	.096	-2.335	1.822	.083	.008	.0088
48	-.083	-.363	.098	-2.328	2.018	.083	.008	.0107
49	-.085	-.391	.102	-2.370	2.286	.085	.009	.0134
50	-.086	-.419	.105	-2.401	2.583	.086	.009	.0167
51	-.085	-.438	.106	-2.360	2.852	.085	.009	.0206
52	-.085	-.467	.110	-2.366	3.226	.085	.009	.0260
53	-.086	-.499	.115	-2.371	3.667	.086	.009	.0332
54	-.089	-.557	.126	-2.481	4.352	.089	.010	.0443
55	-.089	-.599	.133	-2.492	4.984	.089	.010	.0574
56	-.089	-.642	.144	-2.492	5.720	.089	.010	.0752
57	-.091	-.720	.164	-2.605	6.865	.091	.011	.1037
58	-.096	-.825	.195	-2.788	8.402	.096	.012	.1466
59	-.113	-1.054	.263	-3.369	11.344	.113	.014	.2281
60	-.134	-1.335	.357	-4.096	15.000	.134	.015	.3488
61	-.133	-1.425	.422	-4.232	16.644	.133	.013	.4616
62	-.110	-1.279	.452	-3.700	15.558	.110	.008	.5551
63	-.084	-1.059	.514	-3.026	13.314	.084	.004	.7059
64	-.031	-4.457	.531	-1.362	5.733	.031	.001	.8227

A_K , A_L , A_M and $A_{1/E}$ designate the first order sensitivity coefficients associated, respectively, with K , L_∞ , M and $1/E$.

A_{KK} , A_{LL} , A_{MM} designate the second order sensitivity coefficients associated with K , L_∞ and M .

A_{KL} is the coefficient of crossed sensitivity associated with K and L_∞ .

Terminal $E = 0.70$, $K = 0.50$, $L_\infty = 70.0$, $M = 0.20$.

Table 6. Effects upon fishing mortalities of uncertainties in the parameters.

l_i	A_K	A_L	A_{KK}	A_{KL}	A_{LL}	A_M	A_{MM}
14	1.186	1.590	.012	8.021	-.224	-.186	.012
15	1.179	1.604	.011	7.772	-.219	-.179	.011
16	1.172	1.619	.009	7.525	-.215	-.172	.010
17	1.164	1.634	.008	7.276	-.211	-.164	.008
18	1.157	1.651	.007	7.028	-.206	-.157	.007
19	1.149	1.667	.006	6.777	-.202	-.149	.006
20	1.142	1.685	.005	6.519	-.198	-.142	.005
21	1.134	1.703	.004	6.264	-.193	-.134	.004
22	1.126	1.723	.003	6.012	-.189	-.126	.003
23	1.120	1.747	.002	5.819	-.187	-.120	.002
24	1.114	1.772	.001	5.623	-.185	-.114	.002
25	1.109	1.800	.001	5.462	-.184	-.109	.001
26	1.106	1.833	.000	5.371	-.186	-.106	.001
27	1.104	1.872	.000	5.336	-.190	-.104	.001
28	1.100	1.907	-.000	5.227	-.192	-.100	.000
29	1.096	1.946	-.000	5.128	-.195	-.096	.000
30	1.092	1.987	-.001	5.045	-.199	-.092	-.000
31	1.089	2.031	-.001	4.980	-.204	-.089	-.000
32	1.087	2.081	-.001	4.952	-.212	-.087	-.001
33	1.087	2.137	-.001	4.983	-.224	-.087	-.001
34	1.088	2.199	-.001	5.063	-.240	-.088	-.001
35	1.088	2.265	-.001	5.142	-.256	-.088	-.001
36	1.089	2.335	-.002	5.240	-.275	-.089	-.000
37	1.090	2.410	-.002	5.341	-.295	-.090	-.000
38	1.092	2.490	-.002	5.452	-.317	-.092	-.000
39	1.092	2.571	-.002	5.524	-.337	-.092	-.000
40	1.089	2.649	-.002	5.507	-.353	-.089	-.000
41	1.085	2.730	-.003	5.460	-.368	-.085	-.001
42	1.083	2.823	-.003	5.476	-.393	-.083	-.001
43	1.083	2.930	-.004	5.577	-.429	-.083	-.001
44	1.083	3.047	-.004	5.698	-.471	-.083	-.001
45	1.085	3.179	-.005	5.893	-.525	-.085	-.001
46	1.084	3.311	-.005	5.988	-.573	-.084	-.001
47	1.083	3.456	-.006	6.106	-.630	-.083	-.001
48	1.083	3.619	-.008	6.279	-.703	-.083	-.001
49	1.085	3.806	-.009	6.532	-.798	-.085	-.001
50	1.086	4.009	-.011	6.793	-.903	-.086	-.001
51	1.085	4.221	-.014	6.979	-.1005	-.085	-.002
52	1.085	4.467	-.017	7.254	-.1141	-.085	-.002
53	1.086	4.742	-.022	7.562	-.1300	-.086	-.002
54	1.089	5.073	-.028	8.058	-.1525	-.089	-.002
55	1.089	5.426	-.036	8.456	-.1735	-.089	-.002
56	1.089	5.828	-.047	8.893	-.1977	-.089	-.002
57	1.091	6.320	-.064	9.569	-.2311	-.091	-.003
58	1.096	6.912	-.090	10.444	-.2697	-.096	-.003
59	1.113	7.721	-.137	12.085	-.3203	-.113	-.001
60	1.134	8.704	-.205	14.142	-.3381	-.134	.003
61	1.133	9.660	-.272	15.367	-.2881	-.133	.005
62	1.110	10.612	-.330	15.625	-.1983	-.110	.004
63	1.084	11.828	-.423	15.931	-.794	-.084	.003
64	1.031	13.184	-.499	14.971	.291	-.031	.000

Terminal E = 0.70, K = 0.50, L_∞ = 0.70, M = 0.20.

As mentioned earlier, a population in which the K and L_∞ are distributed according to a joint distribution $\Omega(K, L_\infty)$ is still considered. Again the means of K and L_∞ are noted \bar{K} and \bar{L}_∞ , respectively;⁶ the squared coefficient of variation of the individual K is noted V_{KK} , and that of the individual L_∞ V_{LL} , while the coefficient of covariation between the individual K and L_∞ is noted:

$$V_{KL} = \frac{\text{Cov}(K, L_\infty)}{\bar{K} \cdot \bar{L}_\infty}$$

$S(K, L_\infty, l)$ designates the survival rate from l_c to l of all the individuals having given parameter values (K, L_∞) . The number of fish from a cohort surviving at length l is thus:

$$N(l) = \int \int \omega(K, L_\infty) \cdot S(K, L_\infty, l) dK dL_\infty$$

This number differs from the number that would be observed if all the individuals were growing with the same K ($= \bar{K}$) and L_∞ ($= \bar{L}_\infty$), which is $\bar{N}(l) = N_c \cdot \bar{S}$, where $\bar{S} = S(\bar{K}, \bar{L}_\infty, l)$ (it is not an arithmetic mean).

The discrepancy is expressed in relative terms as: $\frac{N(l) - \bar{N}(l)}{\bar{N}(l)}$

Making use once more of second order Taylor's expansion, of $S(K, L_\infty, l)$, as a function of K and L_∞ , one can approximate the true $N(l)$ by the quantity $N^*(l)$, the asterisk being hereafter associated with the approximations obtained through Taylor's expansion. The discrepancy will be approximated through $[N^*(l) - \bar{N}(l)]/\bar{N}(l)$ which will be expressed as a linear function of the coefficients of variation-covariation.

At length l_i , we have $N(l_i) = N_i$, $\bar{N}(l_i) = \bar{N}_i$ and $N^*(l_i) = N^*_i$, and:

$$\frac{N_i^* - \bar{N}_i}{N_i} \approx \beta_{1,i} \cdot V_{KK} + \beta_{2,i} \cdot V_{KL} + \beta_{3,i} \cdot V_{LL}$$

and similarly

$$\frac{\bar{N}^*(\bar{l}_i) - \bar{N}(\bar{l}_i)}{\bar{N}(\bar{l}_i)} \approx \beta_{1,i}^S \cdot V_{KK} + \beta_{2,i}^S \cdot V_{KL} + \beta_{3,i}^S \cdot V_{LL}$$

for the middle of the length intervals.

The coefficients $\beta_{m,i}$ and $\beta_{m,i}^S$ are similar to those previously defined as sensitivity coefficients. They will be called "discrepancy coefficients".

The discrepancy affecting the number of survivors at length will in turn give rise to a discrepancy in the estimate of catches over the corresponding length intervals. The true catches $c(l)$ will again be approximated by the quantities $c^*(l)$. As a matter of fact, if $F(l)$ is constant and equal to F_i within the interval (l_i, l_{i+1}) ,

$$c^*_i = \frac{F_i}{F_i + M} (N^*_i - N^*_{i+1})$$

An additional set of discrepancy coefficients will therefore be defined to estimate $(c^* - \bar{C})/\bar{C}$ (where \bar{C} is the catch predicted when \bar{K} and \bar{L}_∞ hold for all the individuals) by a linear relationship:

$$\frac{c^*_i - \bar{C}_i}{\bar{C}_i} \approx \beta_{1,i}^C \cdot V_{KK} + \beta_{2,i}^C \cdot V_{KL} + \beta_{3,i}^C \cdot V_{LL}$$

The coefficients $\beta_{m,i}$, $\beta_{m,i}^S$ and $\beta_{m,i}^C$ are used in the first part of the analysis of the influence of individual variability in growth. Assuming that the fishing mortalities at length are known, we want to

⁶The values K and L_∞ are generally different from those given by a direct regression of length on age (Sainsbury 1980).

assess the effects on the number of survivors at each length and on the catches at length, as compared to what would be obtained had all the individuals a common pair of growth parameters (K , L_∞). These aspects are addressed below.

Afterwards, attention will be paid to the fishing mortalities estimated through length cohort analysis. In practice, the catches at length $C_i(l)$ are analyzed as if they were associated with common K and \bar{L}_∞ for all animals. The solutions \tilde{F}_i to the catch equations for observed C_i generally differ from the true F_i , unless there is no individual variability in growth.

It is not possible, however, to analyze simply the differences $\tilde{F}_i - F_i$. The reconstructed numbers \tilde{N}_i and the derived fishing mortalities \tilde{F}_i presented earlier should be considered. Reconstructed numbers (see equation 5) are the $\tilde{N}(l)$ derived at each length from catches over larger length intervals.

Again, this cannot be done directly. We must use an additional approximation $\tilde{N}^*(l)$ derived from the $C^*(l)$ obtained by Taylor's expansion as indicated by the asterisk. $\tilde{N}^*(l)$ will in turn be compared to the number $\bar{N}(l)$ which would be obtained in the absence of individual variability in growth, $\tilde{N}^*(l) - \tilde{N}(l)$ being associated with coefficients $\beta_{m,i}^N$ ($m = 1, 2, 3$).

The discrepancy in the reconstructed stock numbers is not of major concern here, but is an intermediate step to assess the errors affecting the fishing mortalities. In fact, instead of the differences $\tilde{F}_i - F_i = \tilde{F}_i - \bar{F}_i$ (where the \tilde{F}_i are produced by a usual cohort analysis of the \bar{C}_i), consider the differences $\tilde{F}_i - \bar{F}_i$, or more precisely $\tilde{F}_i^* - \bar{F}_i$. The F_i^* are derived from the ratios of the C_i^* and the N_i^* . When the discrepancy $N_i - \bar{N}_i$ is moderate, it will be possible to use an additional approximation, leading to number F_i^{**} and the relationship:

$$\frac{\tilde{F}_i^{**} - \tilde{F}_i}{\tilde{F}_i} \approx \beta_{1,i}^F \cdot v_{KK} + \beta_{2,i}^F \cdot v_{KL} + \beta_{3,i}^F \cdot v_{LL}$$

These considerations on reconstructed stock numbers and fishing mortalities are developed below.

To summarize, here are two sets of estimates:

$\bar{N}(l_i)$ = number of survivors at the middle of length interval (l_i, l_{i+1}) when \bar{K} and \bar{L}_∞ are common to all individuals

$N(l_i)$ = true number of survivors

$N^*(l_i)$ = approximation through second order Taylor's expansion of $S(K, L_\infty, l)$

and

$\tilde{N}(l_i)$ = stock number reconstructed from the \bar{C}_i

$\tilde{N}(l_i)$ = stock number reconstructed from the C_i

$\tilde{N}^*(l_i)$ = stock number reconstructed from the C_i^* (derived themselves from the $N^*(l_i)$). and the relationships:

$$\frac{[N^*(l_i) - \bar{N}(l_i)]}{N(l_i)} = \beta_{1,i}^S \cdot v_{KK} + \beta_{2,i}^S \cdot v_{KL} + \beta_{3,i}^S \cdot v_{LL}$$

$$\frac{[\tilde{N}^*(l_i) - \tilde{N}(l_i)]}{\tilde{N}(l_i)} = \beta_{1,i}^N - v_{KK} - \beta_{2,i}^N \cdot v_{KL} + \beta_{3,i}^N \cdot v_{LL}$$

$$\frac{[(\tilde{F}_i^{**} - \tilde{F}_i)]}{\tilde{F}_i} = \beta_{1,i}^F \cdot v_{KK} + \beta_{2,i}^F \cdot v_{KL} + \beta_{3,i}^F \cdot v_{LL}$$

Discrepancies in the Number of Survivors and Catches at Length

Survivors in the continuous case

The number of fish surviving at length l is:

$$N(l) = \iint \omega(K, L_\infty) \cdot S^+(K, L_\infty, l) \cdot dK \cdot dL_\infty$$

By means of Taylor's expansion, we find an approximation $N^*(l)$ of $N(l)$:

$$N(l) \approx N^*(l) = N_C \left(S(\bar{K}, \bar{L}_\infty, l) + \frac{1}{2} \cdot \frac{\partial^2 S(\bar{K}, \bar{L}_\infty, l)}{\partial K^2} \cdot \text{Var}(K) + \frac{\partial^2 S(\bar{K}, \bar{L}_\infty, l)}{\partial K \cdot \partial L_\infty} \cdot \text{Cov}(K, L_\infty) \right. \\ \left. + \frac{1}{2} \cdot \frac{\partial^2 S(\bar{K}, \bar{L}_\infty, l)}{\partial L_\infty^2} \cdot \text{Var}(L_\infty) \right)$$

At that point, we have to develop the derivatives, whence:

$$S(K, L_\infty, l) = \left(\frac{L_\infty - l_C}{L_\infty - l} \right) - \frac{M}{K} \cdot e - \frac{1}{K} \int_{l_C}^l \frac{F(r)}{L_\infty - r} \cdot dr$$

The first factor corresponds to a function the derivatives of which have previously been presented.

$$\left(\frac{L_\infty - l_C}{L_\infty - l} \right)^{-M/K} = \phi(l, l_C, M, K, L_\infty)$$

and the second one will be noted $\psi(l, l_C, K, L_\infty)$

The derivations are carried out by combinations of derivatives of ϕ and ψ . Those of ψ are easy to obtain

$$\begin{aligned} \cdot \frac{1}{\psi} \cdot \frac{\partial \psi}{\partial K} &= \psi_K = \frac{1}{K^2} \int_{l_C}^l \frac{F(r)}{L_\infty - r} \cdot dr \\ \cdot \frac{1}{\psi} \cdot \frac{\partial \psi}{\partial L_\infty} &= \psi_L = \frac{1}{K} \int_{l_C}^l \frac{F(r)}{(L_\infty - r)^2} \cdot dr \\ \cdot \frac{1}{\psi} \cdot \frac{\partial^2 \psi}{\partial K^2} &= \psi_{KK} = -\frac{2}{K^3} \int_{l_C}^l \frac{F(r)}{L_\infty - r} \cdot dr + \psi_K^2 \\ \cdot \frac{1}{\psi} \cdot \frac{\partial^2 \psi}{\partial K \cdot \partial L_\infty} &= \psi_{KL} = -\frac{1}{K^2} \int_{l_C}^l \frac{F(r)}{(L_\infty - r)^2} \cdot dr + \psi_K \cdot \psi_L \\ \cdot \frac{1}{\psi} \cdot \frac{\partial^2 \psi}{\partial L_\infty^2} &= \psi_{LL} = -\frac{2}{K} \int_{l_C}^l \frac{F(r)}{(L_\infty - r)^3} \cdot dr + \psi_L^2 \end{aligned}$$

Thus:

$$\begin{aligned} \frac{1}{S(K, L_\infty, l)} \cdot \frac{\partial^2 S}{\partial K^2} &= \frac{1}{\phi} \cdot \frac{\partial^2 \phi}{\partial K^2} + \frac{1}{\psi} \cdot \frac{\partial^2 \psi}{\partial K^2} + 2 \cdot \frac{1}{\phi} \cdot \frac{\partial \phi}{\partial K} \cdot \frac{1}{\psi} \cdot \frac{\partial \psi}{\partial K} \\ \frac{1}{S(K, L_\infty, l)} \cdot \frac{\partial^2 S}{\partial K \cdot \partial L_\infty} &= \frac{1}{\phi} \cdot \frac{\partial^2 \phi}{\partial K \cdot \partial L_\infty} + \frac{1}{\psi} \cdot \frac{\partial^2 \psi}{\partial K \cdot \partial L_\infty} + \frac{1}{\phi} \cdot \frac{\partial \phi}{\partial K} \cdot \frac{1}{\psi} \cdot \frac{\partial \phi}{\partial L_\infty} + \frac{1}{\phi} \cdot \frac{\partial \phi}{\partial L_\infty} \cdot \frac{1}{\psi} \cdot \frac{\partial \psi}{\partial K} \\ \frac{1}{S(K, L_\infty, l)} \cdot \frac{\partial^2 S}{\partial L_\infty^2} &= \frac{1}{\phi} \cdot \frac{\partial^2 \phi}{\partial L_\infty^2} + \frac{1}{\psi} \cdot \frac{\partial^2 \psi}{\partial L_\infty^2} + 2 \cdot \frac{1}{\phi} \cdot \frac{\partial \phi}{\partial L_\infty} \cdot \frac{1}{\psi} \cdot \frac{\partial \psi}{\partial L_\infty} \end{aligned}$$

Survivors in the discrete case

We consider a discrete scheme in which $F(l) = F_i$ for $l_i \leq l \leq l_{i+1}$ and denote:

$$k_{1,i} = \frac{1}{K} \int_{l_1}^{l_i} \frac{F(r)}{(L_\infty - r)} \cdot dr = \frac{1}{K} \left(\sum_{j=1}^{i-1} F_j \int_{l_j}^{l_{j+1}} \frac{1}{(L_\infty - r)} \cdot dr \right)$$

$$k_{1,i} = \frac{1}{K} \left[\sum_{j=1}^{i-1} F_j \cdot \log \left(\frac{L_\infty - l_j}{L_\infty - l_{j+1}} \right) \right], \quad j = 2, \dots, I$$

$$k_{2,i} = \frac{1}{K} \int_{l_1}^{l_i} \frac{F(r)}{(L_\infty - r)^2} \cdot dr = \frac{1}{K} \left[\sum_{j=1}^{i-1} F_j \left(\int_{l_j}^{l_{j+1}} \frac{1}{(L_\infty - r)^2} \cdot dr \right) \right]$$

$$k_{2,i} = \frac{1}{K} \left[\sum_{j=1}^{i-1} F_j \left(\frac{1}{L_\infty - l_{j+1}} - \frac{1}{L_\infty - l_j} \right) \right], \quad j = 2, \dots, I$$

and, similarly:

$$k_{3,i} = \frac{2}{K} \int_{l_1}^{l_i} \frac{F(r)}{(L_\infty - r)^3} \cdot dr$$

$$k_{3,i} = \frac{1}{K} \left[\sum_{j=1}^{i-1} F_j \left(\frac{1}{(L_\infty - l_{j+1})^2} - \frac{1}{(L_\infty - l_j)^2} \right) \right], \quad j = 2, \dots, I$$

With these notations, at length l_i :

$$\frac{1}{\Psi} \cdot \frac{\partial \Psi}{\partial K} = \frac{k_{1,i}}{K} (= \Psi K_i)$$

$$\frac{1}{\Psi} \cdot \frac{\partial \Psi}{\partial L_\infty} = k_{2,i} (= \Psi L_i)$$

$$\frac{1}{\Psi} \cdot \frac{\partial^2 \Psi}{\partial K^2} = \frac{k_{1,i}}{K^2} \cdot (k_{1,i} - 2) \quad (= \Psi K K_i)$$

$$\frac{1}{\Psi} \cdot \frac{\partial^2 \Psi}{\partial K \cdot \partial L_\infty} = -\frac{k_{2,i}}{K} + \frac{k_{1,i} \cdot k_{2,i}}{K} \quad (= \Psi K L_i)$$

$$\frac{1}{\Psi} \cdot \frac{\partial^2 \Psi}{\partial L_\infty^2} = k_{3,i} + k_{2,i}^2 \quad (= \Psi L L_i)$$

$\phi K_i, \phi L_i, \phi K K_i, \phi K L_i$ and $\phi L L_i$ are defined similarly by derivation of ϕ .

In addition, note: $\bar{S}_i = S(\bar{K}, \bar{L}_\infty, l_i)$ and $\bar{N}_i = N_c \cdot \bar{S}_i$ (these are not means, but quantities calculated using the means of K and L_∞).

Thus

$$\frac{1}{\bar{S}_i} \cdot \frac{\partial^2 S(l_i)}{\partial K^2} = g_{1,i} = \phi_{KK} + \psi_{KK} + 2 \cdot \phi_{Ki} \cdot \psi_{Ki}$$

$$\frac{1}{\bar{S}_i} \cdot \frac{\partial^2 S(l_i)}{\partial K \cdot \partial L_\infty} = g_{2,i} = \phi_{KL} + \psi_{KL} + \phi_{Ki} \cdot \psi_{Li} + \phi_{Li} \cdot \psi_{Ki}$$

$$\frac{1}{\bar{S}_i} \cdot \frac{\partial^2 S(l_i)}{\partial L_\infty^2} = g_{3,i} = \phi_{LL} + \psi_{LL} + 2 \cdot \phi_{Li} \cdot \psi_{Li}$$

and:

$$N(l_i) \approx N_i^* = \bar{N}_i \left(1 + \frac{g_{1,i}}{2} \text{Var}(K) + g_{2,i} \cdot \text{Cov}(K, L_\infty) + \frac{g_{3,i}}{2} \text{Var}(L_\infty) \right)$$

In order to assess the effects of individual variability, reference will again be made to the quantities $\frac{N_i^* - \bar{N}_i}{\bar{N}_i}$ and $\frac{\text{Var}(K)}{\bar{K}^2}$, $\frac{\text{Cov}(K, L_\infty)}{\bar{K} \cdot \bar{L}_\infty}$, $\frac{\text{Var}(L_\infty)}{\bar{L}_\infty^2}$

This implies that one is to consider the coefficients $(\bar{K}^2/2) g_{1,i}$, $\bar{K} \cdot \bar{L}_\infty \cdot g_{2,i}$ and $(\bar{L}_\infty^2/2) g_{3,i}$ respectively noted $\beta_{1,i}$, $\beta_{2,i}$ and $\beta_{3,i}$.

Catches at length

An approximation of the numbers caught can follow directly, using the relationships:

$$C_i \approx C_i^* = (N_i^* - N_{i+1}^*) \cdot \frac{F_i}{F_i + M}, \text{ for } i = 1, \dots, I-1$$

and

$$C_I \approx C_I^* = N_I^* \cdot \frac{F_I}{F_I + M} = N_I^* \cdot E$$

thus

$$C_i^* = \frac{F_i}{F_i + M} \cdot N_c [\bar{S}_i - \bar{S}_{i+1} + B_i]$$

where

$$B_i = \frac{\text{Var}(K)}{2} \cdot (g_{1,i} \cdot \bar{S}_i - g_{1,i+1} \cdot \bar{S}_{i+1})$$

$$+ \text{Cov}(K, L_\infty) \cdot (g_{2,i} \cdot \bar{S}_i - g_{2,i+1} \cdot \bar{S}_{i+1})$$

$$+ \frac{\text{Var}(L_\infty)}{2} \cdot (g_{3,i} \cdot \bar{S}_i - g_{3,i+1} \cdot \bar{S}_{i+1})$$

for $i = 1, \dots, I$ and, by convention, $g_{1,I+1} = g_{2,I+1} = g_{3,I+1} = 0$.

$$\bar{C}_i = \frac{F_i}{F_i + M} \cdot N_c (\bar{S}_i - \bar{S}_{i+1})$$

is the estimate of catches at length i based on survival rates corresponding to \bar{K} and \bar{L}_∞ .

As in previous sections, we are interested in the relationships between $(C_i^* - \bar{C}_i)/\bar{C}_i$ and the coefficients of variation-covariation of K and L_∞ , thus:

$$\frac{C_i^* - \bar{C}_i}{\bar{C}_i} \approx \beta_{1,i}^c \cdot v_{KK} + \beta_{2,i}^c \cdot v_{KL} + \beta_{3,i}^c \cdot v_{LL}$$

where

$$\beta_{1,i}^c = \bar{K}^2 (g_{1,i} - \bar{s}_i - g_{1,i+1} \cdot \bar{s}_{i+1}) / 2 (\bar{s}_i - \bar{s}_{i+1})$$

$$\beta_{2,i}^c = \bar{K} \cdot \bar{L}_\infty (g_{2,i} \cdot \bar{s}_i - g_{2,i+1} \cdot \bar{s}_{i+1}) / (\bar{s}_i - \bar{s}_{i+1})$$

$$\beta_{3,i}^c = \bar{L}_\infty^2 (g_{3,i} \cdot \bar{s}_i - g_{3,i+1} \cdot \bar{s}_{i+1}) / 2 (\bar{s}_i - \bar{s}_{i+1})$$

Reconstructed Stock Numbers and Estimated Fishing Mortalities

Reconstructed stock numbers

If a cohort analysis is carried out using the C_i , it will not provide the actual N_i and F_i , even if the terminal F_I^+ is exact. If there is no individual variability, it estimates:

$$\tilde{N}_i = \sum_{j>i} \frac{v_j}{v_i} \cdot \bar{C}_i \quad \text{and} \quad \tilde{F}_i = \bar{K} \cdot (\bar{L}_\infty - \bar{L}_i) \cdot \frac{\bar{C}_i}{\tilde{N}_i} \cdot \frac{1}{(l_{i+1} - l_i)}$$

The differences between the \tilde{F}_i and the true F_i stem from the approximations due to the discretisation which were studied in previous sections.

When individual variability in growth occurs additional biases occur due to the discrepancies between C_i and \bar{C}_i .

The stock number reconstructed at length \bar{l}_i is:

$$\tilde{N}_i = \sum_{j \geq i} \frac{v_j}{v_i} \cdot C_j$$

which is approximated by:

$$\tilde{N}_i^* = \sum_{j \geq i} \frac{v_j}{v_i} \cdot C_j^*$$

and

$$\tilde{N}_i^* - \tilde{N}_i = \sum_{j \geq i} \frac{v_j}{v_i} (C_j^* - \bar{C}_i)$$

In order to obtain the relative deviation $(\tilde{N}_i^* - \tilde{N}_i)/\tilde{N}_i$, as dealt with previously, estimate the coefficients:

$$\beta_{1,i}^N = \frac{\bar{K}^2}{2} \left(\sum_{j \geq i} v_j \cdot \frac{F_j}{F_j + M} (g_{1,j} \cdot \bar{s}_j - g_{1,j+1} \cdot \bar{s}_{j+1}) \right) / \left(\sum_{j \geq i} v_j \cdot \frac{F_j}{F_j + M} (\bar{s}_j - \bar{s}_{j+1}) \right)$$

$$\beta_{2,i}^N = \bar{K} \cdot \bar{L}_\infty \left(\sum_{j \geq i} v_j \cdot \frac{F_j}{F_j + M} (g_{2,j} \cdot \bar{s}_j - g_{2,j+1} \cdot \bar{s}_{j+1}) \right) / \left(\sum_{j \geq i} v_j \cdot \frac{F_j}{F_j + M} (\bar{s}_j - \bar{s}_{j+1}) \right)$$

$$\beta_{3,i}^N = \bar{L}_\infty^2 \left(\sum_{j \geq i} v_j \cdot \frac{F_j}{F_j + M} (g_{3,j} \cdot \bar{s}_j - g_{3,j+1} \cdot \bar{s}_{j+1}) \right) / \left(\sum_{j \geq i} v_j \cdot \frac{F_j}{F_j + M} (\bar{s}_j - \bar{s}_{j+1}) \right)$$

Mortalities

$$\text{Finally, we have: } \tilde{F}_i^* = \bar{K} \cdot (\bar{L}_\infty - \bar{l}_i) \cdot \frac{C_i^*}{\tilde{N}_i^*} \cdot \frac{1}{(\bar{l}_{i+1} - \bar{l}_i)}$$

and

$$\frac{\tilde{F}_i^*}{\tilde{F}_i} = \frac{C_i^*}{\bar{C}_i} \cdot \frac{\tilde{N}_i}{\tilde{N}_i^*} = \left(1 + \frac{C_i^* - \bar{C}_i}{\bar{C}_i} \right) \left(\frac{1}{1 + \frac{\tilde{N}_i^* - \tilde{N}_i}{\tilde{N}_i}} \right)$$

When (and only when) $\frac{\tilde{N}_i^* - \tilde{N}_i}{\tilde{N}_i}$ is small enough compared to 1, an additional approximation can be used:

$$\frac{1}{1 + \frac{\tilde{N}_i^* - \tilde{N}_i}{\tilde{N}_i}} \approx 1 - \frac{\tilde{N}_i^* - \tilde{N}_i}{\tilde{N}_i}$$

It will lead to a new approximation, \tilde{F}_i^{**} , the approximate term of relative deviation being expressed as:

$$\frac{\tilde{F}_i^{**} - \tilde{F}_i^*}{\tilde{F}_i^*} = \beta_{1,i}^F \cdot v_{KK} + \beta_{2,i}^F \cdot v_{KL} + \beta_{3,i}^F \cdot v_{LL}$$

where

$$\beta_{m,i}^F = \beta_m^C - \beta_m^N, \quad m = 1, 2, 3$$

In fact since we are interested in the fishing mortalities \hat{F}_i obtained from the catch equations, we shall define \hat{F}_i^* and \hat{F}_i^{**} by:

$$\text{and } \frac{\hat{F}_i^*}{\hat{F}_i} = \frac{\tilde{F}_i^*}{\tilde{F}_i}$$

$$\frac{\hat{F}_i^{**}}{\hat{F}_i} = \frac{\tilde{F}_i^{**}}{\tilde{F}_i}$$

Example

An illustration is given in Tables 7, 8 and 9, based upon the basic data set of Table 1.

Table 7 is devoted to catches at length, and gives the values of the corresponding coefficients, as discussed above. The risks of discrepancies appear low up to a length of 50. With a coefficient of variation of 0.2 (which appears quite large) for K, under 50, the maximum relative bias will be about 4%. Discrepancies are slightly bigger for L_∞ , for the same coefficient of variation (which may not be the case). For both parameters, individually, and when taking into account the co-variation effect, the problems increase sharply with lengths beyond 50. It is also worth noticing the overall pattern of the coefficients. They introduce negative biases for medium lengths, where most of the catches are concentrated, and positive biases on both sides. In other words, individual variability in the growth parameters tends to smooth the dome-shaped distribution.

Table 8 is associated with the numbers of survivors at a given length. Several coefficients appear in the different columns.

Comparing $\beta_{m,i}^S$ and $\beta_{m,i}^N$ makes it possible to appreciate the effect of the discrepancy between the real numbers of survivors and the reconstructed numbers, $D_{m,i} = \beta_{m,i}^S - \beta_{m,i}^N$. Table 8 shows that significant differences may appear between $N_c \bar{S}$ and the real numbers of survivors, again mainly beyond length 50. It also shows that the differences between the real numbers and the reconstructed numbers from the real catches will be much smaller; the coefficient $\beta_{m,i}^S$ and $\beta_{m,i}^N$ are quite similar. The $D_{m,i}$ are small. However, for the largest lengths, significant differences may appear.

Finally Table 9 is devoted to the possible discrepancies in fishing mortalities. The coefficient $\beta_{m,i}^F$, defined above described the influence on the difference between \tilde{F}_i and the corresponding estimate \hat{F}_i^{**} and $\tilde{\hat{F}}_i^{**}$.

As mentioned, when the discrepancy $(\tilde{N} - \tilde{\tilde{N}})/\tilde{N}$ becomes too large, the approximations used to define the $\beta_{m,i}^F$ become poor.

Since several approximations have been used, the overall quality of the results has to be checked. Complete simulations would be useful. Very simple ones have only been used here. They consider that two equally numerous components exist. In the first component the growth parameters are K^1 and L^1_∞ ; they are K^2 and L^2_∞ in the second one.

$$\frac{K^1 + K^2}{2} = \bar{K} = 0.5$$

With

$$\frac{L^1_\infty + L^2_\infty}{2} = \bar{L}_\infty = 70.0$$

Three combinations have been used:

In the first one, $K^1 = .56, K^2 = .44, L^1_\infty = L^2_\infty = 70$

In the second one, $K^1 = K^2 = .50, L^1_\infty = 61., L^2_\infty = 79$

In the third one, $K^1 = .56, K^2 = .44, L^1_\infty = 61, L^2_\infty = 79$

The approximations using the coefficients previously defined appear as very satisfactory up to $l = 50$. For instance at length 40, the real fishing mortality being equal to 1.578, and \hat{F}_i^{**} being defined by

$$\frac{\hat{F}_i^{**}}{\hat{F}_i} = \frac{\tilde{F}_i^{**}}{\tilde{\hat{F}}_i}$$

$\hat{F}_i^{**} = 1.556$ while $\hat{F}_i = 1.554$ in the first combination

$\hat{F}_i^{**} = 1.468$ while $\hat{F}_i = 1.480$ in the second one, and

$\hat{F}_i^{**} = 1.616$ while $\hat{F}_i = 1.612$ in the last one.

The discrepancies become, as predicted by the discrepancy coefficients, much larger beyond 50. It is no more possible to use \hat{F}_i^{**} since $\frac{\tilde{N}_i^* - \tilde{\tilde{N}}_i}{\tilde{\tilde{N}}_i}$ is too large.

$$\frac{\tilde{N}_i^*}{\tilde{\tilde{N}}_i}$$

For the interval 55-56 (median 55.5) the real fishing mortality \hat{F}_i is 1.97. A VPA neglecting the individual variability associated to combination 1 would lead to $\hat{F}_i = 1.10$. The discrepancy is large since $\hat{F}_i - F_i / \hat{F}_i = -44\%$. Using $\beta_{3,i}^F$ would lead to predicted discrepancy equal to -70%, with $\hat{F}_i^{**} = .59$. This is a poor approximation, due to the fact that $\tilde{N}_i^* - \tilde{\tilde{N}}_i / \tilde{\tilde{N}}_i$ is equal to 1.05. However, the approximations for the catches and numbers remain good since, for instance, the predicted catches are equal to 35.8, while the real ones are 36.1 (and those that would be obtained with no variability 28.0).

Deriving F_i^* from $\frac{C_i^*}{C_i}$ and $\frac{\tilde{N}_i^*}{\tilde{\tilde{N}}_i}$ from $\frac{\hat{F}_i^*}{\hat{F}_i} = \frac{\tilde{F}_i^*}{\tilde{\hat{F}}_i}$ gives a better approximation

with $\hat{F}_i^* = 1.23$, not so far from $\hat{F}_i = 1.10$. In any case, the difference between \tilde{F}_i and F_i suggests that it will be difficult to get an excellent approximation of \tilde{F}_i from an approximation of F_i .

The possibility of correcting the F_i , in order to take into account the individual variability through the approximations presented here has not been checked. It would require further examination, and some estimate of $\text{Var}(K)$, $\text{Var}(L_\infty)$ and $\text{Cov}(K, L_\infty)$. In any case, it is useful once some estimates of F_i have been obtained to calculate the discrepancy coefficients defined here. They can be used as an indicator of the length ranges where big discrepancies are (or are not) likely to occur.

The example treated here suggests that individual variability in growth, especially the variability in L_∞ , may create serious problems for the estimation of fishing mortalities for length close to L_∞ . These problems do not necessarily affect the estimation of the fishing mortalities in smaller lengths. Among the possible sources of variability differences between males and females can be a major one, when a single length distribution is considered with both sexes pooled together. This should be avoided when possible, especially when length-specific fishing mortalities are calculated, and even if the mortalities are similar at the same length for both sexes.

Discussion

Taylor's expansions, properly used, offer a powerful tool for addressing several problems. They could be combined with simulations, and offer a synoptic view of several sensitivity problems, and sometimes a mathematical understanding of the results observed by simulations. They could be applied not only to the estimation of fishing mortalities, but cover yield-per-recruit or spawning biomass-per-recruit calculations as well (Laurec and Mesnil 1985).

It is possible to perform the sensitivity analyses on a microcomputer; the computer programs developed by the authors are quite simple. It would be wise to conduct such analyses systematically.

Large length grouping intervals should be avoided when possible. Smoothing is preferable. It can be performed on the basic length distribution of the catches, or on the calculated fishing mortalities at length. The choice between these possibilities, or their combination, has not yet been analyzed.

Risks of bias associated with sampling and estimation errors on K , L_∞ and on catches at length are generally moderate. Major problems may, however, occur when approaching L_∞ . The same is true for individual variabilities in growth parameters. In length cohort analyses the last length interval must have an initial length much smaller than L_∞ . In our example it must not be larger than 70% of L_∞ . This is quite close to what is suggested by Pereiro (1984). This problem can and should be however investigated for each specific case.

**Monte-Carlo Testing of Two Methods for
Estimating Growth from Length-Frequency Data
with General Conditions for Their Applicability**

A.A. ROSENBERG
J.R. BEDDINGTON

*Marine Resources Assessment Group
Center for Environmental Technology
Imperial College of Science and Technology
48 Prince's Gardens
London, SW7 1LU
United Kingdom*

Rosenberg, A.A. and J.R. Beddington. 1987. Monte-Carlo testing of two methods for estimating growth from length-frequency data with general conditions for their applicability, p. 283-298. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

In the first section of this paper we consider how variability in length at age can arise assuming von Bertalanffy growth. The variance in length at age is written as a function of age and the mean and variance of the growth parameters for the cases where L_{∞} , K and t_0 are random variables. When L_{∞} is a random variable, variance in length at age increases monotonically with age. With K variable, variance in length first increases, then decreases at older ages. The inflection point is at the end of the period of rapid growth. Variation in t_0 produces high variance at the youngest ages which decreases rapidly with age. The results for L_{∞} random are then applied to provide a simple necessary condition for the distribution of length frequencies in adjacent age groups to have two distinct modes. General predictions for the determination of growth parameters from length-frequency data are made.

Monte-Carlo tests of two methods for the determination of growth parameters from length-frequency data were performed. The ELEFAN I procedure appears to be able to adequately estimate K only if the true value of K is known to within about 25%. In this way it is very sensitive to the range of values to be scanned chosen by the investigator.

A mixture distribution method for analyzing length frequencies requires a much larger amount of input information than ELEFAN I. It performs well provided initial cohort strengths (relative), mortality and the form of variance in length at age can be approximated on input. Both ELEFAN I and the mixture method are sensitive to increasing variance in length at age, ELEFAN I more so. ELEFAN I consistently underestimates the true value of K while the mixture model tends to underestimate L_{∞} . The implications of underestimating K for the estimation of total mortality and the calculation of potential yield are discussed.

Introduction

The recent resurgence of interest in the use of length-frequency information for stock assessment has been directed toward the development of techniques which can be applied when more conventional methods based on age data are inappropriate or unfeasible. Age information is notoriously time consuming and costly to obtain (see Mathews, Part I, this vol.). For many tropical and crustacean fisheries no suitable methods for aging animals are available. Because of this, length-based methods such as the ELEFAN programs (Pauly 1982) have rapidly gained acceptance for determining growth, mortality and recruitment in exploited stocks.

Two major problems arise in the use of methods such as ELEFAN or the more statistically formal techniques employing maximum likelihood or related estimators (MacDonald and Pitcher 1979; Schnute and Fournier 1980):

- 1) There has been little testing of the accuracy of the methods and their sensitivity to input parameters;
- 2) There is no guide for the user to judge when a particular method is appropriate for the data in hand.

Testing accuracy and sensitivity of a proposed method is an essential step before its use can be widely advocated. This requires applying the model to data whose characteristics are known. Simulation of length-frequency data and Monte Carlo tests of the ability of the method to describe the underlying structure of the simulated data can indicate under what conditions a method will or will not perform acceptably. It is also relevant to consider general conditions for obtaining useful estimates from length-frequency data.

In the next section we consider how variability in length-at-age can arise assuming von Bertalanffy growth. The implications of these results for the analysis of length-frequency data are discussed. Then, Monte Carlo sensitivity tests of the ELEFAN I (Pauly and David 1981) and a mixture distribution method for determining von Bertalanffy growth parameters from length-frequency data (MacDonald and Pitcher 1979; Schnute and Fournier 1980; Sparre, Part I, this vol.) are described. The sensitivity of the two methods to data quality and input parameters is investigated. The contributions of Hampton and Majkowski (Part I, this vol.) should be consulted for a parallel investigation of the properties of ELEFAN I.

Sources of Variability in von Bertalanffy Growth

The von Bertalanffy growth equation is the most commonly applied growth model for fish populations. In its simplest form, termed the special von Bertalanffy equation by Pauly (1982), we write the length at time t as

$$l_t = L_\infty (1 - e^{-K(t - t_0)}) \quad \dots 1)$$

where L_∞ is the asymptotic length attainable, K is a growth constant with units of 1/time which determines the shape of the curve, and t_0 is the time at which the fish had length zero (the time origin of the curve). Equation (1) is most often used to describe the mean growth of a population. However, we may also consider that individuals in the population are growing according to (1), such that variability in length-at-age of fish in the population results from differences between individuals in their values of L_∞ , K and t_0 . Sainsbury (1980) has given the mean and variance in length-at-age when L_∞ is a normal random variable and when K is a gamma random variable. It is useful to consider the effects of variations in L_∞ , K and t_0 separately.

For L_∞ , distributed as $N(L_m, S^2_L)$ with K and t_0 constant, mean length-at-age T is given by

$$E[l_t|T] = L_m (1 - e^{-K(T - t_0)})$$

where E denotes expected value. The variance in length-at-age is

$$V[l_t|T] = S^2_L (1 - e^{-K(T - t_0)})^2 \quad \dots 2)$$

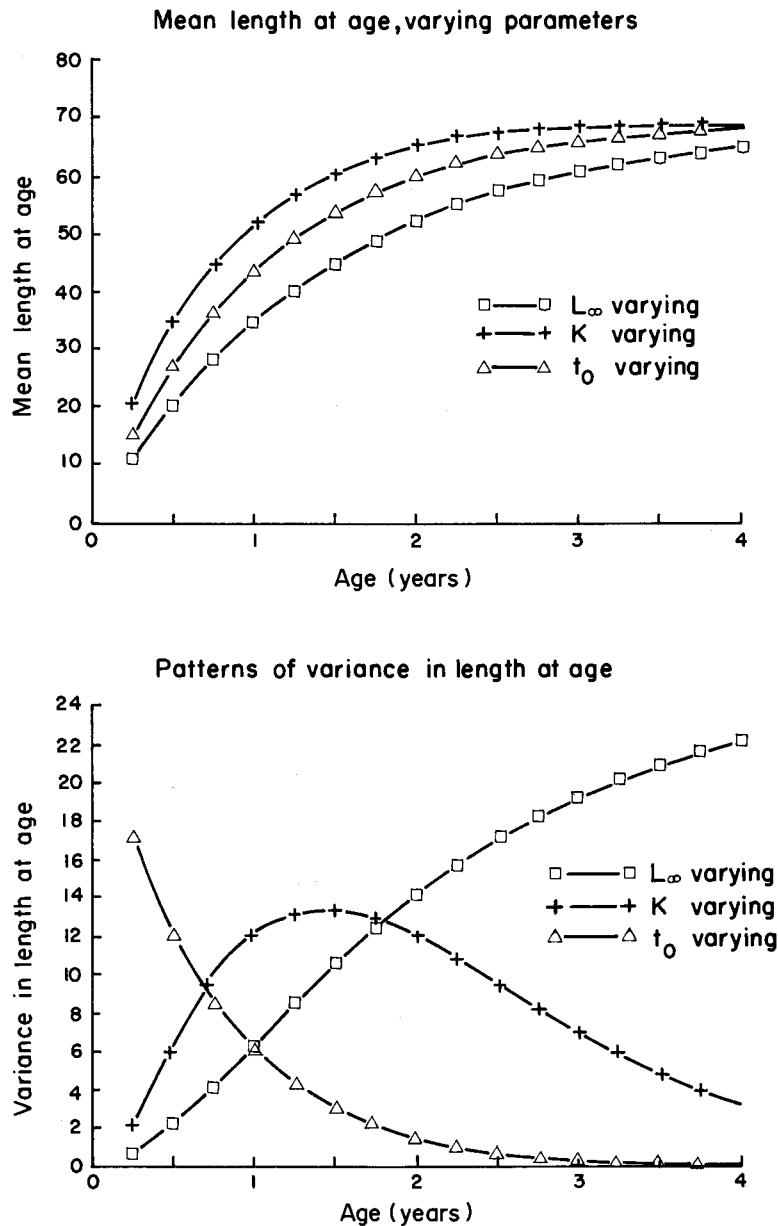


Fig. 1. Patterns of variation in length-at-age induced by allowing von Bertalanffy growth parameters to be random variables. For this example $L_m = 70$, $K_m = .7$, $t_{om} = 0$.

Variance in length-at-age will increase monotonically with age to a limiting value of S^2_L when L_∞ is the only source of variation (Fig. 1). The coefficient of variation in length will be constant. For L_∞ random, the variance in length-at-age increases with age and length. Variation in length-at-age in nature is usually thought to increase with age (Schnute and Fournier 1980; Sparre, Part I, this vol.; Jones, Part I, this vol.), which may indicate that variation in L_∞ between individuals is most important. However, increased error in age determination for older fish (i.e., the tendency for fish of the same size to be assigned the same age) could produce a similar effect.

When L_∞ and t_o are constant and K has a gamma distribution with parameters (A, B) where $A = k^2 m / S^2_k$ and $B = S^2_k / k_m$

$$E[1_t|T] = L_\infty (1 - (1 + B(T - t_o))^{-A})$$

and

$$V[1_t|T] = L_\infty^2 [(1 + 2B(T - t_o))^{-A} - (1 + B(T - t_o))^{-2A}]$$

The gamma distribution was used because it is always positive and provides a flexible two parameter distribution for K (Sainsbury 1980). In this case, variance in length-at-age first increases and then decreases (Fig. 1). The point of inflection in variance is at the end of the fast growth period (De-Angelis and Mattice 1979). The maximum variance in length-at-age under variable K is given by

$$(1 + BT)^{-(2A+1)} - (1 + 2BT)^{-(A+1)} = 0$$

which can be solved numerically to give the age at which the variance in length-at-age starts to decrease. Below are three examples which are indicative of the general pattern.

K_m	S^2_K	Age of maximum variance in length-at-age
1.0	0.01	1.0
0.7	0.01	1.42
0.4	0.01	2.58

Finally, when K and L_∞ are constant but t_o is $N(t_{om}, S^2_t)$

$$E[l_t|T] = L_\infty (1 - \exp \{ K(T - t_{om} - KS^2_t/2) \})$$

and

$$V[l_t|T] = L_\infty^2 [\exp(-2K(T - t_{om} - KS^2_t/2)) - \exp(-K(2T - t_{om} - KS^2_t/2))].$$

Here, the variance in length-at-age is high for the younger age groups and then decreases rapidly with age (Fig. 1).

The distribution of length-at-age at a given age will be normally distributed if variable L_∞ is the major source of variation. To detect two separate age groups or cohorts it is necessary that the combined distribution of adjacent length groups be bimodal. Behboodian (1970) has given a simple sufficient condition for unimodality of a mixture of two normal distributions as

$$|U_1 - U_2| < 2 \min(S_1, S_2)$$

where U_1 and U_2 are the means and S_1 and S_2 are the standard deviations of the two distributions. Substituting the results given above for random L_∞ and reversing the inequality we obtain

$$[L_m \exp -K(t - t_o)](e^{-k} - 1) > 2 S^2_L (1 - \exp(-K(t - t_o)))^2 \quad \dots 3)$$

to give a necessary condition for bimodality for the mixture of the frequencies of two adjacent cohorts.

We can determine the point at which adjacent cohorts will have distinct modes by equating the two sides of the above inequality and plotting the coefficient of variation in L_∞ against age. This has been done for a set of values of L_∞ and K characteristic of several widely exploited species using the values given in Beddington and Cooke (1983). The pattern for all the species looks similar (Figs. 2 a-f). Over a wide range of variation in L_∞ only the first few age groups are distinguishable. Only at a very low coefficient of variation in L_∞ do adjacent older age groups show bimodality.

The question is, how much do the fish have to grow for the age groups to be detectably different in the histogram? This condition for bimodality assumes that the underlying population has been well sampled and that the distribution of length-at-age is normally distributed. The condition as presented here is a necessary condition for bimodality and is conservative for separating modes in length-frequency distributions. The condition is not sufficient for bimodality. In other words, sampling variation or other noise in the data could still obscure the bimodal pattern even if the condition was fulfilled. On the other hand, the ability of a method to follow modes in the length-frequency distribution may be improved in a sequence of samples through time when, occasional,

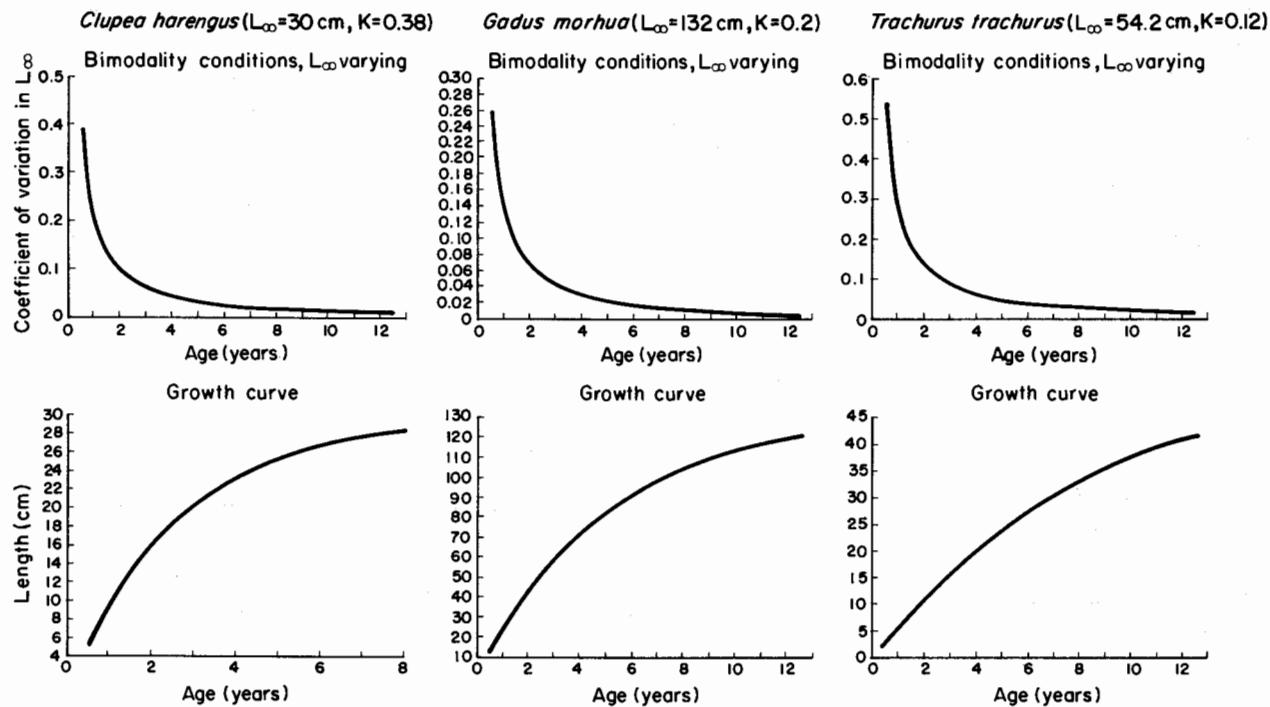


Fig. 2a. Bimodality conditions for *Clupea harengus*, *Gadus morhua* and *Trachurus trachurus*, using the growth parameters given in Beddington and Cooke (1983; see also Table 1). The coefficient of variation in L_∞ is plotted on age. The solid line indicates the locus for bimodality versus unimodality for mixtures of adjacent age groups. Values of the C.V. of L_∞ falling below the curve give bimodality for the appropriate age groups.

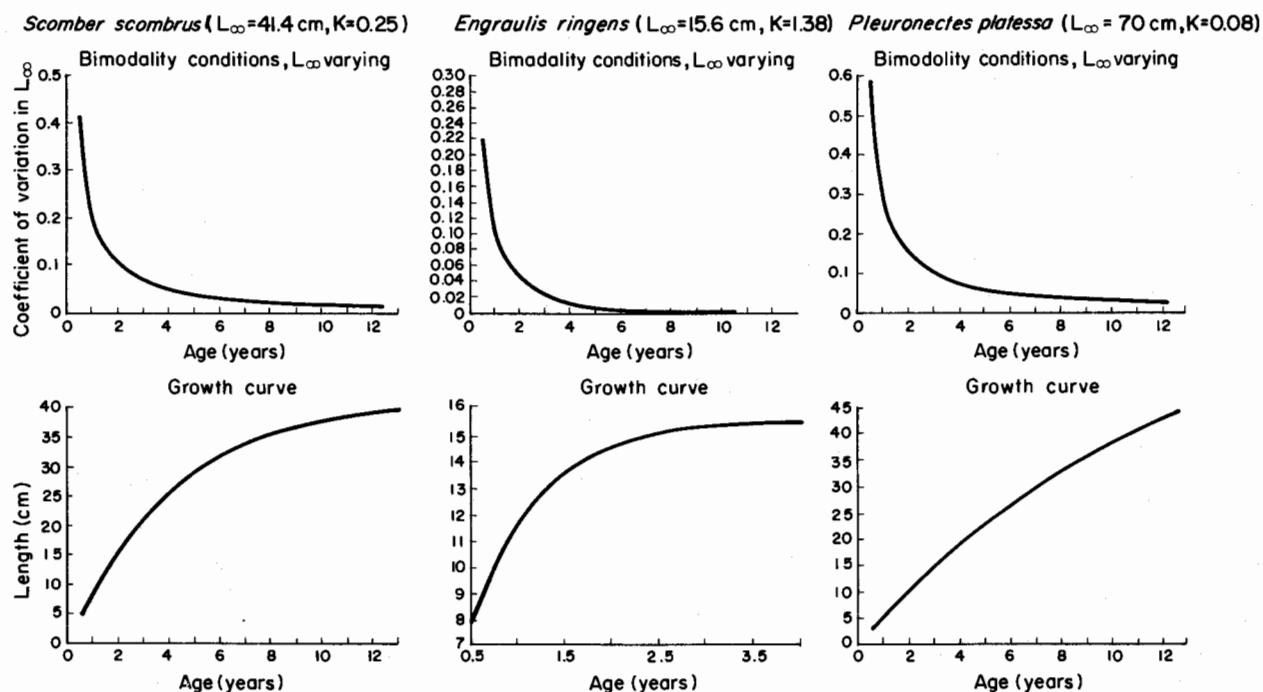


Fig. 2b. Bimodality condition for *Scomber scombrus*, *Pleuronectes platessa* and *Engraulis ringens*, using the growth parameters given in Beddington and Cooke (1983; see also Table 1 for *S. scombrus* and *P. platessa*). The coefficient of variation in L_∞ is plotted on age. The solid line indicates the locus for bimodality versus unimodality for mixtures of adjacent age groups. Values of the C.V. of L_∞ falling below the curve give bimodality for the appropriate age groups.

very large year-classes occur. But methods for length-frequency distribution analysis necessarily operate on the assumption that the means of adjacent age groups are different (i.e., the fish are growing) and a guide to how much the separation needs to be for a particular species is required. There is little information concerning the robustness of this condition to departures from normality although a simple plotting exercise for a range of parameter values when the distributions are not normal indicates that it is in fact quite robust.

The distribution of length-at-age when K or t_0 varies is strongly skewed to the left. Indications of departures from normality for certain species may be a result of the importance of these other sources of variation.

Predictions

Some simple predictions concerning the sensitivity of methods for analyzing length-frequency data to either input parameters or characteristics of the data can be made. The expressions for variance in length-at-age when the parameters of the von Bertalanffy curve are random variables and the conditions for bimodality of adjacent age groups should be used as a guide by the investigator in interpreting length-frequency data. Obviously, one does not have estimates of the variances of L_∞ , K or t_0 between individuals for a population in nature. However, given an estimate of the mean parameter value we must ask how variations in length arise to produce a particular pattern in the length-frequency histogram.

It is possible to calculate the coefficient of variation between years or between populations for some species by examining published estimates of growth parameters (Table 1). The coefficient of variation in L_∞ between years seems to be quite small, but these estimates are made by taking the mean of the mean size at age in the oldest age group, or by averaging several estimates of L_∞ . We expect the resulting coefficients of variation to be smaller than the true value for the variation among individuals in a cohort. If these values are compared with the corresponding figures for the bimodality condition (Fig. 2) it appears that for several of the stocks, at least three or four year-classes will be distinguishable as separate modes in the length-frequency distribution. This is an encouraging result for the analysis of length-frequency data even though the method assumes the population is well sampled and L_∞ is the only source of variation in length-at-age.

Table 1. Rough estimation of the coefficient of variation in L_∞ for several exploited stocks. Method refers to means taken between populations (1) or within a population between years (2). N is the number of years or populations used in the calculation. C.V. is the coefficient of variation.

Species	Method	Mean	S.D.	C.V.	N	Reference
<i>Clupea harengus</i>	1	32.0	5.2	0.16	18	Anthony and Waring (1980)
<i>Clupea harengus</i>	1	28.5	5.98	0.21	5	Pauly (1980)
<i>Clupea harengus</i>	2	31.3	0.57	0.009	19	Molloy (1984)
<i>Clupea harengus</i>	2	30.9	0.62	0.020	20	Hubold (1978)
<i>Clupea harengus</i>	2	36.2	0.64	0.008	18	Jakobsson and Halldorsson (1984)
<i>Gadus morhua</i>	1	110.4	19.83	0.18	9	Pauly (1980)
<i>Trachurus trachurus</i>	2	35.8	0.96	0.027	3	Borges (1984)
<i>Scomber scombrus</i>	1	41.1	1.81	0.044	4	Anderson and Paciorkowski (1980)
<i>Pleuronectes platessa</i>	2	58.8	2.74	0.068	4	Bannister (1978)
<i>Pleuronectes platessa</i>	2	50.6	2.95	0.058	17	Rijnsdorp (1984)
<i>Solea solea</i>	2	39.6	3.52	0.089	15	de Veen (1978)

In general, it will be difficult to distinguish modes in the length-frequency histogram for more than the first few age groups unless there is very little variation in L_∞ between individuals. This means that methods which try to estimate mean growth parameters from length-frequency data will necessarily rely on the movement of the modes of these groups. There will be little information

available for growth at older ages and this may influence the estimation of L_∞ and K. However, on the positive side, it appears that the first age group is distinct under a wide range of variation in L_∞ which gives some hope for the estimation of early growth at least, in most stocks.

When K is large, it will be difficult to distinguish the variance due to t_0 varying from that due to varying K. In other words, for rapidly growing fish, it will be difficult to separate variations in growth from those in recruitment times. Estimation of both K and the pattern of recruitment such as attempted by several length-frequency analysis methods may not be possible or should be interpreted cautiously in these cases.

Under varying recruitment timing, older ages may be more distinct, while variable L_∞ will have the opposite effect. The effects of varying growth rates are more complicated because variance in length-at-age is not a monotonically increasing or decreasing function of age.

Simulation Model

Monte Carlo simulations of length-frequency data from a hypothetical fish population were performed to assess the accuracy and sensitivity of the ELEFAN I and mixture distribution methods under a variety of conditions. The simulation uses a modification of Breen and Fournier's (1984) program. The length-frequency data obtained from a hypothetical fish population is assumed to be a finite mixture of normal distributions, each component of which is centered around a mean size-at-age determined by an input von Bertalanffy growth curve plus a normal random error term. Although any growth curve can be used, here we used the simple form of the von Bertalanffy curve and concentrated on estimating L_∞ and K. The standard deviation of length-at-age for a given age is a linear function of age.

We assume a constant mortality rate Z. Recruitment varies through the year as a gamma function, so that the shape of the recruitment curve can be altered from knife-edge to near normally distributed (Fig. 3). The size of the recruiting year-class is a normal random variate centered on a recruitment factor, R. Recruitment is distributed through the year by multiplying the total recruitment for the year by a gamma variate defined by

$$p(x) = \frac{x^q - 1 e^{-x}}{G(x)}$$

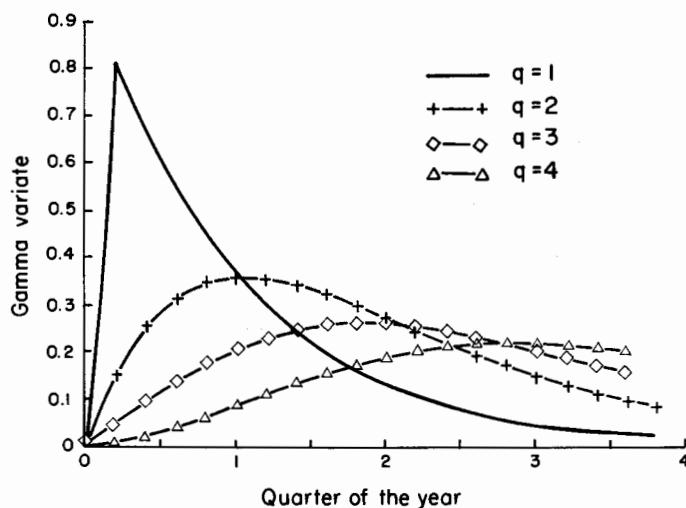


Fig. 3. Shape of the gamma distribution for several values of the parameter q (see text for formulation). X is the quarter of the year.

where G denotes the gamma function and q controls the shape of the distribution and x is the period of the year in which the sample is taken.

Fish are sampled from a population with the input characteristics described above. Each fish in a sample is first assigned to an age class in proportion to the underlying population age structure. It is then assigned a length at the given age, and tabulated in the appropriated length class in the frequency distribution. The input sampling parameters for the simulation are: sample size, the number of age groups, the number of samples taken and their spacing in time. Here, the sample size was 1,000 fish from 0.25 to 4 years old. Four samples were taken a year.

The von Bertalanffy parameters L_{∞} , K and t_0 were input along with the parameters governing the change in variance with size, A the intercept and B the slope. The variance of the means around the von Bertalanffy curve and recruitment variance were fixed for all runs. Finally, the total mortality rate Z was input.

The program outputs four (or a specified number of) sets of length frequencies for input to an analysis program. Two samples are shown in Figs. 4 and 5.

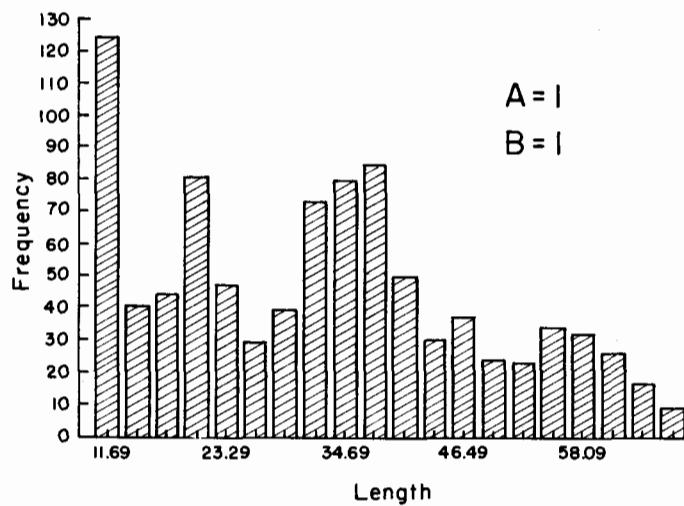


Fig. 4. Example of length-frequency histogram generated by the program with low values of the parameters controlling variance in length-at-age. R was 50, $q = .5$, $k = .7$, $L_{\infty} = 70$.

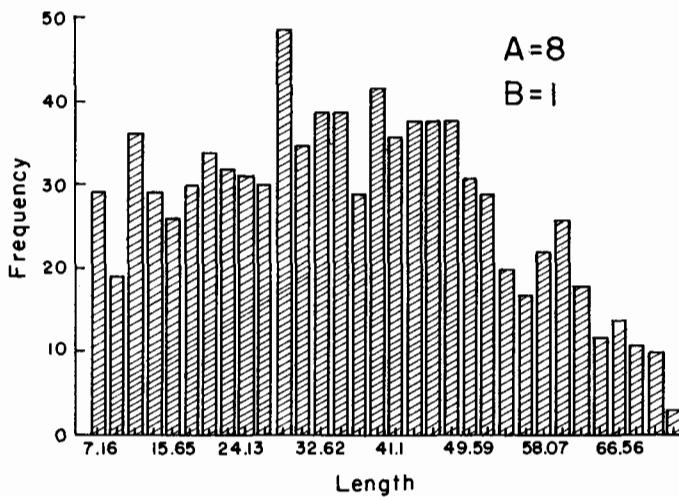


Fig. 5. Example of length-frequency histogram generated by the program with high values for the parameters controlling variance in length-at-age. Other parameter values were as in Fig. 4.

ELEFAN I Simulation Tests

The ELEFAN I method has been well documented elsewhere and we will not repeat a lengthy description here. The program used was a slight modification of Sparre's version (Sparre, pers. comm.) written in Fortran 77. Basically, the program seeks to score peaks and troughs in a set of length-frequency histograms as positive and negative, respectively. The sum of the scores is defined as the Available Sum of Peaks (ASP). Trial von Bertalanffy growth curves are projected through the data and accumulate points according to the numbers of peaks (positive) and troughs (negative) the curves pass through. This gives the Explained Sum of Peaks (ESP). The growth parameters with the highest ESP within a range considered reasonably by the investigator are selected as the best fit to the data.

An additional criterion is used for the estimation of L_∞ , which must be greater than or equal to the largest size group in the data. In most of our trials, the L_∞ value chosen on the basis of the scoring function was smaller than the largest size group in the data, and so the ELEFAN I estimate of L_∞ was the largest size group. In other words, ELEFAN I was actually only estimating K by means of the scoring function. This is not to imply that choosing the largest size group in the data is unreasonable as an estimate of L_∞ for a well sampled stock. However, we have focused on the estimation of K for this reason. The simulation was not designed to include sampling variability, so the estimate of L_∞ by selecting the largest size group is usually close to the true L_∞ .

ELEFAN I is a curve fitting procedure but requires as input a range of values to be scanned for the best estimate of K (and L_∞ given the proviso above). For the program to estimate the true parameter, the range must encompass the true K. An assessment of the importance of the critical choice of range is given below. For the simulations presented here, with a K value of 0.7, a range of 0.2 to 1.1 was scanned by the program in steps of 0.05.

The tests considered variation in 3 parameters: A, the intercept for the relationship between standard deviation of length-at-age and age; R, the recruitment factor and q the parameter of the gamma distribution for recruitment times. Each parameter set was run 10 times and the mean squared error (MSE) and bias computed as

$$\text{MSE} = (\text{estimated} - \text{true})^2 / 10$$

and

$$\text{Bias} = (\text{estimated} - \text{true}) / 10$$

for the 10 runs. Although sample size is small, repeated runs of 10 indicated change in the results was no greater than 10%.

Data with low variance in length-at-age and knife edge recruitment should provide a realistic baseline for the ELEFAN I program since peaks and troughs will be well defined. Note this is not perfect, but good data. It still contains a stochastic component. Tables 2 and 3 indicate that both MSE and bias are still substantial for this parameter combination. The coefficient of variation is about 70% with a negative bias of 30%. As variance in length-at-age increases, both MSE and bias also increase. Note that ELEFAN I consistently underestimates the true value of K. This was a constant feature of the trials.

As recruitment is spread out over the year, simulated by increasing q with low variance in length-at-age (Table 2), the effects on the ability of ELEFAN to accurately estimate K are similar to the effects of increasing variance in length-at-age controlled by the A parameter.

The effects of variable recruitment (Table 3) are very similar to those for variance in length-at-age. As R increases, the bias in 10 runs surpasses 50% of the K value. A large recruiting year class will have a very high contribution to ASP. Its effect will be that of an outlier in the data, having a disproportionately large influence on the ESP.

In each run, the maximum size of the bias is determined by the input range of parameter values to be scanned. With K of 0.7 and a lower limit of the scanning range of 0.2 the bias cannot be greater than -0.5 and MSE of 25. If the range is narrower or wider, the effect is obvious (Table 4).

Table 2. Simulation tests of ELEFAN I: response to variance in length-at-age and varying recruitment times. A denotes the intercept for a linear relationship between variance in length-at-age and age and q is the parameter of the standard gamma distribution controlling the distribution of recruitment through the year. The input parameters were K = 0.7, R = 50, B = 1. The first number in each pair is the MSE of the estimate of K for 10 runs of the simulation. The second entry in each pair is the bias in estimating K in the 10 runs.

		A				
		0.01	0.5	1.0	2.0	3.0
q						
0.5	0.281	0.207	0.250	0.282	0.293	
	-0.205	-0.275	-0.380	-0.285	-0.415	
1.0	0.200	0.228	0.241			
	-0.180	-0.295	-0.490			
1.5	0.191		0.257			
	-0.370		-0.300			
2.0	0.207			0.213	0.236	
	-0.415			-0.435	-0.485	

Table 3. Simulation tests of ELEFAN I: response to increased recruitment variability. R is the recruitment multiplier. q is the parameter of the gamma distribution of recruitment times. The input parameters were K = 0.7, A = 1, B = 1. The first entry in each pair of numbers is the MSE for the estimates of K in 10 runs of the simulation. The second number in the pair is the mean bias in the estimates of K for the 10 runs.

		R			
		1	25	50	75
q					
0.5	0.194	0.227	0.220	0.205	
	-0.175	-0.230	-0.260	-0.340	
1.0	0.208	0.173	0.241		
	-0.110	-0.360	-0.490		
1.5	0.229		0.222		
	-0.350		-0.370		
2.0	0.176			0.217	
	-0.365			-0.325	

We examined the results of several trials of ELEFAN I in terms of the ESP tables for each parameter combination. In fact, there is often a peak in the ESP surface at or close to the true value of K, even when the variance in length-at-age is large. However, this peak is usually a local maximum only. For example, in many runs, if the range of values given to ELEFAN I was between 0.55 and 1.0 say, the program would find the K value close to the true 0.7. However, trial values below 0.55 have higher ESP values and there may be other local maxima. In other words, the program usually estimates the growth parameter K very well when it is already known to within about 20-25% of its true value.

Table 4. Simulation tests of ELEFAN I: response at different K values and increasing variance at age. The first entry in each pair is the MSE for estimating K in 10 runs. The second entry is the bias in the estimates of K in 10 runs. The input parameters were R = 50, q = 1, B = 2.

		K		
		0.4	0.7	1.0
A	0.01	0.185 -0.25	0.296 -0.280	0.270 -0.450
	0.5	0.031 -0.170	0.194 -0.425	
1.0		0.035 -0.130	0.241 -0.490	
	2.0	0.073 -0.100		
3.0		0.123 -0.035		

With respect to L_{∞} , the ESP/ASP criterion usually underestimates the true value and so the size of the largest length group is chosen as the estimate of L_{∞} . Sparre (pers. comm.) has noted that there may be difficulties with choosing L_{∞} as the largest age because several of the growth curves projected through the data will pass through this length class and overinflate the ESP value such that the ratio ESP/ASP is greater than 1. Since this ratio is supposed to be analogous to a coefficient of determination or lack of fit statistic a value greater than 1 has no meaning and the estimates should be rejected. Although this is easy to detect in the resultant estimation run, the remedy for the problem has only recently been incorporated in the ELEFAN I program (see Appendix I and Pauly, Part I, this vol.).

Mean squared error and bias do not show the whole picture. Individual runs sometimes estimated K quite accurately. In our simulation we have assumed that the results were obtained by a naive user. That is, it is not possible to analyze each run individually as a practitioner would do for his or her own data. Subjective decisions by the investigator cannot easily be incorporated into the analysis of each data set in the simulation.

For a set of runs the ELEFAN I method does not appear to reliably estimate growth parameters from length-frequency data unless the input parameter range is narrow and the true value lies within this range. The explained sum of peaks surface appears to have multiple maxima, so that, while a maximum often occurs near the correct parameter value, other values could be selected with no clear means of distinguishing between multiple solutions. ELEFAN I is particularly sensitive to increased variance in length-at-age or increased variance in recruitment times. With respect to the objectivity of the method, it is sensitive to the input parameter range to be scanned. Of course, the same input parameters and data should give different investigators the same answer.

Recent modifications to the ELEFAN I method improve its performance. Runs of the model with the modified method, which corrects for the projection of growth curves through a peak more than once, remove the tendency for ELEFAN to drift toward low K values. The results still indicate some bias in the parameter estimates, however (Appendix I).

Mixture Distribution Method Tests

One way of statistically treating the problem of estimating the growth parameters of a fish population from length-frequency data is based on the idea that a length-frequency sample can be

conceptually thought of as a mixture of probability distributions for length-at-age. That is, at each age we hypothesize that the lengths of the fish in that age class (or cohort) are normally distributed with mean given by the von Bertalanffy or similar function and the variance by a specified form (MacDonald and Pitcher 1979; Schnute and Fournier 1980). To follow the progression of a cohort through time, samples are linked by the growth curve, mortality rate and initial cohort strengths to calculate where each cohort should fall in the next sample and how large its contribution to the overall mixture distribution should be. The growth curve gives the means of the distributions and the variances are often assumed to be some function of the means. The cohort strengths and mortality rate give the mixing proportions for the sample. The expected frequencies computed from a set of input parameter values can then be compared with the observed frequencies via a statistical criterion such as chi squared or a (quite similar) maximum likelihood estimator.

One difficulty with this method is that the number of parameters to optimize over becomes very large if cohort strengths, mortalities and growth parameters are included. Parameters in the set can be fixed to reduce the dimensionality in the problem.

A second constraint on the use of the method is that the number of cohorts in each sample must be specified as input. This information will not be available for many fish stocks for which growth studies have not already been done. MacDonald and Pitcher (1979) note that older age groups can be grouped together without adversely affecting the results. Schnute and Fournier (1980) suggest some goodness of fit tests similar to a likelihood ratio statistic, which can be used to choose between solutions using different numbers of cohorts. However, the need to specify the number of cohorts is a limitation of this method.

In the tests presented below we have focused on the estimation of L_∞ and K for comparison with the results for ELEFAN I. To this end the cohort strengths, parameters describing the variance in length-at-age and mortalities were fixed and optimization using a chi squared criterion was performed over the two parameters L_∞ and K. Some care is required in choosing the fixed values of the cohort strengths and mortality parameters. In our trials cohort strengths which conformed to the pattern of mortality (negative exponential) were input, but they were not the exact cohort strengths from the simulation program. In general, the results given below hold as long as the pattern is correct and the magnitudes are similar to the true cohort strengths and mortality. A modification of Sparre's (Part I, this vol.) program written in Fortran 5 was used with a Gauss-Newton optimization routine from the NAG subroutine library.

The mixture method was most sensitive to variance in length-at-age (Tables 5 and 6). As the intercept for variance in length-at-age (A) increased, the quality of the estimates of L_∞ and K deteriorated. Note that the value of variance in length-at-age at which the mixture method breaks down is well above that for ELEFAN I. By comparison to ELEFAN I, the mixture method is insensitive to variation in recruitment times. However, because the cohort strengths and mortalities are given for these trials, the mixture model contains much more information as input than the ELEFAN I method. Also, the method seemed to perform better at high values of K than at low values (Table 8).

The need to input the number of cohorts in each sample seems a major limiting feature of the mixture method. However, simulations when the input guess at the number of cohorts was incorrect (Table 7) indicate that this factor may not be as crucial as usually assumed. Underestimating the number of cohorts seems to have a more drastic effect than overestimating.

The results of the mixture method were independent of the starting values for the estimation procedure as long as they were in a permissible range (K non-negative). For example, starting values of K = 0.2, L_∞ = 40 when the true values were (0.7, 70) performed as well as starting values of (0.6, 60).

Discussion

The bimodality condition given in (3) provides a guide for determining when either ELEFAN I or a mixture distribution method can accurately estimate growth parameters from length-frequency data. For example, consider the two histograms shown in Figs. 4 and 5. If we assume that the variance in length-at-age results from variation in L_∞ only, we can solve (2) for S^2_L given knowledge

Table 5. Simulation testing of a mixture distribution method of length-frequency analysis: response to variance in length-at-age and variation in recruitment times. Symbols and format as in Table 1. The first entry is the MSE over 10 runs in the estimate of L_∞ . The second entry is the mean bias over 10 runs for the estimates of L_∞ . The input parameters were $L_\infty = 70$, $K = 0.7$, $B = 2$, $R = 50$.

		A				
		4	6	8	10	12
q		37.86	31.38	62.16	95.08	214.48
0.5		-5.79	-5.52	-7.80	-9.65	-14.54
1.0		26.81	37.75			
		-4.46	-6.06			
1.5		17.28	33.16			
		-3.58	-5.70			
2.0		21.90		45.96		
		-4.30		-6.65		
3.0		20.75			95.06	
		-4.00			-9.68	

Table 6. Simulation testing of a mixture distribution method for length-frequency analysis: response to variance in length-at-age and varying recruitment times in the estimation of K. See Table 5 for input parameters and explanation.

		A				
		4	6	8	10	12
q		0.005	0.026	0.105	0.181	0.479
0.5		-0.013	0.136	0.316	0.421	0.687
1.0		0.008	0.029			
		-0.064	0.152			
1.5		0.013	0.138			
		-0.089	0.025			
2.0		0.009		0.068		
		-0.076		0.252		
3.0		0.007			0.179	
		-0.073			0.416	

of $V[l_t]$. Using (3) we predict that at least 3 modes should be evident in Fig. 4 but only 1 in Fig. 5. In this sample case, it is obvious that there are clear modes in the first histogram. For many data sets it will not be so obvious and the condition can indicate whether the underlying structure of the length-frequency histogram is such that analysis for the growth parameters can be fruitful.

Both ELEFAN I and the mixture method are limited in applicability. ELEFAN I seems to require very good starting estimates of K to find an acceptable value. The mixture method requires information on the number of age groups although perhaps less critically than previously thought. In addition, the relative initial sizes of the cohorts and mortality must be approximated.

Table 7. Effect of mis-estimating the number of cohorts in the sample on estimates of K by the mixture method. The input parameter were A = 4, B = 1, R = 50 and q = 1. The true K equal to 0.7. The true number of cohorts was 4.

Input number of cohorts		2	3	4	5	6
MSE		0.119	0.036	0.008	0.004	0.007
Bias		-0.124	-0.151	-0.060	-0.023	0.022

Table 8. Simulation testing of mixture distribution method for length-frequency analysis: performance at different values of K. The input parameters were A = 4, B = 1, R = 50, q = 1, $L_{\infty} = 70$. The first entry in each set is the MSE of the estimates of L_{∞} for 10 runs. The second entry is the bias for L_{∞} , third is the MSE in the estimate of K and fourth is the bias in the K estimates.

K		0.2	0.4	0.7	1.0
MSE L_{∞}		824.54	103.23	26.81	82.91
Bias L_{∞}		-28.47	-8.27	-4.46	-9.06
MSE K		0.118	0.023	0.008	0.012
Bias K		0.330	0.110	-0.064	-0.086

Both methods are most sensitive to variance in length-at-age, ELEFAN I more so than the mixture method. The general reason for sensitivity to variance in l_t is explicable in terms of the bimodality condition (3). Distinct modes in the length-frequency histogram result from two factors, small variance in length-at-age and wide separation of the mean lengths of adjacent ages. Therefore, methods which try to separate adjacent ages (peaks and troughs in ELEFAN I terms) will be sensitive to variance in length-at-age and to the value of K. This is the pattern of the simulation results. The mixture method is less sensitive than ELEFAN I because the pattern of variance in length-at-age is explicitly taken into account as input.

Why is ELEFAN I more sensitive to recruitment parameters than the mixture method? This is in large part due to the additional input information for the mixture method. But also, ELEFAN I will give greater weight to well-structured, large peaks and so depend heavily on the information from the recruiting year-class. The mixture method will give weight to age classes depending on the departure from initial cohort size. If, as done here, the cohort sizes are governed by the exponential mortality rate the influence of the recruiting year-class will not be disproportionate. In addition, the mixture model responds less to recruitment spread through the year because the number of cohorts in the sample is fixed on input.

Overall, the mixture method seems more accurate and robust over a range of data qualities, but the information required for its use is substantially greater. ELEFAN I requires less input, but cannot be counted upon to produce accurate estimates without substantial subjective input by the user. Note that this is not the same as producing estimates that appear reasonable when the true values are unknown. ELEFAN I has the potential to produce accurate estimates because the scoring function often (but not always) contains a local maximum near the true value. Recent modifications to ELEFAN I may improve the ability of the program to choose between multiple maxima (Appendix I and see Morgan, Part I, this vol.). It is interesting to note that the mixture model as implemented here was the more economical in terms of computer time.

Implications for Stock Assessment

The estimation of the parameters of growth is not the end product of investigations of the length-frequency data from fish. Usually, such data are used to estimate other demographic parameters which together determine the potential yield and dynamics of the stock. For example, a common method for estimating the total mortality of the stock is using the age distribution of the sample. Typically the estimation proceeds in two stages. First, using the estimated growth curve, the length sample is converted to an age structure. Variations on such techniques, e.g., Clark (1981) use essentially the same procedure. The mortality rate is then estimated from the age distribution of the catch (Chapman and Robson 1960).

Applying this method to typical data generated from the ELEFAN I program yields the results indicated in Table 9. The MSE and bias indicate that there is both substantial uncertainty and a strong tendency to underestimate the level of mortality. This would imply, for stock assessment, a considerable problem. For example, if VPA is used, terminal F values would be underestimated and hence starting population sizes overestimated. Similarly, yield per recruit calculations will be biased and estimates of the potential yield derived from simple Beverton and Holt or Schaefer type models will be in error.

Table 9. Simulation testing of ELEFAN I: MSE and bias in the estimation of total mortality Z. These simulation runs are the same as from the top line of Table 2.

		A		
		1.0	2.0	3.0
		Z		
0.7	0.176	0.181	0.222	
	-.290	-.268	-.369	
1.0	0.327	0.357	0.488	
	-.400	-.460	-.433	

Acknowledgements

We would like to thank P. Sparre for generously providing copies of his programs for length-frequency analysis and D. Sampson for his helpful discussion and comments on the manuscript.

Appendix I

An important modification to the ELEFAN program was proposed by D. Pauly at the February 1985 conference. He suggested that problems of bias in the estimation of K may be in part corrected by ensuring that as growth curves are traced through the data to calculate the ESP, peaks are flagged when the curve hits them. Then, additional curves striking any part of a flagged peak are not counted in the ESP.

The program we used was modified in this fashion and a small number of additional runs (sets of 10) made after the conference. The results for the estimation of K are presented below. For these runs, K was 0.7, $L_{\infty} = 70$, R = 50 and B = 1. The variance in length-at-age as controlled by the A parameter was varied. The table below should be compared with Table 2. The range in A was wider for these additional runs simply because the program performed better and still gave meaningful (although biased) estimates at A = 7.

A

	.5	1	3	5	7
q	0.045	0.053	0.056	0.024	0.050
1	0.195	0.190	0.225	0.120	0.180

It may be also mentioned that the modification tested here also has the advantage of suppressing "drifting" toward impossibly low values of K, as can occur during the analysis of some data sets (see also Pauly, Part I, this vol.).

References

- Anderson, E.D. and A.J. Paciorkowski. 1980. A review of the north west Atlantic mackerel fishery. Rapp. P.-V. Réun. Cons., Cons. Int. Explor. Mer 177: 175-211.
- Anthony, V.C. and G. Waring. 1980. The assessment and management of the Georges Bank herring fishery. Rapp. P.-V. Réun. Cons., Cons. Int. Explor. Mer 177: 72-192.
- Bannister, R.C.A. 1978. Changes in plaice stocks and plaice fisheries in the North Sea. Rapp. P.-V. Réun. Cons., Cons. Int. Explor. Mer 172: 82-101.
- Beddington, J.R. and J.G. Cooke. 1983. The potential yield from a fishery. FAO Fish. Tech Pap. 242. Rome, 47 p.
- Behboodian, J. 1970. On the modes of a mixture of two normal distributions. Technometrics 12: 131-139.
- Borges, M.F. 1984. Evaluation of the results on horse mackerel, *Trachurus trachurus* of a series of young fish surveys. Int. Counc. Explor. Mer, Council Meeting 1984/H: 26.
- Breen, P.A. and D.A. Fournier. 1984. A user's guide to estimating total mortality rates from length-frequency data with the method of Fournier and Breen. Can. Tech. Rpt. Fish. Aquat. Sci. 1239. 63 p.
- Chapman, D.G. and D.S. Robson. 1960. On the analysis of a catch curve. Biometrics 16: 354-368.
- Clark, W.G. 1981. Restricted least squares estimates of age composition from length composition. Can. J. Fish. Aquat. Sci. 38: 297-307.
- DeAngelis, D.L. and J.S. Mattice. 1979. Implications of partial differential equation cohort models. Math. Biosci. 47: 271-285.
- Hubold, G. 1978. Variations in growth rate and maturity of herring in the northern North Sea in the years 1955-1973. Rapp. P.-V. Réun. Cons., Cons. Int. Explor. Mer 172: 154-163.
- Jakobsson, J. and D. Halldorsson. 1984. Changes in biological parameters in the Icelandic summer spawning herring. Int. Counc. Explor. Mer, Council Meeting 1984/H: 43.
- MacDonald, P.D.M. and 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.
- Molloy, J. 1984. Density dependent growth in Celtic Sea herring. Int. Counc. Explor. Mer, Council Meeting/H: 30.
- Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters and mean environmental temperature in 175 fish stocks. J. Cons., Cons. Int. Explor. Mer 39: 175-192.
- Pauly, D. 1982. Studying single species dynamics in a tropical multispecies context, p. 33-70. In Pauly, D. and G.I. Murphy (eds.) Theory and management of tropical fisheries. ICLARM Conference Proceedings 9, 360 p.
- Pauly, D. and N. David. 1981. ELEFAN I, a BASIC program for the objective extraction of growth parameters from length-frequency data. Meeresforsch. 28(4): 205-211.
- Rijnsdorp, A.D. 1984. On the growth of mature North Sea plaice. Int. Counc. Explor. Mer, Council Meeting 1984/C.M. G: 34.
- Sainsbury, K.J. 1980. Effect of individual variability on the von Bertalanffy growth equation. Can. J. Fish. Aquat. Sci. 37: 241-247.
- Schnute, J. and D. Fournier. 1980. A new approach to length frequency analysis: growth structure. Can. J. Fish. Aquat. Sci. 37: 1337-1351.
- de Veen, J.F. 1978. Changes in North Sea sole stocks. Rapp. P.-V. Réun. Cons. Int. Explor. Mer 172: 124-136.

**A Simple Method for Estimating the
von Bertalanffy Growth Constants for
Determining Length from Age and Age from Length**

M.J. SANDERS
*Marine Resources Service
Fisheries Resources and Environment Division
Fisheries Department
Food and Agriculture Organization of the United Nations
Via delle Terme di Caracalla
00100 Rome, Italy*

Sanders, M.J. 1987. A simple method for estimating the von Bertalanffy growth constants for determining length from age and age from length, p. 299-303. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

A method is presented for estimating the von Bertalanffy growth constants relevant to determining length from age and age from length. Programs based on this method and suitable for use with HP 67/97 calculators have been prepared.

Introduction

The traditional use of the von Bertalanffy growth constants is to estimate mean length from age. When it is necessary to estimate mean age from length, the same values for the constants are usually used (see Gulland 1969; Pauly 1980). It is now believed that the latter procedure is invalid (except when considering the growth of a single fish). When considering a number of fish, the relationship between mean length and age is associated with a lower L_{∞} and higher K, than for the relationship between mean age and length (Jones 1981).

In this paper a simple method is presented for estimating the growth constants relevant to determining length from age and age from length. It is based on the reasoning that the best choice values for the constants will be those providing estimates of the mean length at age (or mean age at length) in closest agreement with the observed data (but see Rosenberg and Pope, Part I, this vol.).

Method

A rearrangement of the von Bertalanffy equation provides the linear relationship:

$$\log_e (1 - \ell_t / L_\infty) = Kt_o - Kt$$

where ℓ_t is the length at age t , and L_∞ , K and t_o are the growth constants. Hence when L_∞ is known or assumed, undertaking a linear (least squares) regression analysis of $\log_e (1 - \ell_t / L_\infty)$ against t provides estimates of K (= slope) and t_o (= x axis intercept). The method was first proposed by von Bertalanffy (1934) for the estimation of K , and by Beverton and Holt (1957) and Ricker (1958) for the estimation of K and t_o .

The amendment proposed here is to assume a preliminary value of L_∞ , undertake the linear regression analysis described above, and then to use the L_∞ , K and t_o combination to obtain estimates of the mean length for each observed age (or the mean age for each observed length). A linear (geometric mean) regression analysis is then undertaken of the estimated against the observed mean lengths (or ages).

Identifying the best choice combination of L_∞ , K and t_o (or L_∞' , K' and t_o') is based on the values obtained for the slope, y-axis intercept and coefficient of determination. The ideal is when the slope is unity, the y-axis intercept is zero and the coefficient of determination is maximized. This procedure for identifying the best choice is preferred to that referred to in Pauly (1984), based on maximizing the coefficient of determination from the relationship of $\log_e (1 - \ell_t / L_\infty)$ against t . The procedure is repeated with a new value for the assumed L_∞ , until the best combination of the constants can be recognized.

Prior to applying the method, it is necessary to arrange the length and age data appropriately. When estimating the growth constants relevant to determining length from age, the input data should be expressed as mean lengths at age. If the constants relevant to estimating age from length are required, however, the input data should be expressed as mean ages at length. These requirements presuppose the availability of data in the form of an age-length key.

HP programs based on this method are given in Morgan and Pauly (Part II, this vol.). The user instructions are the same for both options; option 1 provides estimates of the constants relevant to length at age, and option 2 gives estimates of the constants relevant to age at length.

Application

The method was applied to the length and age data for the Pacific bonito (*Sarda chiliensis*), quoted in Bartoo and Parker (1983) and originating from Campbell and Collins (1975). These data are reproduced as mean lengths at age and mean ages at length in Table 1. Only the mean ages for

Table 1. Length and age data for Pacific bonito (*Sarda chiliensis*).^a

Age groups	Length at age			Length groups	Age at length		
	Mean lengths (cm)	Standard deviations	n		Mean ages (yr)	Standard deviations	n
1	50.98	2.68	424	<58	1.00	0.00	410
2	61.58	2.07	158	58	1.86	0.36	21
3	71.55	2.02	53	60	1.88	0.33	75
4	72.83	1.96	80	62	1.95	0.23	56
5	73.95	1.86	21	64	2.00	0.00	12
				66	2.33	0.52	6
				68	3.00	0.68	14
				70	3.53	0.62	32
				72	3.82	0.65	61
				74	3.92	0.69	36
				76	4.42	0.51	12
				> 76	4.00	0.00	1

^aFrom Campbell and Collins (1975).

(and including) the lengths 58 to 76 cm were used as input data. The mean ages of smaller fish were excluded as being biased, there being no length distribution given for the 0 age category. The mean age for the > 76 cm group was excluded, again on the basis of presumed bias, since there were too few fish in the group.

Results

The best choice combination of the growth constants relevant to estimating lengths from ages, to the nearest whole number value of the asymptotic length, is $L_{\infty} = 75$, $K = 0.81$ and $t_0 = -0.40$. This was determined on the basis of the values for the associated y-axis intercept ($u = 2.4186$) and slope ($v = 0.9675$) being closest to zero and unity, respectively. The maximization of the coefficient of determination ($r^2 = 0.9821$) was taken to be a less useful indicator of the best choice combination in this example.

The best choice combination of the growth constants for estimating ages from lengths to the nearest whole number for the asymptotic length, is $L_{\infty}' = 89$ cm, $K' = 0.29$ and $t_0' = -2.13$, and $L_{\infty}' = 90$ cm, $K' = 0.27$ and $t_0' = -2.25$ yr. Again this is based on the values for the associated y-axis intercept ($u = -0.0707$) and slope ($v = 1.0246$) being closest to zero and unity, respectively. The coefficient of determination ($r^2 = 0.9525$) was also maximized with this combination.

A plot of the two relationships is given in Fig. 1 and the results of the estimations are given in Tables 2 and 3.

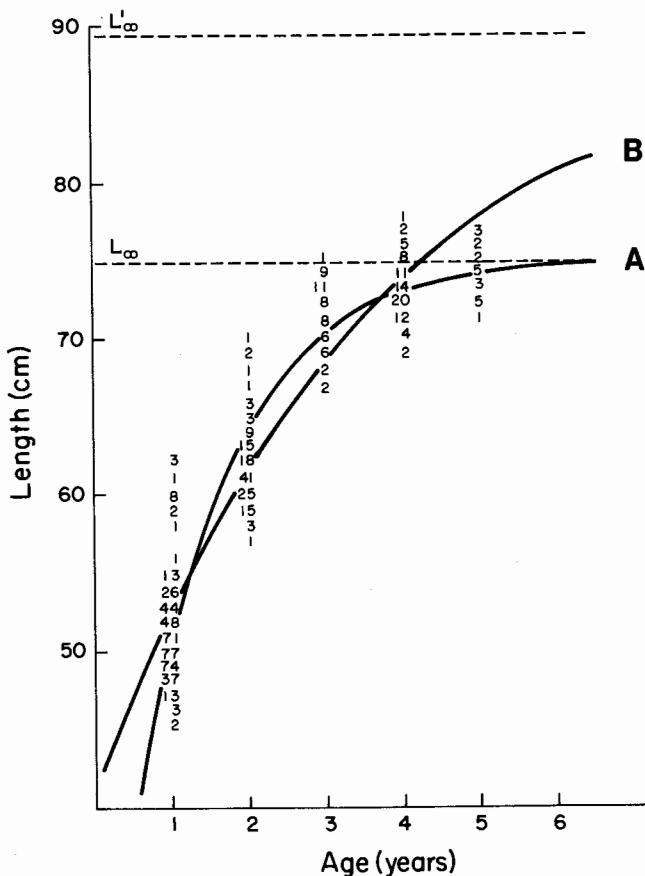


Fig. 1. Plots of (A) mean lengths at age and (B) mean ages at length for Pacific bonito (*Sarda chiliensis*); see Table 1 for source of data. (These plots suggest the possibility of some large fish of ages 1 and 2 having been attributed erroneous ages. Editor's note.)

Table 2. Estimates of the growth constants (L_{∞} , K and t_o), mean lengths at age and regression/correlation parameters.

Observed age groups (years)	Observed mean lengths (cm)	$L_{\infty} = 74.0$	$K = 1.4626$	$t_o = 0.5651$	Estimated lengths (cm) when 75.0	76.0	77.0
1	50.98	34.83			50.86	52.56	53.30
2	61.58	64.93			64.24	63.79	63.45
3	71.55	71.90			70.21	69.64	69.26
4	72.83	73.51			72.86	72.68	72.58
5	73.95	73.89			74.05	74.27	74.47
y-axis intercept, u		-48.0701			2.4186	7.2534	9.0583
slope v		1.6906			0.9675	0.8966	0.8697
coefficient of determination, r^2		0.9091			0.9783	0.9821	0.9793

Table 3. Estimates of the growth constants (L_{∞}' , K' and t_o'), mean ages at length and regression/correlation parameters.

Observed mean ages (years)	Observed length groups (cm)	$L_{\infty}' = 86.0$	$K' = 0.3381$	$t_o' = -1.7432$	Estimated ages (years) when 88.0	89.0	90.0	91.0	95.0
1.86	58	1.58			1.56	1.55	1.54	1.53	1.51
1.88	60	1.80			1.78	1.78	1.78	1.77	1.76
1.95	62	2.03			2.03	2.03	2.03	2.03	2.03
2.00	64	2.29			2.30	2.30	2.30	2.30	2.31
2.33	66	2.57			2.58	2.59	2.59	2.60	2.61
3.00	68	2.88			2.90	2.91	2.91	2.92	2.94
3.53	70	3.23			3.25	3.26	3.26	3.27	3.29
3.82	72	3.63			3.64	3.64	3.65	3.65	3.66
3.92	74	4.08			4.08	4.08	4.08	4.08	4.08
4.42	76	4.62			4.59	4.58	4.57	4.56	4.53
y-axis intercept, u		-0.0719			-0.0708	-0.0707	-0.0707	-0.0709	-0.0721
slope, v		1.0250			1.0247	1.0246	1.0246	1.0247	1.0251
coefficient of determination, r^2		0.9517			0.9524	0.9525	0.9525	0.9524	0.9516

Discussion

The underlying basis of the method (apart from the assumption of von Bertalanffy growth) is that the best choice combinations of the growth constants are those which give estimates of the mean lengths at age (or mean ages at length) in closest agreement with the observed data; this also assumes that the quality of the observed data is sound.

Unfortunately, with respect to determining the growth constants relevant to estimating ages from lengths, the limitations imposed by practical considerations are likely to be substantial. The principal problem is in obtaining mean ages at length having satisfactory precision and minimal bias. To improve precision, sets of length and age data should be collected for as many short time intervals as practical. These should then be separately incorporated within a single age-length key; from these, mean ages can be determined for each length.

The estimated mean ages at length will be dependent on the number of pairs of length and age data in each time interval. If the number in a particular time interval is much larger than for the other intervals, the ages represented in that sample will be over-represented in the means. Minimiz-

ing this bias requires the number in each time interval to be proportional to the stock abundance, or some index of abundance such as catch per unit effort. The less valid alternative would be to have the same numbers in each time interval (which may be achieved indirectly by converting the number in each interval to percentages).

The method as presented relates to the input of data as mean lengths at age (or mean ages at length). It is equally valid when the inputs are the lengths and ages of individual fish, although only for determining the constants relevant to estimating lengths from age. With respect to determining the constants for estimating ages from lengths, undertaking the linear regression of t against $\log_e(1 - l_t/L_\infty)$ gives K' ($= -1/\text{slope}$) and t'_0 ($= y$ axis intercept).

The use of data for individual fish does not avoid the practical problems mentioned earlier in this section. Furthermore, it increases the possibility of encountering lengths greater than L_∞ (or L'_∞), in which case the method fails.

References

- Bartoo, N.W. and K.R. Parker. 1983. Stochastic age-frequency estimation using the von Bertalanffy growth equation. Fish. Bull. (U.S.) 81(1): 91-96.
- Beverton, R.J.H. and S.J. Holt. 1957. On the dynamics of exploited fish populations. Fish. Invest. Ser. II. Vol. 19, 533 p.
- Campbell, G. and R.A. Collins. 1975. The age and growth of the Pacific bonito *Sarda chiliensis*, in the eastern North Pacific. Calif. Fish Game 61: 181-200.
- Gulland, J.A. 1969. Manual of methods for fish stock assessment. Part 1. Fish population analysis. FAO Man. Fish. Sci. 4. 154 p.
- Jones, R. 1981. The use of length composition data in fish stock assessments (with notes on VPA and cohort analysis). FAO Fish. Circ. 734. 55 p.
- Pauly, D. 1980. A selection of simple methods for the assessment of tropical fish stocks. FAO Fish. Circ. 729. 54 p.
- Pauly, D. 1984. Fish population dynamics in tropical waters: a manual for use with programmable calculators. ICLARM Studies and Reviews 8, 325 p. International Center for Living Aquatic Resources Management, Manila, Philippines.
- Ricker, W.E. 1958. Handbook of computations for biological statistics of fish populations. Bull. Fish. Res. Board Can. 119, 300 p.
- von Bertalanffy, L. 1934. Untersuchungen über die Gesetzmäßigkeiten des Wachstums I. Roux' Arch. Entwicklungsmech. 131: 613-652.

Comments on Age-Length vs. Length-Age Relationships

A.A. ROSENBERG

*Marine Resources Assessment Group
Center for Environmental Technology
Imperial College of Science and Technology
48 Prince's Gardens
London, SW7 1LU
United Kingdom*

J.G. POPE

*Ministry of Agriculture, Fisheries and Food
Fisheries Laboratory
Lowestoft, Suffolk NR33 0HT
United Kingdom*

Rosenberg, A.A. and J.G. Pope. 1987. Comments on age-length vs. length-age relationships, p. 305-309.
In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

Two views of the problems related to the estimation of the time interval (Δt) between two length groups by means of growth parameters (L_{∞} and K) are presented. The first of these views leads to an equation useful for correcting estimates of Δt close to L_{∞} . The second leads to the suggestion that the classical estimation from the regression of length on age should continue to be used for estimating age from length based on a calibration experiment.

Introduction

The contributions of Hoenig (Part I, this vol.) and of Sanders (Part I, this vol.), which both presented methods for the estimation of growth parameters suited to computation of Δt values (mainly for use in length-cohort analysis and related methods) resulted, at the conference in a rather excited discussion, and the opportunity was offered to the participants of this discussion to submit, after the conference, comments to be consolidated into a brief paper and included in the proceedings. Two such written comments by A.A. Rosenberg and J.G. Pope were received by the time this volume was completed, and they are presented below (the Editors).

J.G. POPE

It was noted that using a von Bertalanffy growth curve fitted to the average length-at-age data will cause bias in estimates of Δt if K and/or L_∞ are subject to variation. Sainsbury (1980) gives a general description of the problem of estimating von Bertalanffy curves subject to individual variations in K and L_∞ and both Jones (Part I, this vol.) and Laurec and Mesnil (Part I, this vol.) consider the problem in the context of bias in the length-cohort analysis. Laurec and Mesnil (Part I, this vol.) provide formulae for the calculation of biases due to such variation in L_∞ and K . Clearly, it would be a sensible practice to use these formulae when attempting a length-cohort analysis. In what follows, to understand the problem a little more clearly, I have assumed a rather simpler form of variation in parameters and worked out the size of biases that these would create. These suggest that by avoiding a specified region close to L_∞ it is possible to reduce the biases substantially and this may allow length-cohort analysis to be used without serious error. This finding is, of course, very similar to that of Laurec and Mesnil (Part I, this vol.) but, being based on a more limited study of the problem, is perhaps easier to grasp.

Assumptions

To keep assumptions about growth changes simple, I have only considered the case where the distribution of length at age has a rectangular distribution about the average length at age L_t in the range $L_t - R$ to $L_t + R$. I assume that individual fish remain at the same distance from the average growth curve throughout time and that all fish have a constant mortality rate of Z . The same distance assumption means that all the fish at age t have the same growth increment as the average-sized fish. With a rectangular distribution of size about the mean length at age, fish of length L_2 can be found between ages ta and tb where

$$L_2 = L_{ta} + R$$

$$L_2 = L_{tb} - R$$

within this range the probability distribution of fish of length L_2 is

$$-Ze^{-Zt}/(e^{-Z+b} - e^{Zta})$$

A fish of length L_2 at time t will have grown from length L_1 in time Δt . It may be thought of as the survivor of $e^{M\Delta t}$ fish of length L_1 . We need to compute the average value of $e^{M\Delta t}$ to give the appropriate multiplier for the numbers of fish of length L_2 to convert them to equivalent numbers of length L_1 fish. This multiplier will be given by

$$\int_{ta}^{tb} e^{-zt} \left(\frac{L_\infty - L_t + L_2 - L_1}{L_\infty - L_t} \right)^{M/K} dt / \frac{Z}{e^{-zta} - e^{-zb}}$$

converting the variable of integration from t

$$\text{to } X = L_t - L_2$$

enables the integration to be performed by series expansion in X . Both the denominator and the integrand yield odd valued power series in R and, thus, writing the multiplier = $A_0 + A_2 R^2 + A_4 \dots$ etc. we may evaluate A_0, A_2 etc. by inspection.

$$A_0 = \left(\frac{L_{1/2} - L_1}{L_\infty - L_2} \right)^{M/K}$$

which is, of course, Jones result and which would be the answer if $R = 0$.

$$A_2 = \left(\frac{L_\infty - L_1}{L_\infty - L_2} \right)^{M/K} \left[\frac{M(M-K)}{K^2 6 (L_\infty - L_1)^2} + \frac{2M(Z-M-K)}{K^2 6 (L_\infty - L_1)(L_\infty - L_2)} \right. \\ \left. - \frac{M - (2Z - 3K - M)}{K^2 6 (L_\infty - L_2)^2} \right]$$

If we consider that $M \approx K$ as is often the case, then this becomes

$$A_2 = \left(\frac{L_\infty - L_1}{L_\infty - L_2} \right)^{M/K} \left[\frac{(Z - 2K)}{3K} \left(\frac{L_2 - L_1}{(L_\infty - L_1)(L_\infty - L_2)^2} \right) \right]$$

which suggests that the correction will be strongly dependent on the value of

$$\frac{R}{(L_\infty - L_2)}$$

and dependent on the size of

$$\frac{(L_2 - L_1)}{L_\infty - L_1} \text{ and } \frac{Z - 2K}{K}.$$

Since from the rectangular distribution $(-R, +R)$ the variance

$$\sigma^2 = \frac{R^2}{3}$$

we may consider the corrected multiplier of the survivors where $K = M$ as

$$\left(\frac{L_\infty - L_1}{L_\infty - L_2} \right) \left(1 - \frac{\sigma^2 (Z - 2K)(L_2 - L_1)}{K (L_\infty - L_1)(L_\infty - L_2)^2} \right)$$

This is so, provided we ignore the terms A_4 etc. of the expansion.

It is clear that if we keep $(L_\infty - L_2) > 3\sigma$ and $(L_2 - L_1) < \sigma$ then the correction will be no more than

$$-\frac{(Z - 2K)}{K} \frac{1}{27}$$

which should be small for most realistic levels of Z . It, thus, seems that by avoiding the close proximity of L_∞ , we can reduce the bias problem to fairly minor levels at least for the survivors part of the equation.

Discussion

From this rather simplified model of variable growth-at-age and mortality, it is possible to find expressions for the bias in the survivorship multiplier of the length-cohort analysis. This could be used directly as a correction but perhaps it would be more sensible to use it to decide on a sensible largest size for initiating a length-cohort analysis. The size of correction suggests that starting calculation at a largest length group of at least 3 times the standard deviation of length at age smaller than L_∞ and also keeping size classes smaller than this standard deviation should serve to reduce bias to acceptable limits. This result applies to the survivors and an equivalent result for the catch multiplier needs to be developed. Better still is obviously to use the more detailed results of Laurec and Mesnil (Part I, this vol.).

A.A. ROSENBERG

The question of whether the regression of length on age or that of age on length should be used to predict the time necessary for a fish to grow from l_t to l_{t+d} , is a problem of calibration. A substantial literature exists on the statistics of calibration methods, albeit most of it concerns linear calibration (see Hunter and Lamboy (1981) and accompanying discussion for review).

A general calibration problem can be described as follows: we wish to determine a quantity which is difficult or impossible to measure directly. There are two (or more) available measurements we can make. One of these is difficult and expensive but quite accurate; and the other is cheap and simple but contains substantial measurement error. We perform a calibration experiment, where both measurements are taken for a set of samples, and produce a calibration curve of y on x . A further set of samples is then taken and only y is measured. How can we best estimate the underlying quantity of interest?

For the case where all relationships are linear, discussion in the statistics literature over whether the regression of y on x (the classical estimator) or of x on y (the inverse estimator) should be used was instigated by Krutchkoff (1967). He concluded that the inverse estimator had smaller mean squared error than the classical estimator. However, for the linear case, the inverse estimator is not unbiased while the classical is (Williams 1969). Furthermore, Lwin (1981) has shown that the inverse estimator is affected by departures of the residual distributions from normality while the classical estimator is insensitive. Lwin (1981) has provided more detailed comparisons of the two estimators. Of particular interest is the observation that the inverse estimator performs more poorly (in the mean square error sense) than the classical as the point x to be estimated gets farther away from the mean value of x , and that sample size will not alleviate this problem.

For the fisheries problem of interest here, the calibration of length and age is more complicated because of non-linearity. We can consider the calibration experiment is performed by measuring and aging (e.g., with otoliths) a sample of fish. The otolith readings are costly and difficult and the length measurements are easy and cheap. The underlying quantity we want to know about is time. Otolith readings measure age fairly accurately we hope and so can provide a time scale, but length measurements are non-linearly related to age and are in addition highly variable estimates of time.

Many of the results from the study of linear calibration problems apply to the non-linear case. Scheffé (1973) has provided an approach to non-linear calibration problems which uses the classical estimator. Much of his paper is devoted to the construction of interval rather than point estimates. An example is given in Knafl et al. (1984). The inverse estimator has not been used for non-linear calibration.

Based on the results for linear calibration and Scheffé's method the classical estimator from the regression of length on age should continue to be used for estimating age from length based on a calibration experiment.

References

- Hunter, W.G. and W.F. Lamboy. 1981. A Bayesian analysis of the linear calibration problem. *Technometrics* 23: 323-350.
- Knafl, G., C. Spiegelman, J. Sacks and D. Yivisaker. 1984. Nonparametric calibration. *Technometrics* 26: 233-241.
- Krutchkoff, R.G. 1967. Classical inverse regression methods of calibration. *Technometrics* 9: N25-439.
- Lwin, T. 1981. Discussion of paper by Hunter and Lamboy. *Technometrics* 23: 339-341.
- Sainsbury, K.J. 1980. Effect of individual variability on the von Bertalanffy growth equation. *Can. J. Fish. Aquat. Sci.* 37: 241-247.
- Scheffé, H. 1973. A statistical theory of calibration. *Ann. Stat.* 1: 1-37.
- Williams, E.J. 1969. A note on regression methods in calibration. *Technometrics* 11: 189-192.

Use of ELEFAN I for Sampling Design

D. LEVI

*Istituto di Tecnologia
della Pesca e del Pescato (ITPP)
Mazara del Vallo, Italy*

M.G. ANDREOLI*

*Istituto di Zoologia
Università di Parma
Parma, Italy*

L. CANNIZZARO

*Istituto di Tecnologia
della Pesca e del Pescato (ITPP)
Mazara del Vallo, Italy*

Levi, D., M.G. Andreoli and L. Cannizzaro. 1987. Use of ELEFAN I for sampling design, p. 311-319.
In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

Data from one year sampling of sardines (*Sardina pilchardus*) in a port of the Adriatic Sea are used to introduce a new method to design sampling "as sampling goes on".

The method is based on modification of Pauly's ELEFAN I program, used in conjunction with an estimator related to the jackknife technique. Potential uses of the proposed method are discussed.

Introduction

Analytical models in fish population dynamics depend critically on the estimates of parameters of the growth equation, especially K and L_{∞} . It appears of some interest to assess the influence of the number of samples used to estimate these parameters on the confidence of their estimates.

We propose here to use estimates generated by a modified version of the ELEFAN I program and jackknifed estimates of parameter changes ($\Delta\%$) to decide whether sampling should be carried on. A comparison of estimates obtained with and without reference to our new estimator ($\Delta\%$) is also presented.

*Present address: Istituto di Istologia e Embriologia, Università di Palermo, Palermo, Italy.

Material and Methods

MATERIAL

Length data, represented by 79 samples of sardines, *Sardina pilchardus*, (Table 1) collected from the landings in the port of Cesenatico, Central Adriatic, in 1976 (Levi et al. 1985), were available (there were 115 samples originally, but the data of samples collected the same day were pooled resulting in the 79 samples in Table 1).

The port of Cesenatico was selected because it provided most samples on the stocks of small pelagic fish of the Northern and Central Adriatic since 1974. For further information on the sampling program see Levi (1978).

METHODS

We define here as "optimum sampling" a situation where the withdrawal of any sample from a set of samples has no influence on the estimate of a parameter. An estimator of the adequacy of sampling can then be defined which is equal to the percentage difference between the parameter estimated once from all available samples (n) and the same parameter estimated as the mean of n estimates, each based on all samples minus a different sample.

This estimator is related to the jackknife method; in fact the estimator can be expressed as:

$$\Delta\% = \frac{(St - \bar{\varphi}_{n-i}) \times 100}{\bar{\varphi}_{n-i}} \quad \dots 1)$$

where St is the desired statistic (i.e., the growth parameter estimate) based upon the complete sample (Sokal and Rohlf 1981), and $\bar{\varphi}_{n-i}$ is the average of the successive parameter estimates based on all samples minus one.

We now define sampling as optimum when the above estimator is equal to ϕ (not in mathematical terms, obviously, but as a "computer- ϕ " allowing for rounding errors in the calculation procedures for parameter estimation). The higher the percentage difference (= the estimator), the less satisfactory will be the sampling.

To calculate the desired statistics, St , the parameters K , L_∞ and " t_o " on n samples, the ELEFAN I program (Pauly et al. 1983) was used, after it was modified to eliminate any "at least partly human-aided optimization procedure". Note that the " t_o " discussed throughout this contribution is a relative measure and pertains to a birthdate set at 1 January. It is not a "real" t_o (see Pauly, Part I, this vol.).

The modifications introduced in ELEFAN I (translated into Fortran IV) were:

- allowing K and L_∞ to vary between given boundaries, obtained from the paper by Pauly (1980) as characteristic of the family to which the sampled population belongs (i.e., Clupeidae);
- automatically identifying the maximum ESP/ASP ratio found for all possible combinations of the three parameters K , L_∞ and t_o within the boundaries of K and L_∞ .

These modifications were paid, as predictable, in terms of computer time, although there is still a lot to be done to improve fitting procedures beyond the original version of the program, e.g., by optimizing through the program itself the selection of the appropriate step sizes in iterative search.

To run a modified ELEFAN I associated with the above jackknife-like procedure on 79 samples required about 1,664 min. on an IBM 3033, while about 391 min. were necessary to process 48 samples. Hence an increase of number of samples of 1.63 involved an increase of 4.25 in computer time; this suggests that beyond a limited number of samples it is necessary to use powerful computing facilities.

Table 1. Length-frequency data of *Sardina pilchardus* sampled from Cesenatico, Central Adriatic Sea, 1976.

Date	5/1	14/1	17/1	23/1	4/2	11/2	26/2	27/2	3/3	16/3	17/3	22/3	25/3	6/4	13/4	14/4	16/4	22/4	23/4	26/4
Sample No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Length classes (cm)																				
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	2	0	0	0	5	2	0	4	0	0	0	1	3	0	3	0	6	5	1	3
14	5	1	0	6	10	8	6	17	2	2	0	0	2	1	3	1	4	0	10	5
15	26	1	6	44	42	29	25	58	33	6	7	1	16	13	17	0	13	11	46	18
16	128	71	78	98	101	57	86	67	149	101	41	39	148	115	74	16	13	42	87	90
17	61	81	87	52	154	73	77	27	178	178	56	101	361	202	104	69	96	80	52	67
18	13	23	37	6	53	40	6	3	21	96	21	28	80	53	15	22	19	2	18	15
19	2	3	5	1	8	6	0	1	1	16	4	3	3	7	1	3	0	0	1	1
20	0	1	1	0	0	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total	237	181	214	207	373	217	200	177	384	399	129	173	613	391	218	111	152	140	215	199

Date	28/4	5/5	6/5	7/5	10/5	13/5	18/5	20/5	26/5	28/5	1/6	15/6	18/6	21/6	23/6	25/6	30/6	2/7	8/7	16/7
Sample No.	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Length classes (cm)																				
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	24	7	14	0	2	1	0	4	0	0	10	0	1	8	2	1	0	2	0	0
14	43	30	11	3	12	3	18	28	0	2	41	8	10	31	1	0	9	4	8	0
15	163	21	17	52	128	24	27	186	2	18	49	46	31	40	13	13	20	30	20	8
16	279	61	44	182	215	47	31	273	23	33	150	255	86	104	73	59	77	67	67	98
17	255	51	66	165	86	81	40	107	74	35	140	221	57	83	43	55	55	29	63	77
18	49	34	37	25	22	37	28	16	29	9	20	17	4	10	2	8	7	0	9	4
19	5	0	7	0	1	1	4	0	8	1	0	3	0	0	0	2	0	0	1	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total	825	204	196	427	466	194	148	614	136	98	410	550	189	276	134	138	168	132	168	187

Table 1. Continued.

Date	19/7	20/7	27/7	5/8	11/8	18/8	27/8	6/9	11/9	15/9	17/9	24/9	25/9	29/9	11/10	13/10	20/10	21/10	25/10	29/10
Sample No.	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Length classes (cm)																				
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
11	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
12	0	0	0	0	0	0	1	1	0	0	2	0	2	0	0	0	0	0	0	
13	0	1	4	12	0	3	56	5	4	0	21	5	9	2	3	0	0	1	3	2
14	5	1	42	42	0	27	36	1	37	0	16	19	11	2	21	0	6	13	30	24
15	17	8	85	124	6	112	24	19	23	8	10	19	18	8	27	1	4	6	28	24
16	75	60	107	155	31	118	84	56	64	41	36	55	101	44	76	13	20	42	57	96
17	56	68	88	46	57	74	74	53	45	73	55	35	223	63	142	70	117	136	174	93
18	6	13	16	9	24	21	38	9	14	15	22	6	105	38	61	43	129	65	45	65
19	2	1	2	0	2	1	6	2	2	4	3	0	21	4	8	15	23	14	10	6
20	0	0	0	0	0	0	1	0	1	0	0	0	1	1	0	4	4	0	0	2
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Total	161	152	344	388	120	356	321	146	190	141	165	139	491	162	338	146	303	277	347	312

Date	3/11	4/11	5/11	9/11	10/11	11/11	17/11	24/11	25/11	26/11	30/11	7/12	9/12	19/12	14/12	16/12	17/12	21/12	22/12
Sample No.,	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
Length classes (cm)																			
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	4	0	5	0	0	0	1	0	1	0	1	0	1	0	0	1	0	2	1
14	37	1	21	0	1	2	12	7	5	4	69	0	3	5	11	0	1	2	5
15	15	1	22	0	4	3	11	10	2	3	111	1	4	2	14	0	1	2	2
16	48	16	34	2	21	13	65	60	16	8	67	10	43	12	70	26	26	31	46
17	98	45	112	56	121	51	152	94	64	37	90	79	135	47	104	71	77	97	178
18	43	35	67	62	112	39	48	35	27	56	14	61	75	34	17	18	20	31	35
19	9	12	17	14	25	9	6	5	5	6	2	5	8	14	0	3	4	2	6
20	1	2	1	4	1	2	0	0	0	1	0	4	1	1	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Total	255	112	279	138	285	119	295	211	120	115	354	160	270	116	216	119	129	167	273

Results

A first exercise was to run the above procedure on all available samples, under the assumption that year 1976 had been, as a whole, oversampled. Estimates of growth parameters so obtained appear to be very reliable. Our estimator ($\Delta\%$), defined as the difference between a single run of the modified ELEFAN I on all samples and the average of 79 runs dropping a sample in turn (thus expressing the average variability introduced by different samples) was $= \phi$ for all three parameters: i.e., taking out *any* sample had no influence on the final estimation (Table 2).

Table 2. Values of annual growth parameters (L_∞ , K, " t_o ") calculated with modified ELEFAN I (St), with jackknife method ($\bar{\varphi} = \hat{S}_t$), and precision estimators ($\Delta\%$, C.V.) for different monthly sampling rates and for all available samples.

	A ^a	B	C	D	E ^b
n ^c	12	24	36	48	79
St ^d L _∞	24.00	24.00	25.00	24.00	25.00
$\bar{\varphi}$ ^e L _∞	22.167	13.938	34.236	11.760	25.00
$\Delta\%$ ^f L _∞	0.69	1.79	1.07	1.07	0.00
C.V. ^g L _∞	13.08	25.48	4.31	23.79	0.00
St K	0.200	0.400	0.600	0.400	0.350
$\bar{\varphi}$ K	-5.392	-1.708	1.621	1.134	0.350
$\Delta\%$ K	71.76	18.64	5.11	4.07	0.00
C.V. K	18.36	79.68	66.00	14.01	0.00
St " t_o "	-6.181	-1.268	-1.201	-1.268	-2.459
$\bar{\varphi}$ " t_o "	-53.431	28.422	9.219	15.721	-2.476
$\Delta\%$ " t_o "	227.80	50.45	19.87	22.18	0.01
C.V. " t_o "	14.63	40.10	45.05	23.44	0.65

^aColumns (A) to (D) correspond to sampling rates of one to four times per month.

^bColumn (E) is based on all samples of Table 1.

^cn = number of samples.

^dSt = best estimated obtained through ELEFAN I applied to all samples.

^e $\bar{\varphi}$ = jackknifed estimate.

^f $\Delta\%$ = percent difference between St and the average of all pseudovalues (equation (1)).

^gC.V. = coefficient of variation of jackknifed estimate (equation (2)).

We then estimated the reliability of growth parameter estimates based on different monthly rates of sampling during the year. The outputs of the modified ELEFAN I and of jackknifed estimates (i.e., St and $St = \bar{\varphi}$) are summarized in Table 2, together with their $\Delta\%$ and the jackknifed coefficient of variation:

$$C.V. = \sqrt{\frac{\sum (\varphi_i - \bar{\varphi})^2}{n(n-1)}} \cdot 100 \quad \dots 2)$$

for monthly sampling rates of one to four times a month evenly spaced in time (i.e., every first week; every first and third weeks; first three weeks; every four weeks). In the rare case when one week had not been sampled, the two nearby weekly samples were interpolated. If more than one sample was available in a given week, frequencies were pooled and the resulting sample was given the date of the first sample in the week.

It is quite evident that, in real life, a simple increase of sampling intensity does not provide a parallel increase of reliability of estimates. The improvement depends very much on the amount of variation introduced by new samples every month. There will be "stable" months, where the length-frequency distributions do not change very much from one sample to the other, and conversely months when the variability is so high that any new sample introduced in the calculations calls for further sampling. Hence, monthly sampling rate should be variable.

A closer look at what happened within months seemed, thus, necessary. A comparison of $\Delta\%$ s was attempted between estimates obtained on all samples available in a given month and estimates obtained by increasing the number of samples from two up to the maximum number of samples available in that same month. Results are summarized in Tables 3A-C (A for L_{∞} , B for K and C for " t_0 "). The values for a given parameter reported in the last row of each table are the estimates of that parameter which would have been obtained if *all* samples available in that month had been

Table 3A. Percent differences between estimates of L_{∞} obtained by using an increasing number of the available samples in a given month ($n = 2, 3 \dots 11$).

	Σ^a	J	F	M	A	M	J	J	A	S	O	N	D	All 1976 samples ^b
n	4	8	13	21	30	37	43	47	54	60	71	79	79	
2	4.44	0.00	2.22	0.00	1.10	7.87	0.00	0.00	1.08	4.17	1.10	0.00	—	
3	1.44	0.00	0.00	1.47	1.43	5.11	2.94	0.00	1.46	0.73	0.00	2.08	—	
4	0.53	3.87	1.10	0.55	5.26	1.59	0.00	0.00	3.78	1.10	0.00	1.55	—	
5			0.44	0.00	4.96	4.80	0.00		2.54	0.84	0.00	1.24	—	
6				0.00	2.22	1.41	0.00		0.74	0.36	0.00	1.03	—	
7					0.00	2.64	1.20				0.00	0.88	—	
8						0.00	2.56			2.64		0.00	0.77	
9							1.58				0.00		—	
10											0.00		—	
11											0.00		—	
$\hat{S}^c L_{\infty}$	23.875	26.125	22.100	23.00	28.111	25.714	22.00	23.00	28.857	22.583	23.00	25.813	25.00	
$S^d L_{\infty}$	23.50	23.50	22.50	23.00	25.00	24.00	22.00	23.00	25.00	23.00	23.00	24.50	25.00	

^a Σ = cumulative number of samples over time.

^bEstimate based on all samples in Table 1.

^c \hat{S} = jackknifed estimate using all samples of a given month (or all 1976 samples in last column).

^d S = estimate using all samples of that month (or all 1976 samples in last column).

Table 3B. Percent differences between estimates of K obtained by using an increasing number of the samples available in a given month ($n = 2, 3 \dots 11$).

	Σ^a	J	F	M	A	M	J	J	A	S	O	N	D	All 1976 samples ^b
n	4	8	13	21	30	37	43	47	54	60	71	79	79	
2	0.00	25.00	86.67	3.45	71.43	65.22	89.47	44.00	72.73	89.47	28.57	0.00	—	
3	3.57	15.38	0.00	82.35	40.00	55.56	134.78	28.57	20.00	3.57	0.00	1.79	—	
4	1.33	6.49	18.92	82.61	65.22	5.88	0.00	33.33	24.53	16.13	0.00	1.33	—	
5			20.00	77.27	11.11	58.33	1.12		21.62	28.57	0.00	1.06	—	
6				73.91	73.33	14.29	3.64		3.45	10.20	0.00	0.88	—	
7				70.83	26.32	12.00			1.45		0.00	0.76	—	
8				81.40	35.59					0.00	0.66	—		
9					3.45					0.00		—		
10										0.00		—		
11										1.49		—		
$\hat{S}^c K$	0.988	0.712	1.000	-1.481	1.267	0.329	1.117	1.750	1.086	1.317	0.764	0.994	0.35	
$S^d K$	0.95	0.90	0.60	0.05	1.00	0.20	0.95	1.00	1.00	0.90	0.90	0.95	0.35	

^a Σ = cumulative number of samples over time.

^bEstimates based on all samples in Table 1.

^c \hat{S} = jackknifed estimate using all samples of a given month (or all 1976 samples in last column).

^d S = estimate using all samples of a given month (or all 1976 samples in last column).

Table 3C. Percent difference between estimates of " t_o " obtained by using an increasing number of the samples available in a given month ($n = 2, 3 \dots 11$).

n	Σ^a	J	F	M	A	M	J	J	A	S	O	N	D	All 1976 samples ^b
		4	8	13	21	30	37	43	47	54	60	71	79	79
2		12.97	22.74	87.93	4.52	58.69	81.91	93.73	48.12	75.82	93.71	25.70	0.49	
3		3.09	16.40	0.61	45.91	59.98	42.66	91.93	52.71	8.44	4.37	0.15	20.52	
4		4.02	7.29	75.30	31.67	11.60	2.77	1.48	51.39	42.72	26.02	0.12	17.69	
5				83.56	23.76	9.32	50.63	1.43		32.39	15.20	0.09	14.15	
6					19.05	73.82	44.39	8.44		2.78	23.22	1.40	11.20	
7						15.89	5.76	40.89			21.04	0.07	9.76	
8							43.49	83.26				0.06	8.64	
9								16.86				0.05		
10												0.05		
11												6.17		
St "t _o "		-1.018	-0.896	23.497	-94.384	0.422	15.788	-0.440	1.292	-0.394	0.366	-1.031	-0.105	-2.476
St "v _o "		-1.163	-1.173	-1.215	-30.236	-0.679	-5.011	0.817	-0.595	0.236	-0.715	-0.652	-0.312	-2.459

^a Σ = cumulative number of samples over time.

^bEstimates based on all samples in Table 1.

^cSt = jackknifed estimate using all samples in corresponding and preceding month(s) (or whole of 1976 in last column).

^dSt = estimate using all samples in that month (or whole of 1976, in last column).

used as input for the modified ELEFAN I program. The row before the last reports parameters obtained through the modified jackknife method. The last column reports the best estimate of the same parameter obtained by running the program on all samples (79) available in that year. In most cases, these values are different between months, and between monthly and yearly estimates. Thus, the values of $\Delta\%$ can be viewed as estimators of goodness of sampling only within a month. Months where $\Delta\%$ remained high or suddenly increased again, should have been sampled more. Conversely, sampling could have been stopped earlier without loss of information when $\Delta\%$ had quickly approached ϕ for all three growth parameters.

Discussion

It may not be fully clear at this point why we trust the parameter estimations on 79 samples taken along the year more than those based on monthly samples alone. The reason is very simple. In the case of 79 samples, 79 pseudovalues, computed such as to provide "an unbiased estimate" of the growth parameters, are practically all the same (" t_o " is a minor exception; \bar{v} is -2.476 while St is -2.459). Within each month, on the other hand, the pseudovalues keep fluctuating, sometimes with one order of magnitude. Such large fluctuations seem likely to be due to the fact that our modified version of ELEFAN I has not yet incorporated a routine to eliminate the computation of ESP/ASP ratios higher than one, corresponding to unreasonably small values of K and incredibly large " t_o "s, by putting "flags" on peaks already counted once (see Pauly, Part I, this vol.).

In fact in all cases where ESP/ASP is smaller than one (e.g., January, February, August, September, October, November, December of Table 3 and columns (1), (2), (3) of Table 2), the values of K, L_∞ and " t_o " are acceptable. The month of May is an exception, showing anomalous values for t_o with ESP/ASP still < 1 , but there the K value is very low indeed.

The consequence is that the standard deviation remains very high, and so does the coefficient of variation of St. In those months when this fluctuation does not occur, however, both the jackknifed estimate and the St obtained through ELEFAN I are, for all parameters, very different from the yearly estimate (see, e.g., the November values in Table 3).

These are also the reasons why we devised the new estimator $\Delta\%$ for use *within* each month. The jackknife could not be used traditionally as suggested by Pauly et al. (1983) to estimate the variance of the parameters being calculated, as it is both sensitive to sampling error influence when

dealing with few samples, and to the number of observations when dealing with many samples. This can be shown by transforming Tukey's equation as given by Sokal and Rohlf (1981) into:

$$\bar{\varphi} = n (S_t - \bar{\varphi}_{n-i}) + \bar{\varphi}_{n-i} \quad \dots 3)$$

Table 2 shows that the differences between \hat{S}_t and S_t can only be explained by $\Delta\%$. In these cases ESP/ASP ratios are always < 1 , n is fairly high, yet the jackknifed estimates (\hat{S}_t) produce unacceptable values.

This phenomenon is also apparent, although for smaller differences, in Table 3 (see values for February, August, September and November).

From the above remarks one could easily reach the conclusion that the sampling design for sardines landings in Cesenatico in 1976 was not very good, and that one should try to better allocate between months the sampling effort. This is the use commonly made of examination (generally by ANOVA) of historical data (e.g., Gulland 1955). However, that contribution does not propose a method for use when no historical data are available.

A sampling design cannot be planned once forever; rather, it should be changed when the estimated parameters vary in time.

In developing fisheries, explicit consideration of the variation in time of biological parameters is necessary if parameters are to be properly estimated, rather than guessed.

Based on the experience presented in this paper, we would like to propose the following procedure to optimize sampling "as sampling goes on".

I — Start sampling in a given month.

Run the ELEFAN I program on samples collected, and calculate S_t . Also calculate pseudovalues and their average. Go on sampling and stop when the $\Delta\%$ between the mean of pseudovalues and S_t is close to ϕ and does not fluctuate.

II — When sampling is reinitiated the following month for continuous spawners, possibly after several months for other fish, step I is repeated, but, also, modified ELEFAN I is run on all samples available so far. Sampling is then stopped when S_t and \hat{S}_t ($= \bar{\varphi}$) converge while $\Delta\%$ tends to decrease.

Convergence should occur when the cycle of changes in the population structure begins to repeat itself and has been captured. This should rarely occur in less than a year.

III — When the method proposed here is applied to data obtained from trawl surveys, "stratum" might be substituted for "month".

Acknowledgements

We wish to express our gratitude to the I.R.P.E.M. (CNR — Ancona) team who provided the sample data, namely N. Cingolani, G.F. Giannetti, R. Scorceti and E. Peloni, and to Prof. G. Bombace for permitting the release of the data.

This work could be completed in time for the meeting only due to the willing collaboration of Prof. P.M. Angelo and Dr. G.C. Ferrara who helped in data processing on more powerful computing facilities than those directly available to us. To them go our sincerest thanks.

References

- Levi, D. 1978. A summary report on the routine sampling programme for sardine and anchovy in the Adriatic. FAO—Rapport de la septième session du group de travail sur l'évaluation des ressources et les statistiques de pêche du Conseil General des Pêches pour la Méditerranée (CGPM). FAO—Rapport sur les pêches 204: 115-117.
- Levi, D., N. Cingolani, G.F. Giannetti, R. Scorceti, E. Peloni and N. Frontini. 1985. Stock assessment of *Sardina pilchardus* (Wald.) and *Engraulis encrasicolus* (L.) of the Northern-Central Adriatic Sea: some preliminary results. Quad. Ist. Ric. Pesca Maritt. Ancona 4(1): 1-34.
- Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. J. Cons. CIEM 39: 175-192.
- Pauly, D., N. David and J. Ingles. 1983. ELEFAN I: User's instruction and program listings (Rev. 2). ICLARM MS, 28 p.
- Sokal, R.R. and J. Rohlf. 1981. Biometry: the principles and practice of statistics in biological research. 2nd Edition. W.H. Freeman and Co., San Francisco.

Fisheries Management in a Developing Country: the Most Appropriate Balance of Size- and Age-Related Methods for Practical Assessments

C.P. MATHEWS

*Mariculture and Fisheries Department
Kuwait Institute for Scientific Research
P.O. Box 24885, Safat
Kuwait*

Mathews, C.P. 1987. Fisheries management in a developing country: the most appropriate balance of size- and age-related methods for practical assessments, p. 321-334. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

This paper reviews growth and mortality analyses presently completed or underway in Kuwait and shows how the use of a combination of purely size-based and age-length-key based approaches has enriched stock assessment work and produced more flexible, cost-effective and timely output, without detracting from the longer-term work required for successful application of classical age-based techniques.

Introduction

Stock assessment has traditionally been based on the techniques initially developed in northern Europe from the turn of the century onwards. These techniques rely heavily on the aging of fish by means of examination of scales, otoliths and other hard parts, sometimes after simple cleaning, sometimes after much careful preparation. The work on aging of fish has generated a copious literature which lies outside the scope of this paper. Holden and Raitt (1974), however, give a useful summary of the techniques used, while Williams and Bedford (1974) examine the aging of otoliths in some detail. Christensen (1964) first reported the use of heat in preparing otoliths for examination. Other parts of fish, e.g., vertebrae, opercular bones and spines have also been used, including especially scales which are particularly useful in the case of many highly priced species such as salmon that cannot be sacrificed or bought for age determination. In spite of the development of a widely accepted technology for age determination (Bagenal 1974), this approach has been generally difficult to apply in tropical and equatorial waters because many fish in these areas tend not to form annual marks. Marks are either absent, or they are present but are formed irregularly. They are usually also much more difficult to distinguish than the annual marks found in the hard parts of temperate species.

A new approach was developed by Panella (1974) who showed that marks were formed in otoliths at daily intervals; a rigorous demonstration of the daily occurrence of microstructures in otoliths was provided by Brothers et al. (1976) who applied this technique to the validation of aging by means of annual marks in a tropical species of hake. This technique, suited for tropical and warm

water fishes, has been used for stock assessment purposes, but only with difficulty because high quality equipment and well-trained personnel are needed, and because it is laborious and costly.

In general, the high cost of establishing a system for fish aging, coupled with general difficulties of applying this approach in tropical countries, most of which are economically less able to support costly research, has caused great difficulties in carrying out stock assessment in tropical and equatorial waters. Morgan (1983) provided provisional cost estimates for aging fish and noted how expensive aging by means of otoliths could be.

The difficulties in aging fish by means of their hard parts lead to work being done on the separation of age groups by other techniques. One of the earlier attempts to do this on fish was carried out by Cassie (1954) who separated age groups by means of otolith examination; Mathews (unpublished data) carried out a similar analysis for a population of *Epinephelus morio* from the Campeche Bank, Mexico, and found that one age group was overlooked because one of the inflection points used to separate the age groups was difficult to distinguish. Abramson (1971) developed a computer-based technique for separating age groups but the NORMSEP program he developed required as input information the limits of the age-specific size-frequency distributions and this limited its usefulness since the age limits chosen for each age group must often be based on subjective considerations. Mathews (1974) compared growth rates for four species of tropical fish determined by these techniques of mathematical separation and by means of studying otoliths. He concluded that the study of hard parts was usually more reliable, but that a combination of techniques was often very helpful.

More recently, Pauly and David (1981) and Pauly (1982) have developed a package of programs (ELEFAN) which, in its various forms, provides an exceptionally versatile tool for stock assessment, and has been since applied to data from tropical fish and shrimp populations. Pauly (Part I, this vol.) discusses the methodology for obtaining estimates of growth, total mortality and other relevant parameters required for population analysis from size-frequency data only. If the empirical equation of Pauly (1980) can be relied upon to provide estimates of natural mortality, then a complete assessment of a fish stock may be carried out by analyzing only a suitable set of size-frequency data. This potential makes the ELEFAN techniques particularly attractive for application to warm-water populations in developing countries.

The Establishment of a Fisheries Management and Stock Assessment Capability in Kuwait

Work on shrimp stocks in Kuwait was initiated in 1978, and on fish stocks in 1980. Results of work on shrimp have been reported elsewhere (see Mathews et al., Part I, this vol.). Catch-and-effort data on finfish are obtained routinely through monthly interviews of fishermen at landing places (Morgan 1982a). In 1981, routine measurements of samples of some species of fish in the markets were introduced on a monthly basis, and were extended to cover more species in 1982. Data on size frequencies have been obtained regularly at sea on R/V Oloum 1 since 1978; although it was only possible to sample fish populations from 4 to 8 times per year, these cruises provide a useful source of data for additional assessment work.

Starting in 1981, a capability for aging fish was introduced in Kuwait. Aging of grouper or "hamoor" (*Epinephelus tauvina*), red snapper or "hamra" (*Lutjanus coccineus*) and croaker or "newaiby" (*Otolithes argenteus*) has been carried out routinely and Age-Length Keys have been constructed according to the classical approach of Holden and Raitt (1974). Because Age-Length Keys are actually based on age and length, catch-and-effort data, they will be referred to here as "age-length catch-effort keys" (ALCEKs) as the term "age-length key" is deceptive in this context. This work now proceeds routinely. Reliable ALCEK analyses have, however, only started to become available in late 1984, whereas assessment advice was required much earlier. Therefore ELEFAN-based analyses were carried out on several stocks. These include pomfret or "zobeidy" (*Pampus argenteus*) (Morgan 1985) and newaiby (*Otolithes argenteus*), both of which are fast growing species; hamoor, a slow growing species (Baddar and Morgan 1984); and hamra, a very slow growing species (see Morgan, Part I, this vol.). Actually, the ELEFAN technique should not have been used in conjunction with these last two, high-longevity species (Pauly 1982), but it was in fact applied in an effort to obtain preliminary evidence of the condition of the stocks.

Mathews and Samuel (1985a, 1985b) have updated the age-based work on hamoor, hamra and newaiby and present assessments based on ALCEKs after full validations of the ages (Holden and Raitt 1974). However, these assessments were produced two years later than the size-based assessments.

The object of this paper is to review work presently completed or underway in Kuwait, and to show how the use of a combination of purely size based, and of age and age-length key based approaches has enriched stock assessment work in Kuwait and has produced more flexible and timely outputs, without detracting from the higher precision and longer-term work required for successful application of the more established and better known age-based techniques.

The methodology for size-based estimation of growth parameters using ELEFAN I has been described by Pauly and David (1981) and Pauly (1982 and Part I, this vol.), while Morgan (1985 and Part I, this vol.) has described fully its application to Kuwaiti fish stocks. Mathews and Samuel (1985a) describe the application of age validation techniques to hamoor, hamra and newaiby while Mathews and Samuel (1985b) provide ALK-based stock assessments for these species. Prices quoted here are 1983 weighted annual means, derived from data supplied by the Central Statistical Office, Kuwait, converted to US\$ at exchange rates valid in October 1984.

Results of Stock Assessments Carried Out in Kuwait

Various species of fish have been studied. Table 1 summarizes some of the data on the key stocks in Kuwait.

ZOBAIDY

This fish provides a substantial proportion of the total finfish landings (7.7% in 1984), is very high priced ($\sim \$5.07/\text{kg}$) and is one of Kuwait's preferred species. Otoliths were examined and found to provide few identifiable marks (Bedford 1982; Williams 1986). Examination showed that these are unreliable for age determination. The otoliths are thin, fragile and generally unsuited for traditional age determination. Morgan (1985) carried out an ELEFAN-based analysis, and was able to determine L_{∞} , K, selectivity length, age and size at first entry to the fishery, and to construct a yield surface for the fishery. He concluded that it was operating near optimum size and age at entry to the fishery, and that increases in effort would be unlikely to increase landings. He also noted that the structure of the yield surface would lead to a decrease in yield per recruit if size at entry were increased. The growth curve suggests that fish may reach up to ~ 3 years.

NEWAIBY

This species has large, very easily readable otoliths which however, need to be broken and ground (Bedford 1982; Williams 1986); it was, therefore, the first species to be aged routinely. Samuel and Morgan (unpublished data) provided ELEFAN-based analysis and also carried out a preliminary age-based analysis by comparing growth data derived from the mean lengths of samples of known age with an ELEFAN-based growth curve. Very similar estimates of growth and mortality parameters were obtained. Because both age and length data were available from only a single year (1982), mortality rates were based on catch curve analyses only. Results of provisional age and of complete length-based analyses were very similar, and suggested that purely length-based assessments were feasible for newaiby.

Mathews and Samuel (1985b) presented an ALCEK-based analysis, with a yield surface. Their data confirmed previous work which showed that length-based assessments were reliable. Newaiby reaches 5 years, but most of the fishery is based on 1- and 2-year olds with only a few 3- and 4-year olds being taken. They found that the fishery was situated fairly near the eumetric curve and that increases in effort would provide little or no increases in yield, while only a slight benefit would be provided by increasing age and size at entry to the fishery. These results confirmed the previous

Table 1. Species selected for stock assessment in Kuwait; all are commercially important.^a

Latin name	English name	Arabic name in Kuwait	Resource type	Main gear used in landing the species	Maximum length (cm)	Estimated maximum age (years)	Stock assessment technique used
<i>Pampus argenteus</i>	Silver pomfret	Zobaidy	Pelagic	Gillnets	31	3	Size-based only (1981-1983). Aging not possible to date.
<i>Otolithes argenteus</i>	Silvery croaker	Newaiby	Demersal	Trawls, gill nets	55	5	Size-based and preliminary age-based (1982). ALCEK-based (1984).
<i>Epinephelus tauvina</i>	Orange spotted grouper	Hamoor	Demersal, associated with reefs but occasionally trawled	Fish traps (pots)	~110 (reported outside Kuwait to exceed 150)	26	Preliminary size-based (1982-1983). ALCEK-based (1984).
<i>Lutjanus coccineus</i>	Crimson snapper	Hamra	Demersal, associated with reefs, occasionally trawled	Fish traps (pots)	~ 85	~45	Preliminary size-based (1982-1983). ALCEK-based (1984).
<i>Acanthopagrus latus</i>	Yellow-finned black porgy	Sheim	Demersal	Trawl nets, set nets	~ 35	~12	Size- and age-based ELEFAN (1984). ^b ALCEK-based (1985) (projected).
<i>Acanthopagrus cuvieri</i>	Silvery black porgy	Sobaity	Demersal	Trawl nets, hook and line, stake nets	~ 65	~14	Preliminary age-based only (1984).
<i>Pomadasys argenteus</i>	Silvery grunt	Nakroor	Demersal	Fish traps (pots)	~ 82	~20	Preliminary size-based (1983). Both age- and size-based methods difficult to apply.
<i>Arius thalassinus</i>	Giant sea catfish	Chim	Demersal	Trawls and gill nets	~ 65	~13-14	Age-based only (1983-1984).

^aAll names according to Kuronuma and Abe (1972); all lengths are total lengths, except for *Pampus argenteus*, of which fork length is measured because the lobes of the caudal fin are often damaged.

^bUsing the method of Morgan, (Part I, this vol.).

ELEFAN-based analyses. It was particularly significant that cohort-based estimates of Z , taken from ALCEKs and mortalities obtained via the length converted catch curve routine of ELEFAN II were similar.

About 4% of the 1983 fish landings were newaiby; they have fluctuated from ~350 to ~580 t during the last 3 years. Newaiby is also a moderately priced fish (US\$3.63/kg); therefore, it was decided to suspend aging work temporarily, and to use available manpower for working on other species about which less was known. Data available, however, suggested that further work on newaiby by means of quarterly ALCEKs would be justified in the future because growth rates varied between cohorts by amounts sufficient to cause potentially serious biases in growth estimates, and probably also in mortality estimates.

HAMOOR

This species provided 18% of the 1983 landings (the single biggest contribution by weight for any one species) and is traditionally one of Kuwait's most prestigious food species. It is relatively high priced (US\$4.09/kg) and much attention was, therefore, given to stock assessment of this species. Measurement of regular monthly length samples was started in late 1981 and has been carried out since then. The otoliths of this species are large and heavy and need to be broken, ground and heated in an alcohol flame for annual marks to be distinguished successfully. Although markedly less easy to read than newaiby, age determination is feasible. It was demonstrated that annual marks are formed, and that hamoor may reach over 20 years. Formal age validation was not completed until early 1984 (Mathews and Samuel 1985a).

Baddar and Morgan (1984) carried out a preliminary ELEFAN-based assessment in 1983, using the first full year of length measurements (1982), at a time when the aging of this fish was still in doubt. They concluded provisionally that changes in size and age at entry were unlikely to provide a significant change in yield per recruit, while increases in effort might provide ~40% increases in yield per recruit. They compared results of ELEFAN-based analysis with growth data based on the mean lengths of samples of fish of known (but unvalidated) age, and this suggested some deviation of the growth curve obtained from ELEFAN I from these observed data, which justified further work.

Mathews and Samuel (1984b) provided ALCEKs, estimates of growth and mortality parameters and were able to obtain both catch-curve and cohort-based mortality estimates. These differed substantially ($Z = 0.89$ for cohort-based, and 0.23 for catch-curve based estimates, in both 1982 and 1983). These authors also derived a yield surface and stock assessment advice based on ALCEKs and showed that small changes in age and size at entry or increase of effort would not cause any change in yield per recruit.

The hamoor fishery in Kuwait is situated very near the eumetric curve and at a suitably high level of effort. This confirms earlier work of Morgan (1982b) who, using an effort index based on variability in catches, suggested that increases in effort were unlikely to increase landings of hamoor.

Because hamoor is Kuwait's most important fishery both by volume and total value of the landings and because it is now definitely known to be a long-lived species unsuitable for purely size-based analysis, stock assessment by means of ALCEKs will be continued for the next 3-4 years.

HAMRA

This species is the second most important by weight (18.0%). It is a cheap species (US\$1.51/kg), and is regarded by some Kuwaitis as a poor food species. Nevertheless, it is widely eaten in Kuwait by the foreign and lower income section of the population and this, together with the large weights landed, justified detailed stock assessment of hamra. Regular monthly measurements were initiated in late 1981 and have been continued since. Preliminary length-based analysis (Morgan 1984) showed that serious deviations between derived growth curves based on ELEFAN I and growth based on samples of fish of known age occurred, so that this type of analysis was discontinued until hamra aging was formally validated. Mathews and Samuel (1985a) showed that hamra form annual marks and it was confirmed that these fish probably reach up to 45 years old. Mathews and Samuel (1985b) provided an ALCEK-based assessment with a yield surface. They showed that increases in size and age at entry would be likely to provide some increases in yield per recruit, while increases

of effort—at the present size at entry—would probably decrease yield per recruit. The fishery appeared not to be situated close to the eumetric curve.

Because of its importance, this stock also needs to be studied for the next 3-4 years by means of ALCEKs. The cohort-based mortality estimates gave $Z = 0.242$ while catch-curve based estimates gave a value of $Z = 0.089$ and 0.101 for 1982 and 1983, respectively. Samples of large and old fish are difficult to obtain and further work on the older fish is required so as to confirm the mortality estimates in particular, and indeed the whole assessment. The older, larger fish presently contribute ~40% of the landings.

SHEIM

This species provided only 2% of Kuwait's landings in 1983, but (with the exception of sobaity) is the most expensive fish landed (US\$5.24/kg). Although landings are always low, this is a very important species because of its desirability to some Kuwaitis. Routine monthly measurements were established at the beginning of 1983. Otolith collections were established at the same time.

Unlike the other species so far studied, for which landings occur in all months of the year, sheim is seasonal: landings are negligible from December to February so that neither length measurements nor otoliths can be collected in adequate quantity for this period. However, the otoliths are large and easy to read. They may be aged whole or by means of a polished edge (Bedford 1982; Williams 1986). The omission of the three winter months is not crucial for aging since otoliths usually grow by means of marginal accretion in the summer months. Preliminary otolith reading shows that growth at the margins of otoliths occurs from April-May to September-October. Formal validation is being completed and ALCEKs will, if validation proves (as expected) positive, be constructed within a year. Preliminary age-based growth curves and mortality estimates have already been obtained (Samuel, unpublished data). Preliminary aging indicates that sheim reach 12-14 years so that reliable growth estimation based on ELEFAN I may not be possible. Nevertheless, an assessment based on Morgan's method for the simultaneous analysis of length and age data (Part I, this vol.) was produced in 1984 (Morgan, unpublished data) and will be used to provide provisional assessment advice when needed.

SOBAITY

This species provides less than 1% of the landings. Nevertheless, its high price (highest for 1983: US\$6.27/kg) justified attention. However, even more so than is the case with sheim, sobaity is highly seasonal, being taken mainly in November, December and January. It is rare and usually only a few (2-3 to 10-20) fish are landed in any one day. Often fishermen sell the fish directly to consumers without going through the market. Data must, therefore, be supplemented where possible by means of specimens obtained directly by fishing with hook and line. Reliable size-frequency data are, therefore, unobtainable and only age data can be used to obtain estimates of growth parameters. Only rough estimates of Z may be obtained. Work on the assessment of this species continues. Samuel and Bawazeer (1985) summarize available data for this fishery.

NAKROOR

This species is also traditionally preferred by the established population in Kuwait and provided 10% of the landings in 1983. It is moderately expensive (US\$3.62/kg). Because of the large volume and high value of the landings, monthly size-frequency data were collected in 1982 and 1983, while otoliths were collected for a full calendar year from 1982 to 1983. A preliminary ELEFAN-based analysis was applied and suggested that nakroor reach up to 20 years (Morgan, unpublished data) but this work was not continued because the longevity of the species makes successful ELEFAN application unlikely. Study of the burned otoliths showed clear marks but these marks were irregular in their incidence. The growth patterns typical of annual marks were conspicuously absent. Attempts to validate tentatively estimated ages failed, and it was possible to show that opaque and hyaline zones accrued to the margin at irregular intervals and at various times of year. Formal validation showed that the marks were not annual. Fig. 1 shows the size-frequency distributions

obtained in two successive years. It may be seen that peaks and troughs for the two years did not coincide, which suggests that the data are unsuitable for size-frequency analyses. An additional problem arises with the taxonomy of this species, as it is likely that two similar species occur together and cannot presently be separated.

Since both size and ALCEK-based analyses appeared to be inapplicable to nakroor, attempts to study this stock were temporarily suspended. In July 1983, measurement and otolith collections for nakroor were stopped.

It will be necessary to construct growth curves using daily marks in the otoliths of this species before reliable assessment can be provided.

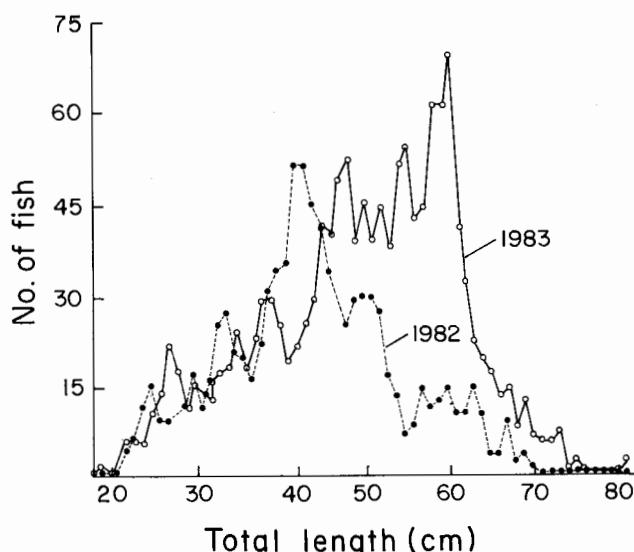


Fig. 1. Size-frequency distribution of nakroor (*Pomadasys argenteus*) sampled in Kuwait waters in successive years.

CHIM

This species is one of the most abundant in the catches obtained on R/V Oloum during the time span 1978-1983, when regular exploratory cruises have been conducted. However, chim tend to occur in schools of different sizes and they are also a highly seasonal species, being more abundant in summer (May-October) than in winter. Adequate samples for size-frequency analyses were difficult to obtain. Apparently reliable growth estimates have been obtained, and may be used for stock assessment once aging is validated. Work continues on this species because Mathews and Samuel (1984a) estimated that chim catches could reach 1,500-2,000 t/year, which could provide a substantial volume of very cheap (~US\$1.80/kg) fish for human consumption.

Costs of Age and Length-Based Assessments

Morgan (1983) provided some estimates of costs based on 1982 data, prior to full establishment of the aging capability. He concluded that age-based analyses were substantially more expensive than length-based analyses. Table 2 presents a full analysis, including hidden costs. Personnel costs include housing and home leave transportation (where supplied by KISR) and are broken up into junior and senior staff. Actual 1984 salaries were used. Variation in personnel costs arises because staff of different experience and seniority were allocated to different activities at different times. Foreign consultants were hired to assist in areas where crucial work required extra inputs, and consultant costs include honoraria, housing and board in Kuwait, and transportation.

Table 2 shows that personnel costs increased up to a relatively constant level of US\$66,000-70,000/year by 1982, although total costs increased slightly because of steady increase in fish

Table 2. Personnel and other costs for the establishment of a stock assessment capability in Kuwait. Personnel in person-months/year, costs in US\$/year; 1984 costs were used in calculations except for fish purchases, for which the actual cost in each year and the conversion rate from Kuwaiti dinars to US\$ for that year were used. (Prices of fish changed markedly in Kuwait from 1981 to 1984). 1984 personnel cost, for first 9 months only.

	Year				Totals
	1981	1982	1983	1984	
A. Fish measurements	(2.50)	2.2	2.5	2.5	
	3,600	3,600	3,600	6,600	14,400
B. Aging of fish					
Training					
Junior	(8.0)	(16.0)	(6.0)		
	11,400	22,800	8,500		
Senior	(1.0) ^a	(1.5) ^a	(1.0) ^a		71,700
	8,500 ^a	12,000 ^a	8,500 ^a		
Validation (informal)					
Junior	(0.5)	(0.5)	(0.5)		
	700	700	700		
Senior	(0.5) ^a	(0.5)	(0.5)		11,300
	4,300 ^a	2,800	2,500		
Validation (formal)					
Junior			(4.0)		
			5,200		
Senior			(2.0)		8,000
			2,800		
Routine reading					
Junior	(3.0)	(1.0)	(8.0)		
	4,300	5,100	11,400		20,800
Senior	(9.0)	(2.0)	(4.0)		
Programming					
Junior		(0.75)			
		1,000			
Senior	(0.5) ^a		(0.75)		18,500
	4,200 ^a		4,000		
ALCEK construction					
Senior	(2.25)	(2.25)	(2.25)		
	6,400	6,400	6,400		19,200
Analysis					
Senior	(2.0)	(2.25)	(2.25)		
	9,800	10,500	10,500		30,800
Management					
Senior	(1.0)	(1.0)	(1.0)		
	6,900	6,900	6,900		20,700
Total personnel cost	32,000	69,300	66,000	48,100	215,400
C. Fish purchases	7,000	10,000	22,000	21,000	60,000
Total operating costs	39,000	79,300	88,000	69,000	275,400
Species covered	Hamoor	Hamoor	Hamoor	Hamoor	Hamoor
	Hamra	Hamra	Hamra	Hamra	Hamra
	Newaiby	Newaiby	Sheim	Sheim	Sheim
		Nakroor	Sobaity	Sobaity	Sobaity
			Nakroor		Various
					<i>Epinephelus</i>
					spp.
					Some other species

^aConsultant.

purchases required for biological (age-related) work. The biological work is broken up into validation, routine aging and construction of age-length keys, programming and analyses. Validation is divided into formal validation, i.e., following the detailed procedures of Holden and Raitt (1974) and informal validation, i.e., preliminary examination and determination of the pattern and appearance of marks in otoliths, and the likelihood of marks being annual. Management advice should not be given without formal age validation although preliminary assessment advice sometimes needs to be provided on the basis of informal validation only. If only length-based assessment techniques had been used, the only costs required would have been those needed to carry out fish measurements and some programming and analysis costs, for a total of about \$30,000, i.e., ~10% of the total expended. Total capital costs are not included but Morgan (1983) showed that they were negligible (~\$1,300).

No capability for aging fish by means of daily marks has been established in Kuwait. This needs to be done if all stocks are to be aged successfully, and capital costs will be higher (at least \$10,000).

Training cost was a significant element during the first three years' work, and this is likely to be true elsewhere. One weakness of the Kuwait project appears in the high cost of constructing ALCEKs. This type of work is still carried out manually. The work is slow, demanding and becomes tedious once initial training is over and staff are able to carry out calculations accurately. The work can be drastically reduced if it is computerized, and there is a plan to do this in the near future in Kuwait.

This analysis excludes all costs associated with running the project's research vessel, and establishing and maintaining the land based monitoring system for obtaining catch-and-effort data, both of which are essential and expensive elements in any fisheries management project.

It is expected that the eventual costs associated with aging fish will decline substantially. In 3-4 years, essential data on growth and at least some data on mortality are likely to be available for all species and stocks which are commercially significant or which could become so in the foreseeable future. It will also be possible to choose the species for which ALCEKs need to be constructed routinely. The application of the method for simultaneous analysis of length and age data developed by Morgan (Part I, this vol.) and of the method's likely extensions will probably allow reductions of fish purchases to lower levels in the next 3-4 years.

In retrospect, it is clear that age validation and ALK mortality estimation for older hamoor and hamra, would have been more precise if larger numbers had been sampled. Extra expenditure on long-lived species may be required in the immediate future; scientists planning such work should ensure that abundant fish samples are taken for long-lived species.

Timely Stock Assessment Advice

Optimal use of ELEFAN (and, where appropriate, other size based techniques) and ALK and other age-based procedures must be decided on a stock by stock basis. Where a hitherto unstudied stock needs to be managed, speedy assessments will always be better than none at all and greater accuracy may be sacrificed in the short run so as to obtain quick outputs.

For instance, in Kuwait the possibility of introducing management of several stocks by means of establishing size limits for landings was considered in late 1983, at a time when no age validations had been completed. Only ELEFAN-based analyses were available. Table 3 shows proposed size limits, approximate ELEFAN-based estimates of optimum size at entry, and accurate estimates of optimum size at entry made later on by means of ALK-based assessment. Only the first two columns of data were available at the time. The estimated value of size at entry for nakroor is questionable because neither ALCEKs nor ELEFAN are likely to produce reliable estimates of growth parameters (although this was not known in late 1983), but both estimates for values of optimum size at entry for newaiby are very close, and for zobeidy, ELEFAN-based estimates are likely to be reliable. For the long-lived hamoor, the ELEFAN technique was thought unlikely to provide reliable stock assessments (Pauly 1982), but in fact little difference was found in optimum size at entry estimated by the two techniques. For hamra, ALCEK estimated optimum size differed from the ELEFAN estimate. However, the data shown in the columns A and B of Table 3 justified the need for further

research before the introduction of new size limits, so that the application of the ELEFAN techniques, even to very long-lived fish populations, helped in the management of Kuwait's fisheries. Data of the kind shown in columns A and B of Table 2 are too weak to justify major changes in management policy but may often be sufficient to show the need for prudence in introducing radically new measures without conducting sound supporting research.

Table 3. Optimum sizes^a and size limits^a proposed for various Kuwaiti stocks in 1983. Data in columns (A) and (B) are from Mathews and Samuel (1984b).

Species	Proposed limit A	Approximate optimum size ^c B	ALCEK-based optimum size C	Actual size at entry
Zobaidy	20	16	NA	~18
Newaiby	27	33	~33	~25-26
Nakroor	43	40	NA	~40
Hamoor	45	46	45	~40-45
Hamra	40	26	40-45	~40-45

^aAll lengths refer to cm below and total length, except for Zobaidy, which is measured in terms of fork length.

^bRough estimates based on the size-frequency distributions of the landings.

^cAs obtained from ELEFAN-based analyses.

A Strategy for Optimizing the Benefits of Length- and Age-Based Stock Assessment

Fisheries management studies were initiated in Kuwait in 1978, with the establishment of a series of annual fishing cruises which are being continued. In 1980, monitoring of fish catch and effort was established and annual fishing statistics have been published ever since (Morgan 1981a; Baddar and Morgan 1984; Hakim et al. 1983 and 1984). Previous experience of size-based analyses (Mathews 1974; Morgan 1982b) showed that, useful as they were, they had many weaknesses. At this stage, it was decided to pursue ALCEK stock assessments and at the same time, to attempt to develop new and quicker methods of stock assessment. Because no effort data existed then, Morgan (1982b) established an effort index based on catch variability which allowed approximate surplus yield curves to be constructed. Morgan (1981) also attempted to provide stock assessment advice for Gulf fish stocks on the basis of length-frequency data.

In 1982, the full impact of the contributions of Pauly and David (1981) and Pauly (1982) was felt in Kuwait. They showed that the ELEFAN technique was much more powerful and flexible than other previously used size-frequency based analyses, but their ideas were still untried. A two-pronged strategy was, therefore, adopted; stock assessments by means of ELEFAN I and II would be attempted as soon as possible, starting on Kuwait's main stocks (zobaidy, hamra and newaiby) and traditional ALCEK methodology, already initiated, would be continued until reliable independent estimates were obtained. Provisional stock assessments by means of ELEFAN would be confirmed by ALCEK studies and after some years of using both approaches, choices as to which were most appropriate for each stock would be made. This strategy has been successful.

The Kuwaiti experience suggests that the following approach should be tried wherever an assessment capability is to be established for stocks that are previously unstudied.

1. Establish a catch/effort data collection system suitable to the area (in Kuwait, a system based on interviews of fishermen was the best option).
2. On the basis of one year's worth of catch-and-effort data, consumer preferences, industry and government priorities (if these can be ascertained), choose the main species for which stock assessment is required.

3. Establish a system for collection of monthly size frequencies involving at least 500 fish of each species being studied per month.
4. Carry out provisional ELEFAN analysis after 12 or more monthly samples for each species have been collected.
5. Carry out provisional stock assessments based on ELEFAN. Where ELEFAN-based assessments suggest that the *status quo* is at or near optimum effort and age and size at entry to the fishery, the assessment may be used as a justification for maintaining the *status quo*. Morgan (1985) showed that the zobaidy fishery was harvested at an optimal size and effort and suggestions that size at entry for this fish be increased were successfully resisted.

If the ELEFAN-based assessments were to suggest that a radical change in the management of a fishery is needed, the assessment should be used as a justification for further research, but no major management changes should be made on the basis of the assessment alone. This caution would be particularly necessary if t_{max} calculated by ELEFAN was 5 years or over (Pauly 1982).

6. Conduct formal age validation by means of annual marks in otoliths or other hard parts (Holden and Raitt 1974; Williams and Bedford 1974; Williams 1986).
7. If validation is successful, carry out full stock assessment using ALCEK-based technology, and provide catch curve and cohort based estimates to total mortality, Z. If the results of ALCEK-based stock assessments differ markedly from those of ELEFAN (e.g., hamoor, hamra), reject ELEFAN-based assessments and discontinue application of purely size-based approach (stocks reaching over 5-10 years old). Continue assessments on the basis of annual (or, if necessary semi-annual or quarterly) ALCEKs.

If ALCEK assessments provide essentially similar results to ELEFAN-based analyses, two options exist:

- a. For small stocks, use only ELEFAN-based assessments. For newaiby in Kuwait, the landings of which fluctuate from 350 to 580 t/year, this procedure was adopted. The stock was too small to justify continuous research because cohort (ALCEK) and catch curve (ELEFAN II) estimates of Z for the one year for which sound data are available (1982) showed very similar values (Mathews and Samuel 1985b). The stock was too small and the fish too low-priced to justify continuous research.
- b. If the stock is sufficiently important, repeat ALCEK assessments over several years and substitute cohort mortality estimates for earlier catch curve estimates. After 3-4 years, reappraise the situation. If ELEFAN-based and ALCEK estimates are similar, and if recruitment is relatively constant, ELEFAN-based analyses could be chosen as the basis for assessment provided they are significantly cheaper, but not otherwise. (This procedure, however, can be dangerous because in the event of subsequent changes in landings the biological base necessary for accurate and detailed assessment is likely to be absent; exclusive use of length-based assessments in an important stock is not prudent). If changes or trends in mortality rates are noticed after 3-4 years, ALCEK estimates should be used all the time for stock assessment.
8. If neither ELEFAN-based nor ALCEK-based approaches seem likely to provide reliable assessments (e.g., for nakroor in Kuwait), stock assessment must be suspended until a growth curve may be established on the basis of daily marks occurring in the otoliths. This technique should eventually be applied to any stock which is being assessed on the basis of ELEFAN only, so as to demonstrate conclusively that the ELEFAN-based advice is sound, and to estimate t_0 reliably.
9. To complete ELEFAN-based and ALCEK assessments, Pauly's (1980) technique for estimating M (natural mortality) needs to be applied unless independent methods of estimating M are available. However, independent estimates should be obtained whenever possible.
10. Morgan (Part I, this vol.) has proposed a combination of ELEFAN I and age-based analyses which allows assessment to be made without the catch-and-effort data required for construction of ALCEKs. His technique is likely to be particularly useful in the early stages of

a fishery, when the catch-and-effort data collection system is still new and untried. This new technique is very flexible and it will be applied to and tested on Kuwaiti populations by contrasting results of the new with the traditional techniques; it will also be used to reduce the cost of sampling where possible.

Discussion

The steps followed in developing Kuwait's assessment and management capability suggest that, in essence, the prudent manager will follow a strategy of conservative common sense. He or she may stop after almost any step from 1 to 10; the progression shows a change from quite timely but potentially inaccurate or even erroneous assessments to more and more accurate, but more costly results, which also take longer to achieve. It takes a long time to establish a capability for producing ALCEK-based assessments; however, once established, the capability may provide timely results. It is in the crucial early stages of installation of a capability that ELEFAN- and other size-based techniques will prove most useful, especially in areas where aging of fish can be done only by means of daily marks. In this case ELEFAN I will probably be the standard method, especially if Morgan's (Part I, this vol.) modification of the original approach can be applied routinely.

The Fisheries Management Project in Kuwait has been relatively costly because of the decision to apply traditional and new methods simultaneously, so as to compare them and to ensure maximum precision and prudence in carrying out assessments and in providing management advice. New projects may draw on this experience. For instance, it might be possible to establish an assessment project initially on the basis of length data only and a minimal aging capability. This minimal capability could be used (perhaps on a consultancy basis) to separate stocks reaching up to 4-5 years from all other stocks. Then the project could confine its attention to stocks of fish reaching 1-5 years. At a later stage, once the administrators are convinced of the usefulness of stock assessment advice for the fast-growing fish stocks, the project could extend the aging capability to cover all stocks.

If such a policy is to be considered, one important step should be taken: the minimal aging capability should ensure that all major stocks where assessment is thought eventually to require age-based analyses, should be subjected to as intense a sampling program as feasible. When, after perhaps 4-5 years, a full-scale aging capability is installed, a backlog of biological material will be available for training, validation and routine aging. This material should include stocks of potential as well as actual importance, if at all possible. This step would provide an invaluable perspective to age-based analysis: instead of having to spend 3 years establishing this capability and waiting for the first ALCEK analyses to become available, it would be possible to produce a number of ALCEK cohort-based mortality estimates for different species after the first 2 years of aging work. These results would include 4-5 cohort-based estimates of Z for each species, instead of only one estimate. This relatively speedy response would be likely to produce better, more timely management advice and to obtain more support more quickly from administrators. Again extension of Morgan's (Part I, this vol.) method may allow ALCEKs to be replaced for some stocks of longer-lived fish and so could lead to a reduction in the number of fish to be aged.

Kuwaiti personnel and fish prices are very high. Both personnel costs and fish prices vary enormously from country to country; the technical basis for the different types of stock assessments, however, is likely to remain similar in different areas. Personnel costs given in Table 2 in person-months, may be used together with personnel costs in different areas to estimate the likely overall costs for each area. Total expenditure on fish samples may be estimated by assuming that an absolute minimum of 50 fish/month should be obtained for each fish species whose age is to be validated (100 fish/month would be better) while routine aging may require 500 fish/year for each species, once aging is validated (more will be required for any long-lived species). These numbers, together with local fish prices, can be used to estimate the total budget required for purchasing adequate numbers of fish in different areas. In Kuwait, it is also a market requirement that fresh fish be sold whole and uncut. In some areas, it may be possible to open the body cavities of fish and to remove otoliths without altering the market value; if so, it will be possible to reduce fish purchases to low levels.

Stock assessment studies in Kuwait were started in 1980 and by 1984 size- and age-based assessments of major stocks were available. Management advice was provided in late 1984 on Arabian Gulf fish stocks for the first time (Mathews and Samuel 1984b). However, Kuwait waters are small, and it is probable that most fish stocks are shared with neighboring Gulf countries.

In the long-term, successful management of Kuwait's fisheries without obtaining catch-and-effort data from neighboring Gulf countries will be difficult; only if effort expended on each fish stock in Kuwait is a constant fraction of total effort expended will mortality rates estimated from year to year in Kuwait be unbiased. A tagging program is being established in Kuwait as one means of determining Gulf-wide mortality rates. However, a detailed knowledge of catch and effort expended in the whole Gulf is needed. Furthermore, events in Kuwaiti fisheries will be deeply influenced by events in neighboring fisheries outside Kuwait's control. Nevertheless, Kuwait's demonstrated shrimp and finfish assessment capability provides the first step towards assessment and management for fisheries in the Gulf. With this experience, and with cooperation from other Gulf countries through the Gulf Cooperation Council and the Food and Agriculture Organization (FAO), the technical capability established in Kuwait can be extended to the Gulf stocks and fisheries as a whole; only then will effective management of Kuwaiti and Gulf fish stocks be possible.

This paper does not address the uses of and need for research vessel data in fisheries assessment and management projects. However, in Kuwait, trawl surveys have been carried out since 1978 and will be continued. The surveys have provided data which show that demersal fish population abundance has decreased markedly over the seven-year period studied and these data have been used to provide estimates of the bycatch of the shrimp fleet (Mathews and Samuel 1984a). The species composition of the demersal populations has not yet been studied in detail, but it is likely that this also will change and may affect the whole fish community. A fisheries assessment and management project would be much cheaper and easier to run without a research vessel, but it would provide less useful assessments in the long run.

Acknowledgements

Results of a long-term stock assessment and management project are necessarily based in the work of a whole team rather than the outputs of one or a few individuals. However, acknowledgement must be given to the key roles played by two people in the development of the Fisheries Management Project in Kuwait: Mathew Samuel, who joined us 6 years ago, and who has been responsible for seagoing work, day-to-day land-based monitoring and the aging of fish; and Gary Morgan, who joined 5 years ago, and who has so successfully established the only Population Dynamics Unit in the Arabian Gulf.

References

- Abramson, N.J. 1971. Computer programmes for fish stock assessment. FAO Fish. Tech. Pap. 101. 148 p.
- Baddar, M.K. and G.R. Morgan. 1983. The fisheries of Kuwait, 1981. Kuwait Institute for Scientific Research, Kuwait.
- Baddar, M.K. and G.R. Morgan. 1984. Stock assessment of hamoor, *Epinephelus tauvina*, in Kuwait waters, p. 499-522. In C.P. Mathews (ed.) Proceedings of Shrimp and Fin Fisheries Management Workshop, 9-11 October 1983, Vol. 2. KISR 1366/MB-48. Kuwait Institute for Scientific Research, Kuwait.
- Bagenal, T.B., editor. 1974. The ageing of fish: proceedings of an international symposium. Unwin Brothers Limited, Surrey.
- Bedford, B. 1982. The use of otoliths for ageing Kuwait's fish species. Kuwait Institute for Scientific Research, Kuwait.
- Brothers, E.B., C.P. Mathews and R. Lasker. 1976. Daily growth increments in otoliths from larval and adult fishes. Fish. Bull. 74(1): 1-8.
- Cassie, R.M. 1954. Some uses of probability paper in the analyses of size frequency distributions. Aust. J. Mar. Freshwat. Res. 5: 513-222.
- Christensen, J.M. 1964. Burning of otoliths, a technique for age determination of soles and other fish. J. Cons., Cons. Int. Explor. Mer 29: 73-81.

- Hakim, S., J. Dashti, J. Saif, M. Baddar and M. Samuel. 1983. The fisheries of Kuwait, 1982. Kuwait Institute for Scientific Research, Kuwait.
- Hakim, S., M. Samuel, M. Baddar, A. Bawazeer and J. Dashti. 1984. The fisheries of Kuwait, 1983. Kuwait Institute for Scientific Research, Kuwait.
- Holden, M.J. and D.F.S. Raitt. 1974. Manual of fisheries science. Part 2. Methods of resource investigation and their application. FAO Fish. Tech. Pap. (115 Rev. 1). 214 p.
- Mathews, C.P. 1974. An account of some methods of overcoming errors in ageing tropical and subtropical fish populations when hard tissue growth markings are unreliable and the data sparse, p. 158-166. In T.B. Bagenal (ed.) Ageing of fish: proceedings of an international symposium. Unwin Brothers Limited, Surrey.
- Mathews, C.P. and M. Samuel. 1984a. The stock of "Mixed Fish" in Kuwait, the "By-Catch" and their potential for exploitation, p. 352-386. In C.P. Mathews (ed.) Proceedings of the Third Shrimp and Fin Fisheries Management Workshop, 4-5 December 1982, Vol. 2. Kuwait Institute for Scientific Research, Kuwait.
- Mathews, C.P. and M. Samuel. 1984b. Size limits for fin fish in Kuwait and related considerations. Kuwait Institute for Scientific Research, Kuwait.
- Mathews, C.P. and M. Samuel. 1985a. Ageing and validation of newaiby, hamra and hamoor in Kuwait, p. 56-66. In C.P. Mathews (ed.) Proceedings of the 1984 Shrimp and Fin Fisheries Management Workshop, 15-16 December 1984, Kuwait. Kuwait Institute for Scientific Research, Kuwait.
- Mathews, C.P. and M. Samuel. 1985b. Stock assessment of newaiby, hamoor and hamra in Kuwait, p. 67-115. In C.P. Mathews (ed.) Proceedings of the 1984 Shrimp and Fin Fisheries Management Workshop, 15-16 December 1984, Kuwait. Kuwait Institute for Scientific Research, Kuwait.
- Morgan, G.R. 1981. The fisheries of Kuwait, 1980. Kuwait Institute for Scientific Research, Kuwait.
- Morgan, G.R. 1982a. Methodology for estimating catch and fishing effort in Kuwait's fin fisheries. Kuwait Institute for Scientific Research, Kuwait.
- Morgan, G.R. 1982b. The use of catch variation in fish stock assessment and its application to the fisheries of Kuwait. Kuwait Institute for Scientific Research, Kuwait.
- Morgan, G.R. 1983. Application of length-based stock assessments to Kuwait's fish stocks. ICLARM Newsl. 6(4): 3-4.
- Morgan, G.R. 1985. Stock assessment of zobaidy (*Pampus argenteus*) in Kuwait waters. J. Cons., Cons. Int. Explor. Mer 42: 3-10.
- Pannella, G. 1974. Otolith growth patterns: an aid in age determination in temperate and tropical fishes, p. 28-39. In T.B. Bagenal (ed.) Ageing of fish: proceedings of an international symposium. Unwin Brothers Limited, Surrey.
- Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters and mean environmental temperature in 175 fish stocks. J. Cons., Cons. Int. Explor. Mer 39(3): 175-192.
- Pauly, D. 1982. Studying single species dynamics in a tropical multispecies context, p. 33-70. In D. Pauly and G.I. Murphy (eds.) Theory and management of tropical fisheries. International Center for Living Aquatic Resources Management, Manila, Philippines and Commonwealth Scientific and Industrial Research Organisation, Cronulla, Australia.
- Pauly, D. and N.C. David. 1981. ELEFAN I, a BASIC program for the objective extraction of growth parameters from length frequency data. Meeresforschung 28: 205-211.
- Samuel, M. and A.S. Bawazeer. 1985. A note on the fishery for sobaity, *Acanthopagrus cuvieri*, in Kuwait, p. 142-147. In C.P. Mathews (ed.) Proceedings of the 1984 Shrimp and Fin Fisheries Management Workshop, 15-16 December 1984, Kuwait. Kuwait Institute for Scientific Research, Kuwait.
- Williams, T. 1986. Ageing manual for Kuwaiti fish. Kuwait Institute for Scientific Research, Kuwait.
- Williams, T. and B.C. Bedford. 1974. The use of otolith for age determination, p. 114-123. In T.B. Bagenal (ed.) Ageing of fish: proceedings of an international symposium. Unwin Brothers Limited, Surrey.

Length-Based Methods in Fisheries Research: from Theory to Application

J.A. GULLAND

*Marine Resources Assessment Group
Centre for Environmental Technology
Imperial College of Science and Technology
48 Prince's Gardens
London, SW7 1LU
United Kingdom*

Gulland, J.A. 1987. Length-based methods in fisheries research: from theory to application, p. 335-342. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

This note summarizes, from the personal and possibly prejudiced viewpoint of the chairman, some of the salient points arising from a discussion at the conference on the "Theory and Application of Length-Based Stock Assessments" held on 11-16 February 1985, in Mazara del Vallo, Italy. The background to the meeting—the increasing costs and technical difficulties in making the large number of age determinations required in traditional age-structured approaches to stock assessment, and the need for simpler and quicker methods particularly in developing countries—is first described. This is followed by discussion of the results of the four working groups on sampling, estimation procedures, sensitivity analysis and computer hardware and software.

The basic sampling problems remain those of using an efficient sampling design and of allocating sufficient resources to routine sampling work. Several estimation procedures are now available to estimate growth and mortality, separately or together, from length data, with or without supplementary information. Many methods are highly sensitive to the input data, and confidence regions or similar indications of the range of possible estimates should be given.

Finally, some comments are given on research policy and on the situations where length-based methods are likely to be most useful.

Background

Scientists assessing fish stocks with the aim of advising governments or industry about the status of exploited fish stocks have traditionally followed, singly or in combination, one or two general techniques. These are the use of either production models (treating the population as a

single mass) or analytic models which examine the composition of the population, treating it as the aggregation of the recruitment, growth and mortality of the individual fish. In the dawn of the age of fish population dynamics, some of the earliest analytic models (e.g., Baranov 1918) looked at the composition of the population in terms of length, but since then for more than half a century analytic models have been based almost exclusively on age. There were good reasons for this; the mathematical expressions and the resulting computations are much easier with a parameter (age) that is a simple linear function of time. In most of the fisheries it was possible, using scales or otoliths, to determine the ages of regular large samples of individual fish.

This situation is changing. With the widespread use of computers, simplicity of calculation is no longer essential and for many stocks, particularly in the tropics, it has often proved difficult to determine the age of individual fish. Where seasonal variations are not marked the traditional annual rings are indistinct or absent. Recently developed techniques have shown that it is possible to count daily rings on the otoliths of most fish, but this is very time-consuming. It can only be used for a few fish and is, therefore, a method for determining the general growth rate and checking other estimates rather than for application to regular large-scale sampling.

Increased attention is, therefore, being paid to the use of length-frequency data in analytic assessment. To some extent this was undertaken because of the difficulties of age-structured methods, and was accompanied by a feeling that length-structured methods were in some way second rate, and at best only approximations to proper age-structured methods. However, there is now a recognition that there may also be good theoretical justifications for preferring lengths.

Analytic methods of stock assessment are concerned with the interactions of changes in numbers through recruitment and mortality and changes in weight through growth. Many of the important biological or fishery characteristics, e.g., fecundity or selection by nets with different mesh sizes are much more closely related to size than to age.

This theoretical advantage of size over age might seem to apply most strongly if size is measured in terms of weight, rather than length. However, there are problems with using weight. Measuring the lengths of large samples of fish is quicker and more reliable, especially at sea, than large-scale weight sampling. If there is marked seasonal variation in condition factor, the weight of an individual fish may decrease for a period. Though it seems that Antarctic krill may also shrink in length, this rarely happens for fish. There may be occasions, e.g., handling statistics of commercial shrimp landings recorded in count-per-pound categories, when weight data may be used but in most cases collecting and using length data will be much more convenient.

The time was, therefore, ripe for an examination of the use of length-structured approaches to fish stock assessment, a review of the validity and precision of some of the existing methods of using length composition data and providing advice, particularly to developing countries that are considering the collection and analysis of length composition data on their program of research.

The meeting which had the task of making this examination and review was organized jointly by the International Center for Living Aquatic Resources Management (ICLARM) and the Kuwait Institute for Scientific Research (KISR), which are two of the bodies which have been greatly concerned in length-structured methods. ICLARM has been particularly active in developing length-structured methods and encouraging their use in tropical countries, while KISR is one of the national research institutes where length-methods are being increasingly used as the practical alternative to traditional age-structured methods.

The first part of the meeting was devoted to the presentation and discussion of papers. These covered a listing by FAO staff of a number of questions raised in the possible application of length-based methods in developing countries, descriptions of several of the length-based methods used in stock assessment, examination of the sensitivity of a number of these methods to the assumptions made concerning data and parameter values and discussion of some of the sampling and other practical problems met in applying length-structured methods (see Csirke et al., Part I, this vol.).

During the second part of the meeting the participants divided into four working groups discussing sampling problems, the analytical procedures involved in different methods of assessment using length data, the sensitivity of these methods, and the computer hardware and software involved in their application. Preliminary reports were presented to the meeting at the end of the week. The final reports were agreed to in correspondence after the meeting (see reports Part II, this vol. by Hoenig et al., Shepherd et al. and Majkowski et al.; see also Morgan and Pauly, Part II, this vol.).

Parameter Estimation—General

The process of assessing fish stocks in order to provide advice on, e.g., long-term effects on catch rates and total catches of adding new vessels to a fishing fleet, can seldom be done in a single operation. When using analytic models it is necessary to estimate certain basic parameters of the population (e.g., growth and mortality rates) before using these values in models (e.g., of yield per recruit) to provide the necessary assessments and advice. Some of the stages involved (sampling of catches by research vessels or by the commercial fishing fleet or calculating yield-per-recruit curves) are in principle the same whether age or length is used as the basic input to the analysis.

Growth Alone

The most important change that has to be made in using a length-based approach is in the way that growth is estimated and how von Bertalanffy or other growth curves are fitted. Much of the attention of the meeting was, therefore, focused on the various methods of fitting growth either alone or describing the mean growth pattern of individuals in the population in combination with the estimation of related parameters (mortality or the relative strengths of different year-classes). This should not be taken as implying that in any research program lesser priority should be given to the other steps involved—collection of basic data, studies of the biology of the different species, uses of growth and other parameters to estimate the current state of exploitation of the stock, etc.—but as far as these processes are involved no major new principle is introduced when changing from age to length and the problems involved have been examined at other meetings.

A number of methods of estimating growth were presented (see papers by Pauly, Shepherd and Sparre, Part I, this vol.). The general principles followed were similar. First, individual length samples are examined to identify modes. These modes are taken as representing peaks in spawning (and hence, presuming no big differences in individual growth rates, also in recruitment)—once annually in the case of temperate fish but possibly more in the case of tropical fish. A range of alternative sets of growth parameters are then examined to determine which set best fitted the observed modes. The methods differ in the way of identifying the position of the modes, the growth curve used (e.g., whether a factor for seasonal changes in growth rates is included) and the way in which the fit of data is scored. These differences and other differences as in the power or user-friendliness of the various computer programs are discussed in more detail in Morgan and Pauly (Part II, this vol.).

Under favorable circumstances, with good sampling and clear modes that progress from sample to sample in a consistent manner, all methods work well. With poor data and poor differentiation between modes few work well. For most methods there is usually a banana-shaped plateau covering a range of values of parameters (e.g., the K and L_∞ of the von Bertalanffy equation) over which the scores (fits of data) obtained differ little. Though there will be one specific set that gives the highest score (e.g., the highest ESP/ASP ratio using the ELEFAN I program), there is little significance, in the statistical or biological sense, to distinguish this set from other combinations of parameter values within the central plateau of the response surface.

This is not surprising. The progression of modes gives the absolute rate of growth (as cm per year) over only the range of sizes within which clear modes can be distinguished. This can be small, excluding the pre-recruit fish, and the larger sizes for which the distinction between adjacent modes become indistinct. The rate of growth can be expressed as $dl/dt = KL_\infty (1 - l_t/L_\infty)$ and if l/L_∞ does not vary much, all that can be estimated well is KL_∞ . For example if $l/L_\infty = 0.5$, then the rate of growth will be $0.5 K \cdot L$. To give a proper idea of the range of possible values, and to avoid undue attention being paid to one particular set of parameter values, it is desirable to express the results of any analysis that fits growth curves to length data in the form of response surfaces (see Shepherd, Part I, this vol., Table 2).

Any additional information that can sharpen the peaks in the response surface, particularly one that tends to produce separate estimates of K and L_∞ , or to give rise to a ridge in the response

surface that lies at right angles to the elongated plateau ("banana") generally produced by length-frequency data, is therefore welcome. Guesses of L_{∞} from the sizes of the biggest fish have proved useful. Most useful are age-determinations of the oldest fish, even if they are very few—as might be the case of a program of counting daily rings. It must be admitted that there is a logical impracticability in the preceding sentence. The question of whether age-determinations of the *largest* fish (which are the only ones that can be identified before aging) does in fact tell one much about the oldest fish or only something about the fastest growing fish is not trivial. Methods for the formal incorporation of length-at-age data obtained, e.g., from otolith readings, into length methods such as ELEFAN are described by Morgan (Part I, this vol.). Long-term tagging returns should be useful but short- to medium-term returns may give information only on the rate of growth of fish in the size range already well-covered by length-mode analysis.

During the meeting, little attention was paid to some methods of sharpening the distinction between adjacent modes by considering the departures of an observed length distribution at a particular moment from the typical or average distribution over a long period. This approach is of little value when sampling is poor or covers only a short period but is likely to be particularly useful when there has been a good long-term program of length sampling, and where the year to year changes in year-class strength are large (see, e.g., Gulland 1983, page 86 and example 4.1).

Mortality and Growth

The slope of the right hand side of any length distribution can tell something about the mortality. The higher the mortality, the fewer old (and presumably large) fish, and the steeper the slope of the right hand limb. The principles involved in estimating mortality from a length composition are basically similar to those in using the slope of an age-composition (i.e., a catch curve) but are more complex.

The numbers in a length interval of a given width (e.g., a 2-cm group) depend not only on the mortality rates but also how long, on the average, it takes a fish to grow through the length interval, i.e., how many year-classes are represented in the interval. This will increase as the fish approach their maximum size.

The question of the correct age-length relation to use in this and similar situations is not simple. As pointed out by Sanders (Part I, this vol.) the curve of mean length at a given age is not the same as that of mean age at a given length, especially near L_{∞} , or where there is considerable variability between the growth curves of individual fish. It is, however, not clear which curve, if any, is best to use. The matter was not resolved at the meeting (see Rosenberg and Pope, Part I, this vol.) and for the present any relation should be used with caution. In practice the only relation likely to be obtainable when using length-based methods will be an estimate of mean lengths at age, e.g., from progressions of modes.

In any case, the immediate output from any examination of a length-composition, e.g., mean length or slope of the right hand limb, will be influenced by both growth and mortality. An accumulation of fish near L_{∞} can occur either because mortality (Z) is low, or because the fish grow quickly towards their limiting size (i.e., K is high). More exactly stated, estimates of K and Z are likely to be highly correlated. Analyses of the right hand part of any length composition will be better at estimating the ratio Z/K rather than Z or K separately. This situation is analogous to the progression of modes among the smaller fish, which produce better estimates of the product KL_{∞} than of K or L_{∞} separately.

Because of the interactions of growth and mortality it may be advantageous to use methods of analysis that estimate growth and mortality simultaneously, that is, to examine sets of values of K , L_{∞} , t_0 and Z (and also if necessary, N_t , the strength of recruitment in different years) to determine which set gives the best fit (see papers by Sparre and Pope, Part I, this vol.). This approach might be expected to give some improvements in the estimate of growth, as compared with the simpler methods (see Rosenberg and Beddington, Part I, this vol.). By putting constraints on the probable relative magnitudes of successive modes (from Z), as well as on their location (from K and L_{∞}), it should be possible to reject some of the combinations of growth parameters and their associated modes.

The disadvantage of most of these methods is that they have to make the assumption of steady state, i.e., that Z does not change with time. This is unfortunate, because one of the biggest problems in stock assessment is in determining the impact of fishing on the stock to separate the contributions of natural and fishing mortalities (M and F) to the total mortality (Z). The assessment scientist therefore often hopes for the maximum possible variation in F , and hence in Z , so as to generate the strongest possible signal.

One method that does not depend on an assumption of constant Z is the length-based version of cohort analysis. This follows the same principles as the age-based version and like it can be extended to a multispecies situation to take account of predation (see Pope and Yang, Part I, this vol.). The major practical problem of estimating terminal fishing mortality, natural mortality and composition of the total catch are the same in the length-based and age-based version, though the former also has the problem already mentioned above of knowing the appropriate length-age relation to use in dividing the length-composition up into "cohorts".

Assessments

In general the final stages of producing assessments and advice to managers are exactly the same whether the original data on the composition of the catches or the stock are in the form of length or age. Parameters of growth, selection, mortality, etc. are used to estimate, for example, the effect of increasing fishing effort by 10%. However, there are some occasions when length composition data can be used directly. This has been the case for one method of calculating the effect of changes in mesh size (Gulland 1961). Shepherd (Part I, this vol.) gives a method of predicting next year's stock (and hence catch rate) directly from the length composition in the current year using a modification of a Leslie matrix.

Since these methods eliminate some intermediate steps and consider processes (selection, amount of growth from one year to the next) which are probably more closely related to length than to age, they should be more reliable as well as simpler than methods to obtain the same result using age-based methods. Thus, age-structured methods of predicting the 1986 stock from 1985 data have always suffered when the weight of say, a five-year old fish in 1986 is appreciably larger or smaller than average. To the extent that an unusually large five-year old fish in 1986 was also an unusually large four-year old fish in 1985, and, therefore, grows in accordance with the normal pattern of a fish of that size, this particular problem is avoided. In these and other cases it can be seen that the use of length-structured methods is an improvement rather than an unavoidable approximation to the ideal method.

Sampling

In many developing countries—and in not a few developed countries—the biggest single obstacle to the effective application of length-structured models, or any other kinds of models for that matter, is the absence of adequate samples. When there are doubts about, for example, how to match an observed mode at 15 cm in samples in January to later modes at 23 cm and 28 cm in July, additional samples say in March and May are much better than the most sophisticated models or the most powerful computer.

In general, therefore, the first and most important action to be taken in connection with length-structured methods is to ensure that there is adequate sampling. This will almost always imply an increase in the amount of sampling, often without too much concern about developing a sophisticated sampling design. Provided the sampling is carried out in a sensible manner, being well spread out to give a good coverage in time and space and, if a knowledge of the total catch is important, also coverage of catches by all types of gear, the actual sampling design is not of high priority.

At some stage, however, attention should be given to the design of the sampling scheme. Is enough (or perhaps too much) sampling being done? Is what sampling that is being done carried

out in the most efficient manner? These are not, or at least should not be, new questions. When age-structured models have been used, the basic age-composition data have usually been obtained in two stages by applying an age-length key to an estimated length composition. It has sometimes been assumed that because the numbers of length samples usually greatly exceed the number of age-determinations, the sampling errors in the age sampling contribute the greatest amount to the final variance and thus sampling errors in the length sampling are of a minor importance. This should, however, not be taken to imply that they can be ignored.

If a particular use for length sampling is assumed, e.g., application of the ELEFAN package, then there are methods (e.g., the jack-knife) to determine when further samples add no further information (see Levi et al., Part I, this vol.). This gives an absolute upper limit to the amount of sampling. In practice sampling should probably be stopped earlier at the point at which the additional contribution to the precision of the final output (e.g., management advice) of additional samples is less than the additional precision that would be achieved by putting the resources (e.g., personnel) required for the additional length sampling to other purposes, e.g., improving the statistics of catch and fishing effort.

Some guidance on desirable levels of sampling were suggested by Hoenig et al. (Part II, this vol.) but because of the differences between fisheries in the degree of variability in space and time, the nature of the length composition (one or many modes, clear or indistinct), the costs of sampling, alternative uses of the resources used in sampling and the uses to which the samples and the analyses will ultimately be put, these suggestions can be no more than rough rules of thumb. There are limits below which sampling is so poor that it would be dangerous to draw more than the most tentative conclusions and above which further sampling would be hardly worthwhile. These limits are indistinct and wide and the values suggested are only indications though any research scientist using results from data towards the lower suggested limit should always give serious attention to increasing the intensity of sampling.

Suggestions were also made for the sampling design. These follow well-known principles. Most sampling for length is the result of several stages. Measurements are made of a number of fish taken from one of several boxes or baskets of fish landed by one of several vessels landing at one of several ports on a particular day. Because the between-ship or between-port variance is normally much more than the within-ship or within-port variance samples should be small and frequent.

A lower limit on the size of the individual sample is set by the difficulty in taking a truly random sample from a large pile of fish. The natural bias is to take the bigger fish, though anyone aware of this may over compensate and take too many small fish. This bias is best avoided by sampling complete boxes or baskets of fish. In routine sampling the best sample size may be around 50 fish or less (if this corresponds to the typical number of relatively large fish in a box), but in the early stages of an investigation it may be desirable to let each sample stand by itself. In that case it may be better to have larger samples, e.g., 200 in which the modes, if any, will be clearer.

Research Policy

The meeting showed clearly that the analysis of length composition data is a powerful and potentially reliable method of assessing fish stocks. What then should the director of research do? Should future research work in the field concentrate on massive length sampling, backed up by the use and development in the office of the type of models discussed at this meeting? If so, what additional scientific observations, if any, are needed? Alternatively, bearing in mind the cautions expressed during the meeting, should collection and analysis of length data be given lower priority than other methods? What are the conditions under which one or another approach should be followed. These questions were looked at during the meeting, especially by working group II, and some clear general guidelines emerged.

First, a preliminary set of length samples should be taken, spread as far as possible throughout the fishing season in all major areas and taken from all the important types of fishing gear. This need not be a big task. For a simple fishery without a great variety of gears and not spreading over a large area, one or two samples a month would be perfectly adequate.

This preliminary set of samples should enable each fishery to be placed in one or other of the entries in a 2 x 2 table (see Shepherd et al., Part II, this vol.). The two classifications are whether or not there is a great range of sizes of fish in the fishery and whether or not there are clear modes with a clear progression. If there is a big range in sizes, e.g., the biggest fish are ten times or more the weight of small fish, obviously there should be favorable opportunities for management schemes (e.g., closure of nursery areas or mesh regulation) that protect small fish in order to increase the catches by catching bigger (though fewer) fish later. Such schemes will be directly related to the sizes of fish caught, e.g., on the nursery areas and the output from length-structured analysis is likely to be of immediate value to the manager. If the modes are clear and progress in a reasonable manner, then it should be relatively easy to use age-structured methods.

The most favorable situation is therefore where the size range is wide and the modes clear (Type A, see Shepherd et al., Part II, this vol.). In this situation it should be possible to obtain most of the information needed for stock assessment including growth and mortality estimates. For such a fishery it would be reasonable to set up an extensive program of length sampling for two years (to allow for possible year-to-year differences) in the expectation that the analysis of the data obtained would provide many of the answers required. At the end of two years it would be possible to review progress and identify what important question remain unanswered and modify the program accordingly. It is likely that the length sampling could be reduced perhaps to a level providing no more than routine monitoring with the main attention switched to other lines of research. The choice of line will vary from fishery to fishery.

If the size range is large but modes indistinct (Type B), it will be more difficult to use length-based methods but any output is likely to be useful. A long period of continuing observations has a better chance of producing results (e.g., by following modes or anti-modes due to unusually strong or weak year-classes) than a short period of intense study. Length sampling should therefore be given priority over several years but only on such a scale as can reasonably be carried on without interference with other programs. Among these, priorities should be given to those likely to provide independent estimates of growth, e.g., tagging or otolith readings, even if the latter can only be done by counting daily rings on the otoliths from a few fish.

If there is a single narrow mode that does not progress sensibly and in the extreme case stays in the same position month after month (Type D), use of length-based methods is difficult and can often be misleading. The observed catches are clearly a highly selected sample of the population of fish in the sea either because of gear selection and/or due to the behavior of the fish. [Gear selection is highly unlikely to be the whole explanation. Fishermen will not use a very selective gear unless the fish available to them are already tightly selected by behavior into a narrow length range].

With such selected sampling, examination of length data is not likely to tell one much about the underlying population though some information may be estimated if sampling is very intensive (see Levi et al., Part I, this vol.). Length sampling should probably be given low priority except for occasional samples to check whether the position of the mode is staying roughly the same and attention given to other methods.

Situations in which a single narrow mode progresses in apparently sensible manner (Type C) probably belong to one or other of two distinct subtypes corresponding to either Type A or Type D. Good examples of the former are provided by many penaeid shrimp stocks in which recruitment is concentrated in a short season. At least when heavily fished, few shrimp live to more than one year old so that when the new year-class enters the fishery, few of the previous year-classes are left alive. Nevertheless for a short period there should be two well-separated modes and most of the conclusions for Type A apply. The exception concerns mortality rates. Since each size group will be taken predominantly at one particular time of year, when considering changes in catch per unit effort it may be difficult to separate the effects of mortality from those of changes in seasonal availability.

In this favorable subtype, the progression of modes is discontinuous, or saw-toothed, with a regular increase through most of the year and an abrupt return to a minimum as a new batch of recruits enters (see Ebert, Part I, this vol.). In the other subtype the progression is less logical and only matches the natural growth of a pulse of recruitment for at best part of the year. This is really a less extreme version of Type D and as in that type length-based analysis should be done with caution and collection of length data probably deserves low priority.

References

- Baranov, F.I. 1918. On the question of the biological basis of fisheries. *Nauch Issled. Ikhtiol. Inst. Izv.* 1(1): 81-128. (In Russian)
- Gulland, J.A. 1961. The estimation of the effect on catches in gear selectivity. *J. Cons., Cons. Int. Explor. Mer* 26(2): 204-214.
- Gulland, J.A. 1983. Fish stock assessment. A manual of basic methods. FAO/Wiley Series in Food and Agriculture. Vol. 1. Wiley, Chichester.

**Data Acquisition for Length-Based Stock Assessment:
Report of Working Group I**

J.M. HOENIG*

*Minnesota Department of Natural Resources
Section of Fisheries, Box 12
500 Lafayette Road
St. Paul, Minnesota 55110, USA*

J. CSIRKE and M.J. SANDERS

*Marine Resources Service
Fisheries Resources and Environment Division
Fisheries Department
Food and Agriculture Organization of the United Nations
Via delle Terme di Caracalla
00100 Rome, Italy*

A. ABELLA, M.G. ANDREOLI, D. LEVI and S. RAGONESE

*Istituto di Tecnologia della Pesca e del Pescato (ITPP)
via Luigi Vaccara 61
91026 Mazara del Vallo, Italy*

M. AL-SHOUSHANI and M.M. EL-MUSA

*Kuwait Institute for Scientific Research
P.O. Box 24885, Safat
Kuwait*

Hoenig, J.M., J. Csirke, M.J. Sanders, A. Abella, M.G. Andreoli, D. Levi, S. Ragonese, M. Al-Shoushani and M.M. El-Musa. 1987. Data acquisition for length-based stock assessment: report of working group I, p. 343-352. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

Sampling design considerations and sampling effort for length-based stock assessment are often as demanding as those for traditional age-based assessment. Four factors have been identified which should be considered in designing a program for collecting length-frequency data: 1) period of time over which data are aggregated, 2) width of length interval for grouping data, 3) span of time over which data are collected (i.e., length of study) and 4) uniformity of sampling frequency. Guidelines are presented for collecting length data.

*Present address: RSMAS/CIMAS, University of Miami, 4600 Rickenbacker Causeway, Miami, Florida 33149, USA.

Introduction

Working Group I on data acquisition was requested to formulate advice on sampling procedures for users of length-based assessment methods. The guidelines which follow are based on a review of the data requirements of the principal methods. No attempt was made to formulate specific sampling designs. For this the reader is referred to standard fishery manuals (Gulland 1966; Banerji 1974; Bazigos 1974, 1976; FAO 1981) and sampling texts (Mendenhall et al. 1971; Cochran 1977).

Length-based assessment methods are sometimes proposed as simpler and shorter-term alternatives to assessment methods based on age frequency and catch-per-unit-effort data. However, in terms of sampling design and required sampling effort, length-based assessment is often as demanding as age-based assessment and can be equally or even more demanding than assessment based on catch-per-unit-effort information.

The length structure of a population is likely to vary with geographic area, habitat, and over time due to behavioral factors and/or the dynamics of the population. Sampling schemes must produce results which are representative of the statistical population of interest. This means that proper weights and procedures should be used to combine samples. For example, in a research trawl survey designed to obtain an estimate of the mean length in the population using stratified random sampling, samples from different strata should be combined using weights which reflect the size of the strata. Thus, estimates from different strata should be combined using weights proportional to the product of (geographical extent of stratum) \times (catch per unit effort in stratum). If sampling effort in each stratum is proportional to the geographical extent of the stratum (i.e., if the same number of hauls is made at each station, a constant percentage of each haul is measured, and the number of hauls is proportional to the area of the stratum) then the estimates are self-weighting and the CPUE is (implicitly) accounted for. Similarly, in studying the length composition of the commercial catch, length frequencies from different sampling units should be weighted by the total catches in the units. Thus, length-based assessment does not necessarily avoid having to collect catch and catch-per-effort data.

Use of modal analysis to estimate growth parameters can impose a different set of constraints on the collection of length-frequency data. For example, a species which occurs over a wide geographical range may have different growth rates and/or different spawning periods in different areas. Combining samples from different areas may obscure or create additional modes. Furthermore, behavioral characteristics of some species lead to schools of fish with nearly uniform body size. This means that a given cohort of fish might separate into schools on the basis of body size, and a given school of fish may contain several age groups, all of which have approximately the same body size (Fréon 1984, 1985). In this case it may be necessary to define the statistical population of interest in a narrow manner as a subset of the biological population of ultimate concern.

From the above, it can be seen that sampling for length frequencies requires care. The intended use of the data has a direct bearing on how the data should be collected. Studying variability between samples is of interest for its own sake but should also be used to define the population under study.

The preceding discussion is not intended to discourage the use of length-based assessment but rather to caution the reader about some of the pitfalls. The main goal should be to measure some fish and make assessments. Improvements in sampling design can be made on an ongoing basis.

Strategies

The first steps in setting up an assessment program are generally to review the available information, formulate goals and design a research program. A pilot survey is extremely useful in refining or altering the above and can be used to help identify minimum data requirements for both immediate and long-term assessments and to design the sampling programs. For example, a few length-frequency samples may suggest whether modal analysis is likely to be useful (see Guidelines section).

Different methods of assessment require different data inputs. The assessment methods to be tried should be carefully considered in designing the research program. It is well to collect data with future data needs in mind.

If no pilot survey is available to aid in designing sampling and research programs, the course of action will be dictated by the urgency of the need for information and the availability of personnel and funds. One strategy is to begin by collecting data for the most demanding assessment procedures and models and assess the utility/adequacy of the data as they accumulate. This might include collecting otoliths in case they are needed in the future. This strategy may be expensive but it enables the research program to begin immediately and maximizes the chances for success. An alternative strategy, when personnel or other resources are limited, is to attempt to obtain the minimum data required for assessment and then use this information to design future research programs.

Research programs for assessment should in general have the additional goals of learning about the variability in the fish population and the fishery. This allows one to evaluate the efficiency of the sampling design, to relate variability in sampling results to variability in assessment results, and ultimately to formulate rules of thumb for deciding when an assessment might work, how much sampling effort is needed and how sampling should be conducted.

Variability in the system can be studied using analysis of variance and variance components to estimate effects of various factors (and interactions of factors) and to test for significance. For this, the sampling scheme used must be well specified and should be as simple as possible. The jackknife procedure can also be used to analyze complicated survey designs. (For an introduction to the jackknife technique, see Sokal and Rohlf 1981; Pauly 1984; for advanced theory see Gray and Schucany 1972; Miller 1974; Efron 1982; for applications in fishery biology and ecology see Levi et al., Part I, this vol.; Smith 1980; Heltshe and Forrester 1983; Pauly 1984).

Survey results can be evaluated as sampling is being conducted. For example, for estimating mean length of the catch in a fishing port, one can continue visiting ships (chosen at random) until the estimate becomes stable. (Again, the sampling scheme must be considered in the evaluation).

It is difficult to know what precision in length data is required to have a given precision in various final assessment estimates. Levi et al. (Part I, this vol.) proposed a jackknife-like procedure to determine when enough samples have been collected that ELEFAN (L_{∞} and K) estimates are stable. The key point is that no simple statistic has yet been found which predicts the stability of ELEFAN I estimates. As experience accumulates, it may be possible to formulate simple rules of thumb for predicting stability from the coefficients of variability of various simple ancillary statistics like the mean, mode, median, etc.

When the uses for the data are not clearly specified in advance, it is impossible to know what precision is required. The group felt that a rough rule of thumb might be to attempt a 10 to 50% coefficient of variability in the percentage of fish in each length interval of interest. The acceptable range of variability will depend on the relative importance of the species. This guideline is arbitrary and is intended as a starting point rather than as a fixed rule. It reflects the precision the group felt might be attainable rather than the expected reliability of the assessment.

Mathews (Part I, this vol.) and Morgan (Part I, this vol.) discuss strategies for designing and modifying assessment programs which utilize ELEFAN I and II as key components. They suggest situations in which supplemental length-at-age data are indicated or essential.

Data Requirements

Data requirements of the length-based assessment methods must be considered when designing a sampling program. The following factors have been identified as important: 1) aggregation period; 2) length interval; 3) time span of study; and 4) uniformity of sampling frequency.

When sample sizes are small, it may be necessary to pool or aggregate samples collected over a period of time in order to have an adequate sample. Similarly, if it takes a long period of time to cover a geographic stratum then the stratum sample will be smoothed over time. For modal analysis, this means that modes will become progressively more blurred as the aggregation period is increased.

For some types of analysis it is important to record lengths either as raw measurements or in a histogram with narrow length intervals. For example, the results from the ELEFAN I program will vary as the width of the length interval is varied. Since it is currently impossible to know in advance which interval width is best, it is best to preserve the option of experimenting with intervals of different widths by recording data as precisely as possible. Also, the validity of certain approximations in length-converted catch curve and length cohort analyses depends on the width of the length

classes used (see Pauly 1984). To avoid the necessity of using iterative methods for these analyses, the length intervals must be kept narrow. (Increasing the number of length intervals does not however, improve the backward convergence of a VPA or cohort analysis (Jones 1984)).

Some methods require length data representing the average composition over an annual (or other) cycle. For example, in length-based cohort analysis, the total catch by length group is required. This means that sampling must be carried out over the entire cycle (though not necessarily with constant sampling effort if proper sampling and weighting is used). For modal analysis, there can be temporal gaps in sampling but the time span of the study should be long enough to allow examination of the progression of modes. Similarly, Ebert's (Part I, this vol.) and Damm's (Part I, this vol.) regression techniques require observations of mean length to be obtained at several times.

Data requirements of four classes of assessment models are summarized in Table 1.

Table 1. Sampling requirements for four classes of length-based assessment model.

Mean length type mortality estimators	Modal analyses for growth study
<ul style="list-style-type: none"> — Length-frequency data representative of the fish population are needed. — Total catch data are not needed. — Sampling plan depends on model to be used: <ul style="list-style-type: none"> ● Aggregate over a year for Beverton-Holt continuous spawning model. ● Sample one month for discrete spawning model. ● Sample several times spread out over a year for Ebert-Damm methods. Samples do not need to be uniformly spaced. ● See also section on sampling shrimp landings. 	<ul style="list-style-type: none"> — Samples representative of the fish population or commercial catch are not needed provided the modes are evident and representative of age classes. (For mortality estimation, the relative frequencies must represent the population.) — Total catch data are not needed. — A single sample can be used, but it is better to have sequential samples. Samples need not be taken according to any particular time schedule. — Samples should not be pooled over long periods of time as this blurs modes. — Lengths should be measured and recorded precisely.
<ul style="list-style-type: none"> — Length-converted catch curves — Length-frequency data representative of the fish population are needed. — Total catch data are not needed. — Continuous spawning model is assumed; thus, it is best to sample over a whole year (or longer) to smooth curve. — If lengths are grouped into narrow length intervals then simple, approximative methods can be used to estimate Z; otherwise, for wide intervals iterative techniques should be used (Pauly 1984). 	<ul style="list-style-type: none"> — Length-based VPA and cohort analysis — Samples representative of the catch must be taken. — Total catch data (including discards) are needed. — Estimated catches at length should pertain to entire annual (or other) cycle. — If lengths are grouped into narrow length intervals then the cohort analysis approximation can be used; otherwise, for wide intervals the iterative length-based virtual population analysis should be used (Pauly 1984). — The number of length classes used (i.e., the width of the intervals) does not affect the convergence properties of the cohort analysis (Jones 1984).

Guidelines

The following guidelines apply in general:

- 1) Collect and preserve raw length measurements if possible (i.e., length of individual fish). A stem-and-leaf plot (Tukey 1977; Chambers et al. 1983) is a convenient way to record data in great detail with a minimum of writing (see Fig. 1) and enables a quick, visual analysis of the data at the time of collection. If data must be collected as a histogram, use as narrow a length interval as possible. It is always possible to regroup data into longer length intervals at the time of data analysis. If data are stored as summaries of pooled

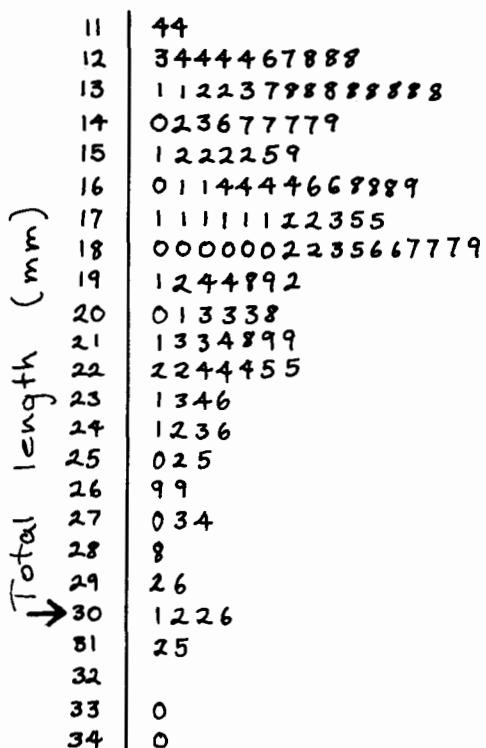


Fig. 1. Example of a stem-and-leaf diagram for recording detailed length-frequency data in a compact form. Digits to the left of the vertical line (stems) are the leading digits of the observations; digits to the right (leaves) are the individual trailing digits. Thus, the line marked with an arrow represents the four lengths: 301, 302, 302 and 306 mm. This diagram contains 184 digits whereas writing out each number would require 408 digits. See Chambers et al. (1983) for variations.

samples, one loses the opportunity to arrange the data into shorter length intervals and to pool data over different periods so as to, for example: a) try new methods with more demanding data requirements; b) study further hypotheses that may arise; c) study variability in the system and evaluate changes in sampling design.

- 2) When collecting commercial catch data, document the fishery including, if possible, target species, number and type of gear, mesh size, fishing location and quantity and description of discards. Individual vessel trip catch rates may also be collectible. When discards are substantial (e.g., Saila 1983) gear selectivity parameters may not be adequate for adjusting length-frequency distributions. The sampling scheme used should also be documented.
- 3) Variability between sampling units is almost universally greater than variability within units. Thus two boxes examined from each of two ships is generally better than four boxes examined from one ship.
- 4) Provide the person doing the sampling with simple, objective rules for choosing sampling units so that errors and biases are kept to a minimum. For example, if only half of each

box of fish is to be measured, have the sampler always examine the right side of the box. If fish are examined at sea, the sampler can draw randomly placed circles on the deck and measure only those fish which fall on a circle. Care must be exercised in formulating rules and sampling instructions should be simple and clear. Instructions should be pretested before being employed in an actual sampling program. In one case, port samplers were told that a minimum of 50 fish had to be measured for the data to be usable. The samplers therefore only measured certain fish species when they could be sure of obtaining 50 lengths. This happened to occur primarily when there were a lot of small fish present which biased the results (D. Pauly, pers. comm.).

- 5) Monitor sampling results regularly (e.g., by graphing the data). If data are plotted soon after collection, one can modify the sampling plan in response to unexpected findings. For example, one might wish to increase the sampling frequency to be able to monitor the progression of an unusual recruitment event (a peak) through a series of length-frequency samples. Whenever possible, one should graph the data at the time of analysis to get a feel for the relationships and to visually check assumptions. The most common types of plot for length-frequency analysis are the plain histogram and the histogram smoothed by a running average. An alternative method of smoothing and presenting the data would be a density trace (see Chambers et al. 1983).

For modal analysis, the following guidelines are suggested:

- 1) Normally, one can pool samples collected within a given time period (e.g., a month, a quarter year) to boost sample sizes, but for very small (i.e., short-lived) species like shrimp one may need smaller time intervals to avoid missing important events in the life history. Thus, use an estimate of the life span or, if not available, the maximum length as a guide to the aggregation period.
- 2) For protracted spawners, it is most important to keep the data summarized by short time intervals since modes will be more subtle and aggregating samples over time blurs modes.
- 3) Usually, the bigger the maximum size the more age groups and, hence, the more length intervals should be used to record the data. (During analysis, length intervals can always be combined if desired.)
- 4) The larger the maximum size the larger the sample should be (though this is admittedly difficult).
- 5) Check for heterogeneity among different samples of schooling fish and fish which occur over a wide geographical area (see Introduction).
- 6) The group felt that the heuristic sampling recommendations shown in Table 2 (modified from Pauly 1984) appear to be appropriate. Mathews (Part I, this vol.) also recommends monthly samples of 500 lengths be collected for 12 months or more.

Table 2. Suggested sample sizes for modal analysis. (After Pauly 1984 based on Munro 1980).

Total sample size (no. fish)	Time (in months) over which data for total sample were accumulated*				
	1	2	4	6	12
1 — 99	0	0	0	0	0
100 — 499	0	0	1	2	2
500 — 999	1	1	2	3	4
1,000 — 1,499	1	2	3	4	5
1,500 — ∞	2	3	4	5	5+
0 = not usable		2 = fair		4 = very good	
1 = poor		3 = good		5 = excellent	

*It is here assumed (1) that the samples cover a wide range of lengths, (2) that gear selection is accounted for and (3) that the sizes of the monthly samples are more or less equal if the total sample is accumulated over more than one month.

- 7) A better idea of the likelihood of success of modal analysis can be obtained if one or more preliminary samples are available. Four patterns may be identified (see also Shepherd et al., Part II, this vol.):
- a single mode is evident which appears stationary over time;
 - a single mode is evident but the mode seems to progress over time;
 - the length distribution is clearly polymodal;
 - the length distribution appears to be vaguely polymodal or nondescript.
- Pattern (c) is the ideal situation for modal analysis and requires few samples. Pattern (b) is usable but requires samples to be collected over a longer span of time to see the modal progression. Pattern (d) and to a lesser extent (a) can also be utilized but with more difficulty. If samples are taken over a period that is long enough, one may observe an unusual recruitment event (i.e., extra high or low recruitment) which puts a "marker" in the length distribution. This marker may then be followed over time.
- 8) The bigger an individual sample, the more likely it is to reflect accurately the population length frequency (assuming sampling is done in an appropriate manner). However, two samples taken, say, six months apart describe two different statistical populations. Both populations are important for the ELEFAN I analysis and the samples could be given equal weights even though one sample may provide more precise information than the other. The easiest and probably the most satisfactory procedure would be to use approximately uniform sample sizes (and uniform weighting). However, sampling dates separated from other sampling dates by large temporal gaps should probably receive extra sampling effort since isolated dates will tend to have a high influence on the ELEFAN I estimates.
- 9) Users of modal analysis techniques should be aware that immigration or emigration of segments of a population can change the shape of a length-frequency distribution in a manner that may resemble modal progression due to growth and mortality. This is difficult to guard against but may be approached by collecting and examining separately samples from different areas.

Multispecies/Multiple Objectives

In general, it will not be possible to design a sampling scheme which is optimal for each of several objectives (e.g., species) or even for some clearly specified but complex goal involving several objectives. In these situations one may be best off stratifying the sampling simply for convenience (e.g., for logistic or administrative reasons) and using proportional allocation. One can then use post-stratification to increase precision of estimates. Cochran (1977) points out that the strata to be used in post-stratification must be decided upon in advance to avoid human bias.

Strategies for the dual goals of sampling lengths and ages are discussed in Morgan (Part I, this vol.) and Mathews (Part I, this vol.). Age data appear to be most critical for long-lived species which can be identified by having a longevity of five or more years predicted by preliminary ELEFAN I (or other modal) analysis or on the basis of life history characteristics (large size, predaceous habits, etc.).

Age data can be incorporated into an ELEFAN I data set (Morgan, Part I, this vol.), used to validate or extend ELEFAN I results, or used to develop an age length key. If large fish are examined so that there is a reasonable estimate of longevity, the longevity can be used to estimate mortality directly (Alverson and Carney 1975; Hoenig and Lawing 1983; Hoenig 1983).

The Special Case of Commercial Size Categories

The catch statistics for some groups of animals, such as shrimp, are often reported as weight landed in each of several commercial grades, e.g., number of tails per pound or kilogram. These statistics may constitute an important part of the available database. Thus, it is important to consider how to collect and analyze this type of catch data.

When possible, the best procedure would be to use stratified random sampling and sample each commercial size category (stratum) independently for length frequencies. An overall length-frequency distribution could then be obtained by combining the length frequencies from each

stratum using weights proportional to the size of the stratum (i.e., proportional to the quantity landed in the stratum).

The simplest method would be to estimate the mean weight in the catch by

$$\bar{w} = \frac{\sum w(i)}{\sum w(i) s(i)} \quad \dots 1)$$

where $w(i)$ is the total weight in the i -th size category and $s(i)$ is the number of tails per unit weight in the i -th size category.

The sample mean weight can then be equated with the expression for mean weight of animals older than some age t_c

$$\bar{w} = ZW_{\infty} \sum_{n=0}^3 \frac{U_n \exp(-nK(t_c - t_o))}{Z + nK} \quad \dots 2)$$

where W_{∞} , K and t_o are the von Bertalanffy growth parameters, Z is the total instantaneous mortality rate, t_c is the mean age of first capture, and $U_n = 1, -3, 3, -1$ for $n = 0, 1, 2, 3$, respectively (Beverton and Holt 1957). This is solved iteratively for Z as suggested by Munro (1980) and Pauly (1984). Since weight at any time t is

$$W_t = W_{\infty} (1 - \exp(-K(t - t_o)))^3 \quad \dots 3)$$

it follows that the weight at time t_c is

$$W_c = W_{\infty} (1 - \exp(-K(t_c - t_o)))^3 \quad \dots 4)$$

This can be solved for $K(t_c - t_o)$ giving

$$K(t_c - t_o) = -\log_e(1 - (W_c/W_{\infty})^{1/3}) \quad \dots 5)$$

Substituting the above into the expression for mean weight gives

$$\bar{w} = ZW_{\infty} \sum_{n=0}^3 \frac{U_n (1 - (W_c/W_{\infty})^{1/3})^n}{Z + nK} \quad \dots 6)$$

which eliminates the need to know t_c and t_o .

There are two difficulties with this approach. First, like other estimators based on sample moments, it is not possible to check assumptions with a goodness-of-fit test or by graphing a relationship. Second, it is difficult to know what value of W_c to use. One possibility is to eliminate the smallest size group and use W_c corresponding to some value between 1/(size group eliminated) and 1/(new first size). Regardless of the value of W_c used, the resulting mortality estimates will be a function of the true mortality rate. That is, the mortality rates computed as above can be

used as an index of mortality provided the selectivity of the gear and the value of W_c used in the computations are not changed. An index of mortality is sufficient for some assessment models (Hoenig, Part I, this vol.). Mortality estimates from the mean weight method should not be used in conjunction with estimates obtained by other methods without first performing a sensitivity analysis of the importance of W_c .

It is tempting to try to convert mean weights in size classes to mean lengths using the allometric formula. Since weight is not a linear function of length, converting the mean weight to length does not give an unbiased estimate of mean length (Pienaar and Ricker 1968; Houghton and Flatman 1978; Nielsen and Schoch 1980). However, over a short interval of weight, length is an almost linear function of weight and thus converting shrimp weight classes to length may not cause a serious bias provided there are enough weight categories. The bias will depend on how finely the shrimps are graded and how accurately the shrimps are sorted into the appropriate weight categories.

References

- Alverson, D.L. and M.J. Carney. 1975. A graphic review of the growth and decay of population cohorts. *J. Cons., Cons. Int. Explor. Mer* 36(2): 133-143.
- Banerji, S.K. 1974. Frame surveys and associated sample survey designs for the assessment of marine fish landings. FAO Indian Ocean Programme IOFC/DEV/74/39: 15 p.
- Bazigos, G.P. 1974. Design of fisheries statistical surveys—*inland waters*. FAO Fish. Tech. Pap. 133.
- Bazigos, G.P. 1976. The design of fisheries statistical surveys—*inland waters*. Populations in non-random order, sampling methods for echo surveys, double sampling. FAO Fish. Tech. Pap. 133, Suppl. 1, 46 p.
- Beverton, R.J.H. and S.J. Holt. 1957. On the dynamics of exploited fish populations. *Fish. Invest. Lond.* 2(19), 533 p.
- Chambers, J.M., W.S. Cleveland, B. Kleiner and P.A. Tukey. 1983. Graphical methods for data analysis. Duxbury Press, Boston, 395 p.
- Cochran, W.G. 1977. Sampling techniques. Third edition. John Wiley & Sons, New York. 428 p.
- Efron, B. 1982. The jackknife, the bootstrap and other resampling plans. CBMS-NSF Regional Conf. Series in Applied Math. Soc. Indus. Appl. Math., Philadelphia. 92 p.
- FAO. 1981. Methods of collecting and analyzing size and age data for fish stock assessment. FAO Fish. Circ. 736. 100 p.
- Fréon, P. 1984. La variabilité des tailles individuelles à l'intérieur des cohortes et des bancs de poissons 1: Observations et interprétation. *Oceanologica Acta* 7(4): 457-468.
- Fréon, P. 1985. La variabilité des tailles à l'intérieur des cohortes et des bancs de poissons 2: Applications à la biologie des pêches. *Oceanologica Acta* 8(1): 87-99.
- Gray, H.L. and W.R. Schucany. 1972. The generalized jackknife statistic. Marcel Dekker, Inc., New York. 308 p.
- Gulland, J.A. 1966. Manual of sampling and statistical methods for fisheries biology. Part 1. Sampling methods. FAO Manual Fish. Sci. 3. 87 p.
- Heltshe, J.F. and N.E. Forrester. 1983. Estimating species richness using the jackknife procedure. *Biometrics* 39: 1-11.
- Hoenig, J.M. 1983. Estimating total mortality rate from longevity data. University of Rhode Island, Kingston. 92 p. Ph.D. dissertation.
- Hoenig, J.M. and W.D. Lawing. 1983. Estimating the total mortality rate using the k largest order statistics for age. ICES C.M. 1983/D: 22, 16 p.
- Houghton, R.G. and S. Flatman. 1978. A bias in calculating mean weight from a mean length and a discussion of the method used in working groups. ICES C.M. 1978/G: 18. 9 p.
- Jones, R. 1984. Assessing the effects of changes in exploitation pattern using length composition data (with notes on VPA and cohort analysis). FAO Fish. Tech. Pap. 256. 118 p.
- Mendenhall, W., L. Ott and R.L. Scheaffer. 1971. Elementary Survey Sampling. Duxbury Press, Belmont, California, USA.
- Miller, R.G. 1974. The jackknife—a review. *Biometrika* 61: 1-17.
- Munro, J.L. 1980. Stock assessment models: applicability and utility in tropical small-scale fisheries, p. 35-47. In S.B. Saila and P.M. Roedel (eds.) Stock assessment for tropical small-scale fisheries. University of Rhode Island, Kingston. 198 p.

- Nielsen, L.A. and W.F. Schoch. 1980. Errors in estimating mean weight and other statistics from mean length. *Trans. Amer. Fish. Soc.* 109: 319-322.
- Pauly, D. 1984. Fish population dynamics in tropical waters: a manual for use with programmable calculators. ICLARM Studies and Reviews 8. International Center for Living Aquatic Resources Management, Manila, 325 p.
- Pienaar, L.V. and W.E. Ricker. 1968. Estimating mean weight from length statistics. *J. Fish. Res. Board Can.* 25: 2743-2747.
- Saila, S.B. 1983. Importance and assessment of discards in commercial fisheries. FAO Fish. Circ. No. 765. 62 p.
- Smith, S.J. 1980. Comparison of two methods of estimating the variance of the estimate of catch per unit effort. *Can. J. Fish. Aquat. Sci.* 37: 2346-2351.
- Sokal, R.R. and F.J. Rohlf. 1981. Biometry. Second edition. Freeman and Co., San Francisco, USA.
- Tukey, J.W. 1977. Exploratory data analysis. Addison-Wesley, Reading, Massachusetts, USA.

**Methods of Analysis and Assessment:
Report of Working Group II**

J.G. SHEPHERD

*Ministry of Agriculture, Fisheries and Food
Fisheries Laboratory
Lowestoft, Suffolk NR33 OHT
United Kingdom*

G.R. MORGAN*

*Kuwait Institute for Scientific Research
P.O. Box 24885, Safat
Kuwait*

J.A. GULLAND

*Marine Resources Assessment Group
Centre for Environmental Technology
Imperial College of Science and Technology
48, Prince's Gardens
London, SW7 1LU
United Kingdom*

C.P. MATHEWS

*Mariculture and Fisheries Department
Kuwait Institute for Scientific Research
P.O. Box 24885, Safat
Kuwait*

Shepherd, J.G., G.R. Morgan, J.A. Gulland and C.P. Mathews. 1987. Methods of analysis and assessment: report of working group II, p. 353-362. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

A classification of methods for the analysis of length-frequency data, and their main advantages and disadvantages are discussed.

Two broad classes that emerged were: (1) *analytic* methods (i.e., methods for the estimation of growth from shifts in length-frequency distributions, estimation of mortalities from catch composition

*Present address: c/o W.A. Marine Research Laboratories, P.O. Box 20, North Beach 6020, W.A., Australia.

data, etc.) and (2) *synthetic* methods (i.e., yield-per-recruit models and catch prediction models). A comparison of length-based and traditional (i.e., age-based) methods is presented. Recommendations are given regarding the choice of analytical methods appropriate for four different, idealized types of length-frequency data sets.

Introduction

Methods for the assessment and utilization of length composition data may be divided into a broad classification according to their purpose. First, there are methods which are essentially *analytic*, i.e., intended for the determination of vital parameters such as rates of growth and mortality, and the elucidation of population structure. Second, there are methods which are primarily *synthetic*, i.e., those intended for the combination of raw data and vital parameters to produce estimates of quantities of interest for practical purposes such as assessment of catch rates, the effects of mesh changes, determination of long-term yields, etc.

The analytic methods in practice may be subdivided usefully into those which determine only growth parameters, and those which also determine quantities related to population size and its changes (mortality). The second class of synthetic methods may also be subdivided into those which assume or are only applicable to steady-state calculations, and those which are intended to take account of transient effects (such as short-term catch forecasts).

In practice, and largely for historical reasons, the first subclass in each case (i.e., analytical methods for determination of growth parameters, and steady-state methods for stock assessment) has been more highly developed, but there is no compelling reason for this situation, and it is now changing quite rapidly. Several papers presented here document this recent trend towards more balanced development of the whole spectrum of potential methods (see, e.g., Sparre, Part I, this vol.; Shepherd, Part I, this vol.; Pope, Part I, this vol.).

The methods of assessment available are summarized in some detail in Table 1. A more detailed discussion of the most important points is given below, concentrating particularly on those which emerged during the discussions at this conference.

Methods of Assessment: Analytic

Until recently the principal methods available for the determination of growth parameters from length compositions were those based on the method of mixtures, originally due to Hasselblad (1966), developed and described by Abramson (1971), but epitomized by the work of MacDonald and Pitcher (1979). These essentially analyze the content of length compositions without reference to biological constraints and, thus, whilst versatile require considerable interpretative skill and care if they are to provide credible results. The only significant competitor to these techniques (other than subjective analysis of modal sequences using pencil and paper) has been the method developed by Pauly and David (1981) known as ELEFAN I. This is quite different (see Pauly, Part I, this vol.), relying on the optimization of a criterion of goodness-of-fit of assumed growth parameters and the data, based on coincidences of expected and observed modes.

More recently Schnute and Fournier (1980) and Fournier and Breen (1983) have developed an extended version of the method of mixtures which constrains the modal sequence to conform (more or less) to a specified growth model, thereby greatly stabilizing the technique and reducing the scope for error. Their work has been further developed by Sparre (Part I, this vol.) and Pope (Part I, this vol.) who have allowed for the analysis of several (sequential) length compositions simultaneously and, thus, also for the reliable estimation of cohort sizes as well as mortality. This work brings the method of mixtures to a high state of development and the results obtained seem to be quite reliable. These methods are characterized by being quite highly parametric (requiring a

Table 1. Characteristics of some methods for analysis of length-frequency data.^a

Author and/or method	Parameters estimated						Assumptions and data requirements								Hardware requirements		Remarks	
	K	L _∞	t ₀	Seasonal growth	Z	L _c	Recruitment pattern	strength	Stable age dist.	Gen. v. Bert. growth	Normal length at age	Pulse season recruit- ment	Samples needed at least one	regular	C/f or C at L	Other input parameters		
ELEFAN I	+	+	+	+	—	—	—	—	N	Y	N	Y	Y	N	N	—	Micro	Needs seasonal samples to estimate seasonal growth-ELEFAN I and II provide most information for Y/R assessments
ELEFAN II	—	—	—	—	+	+	+	—	Y	Y	N	N	Y	N	N	Growth	Micro	
ELEFAN III (VPA II)	—	—	—	—	+	+	+	+	Y	Y	N	N	Y	N	Y	Growth, M	Micro	Estimates F at length
ELEFAN III (VPA III)	—	—	—	—	+	+	+	+	N	Y	N	N	N	Y	Y	Growth, M	Micro	Estimates F at length and F over time
ELEFAN IV	—	—	—	—	—	—	+	—	Y	Y	N	N	Y	N	N	Growth, selec. curve	Micro	Estimates M (& s.e.)
Ebert/Damm	+	+	—	—	+	—	—	—	Y	Y	Y	Y	Y	N	N	\bar{L}, L_c	Micro	Mean length instead of length distribution requires parameter estimates dependent on number of samples
Wetherall et al.	+	+	—	—	+	—	—	+	Y	Y	Y	N	Y	N	N	—	Micro	A variety of methods based on stable age distribution
Sparre	+	+	+	+	+	+	+	—	N	Y	Y	Y	Y	N	Y	CPUE at L	Large	Also requires estimate of number of cohorts in a sample
Pope I	+	+	+	—	+	+	—	+	N	Y	Y	Y	N	Y	Y	CPUE at L	Large	Simultaneous estimate of parameters
Pope (ANOVA)	+	+	+	—	+	—	+	+	N	Y	N	Y	N	Y	Y	CPUE at L	Large	Estimate Z by length group
Shepherd's SRLCA	+	+	+	—	—	—	—	—	N	Y	N	Y	Y	N	N	—	Micro	Seasonal growth extension may be possible
Hoenig	—	—	—	—	—	—	—	—	n.a.	N	N	N	N	N	Y	Yield, Z	Calculator	Estimates (Z/K) or Z; estimates yield/Z or yield/F given M
Phalanx	—	—	—	—	+	+	—	—	Y	Y	N	N	N	Y	Y	Growth, food consumption	Large	Multispecies model

Continued

Table 1. Continued

Author and/or method	Parameters estimated						Assumptions and data requirements						Hardware requirements	Remarks			
	K	L _∞	t ₀	Seasonal growth	Z	L _c	Recruitment pattern	strength	Stable age dist.	Gen. v. Bert. growth	Normal length at age	Pulse season recruit- ment	Samples needed at least one regular	C/f or C at L	Other input parameters		
Shepherd's predictive	—	—	—	—	—	—	—	—	Y	Y	N	N	Y	N	Y	Growth	Micro
Sanders' sex ratio	—	—	—	—	—	—	—	—	Y	Y	N	N	Y	N	Y	Z, growth, catch-at-length by sex	Calculator
Schnute and Fournier	+	+	+	+	+	+	+	—	N	Y	Y	Y	Y	N	Y	—	Large
Length VPA	—	—	—	—	+	—	—	—	Y	Y	N	N	Y	N	Y	L _∞ , M/K	Calculator
Sanders' age at length	+	+	+	+	—	—	—	—	N	Y	Y	N	n.a.	N	N	Age/length	Calculator

^aY = yes; N = no; + = parameter is estimated; — = parameter is not estimated; n.a. = not applicable; C/f = catch/effort data; C at L = catch-at-age data.

fairly large number of assumptions about—for example—the number and normality of modes, and constancy or otherwise of mortality). The method of Sparre, for example, may require ten or more parameters to be determined, many of them incidental or “nuisance” parameters. This is not a particular disadvantage in principle, but it does consume degrees of freedom and also means that such methods tend to be quite computationally demanding. The need for efficient minimization routines also means that they are most easily implemented on minicomputers, although they could be implemented on fairly powerful microcomputers preferably with a floating point coprocessor.

At the other end of the computational spectrum, ELEFAN I requires very few assumptions to be made about the structure of the data, and can be run (in BASIC) on almost any microcomputer. However, several users have reported difficulty in obtaining convergence on some data sets, and the method seems to have a tendency to underestimate L_∞ , and possibly also K. Shepherd (Part I, this vol.) has recently developed a method which is very similar in spirit to ELEFAN I, but uses a different criterion for the quality of modal matching. This method has not yet been subjected to extensive testing “in the field”, but the tests carried out at this meeting (e.g., Table 2) suggest that although it may have a slight tendency to overestimate K and L_∞ , it performs quite reliably even on fairly “difficult” data sets. The method does not yet allow for seasonal growth, but the modifications should be possible and further tests are to be encouraged. The method is presently available only in FORTRAN 77 (see Morgan and Pauly, Part II, this vol.), but could be run on any of the many microcomputers for which a suitable compiler is available, or translated into BASIC if necessary. It may well prove useful as a preprocessor for the more elaborate (and informative) methods such as those of Pope (Part I, this vol.) and Sparre (Part I, this vol.).

There is also a whole class of methods, ranging from extreme to moderate simplicity, depending essentially on the interpretation of the mean length in the stock as related quantities. Such methods invariably assume a steady-state (equilibrium) age composition, which usually requires both constant mortality with both age and time, and constant recruitment. Situations where these conditions are all fulfilled are fairly rare, and since these methods are quite sensitive to violations of the assumptions, their use cannot be generally recommended except under especially favorable conditions or for very preliminary estimates, for which they are of course still useful.

It has been remarked by several authors that when analyzing length compositions alone it can be quite difficult to choose between different and apparently equally plausible alternative interpretations. It is indeed clear (see Table 2) that most of the methods really determine only the product of K and L_∞ at all accurately even in favorable circumstances. The selection of L_∞ in particular may have a major effect on the estimation of mortality on larger fish and, thus, the addition of even a limited amount of age/length data may have a considerable benefit. Thus, aging a selection of the largest fish may (if they prove to be old relative to $1/K$) rapidly give a reasonable estimate or a lower limit L_∞ while Wetherall et al. (Part I, this vol.) provide a rigorous method for estimation of L_∞ without growth data. Morgan (Part I, this vol.) has also developed a method for combining limited age/

Table 2. Length composition analysis by the SRLCA program of Shepherd (Part I, this vol.) for *Otolithes argenteus* data of 1982: table of score function.

L_∞	Value of K										
	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60
40	4.1	4.5	3.7	5.7	3.7	3.5	9.6	14.1	12.7	9.5	12.0
45	2.1	2.8	3.5	3.5	12.8	11.3	11.1	11.9	11.3	11.2	15.1
50	3.9	4.9	4.8	11.3	12.3	10.9	10.2	14.4	13.3	16.6	21.0
55	1.7	1.8	8.5	12.6	8.3	11.5	11.1	20.5	23.6	24.6	23.7
60	3.3	5.8	12.2	9.8	9.3	15.1	23.0	25.9	19.7	20.7	24.9
65	7.0	5.2	8.7	8.6	13.4	25.1	19.6	25.6	27.5	25.1	30.8
70	2.3	11.1	10.4	9.5	21.2	22.2	28.5	28.5	23.2	16.4	8.0
75	5.5	7.7	9.0	20.0	21.4	30.0	30.0	22.6	12.2	5.8	13.2
80	3.3	9.6	12.8	18.0	29.5	32.3	23.0	10.6	9.0	19.3	29.7
85	8.7	15.5	16.8	24.0	34.6	26.1	11.5	10.6	22.2	33.2	42.7
90	7.6	3.3	15.8	32.8	31.9	15.5	9.8	22.4	34.3	44.2	52.2

length data with length compositions, which shows considerable promise and warrants further investigation.

It has become apparent that a single point estimate of the growth parameters rarely results from the analysis of size composition data but rather a range of K and L_∞ values are produced over a plateau on the goodness-of-fit criteria response surface.

As an example, Table 2 presents the response surface produced when Shepherd's method is applied to monthly 1982 size composition data of *Otolithes argenteus* from Kuwait. Three plateaus result (indicated in the table) from which it is not possible to distinguish the correct growth parameter combination on the basis of the length data alone. Additional data are, therefore, required in this example in order to locate the plateau upon which the correct parameter combination may be found. On the basis of length-at-age data, the correct combination appears to be in the region of $L_\infty = 60$, $K = 0.28$.

It is interesting to note that in Table 2, successive plateaus are produced about multiple values of K . This harmonic property of Shepherd's method appears common and may be a result of mismatching modes. The evaluation of such mismatching may be examined if a knowledge of the number of recruited cohorts per year is available, or if other (e.g., age-at-length, tagging) data exist.

It is, therefore, apparent that an examination of the goodness-of-fit response surface is an essential element in any assessment technique utilizing size composition data (and should also be examined for data sets which utilize age or tagging data) since this surface will contain at least qualitative information on the confidence region of the parameter values and their degree of interdependence.

The more elaborate methods such as those of Sparre (Part I, this vol.) and Pope (Part I, this vol.) have the apparent advantage that mortality and population size may be estimated simultaneously (perhaps in accordance with some more or less restrictive model such as constant Z with age or time). Where growth parameters only are estimated (using e.g., ELEFAN I or the SRLCA programs), they must be used as input parameters to further computations of population structure and, thus, Z and cohort size. There are various methods for doing this, notably the use of length-converted catch curves (Pauly 1984), or cohort slicing and the application of catch curve or log ratio methods. These are relatively straightforward, but were not actively pursued at this meeting and are not discussed in detail here. If total international catch compositions are available, traditional VPA techniques may be applied (as in ELEFAN III, see Pauly, Part I, this vol.). An interesting development is the use of linear factorial modelling ("analysis of variance") techniques, using the computed ages determined from length data, as presented by Pope (Part I, this vol.). This effectively provides a catch curve corrected for varying cohort size, and is potentially a powerful method worthy of further investigation.

It is not yet clear whether the simultaneous or sequential methods of proceeding are preferable. The former have the merit of intellectual rigor and convenience, whilst the latter are computationally simpler and, requiring human intervention at a critical stage, may be more informative (especially of problems with the data).

There is also a class of methods (see Table 1) generally only applicable if a stable age composition can be assumed validly, which do not estimate growth and mortality coefficients separately but only in combination (e.g., as Z/K). These are distinctive in that they do not rely on modal analysis, but regrettably are of limited application because of the rarity of stable age compositions.

Finally, a method based on analysis of sex ratios was presented by Sanders (Part I, this vol.). This is clearly only applicable where sex can be easily determined during sampling, but in such cases provides a quite distinctive and possibly valuable approach.

Methods of Assessment: Synthetic

The techniques for computing quantities such as yield and biomass per recruit once growth parameters and selection curves have been determined are classical (Beverton and Holt 1956; Ricker 1975; Gulland 1983) and seem to be perfectly adequate for the purpose. The utility of yield-per-recruit as a tool for management is, however, being increasingly questioned in temperate areas, and there is no reason to suppose that it is any more useful elsewhere. As discussed, above techniques are now available which would permit the analysis of time series of research vessel survey or CPUE

length composition to provide time series of indices of stock and recruitment, and it could well be that these, together with analysis of the tradeoff between catches in the short term and catch rates in the long term (using biomass-per-recruit estimates) could be more useful. A discussion of some nonparametric methods for interpretation of stock and recruit data may be found in the reports of the ICES working group on the Methods of Fish Stock Assessment (Anon. 1983, 1984).

Methods for the preparation of short-term forecasts of catches or catch rates seem not to have been much developed unless they have been forgotten, but present no great technical difficulties. The simple procedure proposed by Shepherd (Part I, this vol.) is not yet ready for practical application, but requires only to be made more sensitive to signal and less sensitive to noise in the data, which should be possible. Whether such calculations are of use for management or advice to fishermen is not clear.

Comparison with Traditional (i.e., Age-Based) Methods

As a brief and obviously nonexhaustive test of some of the length-based methods, monthly size composition data on newaiby (*Otolithes argenteus*) from Kuwait and North Sea haddock (*Melanogrammus aeglefinus*) were examined. In addition, age-length keys for the two species were available for comparison of growth and mortality parameter estimates.

Application of some of the simpler analytic length-based methods for growth parameter estimation provided response surfaces which included the "correct" estimates as derived from age-at-length data although these estimates were commonly at local maximum and rarely at the global maxima. From the length-based methods alone it was, therefore, not possible to distinguish the correct parameter combination unless additional information (e.g., age-at-length data, independent estimates of L_∞) was available.

It should be emphasized that the data sets used in these cases were such that modes in the size composition were not readily apparent (e.g., see Fig. 1 for newaiby data) and it is expected that the better the separation of modes in a sample, the less extensive the plateau in the associated response surface would be, resulting in better estimates of the growth parameter values. It has been shown both at the meeting and previously (e.g., Morgan 1983) that at least the ELEFAN I technique provides almost identical growth parameter estimates to traditional age-at-length methods when modes in the length sample are well separated.

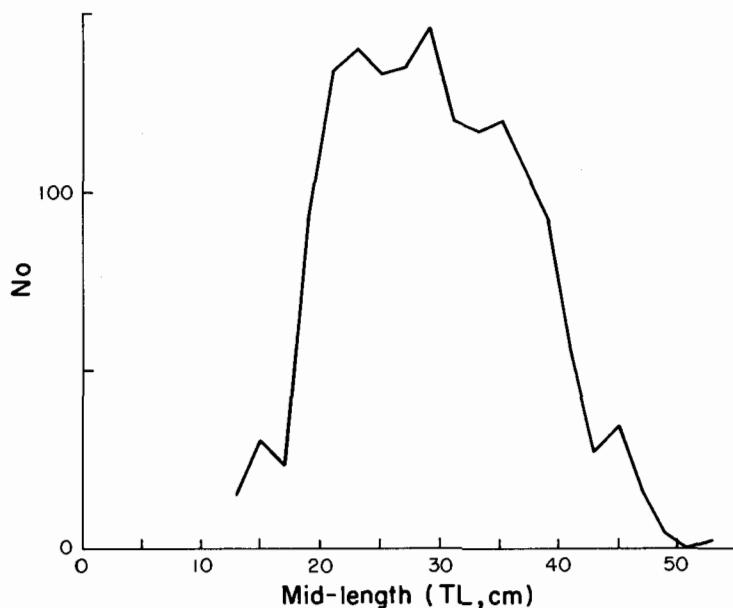


Fig. 1. Size composition of newaiby (*Otolithes argenteus*) from Kuwait, 4th quarter 1982.

In the examples studied, mean size did not change significantly with time (e.g., Fig. 2) and this limited the usefulness of Ebert's regression method in providing growth parameter estimates.

While not tested, the mixture models of Sparre (Part I, this vol.) and Schnute and Fournier (1980) might have been expected to also produce estimates within the correct region for these data sets. However, it should be recognized that both of the methods require additional inputs (e.g., number of cohorts present in the sample) for separation of the cohort structure.

Because of the absence of well defined modes in the data, none of the methods for the simultaneous estimation of mortality and growth parameters (e.g., Pope, Part I, this vol.; Ebert, Part I, this vol.) appear likely to give adequate results with the data sets examined. It was, therefore, essential in this instance to have external inputs of growth parameters for mortality rate estimation. However, with these additional inputs, the methods examined all gave realistic and consistent estimates of mortality.

Brief examination of effects of incomplete sampling was undertaken by truncating the size distributions at the lower and upper ends, adjusting the age distribution accordingly and re-analyzing the data. In general, the effects on the length-structured models of such perturbation were limited when the distribution was truncated at the upper end, thus reflecting the relatively poor information content of these larger fish. The estimates of growth and mortality parameters from the age-structured models were more influenced by this length truncation (simulating selection) since the older (and larger) cohorts contained relatively more information than the equivalent size distribution.

Truncation of the size distribution at the lower end has profound effects on the ability of the size-related methods to identify parameter values while this effect was not so pronounced with the age-based methods. This effect apparently results from the larger information content for length-based analysis of the smaller size groups (where the modes were more distinct) when compared with the information content of the smaller (and younger) size groups for age-based analyses.

It was, therefore, generally concluded that when using length-based methods, information on the smaller size groups (where the modes are usually more distinct) is of greater importance for growth and mortality parameter estimation than it is for age-based methods where, apart from generally declining sample sizes, the age structure of the older fish has equivalent information content as the age structure of the younger fish.

In this context, length-related assessment methods may be particularly appropriate to situations such as survey data which often have more information on the smaller and younger group than commercial catch data.

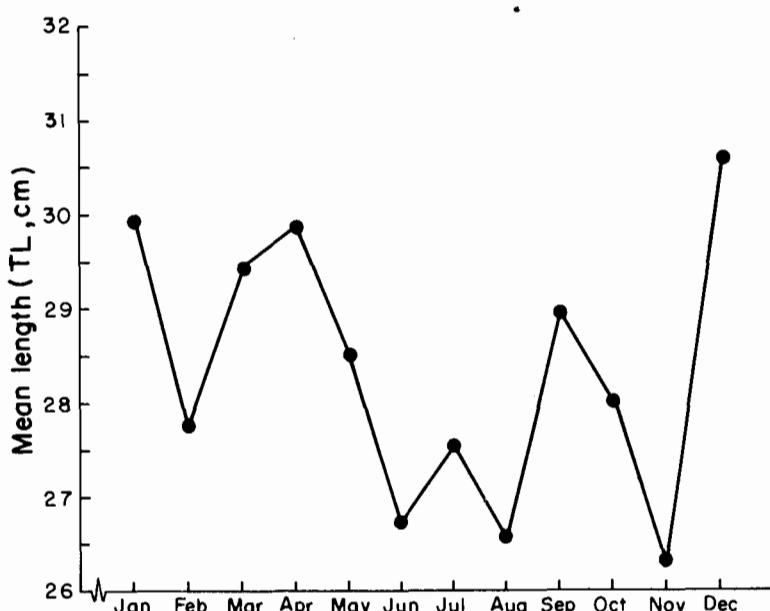


Fig. 2. Mean length of newaiby (*Otolithes argenteus*) landings in Kuwait in 1982.

Conclusions

For assessment purposes, length-structured models have an important role in providing some or all of the information necessary for management purposes. The application of length-related techniques and the reliability with which these techniques will perform is, however, closely linked to the characteristics of the data available from a fishery and their quality.

Fig. 3 attempts to summarize, in simple terms, the possible variations in sequences of size composition data. Type A involves data where there is only one observable mode in the data and this mode does not change with time; Type B represents data where a single mode is observable but this mode shifts with time as fish grow. This type of data is usually a result of a short recruitment period and one cohort per year. In Type C data a situation occurs where there are obvious multiple modes in the size distribution which correspond to cohorts while type D is the situation of multiple overlapping cohorts such that no distinct modes in the distribution are apparent.

It should be emphasized that these four possible data types represent extreme situations and in practice there will be a gradation between each type.

The proper roles of length-based assessment techniques in these four instances will differ and are summarized as follows:

Data type	Priorities for data collection		Most appropriate length-based method for estimating:	
	Length	Other type	Growth	Mortality
A	Low	High — Selection, life history, growth, mortality	None	None
B	High	Low — Growth, mortality recruitment period data	ELEFAN I, SRCL and similar methods	Ebert if stable-age assumption reasonable
C	High, intensive for one year	Low — Growth data for confirmation	Mode separation techniques, ELEFAN I, Sparre, Pope, etc.	Length-converted catch curve, mean length methods, Pope, Sparre, etc.
D	High, extensive over many years	High — Growth information	Sparre, Pope, Morgan, Schnute/Fournier	Length-converted catch curve, Pope and Sparre

Catch and fishing effort data have not been included in the table since these data are usually essential information for management purposes and are collected as a matter of course in most fisheries. However, where additional information is required to support purely length-based methods, catch-at-length data provide valuable information on mortality rates in particular.

The table highlights the critical position of the growth parameters in length-based assessment techniques since in the instances where supporting information is desirable for length-based assessments such data are always the growth parameters (or at least growth increment data). Obviously, in instances where length-based assessment methods do not contribute greatly to assessment techniques, growth and mortality need to be estimated independently although given the growth parameters, size composition data can be a useful tool for estimating mortality rates.

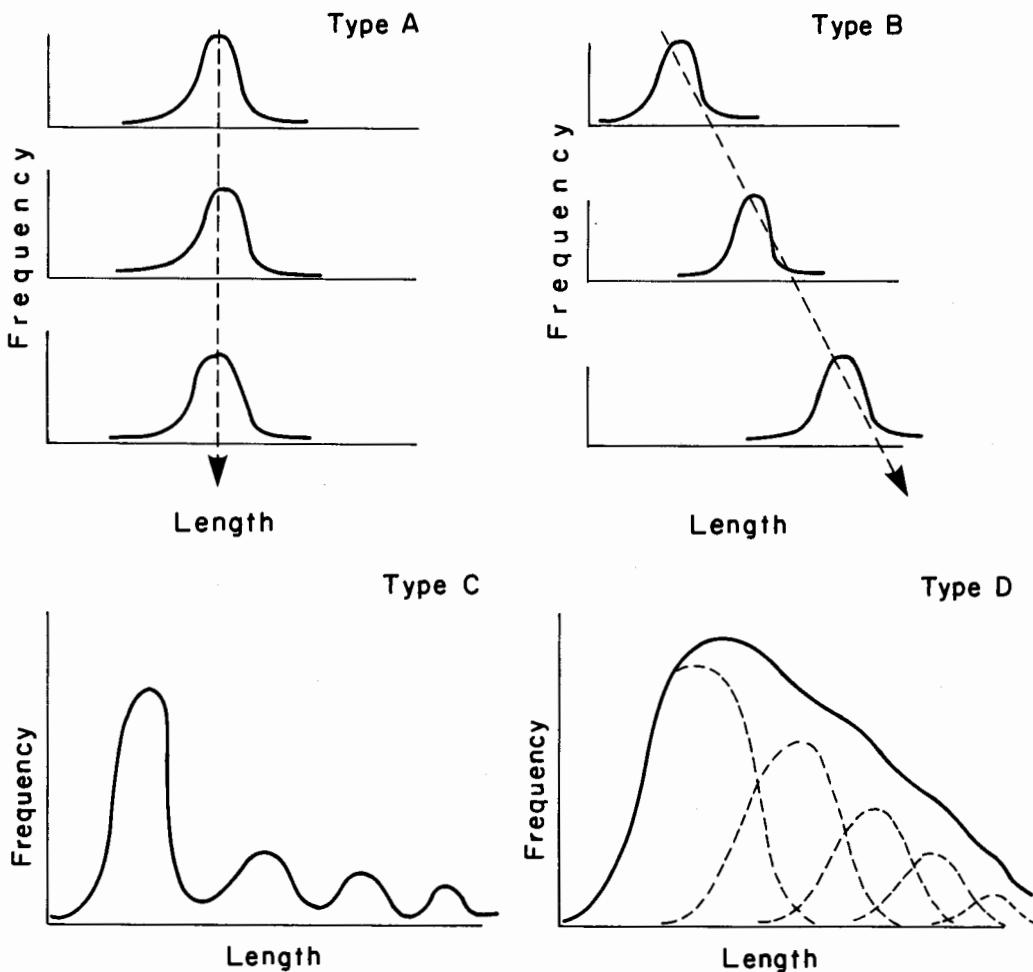


Fig. 3. Types of size composition data.

References

- Abramson, N.J. 1971. Computer programs for fish stock assessment. FAO Fish. Tech. Pap. 101, pag. var., Rome.
- Beverton, R.J.H. and S.J. Holt. 1956. A review of methods for estimating mortality rates in fish populations, with special references to sources of bias in catch sampling. Rapp. P.-v. Réun. Const. Int. Explor. Mer 140: 67-83.
- Fournier, D.A. and P.A. Breen. 1983. Estimation of abalone mortality rates with growth analysis. Trans. Amer. Fish. Soc. 112: 403-411.
- Gulland, J.A. 1983. Fish stock assessment. A manual of basic methods. John Wiley and Sons, New York.
- Hasselblad, V. 1966. Estimation of parameters for a mixture of normal distributions. Technometrics 8(3): 431-440.
- Macdonald, P.D.M. and 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.
- Morgan, G.R. 1983. Application of length-based stock assessments to Kuwait's fish stocks. ICLARM Newsletter 6(4): 3-4.
- Pauly, D. 1984. Fish population dynamics in tropical waters: a manual for use with programmable calculators. ICLARM Studies and Reviews 8, Manila, 325 p.
- Pauly, D. and N. David. 1981. ELEFAN I, a BASIC program for the objective extraction of growth parameters from length-frequency data. Meeresforsch. 28(4): 205-211.
- Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191. 382 p.
- Schnute, J. and D. Fournier. 1980. A new approach to length-frequency analysis: growth structure. Can. J. Fish. Aquat. Sci. 37: 1337-1351.

**Sensitivity of Length-Based Methods for Stock Assessment:
Report of Working Group III**

J. MAJKOWSKI

J. HAMPTON

*Division of Fisheries Research, Marine Laboratories
Commonwealth Scientific and Industrial Research Organisation
GPO Box 1538, Hobart, Tasmania 7001, Australia*

R. JONES

*Marine Laboratories
Department of Agriculture and Fisheries for Scotland
PO Box 101, Aberdeen ABG 8DB, Scotland, United Kingdom*

A. LAUREC

*Institut Scientifique et Technique des Pêches Maritimes
Nantes, Cedex, France*

A.A. ROSENBERG

*Marine Resources Assessment Group
Centre for Environmental Technology
Imperial College of Science and Technology
48 Prince's Gardens
London, SW7 1LU
United Kingdom*

Majkowski, J., J. Hampton, R. Jones, A. Laurec and A.A. Rosenberg. 1987. Sensitivity of length-based methods for stock assessment: report of working group III, p. 363-372. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

The sources of errors related to the application of a number of length-based stock assessment methods are reviewed. The methods reviewed include approaches for estimating seasonal growth parameters, total, natural and fishing mortalities, yield per recruit and for making yield predictions. Available techniques for examining the sensitivity of length-based methods are simulation modelling, sensitivity analysis, jackknife and bootstrap procedures and the examination of response surfaces. The applicability of these techniques to length-based stock assessments is discussed.

Introduction

The objective of this paper is to report on the deliberations of the working group created during the conference to discuss the sensitivity of length-based methods for stock assessment and procedures for examining this sensitivity.

Length-based stock assessment is defined here as determining, from catch length frequencies, the present biological state of a population (i.e., the present abundance, age structure and parental biomass compared to those under pre-exploitation conditions), predicting the effect of different fishing regimes on future population states and resulting catches, and determining biologically optimal fishing regimes. The group examined methods for estimating growth parameters, mortality rates, population abundance and yield per recruit, as well as some methods originally developed for age-based stock assessment but which can be used for length-based stock assessment in some circumstances.

The group considered the sensitivity of the methods from a broad perspective, including responses to (i) using input data collected in various ways (including incomplete, inaccurate, biased or noisy data) from fisheries systems having different types of characteristics, (ii) making imprecise or incorrect assumptions regarding the systems studied and (iii) applying input values of parameters imprecisely estimated, uncertain, incorrectly assumed or stochastically variable.

In this report the sensitivity of the stock assessment methods is discussed first. A review of procedures for examining this sensitivity then follows.

Methods for Estimating Growth Parameters

General Comments

The first step in length-based assessment of a fish stock is usually to estimate growth parameters. Sometimes this is done simultaneously with the determination of other parameters, such as the number of fish caught from individual age classes or mortality rates (for examples, see Powell 1979; Jones 1981; Wetherall et al., Part I, this vol.).

Successful estimation of growth parameters from catch length frequencies alone is highly dependent on the extent to which the modes associated with individual cohorts (i.e., year-classes) in these frequencies are distinguishable. As the cohorts are usually assumed to be represented by a weighted sum of overlapping normal distributions, we can formulate a necessary condition of the separability of two cohorts as

$$|\mu_i - \mu_{i+1}| > 2 \min(\sigma_i, \sigma_{i+1}) \quad \dots (1)$$

where μ_i and σ_i are the mean length and the standard deviation, respectively, associated with cohort i (Behboodian 1970; Rosenberg and Beddington, Part I, this vol.).

The separability of cohorts is most difficult for older fish, which may grow more slowly and be more variable in length at a given age than younger ones (Sainsbury 1980; Jones, Part I, this vol.; Rosenberg and Beddington, Part I, this vol.). Methods vary in their effectiveness in extracting growth parameters from catch length frequencies, but if cohorts in length frequencies are not distinguishable, at least for young fish, all of them will fail.

We believe that the inclusion of those parts of catch length frequencies that fail to satisfy condition (1) will not improve the estimation of growth parameters, because sampling errors may create artificial peaks. These peaks may lead to a considerable bias in the estimates. As preliminary knowledge of μ_i 's and σ_i 's is necessary for the exclusion of the above-mentioned parts of length frequencies, an iterative procedure for estimating growth parameters may be appropriate.

It is very important that as much reliable *a priori* biological information as possible is taken into account in the estimation of growth parameters. For example, because growth of fish seems to be often sufficiently approximated by the von Bertalanffy curve modified for seasonal oscillation (see Pitcher and MacDonald 1973; Pauly 1982, Part I, this vol.), we suggest a constraint in the form of this curve be included where appropriate.

If catch length frequencies are associated with different seasons of the year, it is important to account for seasonal oscillations in growth. By doing so, we increase the number of growth parameters but these additional parameters may sometimes be determined or constrained on the basis of *a priori* biological knowledge. If possible, we strongly recommend fixing or constraining them in the process of estimating growth parameters, thereby reducing the number of parameters to be estimated on the basis of length frequencies.

The less easily cohorts are distinguishable in catch length frequencies, the more pressing is the need to include information on fish growth that is independent of the length frequencies. This information can be provided by tagging experiments or reading hard parts for age determinations (see review in Prince and Pulos 1983). The improvement in growth parameter estimates based on length frequencies may be substantial, but only if the independent information on growth rates is provided for ages that are associated with poorly distinguishable age class modes in the catch length frequencies.

The selection of curves having more parameters than the von Bertalanffy curve, such as the Richards (1959) curve, is not recommended, even if modified for seasonal oscillations, unless growth of a particular species is described considerably better (in the statistical sense) by such a curve. It is usually difficult to uniquely determine the parameters of the von Bertalanffy curve modified for seasonal oscillations; the inclusion of more parameters may worsen the problem.

Because the L_{∞} and K estimates of the von Bertalanffy curve are strongly correlated, they are difficult to estimate uniquely, even if age class modes are easily distinguishable in catch length frequencies, unless some independent information on growth is available. Usually several combinations of L_{∞} and K values can be nearly equally well fitted to length frequencies. Not having unique values of L_{∞} and K may only be a major problem if either the growth curve is used for the extrapolation of the observed growth pattern or the L_{∞} and K values are utilized for the estimation of the instantaneous rate of natural mortality, M, on the basis of Pauly's (1982) formula. In recognizing the problem, we must warn against the common practice of using the values of L_{∞} and K in these ways.

NONPARAMETRIC METHODS

Nonparametric methods for estimating growth parameters, unlike the parametric methods (see below), utilize only a part of the information contained in catch length frequencies, namely, the number and positions of apparent peaks in these frequencies.

ELEFAN I (see Pauly and David 1981) is a now commonly used computer procedure for estimating the von Bertalanffy growth parameters from catch length frequencies. Estimates resulting from this technique may be strongly influenced by (i) sampling errors, (ii) variation in growth rates among individual fish and among cohorts, (iii) variation in the time of recruitment of individual fish, (iv) size-selective fishing and (v) the way the length measurements are grouped in length classes, and time intervals for the analyses. Some of these influences are addressed by Hampton and Majkowski (Part I, this vol.) and Rosenberg and Beddington (Part I, this vol.). Here, we note only that Pauly and David's (1981) criteria for fitting the growth parameters cannot be fully justified on statistical grounds.

Morgan's (Part I, this vol.) modification of ELEFAN I can be applied if growth information independent of the length frequencies is available. It may lead to a considerable improvement in the estimation of growth parameters, especially if reliable information on growth is available for old fish. Otherwise, all the above comments on ELEFAN I also apply to Morgan's modification.

Year-to-year variability in recruitment and variability in growth parameters will influence the reliability of results of Shepherd's (Part I, this vol.) method. Growth parameter estimates derived by this method, as it stands at the moment, could be significantly biased by seasonal oscillations in growth if catch length frequencies for different seasons of the year are used as input data. However, this bias could be eliminated by modifying the method. Sampling errors will affect estimates derived from the method, but possibly to a lesser extent than those obtained from ELEFAN I.

PARAMETRIC METHODS

Parametric methods utilize not only information on the number and positions of apparent peaks in catch length frequencies, as do the nonparametric methods, but also information on their shape. In other words, these methods utilize more information contained in length frequencies than do the nonparametric methods.

Parametric methods usually assume that age class peaks in catch length frequencies can be approximated by normal distributions. The validity of this assumption has been questioned by some authors (for example, by McNew and Summerfelt 1978). The log normal distributions can also be easily used, but the validity of this approximation can be similarly questioned.

Nonparametric methods may be suitable for estimating growth parameters only. The parametric methods to be reviewed in this section additionally provide estimates of the variance of lengths within individual age classes, the number of fish caught from each age class and, in the case of more sophisticated methods, the instantaneous rates of fishing and natural mortalities, the population abundance and some other population parameters.

The graphical and computer methods developed and applied by Harding (1949), Cassie (1954), Bhattacharya (1967), Hasselblad (1966), MacDonald (1969, 1975), Abramson (1971), Kumar and Adams (1977) and MacDonald and Pitcher (1979) do not constrain fish growth by any curve. We believe that the lack of this constraint often makes the estimation of growth rate more difficult and, in some cases, less reliable.

The method developed by Schnute and Fournier (1980) takes into account *a priori* information on the form of a growth curve, the form of the relationship of the variation in length at a given age to age and their parameters. For this reason we consider this method to be superior to the methods considered above in this section, provided such *a priori* information exists and such assumptions are appropriate. The Schnute and Fournier (1980) method does not allow for year-to-year changes in the temporal pattern of recruitment. The Pope (Part I, this vol.) and Sparre (Part I, this vol.) methods are similar to that developed by Schnute and Fournier (1980). The last method is extended by Fournier and Breen (1983) and Breen and Fournier (1984) to simultaneously estimate mortality rates as well.

The parametric methods discussed, unlike some nonparametric methods, have criteria for fitting growth and other parameters which are fully supportable from the statistical point of view. In practice some of the theoretical conditions justifying these criteria may not be fulfilled (for example, simple random sampling). But if not optimal in the precise statistical sense, these criteria will likely remain reasonable for most applications.

All parametric techniques mentioned in this section result in large numbers of parameters being simultaneously estimated. Hence, the precise estimation of these parameters may be difficult, especially if the plausible ranges of the parameters to be estimated are very wide or entirely unknown. On the other hand, it should be realized that these methods utilize more information from length frequencies than do nonparametric ones, which may compensate for the larger number of parameters to be estimated. Their main advantage is that they allow for the variability of lengths at a fixed age in the estimation process.

Methods for Estimating Mortality Rates

METHODS ASSUMING CONSTANT MORTALITY

The Ebert (Part I, this vol.) and Damm (Part I, this vol.) methods require only information on the changes of mean length over time. Hence, they do not fully utilize all the information in catch length frequencies. However, sometimes only information on mean length is available and in such a case we recommend using these methods, if all assumptions associated with them are satisfied (which is unlikely for many species).

Several methods have been proposed for estimating from catch length frequencies, the instantaneous rate of fishing mortality (F) or the quantity Z/K , where Z is the instantaneous rate of total mortality and K is a parameter of the von Bertalanffy growth equation. These methods are reviewed

in detail by Wetherall et al. (Part I, this vol.). The reliability of the Beverton and Holt estimator of Z/K is strongly dependent on the difference of the mean length, \bar{L} , from either L_∞ or L_c , the length at recruitment to the fishable stock. As \bar{L} approaches either of these limits, the method's ability to estimate Z/K decreases rapidly. The sensitivity of the method to uncertainties in the input parameters L_∞ and L_c was examined by Majkowski (1982a) for a tropical fish species.

The effect of individual variability in fish growth on results of the methods for estimating Z/K has not been extensively studied. Laurec and Mesnil (Part I, this vol.) show that the Beverton and Holt estimator is quite robust to individual variability, because it utilizes only \bar{L} . Studies carried out by Jones (Part I, this vol.) suggest that methods that transform the right-hand part of length frequencies may give biased results. Further studies on the reliability of the methods for estimating F and Z/K are required.

Sanders' (Part I, this vol.) sex-ratio methods for estimating natural and fishing mortality rates require quite strict conditions to be applied. However, in particular situations when these requirements hold, these methods may provide useful estimates. Their sensitivity to input parameters, sampling or individual variability has not been investigated.

LENGTH-BASED COHORT ANALYSIS

On the basis of catch length frequencies, the von Bertalanffy growth parameters, and the instantaneous rates of natural and so-called terminal fishing mortalities, length-based cohort analysis (Jones 1974, 1981, Part I, this vol.; Pauly, Part I, this vol.) estimates both the number of fish by length class and the instantaneous rate of fishing mortality by length class. The sensitivity of the method to uncertainties in the instantaneous rates of terminal fishing and natural mortalities is expected to be similar to that for age-based cohort analysis (Pope 1972; Hampton et al. 1984). It is always strongly advisable to use sensitivity analysis (see below) in conjunction with cohort analysis.

Problems will arise with the method if length frequencies of samples do not fully reflect the characteristics of the entire catch (Jones, Part I, this vol.; Laurec and Mesnil, Part I, this vol.). If there are substantial errors in catch length frequencies estimated on the basis of sampling, the estimation of the instantaneous rate of fishing mortality and the fish abundance may be very unreliable, especially at lengths close to L_∞ . This is partially due to low frequencies of large fish in the catch. In some circumstances smoothing the data may reduce these problems. Smoothing may also be justified under the assumption that the fishery is operating at a steady state.

Results of length-based cohort analysis are sensitive to changes in the input estimates of the growth parameters L_∞ and K , but an uncertainty in the value of t_o is not important to this analysis. Laurec and Mesnil (Part I, this vol.) have approximated the effects of uncertainties in L_∞ and K on the estimation of the instantaneous rate of fishing mortality. Estimates of F for length classes associated with small fish are robust to uncertainties in these growth parameters as long as the values of L_∞ and K are compatible (see above). However, estimates of F for lengths close to L_∞ are strongly affected, which confirms previous conclusions (Pereiro and Sallares 1984). One way of determining the extent of the problem in a particular application is to analyze the catch data systematically over a plausible range of growth parameter values.

Preliminary analyses of the sensitivity of the method to individual variability have been carried out by Jones (Part I, this vol.) and Laurec and Mesnil (Part I, this vol.). Estimation problems arise in the region close to L_∞ , which emphasizes the necessity of defining a terminal group whose upper bound is sufficiently smaller than L_∞ . Further studies are required on this problem. Pronounced individual variability in growth will make it difficult to use length-based cohort analysis when the fishery is not in a steady state. Individual or year-to-year variability in growth will result in problems in separating cohorts and, consequently, tends to smooth year-to-year fluctuations in recruitment and F . In this case, only general trends (such as a moving average) may be observed. The extent to which these estimates may be smoothed is unknown at present.

The sensitivity of length-based cohort analysis to changes in the length interval is caused by sampling noise. Smoothing the length frequencies may help alleviate this problem. There is no apparent advantage to wider grouping versus smoothing particularly in steady-state conditions.

Methods for Predicting Future Yields and Estimating Total Allowable Catches

YIELD-PER-RECRUIT ANALYSIS

The relative sensitivity of results of yield-per-recruit (Y/R) analysis to uncertainties in input parameter values may be different from the sensitivity of intermediate results of length-based stock assessment to the same uncertainties. A bias in an estimate of a parameter used in the Y/R analysis may propagate through the assessment procedures with the ultimate effect that the Y/R calculation may be magnified or reduced. This process can be illustrated by assuming M to be overestimated and analyzing its effect on Y/R predictions in the case of increasing a mesh size. The overestimated M will cause this effect to be underestimated. Also, as a consequence of the overestimated M, F will be underestimated by length-based cohort analysis. The underestimated F will cause this effect to be even further underestimated. Therefore, the final effect of overestimating M on the Y/R calculation will be magnified.

To our knowledge, there had been no systematic investigation of the sensitivity of Y/R results derived in length-based assessment before this conference but subsequently this problem has been addressed by Laurec and Mesnil (Part I, this vol.). The problems anticipated involve biases and inaccuracies occurring in the estimation of input parameters to Y/R analysis. We would like to note that the uncertainty in t_o , while unimportant for most analyses, may affect Y/R because t_o determines the age at recruitment. This is only true if the age at first capture is not equal to the age at recruitment.

OTHER METHODS

In the previous sections we have reviewed the sensitivity of stock assessment methods developed specifically for analyzing catch length frequencies. Here, we would like to note that all traditional methods developed specifically for analyzing catch age frequencies (see reviews in Ricker 1975; Gulland 1977, 1982) may also be applicable to catch length frequencies, if these frequencies can be reliably transformed to age frequencies.

Length frequencies can be transformed to age frequencies using the methods reviewed above. Alternatively, age for each fish caught can be deterministically estimated on the basis of a length-age relationship, if parameters of such relationship can be estimated. The last method is described and its reliability examined by Majkowski and Hampton (1983a, 1983b). This method is routinely used, for example, for southern bluefin tuna (see Majkowski 1982a; Hampton et al. 1984) and for tuna stocks relevant to the International Commission for Conservation of Atlantic Tunas. Note that the length-age relationship is not necessarily the same as the age-length relationship (see Kirkwood 1983). The problem with this aging method is related to individual variability in growth, which is not accounted for and which may lead to the underestimation of the instantaneous rate of fishing mortality for large fish.

Traditional age-based stock assessment methods include age-based cohort and classical Y/R analyses, least-square stock assessment methods, simulation methods used for predicting future catches and population parameters and procedures for estimating total allowable catches (e.g., maximum sustainable yields, catches stabilizing the level of parental biomass). These methods have been routinely applied for analyzing catch length frequencies of, for example, southern bluefin tuna (see Murphy and Majkowski 1981; Majkowski and Hampton 1983c; Hampton et al. 1984; Majkowski and Caton 1984; Majkowski 1985; Hampton and Majkowski 1986).

Although the age-based methods have been applied for a much longer period than the length-based methods, their sensitivity has not been extensively studied either. Some examples of such studies include Mandecki (1976), Mohn (1979), Orth (1979), Moreau (1980), Majkowski and Hampton (1983c), Hampton et al. (1984) and Hampton and Majkowski (1986).

Procedures for Examining the Sensitivity of Length-Based Methods

Various procedures of very different complexity may be used for examining the sensitivity of length-based stock assessment methods. They include simulation modelling, so-called sensitivity analysis, jackknife and bootstrap techniques and examinations of response surfaces.

SIMULATION MODELLING

Simulation modelling techniques can be used for examining responses of length-based stock assessment methods to using input data collected in different ways (including incomplete, inaccurate, biased or noisy data) from fisheries systems having different types of characteristics and making imprecise or incorrect assumptions regarding the systems studied. Simulations incorporating some stochastic processes (for example, see Jones, Part I, this vol. and Rosenberg and Beddington, Part I, this vol.) may be appropriate for investigating the effects of specific sources of errors, such as different sampling schemes, on particular methods. The approach of holistically simulating a fish population (i.e., following individual fish through their lives) incorporating a number of stochastic processes (for an example see Hampton and Majkowski, Part I, this vol.), whilst time consuming, may be used to assess the overall reliability of various methods of analyzing length frequencies.

When using simulation techniques, it is important to reconstruct, as closely as possible, the sampling scheme that occurs in the real world. The simulation model should define strata and clusters within those strata for the purposes of sampling. Oversimplistic assumptions may result in misleading conclusions. If the real sampling scheme is stratified, simulated simple random sampling that ignores stratification would result in overestimates of variances of parameters. If clustering occurred and it was neglected in simulations, these variances might be underestimated.

If the exact sampling scheme cannot be simulated, simulations with simple random sampling may provide an approximate indication of the robustness of a particular method. Certainly, if the lack of robustness of a method is indicated by using simple random sampling, it would prove to be even less robust if the exact sampling scheme with a strong clustering effect could be reproduced in the simulations.

SENSITIVITY ANALYSIS

Sensitivity analysis is a technique developed for examining the sensitivity of results of mathematical models or procedures to applying input values that are imprecisely estimated, uncertain, stochastically variable or incorrectly assumed (see reviews in Majkowski 1982b, 1983; Majkowski and Hampton 1983a). This technique involves simple, repetitive executions of mathematical models or procedures with parameter values perturbed (for example see Majkowski and Hampton 1983c; Hampton et al. 1984; Hampton and Majkowski 1986), Monte-Carlo simulations (for example see Majkowski and Hampton 1983b) or their analytic approximations usually based on Taylor expansions (for examples, see Majkowski et al. 1981; Majkowski and Hampton 1983a; Laurec and Mesnil, Part I, this vol.). The use of such approximations may be quite adequate if values of input parameters are relatively well known, but even in such cases they may lead to misleading results (for an example, see Majkowski and Hampton 1983a). When using approximation techniques, it is highly desirable to validate them (for the methods see Majkowski et al. 1981; Majkowski and Hampton 1983a).

JACKKNIFE AND BOOTSTRAP TECHNIQUES

The jackknife and bootstrap techniques may be used for reducing the bias in and evaluating the statistical accuracy of estimates from length-based stock assessment methods. They can be quite powerful, if correctly applied (see Efron 1982; Levi et al., Part I, this vol.).

In applying the jackknife and bootstrap techniques it is important to take due consideration of the sampling scheme involved. For example, one must jackknife according to some specific source of variability and remove, in turn, all samples relating to a particular stratum. Thoughtless jackknifing of single elements of data will provide little information and prove computer time-consuming. Care must be exercised in the selection of the units for elimination in the jackknife procedure to ensure that the assumption of identical distributions of observations eliminated is met. The method may be very sensitive to violation of this assumption, and seasonal observations if growth is seasonally varying may provide a significant violation. The computer time necessary to employ the jackknife and bootstrap techniques may be considerable (see Levi et al., Part I, this vol.).

EXAMINATION OF RESPONSE SURFACES

The examination of response surfaces may be a useful technique for the evaluation of effects of different parameter combinations (for example, L_∞ and K). Even if they cannot be used to construct confidence limits, they will always be useful for demonstrating the nature of correlation problems.

Conclusions

Not recognizing the existence of errors likely affecting any length-based stock assessment involves three types of risk:

- (i) wrong advice on fisheries management may be given due to inaccuracies of results of analyses which are unknown or ignored;
- (ii) conversely, data collections may be unnecessarily too intense (oversampling) and the methods of their analysis may be too complex resulting in the accuracy of results being better than that required to eliminate the possibility of providing wrong advice; and
- (iii) the system used for data collection and analysis may not be most optimal for the type of advice required. Other less expensive schemes may lead to more reliable advice.

The costs involved in biological programs, especially field data collections, are sufficiently high to justify a careful examination of their effectiveness and cost-benefit relationships. Various new methods for length-based stock assessment are still being developed without such rigorous examinations being made. It is impossible to determine their superiority/inferiority to the existing methods without these examinations carried out for both existing and new methods.

The need for more adequate procedures for examining the sensitivity of methods of length-based stock assessment also exists. Such procedures range from simple recalculations of stock assessment results using possible alternative values of selected input parameters (such as the instantaneous rate of natural mortality) to the estimation of probability distributions of "target quantities" such as the yield per recruit. Even a very approximate determination of the sensitivity is better than completely ignoring the problem. Some approximations and, from a more general point of view, pragmatism will often be required in determining the sensitivity. Whatever the technique used for examining the sensitivity, it will be productive only if consideration is given to the entire stock assessment process from the data collection to the calculation of final results.

Any examination of the sensitivity of length-based stock assessment will require some specific input information in addition to that essential for stock assessment. At least, a range of possible values for input parameters for stock assessment will be required. If possible more elaborate inputs such as an estimated variance/covariance matrix or probability distributions of input parameters should be used.

Acknowledgements

We are thankful to Dr. G. Eckert, Dr. F.R. Harden Jones, Mr. W.S. Hearn, Dr. T. Kenchington, Dr. G. Kirkwood, Dr. K.J. Sainsbury, Ms. S. Wayte (CSIRO Division of Fisheries Research, Hobart, Tasmania, Australia), Mr. G. Leigh (CSIRO Division of Mathematics and Statistics, Hobart, Tasmania, Australia) and Mr. I. Somers (CSIRO Division of Fisheries Research, Cleveland, Queensland, Australia) for comments and suggestions for improvements to an earlier version of this paper.

References

- Abramson, N.J. 1971. Computer programs for fish stock assessment. FAO Fish. Tech. Rep. 101. 96 p.
- Behboodian, J. 1970. On modes of a mixture of two normal distributions. Technometrics 12: 131-139.
- Bhattacharya, C.G. 1967. A simple method of resolution of distribution in Gaussian components. Biometrics 23: 115-135.
- Breen, P.A. and D. Fournier. 1984. A user's guide to estimating total mortality rates from length frequency data with the method of Fournier and Breen. Can. Tech. Rep. Fish. Aquat. Sci. 1239. 63 p.
- Cassie, R.M. 1954. Some uses of probability paper in the analysis of size frequency distributions. Aust. J. Mar. Freshwat. Res. 5: 513-522.
- Efron, B. 1982. The jackknife, the bootstrap and other resampling plans. CBMS-NSF Regional Conference Ser. 38. 92 p.
- Fournier, D.A. and P.A. Breen. 1983. Estimation of abalone mortality rates with growth analysis. Trans. Amer. Fish. Soc. 112: 403-411.
- Gulland, J.A., editor. 1977. Fish population dynamics. John Wiley & Sons, New York. 372 p.
- Gulland, J.A. 1982. Fish stock assessment—a manual of basic methods. FAO/Wiley Series on Food and Agriculture, vol. 1. 223 p.
- Hampton, J. and J. Majkowski. 1986. Computer simulations of future southern bluefin tuna parental biomass, recruitment and catches under the 1982 fishing regime. N. Amer. J. Fish. Mangmt. 6: 77-87.
- Hampton, J., J. Majkowski and G.I. Murphy. 1984. The assessment of the southern bluefin tuna (*Thunnus maccoyii*) population and the determination of catch levels which stabilize the parental biomass. CSIRO Mar. Lab. Rep. 165. 27 p.
- Harding, J.P. 1949. The use of probability paper for graphical analysis of polymodal frequency distributions. J. Mar. Biol. Assoc. UK 28: 141-153.
- Hasselblad, V. 1966. Estimation of parameters for mixture of normal distributions. Technometrics 8: 431-444.
- Jones, R. 1974. Assessing the long term effects of changes in fishing effort and mesh size from length composition data. J. Cons., Cons. Int. Explor. Mer C.M. 1974/F: 33. 13 p.
- Jones, R. 1981. The use of length composition data in fish stock assessments (with notes on VPA and cohort analysis). FAO Fish. Circ. 734. 55 p.
- Kirkwood, G.P. 1983. Estimation of von Bertalanffy growth curve parameters using both length increment and age-length data. Can. J. Fish. Aquat. Sci. 40: 1405-1411.
- Kumar, D.K. and S.M. Adams. 1977. Estimation of age structure of fish populations from length frequency data, p. 256-281. In W. Van Winkle (ed.) Proceedings of the Conference on Assessing the Effect of Power-Plant-Induced Mortality on Fish Populations. Pergamon Press, Oxford.
- MacDonald, P.D.M. 1969. FORTRAN programs for statistical estimation of distribution mixtures: some techniques for statistical analysis of length frequency data. Fish. Res. Board Can. Tech. Rep. 149. 45 p.
- MacDonald, P.D.M. 1975. Estimation of fine distribution mixtures, p. 231-245. In R.P. Gupta (ed.) Applied Statistics. North Holland, Amsterdam.
- MacDonald, P.D.M. and 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.
- Majkowski, J. 1982a. Estimation of the abundance of the southern bluefin tuna (*Thunnus maccoyii*) (Castlenau) subpopulation exploited by the Australian fishery. CSIRO Mar. Lab. Rep. 143. 11 p.
- Majkowski, J. 1982b. Usefulness and applicability of sensitivity analysis in multispecies approach to fisheries management, p. 149-165. In D. Pauly and G.I. Murphy (eds.) Theory and management of tropical fisheries. ICLARM Conference Proceedings 9. International Center for Living Aquatic Resources Management, Manila, Philippines and Division of Fisheries Research, CSIRO, Cronulla, Australia.

- Majkowski, J. 1983. Usefulness and applicability of sensitivity analysis in resource management, p. 249-261. In Wadhwani, L.C., M.K. James, A.B. Pomeroy and K.P. Stark (eds.) Resource management: issues, techniques, applications. Proceedings of the Workshop on Resources Management held at the James Cook University, 14-17 July 1981, James Cook University of North Queensland, Department of Civil and System Engineering.
- Majkowski, J. 1985. Fisheries scientists report on SBT. Aust. Fish. 44(11): 16-18.
- Majkowski, J. and A. Caton. 1984. Scientists' concerns behind tuna management. Aust. Fish. 43(9): 5-6.
- Majkowski, J. and J. Hampton. 1983a. The effect of parameter uncertainties in an age-length relationship upon estimating the age composition of catches. Can. J. Fish. Aquat. Sci. 40: 272-280.
- Majkowski, J. and J. Hampton. 1983b. Deterministic partitioning of the catch of southern bluefin tuna, *Thunnus maccoyii*, into age classes using an age-length relationship, p. 97-90. In E.D. Prince and L.M. Pulos (eds.) Proceedings of the International Workshop on "Age Determination of Oceanic Pelagic Fishes: Tunas, Billfishes and Sharks", 15-18 February 1982, Miami, Florida. US Department of Commerce, NOAA Tech. Rep. NMFS 8. 211 p.
- Majkowski, J. and J. Hampton. 1983c. The estimation of a catch level which stabilizes the parental biomass of an exploited fish stock. Fish. Bull. 81: 723-732.
- Majkowski, J., J. Ridgeway and D.M. Miller. 1981. Multiplicative sensitivity analysis and its role in development of simulation models. Ecol. Modelling 12: 191-208.
- Mandecki, W. 1976. A computer simulation model for interacting fish population studies, applied to the Baltic. Pol. Arch. Hydrobiol. 23: 281-307.
- McNew, R.W. and R.C. Summerfelt. 1978. Evaluation of a maximum-likelihood estimator for analysis of length-frequency distributions. Trans. Amer. Fish. Soc. 107: 730-736.
- Mohn, R.K. 1979. Sensitivity analysis of two harp seal (*Pagophilus groenlandicus*) population models. J. Fish. Res. Board Can. 36: 404-410.
- Moreau, J. 1980. Influence des divers paramètres sur l'estimation du rendement par recrue; application aux pêches continentales tropicales. Cybium, 3è serie 8: 67-75.
- Murphy, G.I. and J. Majkowski. 1981. State of the southern bluefin tuna population: fully exploited. Aust. Fish. 40(11): 5-10.
- Orth, D.J. 1979. Computer simulation model of the population dynamics of large mouth bass in lake Carl Blackwell, Oklahoma. Trans. Amer. Fish. Soc. 108: 229-240.
- Pauly, D. 1982. Studying single-species dynamics in a tropical multispecies context, p. 33-70. In D. Pauly and G.I. Murphy (eds.) Theory and management of tropical fisheries. ICLARM Conference Proceedings 9. International Center for Living Aquatic Resources Management, Manila, Philippines and Division of Fisheries Research, CSIRO, Cronulla, Australia.
- Pauly, D. and N. David. 1981. A BASIC program for the objective extraction of growth parameters from length frequency data. Meeresforsch. 28: 205-211.
- Pereiro, J.A. and S. Sallares. 1984. Consequences of introducing errors in the values of the input parameters in the length distribution analysis (R. Jones) as applied to the Northern stock of Hake. J. Cons., Cons. Int. Explor. Mer CM 1984/G: 61. 17 p.
- Pitcher, T.J. and P.D.M. MacDonald. 1973. Two models for seasonal growth in fishes. J. Appl. Ecol. 10: 599-606.
- Pope, J. 1972. An investigation of the accuracy of Virtual Population Analysis. Int. Comm. Northw. Atl. Fish. Res. Bull. 9: 65-74.
- Powell, D.G. 1979. Estimation of mortality and growth parameters from the length frequency of a catch. Rapp. P.-V. Réun. Cons. Int. Explor. Mer 175: 167-169.
- Prince, E.D. and L.M. Pulos, editors. 1983. Proceedings of the International Workshop on "Age Determination of Oceanic Pelagic Fishes: Tunas, Billfishes and Sharks," 15-18 February 1982, Miami, Florida. US Department of Commerce, NOAA Tech. Rep. NMFS 8. 211 p.
- Richards, F.J. 1959. A flexible growth function for empirical use. J. Exp. Bot. 10: 290-300.
- Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191. 382 p.
- Sainsbury, K.J. 1980. The effect of individual variability on the von Bertalanffy growth equation. Can. J. Fish. Res. Aquat. Sci. 37: 241-247.
- Schnute, J. and D.A. Fournier. 1980. A new approach to length-frequency analysis: growth structure. J. Fish. Res. Board Can. 37: 1337-1351.

**Analysis of Length-Frequency Data:
Some Available Programs and User's Instructions***

G.R. MORGAN^a

*Kuwait Institute for Scientific Research
P.O. Box 24885, Safat 13109
Kuwait*

D. PAULY

*International Center for Living Aquatic Resources Management
MC P.O. Box 1501, Makati, Metro Manila
Philippines*

Morgan, G.R. and D. Pauly. 1987. Analysis of length-frequency data: some available programs and user's instructions, p. 373-462. In D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Abstract

The computer and calculator programs for the analysis of length-frequency data presented by various authors at the ICLARM/KISR Conference on the "Theory and Application of Length-Based Stock Assessment" held at Mazara del Vallo, Sicily, Italy, on 11-16 February 1985, are documented, based in part on the work and report of Working Group IV (Chairman: J.G. Pope) of that conference.

Given are, for each of the thirteen programs documented here, (1) author(s) and (their) address(es), (2) key reference(s), (3) a brief program description, (4) implementation language, (5) hardware and memory requirements and (6) sources of listings and test data. Complete listings in BASIC, FORTRAN or HP 67/97 calculator codes as well as test data are given for six of the programs documented here. General recommendations regarding program implementation are provided.

Introduction

Use of length data to provide information for management goes back several decades, but analysis has tended to be by hand and frequently has involved tedious book-keeping techniques

*ICLARM Contribution No. 400.

^aPresent address: c/o W.A. Marine Research Laboratories, P.O. Box 20, North Beach 6020, W.A., Australia.

using probability paper. The recent availability of modestly priced microcomputers has provided the appropriate environment for estimation and extension of methods of analysis. Most of the programs in this contribution can be implemented on the Apple II or IBM PC and their compatibles or similar machines.

There are problems with very small computers and if sufficient funds are available, larger microcomputers or small mainframe machines (minicomputers) may be considered. The points to keep in mind are not just initial costs of the computer but also supporting hardware such as disk drives and a printer, software costs such as buying programming languages (BASIC, FORTRAN, PASCAL), operating costs such as buying printer paper, discs, salary for an operator (if a mainframe computer is purchased), and availability of local support for the computer. Operation and support of a large computer is expensive.

The faults of small computers on the other hand, are that they are slow in operation compared with mainframe machines, and that because they store fewer significant figures during number manipulation, they may accumulate errors in programs that do large numbers of mathematical operations. However, even with their obvious problems of relatively slow speed, possible accumulation of errors, and relatively small memory, microcomputers are very sophisticated machines that are capable of carrying out data analysis in minutes that would have been virtually impossible just a few years ago.

Fisheries management of stocks in tropical areas of the world is complicated by the high species diversity in the catches and by a lack of resources, including time, for using traditional methods of tagging or aging fish from hard parts. Here, length-frequency analysis is required and small computers make it possible. The programs in this contribution are offered in the spirit of improving management of complex fisheries by improving analysis of length-frequency data.

The computer programs contained in this paper are, to the best abilities of the contributing authors, error-free. It is naturally recognized that such a goal is seldom obtained and some problems must be expected. The contributing authors have attempted to anticipate difficulties by including a short introduction to each program that provides details concerning data preparation, restrictions on types of data that can be used, and their experience with resolving problems that they have experienced.

During the conference, a computer working group (Working Group IV, Chairman: J.G. Pope), was established which was responsible for running as many as possible of the methods described in Part I of this volume. This was done in order to both service Working Groups I to III and to gain an appreciation of the software available at the meeting. The problems facing the group in implementing the various methods on a PDP, an Apple II+, an Apple IIc and an ACT Apricot computer and on various programmable calculators gave an appreciation of the problems to be expected in practice. Similarly, the problems of running the various programs gave some appreciation of how user-friendly and fast or slow the various programs were. A distinct impression was formed that much of the software was at a development stage where more concern had been given to the methodology than to its ease of use. If software is to be safely exported, professional standards of documentation, test data and user instructions are essential. While international organizations such as ICLARM or FAO can provide a useful extension service, the ultimate responsibility must lie with authors. They should try to regard adequate packaging of their techniques as essential if they want them to be extensively and safely used. This is particularly important for techniques that may be used by less critical users than those usually found at working group meetings of the major international fishery commissions. The current problems are to some extent a natural feature of a rapidly expanding subject area but consolidation is important as well as innovation.

Working Group IV, whose report was incorporated into the present contribution, prepared Table 1 which shows the author, language, implementation, development level, distribution, support available, input and output and underlying model of programs designed to process a number of the length-based techniques. This is intended to provide a quick reference to potential users of what is available and what chances they would have of a successful implementation.

If you encounter problems that cannot be resolved, the usual cause will be improper program or data entry. To aid you in getting a program running, the contributing authors have provided sample data sets together with output so you can test your implementation of their program.

Table 1. Summary of software for a number of length-based assessment methods discussed and/or run at conference.

Program number	Method	(Senior) author/program	Language	Implementation	Development level	Additional documentation	Distribution	Support	Input	Output	Underlying model	Friendliness
1	ELEFAN I	Pauly and David ELEFAN I	BASIC	Several micro and mini comp.	***	A	ICLARM +	Full	Length-frequency data, any interval	L_{∞} , K, t_0 (C, WP) goodness of fit	Modal progression analysis	VG
2	ELEFAN II	Pauly and David ELEFAN II	BASIC	Several micro	***	A	ICLARM +	Full	Length-frequency data, any interval	Z, L_c and recruitment pattern	Length-converted catch curve	VG
3	ELEFAN III	Pope et al. ELEFAN III	BASIC	Several micro	**	A	ICLARM	Full	Monthly total catch at length + other items	P, F by length and month	Virtual population analysis	G
4	ELEFAN IV	Munro ELEFAN IV	BASIC	Several micro	**	A	ICLARM	Full	Length-frequency data, selection curve	Natural mortality (M)	Plot of Z on probability of capture	G
5	ELEFAN V	Morgan ELEFAN V	BASIC+	HP 9845B,	*	N.A.	From author	Yes	Length-frequency and age data	L_{∞} and K, compatibility tests	Concordance of ELEFAN I and age data	G
6	Ebert	Ebert/ZKAN	Applesoft BASIC	Apple II, II+, IIc	**	A	Author 5½" disk	Yes	Mean lengths, time and size at recruitment	Z, L_{∞} , K and growth curve shape	Average size progression	VG
6	Damm	Damm/ZKA/EZ	Applesoft BASIC	Apple II, II+, IIc	**	A	Author 5½" disk	Yes	Mean lengths size, time and size of recruitment	Z, L_{∞} , K and relative recruitment	Average size progression	G
7	Sparre	Sparre/LGTFRQ 77	FORTRAN	VAX	*	A	Author Mag. tape	Yes	Length-frequency time series, CPUE	Z, L_{∞} , K , t_0 and σ	Least squares fit of growth/mortality to data	G
	Pope I	Pope/SAPNJ	FORTRAN 77	HP 1000 Apricot	STILL	UNDER	DEVELOPMENT		Length CPUE annual time series	Z, L_{∞} , K , t_0 , σ relative recruits	Least squares fit of growth/mortality to data	P
	Pope ANOVA	Pope/standard stats pack.	Any good stats pack.	Widespread	***	A	Software vendor	from vendor	Length CPUE annual time series	Z by age, K , L_{∞} , t_0 , relative recruitment	ANOVA of length CPUE converted to age	G
8	Shepherd SRLCA	Shepherd SRLCA	FORTRAN 77	Apricot	**	A	Author 5½" or 3½" disk or tape	No	Length-frequency data	L_{∞} , K, t_0 , goodness of fit	Modal progression; weakly parametric	P
9	Phalanx	Pope/SAPPH	FORTRAN 77	HP 1000	*	N.A.	Author	Some	Total catch at length by species, K , L_{∞}	F and Z by length group by species, Y/R	Multispecies length cohort analysis	P
10	Shepherd Predictive	Shepherd HRLCF	FORTRAN 77	Apricot	*	N.A.	N.A.	No	Time series length, CPUE, K , L_{∞} , t_0	Catch forecast	Growth and mortality projection	P
11	Sanders Sex ratio	Sanders Sex ratio	Calculator program	HP 67/97	**	N.A.	—	—	Catch numbers by length, Z, t_c by sex	F and M by sex	Change in sex ratio with differential mortality	VG
12	Length VPA	Jones/RJ33	BASIC	PDP 11/70	**	A	Author	Some	Total catch at length L_{∞} , M/K F/Z, Selectivity	F and P by length group; Y/R	Length cohort analysis	P
—	Bhattacharya	Pauly/Caddy ^a	Calculator program	HP 67/97	***	A	FAO	errata sheet available from authors	Length-frequency data	Normal components Chi-Squared Test	Mixture of normal distributions	VG
13	Simulation	Hampton/Majkowski	FORTRAN 77	VAX/DEC	**	A	Author	Yes	Nil	Simulated length-frequency data	Stochastic generator	G

+ Many versions exist.

* Runs

** Recent final version

*** Version tested by several users.

A Available

NA Not available/Not applicable

P Poor

G Good

VG Very good

^aNot discussed here, see D. Pauly and J. Caddy (1985). A modification of Bhattacharya's method for the separation of normal distributions. FAO Fish. Circ. 781, Rome, 16 p. (with Errata sheet)

Program Number 1

Title	ELEFAN I
Authors	D. Pauly and N. David
First Author's Address	Daniel Pauly International Center for Living Aquatic Resources Management (ICLARM) MC P.O. Box 1501, Makati, Metro Manila, Philippines

References Pauly, D. and N. David. 1981.—ELEFAN I, a BASIC program for the objective extraction growth parameters from length-frequency data. *Meeresforsch.* 28(4): 205-211.
 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. f. Meereskd. Univ. Kiel*, No. 149, 76 p.
 Brey, T. 1986.—Estimation of annual P/B-ratio and production of marine benthic invertebrates from length-frequency data. *Ophelia*, Suppl. 4: 45-54.
 Pauly, D.—A review of the ELEFAN system for analysis of length-frequency data on fish and aquatic invertebrates. Part I, this vol.

Description ELEFAN I can be used to estimate values of the parameters of a seasonally oscillating version of the von Bertalanffy growth function, given well-structured length-frequency data representative of a population of fish or aquatic invertebrates. The data need not have been sampled at regular time intervals, nor need to be weighted by CPUE. However the sample(s) investigated must contain at least four "peaks" representative of successive cohorts (age groups). Detailed descriptions of the program's mode of operation, assumptions and pitfalls are given in the references above, and in other papers cited therein. ELEFAN I is used in conjunction with a separate routine for data entry called ELEFAN 0 (Pauly, Part I, this vol. and see Table 1).

Listing Notes

- a. **Language.** Various versions of ELEFAN I exist in various dialects of BASIC and FORTRAN. The versions supported by ICLARM (see below) are all written in BASIC.
- b. **Implementations.** ELEFAN I has been implemented by various authors onto a wide range of computers from a tiny Radio Shack TRS-80 Model I to a CRAY II Supercomputer. The versions of ELEFAN I presently supported by ICLARM, which incorporate all improvements suggested at the conference of which the present volume is the proceedings, are:
 - i) the "KIEL Package" for IBM PC and compatibles, written in BASICA and compiled using Microsoft Quickbasic. The package, (available from ICLARM for US\$15 to cover costs)** consists of one diskette (with both compiled and uncompiled codes) and the manual of Brey and Pauly (1986). A revised, second edition of this manual will be issued in early 1988.
 - ii) the "KIEL Package" for the Apple II family of computers and their compatibles (CP/M version). This package, which runs in compiled MBASIC costs the same, and consists of the same elements as the package in (i). A complementary diskette implementing high resolution graphics for this package (Liew 1986)* is available for US\$15.**
 - iii) the "KIEL Package" for Radio Shack TRS 80 Model I (with 32 K Expansion Interface) or Model III. Same costs and manual as versions in (i) and (ii) except that program is not compiled (and hence rather slow).

*The references to this and other citations in Program Numbers 1-5 may be found on p. 394.

**Cheque must be drawn to a U.S. bank; UNESCO coupons (obtainable from UN outlets) are acceptable from countries where foreign exchange restrictions prevent issuance of US\$ cheques.

Table 1. Example of a length-frequency data file entered in the ELEFAN 0 program of the "Compleat ELEFAN" package as output through this same program (data from Corpuz et al. 1985); ML refers to "midlength", i.e., class midpoint.

Filename: MUMOR2

Upeneus moluccensis RAGAY GULF 1980

ML/Date	1/20/80	3/2/80	5/21/80	7/23/80	9/27/80	11/20/80	12/11/80
5.5	1						
6.5	1					3	
7.5	4					9	4
8.5	22		3			9	3
9.5	213		0	6	19	4	12
10.5	105	77	26	26	13	108	33
11.5	103	198	423	61	37	245	42
12.5	85	62	415	148	44	235	101
13.5	68	58	289	130	115	133	97
14.5	53	24	185	75	80	107	42
15.5	15		40	32	38	30	31
16.5	3		36	21	8	10	21
17.5			3			4	5
18.5							16
Sum	673	419	1,420	499	354	897	407

n = 4,669

- iv) the "Compleat ELEFAN" for IBM XT and compatibles; a fully integrated package implementing ELEFAN I, all other routines discussed in Pauly (Part I, this vol.) and many more, and structured around hi-resolution graphics for screen, printer and plotter output (see Figs. 1 and 2). Consists of 12 diskettes (US\$50) and a manual (date of issue: early 1988).
- c. Hardware Requirements. The "KIEL Package" (IBM PC, Apple II(CP/M) and TRS-80 versions) require two disk drives, video display and a printer (optional). The printer must be capable of printing graphics if the complementary diskette of the Apple II version is to be used.

The Compleat ELEFAN requires two 360 K disk drives or (preferably) one 360 K drive and one hard disk, a video display (with either a Hercules card or a color graphics adapter, or their compatibles), a graphics-capable printer (optional) and/or a two-pen plotter (HP7470A or IBM 7371, optional).

- d. RAM Requirements. For the "KIEL Package", the RAM requirements are as follows:

Version	RAM
IBM PC	102 K
Apple II (+,e,c) (with CP/M card)	40 K
TRS 80 Model I/III	38K

The RAM requirements for the "Compleat ELEFAN" are 512 K for all routines except VPA III of ELEFAN III (see Program Number 3).

Listing Earlier versions of ELEFAN I (and related programs) were disseminated in the form of photocopies of BASIC listings. This has now been abandoned, due to the length of the programs. Interested users may generate listings of the "KIEL Package" from the supplied disk(s). The Compleat ELEFAN is supplied only in compiled form, and readers are encouraged to report problems and/or bugs such that the master listings kept at ICLARM can be improved and/or corrected.

Test Data Tables with length-frequency data sets and sample runs are provided in the manuals sent by ICLARM along with the ELEFAN diskettes.

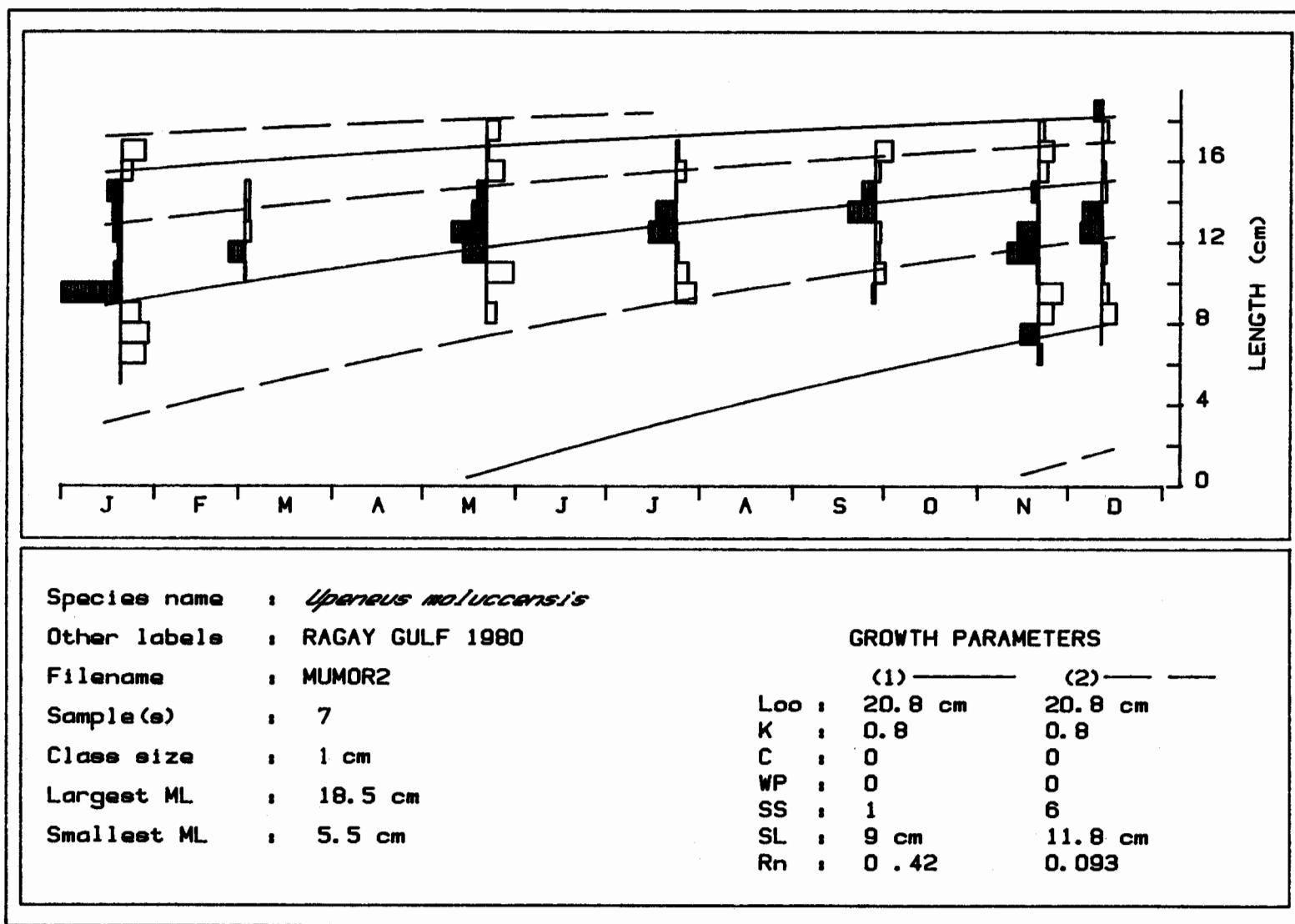


Fig. 1. Facsimile of a plotter output obtained with the ELEFAN I program of the "Compleat ELEFAN" package and representing growth curves of *Upeneus moluccensis* as estimated by ELEFAN I and superimposed on the "restructured" length-frequency data. Note presence of secondary growth curve (dotted line), pertaining to minor cohort. Based on length-frequency data in Table 1.

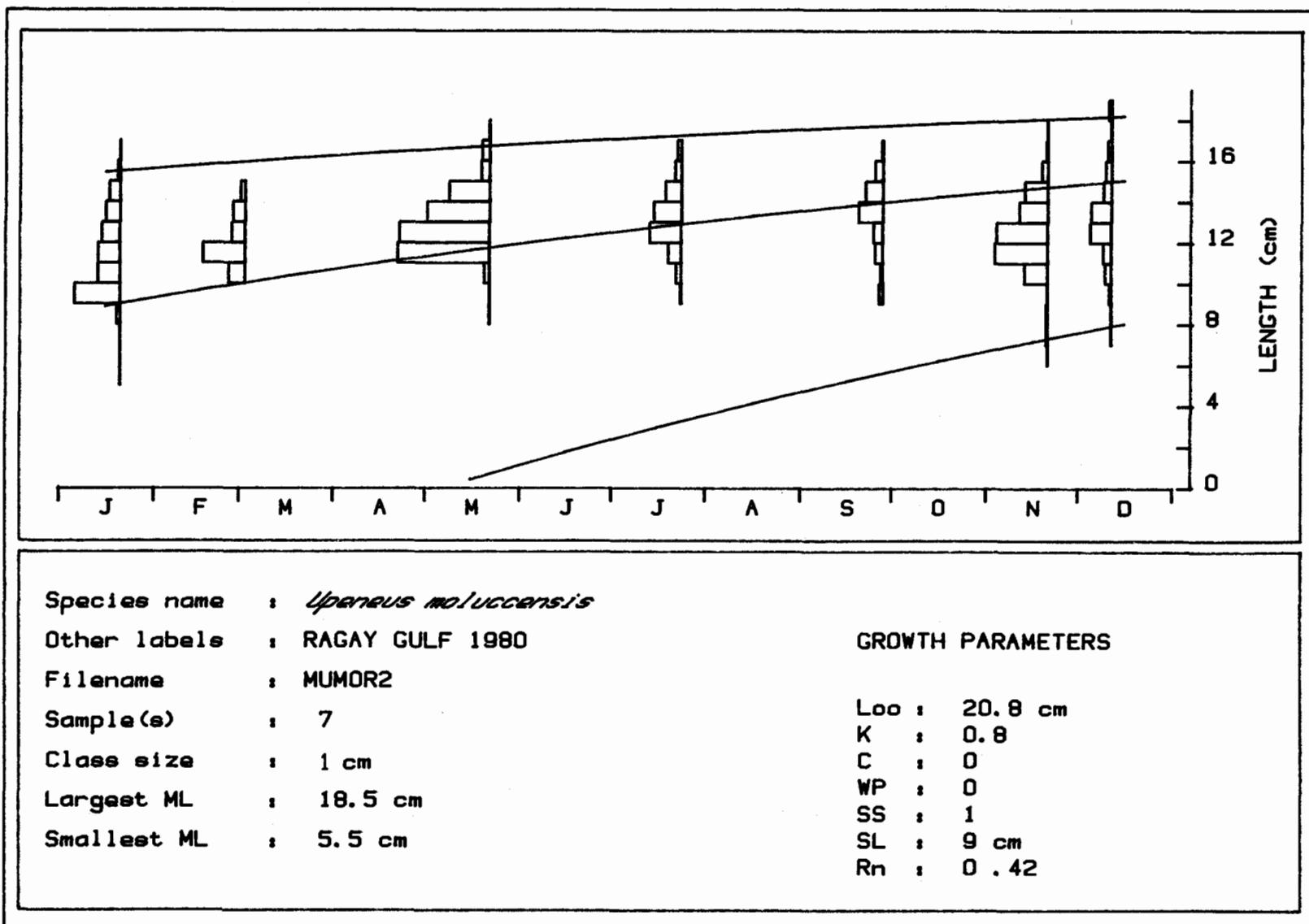


Fig. 2. Facsimile of a plotter output obtained with the ELEFAN I program of the "Compleat ELEFAN" package and presenting a growth curve for *Upeneus moluccensis* superimposed on the length-frequency data from which it was derived (see Table 1).

Program Number 2

Title ELEFAN II

Author D. Pauly

Author's Address Dr. Daniel Pauly
International Center for Living Aquatic Resources Management
(ICLARM)
MC P.O. Box 1501, Makati, Metro Manila, Philippines

References Pauly, D. 1982.—Studying singlespecies dynamics in a tropical multispecies context, p. 33-70. In D. Pauly and G.I. Murphy (eds.) Theory and management of tropical fisheries. ICLARM Conference Proceedings 9.

Pauly, D., J. Ingles and R. Neal. 1984.—Application to shrimp stocks of objective methods for the estimation of growth, mortality and recruitment-related parameters from length-frequency data (ELEFAN I and II), p. 220-234. In J.A. Gulland and B.I. Rothschild (eds.) Penaeid shrimps—their biology and management. Fishing News Books, Farnham, England. 312 p.

Pauly, D.—A review of the ELEFAN system for analysis of length-frequency data on fish and aquatic invertebrates. Part I, this vol.

Description ELEFAN II is a collection of routines for analysis of length-frequency data following, or immediately preceding estimation of growth parameters using ELEFAN II. The version of ELEFAN II included in the “KIEL Package” (see description of program number 1) includes:

- estimation of total mortality using length-converted catch curves and mean length in catch samples (Fig. 3)
- estimation of length-specific probabilities of capture based on detailed examination of ascending, left side of length-converted catch curves
- estimation of L_{∞} and Z/K using a modified version of the method of Wetherall et al. (Part I, this vol.).
- derivation of recruitment patterns as well as some miscellaneous other routines, e.g., for aggregating samples.

The version of ELEFAN II incorporated into the “Compleat ELEFAN” includes, in addition to the routines mentioned above (see Figs. 3, 4, 5 and 6) yield-per-recruit analysis for use with any pattern of selection (i.e., knife-edged, logistic, etc., see Fig. 7).

Listing Notes

- a. Language
- b. Implementation
- c. Hardware Requirements
- d. RAM Requirements

} same as for
ELEFAN I
(see Program Number 1)

Listing**Test Data**

LENGTH - CONVERTED CATCH CURVE

Filename : MUMOR2
L₀ : 20.8 cm
K : 0.8
to : 0

Cutoff length (L')
: 12 cm

Mean length (from L')
: 13.6 cm

Z from mean length
: 3.612

Mean temperature (C)
: 26.5

Natural mortality (M)
: 1.666

Fishing mortality (F)
: 3.067

Total mortality (Z)
: 4.733

Exploitation rate (E)
: 0.648

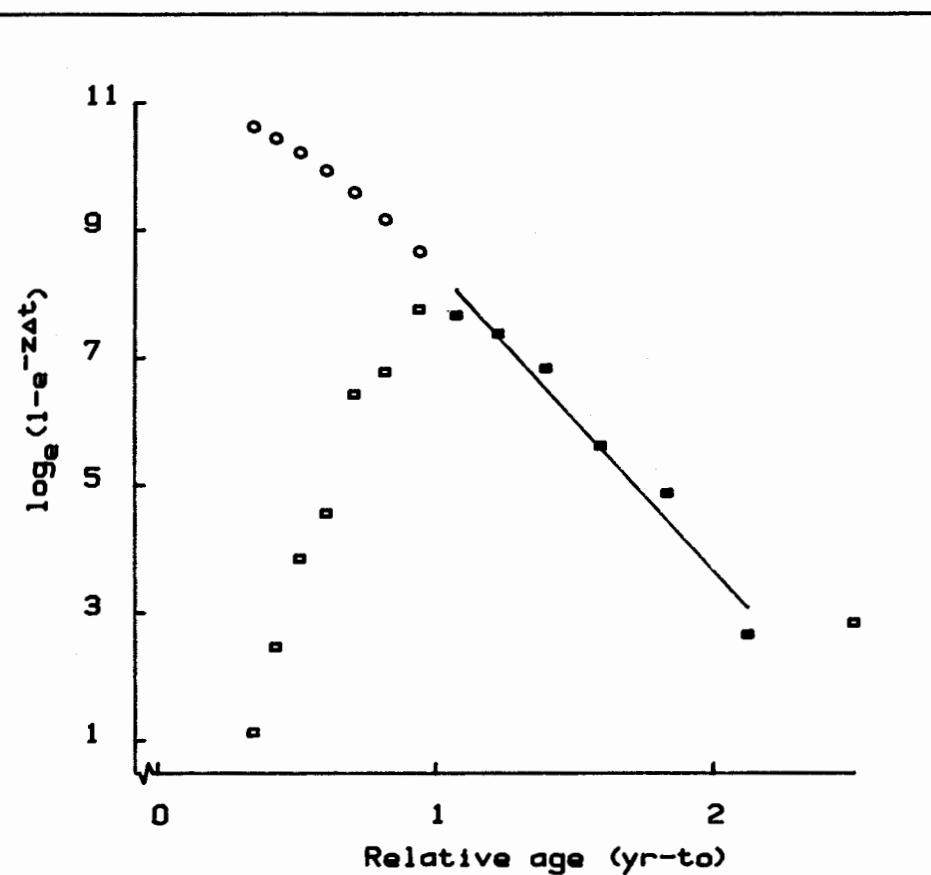


Fig. 3. Facsimile of a plotter output obtained with the ELEFAN II program of the "Compleat ELEFAN" package and pertaining to a length-converted catch curve based on the data in Table 1. Note that the natural mortality value used here was obtained through the empirical equation of Pauly (1980), which was incorporated into the "Compleat ELEFAN" package.

PROBABILITIES OF CAPTURE

Filename : MUMOR2
L_{oo} : 20.8 cm
K : 0.8
t_o : 0

Smallest midlength
: 5.5 cm

Largest midlength
: 18.5 cm

Class interval
: 1 cm

No. of classes
: 14

L₂₅ : 11.1 cm

L₅₀ : 11.83 cm

L₇₅ : 12.57 cm

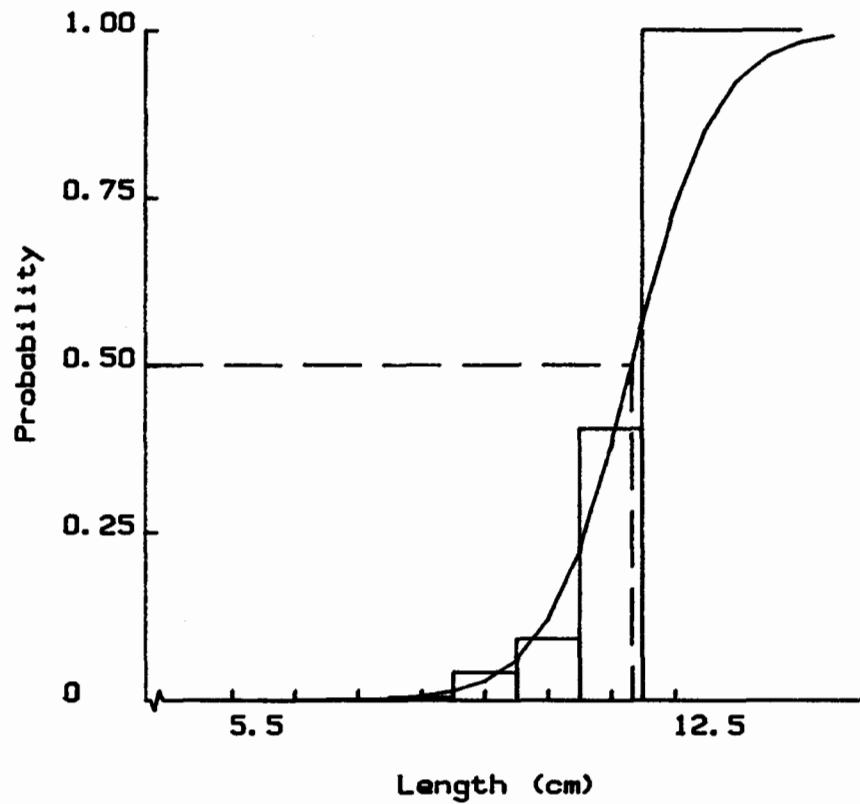


Fig. 4. Facsimile of a plotter output obtained with the ELEFAN II program of the "Compleat ELEFAN" package and pertaining to a resultant curve (the product of a selection and a recruitment curve), as derived from the left side of the length-converted catch curve in Fig. 3.

RECRUITMENT PATTERN

Filename : MUMOR2

Loo : 20.8 cm

K : 0.8

C : 0

WP : 0

to : 0

Smallest midlength: 5.5 cm

Largest midlength : 18.5 cm

Class interval : 1 cm

No. of classes : 14

Caution :

Recruitment patterns as generated by ELEFAN II:

1) suggest that recruitment pulses have a wider 'spread' in time than they really do.

2) allow for differentiating 1 or 2 recruitment pulses per year and not more.

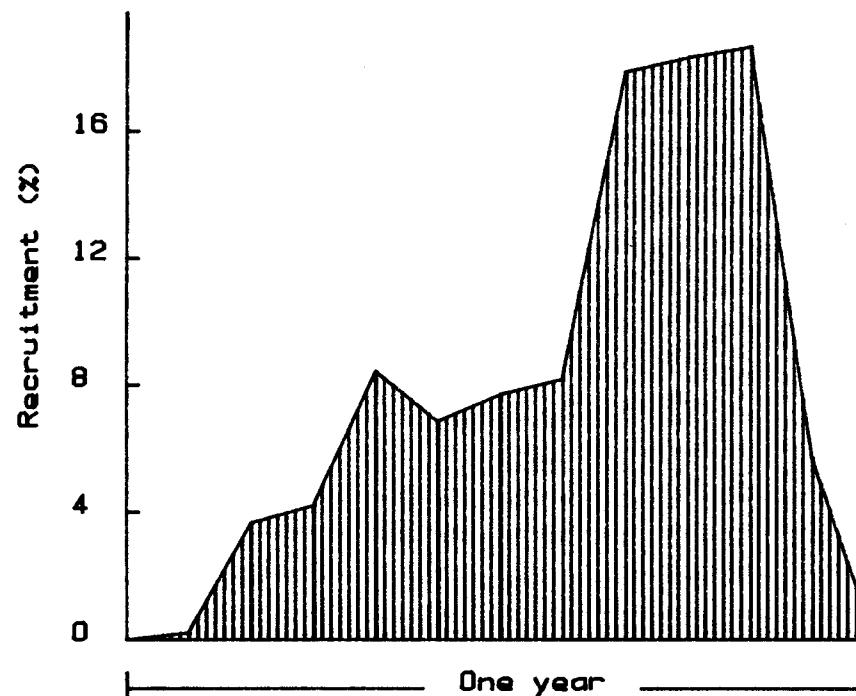


Fig. 5. Facsimile of a plotter output obtained with the ELEFAN II program of the "Compleat ELEFAN" package and pertaining to a recruitment pattern (as defined in Pauly, Part I, this vol.). Based on data in Table 1. Note caveats concerning interpretations of this output.

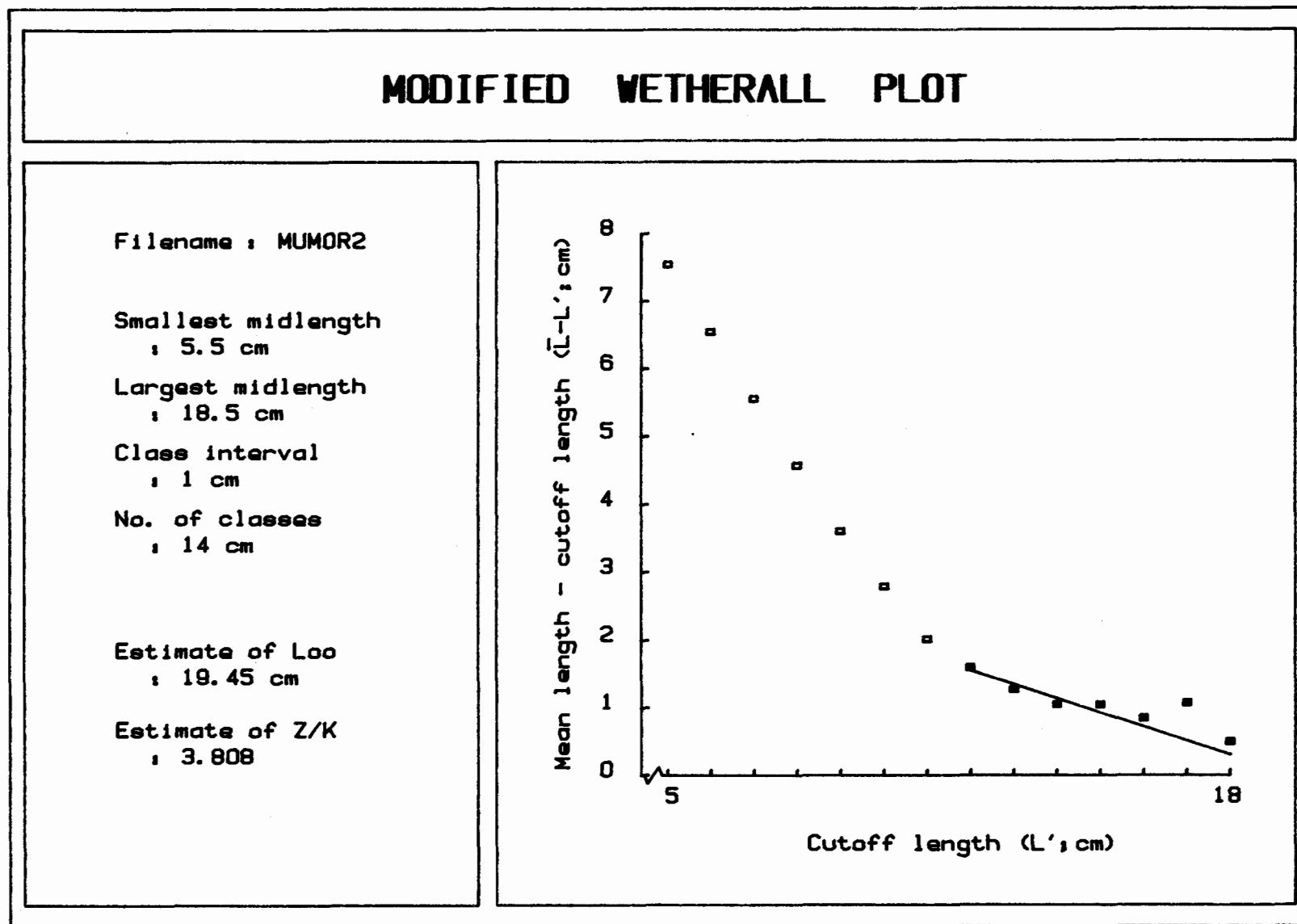


Fig. 6. Facsimile of a plotter output obtained with the ELEFAN II program of the "Compleat ELEFAN" package and presenting a modified Wetherall Plot (see Wetherall 1986, Pauly 1986 and Wetherall et al., Part I, this vol.) based on data in Table 1. Note that deviations from linearity by right-most points of this plot have little effect on results, due to ponderation scheme (which follows Wetherall et al., Part I, this vol.). Initial estimate of L_∞ obtained using this routine can be "fixed" in ELEFAN I, thus reducing the number of "free" parameters to be estimated by that program.

YIELD AND BIOMASS PER RECRUIT

E_{max} : 0.7005

$E_{.10}$: 0.6984

$E_{.50}$: 0.4245

L_c/L_{oo} : 0.5

M/K : 0.2

SELECTION :

Knife-edge
 Sel. curve

WARNING : Y/R and B/R are overestimated- especially at high values of exploitation rate (E) - when the 'knife-edge' assumption is unrealistic (which will most often be the case).

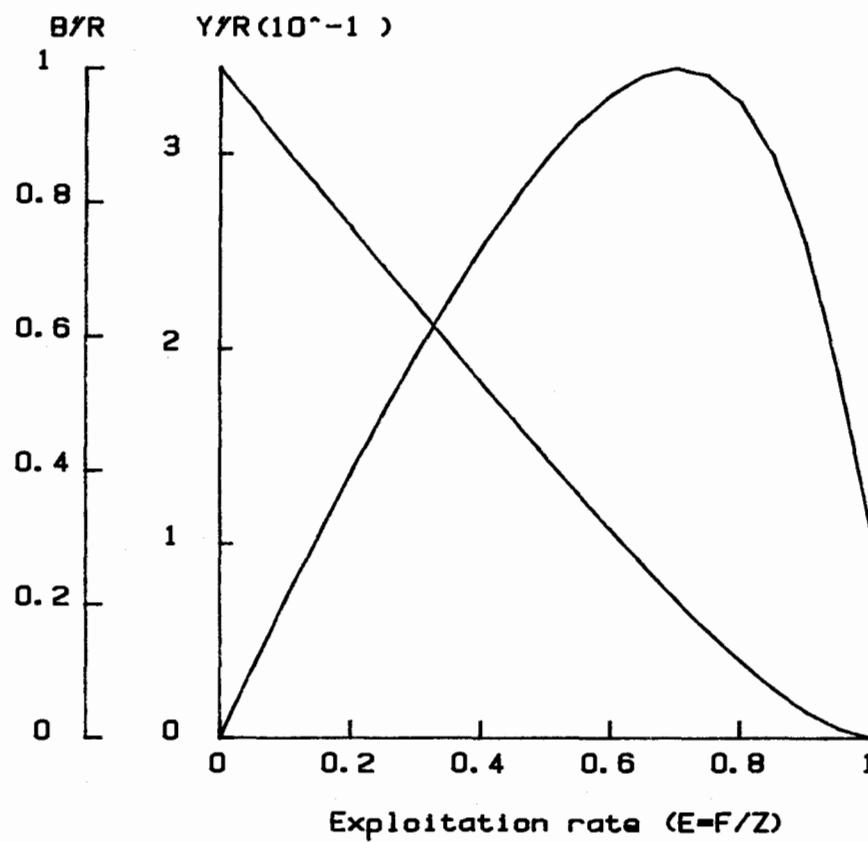


Fig. 7. Facsimile of a plotter output obtained with the ELEFAN II program of the "Compleat ELEFAN" package. This routine for (relative) yield- (and biomass)-per-recruit computations can take account of logistic (or other) selection patterns (as shown in Fig. 4) using a model presented in Pauly and Soriano (1986) and hence provide more realistic (and conservative) parameter estimates than using the standard "knife-edge" assumptions.

Program Number 3

Title ELEFAN III

Authors J.G. Pope, F.C. Gayanilo, Jr., D. Pauly

Authors' Address J.G. Pope
Fisheries Laboratory, Ministry of Agriculture, Fisheries and Food
Lowestoft, Suffolk NR33 OHT
England

F.C. Gayanilo, Jr. and D. Pauly^a
International Center for Living Aquatic Resources Management
(ICLARM)
MC P.O. Box 1501, Makati, Metro Manila
Philippines

References Pope, J.G., D. Pauly and N. David. 1981.—ELEFAN III, a BASIC program for the detailed analysis of catch at length-data using Virtual Population Analysis. ICLARM (MS).
Pauly, D., M.L. Palomares and F.C. Gayanilo, Jr. 1987.—VPA estimates of the monthly population, length composition, recruitment, mortality, biomass and related statistics of Peruvian anchoveta, 1953 to 1981, p. 142-166. In D. Pauly and I. Tsukayama (eds.) The Peruvian anchoveta and its upwelling ecosystem: three decades of change. ICLARM Studies and Reviews 15, Manila, 351 p.
Pauly, D.—A review of the ELEFAN system for analysis of length-frequency data on fish and aquatic invertebrates. Part I, this vol.

Description ELEFAN III as presently available only as a part of the “Compleat ELEFAN” package and can be used to run three different forms of Virtual Population Analysis (VPA), i.e.

- VPA I: the original version of (age-structured) VPA, as formulated by J.A. Gulland in 1965 (Fig. 8)
- VPA II: the VPA equivalent of R. Jones' length cohort analysis (Fig. 9)
- VPA III: an age-structured VPA run on “catch-at-age” data obtained by “slicing” from a catch-at-length matrix a set of “pseudocohorts”, using growth parameter estimates obtained with ELEFAN I (Figs. 10, 11)

The references given above give further details, notably on data requirements, assumptions and potential pitfalls.

Listing Notes

- a. Language
b. Implementation
c. Hardware Requirements
d. RAM Requirements

} same as “Compleat ELEFAN”
version of ELEFAN I

640 K

Listing

} same as “Compleat ELEFAN”
version of ELEFAN I

Test Data

(see Program Number 1)

^aTo whom all correspondence concerning this program should be addressed.

VIRTUAL POPULATION ANALYSIS I

File ID
- Sardine (Mor.)

Natural mortality
- 0.8

Term. f. mortality
- 0.8

Time interval
- 0.25

No. of age grps.
- 17

LEGENDS:
 catch
 Nat. losses
 survivors
 fishing mortality

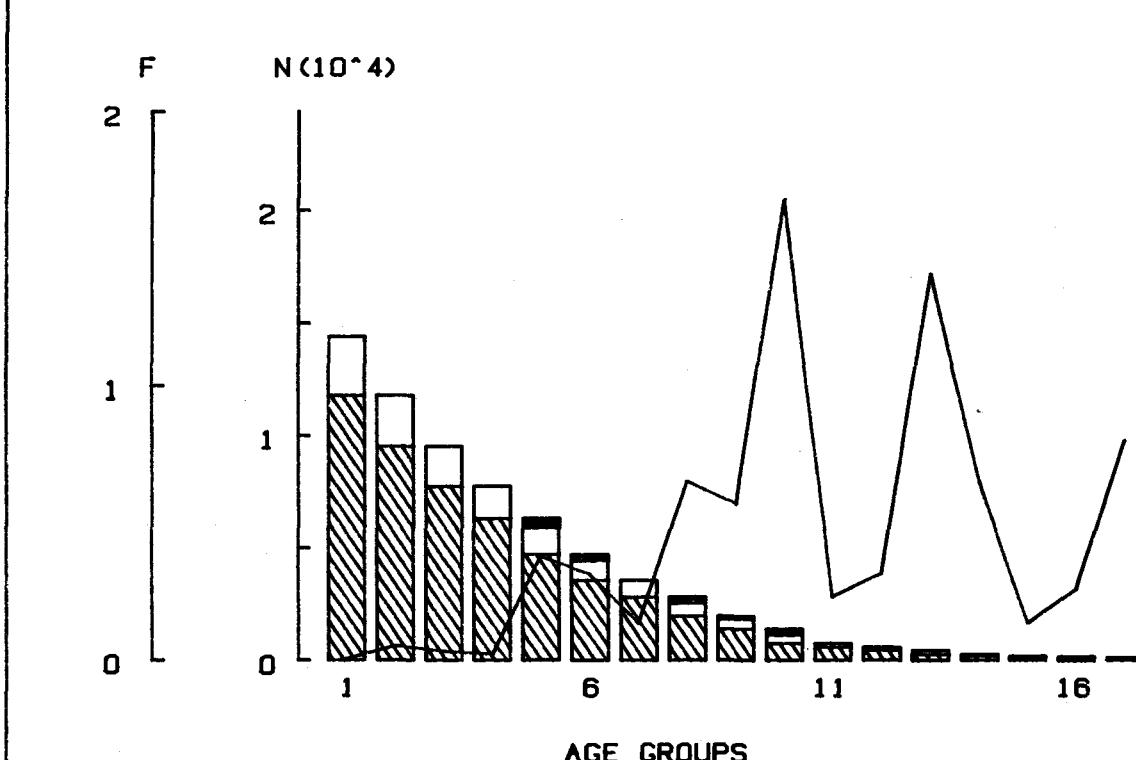


Fig. 8. Facsimile of a plotter output generated by the ELEFAN III program of the "Compleat ELEFAN" package, and representing a standard age-structured VPA (here VPA I). The data used (from Anon. 1978a, Table 1, p. 33) pertain to Moroccan sardine *Sardina pilchardus*. Note that rates and the time interval refer to one year.

VIRTUAL POPULATION ANALYSIS II

Filename
= MERSEN78

Natural mortality
= 0.28

Term. f. mortality
= 0.28

L_{∞} (Loo)
= 130 cm

K growth constant
= 0.1

Mean E
= 0.620

Mean F
= 0.457

LEGENDS:

- catch
- Nat. losses
- ▨ survivors
- fishing mortality

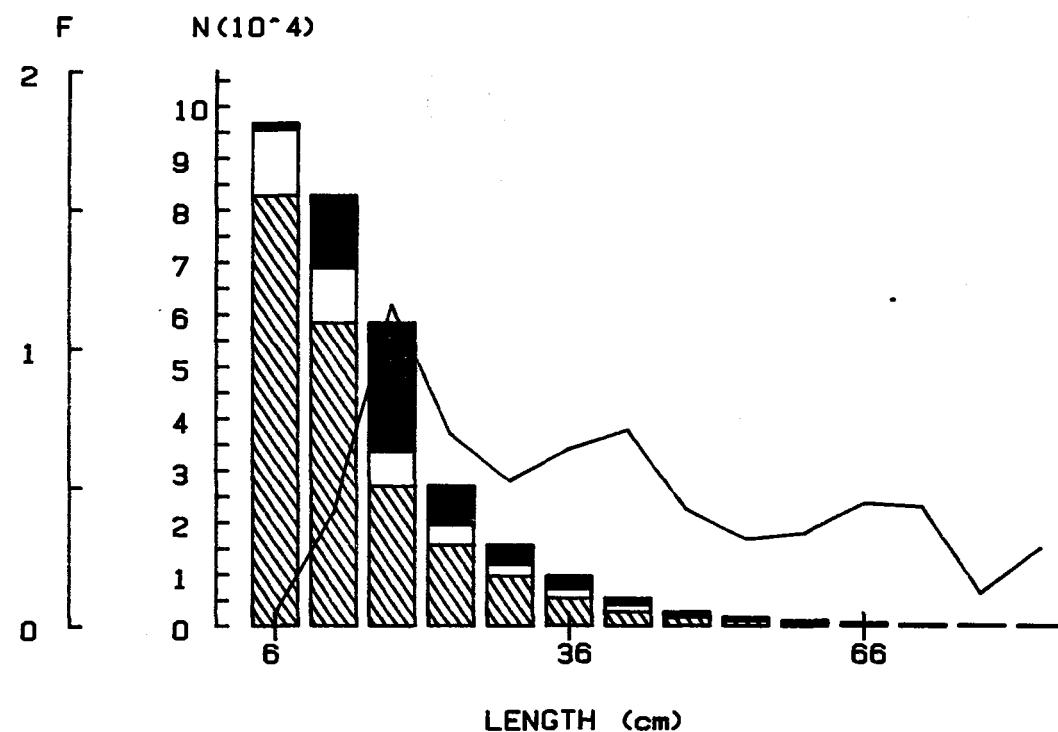


Fig. 9. Facsimile of a plotter output generated by the ELEFAN III program of the "Compleat ELEFAN" package, and representing a form of VPA analogous to R. Jones' length-cohort analysis. The data used (from Anon. 1978b, Table 6, p. 78) pertain to a West African hake *Merluccius merluccius* stock. Note that rates are annual.

VIRTUAL POPULATION ANALYSIS III

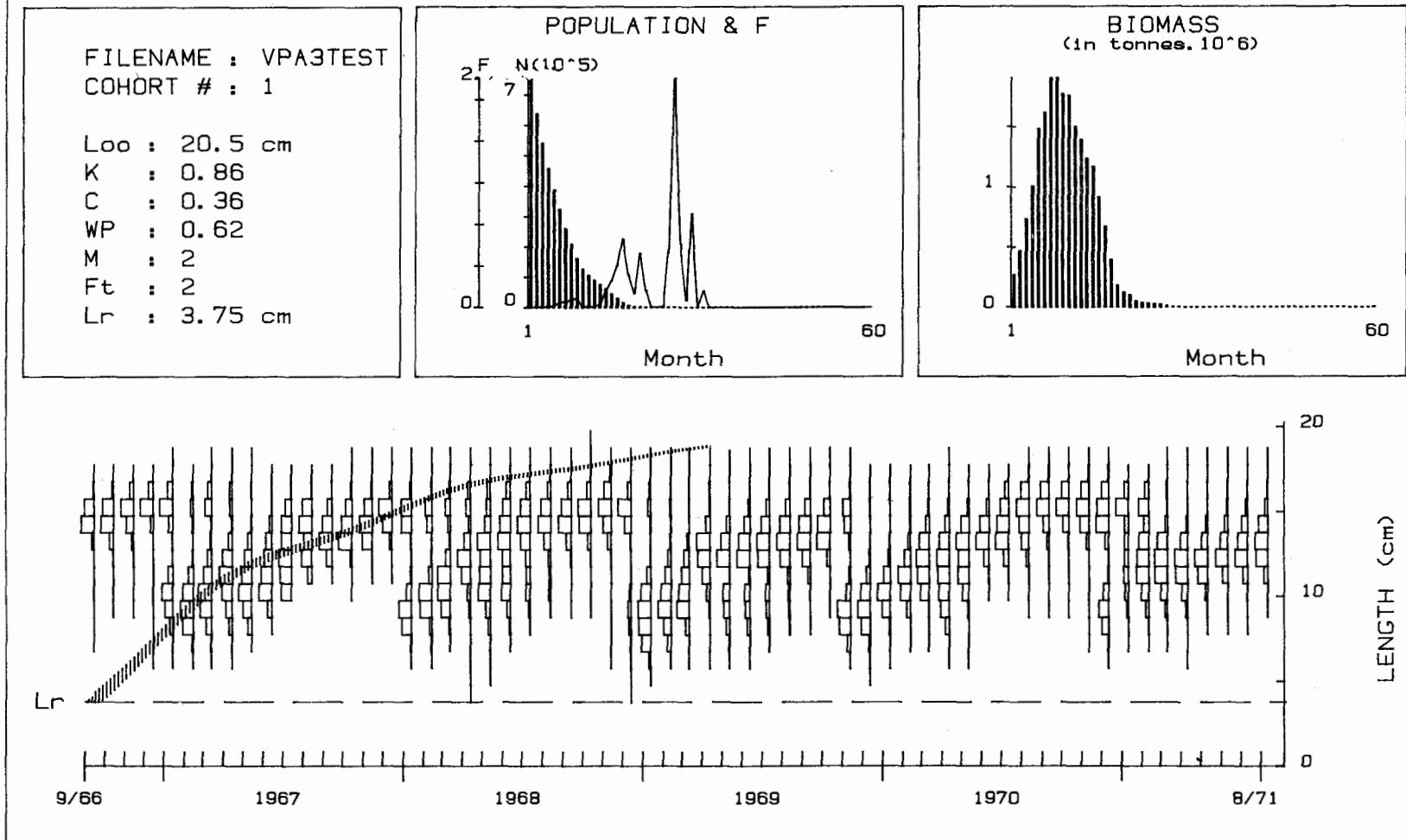


Fig. 10. Facsimile of a plotter output generated by the ELEFAN III program of the "Compleat ELEFAN" package, and representing a "pseudocohort" of Peruvian anchoveta (*Engraulis ringens*) sliced by ways of a set of growth parameters from a matrix of catch-at-length data. Left upper box summarizes input parameters, central and right upper boxes present monthly population, F and biomass estimate for this pseudocohort. Complete analysis involves successive analyses of all adjacent pseudocohorts (see references for details on method).

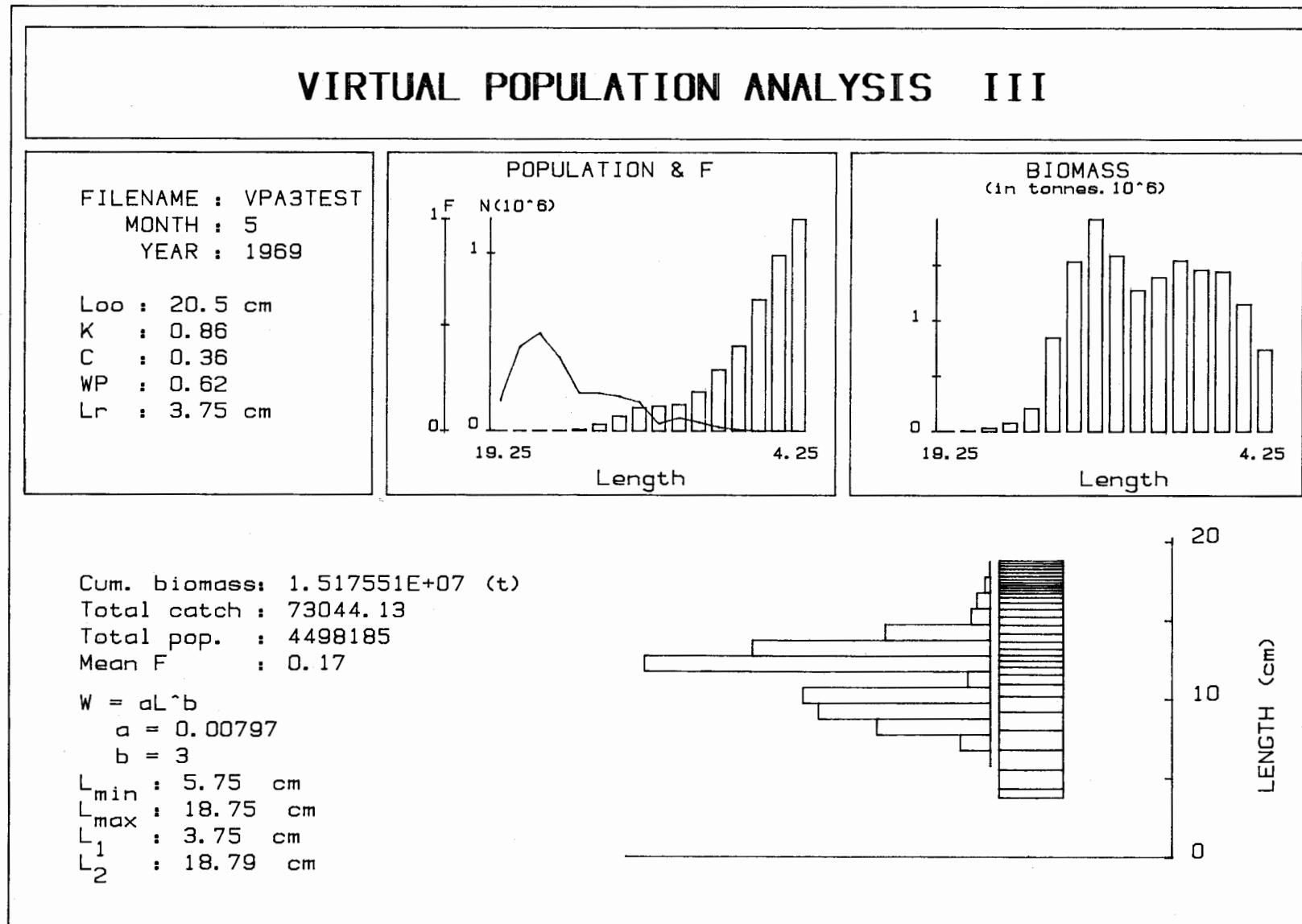


Fig. 11. Facsimile of a plotter output generated by the ELEFAN III program of the "Compleat ELEFAN" package, and representing a reconstruction of the population structure of the Peruvian anchoveta (*Engraulis ringens*) during May 1969, as obtained by combining the population estimates for adjacent "pseudocohorts" (see shaded area, right side of lower box). Details on this approach may be found in the references (see also Fig. 10).

Program Number 4

Title ELEFAN IV

Authors J.L. Munro, F.C. Gayanilo, Jr., M. Soriano and D. Pauly

Authors' Address J.L. Munro
ICLARM South Pacific Office
P.O. Box 438
Honiara, Solomon Islands

F.C. Gayanilo, Jr., M. Soriano and D. Pauly^a
International Center for Living Aquatic Resources Management
(ICLARM)
MC P.O. Box 1501, Makati, Metro Manila, Philippines

References Munro, J.L. 1984.—Estimation of natural mortality rates from selectivity and catch length-frequency data. Fishbyte 2(1): 11-14.
Pauly, D.—A review of the ELEFAN system for analysis of length-frequency data on fish and aquatic invertebrates. Part I, this vol.

Description This program uses length-frequency data representing a steady-state population (as can be obtained, e.g., by pooling a long series of monthly samples) and selection curves obtained from selection experiments to estimate the total natural and fishing mortality rates of stocks of fish and invertebrates. (See Pauly, Part I, this vol.). The program is available only as a part of the “Compleat ELEFAN” package.

Listing Note

- a. Language
- b. Implementation
- c. Hardware Requirements
- d. RAM Requirements

} same as “Compleat ELEFAN”
version of ELEFAN I
(Program Number 1)

Listing

Test Data Presently, only an artificial data set is available to test and demonstrate the program. The manual of the “Compleat ELEFAN” will contain, however, at least one set of actual data (courtesy J. Moreau, Toulouse) to illustrate both the versatility and the potential pitfalls of the program.

^aTo whom all correspondence concerning this paper should be addressed.

Program Number 5

Title ELEFAN V

Authors G.R. Morgan, M. Soriano and D. Pauly

Authors' Address G.R. Morgan^a
Kuwait Institute for Scientific Research
P.O. Box 24885 Safat, Kuwait

M. Soriano and D. Pauly
International Center for Living Aquatic Resources Management
(ICLARM)
MC P.O. Box 1501, Makati, Metro Manila, Philippines

References Morgan, G.R.—Incorporating age data into length-based stock assessment methods. Part I, this vol.

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. f. Meereskd. Univ. Kiel No. 149, 76 p.

Description As originally developed and implemented by G.R. Morgan, this program allows for the joint analysis of length-frequency data and length-at-age data. The original implementation is available in the form of HP 9845B and IBM PC programs from G.R. Morgan.

A slightly modified version of G.R. Morgan's program has been incorporated in the "KIEL Package" (see Programs Number 1 and 2). The package thus allows entry and analysis (jointly with L/F data) of length-at-age data and/or tagging-recapture data.

The ELEFAN V program as such is a part of the "Compleat ELEFAN" package, structured around the estimation of growth parameters through joint analysis of length-frequency and growth increment data. Growth increment data may come from three sources:

- i) tagging-recapture experiments,
- ii) transformed length-at-age data,
- iii) successive mean lengths of age groups identified through separation of length-frequency samples into normally distributing components.

A routine for entry and manipulation of data in (i) and (ii) is incorporated in the ELEFAN 0 routine of the "Compleat ELEFAN", which also contains a routine ("Modal Progression Analysis" or MPA) for decomposition of multimodal length-frequency samples into their normally distributed components (Fig. 12).

^aPresent address: c/o W.A. Marine Research Laboratories, P.O. Box 20, North Beach 6020, W.A. Australia.

BHATTACHARYA'S METHOD OF SEPARATING NORMAL DISTRIBUTIONS

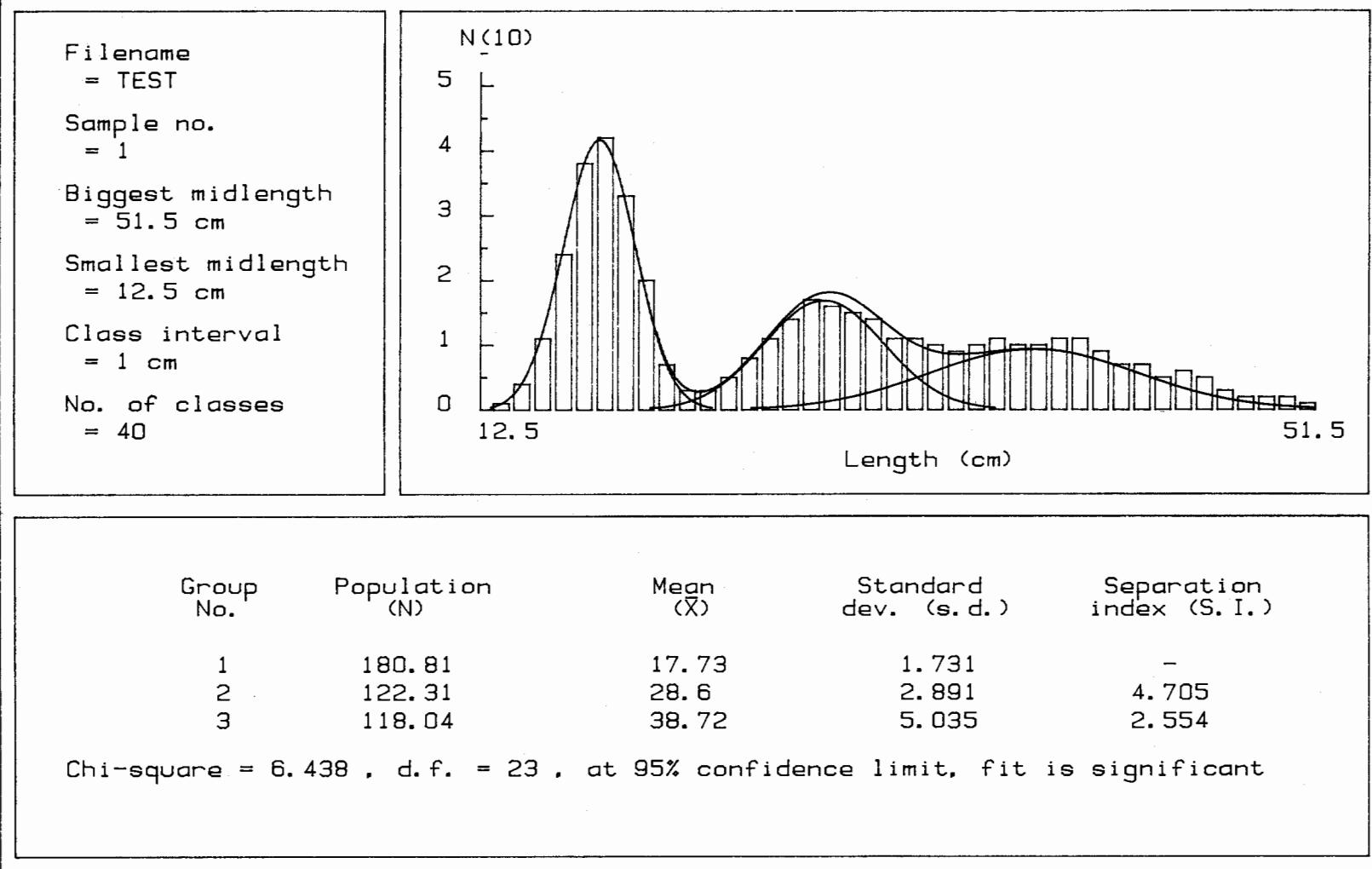


Fig. 12. Facsimile of a plotter output generated by the MPA program of the "Compleat ELEFAN" package, and showing the decomposition, using the method of Bhattacharya (1967), of an artificial length-frequency sample into its component normal distributions. Note summary of information on decomposition, allowing quality control (separation index should be > 2 and χ^2 lower than critical value).

The ELEFAN V program also allows estimation of growth parameters from growth increment data only (i.e., without simultaneously using L/F data). The appropriate routines—several of them new developments—allow explicit consideration of seasonal growth, and estimation of standard errors for all parameter estimates.

Listing Notes

- a. Language
- b. Implementation
- c. Hardware Requirements
- d. RAM Requirements

Listing

Test Data



Consult G.R. Morgan for documentation of original (HP9845B) version, and see ELEFAN I (Program Number 1) for details on the package of which ELEFAN V is a part

Literature Cited (Program Numbers 1-5)

- Anon. 1978a. Report of the *ad hoc* Working Group on Sardine (*Sardina pilchardus* Wald.). CECAF/ECAF Ser./78/7, 35 p. FAO, Rome.
- Anon. 1978b. Report of the *ad hoc* Working Group on Hake (*Merluccius merluccius*, *M. senegalensis*, *M. cadenati*) in the Northern Zone of CECAF. CECAF/ECAF Ser./78/19, 93 p. FAO, Rome.
- Bhattacharya, C.G. 1967. A simple method of resolution of a distribution into Gaussian components. *Biometrics* 23: 115-135.
- Corpuz, A., J. Saeger and V. Sambilay, Jr. 1985. Population parameters of commercially important fishes in Philippine waters. University of the Philippines in the Visayas, College of Fisheries. Tech. Rep. Dept. Mar. Fish 6. 99 p.
- Liew, H.C. 1986. ELEFAN with hi-res graphics for Apple II (CP/M). *Fishbyte* 4(3): 10-12.
- Pauly, D. 1986. On improving operation and use of the ELEFAN programs. Part II: improving the estimation of L_∞ . *Fishbyte* 4(1): 18-20.
- Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters and mean environmental temperature in 175 fish stocks. *J. Cons., Cons. Int. Explor. Mer* 39(3): 175-192.
- Pauly, D. and M. Soriano. 1986. Some practical extensions to Beverton and Holt's relative yield-per-recruit model, p. 491-496. In J.L. Maclean, L.B. Dizon and L.V. Hosillos (eds.) *The First Asian Fisheries Forum*. Asian Fisheries Society, Manila.
- Wetherall, J.A. 1986. A new method for estimating growth and mortality parameters from length-frequency data. *Fishbyte* 4(1): 12-14.

Program Number 6

Title ZKAN and ZKA E2

Authors T.A. Ebert. Modification by U. Damm.

Authors' Address

Dr. T.A. Ebert
 Department of Biology
 San Diego State University
 San Diego, California 98182
 USA

and

Dr. U. Damm
 Institut für Küsten-und
 Binnenfischerei
 Laboratorium Cuxhaven
 Bei der Alten Liebe 1
 D-2190 Cuxhaven
 Federal Republic of Germany

References Ebert, T.A.—Estimating growth and survival parameters by nonlinear regression using average size in catches. Part I, this vol.

Damm, U.—Some modifications of Ebert's method to calculate growth and mortality parameters from average lengths in a population. Part I, this vol.

Description Ebert's program uses the change in average size during a year to estimate, using a nonlinear approach, growth parameters and the total mortality rate of a population when recruitment is periodic. Mortality rate is assumed constant over the period considered while the growth parameters estimated are those of the generalized Richards function with the von Bertalanffy parameters being a special case. The data must be a sequence of sizes that show a positive slope; average size must increase with time. The month with the smallest average size usually should be selected as the time of recruitment, $t = 0$ and the estimated mean of the first mode of the size distribution of this same distribution should be selected as the size at recruitment, L_R . Other times of recruitment are assigned as fraction of a year relative to $t = 0$.

Damm's modification utilizes a linear regression of average size against time to estimate the parameters of the von Bertalanffy growth equation in addition to allowing for the possibility of two recruitment episodes per year.

Listing Notes

- a. Language Applesoft BASIC (DOS 3.3)
- b. Implementation Apple II+
- c. Hardware Requirements Single disk drive, video display, printer (optional). If a printer is used the card is assumed to be in slot number 1, the disk controller card is assumed to be in slot number 6.
- d. RAM Requirements Ebert about 38 K, Damm about 26 K.

Listing ZKAN

The following listing is for the version of the program revised March 1985 for running on an Apple II+ machine.

Program can be compiled with a TASC compiler (Microsoft). The program provides error codes and the line where the error occurred. The codes refer to p. 136 in the 1981 APPLESOFT II BASIC Programming Reference Manual, APPLE product #A2L0006 (030-0013-E), or Table C-1 (p. 331) in APPLE II User's Guide by L. Poole, M. McNiff and S. Cook (1981) published by Osborne/McGraw-Hill, Berkeley, California, USA. Common error codes are 53 (illegal quantity), 69 (overflow), 133 (division by zero).

Fig. 13. Program Listing ZKAN.

```

10 REM LINE 30 IS AN ACTIVE REM USED BY THE TASC COMPILER.
20 REM LINE 30 IS NOT NEEDED FOR AN INTERPRETED BASIC.
30 REM I,N,R,C1,J,KK,L,IC,ID,IE,IG,M
40 D$ = CHR$(4)
50 DIM T(50),S(50),F(50,6),A(6,6),M(6,6),B(6,6),E(6),C(50)
60 ONERR GOTO 1800
70 HOME : VTAB 10: HTAB 3: PRINT "PROGRAM TO ESTIMATE Z, K,
      ASYMPTOTIC"
80 HTAB 3: PRINT "SIZE, AND THE SHAPE PARAMETER OF THE"
90 HTAB 3: PRINT "RICHARDS FUNCTION BY NONLINEAR"
100 HTAB 3: PRINT "REGRESSION."
110 PRINT
120 HTAB 5: PRINT "T. A. EBERT, JANUARY 1984"
130 PRINT : PRINT : INPUT "<RETURN> TO CONTINUE....";F$
140 HOME : VTAB 10
150 MG = - LOG (.00001)
160 PRINT "DO YOU WANT TO ENTER DATA FROM DISK (D)"
170 INPUT "OR KEYBOARD (K)? (D/K) ";Y$
180 IF Y$ = "D" GOTO 270
190 HOME : VTAB 3: INPUT "NUMBER OF DATA PAIRS = ";N
200 FOR I = 1 TO N: PRINT : INPUT "AVERAGE SIZE = ";S(I)
210 INPUT "AT T = ";T(I): PRINT : NEXT I
220 HOME : VTAB 10: INPUT "DO YOU WANT DATA STORED ON DISK? (Y/N) ";
      Y$
230 IF Y$ = "N" GOTO 340
240 PRINT: INPUT "FILENAME FOR DATA = ";F$: INPUT "ON DRIVE NUMBER = ";
      D
250 PRINT D$;"OPEN";F$;"D";D: PRINT D$;"WRITE";F$: PRINT N: FOR I = 1 TO
      N: PRINT S(I): PRINT T(I): NEXT I: PRINT D$;"CLOSE";F$
260 GOTO 340

```

Fig. 13. Continued

```

270 HOME : VTAB 10
280 INPUT "NAME OF FILE WITH DATA? ";F$
290 INPUT "ON DRIVE #? ";D
300 PRINT D$;"OPEN";F$;";D";D
310 PRINT D$;"READ";F$
320 INPUT N: FOR I = 1 TO N: INPUT S(I): INPUT T(I): NEXT I
330 PRINT D$;"CLOSE";F$
340 HOME : VTAB 10
350 PRINT "DO YOU WANT TO CONTINUE": INPUT "WITH THIS SAME DATA SET
(Y/N)? ";F$
360 IF F$ = "Y" GOTO 400
370 HOME : VTAB 10: PRINT "DO YOU WANT TO CONTINUE": INPUT "WITH A
NEW FILE (Y/N)? ";F$
380 IF F$ = "N" GOTO 1850
390 GOTO 140
400 IC = 0:ID = 0:IE = 0:IG = 0: HOME : VTAB 10
410 R = 0: INPUT "DO YOU WANT TO ESTIMATE Z (Y/N)? ";Y$
420 IF Y$ = "Y" THEN IC = R + 1:R = R + 1
430 HOME : VTAB 10
440 INPUT "DO YOU WISH TO ESTIMATE K (Y/N)? ";Y$
450 IF Y$ = "Y" THEN ID = R + 1:R = R + 1
460 HOME : VTAB 10
470 INPUT "DO YOU WANT TO ESTIMATE ASYMPTOTIC SIZE (Y/N)? ";Y$
480 IF Y$ = "Y" THEN IE = R + 1:R = R + 1
490 HOME : VTAB 10
500 INPUT "DO YOU WANT TO ESTIMATE N, THE SHAPE PARAMETER? (Y/N) ";
Y$
510 IF Y$ = "Y" THEN IG = R + 1:R = R + 1
520 HOME : VTAB 5
530 HTAB 5: PRINT "INITIAL VALUES OF PARAMETERS"
540 HTAB 10: INPUT "Z = ";Z: HTAB 10: INPUT "K = ";K
550 HTAB 10: INPUT "MAXIMUM SIZE = ";A
560 HTAB 10: INPUT "SIZE AT RECRUITMENT = ";SR
570 HTAB 10: INPUT "SHAPE PARAMETER, N = ";NR
580 REM F(I,C1) IS S(I) - F, THE VALUE OF THE FUNCTION
590 REM F(I,IC) IS THE PARTIAL WITH RESPECT TO Z
600 REM F(I,ID) IS THE PARTIAL WITH RESPECT TO K
610 REM F(I,IE) IS THE PARTIAL WITH RESPECT TO A, ASYMPTOTIC SIZE
620 REM F(I,IG) IS THE PARTIAL WITH RESPECT TO N, THE SHAPE
PARAMETER

```

Continued

Fig. 13. Continued.

```

630 HOME : VTAB 10
640 PRINT "DO YOU WANT RESULTS SENT TO A": INPUT "PRINTER? (Y/N) ";
P$
650 IF P$ < > "Y" THEN GOTO 680
660 HOME : VTAB 10: INPUT "TURN ON PRINTER, THEN <RETURN>";Y$
670 PRINT D$;"PR#1"
680 C1 = R + 1
690 ZS = 1000:KS = 1000:AS = 1000:NS = 1000
700 HTAB 5: PRINT "ESTIMATES OF PARAMETERS"
710 HOME : PRINT SPC( 3); "Z"; SPC( 8); "K"; SPC( 8); "A"; SPC( 8); "N"
720 PRINT
730 P1 = EXP (Z) - 1:P3 = 1 - EXP ( - Z)
740 D = - 1 / NR
750 A9 = LOG (A / SR):A8 = (A / SR) ^ ( - D)
760 FOR I = 1 TO N: FOR J = 0 TO C1:F(I,J) = 0: NEXT J: NEXT I
770 B = (A ^ D - SR ^ D) / A ^ D
780 KK = INT (MG / Z + .5)
790 FOR J = 1 TO N
800 FOR I = 0 TO KK
810 A1 = EXP ( - Z * I)
820 A6 = EXP ( - K * (I + T(J)))
830 B1 = 1 - B * A6
840 A2 = B1 ^ ( - NR)
850 F(J,C1) = F(J,R + 1) + A1 * A2
860 A3 = I * A1
870 F(J,IC) = F(J,IC) + A3 * A2
880 A4 = B1 ^ ( - (NR + 1))
890 A5 = (I + T(J)) * A6
900 F(J,ID) = F(J,ID) + A1 * A4 * A5
910 F(J,IE) = F(J,IE) + A1 * A6 * A4
920 A7 = - LOG (B1)
930 F(J,IG) = F(J,IG) + A1 * A2 * (A7 + A6 * A8 * A9 / (NR * B1))
940 NEXT I
950 F(J,IC) = A * (F(J,R + 1) - P1 * F(J,IC)) / EXP (Z)
960 F(J,C1) = P3 * F(J,C1)
970 F(J,ID) = - A * NR * B * P3 * F(J,ID)
980 F(J,IE) = - A * P3 / SR * (A / SR) ^ (1 / NR - 1) * F(J,IE) + F(J,C1)
990 F(J,IG) = A * P3 * F(J,IG)
1000 F(J,C1) = S(J) - A * F(J,C1)
1010 NEXT J

```

Continued

Fig. 13. Continued

```

1020 REM MULTIPLY THE MATRIX F BY ITS TRANSPOSE.
1030 FOR I = 1 TO R: FOR J = 1 TO C1:A(I,J) = 0: NEXT J: NEXT I
1040 FOR I = 1 TO R: FOR J = 1 TO C1: FOR L = 1 TO N
1050 A(I,J) = A(I,J) + F(L,I) * F(L,J): NEXT L: NEXT J: NEXT I
1060 GOSUB 1160
1070 FOR I = 0 TO R:E(I) = 0.: NEXT I
1080 FOR I = 1 TO R
1090 FOR G = 1 TO R
1100 E(I) = E(I) + B(I,G) * A(G,C1): NEXT G: NEXT I
1110 Z = Z + E(IC):K = K + E(ID):A = A + E(IE):NR = NR + E(IG)
1120 PRINT INT (Z * 10000 + .5) / 10000; SPC( 3); INT (K * 10000 + .5) /
    10000; SPC( 3); INT (A * 10000 + .5) / 10000; SPC( 3); INT (NR *
    10000 + .5) / 10000
1130 IF ABS (ZS - Z) < .00001 AND ABS (KS - K) < .00001 AND ABS (AS -
    A) < .00001 AND ABS (NS - NR) < .00001 THEN GOTO 1340
1140 ZS = Z:KS = K:AS = A:NS = NR
1150 GOTO 730
1160 FOR I = 1 TO R: FOR J = 1 TO R:B(I,J) = 0: NEXT J: NEXT I
1170 FOR J = 1 TO R
1180 B(J,J) = 1: NEXT J
1190 FOR J = 1 TO R: FOR I = J TO R
1200 IF A(I,J) < > 0 THEN 1230
1210 NEXT I
1220 PRINT D$;"PR#0": PRINT "SINGULAR MATRIX": GOTO 1770
1230 FOR M = 1 TO R
1240 S = A(J,M):A(J,M) = A(I,M):A(I,M) = S:S = B(J,M):B(J,M) = B(I,M):B(I,M) =
    S: NEXT M
1250 T = 1 / A(J,J)
1260 FOR M = 1 TO R:A(J,M) = T * A(J,M):B(J,M) = T * B(J,M): NEXT M
1270 FOR L = 1 TO R
1280 IF L = J THEN 1310
1290 T = - A(L,J)
1300 FOR M = 1 TO R:A(L,M) = A(L,M) + T * A(J,M):B(L,M) = B(L,M) + T *
    B(J,M): NEXT M
1310 NEXT L
1320 NEXT J
1330 RETURN
1340 PRINT : PRINT "-----": PRINT : PRINT
    SPC(10); "FINAL VALUES"
1350 IF IC < > 0 THEN PRINT "Z = ";Z

```

Fig. 13. Continued

```

1360 IF ID < > 0 THEN PRINT "K = ";K
1370 IF IE < > 0 THEN PRINT "A = ";A
1380 IF IG < > 0 THEN PRINT "N = ";NR
1390 PRINT : PRINT
1400 IF N - R < 1 GOTO 1530
1410 SS = 0
1420 FOR I = 1 TO N:SS = SS + F(I,C1) * F(I,C1): NEXT I
1430 PRINT "RESIDUAL SS = ";SS
1440 MS = SS / (N - R)
1450 PRINT : PRINT "MEAN SQUARE ERROR = ";MS
1460 FOR I = 1 TO R:M(I,I) = SQR (B(I,I) * MS): NEXT I
1470 PRINT : PRINT "STANDARD DEVIATIONS OF PARAMETERS": PRINT
    "WITH ";N - R;" DF"
1480 PRINT : PRINT
1490 IF IC < > 0 THEN PRINT "SD FOR Z = ";M(IC,IC)
1500 IF ID < > 0 THEN PRINT "SD FOR K = ";M(ID,ID)
1510 IF IE < > 0 THEN PRINT "SD FOR A = ";M(IE,IE)
1520 IF IG < > 0 THEN PRINT "SD FOR N = ";M(IG,IG)
1530 IF P$ = "Y" THEN PRINT : PRINT : GOTO 1560
1540 PRINT : INPUT "<RETURN> TO CONTINUE...";F$
1550 HOME : VTAB 5
1560 PRINT "      VALUES FOR PLOTTING": PRINT
1570 PRINT " TIME"," SIZE"
1580 PRINT
1590 FOR J = 0 TO 11
1600 F = 0
1610 FOR I = 0 TO KK
1620 A1 = EXP (- Z * I)
1630 A6 = EXP (- K * (I + J / 12))
1640 B1 = 1 - B * A6
1650 A2 = B1 ^ (- NR)
1660 F = F + A1 * A2: NEXT I
1670 F = F * P3 * A
1680 PRINT INT (J / 12 * 100 + .5) / 100, INT (F * 1000 + .5) / 1000
1690 NEXT J
1700 IF P$ = "Y" THEN PRINT : PRINT : GOTO 1730
1710 PRINT : INPUT "<RETURN> TO CONTINUE...";F$
1720 HOME : VTAB 5
1730 PRINT " RESIDUALS": PRINT
1740 PRINT "OBS","CALC","RESID."

```

Fig. 13. Continued

```

1750 PRINT : FOR J = 1 TO N:R1 = INT (S(J) * 1000 + .5) / 1000:R2 = S(J) -
F(J,C1):R2 = INT (R2 * 1000 + .5) / 1000:R3 = INT (F(J,C1) * 10000
+ .5) / 10000
1760 PRINT R1,R2,R3: NEXT J
1770 PRINT D$;"PR#0"
1780 PRINT : PRINT
1790 GOTO 350
1800 PRINT "SYSTEM WILL NOT CONVERGE.": PRINT : PRINT "TRY OTHER
INITIAL ESTIMATES"
1810 PRINT "OR FIX ONE OR MORE PARAMETERS"
1820 POKE 216,0:ER = PEEK (222): PRINT "ERROR NUMBER IS ";ER:PK =
PEEK (219)* 256 + PEEK (218): PRINT "IN LINE NUMBER ";PK
1830 PRINT : PRINT : PRINT : INPUT "<RETURN> TO CONTINUE...";Y$
1840 GOTO 340
1850 END

```

Test Data ZKAN

Test data are provided in Ebert (Part I, this vol.) and a test run is provided here using data for female shrimp (*Penaeus semisulcatus*) in Kuwait (Mathews et al., Part I, this vol.). The data are presented in the accompanying table. The 1981 data were first run and they converged on solutions for three parameters, Z, K and L_{∞} ; n was fixed at -1 (i.e., von Bertalanffy curve). The estimate of L_{∞} was 55.0 mm, $K = 1.31 \text{ year}^{-1}$ and $Z = 3.72 \text{ year}^{-1}$. The 1979 data would not converge on simultaneous estimates of three parameters but by fixing $L_{\infty} = 55 \text{ mm}$, Z and K could be estimated as shown in the sample output: $Z = 2.473 \text{ year}^{-1}$ and $K = 0.725 \text{ year}^{-1}$.

Because estimates are found numerically, initial guesses are required and these initial values can be very important in achieving convergence. With the 1979 data, initial estimates of $Z = 2$ and $K = 1$ lead to a solution as will $Z = K = 0.1$. However, $Z = K = 2$ will not lead to convergence.

If parameter estimates show divergence over several iterations and Z or K become negative then try new initial estimates using the *average* of your initial guesses and the results of the *first* iteration. Frequently these new guesses will prevent overshooting the solutions with increasing amplitude.

If there is a clear directional trend in the changes in parameters and convergence is not achieved, then a unique solution probably is not possible and it will be necessary to fix one or more of the parameters. The first parameter to fix probably should be n, the shape parameter. Fix it at -1 so the growth model is von Bertalanffy. The second parameter to fix probably should be L_{∞} , which can be done using the largest observed size or some independent estimate, as obtained e.g., using the method of Wetherall et al. (Part I, this vol.).

The sample output provides printer output for data and solutions for the 1979 *Penaeus semisulcatus* data shown in the table below.

Data for female *Penaeus semisulcatus* in the Kuwait fishery (Mathews et al., Part I, this vol.), \bar{L} = mean length, n = number in sample, t = time since recruitment, which is the time since the minimum mean size for the year, recruitment length, L_R , is 22 mm for the 1979 data and 26 mm for the 1981 data, $L_{\infty} = 55$ mm was estimated for the 1981 data.

Month	1979			1981		
	\bar{L}	n	t	\bar{L}	n	t
Jan	42.533	30	—	20	3	—
Feb	35.458	107	—	—	—	—
Mar	34.672	122	—	—	—	—
Apr	26.191	105	—	—	—	—
May	23.639	643	0.000	—	—	—
Jun	25.266	924	0.083	—	—	—
Jul	27.983	586	0.167	27.835	593	0.000
Aug	28.590	1,383	0.250	27.524	542	0.083
Sep	29.467	821	0.333	31.049	553	0.167
Oct	30.513	452	0.417	34.714	392	0.250
Nov	32.327	202	0.500	40.357	370	0.333
Dec	36.275	240	0.583	36.077	26	0.417

Sample Output

Program to estimate Z, K, asymptotic size, and the shape parameter of the Richards function by nonlinear regression.

T. A. Ebert, January 1984

Return to continue....

Do you want to enter data from disk (D)
or keyboard (K)? (D/K) K

Number of data pairs = 8

Average size = 23.639
at T = 0

Average size = 25.266
at T = .08333

Average size = 27.983
at T = .16667

Average size = 28.590
at T = .25000

Average size = 29.467
at T = .33333

Average size = 30.513
at T = .41667

Average size = 32.327
at T = .50000

Average size = 36.275
at T = .58333

Do you want data stored on disk? (Y/N) Y

Filename for data = P.semisulcatus F79
on drive number = 1

Do you want to continue

with this same data set (Y/N)? Y

Do you want to estimate Z (Y/N)? Y

Do you wish to estimate K (Y/N)? Y

Do you want to estimate asymptotic size (Y/N)? N

Do you want to estimate N, the shape parameter? (Y/N) N

Initial values of parameters

Z = 2

K = 1

Maximum size = 55

Size at recruitment = 22

Shape parameter, N = -1

Do you want results sent to a
printer? (Y/N) Y

Turn on printer, then return
estimates of parameters

Z	K	A	N
---	---	---	---

2.289	.6846	55	-1
2.4646	.7252	55	-1
2.4727	.7247	55	-1
2.4728	.7247	55	-1
2.4728	.7247	55	-1

Final Values

Z = 2.4727882

K = .724702148

Residual SS = 7.14000006

Mean square error = 1.19000001

Standard deviations of parameters
with 6 DF

SD for Z = .543618308

SD for K = .0795544698

Values for plotting	
Time	Size
0	23.496
.08	25.342
.17	27.08
.25	28.717
.33	30.257
.42	31.707
.5	33.072
.58	34.357
.67	35.567
.75	36.706
.83	37.778
.92	38.787

Residuals

Obs.	Calc.	Resid.
23.639	23.496	.1429
25.266	25.342	-.0763
27.983	27.081	.9025
28.59	28.717	-.1267
29.467	30.257	-.7899
30.513	31.707	-1.1941
32.327	33.072	-.7451
36.275	34.357	1.9179

Listing E2

The program "E2" was written for an "Apple II+" microcomputer in "Applesoft" BASIC. Beyond the input formatting, no use is made of special features of this version of BASIC. On line 202, a logical expression ($T(i) \geq D$) is used in a calculation, which might not be supported by all versions of BASIC. This expression needs to be 1 if true, and 0 if false.

The program prompts for the input values; note that all time-related inputs are expressed in fractions of a year; correspondingly, output values of K and Z are in year^{-1} .

Weights may be given to the data points; this can reduce variability due to small sample size. If weighting is not desired, weights of 1 can be entered.

The search algorithm which is applied is rather straightforward. K is incremented in steps of 1 and the response function is calculated. When the response increases, i.e., when the desired minimum is surpassed, the search direction is reversed and the step size is reduced by an arbitrary amount. The procedure is repeated until the search reaches a minimum. This algorithm is simple and rather robust, though not very fast. It may come to a premature stop in the case of a wide, ill-defined minimum, where function response depends largely on rounding errors. This, however, is a difficult case for any minimization algorithm. During the iteration, the current value of K is printed on the video screen, along with the corresponding value of the response function, which may be the sum of squares of deviations or the difference between the current value of L_∞ and the input value—depending on the option chosen.

Test Data E2

Full test data and outputs are given in Damm (Part I, this vol.).

Fig. 14. Listing Program E2.

LISTING PROGRAM E2

```

1 REM PROGRAM E2
2 REM PROGRAMMER U.DAMM DECEMBER 1984
10 DIM L(30),T(30),W(30),T2(30)
20 HOME : PRINT "THIS PROGRAM ESTIMATES GROWTH AND MORTA"::PRINT
21 PRINT "LITY PARAMETERS FROM THE REGRESSION OF ":PRINT
22 PRINT "AVERAGE POPULATION LENGTH ON TIME":PRINT
23 PRINT "FOR THE CASE OF TWO RECRUITMENTS PER ":PRINT
24 PRINT "YEAR, AND GIVES THE RELATIVE AMOUNT OF ":PRINT
25 PRINT "EACH RECRUITMENT.": PRINT
26 PRINT "BERTALANFFY GROWTH IN LENGTH IS ASSUMED": PRINT
27 PRINT "AND LENGTH AT RECRUITMENT MUST BE KNOWN."
29 PRINT "-----"
30 PRINT : INPUT "NUMBER OF DATA POINTS ?";N
40 HOME : PRINT "ENTER TIME, LENGTH AND WEIGHTING FACTOR": PRINT
41 PRINT "TIME MUST BE GIVEN IN FRACTIONS OF A": PRINT
42 PRINT "YEAR, T=0 AT TIME OF FIRST RECRUITMENT.": PRINT
43 PRINT "WEIGHTING FACTOR MAY BE SAMPLE SIZE": PRINT
44 PRINT "OR MAY BE 1 IF WEIGHTING IS NOT WANTED.": PRINT : PRINT
50 FOR I = 1 TO N
51 INPUT "T = ";T(I)
52 VTAB PEEK (37) :HTAB 12: INPUT "L= ";L(I)
53 VTAB PEEK (37) :HTAB 24: INPUT "W= ";W(I)
55 NEXT
60 SW = 0: FOR I = 1 TO N : SW = SW + W(I) : NEXT
62 FOR I = 1 TO N: W(I) = W(I) /SW * N: NEXT
133 PRINT : PRINT "YOU MAY HAVE ": PRINT
134 PRINT TAB(9) ;"ESTIMATES OF ALL PARAMETERS (1)"
135 PRINT TAB(8) ;"ESTIMATES FOR A GIVEN L-INF. (2)"
136 PRINT TAB(4) "ESTIMATES FOR GIVEN L-INF. AND K(3)":PRINT
137 INPUT "WHAT DO YOU LIKE ?";Q
138 IF Q >3 OR Q < 1 THEN 137
139 HOME : ON Q GOTO 145, 141, 140
140 INPUT "K= ";K
141 INPUT "L-INF= ";L8
145 PRINT:PRINT:PRINT"TIME FROM FIRST TO SECOND RECRUITMENT = ":INPUT D
150 IF Q = 3 THEN 200
160 SS = 1000000!:K = 0 :LU = 1000000!:ST = .1
180 S0=0 :S1 =0:S2 =0:P0 =0: P1 =0: P2 =0: P3 = 0: P4 =0:P5 =0:P6 =0
190 IF Q = 1 THEN CO = SS
191 IF Q = 2 THEN CO = ABS (L8 - LU)
192 PRINT K,CO
193 K = K+ST
200 FOR I = 1 TO N
202 T2(I) = 1 -(T(I) > = D) + T(I) - D
210 E1 = EXP (-K * T(I))

```

Fig. 14. Continued

```

220 E2 = EXP (-K * T2(I))
225 X1 = E1 -E2:X2 = E2
230 S0 = S0 +L(I) *W(I)
240 S1 = S1 + X1 *W(I)
250 S2 = S2 + X2 *W(I)
260 P0 = P0 + L(I) * L(I) * W(I)
270 P1 = P1 + X1 * X1 * W(I)
280 P2 = P2 + X2 * X2 * W(I)
290 P3 = P3 + L(I) * X1 * W(I)
300 P4 = P4 + L(I) * X2 * W(I)
310 P5 = P5 + X1 * X2 * W(I)
390 NEXT I
400 Q0 = P0 - S0 * S0 / N
428 Q1 = P1 - S1 * S1 / N
430 Q2 = P2 - S2 * S2 / N
440 Q3 = P3 - S0 * S1 / N
450 Q4 = P4 - S0 * S2 / N
460 Q5 = P5 - S1 * S2 / N
500 CC = Q1 * Q2 - Q3 * Q5
510 B1 = (Q3 * Q2 - Q4 * Q5) /CC
520 B2 = (Q4 * Q1 - Q3 * Q5) / CC
530 DD = B1 * Q3 + B2 * Q4
531 IF ABS (LU - LB) < CO THEN 180
540 R2 = DD / Q0
600 LU = (S0 - B1 * S1 - B2 * S2) / N
610 SS = P0+ N * LU * LU + B1 * B1 * P1 + B2 * B2 * P2
620 S = -LU*S0-B1*P3-B2*P4+LU*B1*S1+LU*B2*S2+B1*B2*P5
621 SS = SS + S + S
629 IF Q = 3 THEN LU = LB : IF Q = 3 THEN 640
630 IF Q = 1 THEN 633
631 IF ABS(LU - LB) < CO THEN 180
632 GOTO 634
633 IF SS < CO THEN 180
634 ST = ST * -7 / 19
635 IF ABS (ST) < .0000001 THEN 640
636 GOTO 180
640 P = B1 / B2
650 PRINT : PRINT : PRINT
651 PRINT "K= ";K;" L-INF.= ";LU: PRINT
653 PRINT "FIRST RECRUITMENT : ";P* 100;" %": PRINT : IF Q <> 1 THEN 660
654 PRINT "SUM SQUARES = ";SS;" R^2= ";R2
660 PRINT : PRINT : PRINT "ENTER A LENGTH-AT-AGE DATA PAIR WHICH"
661 PRINT " WILL BE CONVERTED TO LR (LENGTH AT T =0)"
664 INPUT "T = ";T
665 VTAB PEEK (37):HTAB 15: INPUT "L= ";L
670 LR = LU - (LU - L) * EXP (K * T)
680 VTAB PEEK (37): HTAB 25: PRINT "LR= ";LR: PRINT
690 Z = LOG (( - B2 - EXP (K) * (LU - LR)) / ( - B2 - LU + LR)) - K
700 PRINT "Z = ";Z
800 PRINT : PRINT "DO YOU WANT TO ": PRINT
801 PRINT " SEE INPUT TIMES AND LENGTHS (1)"
802 PRINT "SEE EXPECTED LENGTHS AND RESIDUALS (2)"
803 PRINT " TRY OTHER PARAMETER VALUES (3)"
804 PRINT TAB(26); "TERMINATE (4)"
810 PRINT : INPUT Q
820 IF Q < 1 OR Q > 4 THEN 810

```

Fig. 14. Continued

```
830 ON Q GOTO 900,940, 133,999
900 HOME : PRINT "    T1      "    T2      "    OBS. LENGTH": PRINT
910 FOR I = 1 TO N
911 PRINT T(I); TAB(12);T2(I); TAB(24);L(I)
912 NEXT
915 GOTO 800
940 HOME : PRINT "CALC. LENGTH RESIDUAL WEIGHTED RESID .": PRINT
950 FOR I = 1 TO N
951 L = LU+B1*(EXP(-K*T(I))-EXP(-K*T2(I)))+B2*EXP(-K*T2(I))
952 PRINT L;TAB(12);L(I)-L;TAB(25);L(I)-L)*W(I)
953 NEXT I
955 GOTO 800
999 END
```

Program Number 7

Title	LGTFRQ
Author	P. Sparre
Author's Address	P. Sparre Danish Institute for Fisheries and Marine Research Charlottenlund Slot 2920 Charlottenlund Denmark

Reference Sparre, P.—A method for estimation of growth, mortality and gear selection/recruitment parameters from length-frequency samples weighted by catch per effort. Part I, this vol.

Description The program utilizes data from multiple length-frequency samples (in form of catch per unit of effort) to estimate the von Bertalanffy growth parameters, total mortality rates and gear selection/recruitment parameters for any given fish stock.

The length frequencies are separated into normally distributed components and the mean length values of each normally distributed component used to estimate the growth parameters. A measure of goodness of fit is provided using a chi-squared criterion.

The method allows for continuous recruitment to the stocks although a pronounced seasonality in recruitment is required in order to separate components of the length-frequency samples.

Listing Notes

- a. Language FORTRAN 77
- b. Implementation VAX 11/750
- c. Hardware Requirements Printer
- d. RAM Requirements Large in VAX version but can be reduced to microcomputer capability by replacement of the NAG minimization subroutine EO4FCF.

Listing Because of the length of this program and since many subroutines are available as NAG library functions, a complete listing of the program is not given here.

A comprehensive user's manual is available upon request from the author.

In order to present the salient features of this program, a detailed flowchart of the operations involved is included here (Figs. 15, 16, 17 and 20).

Flowchart

The central part of the LGTFRQ program is formed by the NAG minimization subroutine EO4FCF and the subroutine LSQFUN which defines the object function to be minimized by EO4FCF.

Subroutine LSQFUN is extensively described in Sparre (Part I, this vol.) and will not be further discussed here.

Some features of EO4FCF are given a superficial discussion here. For a comprehensive description the reader is referred to the NAG-library manual. However, a detailed knowledge on the method behind EO4FCF is not required for running LGTFRQ.

The present version of the LGTFRQ program constitutes about 30 pages of FORTRAN code. The major part of this deals with input and output, and it is not considered worthwhile to reproduce the complete program here.

The essential part of LGTFRQ is the subroutine LSQFUN.

Fig. 15 shows a summary flowchart for the LGTFRQ program.

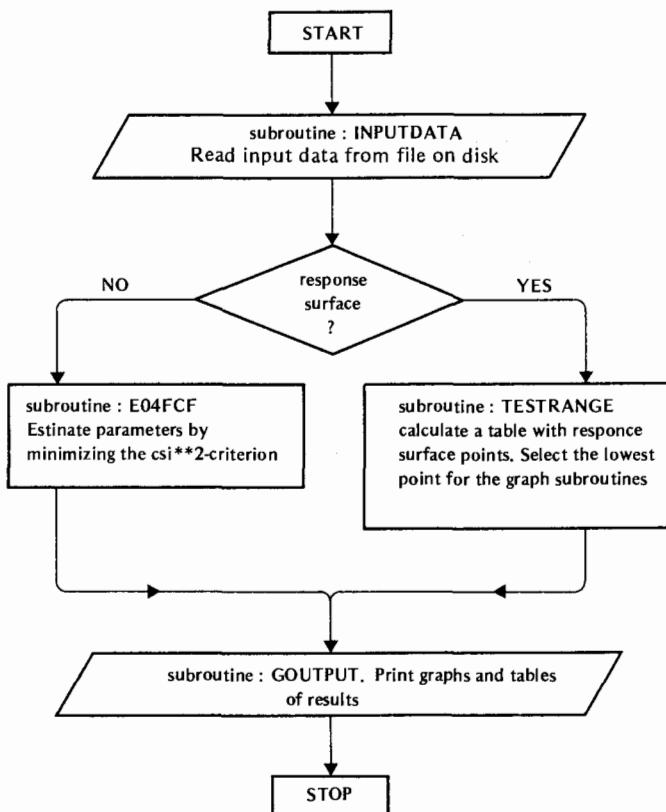


Fig. 15. Flowchart for the main program of LGTFRQ.

Input

Fig. 16 shows the flowchart for the main input subroutine, INPUTDATA. Except for the subroutine by which the initial guesses of parameters are read, PARAMETERS, these subroutines are quite simple ones. The subroutine PARAMETERS is further explained in the flowchart of Fig. 17. The "status" of a parameter can be one of the three alternatives:

1. Basic parameter;
2. Fixed parameter; and
3. Parameter proportional to a basic parameter.

A "basic parameter" is a parameter the value of which is estimated by minimizing the chi-squared criterion. The user must supply an initial guess on a basic parameter. The final value of a basic parameter is calculated by the NAG-subroutine EO4FCF.

A "fixed" parameter will not have its value changed by EO4FCF, e.g., LGTFRQ allows for an estimation of some growth parameters while other growth parameters are kept constant.

A parameter proportional to a basic parameter is a parameter the value of which is estimated by EO4FCF, but under the constraint that it is proportional to a basic parameter. For example, we may estimate the cohort strengths under the constraint that all spring cohort strengths are equal and that all autumn cohort strengths are equal.

The concept of a proportional parameter may also be applied to the assumptions relative to total mortality. For some of the growth and selection parameters, the proportionality concept may be useful as well.

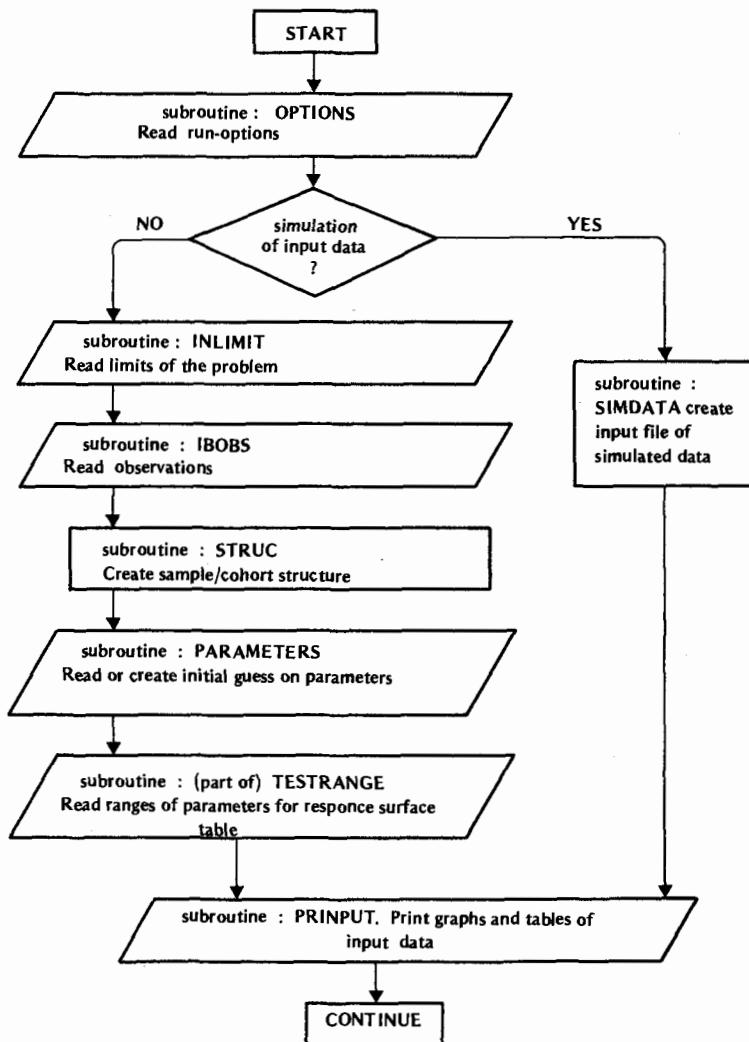


Fig. 16. Flowchart for the main input subroutine, INPUTDATA.

Calculations

The following is a superficial description of some basic features of the model behind the minimization routine EO4FCF and some guidelines for running the program.

EO4FCF minimizes a sum of square. In the present application this sum is the sum of squares of residuals (see Fig. 18 and Sparre, Part I, this vol.). The sum of squares is also called the object function.

The object function is a function of the parameters specified as basic parameters (cf. section on input).

The estimation problem may be considered as that of finding the lowest point on the response surface in the $(n + 1)$ dimensional parameter space (Fig. 19), where n is the number of basic parameters:

$$\{\text{CHI}^{**2}(\text{PARAM}), \text{PARAM}\}$$

Where PARAM is the n -dimensional vector of basic parameters (PARAM is a subvector of THETA).

EO4FCF works in an iterative manner which may be considered a "walk" in the response surface in search of the lowest point.

Fig. 20 shows the flowchart for the search on the response surface.

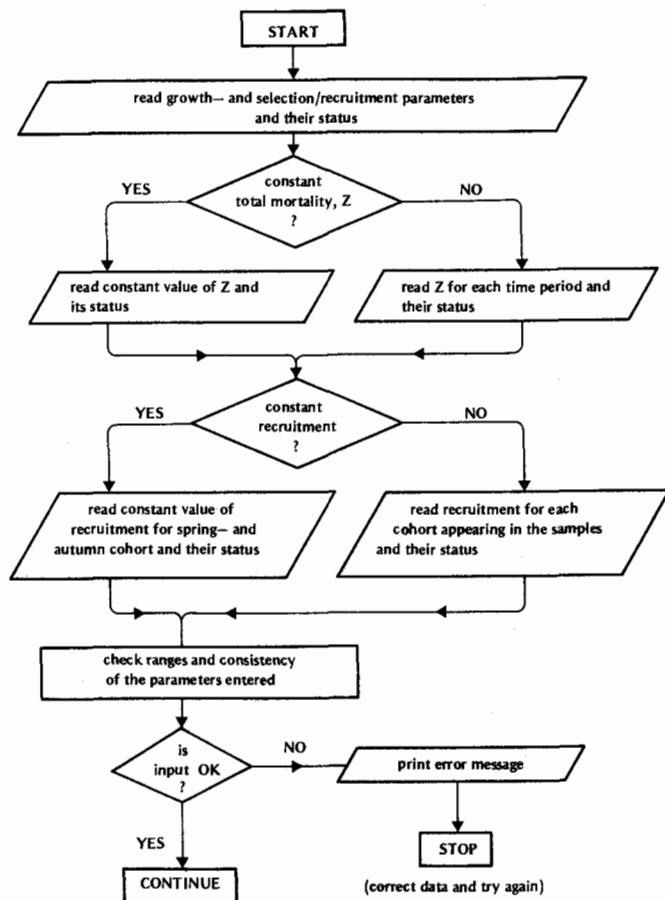


Fig. 17. Flowchart for input of parameters, subroutine PARAMETERS.

Output

There are two kinds of output:

1. Tables and graphs of input data. (Not processed or slightly processed observations and initial guesses of parameter values). This is performed by the subroutine PRINPUT.
2. Tables and graphs of output data. (Estimates of parameters and the estimated dynamics of the fish stock in question). This is performed by the routines PRRESULTS, GOUTPUT and TESTRANGE.

Tables and Graphs of Input Data

The subroutine PRINPUT can produce the following output:

- a. Table of raw data (length frequencies).
- b. Graphs showing the length frequencies as bars.
- c. Sample/cohort structure.
- d. Initial guess on parameters. It also shows the status of the parameters and the corresponding THETA-index.

Tables and Graphs of Results

Subroutine PRRESULTS prints a table with the estimated basic parameters, in case the option for that was chosen.

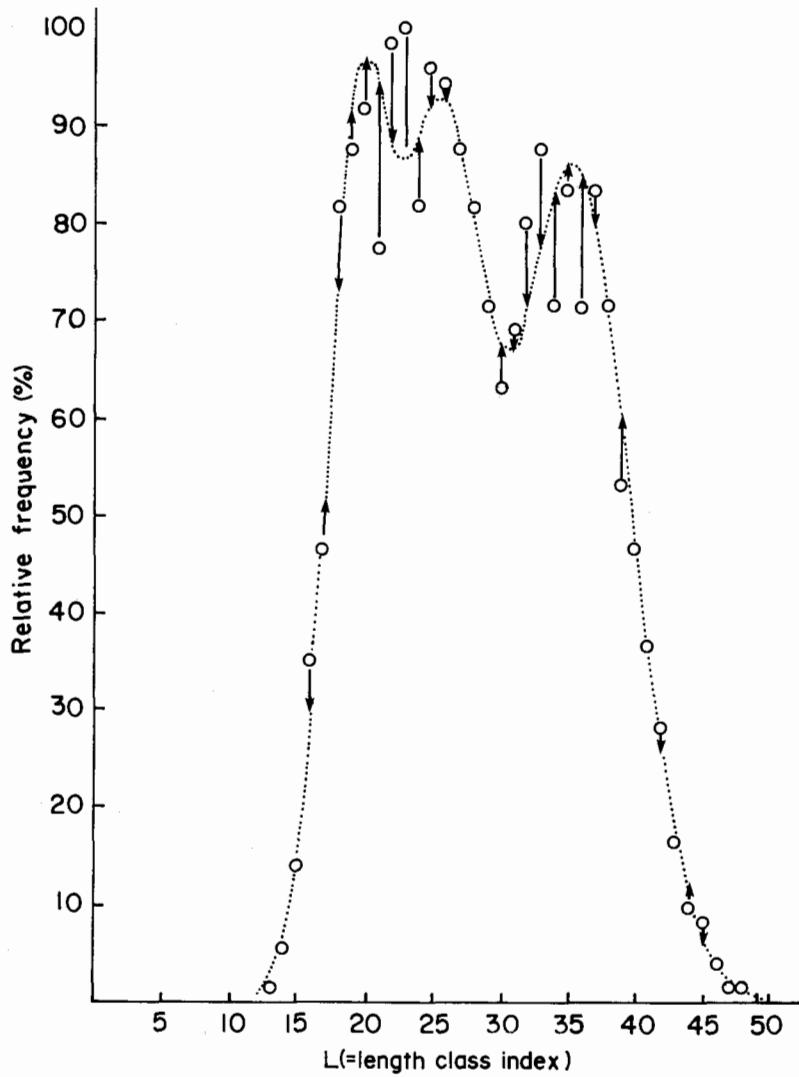


Fig. 18. Showing residuals, i.e., difference between observed and predicted frequencies; the routine EO4FCF is used to minimize the sum of the squares of these residuals.

Subroutine TESTRANGE produces a table with response surface points, in case the option for that was chosen. The subsequent uses of the graph subroutines will be based on the lowest point on the response surface found in the table.

Subroutine GOUTPUT produces three types of graphs together with tables. They are all printed on a lineprinter. The graphs offered by GOUTPUT are:

- a. Growth curves.
- b. Resolution of frequencies into normally distributed components.
- c. Graphs showing the residuals together with the original observations.

The program may provide a large number of messages on input errors or warnings on suspicious input values. These are self-explanatory and are not shown here.

Test Data

Test data for the program are provided in Sparre (Part I, this vol.) and also in the user manual available from the author.

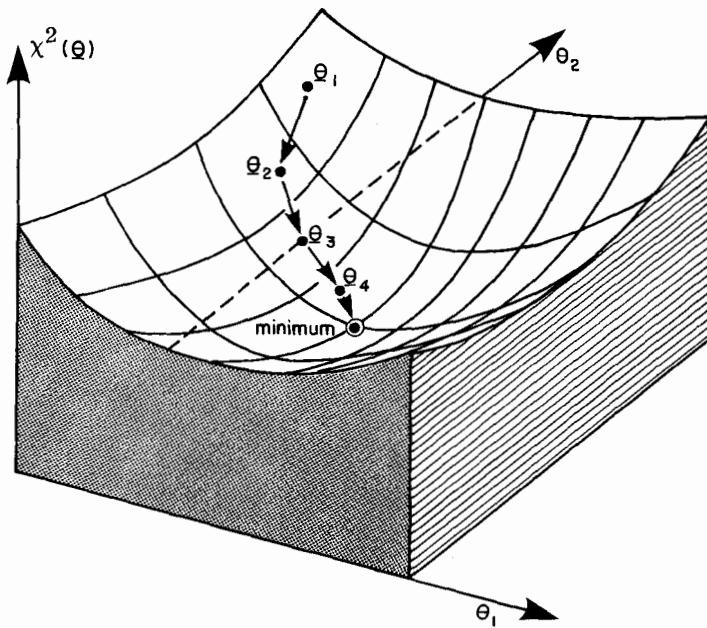


Fig. 19. Schematic representation of the minimization of chi-squared using the NAG subroutine EO4FCF (see text).

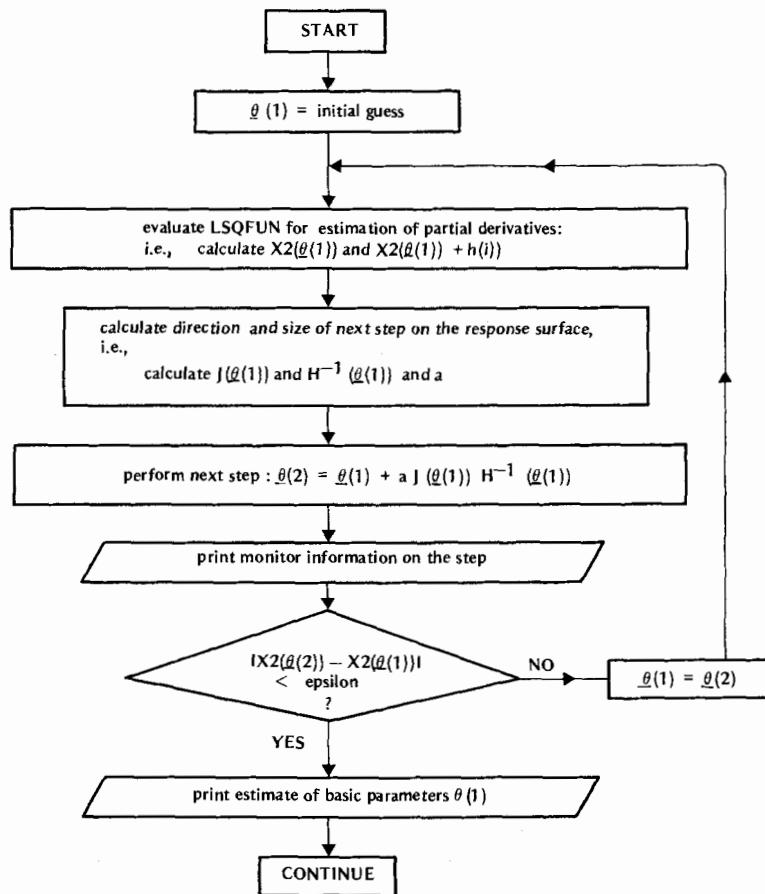


Fig. 20. Flowchart for the estimation process (see also Fig. 19), "the walk on response surface". The stop criterion "epsilon" may be specified by the user. Actually, EO4FCF uses several different stop-criteria.

Program Number 8

Title	SRLCA
Author	J.G. Shepherd
Author's Address	Ministry of Agriculture, Fisheries and Food Fisheries Laboratory Lowestoft Suffolk NR33 0HT United Kingdom

Reference Shepherd, J.G.—A weakly parametric method for estimating growth parameters from length composition data (Part I, this vol.).

Description The program fits growth curves to the modes in a time series of length composition data to estimate the parameters of the von Bertalanffy growth equation. A goodness of fit criterion, which is similar to a correlation coefficient between the data and a test function constructed from the growth curve is used to select the growth curve which most appropriately describes the data. The criterion is designed to take account of data in proportion to its quantity and likely reliability without pre-processing, and not to be unduly sensitive to possible modes created by sampling noise.

Listing Notes

- a. Language FORTRAN 77
- b. Implementation Apricot
- c. Hardware Requirements Printer
- d. RAM Requirements About 48 K

Listing

Fig. 21 gives a complete listing of the SRLCA program.

Fig. 21. Listing Program SRLCA.

```

PROGRAM SRLCA
C
C LENGTH COMPOSITION ANALYSIS PROGRAM
C WRITTEN BY J G SHEPHERD, SEPTEMBER 1984
C (PERFORMS SIMILAR ANALYSIS TO PAULY'S ELEFAN I)
C
C WARNING : PROGRAM UNDER DEVELOPMENT
C NOT GUARANTEED BUG-FREE
C USE AT YOUR OWN RISK

```

Fig. 21. Continued

C

```

REAL TS(12),LENG(50),NUM(12,50),LINF(11),TZERO(2),LMIN,KK
REAL K(11),SCORE(11,11),TMIN(51),NUMA(21,12),NUMT(21),DISCR(12)
REAL INC,LLINF,LLZERO,CUMS(50),LZERO,TSCORE(2),TZ(11,11)
INTEGER AGE,AGE1
CHARACTER FILENAME*10,EXPLORE*2,ANSWER*1,TITLE*80
COMMON NL,LENG,PI,CUMS,NDIST,TS,NUM,RK,TMIN
PI=3.14159
LP=6
OPEN(6,FILE='PRN')
WRITE(*,'('' ENTER NAME OF DATA FILE '')')
READ(*, '(A)') FILENAME
OPEN(200,FILE=FILENAME)
READ(200, '(A)') TITLE
READ(200,*) LMIN,NL,INC
READ(200,*) NDIST,(TS(I),I=1,NDIST)
DO 810 L=1,NL
    READ(200,*) LENG(L),(NUM(I,L),I=1,NDIST)
810 CONTINUE
LENG(NL+1)=LENG(NL)+INC
CLOSE(200)
C -----
1000 WRITE(*,'('' DATA IS FOR '',A80) ')TITLE
WRITE(*, '(A)') EXPlore or EValuate?
READ(*, '(A)') EXPLORE
IF (EXPLORE.EQ.'EX') THEN
    WRITE(*,'(''NO. OF VALUES FOR L-INFINITY? '')')
    READ(*,*) NLINF
ELSE
    NLINF=1
ENDIF
WRITE(*,'(''VALUES FOR L-INFINITY ?'')')
READ(*,*)(LINF(I),I=1,NLINF)
C -----
NLZERO=2
DO 100 I=1,NLZERO
    TZERO(I)=0.25*(I-1)
100 CONTINUE
C -----
IF (EXPLORE.EQ. 'EX') THEN
    WRITE(*,'(''NO. OF VALUES FOR K ? '')')
    READ(*,*) NK
ELSE
    NK=1
ENDIF
WRITE(*,'(''VALUES FOR K ? '')')
READ(*,*)(K(I),I=1,NK)
IF(EXPLORE.EQ. 'EV') THEN
    WRITE(*, '(A)') 'VALUE FOR T-ZERO ?'
    READ(*,*) TZ(1,1)
ENDIF

```

Continued

Fig. 21. Continued

```

C -----
      IF(EXPLORE.EQ.'EX') THEN
      DO 200 ILINF=1,NLINF
         LLINF=LINF(ILINF)
         DO 300 IK=1,NK
            KK=K(IK)
            DO 500 ILZERO=1,NLZERO
               CALL SUMPRD(KK,LLINF,TZERO(ILZERO),TSCORE(ILZERO))
500      CONTINUE
         A=TSCORE(1)
         B=TSCORE(2)
         SCORE(ILINF,IK)=SQRT(A*A+B*B)
         PHI=ATAN2(B,A)
         TTZ=PHI/(2*PI)
         IF(TTZ.LT.0) TTZ=TTZ+1
         TZ(ILINF,IK)=TTZ
300      CONTINUE
200      CONTINUE
C -----
      WRITE(LP,'(''LENGTH COMPOSITION ANALYSIS BY SRLCA FOR '',ABO)
      &TITLE
                  WRITE(LP,'(20X,''TABLE OF SCORE FUNCTION''))
                  WRITE(LP,8000)
                  WRITE(LP,8010),(K(IK),IK=1,NK)
                  WRITE(LP,'(/)')
                  DO 801 ILINF=1,NLINF
                     WRITE(LP,8020)LINF(ILINF),(SCORE(ILINF,IK),IK=1,NK)
801      CONTINUE
                  WRITE(LP,'(1H0/,20X,''TABLE OF T-ZERO'')
                  WRITE(LP,8000)
                  WRITE(LP,8010),(K(IK),IK=1,NK)
                  WRITE(LP,'(/)')
                  DO 800 ILINF=1,NLINF
                     WRITE(LP,8021)LINF(ILINF),(TZ(ILINF,IK),IK=1,NK)
800      CONTINUE
      ENDIF
      IF (EXPLORE.EQ.'EV')THEN
         CALL SUMPRD(K(1),LINF(1),TZ(1,1),SUM)
         WRITE(LP,'(''1 LENGTH COMPOSITION ANALYSIS BY SRLCA FOR '',ABO)
      &TITLE
         WRITE(LP,6000) K(1),LINF(1),TZ(1,1)
         LLINF=LINF(1)
         LEASTAGE=INT(TMIN(1)-0.5)
         INCAGE=1-LEASTAGE
         DO 900 L=1,NL
            IF (LENG(L+1).LT.LLINF) THEN
               TMAX=TMIN(L+1)
               DO 901 ID=1,NDIST
                  AGEMIN=TMIN(L)-TS(ID)+0.5
                  AGEMAX=TMAX-TS(ID)+0.5
                  MINAGE=INT(AGEMIN)
                  MAXAGE=INT(AGEMAX)
                  NAGE=MAXAGE-MINAGE+1

```

Continued

Fig. 21. Continued

```

DO 902 AGE=MINAGE,MAXAGE
  FRACT=1
  IF (NAGE.GT.1) FRACT=1.0/(AGEMAX-AGEMIN)
  IF (NAGE.GT.1.AND.AGE.EQ.MINAGE)
    FRACT=(MINAGE+1-AGEMIN)*FRACT
  IF (NAGE.GT.1.AND.AGE.EQ.MAXAGE)
    FRACT=(AGEMAX-MAXAGE)*FRACT
  AGE1=AGE+INCAGE
  IF (AGE1.GT.20) AGE1=20
  NUMA(AGE1, ID)=NUMA(AGE1, ID)+FRACT*NUM(ID,L)
  NUMT(AGE1)=NUMT(AGE1)+FRACT*NUM(ID,L)
902  CONTINUE
901  CONTINUE
  WRITE(LP,6001) LENG(L),TMIN(L),CUMS(L)
  &           ,(INT(0.5*(TMIN(L)+TMIN(L+1))-TS(ID)+0.5), ID=1,NDIST)
  ENDIF
900  CONTINUE
  WRITE(LP,6002)
  DO 910 AGE=1,20
    WRITE(LP,6003) AGE, (NUMA(AGE, ID), ID=1,NDIST)
    DO 920 ID = 1,NDIST
      NUMA(AGE, ID) = 0.0
920  CONTINUE
910  CONTINUE
  ENDIF
  WRITE(*,'('' DO MORE ? '')')
  READ(*,'(A)') ANSWER
  IF (ANSWER.EQ. 'Y') THEN
    GOTO 1000
  ENDIF
  STOP
999  WRITE(*,'(''ERROR ON OPENING FILE IS '',I3)'')IOS
6000  FORMAT(///,'EVALUATION OF SOLUTION FOR '
  &,/, ' K = ',F10.3
  &,/, 'L-INFINITY = ',F10.1
  &,/, 'T-ZERO = ',F6.2
  &,/, ' LENGTH GRP      AGE      CUM. SCORE'
  &, '      AGE-GROUP PLUS OR MINUS N YEARS')
6001  FORMAT(F10.1,2F8.2,12I8)
6002  FORMAT(///,' AGE COMPOSITION',///,' AGE NUMBERS AT AGE')
6003  FORMAT(I4,12F6.1)
9020  FORMAT(1H0/,F5.1)
8000  FORMAT(1H0/,20X,' VALUE OF K ')
8010  FORMAT(1H0,4X,'L-INF ',11F6.2)
8020  FORMAT(1H ,3X,F6.2),11F6.1)
8021  FORMAT(1H ,3X,F6.2,11F6.2)
  END
C -----
  SUBROUTINE SUMPRD(KK,LLINF,TZERO,SUM)
  REAL LENG(50),TS(12),NUM(12,50),TMIN(51),LLINF,LLZERO,CUMS(50)
  &,DISCR(12),KK
  COMMON NL,LENG,PI,CUMS,NDIST,TS,NUM,RK,TMIN
  RK=1/KK
  LLZERO=LLINF*(1-EXP(KK*TZERO))

```

Continued

Fig. 21. Continued

```

SUM=0.0
DO 400 L=1,NL+1
    IF (LENG(L).LT.LLINF) THEN
        TMIN(L)=RK*ALOG((LLINF-LLZERO)/(LLINF-LENG(L)))
    ENDIF
400      CONTINUE
C -----
DO 600 L=1,NL
    IF (LENG(L+1).LT.LLINF) THEN
        TMAX=TMIN(L+1)
        DELT=TMAX-TMIN(L)
        TBAR=.5*(TMAX+TMIN(L))
        DIFFN=SIN(PI*DELT)/(PI*DELT)
        CUMS(L)=0.0
        DO 700 ID=1,NDIST
            DISCR(ID)=DIFFN*COS(2*PI*(TBAR-TS(ID)))
            DELTAS=SQRT(NUM(ID,L))*DISCR(ID)
            SUM=SUM+DELTAS
            CUMS(L)=CUMS(L)+DELTAS
700      CONTINUE
        IF (L.GE.2) CUMS(L)=CUMS(L)+CUMS(L-1)
    ENDIF
600      CONTINUE
RETURN
END

```

Test of Shepherd's Method (SRLCA)

The method proposed by Shepherd (Part I, this vol.) had not been extensively tested before the meeting, and was run on various data sets during the meeting in order to acquire a better understanding of the virtues and vices of the method.

Tests on simulated data are preferable to real data for testing purposes, since the correct answer is known unambiguously. Unfortunately it was not possible to get any of the simulation programs working in time for new data to be generated, and simulated data were in short supply. However, tests were possible on the crudely simulated data of both R. Jones and P. Sparre (Part I, this vol.), as were tests on data where age information is also available (North Sea Haddock and Kuwaiti Newaiby), which constitute a good second best. Various tests of sensitivity were also conducted.

TEST ON SIMULATED DATA

Jones' Simulation Based on North Sea Haddock

Simulated population length compositions for a haddock-like stock had been presented in Table 2 of the paper by Jones (Part I, this vol.). These simulate a steady state under several levels of

total mortality, assuming constant recruitment and Z , but some variation of length-at-age. The data are therefore quite well-behaved and should not present any problem for any useful method (real data will almost certainly be less well-behaved than this).

SRLCA was run on the single length compositions for $Z = 0.5, 1.0$ and 2.0 , to examine the sensitivity for Z under steady-state conditions. In each case 11 values of L_∞ were specified for the exploratory phase, with K running from 0.1 to 0.6 and L_∞ from 40 to 90 cm.

In each case a well-defined ridge of high values of the score function was obtained, running from high K , low L_∞ to low K , high L_∞ . This typical behavior indicating the strong correlation of the parameters is discussed repeatedly in various parts of this volume.

In some cases there were weak local maxima on the ridge, but as discussed elsewhere, these are of little significance. The results in the text table below therefore indicate the best values of K in each case for a variety of values of L_∞ , together with the true values.

Value of K for various selected L_∞ values.

L_∞	$Z = 0.5$	$Z = 1$	$Z = 2$	Combined	Approx. $K \cdot L_\infty$
				data	
50	0.55	0.60	0.60	0.60	30
60	0.45	0.45	0.50	0.45	27
70	0.35	0.35	0.40	0.35	25
80	0.30	0.30	0.30	0.30	24

Clearly the level of Z has little effect on the results, and since the true value of K is 0.42 for $L_\infty = 58$ cm, the results in the second row are quite close to the correct results (true confidence intervals cannot be given but are almost certainly wider than the ± 0.05 to which K can be determined from the tables used here), although it may well be that K is being overestimated by about 10% even if L_∞ is specified correctly.

As a further test, the three length compositions were combined as though they were three individual sequential estimates in a single data set. This very crudely simulates the effect of changing Z slowly and taking samples at appropriately long intervals of time. The results for these combined data are also given in the table, and are for practical purposes identical to those obtained on the individual compositions.

The available version of SRLCA only provides imputed age compositions by the simple cohort slicing method, but these were for completeness compared with the known true age compositions: the results are given in the text table below.

Comparison of age compositions.

"Age"	$Z = 0.5$		$Z = 1$		$Z = 2$	
	SRLCA	True	SRLCA	True	SRLCA	True
	$L_\infty = 55,$ $K = 0.5$		$L_\infty = 60,$ $K = 0.45$		$L_\infty = 75,$ $K = 0.35$	
3	6,374	6,065	3,860	3,679	1,396	1,353
4	3,350	3,679	1,232	1,353	121	183
5	2,033	2,231	410	498	4	25
6	979	1,353	128	183	0	3
7	543	821	42	67	0	+
8	264	498	16	25	0	+
9	239	302	7	9	0	+

In each case, values of K and L_{∞} were taken from the "ridge" of good solutions giving a high score. For the first few age groups the agreement is rather good, generally to about 10% or so, becoming worse in relative terms as age increases, exactly as would be expected, since numbers at age for large fish must be highly dependent on the value of L_{∞} used. Very similar results were obtained from the combined data set.

These results indicate that SRLCA is capable of determining growth parameters to within about 10% on the rather well-behaved data set, and that the simple decomposition into age groups provided has a similar accuracy for the first two or three age groups only. There seems to be some evidence for a consistent tendency to slightly overestimate KL_{∞} . The results are almost completely unaffected by changing Z either between or within data sets.

P. Sparre's Simulation of a Tropical Fish

Four length compositions simulated for a tropical fish were available in the paper of Sparre (Part I, this vol.). The simulation included sampling variability, variable year class strength, seasonal (bimodal) recruitment and seasonally modulated growth. These data, for a fast-growing fish subject to high mortality, thus constitute a fairly severe test for SRLCA, which does not allow for multiple pulses of recruitment within a year, nor seasonality of growth rate.

SRLCA was run for values of L_{∞} between 30 and 80 ($L_{\max} \sim 55$) and for K between 0.05 and 1.0. A rather broad but fairly distinct ridge of high values of the score function was obtained, without clear local maxima. The crest of the ridge corresponded to the values of K and L_{∞} given below.

Best values of K for selected L_{∞} .

L_{∞}	K	KL_{∞}
50	0.95	48
55	0.75 or 0.80	43
60	0.65	39
65	0.55	36
True value	50	0.8 (mean)
		40

The correct values are $L_{\infty} = 50$ cm and $K = 0.80$. The value of K at the correct L_{∞} is therefore overestimated, but for other (higher) values of L_{∞} the value of KL_{∞} is correct to about 10%, although there is a systematic trend of KL_{∞} with L_{∞} (i.e., the inverse correlation of the parameters is not perfect). The interpretation in terms of ages corresponding to the solution $L_{\infty} = 55$, $K = 0.8$ shows that the modes are correctly identified. These results are satisfactory, surprisingly so considering the strong seasonal modulation of growth and the strong double pulse of recruitment. The reasonably accurate determination of mean K is probably due to the even spread of "samples" over the year. The double recruitment pulse might have been expected to generate an incorrect solution corresponding to half the correct value of K, but there is no sign of good values of the score function corresponding to such a solution. The reason for this unexpectedly desirable behavior is not known, and it should probably not be taken for granted.

The method is at present incapable of detecting or estimating seasonal modulation of growth, but may possibly be modified to allow for this feature. It is not clear whether or how it could be modified to allow for and estimate multiple pulses of recruitment, nor whether such a modification is necessary: if the phenomenon is strong enough to create distinct modes it should be easily visible on the response surfaces.

TEST ON DATA FOR WHICH AGE IS AVAILABLE

Newaiby (*Otolithes argenteus*) from Kuwait

The data for this stock were available and the data from 12 monthly samples in 1982 were analyzed using SRLCA, for L_{∞} between 40 and 90 cm, and K between 0.1 and 0.6. The topography of the response surface in this case is not very clear: there are (as usual) high values of the score function for high K and L_{∞} , but there is a weak outlying "spur" (a ridge joined to the main plateau) running as follows:

Locus of solutions along "spur".

L_{∞}	K	KL_{∞}
40	n/a	—
45	0.30?	13
50	0.25 to 0.30	14
55	0.25	14
60	0.20	12
65	n/a	—

Note: n/a = not apparent.

Since it is known that almost all methods of analysis, modal progression or matching are liable to give spurious solutions at multiples of the true value of K, and there are indeed further ridges on the plateau in positions which correspond approximately to such harmonics, an informed analyst would prefer solutions corresponding to small KL_{∞} among possible alternatives.

The solution at $L_{\infty} = 55$ cm, K = 0.25 was therefore selected, evaluated and compared with the results from age determination. In general, these results were similar to those obtained using age data (Morgan 1984) although the L_{∞} value was a little lower than the 60.0 cm calculated from age data using the same data set.

Modified (truncated) versions of this data set from which all data corresponding to the largest and smallest components of the samples had been removed, were also used to determine the sensitivity of the method to such surgery (see Shepherd et al., Part II, this vol.).

North Sea Groundfish Survey Haddock Data

These annual survey data were analyzed for L_{∞} between 50 and 100 cm and K between 0.1 and 0.6. The response surface shows only a weak ridge running as follows:

L_{∞}	K	KL_{∞}
50	0.40	20
55	0.35	19
60	0.35	21
65	0.30	20
70	0.25	18
75	0.25	19

The solution at $L_{\infty} = 60$, $K = 0.35$ was selected, evaluated and compared with the results from age determination. These results were almost identical to those obtained from age-at-length data.

OTHER DATA SETS

MacDonald's Pike Data

This traditional data set has been analyzed by MacDonald and Pitcher (1979), and by Schnute and Fournier (1980), and age determinations by scale reading are also available.

The analysis by SRLCA for L_{∞} between 50 and 100 cm and K between 0.1 and 0.6 indicates a clear ridge as follows:

L_{∞}	K	KL_{∞}
50	0.5 to 0.6	28
60	0.35	21
70	0.30	21
80	0.25	20
90	0.20	18

Analysis of mean length-at-age data indicates that L_{∞} is probably well beyond L_{max} (74 cm), but previous estimates of K and L_{∞} have not been published. The solution at $L_{\infty} = 90$, $K = 0.20$ leads to modal lengths very close (within a few cm) of the mean lengths at age, and to the modal lengths determined by previous authors.

Using SRLCA the major part of the score is determined by the smaller fish, and the method would be expected to be rather insensitive to truncation of the length composition. To test this, the same data were run after truncation at 60, 50 and 40 cm. In the first two cases a clear ridge is still visible in the response surface, but begins to disappear when the data are truncated at 40 cm (this corresponds to throwing away about half of the data, presumably a fairly severe test).

The results for the crest of the ridge are:

L_{∞}	K		
	< 60 cm	< 50 cm	< 40 cm
60	0.35	0.40	0.30
70	0.30	0.25	0.35
80	0.25	0.30	0.30
90	0.20	0.25	0.25
100	0.20	0.20	0.20

For the larger values of L_{∞} at least the results are fairly stable (bearing in mind that the mapping resolution is only 0.05 in K), and the method is continuing to indicate essentially the same solution (with $KL_{\infty} \approx 20$). Therefore, this test confirms that the method is not unduly sensitive to truncation of the data for larger fish.

It has been observed (P. Sparre, pers. comm.) that ELEFAN I can sometimes be quite sensitive to the size of length interval used. SRLCA, on the other hand, would not be expected to be so sensitive (one would expect loss of precision in location of good solutions rather than systematic bias).

Thus, the data were aggregated into 4-cm groups and re-analyzed.

The results still indicate a fairly clear ridge in the score function, running as follows:

L_∞	K
60	0.45
70	0.35
80	0.25
90	0.20
100	0.25

but the ridge is broader and the crest less well-defined than in the original data grouped into 2-cm intervals.

Penaeid Shrimp Data (Kuwait)

These data are for a very high growth, high mortality stock (normally only one cohort is apparent in the data) and are believed to exhibit strong seasonal modulation of growth rate (C.P. Mathews, pers. comm.). True growth parameters are not known, but results of previous investigations (including tagging) indicate that K is about 1.0 and L_∞ is about 60 mm, although the range of acceptable estimates is wide.

Analysis using SRLCA indicates a broad, weak ridge corresponding roughly to $KL_\infty = 130 \pm 60$ for L_∞ between 60 and 140 mm, with no clear indication of preferred solution. The conventional results lie right on the boundary of this range, and would almost certainly not be selected from the SRLCA analysis.

It is not clear whether this discrepancy indicates a failure of the method, since the conventional analysis is uncertain. However, the indication of high K and/or L_∞ from SRLCA could easily be due to its present failure to allow for seasonal modulation of growth. Since these shrimp are only caught during the growth season any method which ignores seasonal growth modulation is likely to overestimate KL_∞ . The results should therefore probably be taken as cautionary, in that SRLCA may not be suitable in its present form for stock with strong seasonal growth modulation and high mortality.

Conclusions

The results of these tests are very encouraging in that the method seems to have performed well (sometimes surprisingly well) on almost all the data sets presented to it. The tests on simulated data are the most informative, and further such tests should be carried out. Results to date suggest that SRLCA may have a tendency to overestimate K and/or L_∞ . Since the method "fails" when the true solution (typically a ridge of high score values corresponding roughly to constant values of the product KL_∞) merges into the background (high values of score for high K and high L_∞), the tendency to overestimate them may be due to the detailed definition of the score function.

In general the method seems to be quite robust to

- seasonality of recruitment, including (perhaps surprisingly) multiple pulses of recruitment,
- grouping into length intervals,
- truncation of data for large fish.

The method at present does not allow for seasonal modulation of growth rates, and will only correctly estimate the mean growth rate if individuals live for several seasonal cycles, and the data are well spread over the cycle.

The ability to map and scan the "response surface" is very valuable for qualitative interpretation of the results. The lack of proper confidence levels for the comparison of score values is a deficiency (but not so much as no indication at all of precision of the results). Most of the results above have been based on deviations of the score function much less than the 2:1 suggested by Shepherd (Part I, this vol.) as possibly corresponding to a high degree of confidence, and this criterion may be unrealistically stringent.

Literature Cited

- McDonald, P.D.M. and 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.
- Schnute, J. and D. Fournier. 1980. A new approach to length frequency analysis: growth structure. *Can. J. Fish. Aquat. Sci.* 37: 1337-1351.
- Morgan, G.R. 1984. Application of length-based stock assessment to Kuwait's fish stocks. *ICLARM Newsl.* 6(4): 3-4.

Program Number 9

Title	SAPPH
Author	J.G. Pope
Author's Address	J.G. Pope Ministry of Agriculture, Fisheries and Food Directorate of Fisheries Research Fisheries Laboratory Lowestoft Suffolk NR33 OHT United Kingdom

Reference Pope, J.G. and Y. Jiming.—Phalanx analysis: an extension of Jones' length cohort analysis to multispecies cohort analysis. Part I, this vol.

Description The program utilizes catch-at-length data for several species to construct a multispecies extension to Virtual Population Analysis (VPA) which takes some account of predation as a source of natural mortality. The number of fish eaten by other fish included in the analyses between lengths l_1 , and l_2 is calculated using estimates of predator/prey size ratio and the average population numbers of predators and prey species.

The results not only allow estimates of F and Z by length group by species but also allow examination of the effects of management intervention such as mesh size and fishing effort changes. In addition size specific predator/prey changes brought about by fishing can also be examined.

Listing Notes

- a. Language FORTRAN 77
- b. Implementation Apricot
- c. Hardware Requirements Printer
CRT display
Single disk drive
- d. RAM Requirements About 62 K

Listing A flowchart of the program is shown in Pope (Part I, this vol.) together with a summary of test data. Since the program is still under development, a full listing is not given here.

WARNINGS: No solutions are possible with *OPTION 2* when $t_c^A = t_c^B$.

All mortality coefficients should be input as annual values. The units for the ages at first capture should be in years. The program estimation of $C_{t_d}^A \rightarrow t_m$ and $C_{t_d}^B \rightarrow t_m$ is based on $N_o^A = N_o^B = 10,000$.

b. Sex Ratio 2

DESCRIPTION: The program estimates the fishing and natural mortality coefficients (F and M) from knowledge of the catch numbers by length (or age) class, the total mortality coefficient (Z), the age at first capture (t_c), and the age at the end of the first fishing season (t_s), for each sex. Also required is the fishing season duration (Δt) which is taken as being the same for each sex. When the catch numbers are by length, class, the lower length limit for each class (L_d) is required to be converted to age (t_d), using pre-determined values for the growth constants in the 'standard' von Bertalanffy equation.

The operative equations are:

$$\frac{\hat{C}_{t_d \rightarrow t_m}^A}{\hat{C}_{t_d \rightarrow t_m}^B} = \frac{\hat{C}_{t_d \rightarrow t_s}^A + \hat{C}_{t_s \rightarrow t_{s+1}}^A + \hat{C}_{t_{s+1} \rightarrow t_m}^A}{\hat{C}_{t_d \rightarrow t_s}^B + \hat{C}_{t_s \rightarrow t_{s+1}}^B + \hat{C}_{t_{s+1} \rightarrow t_m}^B} \quad \dots(1)$$

$$\hat{C}_{t_d \rightarrow t_s}^A = \frac{(Z-M)/(Z-M(1-\Delta t)) \cdot N_o \exp(-Mt_c - (Z-M(1-\Delta t))(t_d - t_c))}{\Delta t} \cdot \frac{(1 - \exp(-(Z-M(1-\Delta t))(t_s - t_d)))}{\Delta t}$$

in which $t_d = t_s$ when $t_d > t_s$; ...(2)

$$\hat{C}_{t_s \rightarrow t_{s+1}}^A = \frac{(Z-M)/(Z-M(1-\Delta t)) \cdot N_o \exp(-Mt_c - (Z-M(1-\Delta t))(t_d - t_c) - M(1-\Delta t))}{\Delta t} \cdot \frac{(\exp(-(Z-M(1-\Delta t))(t_d - (t_s + (1-\Delta t)))) - \exp(-(Z-M(1-\Delta t))))}{\Delta t} \quad \dots(3)$$

in which $t_d = t_s + (1-\Delta t)$ when $t_d < t_s + (1-\Delta t)$; and $t_d - (t_s + (1-\Delta t)) = \Delta t$ when $t_d > t_s + (1-\Delta t)$;

$$\hat{C}_{t_{s+1} \rightarrow t_m}^A = \frac{(Z-M)/Z \cdot N_o \exp(-Mt_c - (Z-M(1-\Delta t))(t_s - t_c) - \frac{3}{2}M(1-\Delta t) - (Z-M(1-\Delta t)))}{\Delta t} - Z(t_d - (t_s + 1 + (1-\Delta t))) \quad \dots(4)$$

in which $t_d = t_s + 1 + \frac{(1-\Delta t)}{2}$ when $t_d < t_s + 1 + \frac{(1-\Delta t)}{2}$;

and where $\hat{C}_{t_d \rightarrow t_m}^A / \hat{C}_{t_d \rightarrow t_m}^B$ is the catch number ratio, N_o^A and N_o^B are the stock numbers at zero age, t_m^A and t_m^B are the maximum ages, and A and B refers to the sexes.

The program has two options: *OPTION 1* when $M_A = M_B = M$ assumed
OPTION 2 when $F_A = F_B = F$ assumed

Catch number ratios are estimated, after assuming a value of M (*OPTION 1*) or F (*OPTION 2*), for each pair of t_d^A and t_d^B . The extent of proportional 'correlation' between the estimated and observed catch number ratios is then used to identify whether the trial value of M or F is a 'best choice' value. This involves determining the linear (geometric means) correlation and regression parameters; the coefficient of determination (r^2), the y-intercept (u), the slope (v); as well as the sum of the squared differences between the observed and estimated catch number ratios (ΣD^2). The ideal 'best choice' is taken as when the r^2 is maximized, u is zero, v is unity, and ΣD^2 is minimized.

WARNINGS: All mortality coefficients should be input as annual values. The units for the ages at first capture and at the end of the first fishing season should be years. The program estimation of $C_{t_d}^A \rightarrow t_m$ and $C_{t_d}^B \rightarrow t_m$ is based on $N_o^A = N_o^B = 10,000$.

Listing Notes

- a. Language HP 67/97 programming language
- b. Implementation Hewlett Packard
- c. Hardware Requirements HP 67/97
- d. RAM Requirements Nil
- d. RAM Requirements About 12K each program

Listing

A complete listing is given in Fig. 22.

Fig. 22. Program Listing.

			<u>SEX RATIO 1.</u>			<u>Option 1</u>		
STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
001	*LBL1	21 16 11	027	P2S	16-51	053	1	01
002	P2S	15-51	028	ST05	35 05	054	0	00
003	CLRG	16-53	029	R/S	51	055	0	00
004	P2S	16-51	030	*LBL1	21 01	056	0	00
005	CLRE	16-53	031	1	01	057	0	00
006	CLX	-51	032	RCLA	36 11	058	x	-35
007	R/S	51	033	RCL2	36 02	059	RCL0	36 00
008	ST00	35 00	034	÷	-24	060	RCLI	36 46
009	R/S	51	035	-	-45	061	-	-45
010	ST01	35 01	036	LH	32	062	x	-35
011	R/S	51	037	RCL3	36 03	063	RCL0	36 00
012	ST02	35 02	038	÷	-24	064	÷	-24
013	R/S	51	039	CHS	-22	065	PRTX	-14
014	ST03	35 03	040	RTN	24	066	RTN	24
015	R/S	51	041	*LBL2	21 02	067	*LBL3	21 03
016	ST04	35 04	042	RCLB	36 12	068	SPC	16-11
017	R/S	51	043	RCL1	36 01	069	ST00	35 14
018	P2S	16-51	044	-	-45	070	X2Y	-41
019	ST00	35 00	045	RCL0	36 00	071	ST0A	35 11
020	R/S	51	046	x	-35	072	PRTX	-14
021	ST01	35 01	047	CHS	-22	073	GSE1	23 01
022	R/S	51	048	RCLI	36 46	074	RCL4	36 04
023	ST02	35 02	049	RCL1	36 01	075	+	-55
024	R/S	51	050	x	-35	076	PRTX	-14
025	ST03	35 03	051	-	-45	077	ST0B	35 12
026	R/S	51	052	e ^x	33	078	RCLD	36 14

Continued

Fig. 22. Continued

STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
079	STOA	35 11	121	P±S	16-51	163	ENT↑	-21
080	PRTX	-14	122	GSB2	23 02	164	RCL6	36 06
081	P±S	16-51	123	P±S	16-51	165	X²	53
082	GSB1	23 01	124	RCL6	36 15	166	RCL9	36 09
083	P±S	16-51	125	÷	-24	167	÷	-24
084	RCL5	36 05	126	1/X	52	168	RCL7	36 07
085	+	-55	127	PRTX	-14	169	X²Y	-41
086	PRTX	-14	128	STOA	35 11	170	-	-45
087	STOC	35 13	129	R/S	51	171	÷	-24
088	RTN	24	130	*LBL0	21 13	172	STOB	35 12
089	*LBLA	21 11	131	SPC	16-11	173	X	-35
090	DSP4	-63 04	132	ST+8	35-55 08	174	RCL4	36 04
091	GSB3	23 03	133	R↓	-31	175	X²	53
092	RCL7	36 07	134	ST+9	35-55 09	176	RCL9	36 09
093	STOI	35 46	135	RCL9	36 09	177	÷	-24
094	GSB2	23 02	136	PRTX	-14	178	CHS	-22
095	STOE	35 15	137	RCL8	36 08	179	RCL5	36 05
096	RCLC	36 13	138	PRTX	-14	180	+	-55
097	STOB	35 12	139	÷	-24	181	÷	-24
098	P±S	16-51	140	STOD	35 14	182	PRTX	-14
099	GSB2	23 02	141	PRTX	-14	183	JX	54
100	P±S	16-51	142	RCLA	36 11	184	1/X	52
101	RCL6	36 15	143	X²Y	-41	185	RCLB	36 12
102	÷	-24	144	Σ+	56	186	X	-35
103	1/X	52	145	RCLD	36 14	187	STOE	35 15
104	PRTX	-14	146	RCLA	36 11	188	RCL6	36 06
105	STOA	35 11	147	-	-45	189	RCL6	36 06
106	R/S	51	148	X²	53	190	X	-35
107	*LBLB	21 12	149	ST+6	35-55 06	191	RCL9	36 05
108	DSP4	-63 04	150	R/S	51	192	÷	-24
109	STOA	35 11	151	*LBLD	21 14	193	CHS	-22
110	SPC	16-11	152	DSP9	-63 09	194	RCL4	36 04
111	X²Y	-41	153	SPC	16-11	195	RCL9	36 09
112	STOB	35 12	154	P±S	16-51	196	÷	-24
113	PRTX	-14	155	RCL8	36 08	197	+	-55
114	RCL7	36 07	156	RCL6	36 06	198	PRTX	-14
115	STOI	35 46	157	RCL4	36 04	199	RCL6	36 15
116	GSB2	23 02	158	X	-35	200	PRTX	-14
117	STOE	35 15	159	RCL9	36 09	201	P±S	16-51
118	RCLA	36 11	160	÷	-24	202	RCL6	36 06
119	PRTX	-14	161	-	-45	203	PRTX	-14
120	STOB	35 12	162	ENT↑	-21	204	R/S	51

PROGRAM LISTING

SEX RATIO 1Option 2

STEP ENTRY CODE

001 *LBLa 21 16 11
 002 P±S 16-51
 003 CLRG 16-53
 004 P±S 16-51

STEP ENTRY CODE

005 CLRG 16-53
 006 CLX -51
 007 R/S 51
 008 STOB 35 08

STEP ENTRY CODE

009 R/S 51
 010 STOI 35 01
 011 R/S 51
 012 STOB 35 02

Continued

Fig. 22. Continued

STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
013	R/S	51	065	PRTX	-14	117	STOA	35 11
014	ST03	35 03	066	RTN	24	118	SPC	16-11
015	R/S	51	067	*LBL3	21 03	119	X#Y	-41
016	ST04	35 04	068	SPC	16-11	120	STOB	35 12
017	R/S	51	069	STOD	35 14	121	PRTX	-14
018	P#S	16-51	070	X#Y	-41	122	RCL0	36 00
019	ST00	35 00	071	STOA	35 11	123	RCL7	36 07
020	R/S	51	072	PRTX	-14	124	-	-45
021	ST01	35 01	073	GSB1	23 01	125	STOI	35 46
022	R/S	51	074	RCL4	36 04	126	GSB2	23 02
023	ST02	35 02	075	+	-55	127	STOE	35 15
024	R/S	51	076	PRTX	-14	128	RCLA	36 11
025	ST03	35 03	077	STOB	35 12	129	PRTX	-14
026	R/S	51	078	RCLD	36 14	130	STOB	35 12
027	P#S	16-51	079	STOA	35 11	131	P#S	16-51
028	ST05	35 05	080	PRTX	-14	132	RCL0	36 00
029	R/S	51	081	P#S	16-51	133	P#S	16-51
030	*LBL1	21 01	082	GSB1	23 01	134	RCL7	36 07
031	1	01	083	P#S	16-51	135	-	-45
032	RCLA	36 11	084	RCL5	36 05	136	STOI	35 46
033	RCL2	36 02	085	+	-55	137	P#S	16-51
034	÷	-24	086	PRTX	-14	138	GSB2	23 02
035	-	-45	087	STOC	35 13	139	P#S	16-51
036	LH	32	088	RTN	24	140	RCLE	36 15
037	RCL3	36 03	089	*LBLA	21 11	141	÷	-24
038	÷	-24	090	DSP4	-63 04	142	1/X	52
039	CHS	-22	091	GSB3	23 03	143	PRTX	-14
040	RTN	24	092	RCL0	36 00	144	STOA	35 11
041	*LBL2	21 02	093	RCL7	36 07	145	R/S	51
042	RCLB	36 12	094	-	-45	146	*LBLC	21 13
043	RCL1	36 01	095	STOI	35 46	147	SPC	16-11
044	-	-45	096	GSB2	23 02	148	ST+8	35-55 08
045	RCL0	36 00	097	STOE	35 15	149	R↓	-31
046	X	-35	098	RCLC	36 13	150	ST+9	35-55 09
047	CHS	-22	099	STOB	35 12	151	RCL9	36 09
048	RCLI	36 46	100	P#S	16-51	152	PRTX	-14
049	RCL1	36 01	101	RCL0	36 00	153	RCL0	36 08
050	X	-35	102	P#S	16-51	154	PRTX	-14
051	-	-45	103	RCL7	36 07	155	÷	-24
052	e ^x	33	104	-	-45	156	STOD	35 14
053	1	01	105	STOI	35 46	157	PRTX	-14
054	0	00	106	P#S	16-51	158	RCLA	36 11
055	0	00	107	GSB2	23 02	159	X#Y	-41
056	0	00	108	P#S	16-51	160	Z+	56
057	0	00	109	RCLE	36 15	161	RCLD	36 14
058	X	-35	110	÷	-24	162	RCLA	36 11
059	RCL0	36 00	111	1/X	52	163	-	-45
060	RCLI	36 46	112	PRTX	-14	164	X ²	53
061	-	-45	113	STOA	35 11	165	ST+6	35-55 06
062	X	-35	114	R/S	51	166	R/S	51
063	RCL0	36 00	115	*LBLB	21 12	167	*LBLD	21 14
064	÷	-24	116	DSP4	-63 04	168	DSP9	-63 09

Continued

Fig. 22. Continued

STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
169	SPC	16-11	187	÷	-24	205	RCL6	36 06
170	P±S	16-51	188	STOB	35 12	206	x	-35
171	RCL8	36 09	189	x	-35	207	RCL9	36 09
172	RCL6	36 06	190	RCL4	36 04	208	÷	-24
173	RCL4	36 04	191	X ²	53	209	CHS	-22
174	x	-35	192	RCL9	36 09	210	RCL4	36 04
175	RCL9	36 09	193	÷	-24	211	RCL9	36 09
176	÷	-24	194	CHS	-22	212	÷	-24
177	-	-45	195	RCL5	36 05	213	+	-55
178	ENT↑	-21	196	+	-55	214	PRTX	-14
179	ENT↑	-21	197	÷	-24	215	RCL6	36 15
180	RCL6	36 06	198	PRTX	-14	216	PRTX	-14
181	X ²	53	199	JK	54	217	P±S	16-51
182	RCL9	36 09	200	1/X	52	218	RCL6	36 06
183	÷	-24	201	RCL8	36 12	219	PRTX	-14
184	RCL7	36 07	202	x	-35	220	R±S	51
185	X ² Y	-41	203	STOE	35 15			
186	-	-45	204	RCL6	36 15			

PROGRAM LISTING

SEX RATIO, 2Option 1

STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
001	*LBLA	21 11	026	ST+7	35-55 07	051	RCL8	36 08
002	1	01	027	RCL2	36 02	052	+	-55
003	STOC	35 13	028	RCLC	36 13	053	STOB	35 08
004	ST+9	35-55 09	029	x	-35	054	RCL6	36 06
005	R4	-31	030	ST+6	35-55 06	055	X ²	53
006	STO1	35 01	031	RCL1	36 01	056	RCL9	36 09
007	R4	-31	032	X ²	53	057	÷	-24
008	RCLD	36 14	033	RCLC	36 13	058	CHS	-22
009	Y ²	31	034	x	-35	059	RCL7	36 07
010	RCLA	36 11	035	ST+5	35-55 05	060	+	-55
011	RCLD	36 14	036	RCL1	36 01	061	P±S	16-51
012	Y ²	31	037	RCLC	36 13	062	STO1	35 01
013	-	-45	038	x	-35	063	P±S	16-51
014	CHS	-22	039	ST+4	35-55 04	064	RCL4	36 04
015	LN	32	040	ISZI	16 26 46	065	X ²	53
016	STO2	35 02	041	RCLI	36 46	066	RCL9	36 09
017	RCL1	36 01	042	RTN	24	067	÷	-24
018	x	-35	043	*LBLB	21 12	068	CHS	-22
019	RCLC	36 13	044	SPC	16-11	069	RCL5	36 05
020	x	-35	045	RCL4	36 04	070	+	-55
021	ST+8	35-55 08	046	RCL6	36 06	071	P±S	16-51
022	RCL2	36 02	047	x	-35	072	STOB	35 08
023	X ²	53	048	RCL9	36 09	073	RCL1	36 01
024	RCLC	36 13	049	÷	-24	074	P±S	16-51
025	x	-35	050	CHS	-22	075	x	-35

Continued

Fig. 22. Continued

STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
076	ST03	35 03	127	RTN	24	178	RCL4	36 04
077	RCL8	36 00	128	R/S	51	179	X ²	53
078	X ²	53	129	*LBL0	21 13	180	RCL9	36 09
079	RCL3	36 03	130	PRTX	-14	181	÷	-24
080	÷	-24	131	R↓	-31	182	CHS	-22
081	PRTX	-14	132	ST03	35 03	183	RCL5	36 05
082	RCL8	36 11	133	PRTX	-14	184	+	-55
083	PRTX	-14	134	R↑	16-31	185	÷	-24
084	RCL8	36 00	135	RCL2	36 02	186	PRTX	-14
085	P±S	16-51	136	-	-45	187	JX	54
086	RCL8	36 00	137	RCL1	36 01	188	1/X	52
087	P±S	16-51	138	X	-35	189	RCL8	36 12
088	÷	-24	139	RCLD	36 14	190	X	-35
089	ST03	35 03	140	X	-35	191	ST01	35 46
090	CHS	-22	141	CHS	-22	192	RCL4	36 04
091	RCLD	36 14	142	e ^X	33	193	RCL6	36 06
092	÷	-24	143	CHS	-22	194	RCL8	36 12
093	ST08	35 12	144	1	91	195	X	-35
094	PRTX	-14	145	+	-55	196	-	-45
095	RCL4	36 04	146	RCLD	36 14	197	RCL9	36 09
096	RCL9	36 09	147	1/X	52	198	÷	-24
097	÷	-24	148	Y ^X	31	199	ST0A	35 11
098	RCL3	36 03	149	RCL8	36 00	200	PRTX	-14
099	X	-35	150	X	-35	201	RCL8	36 12
100	CHS	-22	151	PRTX	-14	202	PRTX	-14
101	RCL8	36 06	152	RCL3	36 03	203	SPC	16-11
102	RCL9	36 09	153	X ² Y	-41	204	RCLI	36 46
103	÷	-24	154	Z+	56	205	RCL6	36 06
104	+	-55	155	SPC	16-11	206	X	-35
105	RCLA	36 11	156	R/S	51	207	RCL9	36 09
106	RCLD	36 14	157	*LBL0	21 14	208	÷	-24
107	Y ^X	31	158	P±S	16-51	209	CHS	-22
108	LH	32	159	RCL9	36 09	210	RCL4	36 04
109	X ² Y	-41	160	RCL6	36 06	211	RCL9	36 09
110	-	-45	161	RCL4	36 04	212	÷	-24
111	RCL3	36 03	162	X	-35	213	+	-55
112	÷	-24	163	RCL9	36 09	214	PRTX	-14
113	ST0E	35 15	164	÷	-24	215	RCLI	36 46
114	PRTX	-14	165	-	-45	216	PRTX	-14
115	RCLA	36 11	166	ENT↑	-21	217	P±S	16-51
116	RCLB	36 12	167	ENT↑	-21	218	R/S	51
117	RCL8	36 15	168	RCL6	36 06	219	*LBL0	21 16 11
118	CLRG	16-53	169	X ²	53	220	CLRG	16-53
119	P±S	16-51	170	RCL9	36 09	221	P±S	16-51
120	CLRG	16-53	171	÷	-24	222	CLRG	16-53
121	P±S	16-51	172	RCL7	36 07	223	CLX	-51
122	ST02	35 02	173	X ² Y	-41	224	R/S	51
123	R↓	-31	174	-	-45			
124	ST01	35 01	175	÷	-24			
125	R↓	-31	176	ST08	35 12			
126	ST00	35 00	177	X	-35			

Continued

Fig. 22. Continued

PROGRAM LISTING

SEX RATIO 2.Option 2)

STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
001	*LBLA	21 11	048	RCL9	36 09	095	RCL4	36 04
002	1	01	049	÷	-24	096	RCL9	36 09
003	STOC	35 13	050	CHS	-22	097	÷	-24
004	ST+9	35-55 09	051	RCL8	36 08	098	RCL3	36 03
005	R↓	-31	052	+	-55	099	x	-35
006	ST01	35 01	053	ST00	35 00	100	CHS	-22
007	R↓	-31	054	RCL6	36 06	101	RCL6	36 06
008	RCLD	36 14	055	X²	-53	102	RCL9	36 09
009	YX	31	056	RCL9	36 09	103	÷	-24
010	RCLA	36 11	057	÷	-24	104	+	-55
011	RCLD	36 14	058	CHS	-22	105	RCLA	36 11
012	YX	31	059	RCL7	36 07	106	RCLD	36 14
013	-	-45	060	+	-55	107	YX	31
014	CHS	-22	061	P±S	16-51	108	LN	32
015	LN	32	062	ST01	35 01	109	X²Y	-41
016	ST02	35 02	063	P±S	16-51	110	-	-45
017	RCL1	36 01	064	RCL4	36 04	111	RCL3	36 03
018	X	-35	065	X²	53	112	÷	-24
019	RCLC	36 13	066	RCL9	36 09	113	STOE	35 15
020	X	-35	067	÷	-24	114	PRTX	-14
021	ST+8	35-55 08	068	CHS	-22	115	RCLA	36 11
022	RCL2	36 02	069	RCL5	36 05	116	RCLB	36 12
023	X²	53	070	+	-55	117	RCLE	36 15
024	RCLC	36 13	071	P±S	16-51	118	CLRG	16-53
025	X	-35	072	ST00	35 00	119	P±S	16-51
026	ST+7	35-55 07	073	RCL1	36 01	120	CLRG	16-53
027	RCL2	36 02	074	P±S	16-51	121	P±S	16-51
028	RCLC	36 13	075	X	-35	122	ST02	35 02
029	X	-35	076	ST03	35 03	123	R↓	-31
030	ST+6	35-55 06	077	RCL0	36 00	124	ST01	35 01
031	RCL1	36 01	078	X²	53	125	R↓	-31
032	X²	53	079	RCL3	36 03	126	ST00	35 00
033	RCLC	36 13	080	÷	-24	127	RTN	24
034	X	-35	081	PRTX	-14	128	R/S	51
035	ST+5	35-55 05	082	RCLA	36 11	129	*LBLC	21 13
036	RCL1	36 01	083	PRTX	-14	130	PRTX	-14
037	RCLC	36 13	084	RCL0	36 00	131	ST03	35 03
038	X	-35	085	P±S	16-51	132	R↓	-31
039	ST+4	35-55 04	086	RCL0	36 00	133	PRTX	-14
040	ISZI	16 26 46	087	P±S	16-51	134	CHS	-22
041	RCLI	36 46	088	÷	-24	135	RCL0	36 00
042	RTN	24	089	ST03	35 03	136	+	-55
043	*LBLB	21 12	090	CHS	-22	137	1/X	52
044	SPC	16-11	091	RCLD	36 14	138	RCL9	36 09
045	RCL4	36 04	092	÷	-24	139	X	-35
046	RCL6	36 06	093	ST0B	35 12	140	LN	32
047	X	-35	094	PRTX	-14	141	RCL1	36 01

Continued

Fig. 22. Continued

STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
142	RCLD	36 14	168	RCL7	36 07	194	-	-24
143	X	-35	169	X ² Y	-41	195	STOA	35 11
144	-	-24	170	-	-45	196	PRTX	-14
145	RCL2	36 02	171	-	-24	197	RCLB	36 12
146	+	-55	172	STOB	35 12	198	PRTX	-14
147	PRTX	-14	173	X	-35	199	SPC	16-11
148	RCL3	36 03	174	RCL4	36 04	200	RCLI	36 45
149	X ² Y	-41	175	X ²	53	201	RCL6	36 06
150	Z+	56	176	RCL9	36 09	202	X	-35
151	SPC	16-11	177	-	-24	203	RCL9	36 09
152	R/S	51	178	CHS	-22	204	-	-24
153	*LBL0	21 14	179	RCL5	36 05	205	CHS	-22
154	P ² S	16-51	180	+	-55	206	RCL4	36 04
155	RCL8	36 08	181	-	-24	207	RCL9	36 09
156	RCL6	36 06	182	PRTX	-14	208	-	-24
157	RCL4	36 04	183	JX	54	209	+	-55
158	X	-35	184	1/X	52	210	PRTX	-14
159	RCL9	36 09	185	RCLB	36 12	211	RCLI	36 46
160	-	-24	186	X	-35	212	PRTX	-14
161	-	-45	187	STOI	35 46	213	P ² S	16-51
162	ENT↑	-21	188	RCL4	36 04	214	R/S	51
163	ENT↑	-21	189	RCL6	36 06	215	*LBL0	21 16 11
164	RCL6	36 06	190	RCLB	36 12	216	CLRG	16-53
165	X ²	53	191	X	-35	217	P ² S	16-51
166	RCL9	36 09	192	-	-45	218	CLRG	16-53
167	-	-24	193	RCL9	36 09	219	CLX	-51
						220	R/S	51

PROGRAM LISTING

SEX RATIO 2

(Card 1)

STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
001	*LBL0	21 16 11	018	R/S	51	035	-	-45
002	P ² S	16-51	019	STO2	35 02	036	RCLD	36 14
003	CLRG	16-53	020	P ² S	16-51	037	-	-24
004	P ² S	16-51	021	R/S	51	038	X ² 02	16-45
005	CLRG	16-53	022	STO0	35 14	039	0	00
006	CLX	-51	023	R/S	51	040	X	-35
007	R/S	51	024	*LBL1	21 01	041	CHS	-22
008	STO8	35 00	025	RCLB	36 08	042	e ^x	33
009	R/S	51	026	1	01	043	CHS	-22
010	STO1	35 01	027	RCLD	36 14	044	1	01
011	R/S	51	028	-	-45	045	+	-55
012	STO2	35 02	029	RCLI	36 46	046	RCLB	36 12
013	R/S	51	030	X	-35	047	RCL1	36 01
014	P ² S	16-51	031	-	-45	048	-	-45
015	STO8	35 00	032	STOC	35 13	049	RCLD	36 14
016	R/S	51	033	RCL2	36 02	050	-	-24
017	STO1	35 01	034	RCLB	36 12	051	RCLC	36 13

Continued

Fig. 22. Continued

STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
052	x	-35	103	RCLI	36 46	154	x	-35
053	CHS	-22	104	-	-45	155	ST+3	35-55 03
054	e ^x	33	105	x	-35	156	RCL1	36 01
055	x	-35	106	ST+3	35-55 03	157	RCLI	36 46
056	RCLC	36 13	107	1	01	158	x	-35
057	÷	-24	108	.	-62	159	CHS	-22
058	RCL0	36 00	109	5	05	160	e ^x	33
059	RCLI	36 46	110	RCLD	36 14	161	1	01
060	-	-45	111	2	02	162	0	09
061	x	-35	112	÷	-24	163	ENT↑	-21
062	ST03	35 03	113	-	-45	164	4	04
063	RCLB	36 12	114	RCL2	36 02	165	Y ^x	31
064	RCL2	36 02	115	+	-55	166	x	-35
065	1	01	116	CHS	-22	167	RCL3	36 03
066	+	-55	117	RCLB	36 12	168	x	-35
067	-	-45	118	+	-55	169	PRTX	-14
068	X>0?	16-44	119	X<0?	16-45	170	RTN	24
069	0	00	120	0	00	171	*LBLA	21 11
070	RCLD	36 14	121	RCL0	36 00	172	DSF4	-63 04
071	+	-55	122	x	-35	173	STOA	35 11
072	X<0?	16-45	123	RCL2	36 02	174	SPC	16-11
073	0	00	124	RCL1	36 01	175	X ^Y	-41
074	RCLD	36 14	125	-	-45	176	STOB	35 12
075	÷	-24	126	RCLD	36 14	177	FRTX	-14
076	RCLC	36 13	127	÷	-24	178	RCL7	36 07
077	x	-35	128	RCLC	36 13	179	STOI	35 46
078	CHS	-22	129	x	-35	180	GSB1	23 01
079	e ^x	33	130	+	-55	181	STOE	35 15
080	RCLC	36 13	131	1	01	182	RCLA	36 11
081	CHS	-22	132	RCLD	36 14	183	PRTX	-14
082	e ^x	33	133	-	-45	184	STOB	35 12
083	-	-45	134	RCLI	36 46	185	P ^z S	16-51
084	RCL2	36 02	135	x	-35	186	GSB1	23 01
085	RCL1	36 01	136	+	-55	187	P ^z S	16-51
086	-	-45	137	RCLC	36 13	188	RCLE	36 15
087	RCLD	36 14	138	+	-55	189	÷	-24
088	÷	-24	139	1	01	190	1/X	52
089	RCLC	36 13	140	RCLD	36 14	191	PRTX	-14
090	x	-35	141	-	-45	192	STOA	35 11
091	1	01	142	2	02	193	R/S	51
092	RCLD	36 14	143	÷	-24	194	*LBLB	21 12
093	-	-45	144	RCLI	36 46	195	SFC	16-11
094	RCLI	36 46	145	x	-35	196	ST+2	35-55 02
095	x	-35	146	+	-55	197	R↓	-31
096	+	-55	147	CHS	-22	198	ST+9	35-55 09
097	CHS	-22	148	e ^x	33	199	RCL9	36 09
098	e ^x	33	149	RCL0	36 00	200	PRTX	-14
099	x	-35	150	÷	-24	201	RCL3	36 08
100	RCLC	36 13	151	RCL0	36 00	202	PRTX	-14
101	÷	-24	152	RCLI	36 46	203	÷	-24
102	RCL0	36 00	153	-	-45	204	STOE	35 15

Continued

Fig. 22. Continued

STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
205	FRTX	-14	209	RCLB	36 15	212	X ²	53
206	RCLA	36 11	210	RCLA	36 11	213	ST+6	35-55 06
207	X ² Y	-41	211	-	-45	214	R/S	51
208	S+	56						

PROGRAM LISTING

SEX RATIO 2

(Card 2)

STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
001	*LBL _a	21 16 11	041	CHS	-22	081	CHS	-22
002	P ² S	16-51	042	e ^x	33	082	e ^x	33
003	CLRG	16-53	043	CHS	-22	083	-	-45
004	P ² S	16-51	044	1	01	084	RCL2	36 02
005	CLRG	16-53	045	+	-55	085	RCL1	36 01
006	CLX	-51	046	RCLB	36 12	086	-	-45
007	R/S	51	047	RCL1	36 01	087	RCLD	36 14
008	ST00	35 00	048	-	-45	088	÷	-24
009	R/S	51	049	RCLD	36 14	089	RCLC	36 13
010	ST01	35 01	050	÷	-24	090	X	-35
011	R/S	51	051	RCLC	36 13	091	1	01
012	ST02	35 02	052	X	-35	092	RCLD	36 14
013	R/S	51	053	CHS	-22	093	-	-45
014	P ² S	16-51	054	e ^x	33	094	RCLI	36 46
015	ST00	35 00	055	X	-35	095	X	-35
016	R/S	51	056	RCLC	36 13	096	+	-55
017	ST01	35 01	057	÷	-24	097	CHS	-22
018	R/S	51	058	RCLB	36 00	098	e ^x	33
019	ST02	35 02	059	RCLI	36 46	099	X	-35
020	P ² S	16-51	060	-	-45	100	RCLC	36 13
021	R/S	51	061	X	-35	101	÷	-24
022	ST00	35 14	062	ST03	35 03	102	RCLB	36 00
023	R/S	51	063	RCLB	36 12	103	RCLI	36 46
024	*LBL1	21 01	064	RCL2	36 02	104	-	-45
025	RCL0	36 00	065	1	01	105	X	-35
026	1	01	066	+	-55	106	ST+3	35-55 03
027	RCLD	36 14	067	-	-45	107	1	01
028	-	-45	068	X<0?	16-44	108	-	-62
029	RCLI	36 46	069	0	00	109	5	05
030	X	-35	070	RCLD	36 14	110	RCLD	36 14
031	-	-45	071	+	-55	111	2	02
032	ST00	35 13	072	X<0?	16-45	112	÷	-24
033	RCL2	36 02	073	0	00	113	-	-45
034	RCLB	36 12	074	RCLD	36 14	114	RCL2	36 02
035	-	-45	075	÷	-24	115	+	-55
036	RCLD	36 14	076	RCLC	36 13	116	CHS	-22
037	÷	-24	077	X	-35	117	RCLB	36 12
038	X<0?	16-45	078	CHS	-22	118	+	-55
039	0	00	079	e ^x	33	119	X<0?	16-45
040	X	-35	080	RCLC	36 13	120	0	00

Continued

Fig. 22. Continued

STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
121	RCL0	36 00	155	ST+3	35-55 03	189	P#S	16-51
122	x	-35	156	RCL1	36 01	190	RCL7	36 07
123	RCL2	36 02	157	RCLI	36 46	191	-	-45
124	RCL1	36 01	158	x	-35	192	ST01	35 46
125	-	-45	159	CHS	-22	193	P#S	16-51
126	RCLD	36 14	160	e ^x	33	194	GSB1	23 01
127	÷	-24	161	1	01	195	P#S	16-51
128	RCLC	36 13	162	0	00	196	RCLC	36 15
129	y	-35	163	ENT↑	-21	197	÷	-24
130	+	-55	164	4	04	198	1/X	52
131	1	01	165	Y ^x	31	199	PRTX	-14
132	RCLD	36 14	166	x	-35	200	ST0A	35 11
133	-	-45	167	RCL3	36 03	201	R/S	51
134	RCLI	36 46	168	x	-35	202	*LBLB	21 12
135	x	-35	169	PRTX	-14	203	SPC	16-11
136	+	-55	170	RTN	24	204	ST+8	35-55 08
137	RCLC	36 13	171	*LSLA	21 11	205	R4	-31
138	+	-55	172	DSP4	-53 04	206	ST+9	35-55 09
139	1	01	173	ST0A	35 11	207	PCL9	35 09
140	RCLD	36 14	174	SPC	16-11	208	PRTX	-14
141	-	-45	175	X#Y	-41	209	RCL8	36 08
142	2	02	176	ST0B	35 12	210	PRTX	-14
143	÷	-24	177	PRTX	-14	211	÷	-24
144	RCLI	36 46	178	RCL0	36 00	212	ST0E	35 15
145	x	-35	179	RCL7	36 07	213	PRTX	-14
146	+	-55	180	-	-45	214	RCLA	36 11
147	CHS	-22	181	ST01	35 46	215	X#Y	-41
148	e ^x	33	182	GSB1	23 01	216	Σ+	56
149	RCL0	36 00	183	ST0E	35 15	217	RCLC	36 15
150	÷	-24	184	RCLA	36 11	218	RCLA	36 11
151	RCL0	36 00	185	PRTX	-14	219	-	-45
152	RCLI	36 46	186	ST0B	35 12	220	X ²	53
153	-	-45	187	P#S	16-51	221	ST+6	35-55 06
154	x	-35	188	RCL0	36 00	222	R/S	51

PROGRAM LISTING

SEX RATIO 2

(Card 3)

STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
001	*LBLa	21 16 11	009	R/S	51	017	ST01	35 01
002	P#S	16-51	010	ST01	35 01	018	R/S	51
003	CLRG	16-53	011	R/S	51	019	ST02	35 02
004	P#S	16-51	012	ST02	35 02	020	R/S	51
005	CLRG	16-53	013	R/S	51	021	*LBL1	21 01
006	CLX	-51	014	P#S	16-51	022	RCL0	36 00
007	R/S	51	015	ST0B	35 00	023	÷	-24
008	ST0B	35 00	016	R/S	51	024	CHS	-22

Continued

Fig. 22. Continued

STEP	ENTRY	CODE	STEP	ENTRY	CODE	STEP	ENTRY	CODE
025	1	01	050	RCL8	36 08	075	+	-55
026	+	-55	051	RCL6	36 06	076	÷	-24
027	LH	32	052	RCL4	36 04	077	PRTX	-14
028	RCL1	36 01	053	x	-35	078	JX	54
029	÷	-24	054	RCL9	36 09	079	1/X	52
030	CHS	-22	055	÷	-24	080	RCLB	36 12
031	RCL2	36 02	056	-	-45	081	x	-35
032	+	-55	057	ENT↑	-21	082	STOE	35 15
033	PRTX	-14	058	ENT↑	-21	083	RCLE	36 15
034	SPC	16-11	059	RCL6	36 06	084	RCLE	36 06
035	RTN	24	060	X ²	53	085	x	-35
036	*LBLA	21 11	061	RCL9	36 09	086	RCL9	36 09
037	DSP4	-63 04	062	÷	-24	087	÷	-24
038	PRTX	-14	063	RCL7	36 07	088	CHS	-22
039	GSB1	23 01	064	X ² Y	-41	089	RCL4	36 04
040	R↓	-31	065	-	-45	090	RCL9	36 09
041	PRTX	-14	066	÷	-24	091	÷	-24
042	P ² S	16-51	067	STOB	35 12	092	+	-55
043	GSB1	23 01	068	x	-35	093	PRTX	-14
044	P ² S	16-51	069	RCL4	36 04	094	RCLE	36 15
045	R/S	51	070	X ²	53	095	PRTX	-14
046	*LBLB	21 12	071	RCL9	36 09	096	P ² S	16-51
047	DSP9	-63 03	072	÷	-24	097	RCL6	36 06
048	SPC	16-11	073	CHS	-22	098	PRTX	-14
049	P ² S	16-51	074	RCL5	36 05	099	R/S	51

User Instructions — Sex Ratio 1

Step	Instructions	Input data/units	Keys	Output data/units
1	Enter both sides of appropriate option card			
2	Clear storages		a	
3	Enter parameter values			
	— Tot. mortality coef.	Z _c ^A	R/S	
	— Age at first capture	t _c ^A	R/S	
	— Growth constants	L _∞ ^A	R/S	
		K ^A	R/S	
		t _o ^A	R/S	
	— Tot. mortality coef.	Z _B	R/S	
	— Age at first capture	t _c ^B	R/S	
	— Growth constants	L _∞ ^B	R/S	
		K ^B	R/S	
		t _o ^B	R/S	Continued

User Instructions: Sex Ratio 1 (Continued)

Step	Instructions	Input data/units	Keys	Output data/units
4	Enter trial value of M (OPTION 1) or F (OPTION 2)	M or F	STO 7	
5(a)	Enter lower length limits of length class; commencing from the class containing the longer fish.	L_d^A L_d^B	A ↑	L_d^A \hat{t}_d^A L_d^B \hat{t}_d^B \hat{C}_A^* \hat{C}_B \hat{C}_A / \hat{C}_B
	or			
5(b)	Enter lower age limits of age class; commencing from the class containing the oldest fish.	t_d^A t_d^B	B ↑	t_d^A \hat{C}_A t_d^B \hat{C}_B \hat{C}_A / \hat{C}_B
6	Enter C_A and C_B for length or age class	C_A C_B	C ↑	C_A C_B
7	Repeat steps 5 and 6 for each length or age class; progressing towards the class containing the smallest or youngest fish			C_A / C_B
8	Estimate the linear (geometric mean) correlation and regression parameters, and sum of squared deviation of estimated against observed catch number ratios.		D	r^2 u v ΣD^2
9	Repeat steps 2 to 8 for other trial values of M or F until 'best choices' identified from value of r^2 , u, v and ΣD^2 .			

* $C_{t_d}^A \rightarrow t_m$ and $C_{t_d}^B \rightarrow t_m$ abbreviated to C_A and C_B

User Instructions: Sex Ratio 2

Step	Instructions	Input data/units	Keys	Output data/units
1*	Enter both sides of appropriate option card (Card 1 or Card 2)			
2	Clear storage keys		a	
3	Enter parameter values			
	— Tot. mortality coef.	Z_A	R/S	
	— Age at first capture	t_c^A	R/S	
	— Age of end first season	t_s^A	R/S	
	— Tot. mortality coef.	Z_B	R/S	
	— Age at first capture	t_c^B	R/S	
	— Age at end first season	t_s^B	R/S	
	— Season duration	Δt	R/S	
4	Enter trial value of M (OPTION 1) or F (OPTION 2)	M or F	STO 7	
5	Enter lower age limits of age class; commencing from the class containing the oldest fish	t_d^A t_d^B	B	t_d^A \hat{C}_A^{**} t_d^B \hat{C}_B \hat{C}_A / \hat{C}_B
6	Enter C_A and C_B for age class	C_A C_B	C	C_A C_B C_A / C_B
7	Repeat steps 5 and 6 for each age class; progressing towards the class containing the youngest fish			
8	Enter side 1 of Card 3			
9	Estimate linear (geometric mean) correlation and regression parameters, and sum of squared deviations of estimated against observed catch number ratios		B	r^2 u v ΣD^2

Continued

User Instructions: Sex Ratio 2 (Continued)

Step	Instructions	Input data/units	Keys	Output data/units
10	Repeat steps 2 to 8 for other trial values of M or F until 'best choices' identified from the values of r^2 , u, v, and ΣD^2 . * When the data are as frequencies by length class, the lower length of the class must be converted to its age equivalent before proceeding through the above step 1. This can be achieved by the following:			
1	Enter side 1 of Card 3			
2	Clear storages		a	
3	Enter parameter values — Growth constants	L_{∞}^A	R/S	
		K^A	R/S	
		t_o^A	R/S	
	— Growth constants	L_{∞}^B	R/S	
		K^B	R/S	
		t_o^B	R/S	
4	Enter lower length limits of length class; proceeding through all classes as required	L_d^A	A	L_d^A
		L_d^B		t_d^A
				L_d^B
				t_d^B

** $C_{t_d}^A \rightarrow t_m$ and $C_{t_d}^B \rightarrow t_m$ abbreviated to C_A and C_B

Test Data - Sex Ratio 1

Males

F = 2.25
M = 0.6
Z = 2.85
 $t_c = 1.666$

Females

F = 1.8
M = 0.6
Z = 2.4
 $t_c = 1.166$

$L_\infty = 25.0$.
K = 0.4 (both sexes)

Simulated and estimated catch number and catch number ratios and 'correlation' parameters; | non-seasonal fishery where $F_m \neq F_f$, $M_m = M_f$ and $t_c^m \neq t_c^f$

Lower Length Limits L_d^A	L_d^B	Age t_d^B	Lower Limits t_d^B	Catch Numbers		Cumulative Catch Numbers C_A	C_A/C_B	When Ass. F=3.6		When Ass. F=3.7		When Ass. F=3.8		When Ass. M=0.5		When Ass. M=0.6		When Ass. M=0.7		
				c_A	c_B			\hat{c}_A	\hat{c}_B	\hat{c}_A/\hat{c}_B	\hat{c}_A	\hat{c}_B	\hat{c}_A/\hat{c}_B	\hat{c}_A	\hat{c}_B	\hat{c}_A/\hat{c}_B	\hat{c}_A	\hat{c}_B	\hat{c}_A/\hat{c}_B	
14.491	12.165	1.666	1.166	-	1227.9	2904	3724	0.78	44089	60828	0.72	53532	70254	0.76	64950	81081	0.80	3584	4418	0.81
14.491	12.992	1.666	1.333	-	823.1	2904	2497	1.16	44089	40774	1.08	53532	47093	1.14	64950	54351	1.19	3584	2961	1.21
14.491	13.767	1.666	1.500	-	551.7	2904	1674	1.74	44089	27332	1.61	53532	31567	1.70	64950	36432	1.78	3584	1985	1.81
14.491	14.491	1.66	1.666	1098.2	369.8	2904	1122	2.59	489	18321	2.41	53532	21160	2.53	64950	24422	2.66	3584	1331	2.69
15.169	15.169	1.833	1.833	682.9	247.9	1806	752	2.40	27418	12281	2.23	33291	14184	2.35	40391	16370	2.47	2229	892	2.50
15.803	15.803	2.000	2.000	424.7	166.2	1123	504	2.23	17051	8232	2.07	20703	9508	2.18	25118	10973	2.29	1386	598	2.32
16.396	16.396	2.166	2.166	264.1	111.4	699	338	2.07	10604	5518	1.92	12875	6373	2.02	15621	7356	2.12	862	401	2.15
16.951	16.951	2.333	2.333	164.3	74.7	434	226	1.92	6594	3699	1.78	8007	4272	1.87	9714	4931	1.97	536	269	2.00
17.470	17.470	2.500	2.500	102.2	50.1	270	152	1.78	4101	2479	1.56	4979	2864	1.74	6041	3305	1.83	333	180	1.85
17.956	17.956	2.666	2.666	63.5	33.5	168	103	1.65	2550	1662	1.53	3097	1920	1.61	3757	2215	1.70	207	121	1.72
18.410	18.410	2.833	2.833	39.5	22.5	104.5	68.2	1.53	1586	1114	1.42	1926	1287	1.50	2336	1485	1.57	129	81	1.59
18.835	18.835	3.000	3.000	24.6	15.1	65.0	45.7	1.42	986	747	1.32	1198	863	1.39	1453	995	1.46	80	54	1.48
19.233	19.233	3.166	3.166	15.3	10.1	40.4	30.6	1.32	613	501	1.23	745	578	1.29	904	667	1.35	50	36	1.37
19.605	19.605	3.333	3.333	25.1	20.5	25.1	20.5	1.22	381	336	1.14	463	388	1.19	562	447	1.26	31	24	1.27
				$r^2 = 1.00000$		$r^2 = 1.00000$		$r^2 = 1.00000$		$r^2 = 1.00000$		$r^2 = 1.00000$		$r^2 = 1.00000$		$r^2 = 1.00000$				
				$u = 0.00153$		$u = 0.00152$		$u = 0.00155$		$u = 0.00152$		$u = 0.00145$		$u = 0.00154$						
				$v = 1.07524$		$v = 1.02282$		$v = 0.97292$		$v = 0.96079$		$v = 0.99944$		$v = 1.03842$						
				$\Sigma D^2 = 0.21964$		$\Sigma D^2 = 0.02345$		$\Sigma D^2 = 0.03192$		$\Sigma D^2 = 0.07005$		$\Sigma D^2 = 0.00001$		$\Sigma D^2 = 0.06272$						

Test Data - Sex Ratio 2

Males		Females		
$F = 0.3$		$F = 0.3$		$L_{\infty} = 25.0$
$M = 0.6$		$M = 0.3$		
$Z = 0.9$		$Z = 0.6$		
$t_c = 1.25$		$t_c = 1.416$		$M = 0.4$ (both sexes)
$t_s = 1.916$		$t_s = 1.916$		
$\Delta t = 0.666$				

Simulated and estimated catch numbers and catch number ratios and 'correlation' parameters; seasonal fishery where $F_m = F_f$, $M_m \neq M_f$ and $t_c^m \neq t_c^f$

Length L_d^A	Lower Limit L_d^B	Age Limits		Catch Numbers		Cumulative Catch Num bers			When Ass.F=0.2		When Ass.F=0.3		When Ass.F=0.4		When Ass.M=1.2		When Ass.M=1.3		When Ass. M=1.4							
		t_d^A	t_d^B	c_A	c_B	C_A	C_B	C_A/C_B	C_A	C_B	C_A/C_B	C_A	C_B	C_A/C_B	C_A	C_B	C_A/C_B	C_A	C_B	C_A/C_B						
12.585	13.386	1.250	1.416	1091.1	1226.6	1717	3354	0.51	1028	1962	0.52	1721	3359	0.51	2562	5114	0.50	-891	-2043	0.44	-1065	-2089	0.51	-1193	-2092	0.57
16.104	16.104	2.083	2.083	414.3	959.8	698	2127	0.33	419	1244	0.34	702	2133	0.33	1044	3251	0.32	-364	-1280	0.28	-436	-1306	0.33	-488	-1306	0.37
19.037	19.037	3.083	3.083	168.5	526.8	284	1168	0.24	172	684	0.25	288	1173	0.25	427	1787	0.24	-150	-707	0.21	-180	-721	0.25	-202	-722	0.28
21.003	21.003	4.083	4.083	68.5	289.1	115	641	0.1	70	376	0.19	117	644	0.18	174	981	0.18	-61	-388	0.16	-73	-396	0.18	-82	-396	0.21
22.321	22.321	5.083	5.083	27.8	158.7	47	352	0.13	28	206	0.14	48	353	0.13	71	538	0.13	-25	-213	0.12	-30	-217	0.14	-33	-217	0.15
23.204	23.204	6.083	6.083	11.3	87.1	19	193	0.10	12	113	0.10	19	194	0.10	29	295	0.10	-10	-117	0.09	-12	-119	0.10	-14	-119	0.11
23.796	23.796	7.083	7.083	4.6	47.8	7.7	106	0.07	5	62	0.08	8	106	0.07	12	162	0.07	-4	-64	0.06	-5	-65	0.08	-6	-65	0.08
24.193	24.193	8.083	8.083	1.9	26.2	3.2	58.0	0.05	2	34	0.06	3	58	0.05	5	89	0.05	-2	-35	0.05	-2	-36	0.06	-2	-36	0.06
24.459	24.459	9.083	9.083	1.25	31.82	1.2	31.8	0.04	1	19	0.04	1	32	0.04	2	49	0.04	-1	-19	0.04	-1	-20	0.04	-1	-20	0.05
$r^2 = 0.99998$																					$r^2 = 0.99972$					
$u = -0.00144$																					$u = -0.00359$					
$v = 0.97800$																					$v = 0.89667$					
$\Sigma D^2 = 0.00039$																					$\Sigma D^2 = 0.00836$					

Program Number 12

Title	RJ33
Author	R. Jones
Author's Address	R. Jones Marine Laboratories Department of Agriculture & Fisheries for Scotland PO Box 101, Aberdeen ABG 8DB Scotland, United Kingdom

Reference Jones, R. — An investigation of length composition analysis using simulated length compositions (Part I, this vol.).

Program Description: Length Composition Analysis

This program processes length composition data to give estimates of the long-term effects on yield per recruit of changes in mesh size and/or fishing effort.

The outputs also include:

1. Length cohort analysis, including estimates of mortality rate by length group.
2. The cohort analysis table includes estimates of the values of Z/K by the Beverton and Holt method. For each length group, a value of Z/K is calculated from the relationship:

$$Z/K = (L_{\infty} - \bar{L}) / (\bar{L} - L_c)$$

- Where L_c is the lower limit of the length group concerned and \bar{L} is the mean value for all lengths $\geq L_c$.
3. The cohort analysis table includes estimates of F/S, for each length group, where S is the proportion retained, at the length concerned, by the mesh size in use.
 4. Immediate effects of changes in mesh size and/or fishing effort.
 5. Intermediate effects of changes in mesh size and/or fishing effort.

Inputs to the program require:

1. A data file, information about which is given in program lines 370-630.
2. Growth and natural mortality parameters.
3. Selectivity data (selection factor and slope of selection curve at the 50% length).

Input options include:

1. An allowance for discarding of fish and the possibility that some discarded individuals may survive.
2. The adjustment of numbers landed at each length in such a way that the sum of products of numbers multiplied by average weights adds up to some input value of the weight landed.

3. Various options for grouping large individuals that may or may not be larger than the input value of L_∞ .

This version handles one length composition only, and test data and worked numerical examples suitable for testing the program can be found in Jones (1984). Selectivity curves are calculated using a logistic function.

Listing

A complete listing of the program is given in Fig. 23.

Fig. 23. Listing of RJ33 program.

```

FILE NAME --- RJ33.BAS          PRINTED AT 10:21      ON 25-Sep-86

10 ! ****
20 ! LENGTH COMPOSITION ANALYSIS-CALCULATES LONG TERM EFFECTS OF
30 ! CHANGES IN FISHING EFFORT AND MESH SIZE FROM LENGTH COMP.DATA
35 ! THIS VERSION HANDLES ONE LENGTH COMPOSITION ONLY
40 ! INCLUDING INTERMEDIATE EFFECTS AND ALLOWANCE FOR DISCARDING
50 ! COHORT ANALYSIS TABLE INCLUDES VALUES OF Z/K BY BEV.AND HOLT METHOD.
60 ! ****
70 EXTEND
80 INPUT"OUTPUT TO FILE(F) OR KB(K)";X$
90 IF X$="K" THEN F$="KB:" \ GOTO 110
100 INPUT"ENTER OUTPUT FILE NAME";F$
110 OPEN F$ FOR OUTPUT AS FILE #1%
120 PRINT#1% "RJ31"
130 PRINT#1%"LENGTH COMPOSITION ANALYSIS"
140 PRINT#1%"ALLOWING FOR INTERMEDIATE EFFECTS"
150 PRINT"-----"
160 DIM W(50),H(50)
170 DIM P(50),R(50)
180 DIM S(50%),G(50%),S5(50%),S6(50%)
190 DIM V4(50),V5(50),E3(50)
200 DIM DT(50),F2DT(50)
210 DIM P2(50%),T2(50%),LL(50%)
220 DIM RET(50%),S88(50%),S99(50%)
230 Z1$='***   ***   *****  .**  .***  .***  **.***  **.***  .***  **
240 Z2$='***   ***   *****  .***  .***  .***  ***.***  ***.***  ***.***  *
250 Z3$='   *****.**
260 Z4$=*****.**
270 Z5$='*****.**'
280 Z6$='*****.**'
290 Z7$='*****.**'
295 Z8$='*****.**'
300 P1$='CM    N.LANDED BIOMASS  DT    F1DT   F2DT   SF1DT   SF2DT  S2/S1  C.
     NEW BIOMASS NEW CATCH N.LANDS WT.LANDED AV.WT'
310 P2$=' CM      REMOVALS    NUMBERS   E/Z      FDT      DT      F
/S1      Z      NUMBERS      BIOMASS      Z/K'
320 P3$='          ATTAINING
IN SEA'
330 P4$='          LENGTHS'
340 P5$='-----'
-----'
```

Fig. 23. Continued

```

350 PG$=' CM      N.LANDED      AV.WTS      TOTAL WTS      PROPS.RET
      N.CAUGHT      N.REMOVED FROM SEA (NET)'
355 P7$='      CM      N.LANDED      N.CAUGHT      NEW LANDINGS BY NO.
W LANDINGS BY WEIGHT '
357 P8$='          A          B
      A          B
360 ! ****
370 ! INPUT NOTES
380 ! LINES 510 AND 520 ARE FILE HEADINGS
400 ! N1% = NO. OF LENGTH GROUPINGS
410 ! S(L%) = NOS. PER LENGTH GROUP -OF LANDINGS NOT CATCHES
420 ! RET(L%) = PROPORTIONS RETAINED ON BOARD PER LENGTH GROUP (NO ZEROS ALLOWED)
430 ! BB=WEIGHT LANDED
440 ! CC= ORIGINAL MESH SIZE
450 ! A1=SMALLEST LENGTH FOR SMALLEST LENGTH GROUP
460 ! G1=GROUPING INTERVAL
470 ! A2 AND A3 ARE WT./LENGTH PARAMETERS
480 ! i.e. WT.= A3\LENGTH\*A2
490 ! PROGRAM ASSUMES LENGTHS IN CM. AND MESH SIZES IN MM.
500 ! IF LENGTHS ARE IN MM OMIT '10' FROM LINES 2070, 2120 AND 2776
510 ! ****
520 INPUT "NAME OF INPUT FILE"; B$
530 OPEN B$ FOR INPUT AS FILE 2
540 INPUT #2, H1$
550 INPUT #2, H2$
557 INPUT #2, N1%
580 INPUT #2, S(L%) FOR L% = 1% TO N1%
590 INPUT #2, RET(L%) FOR L% = 1% TO N1%
600 INPUT #2, BB
610 INPUT #2, CC
620 INPUT #2, A1 \ INPUT #2, G1 \ INPUT #2, A2 \ INPUT #2, A3
630 CLOSE 2
640 PRINT #1% H1$
650 PRINT #1% H2$
660 N1=N1% \ M4=N1%+1%
680 INPUT "DO YOU WANT TO ALLOW FOR DISCARDS (YES=1, NO=0)"; J5
690 IF J5=1 GOTO 710
700 RET(L%)=1. FOR L% = 1% TO N1%
710 INPUT "DO YOU WANT SUM OF PRODUCTS TO BE COMPATIBLE WITH INPUT WEIGHT"; J7
715 PROP=1 \ IF J5=0 THEN 740
720 INPUT "WHAT PROPORTION OF DISCARDS SURVIVES", PROP
740 !
745 SP=0 \ SPR=0
750 ! ****
760 ! OPTION TO REVISE INPUT NOS. ACCORDING TO SUM OF PRODUCTS
770 W(L%)=A3*(A1+G1*(L%-0.5))\*A2 FOR L% = 1% TO N1%
780 U3=0
800 T1=0
810 T1=T1+W(L%)*S(L%) FOR L% = 1% TO N1%
820 IF J7=0 THEN T1=1 ELSE T1=BB/T1
830 BB=0
840 VV=0
850 S(L%)=T1\*S(L%) FOR L% = 1% TO N1%
860 VV=VV+S(L%) FOR L% = 1% TO N1%
870 BB=BB+W(L%)\*S(L%) FOR L% = 1% TO N1%
880 U3=U3+VV
900 A4=0
910 FOR L1=1 TO N1
930 P(L1)=S(L1)
940 A4=A4+P(L1)\*W(L1)

```

Fig. 23. Continued

```

950 P(L1)=P(L1)/RET(L1)
960 NEXT L1
970 PRINT#1%,P6$
980 FOR L1=1 TO N1
990 SG(L1)=S(L1)/RET(L1)
1000 C1=A1+G1*(L1-1)
1010 S5(L1)=SG(L1)*(1-PROPA*(1-RET(L1)))
1030 PRINT#1%,USING Z4$,C1;S(L1);W(L1);S(L1)*W(L1);RET(L1);SG(L1);S5(L1)
1050 NEXT L1
1060 LL(LZ)=S(LZ) FOR LZ=1% TO N1%
1070 PRINT#1% VV;
1080 PRINT#1%
1090 PRINT#1%"TOTAL WT.      ";
1100 PRINT#1% BB;
1110 PRINT#1%"LANDED"
1120 PRINT#1%"MESH SIZE      "; \ PRINT#1% CC; \ PRINT#1%
1140 PRINT#1%"PROP. OF DISCARDS THAT SURVIVE =";PROP
1150 FOR LZ=1% TO N1%
1160 P(LZ)=0
1170 P(LZ)=P(LZ)+S5(LZ)
1180 NEXT LZ
1190 N9%,N8%=0
1200 FOR LZ=1% TO N1%
1210 L5Z=N1%-LZ+1%
1220 IF P(L5Z)>0 THEN N9%=1%
1230 IF N9%<=0 THEN N8%=N8%+1%
1240 NEXT LZ
1250 N1%=N1%-N8% \ N1=N1%
1260 ! *****INPUT OPTIONS*****
1270 ! INPUT OPTIONS
1273 INPUT'DO YOU WANT TO GROUP LARGE INDIVIDUALS ',JJ8
1274 IF JJ8=0 THEN 1280
1275 INPUT'LARGEST LENGTH GROUP REQUIRED ',XXX
1276 GOSUB 4320
1280 INPUT'WHAT IS LINEF',Y1
1290 C1=A1+G1*N1
1300 IF Y1>C1 THEN 1330
1310 PRINT'LINE IS NOT LARGER THAN LARGEST LENGTH'
1320 GOSUB 4290
1330 PRINT'CHOOSE FROM THE FOLLOWING OPTIONS'
1340 PRINT'1.TO INPUT M/K AS A SINGLE VALUE'
1350 PRINT'2.TO CALCULATE K FROM AN ANNUAL GROWTH INCREMENT'
1360 PRINT'3.TO CALC. K FROM CONVENTIONAL LINE & K VALUES'
1370 PRINT'4.TO INPUT A VALUE OF K'
1380 INPUT'WHAT IS YOUR CHOICE',J5
1390 IF J5=1 GOTO 1620
1400 INPUT'WHAT IS M',M5
1410 IF J5<>2 GOTO 1480
1420 PRINT'INPUT 2 LENGTHS SEPARATED BY A YEAR'S GROWTH'
1430 INPUT'FIRST LENGTH',L6
1440 INPUT'SECOND LENGTH',L7
1450 K1=(Y1-L6)/(Y1-L7)
1460 K1=LOG(K1)
1470 GOTO 1590
1480 IF J5<>3 GOTO 1580
1490 PRINT'INPUT CONVENTIONAL VALUES OF LINE & K'
1500 INPUT'WHAT IS LINEF',G2
1510 INPUT'WHAT IS K',G3
1520 PRINT'AT WHAT LENGTH DO YOU WANT THE 2 GROWTH CURVES'
1530 INPUT'TO COINCIDE',LG

```

Fig. 23. Continued

```

1540 L7=Y1-G2+(G2-L6)*EXP(-G3)
1550 L7=L7/(Y1-L6)
1560 K1=-LOG(L7)
1570 GOTO 1590
1580 INPUT "WHAT IS K",K1
1590 X1=M5/K1
1600 PRINT "K =";K1      M/K =";X1
1610 GOTO 1630
1620 INPUT "WHAT IS M/K",X1
1624 PRINT "INPUT A VALUE OF M FOR ADDITIONAL COMPUTATIONS"
1626 INPUT M5
1630 INPUT "WHAT IS FINAL E/Z",E1
1632 A89=0 \ M2=0 \ A5=0 \ L33=S22
1640 PRINT "ENTER MINIMUM LANDING SIZE FOR NEW MESH"
1642 INPUT "USE DUMMY VALUE IF NOT REQUIRED";MLS
1650 INPUT "WHAT IS SELECTION FACTOR(LENGTHS IN CM. MESH SIZE IN MM.)",S1
1660 PRINT "WHAT IS SLOPE OF SELECTION CURVE FOR ORIGINAL MESH SIZE"
1670 PRINT "(MESH SIZE IN MM.)"
1680 INPUT S22
1700 INPUT "DO YOU WANT COHORT ANALYSIS ONLY(YES=1,0=NO)",I5
1750 IF I5=1 THEN A89=1 \ GOTO 1980
1760 INPUT "HOW MANY INTERMEDIATE YEARS DO YOU WANT";NINT
1800 INPUT "IS THERE A CHANGE IN MESH SIZE";M2
1810 IF M2<=0 GOTO 1910
1820 PRINT "WHAT IS NEW MESH SIZE ?"
1840 INPUT X22
1860 PRINT "WHAT IS SLOPE OF SELECTION CURVE FOR THIS MESH SIZE ?"
1880 INPUT L33
1900 K1=M5/X1
1910 INPUT "WHAT IS PERCENTAGE CHANGE IN EFFORT",AS
1920 A5=A5/100%
1930 PRINT "CHOOSE FROM THE FOLLOWING OPTIONS"
1932 PRINT "0=MINIMUM OUTPUT"
1933 PRINT "1=MAXIMUM OUTPUT"
1938 INPUT "WHAT IS YOUR CHOICE",I5
1940 ! ***** PROPORTIONS RETAINED ,S6 AND S7. ALSO RATIOS S6/S7=G(M%,L%)
1950 ! PROPORTIONS RETAINED ,S6 AND S7. ALSO RATIOS S6/S7=G(M%,L%)
1980 PRINT #1%
1990 PRINT #1% P5$          SELECTION CURVES*
2000 PRINT #1%              'MESH SIZE    ';CC;'     'X22 \ GOTO 2020
2010 PRINT #1%, 'MESH SIZE    ';CC
2012 PRINT #1%, 'MESH SIZE    ';CC
2020 FOR LZ=1% TO N1%
2030 LBAR=A1+G1*(LZ-0.5)
2050 PRINT #1%, USING "###.###",A1+G1*(LZ-1%);
2070 L50=S1*X22/10%
2080 SLP=L33
2090 GOSUB 4590
2100 SG=PR
2110 SLP=S22
2120 L50=S1*CC/10%
2130 GOSUB 4590
2140 S7=PR
2150 IF S7>0 THEN G(L%)=S6/S7 ELSE G(L%)=1.0
2160 IF M2=0 THEN 2172
2170 PRINT #1% USING "###.### ##.###",S7,S6; \ GOTO 2190
2172 PRINT #1% USING "###.###",S7;

```

Fig. 23. Continued

```

2174 G(LZ)=1.0
2190 PRINT#1%
2200 NEXT LZ
2240 !
2250 ! ****
2260 ! LENGTH COHORT ANALYSIS
2270 IF IS<>1 GOTO 2340
2280 PRINT#1% PS$
2290 PRINT#1%"COHORT ANALYSIS"
2295 PRINT#1%
2300 PRINT#1%"LINEF = "Y1;"M/K = "X1;"FINAL F/Z = "E1;"M = "M5
2305 PRINT#1%"SELECTION FACTOR=";S1;"SLOPE AT 50% POINT=";S22;"MESH SIZE=";CC
2307 PRINT#1%
2310 PRINT#1% P2$
2320 PRINT#1% P3$
2330 PRINT#1% P4$
2340 X3=P(N1)/E1
2350 C2=A1+N1*G1
2360 C1=C2-G1
2365 SUMP=SP \ SOPPL=SPR
2370 IF JJ8=1 THEN 2400
2380 SUMP=SUMP+P(N1)
2390 SOPPL=SOPPL+0.5*P(N1)*(C1+C2)
2400 IF IS<>1 GOTO 2420
2410 PRINT#1% USING Z2$,C1;P(N1),X3
2420 S4=P(N1)
2430 S5=0
2440 V4=0
2450 Z2=M5/(1-E1)
2460 F2=Z2-M5
2470 E1=0
2480 F3=3
2490 FOR I1=2 TO N1
2500 J1=N1-I1+1
2510 J2=J1+1
2520 C2=C2-G1
2530 C1=C1-G1
2540 A6=Y1-C1
2550 A7=Y1-C2
2560 B1=A6/A7
2570 D1=X1/2
2580 B1=B1**D1
2590 X4=X3
2600 X3=B1*(X4*B1+P(J1))
2610 S4=S4+P(J1)
2620 IF T1>=C2 GOTO 2640
2630 Q9=S4/X3
2640 E2=P(J1)/(X3-X4)
2650 H1=X3/X4
2660 Z1=LOG(H1)
2670 R(J1)=E2**Z1
2680 Z2=M5/(1-E2)
2690 F2=Z2-M5
2700 DT(J1)=Z1/Z2
2710 SUMP=SUMP+P(J1)
2720 SOPPL=SOPPL+0.5*P(J1)*(C1+C2)
2730 BVBAR=SOPPL/SUMP
2740 BEV=(Y1-BVBAR)/(BVBAR-C1)
2750 V4(J1)=(X3-X4)/Z2
2760 V4(J1)=V4(J1)*W(J1)

```

Fig. 23. Continued

```

2770 V4=V4+V4(J1)
2772 LBAR=C1+0.5*G1
2774 SLP=S22
2776 L50=S1*CC/10%
2778 GOSUB 4590
2780 IF I5<>1 GOTO 2800
2790 PRINT#1% USING Z2$,C1;P(J1);X3;E2;R(J1);DT(J1);F2;F2/PR;Z2;V4(J1)/W(J1);V
J1>;BEV
2800 NEXT I1
2810 DT(N1)=DT(N1-1)
2840 C1=A1
2850 M6=N1-1
2860 V4(N1)=V4(M6)
2870 V4=V4+V4(M6)
2871 IF I5<>1 THEN 2880
2872 PRINT#1%"      Z/K IN FINAL COL.CALCULATED BY BEV. AND HOLT METHOD"
2873 PRINT#1%"      S1 IN COL. F/S1 REFERS TO PROPS RETAINED IN SELECTION CURVE T
LE"
2874 PRINT#1%
2880 IF AB9=1 GOTO 4170
2890 !
2900 ! *****NEW VALUES OF FDT, H(L%)*****
2910 ! NEW VALUES OF FDT, H(L%)
2920 C1=A1
2930 NG=N1-1
2940 FOR L1=1 TO NG
2950 C2=C1+G1
2960 H(L1)=0
2980 IF P(L1)=0 THEN X=0\GOTO 3000
2990 X=R(L1)*S5(L1)*G(L1)*(1+A5)/P(L1)
3000 H(L1)=H(L1)+X
3020 C1=C2
3030 NEXT L1
3040 C2=C1+G1
3050 H(N1)=H(NG)
3060 !
3070 ! *****NEW VALUES OF FDT FOR CALCULATING INTERMEDIATE EFFECTS H(L%)*****
3080 ! NEW VALUES OF FDT FOR CALCULATING INTERMEDIATE EFFECTS H(L%)
3090 F2DT(L1%)=H(L1%) FOR L1%=1% TO N1%
3100 FOR YOPT=1% TO NINT+
3110 IF YOPT=NINT+1 GOTO 3310
3120 FOR I=1% TO N1%
3130 J1=N1%+1%-I
3140 X=0.5*DT(J1)
3150 H(J1)=F2DT(J1)
3160 IF J1=1% GOTO 3280
3170 J2=J1
3180 J2=J2-1%
3190 X=X+DT(J2)
3200 IF X>YOPT GOTO 3230
3210 H(J2)=F2DT(J2) \ IF J2=1% GOTO 3280
3220 GOTO 3180
3230 X=X-DT(J2)
3240 X=X/DT(J2)
3250 H(J2)=X*R(J2)+(1-X)*F2DT(J2)
3260 IF J2=1 GOTO 3280
3270 H(K%)=R(K%) FOR K%=1% TO J2-1%
3280 GOSUB 4720
3290 NEXT I

```

Fig. 23. Continued

```

3300 IF YOPT<NINT+1 GOTO 3380
3310 H(L1%)=F2DT(L1%) FOR L1%=1% TO N1%
3320 J1=N1
3330 GOSUB 4720
3340 ! ****
3350 ! DETERMINATION OF NEW LANDINGS
3360 ! BY WT. TW AND B2
3370 ! BY NO. U2
3380 IF I5<>1 GOTO 3410
3400 PRINT#1% P1$
3410 UU=0 \ XX=0
3420 TW=0
3430 FOR L1=1% TO N1%
3440 C2=C1+G1
3450 L2=L1+1
3460 B3=W(L1)*F3(L1)
3470 B4=0
3480 U1=0
3490 V1=V4(L1)*F3(L1)
3500 V5(L1)=V1
3520 B5=B3*G(L1)*S5(L1)*(1+A5)
3530 B5=B5*RET(L1)/(1-PROP*(1-RET(L1)))
3540 B6=B5/W(L1)
3550 CG=G(L1)*LL(L1)*W(L1)*(1+A5)
3560 U1=U1+B6
3570 UU=UU+CG
3580 TW=TW+B5
3590 XX=XX+S5(L1)*G(L1)*RET(L1)
3600 B4=B4+B5
3610 V2=V2+V1
3630 P2(L1)=U1
3640 IF I5<>1 GOTO 3660
3650 PRINT#1% USING Z1$,C1;LL(L1);V4(L1);BT(L1);R(L1);H(L1);S88(L1);S99(L1);G(I
 );F3(L1);V5(L1);P2(L1)/RET(L1);P2(L1);P2(L1)*W(L1);W(L1)
3660 B2=B2+B4
3670 U2=U2+U1
3680 V3=V3+V1
3690 C1=C2
3700 NEXT L1
3710 IF I5<>1 GOTO 3740
3720 PRINT #1% P5$
3730 PRINT#1% USING Z3$,V4,V3,U2,B2
3731 PRINT#1%
3732 PRINT#1% "SEE TAB. 7, JONES 1984 FAO FISH. TECH. PAP. 256 FOR COL. HEADINGS"
3733 PRINT#1% P5$
3734 PRINT#1%
3740 IF YOPT=NINT+1 GOTO 3770
3750 PRINT#1%"NEW LANDINGS AFTER",YOPT,"YEARS"
3760 GOTO 3790
3770 PRINT#1% P5$
3771 PRINT#1%"NEW LONG TERM LANDINGS"
3780 PRINT#1%-----
3790 IF M2=0 GOTO 3840
3800 PRINT#1%"SEL. FACTOR",S1"PERC. CHANGE IN EFFORT",100*A5
3810 PRINT#1%"OLD MESH",CC,"SLOPE",S22
3820 PRINT#1%"NEW MESH",X22,"SLOPE",L33
3830 GOTO 3850
3840 PRINT#1%"PERC. CHANGE IN EFFORT",100*A5
3845 PRINT#1%

```

Fig. 23. Continued

```

3850 PRINT#1%"PERC. CHANGES IN LANDINGS"
3870 PRINT#1%" BY WEIGHTS BY NUMBERS"
3880 B2=100*(B2-A4)/A4
3890 U4=100*(U2-U3)/U3
3900 V3=100*(V3-V4)/V4
3920 TW=100*(TW-BB)/BB
3950 PRINT#1%" B2,U4
3960 PRINT#1%"PERCENTAGE CHANGE IN BIOMASS =",V3
3975 PRINT#1%,P5$
3977 NEXT YOPT
3980 PRINT#1%,'IMMEDIATE EFFECTS'
3981 PRINT#1%,'-----'
3982 PRINT#1%,"A ASSUMING NO CHANGE IN PROPORTIONS RETAINED."
3984 PRINT#1%,"B ASSUMING PROPORTION RETAINED=0 BELOW MINIMUM LANDING SIZE A
1.0 ABOVE MINIMUM LANDING SIZE"
3988 IF M2=0 THEN 4170
4000 YY=100*(BB-UU)/BB
4030 C1=A1
4040 SUM=0\SUMW=0\SX3=0\SX4=0
4044 PRINT#1%,P7$
4046 PRINT#1%,P8$
4050 FOR L1=1% TO N1%
4060 C2=C1+G1
4070 IF C1<MLS THEN X=0 ELSE X=1.0
4080 X11=G(L1)*S6(L1)*RET(L1)*(1+A5)
4090 X2=X11*W(L1)
4092 X3=G(L1)*X*X*S6(L1)*(1+A5)
4094 X4=X3*W(L1)
4100 PRINT#1%,USING Z5$,C1;S(L1);S6(L1);X11;X3;X2;X4
4110 SUM=SUM+X11
4120 SUMW=SUMW+X2
4122 SX3=SX3+X3
4124 SX4=SX4+X4
4130 C1=C2
4140 NEXT L1
4145 PRINT#1%,P5$
4147 PRINT#1%,'TOTAL';
4150 PRINT#1%,USING Z6$,SUM;SX3;SUMW;SX4
4152 X4=100*(BB-SX4)/BB
4154 PRINT#1%"IMMEDIATE LOSSES FROM CHANGES IN MESH SIZE"
4155 PRINT #1%,"MINIMUM LANDING SIZE =",MLS
4156 PRINT#1%,"A ";YY
4158 PRINT#1%,"B ";X4
4170 PRINT'CHOOSE FROM THE FOLLOWING OPTIONS'
4180 PRINT'1.TO VARY EXPLOITATION PATTERN ONLY'
4190 PRINT'2.TO VARY GROWTH PARAMETERS'
4200 PRINT'3.RETURN TO BEGINNING OF INPUT OPTIONS'
4205 PRINT'4.TO VARY DETAIL OF OUTPUT ONLY'
4210 PRINT'54.STOP PROGRAM'
4220 INPUT"What IS YOUR CHOICE",J7
4230 IF J7=1 THEN 1632
4235 IF J7=2 THEN 1280
4237 IF J7=4 THEN 1930
4240 IF J7=5 GOTO 4990
4250 GOTO 1280
4260 ! ****
4270 ! ENSURES THAT LAST LENGTH INTERVAL IS LESS THAN 0.9*LINEFINITY
4280 ! ****
*****
```

Fig. 23. Continued

```

4290 PRINT"DO YOU WANT PROGRAM TO ENSURE THAT LARGEST GROUP DOES NOT EXCEED 0.
LINEINITY(YES=1,NO=0)"
4300 INPUT JJ8
4310 IF JJ8=1 THEN XXX=0.9*Y1
4314 GOTO 4320
4316 RETURN
4320 C1=A1 \ TOT=0 \ NG=N1%
4325 XXX=XXX+4
4330 TOS=0
4340 TOS5=0
4350 TOS6=0
4360 FOR L1=1 TO NG
4370 C2=C1+G1
4380 IF C2<XXX THEN 4460
4390 TOT=TOT+P(L1)
4400 TOS=TOS+S(L1)
4410 TOS5=TOS5+S5(L1)
4420 TOS6=TOS6+S6(L1)
4430 SOPPL=SOPPL+0.5*P(L1)*(C1+C2)
4440 SUMP=SUMP+P(L1)
4450 N1%=N1%-1%
4460 C1=C2
4470 NEXT L1
4480 N1=N1%
4490 SUMP=SUMP+P(N1%)
4500 SOPPL=SOPPL+P(N1%)*((C1-(0.5*G1))-((L1-N1%)*G1))
4505 SP=SUMP \ SPR=SOPPL
4510 P(N1%)=TOT+P(N1%)
4520 S5(N1%)=TOS5+S5(N1%)
4530 S(N1%)=TOS+S(N1%)
4540 SG(N1%)=TOS6+SG(N1%)
4550 LL(N1%)=S(N1%)
4560 LBAR=SOPPL/SUMP
4570 W(N1%)=A3\LBAR**A2
4580 RETURN
4590 !
4600 ! ****
4610 ! PROPORTIONS RETAINED BY CODEND (PR) USING LOGISTIC SELECTION CURVE
4620 R1= SLP*(LBAR-L50)/0.2276
4630 IF R1>50 GOTO 4680
4640 IF R1<-30 GOTO 4670
4650 PR=EXP(R1)/(1+EXP(R1))
4660 GOTO 4690
4670 PR=0 \ GOTO 4690
4680 PR=1.0
4690 RETURN
4700 ! ****
4710 ! CONVERSION FACTORS F3(LX)
4720 R(N1)=R(NG)
4730 S8=0
4740 S9=0
4750 N7=N1+1
4760 R(N7)=R(N1)
4770 B2=0
4780 U2=0
4790 V3=0
4800 C1=A1
4810 X5=0
4820 X9=0
4830 XG=R(1)

```

Fig. 23. Continued

```
4840 X8=H(1)
4850 FOR L1=1 TO J1
4860 C2=C1+G1
4870 L2=L1+1
4880 E4=0.5*(X5+X6)
4890 S8=S8+E4
4900 E5=0.5*(X8+X9)
4910 S9=S9+E5
4920 R2=S8-S9
4930 S88(L1)=S8
4940 S99(L1)=S9
4950 F3(L1)=EXP(R2)
4960 X5=R(L1) \ X6=R(L2) \ X9=H(L1) \ X8=H(L2)
4970 NEXT L1
4980 RETURN
4990 END
```

Literature Cited

Jones, R. 1984. Assessing the effects of change in exploitation pattern using length composition data. FAO Fish. Tech. Pap. 256, 118 p.

Program Number 13

Title	LFSIM
Author	J. Hampton Division of Fisheries Research Commonwealth Scientific and Industrial Research Organisation GPO Box 1538 Hobart, Tasmania 7001 Australia

Reference Hampton, J. and J. Majkowski. — A simulation model for generating catch length-frequency data (Part I, this vol.).

Program Description Program LFSIM generates catch length-frequency data for an exploited fish population. Presently, 8 year classes are simulated which produce 48 monthly length frequency samples each containing 5 age classes. These specifications can be easily altered by the user.

OUTPUT: Two data files are produced.

- (i) COHORT.DAT — contains LF data for separate year classes.
- (ii) LFDATA.DAT — contains LF data for 48 monthly samples. These data represent catch-at-length data and are formatted for direct input, e.g., to the ELEFAN I and II programs.

Modules

FUNCTION GAUSS

This function generates random normal variates.

Arguments : a — mean
 b — standard deviation
Variables : a, b, c, ix

Associated modules : RAN (intrinsic FORTRAN function)

FUNCTION GROWTH

This function determines the length of a fish at a given age using the seasonally oscillating VBGF of Pauly and Gaschütz (1979)

Arguments : age — age of fish in years
 t_s — age at the onset of the first seasonal oscillation
 L_∞ — asymptotic length for this fish
 K — growth constant for this fish
Variables : age, C, D, K, L_∞ , t_o , t_s , twopi

SUBROUTINE SELECT

This module is the "workhorse" of the program. It increments the two LF matrixes LFREQ and LF. Within this module, the following sequence of operations is undertaken:

- (i) Determine a time of encounter/natural death, x.
- (ii) If $x > t$ (see master list), disregard this fish and return to (i).
- (iii) Decide whether this fish died due to natural causes or was encountered by the fishing gear. If the former, return to (i).
- (iv) Determine the age at the time of recruitment.
- (v) Determine the age at encounter.
- (vi) Determine the growth parameters l_{\max} , k and t_s .
- (vii) Determine the length at encounter.
- (viii) Determine the probability of retention by the gear.
- (ix) Decide if the fish is caught or escapes. If caught, store data.
- (x) Determine time of next encounter/death. Repeat steps (ii), (iii), (v), and (vii) to (x) until the fish is caught, dies naturally or lives beyond the maximum time t.

Arguments:	:	ix0	— year class strength
		am1	— mean age (yrs) at recruitment for fish spawned from major peak
		am2	— mean age (yrs) at recruitment for fish spawned from minor peak
		asd1	— SD (yrs) of major peak
		asd2	— SD (yrs) of minor peak
		p	— proportion of total recruits contributed by major peak
		a and b	— parameters of the selection curve
		f	— instantaneous rate of encounter
		m	— instantaneous rate of natural mortality
		t	— maximum age of fish contributing to the fishery
		i	— number of year class
		islec	— flag: 0 = no selection; 1 = selection
		nc	— number caught from this cohort (output)
Variables	:	a, age, alpha, am1, am2, asd1, asd2, b, f, imth, iselec, iset, ix0, ix, k, km, ksd, kx, len, len2, lf, lfreq, lmax, lmaxm, lmaxsd, m, nc, ntp, pc, pctest, pm, pmtest, t, ts, x, y	

Associated modules : GAUSS, GROWTH, RAN, SPAWN

FUNCTION SPAWN

This function determines the age of individual fish at recruitment ($t = 0$). Presently, all fish should have a positive age at $t = 0$; however, this could easily be changed if code were added to ensure that fish were not encountered or did not die with a negative age. Initially, a decision is reached using RAN as to which spawning distribution (cohort) the fish comes from. Age is then randomly sampled from this normal distribution.

Arguments : am1, am2, asd1, asd2, p — see above for definitions

Associated modules : GAUSS

Listing Notes

- | | |
|--------------------------|--|
| a. Language | FORTRAN 77 |
| b. Implementation | Currently on VAX but should run on any DEC machine |
| c. Hardware Requirements | Printer, Plotter |
| d. RAM Requirements | About 84 K |

Listing

A complete listing of this program is given in Fig. 24.

Fig. 24. Listing of LFSIM program.

```

c      * LIST OF MAJOR VARIABLES (not including indexes) *
c      * -----
c      * a (SELECT) - parameter of the selection curve      *
c      * a (GAUSS) - mean of a normal distribution      *
c      * age - age at encounter of an individual fish      *
c      * alpha - random number for survivorship function (SELECT) *
c      * am1 - mean age at recruitment of fish from major spawning peak *
c      * am2 - mean age at recruitment of fish from minor spawning peak *
c      * asd1 - standard deviation of major spawning peak      *
c      * asd2 - standard deviation of minor spawning peak      *
c      * b (SELECT) - parameter of the selection curve      *
c      * b (GAUSS) - standard deviation of a normal distribution      *
c      * c (GROWTH) - coefficient of seasonal growth      *
c      * c (GAUSS) - sum of 12 random numbers      *
c      * d - growth parameter (set to 1.0)      *
c      * f - instantaneous rate of encounter by the fishing gear      *
c      * iaget - total catch by age class for a single year class      *
c      * imth - month number for LF      *
c      * iselec - flag to indicate whether selection operates      *
c      * iset - flag to indicate whether a fish has been previously      *
c          encountered      *
c      * ix - seed to RAN (random number generator)      *
c      * ix0 - number of recruits for a year class      *
c      * k - growth constant for a fish      *
c      * km - mean value of growth constant      *
c      * ksd - standard deviation of growth constant      *
c      * len - length of an encountered fish      *
c      * len2 - length class number for monthly LF      *
c      * lf - catch number by length class by month of capture      *
c      * lfreq - catch number by length class by month for a year class      *
c      * lmax - asymptotic length for a fish      *
c      * lmaxm - mean asymptotic length      *
c      * lmaxsd - standard deviation of asymptotic length      *
c      * ltot - total catch by month for a year class      *
c      * m - the instantaneous rate of natural mortality      *
c      * nc - total catch from a year class      *
c      * ntp - month number from recruitment for a year class      *
c      * nyc - number of year classes simulated      *
c      * p - proportion of total recruits contributed by major spawning peak *
c      * pc - probability of retention by the fishing gear      *
c      * pctest - random number to determine if fish is retained      *
c      * pm - probability of a given death being from natural causes      *
c      * pmtest - random number to determine if fish dies from natural causes *
c      * q - random number to determine to which spawning peak a fish belongs *
c      * rm - mean year class strength      *
c      * rsd - standard deviation of year class strength      *
c      * t - maximum age of fish contributing to the fishery      *
c      * t0 - growth parameter      *
c      * ts - age of fish at onset of the first seasonal growth oscillation *

```

Fig. 24. Continued

```

c      * twopi - constant
c      * x - time from recruitment to encounter/death for a fish
c      * x0 - number of recruits for a year class
c      * y - the age of a fish at recruitment (t=0)
c      *
c      ****
c
c      program lfsim
real km,lmaxm,lmaxsd,ksd,m
integer ltot(60),iaget(5)
common/block/ d,t0,c,lfreq(60,60),lf(48,60),lmaxm,km,lmaxsd,ksd
common/jjj/ix
open(unit=1,name='lfdata.dat',type='new',access='sequential',
$ recordsize=80)
open(unit=2,name='cohort.dat',type='new',access='sequential',
$ recordsize=120)
ix=134567
c
c      ***** input parameters for spawning activity *****
c
am1=0.5
am2=0.75
asd1=0.083
asd2=0.083
p=0.75
c
c      ***** input parameters for year class strength determination *****
c
rm=4000.
rsd=1000.
c
c      ***** input parameters for growth *****
c
lmaxm=40.
lmaxsd=4.
km=0.5
ksd=0.05
t0=0.0
c=0.4
d=1.0
c
c      ***** population and fishery parameters *****
c
m=0.5
f=1.0
t=5.0
nyc=8
c
c      ***** parameters for gear selectivity (logistic curve) *****
c
a=-10.0
b=0.666667
iselect=1
c
c
do 100 i=1,nyc
do 20 kx=1,60
ltot(kx)=0
do 10 jx=1,60
10 lfreq(kx,jx)=0
20 continue
do 30 kx=1,5
30 iaget(kx)=0
c
c      ***** year class strength determination *****

```

Continued

Fig. 24. Continued

```

c
x0=gauss(rm,rsd)
ix0=x0+0.5

c
c      ***** calculation of length and age at capture for all fish *****
c      in the current year class
c
call select(ix0,am1,am2,asd1,asd2,p,a,b,f,m,t,i,iselect,nc,ix)
write(2,1002)i,x0,nc
write(2,1003)(ntp,ntp=1,24)
do 60 len=1,60
do 55 ntp=1,24
ltot(ntp)=ltot(ntp)+lfreq(ntp,len)
55 continue
write(2,1004)len,(lfreq(ntp,len),ntp=1,24)
60 continue
write(2,1005)(ltot(ntp),ntp=1,24)
write(2,1003)(ntp,ntp=25,48)
do 70 len=1,60
do 75 ntp=25,48
ltot(ntp)=ltot(ntp)+lfreq(ntp,len)
75 continue
write(2,1004)len,(lfreq(ntp,len),ntp=25,48)
70 continue
write(2,1005)(ltot(ntp),ntp=25,48)
write(2,1003)(ntp,ntp=49,60)
do 80 len=1,60
do 85 ntp=49,60
ltot(ntp)=ltot(ntp)+lfreq(ntp,len)
85 continue
write(2,1004)len,(lfreq(ntp,len),ntp=49,60)
80 continue
write(2,1005)(ltot(ntp),ntp=49,60)
do 95 iytot=1,5
m1=(iytot-1)*12+1
m2=m1+11
do 90 ntp=m1,m2
iaget(iytot)=iaget(iytot)+ltot(ntp)
90 continue
95 continue
write(2,1007)(kx,iaget(kx),kx=1,5)
100 continue
c
c      ***** the following code compiles if data in months *****
c
write(1,1009)
mmax=48
do 120 imth=1,mmax
mult=imth/12
irem=imth-mult*12
if(irem.eq.0)irem=12
do 110 len2=1,30
xm1=len2*2-0.5
if(lf(imth,len2).eq.0)go to 110
write(1,1006)imth,xm1,lf(imth,len2),irem
110 continue
120 continue
write(1,1008)
1009 format(' 48    19    2    7.5   43.5')
1008 format(' 9999 0 0 99 99')
1002 format('1year class number',i3,' recruitment',f7.0,
$ ' number caught',i5)
1003 format(55x,'time period',/, ' length   '12i4,4x,12i4/)

```

Continued

Fig. 24. Continued

```

1004  format(i7,3x,12i4,4x,12i4)
1005  format(1x,109('')/-10x,12i4,4x,12i4//)
1006  format(1x,i2,f6.1,i6,4x,i2,4x,'15')
1007  format(///28x,'catch by age class',//1x,5('year',i2,i6,3x))
      stop
      end
c
c      ***** FUNCTION SPAWN *****
c
      function spawn(am1,am2,asd1,asd2,p)
common/jjj/ix
q=ran(ix)
if(q.le.p)spawn=gauss(am1,asd1)
if(q.gt.p)spawn=gauss(am2,asd2)
return
end
c
c      ***** FUNCTION GROWTH *****
c
      function growth(age,ts,lmax,k)
real lmax,k
common/block/d,t0,c
twopi=6.283185
growth=lmax*(1-exp(-(k*d*(age-t0)+c*k*d/twopi*sin(twopi*(age-ts)))
$ ))**(1/d)
return
end
c
c      ***** SUBROUTINE SELECT *****
c
      subroutine select(ix0,am1,am2,asd1,asd2,p,a,b,f,m,t,i,iselect,nc)
real m,lmaxm,km,lmaxsd,ksd,lmax,k
common/block/d,t0,c,lfreq(60,60),lf(48,60),lmaxm,km,lmaxsd,ksd
common/jjj/ix
j=0
nc=0
do 20 kx=1,ix0
iset=0
x=0.
y=spawn(am1,am2,asd1,asd2,p)
go to 11
10
iset=1
alpha=ran(ix)
x=x+(-(alog(1-alpha))/(f+m))
if(x.gt.t)go to 20
pm=m/(f+m)
pmtest=ran(ix)
if(pmtest.lt.pm)go to 20
if(iset.eq.0)j=j+1
age=x+y
ts=y-0.5
if(iset.eq.1)go to 12
lmax=gauss(lmaxm,lmaxsd)
k=gauss(km,ksd)
12
continue
len=growth(age,ts,lmax,k)+0.5
pc=1/(1+exp(-(atb*len)))
if(iselect.eq.0)pc=1
pctest=ran(ix)
if(pctest.gt.pc)go to 10
nc=nc+1
if(len.gt.60.or.len.lt.1)print *,len,age,lmax,k

```

Fig. 24. Continued

```

ntp=x*12+1
if(ntp.gt.60)print *,ntp,age,x,y
lfreq(ntp,len)=lfreq(ntp,len)+1
imth=(i-t)*12+ntp
len2=(len+1)/2
if(imth.ge.1.and.imth.le.48)lf(imth,len2)=lf(imth,len2)+1
20 continue
return
end

c      ***** FUNCTION GAUSS *****
c
function gauss(a,b)
common/jjj/ix
c=0.
do 1 i=1,12
c=c+ran(ix)
1 continue
gauss=(c-6.)*b+a
return
end

```

Literature Cited

Pauly, D. and G. Gaschütz. 1979. A simple method for fitting oscillating length growth data, with a program for pocket calculators. ICES, C.M. 1979/G: 24, Demersal Fish Cttee, 26 p.

Author Index

- Abdul-Ghaffar, A.R. 19, 147, 148, 151-153, 157, 159, 164-165, 166, 322, 401, 402
 Abella, A. 22, 336, 340, 343
 Abramson, N.J. 8, 13, 31, 148, 164, 322, 333, 354, 362, 366, 371, 394
 Adams, S.M. 366, 371
 Allen, K.R. 47, 48, 51
 Almatar, S. 19, 31
 Alverson, D.L. 349, 351
 Al-Attar, M.A. 151, 162-164
 Al-Hossaini, M. 19, 147, 148, 151-153, 157, 159, 164-166, 322, 401, 402
 Al-Shoushani, M. 19, 22, 147, 322, 336, 340, 343, 401, 402
 Anderson, E.D. 288, 298
 Andreoli, M.G. 22, 311, 336, 340, 341, 343, 345, 369
 Anonymous 23, 31, 167, 176, 359, 387, 388, 394
 Anthony, V.C. 288, 298
 Atalo, P. 19, 31
 Baddar, M. 149, 165, 322, 325, 330, 333, 334
 Bagenal, T.B. 32, 165, 321, 333, 334
 Bakkala, R.G. 99
 Bakun, A. 17, 31
 Banerji, S.K. 344, 351
 Bannister, R.C.A. 288, 298
 Baranov, F.I. vii, 22, 31, 67, 74, 336, 342
 Bartoo, N.W. 300, 303
 Bawazeer, A.S. 330, 334
 Bayer, J. 17, 31
 Bazigos, G.P. 344, 351
 Beamish, R.J. 10, 31
 Beddington, J.R. 13, 17, 23, 25, 30, 31, 145, 193, 283, 286, 287, 298, 338, 364, 365, 369
 Bedford, B.C. 164, 165, 321, 323, 326, 331, 333, 334
 Behboodian, J. 286, 298, 364, 371
 Berg, C.J., Jr. 19, 31
 Bertalanffy, L. von 31, 46, 51, 87, 99, 208, 216, 239, 300, 303
 Bevertton, R.J.H. vii, 23, 31, 54, 56, 74, 123, 128, 204, 216, 232, 235, 238, 281, 300, 303, 350, 351, 358, 362
 Bhattacharya, C.G. 8, 13, 33, 76, 99, 366, 371, 393, 394
 Birkeland, C.E. 33
 Bishop, J. 160
 Bloomfield, P. 114, 119
 Borges, M.F. 288, 298
 Brander, K. 116, 119
 Branford, J.R. 159, 165
 Breen, P.A. 76, 83, 99, 140, 146, 289, 298, 354, 362, 366, 371
 Brey, T. 141, 146, 376, 392
 Brothers, E.B. 321, 333
 Brouard, F. 19, 31
 Buddemeyer, R.W. 33
 Cabanban, A.S. 19, 31
 Caddy, J.F. 1, 13, 23, 31, 33, 125, 128, 134, 136, 336, 375
 Campbell, G. 300, 303
 Cannizzaro, L. 311, 340, 341, 345, 369
 Carney, M.J. 349, 351
 Casey, J. 171
 Cassie, R.M. 8, 31, 76, 99, 148, 165, 322, 333, 366, 371
 Caswell, H. 41, 42
 Caton, A. 368, 372
 Chambers, J.M. 347, 348, 351
 Chapman, D.G. 204, 216, 297, 298
 Christensen, J.M. 321, 333
 Chuman, E. 21, 33
 Cillaurren, E. 19, 31
 Cingolani, N. 312, 319
 Clark, W.G. 76, 78, 98, 99, 297, 298
 Cleveland, W.S. 347, 348, 351
 Cochran, W.G. 344, 349, 351
 Collins, R.A. 300, 303
 Cooke, J.G. 23, 31, 286, 287, 298
 Cook, S. 396
 Corpuz, A. 377, 394
 Cox, D.R. 62, 74
 Cramer, H. 78, 95, 99
 Csirke, J. 1, 22, 23, 31, 33, 51, 121, 125, 128, 216, 336, 340, 343
 Cushing, D.H. 7, 31
 Daan, N. 180, 187
 Dalzell, P. 19, 31
 Damm, U. 36, 42, 45, 49, 124, 346, 355, 366, 375, 395, 404, 405
 Dashti, J. 330, 334
 David, N. 14, 20, 33, 42, 43, 53, 74, 76, 100, 113, 119, 138-140, 142, 145, 146, 148, 149, 166, 167, 176, 193, 202, 204, 208, 216, 284, 298, 312, 317, 319, 322, 323, 330, 334, 354, 362, 365, 372, 375, 376, 386
 De Angelis, D.L. 286, 298
 de Veen, J.F. 288, 298
 Dixon, W.J. 38, 42, 43
 Dizon, L.B. 394
 Draper, N. 78, 82, 99
 Ebert, T.A. 16, 31, 35, 36, 39, 40, 42, 43, 45-48, 51, 341, 346, 355, 360, 366, 375, 395, 396, 401, 402
 Edser, T. 22, 31, 67, 74
 Efron, B. 345, 351, 369, 371
 El-Musa, M. 22, 148, 149, 151-153, 157, 159, 164-166, 336, 340, 343
 FAO 150, 151, 157, 162, 165, 344, 351
 Farmer, A.S.D. 148, 159, 165
 Flatman, S. 351

- Forney, J.L. 179, 187
 Forrester, N.E. 345, 351
 Foucher, R.P. 99
 Fournier, D. 53, 74, 76, 78, 82, 83, 95, 97, 99, 100,
 140, 146, 193, 202, 284, 285, 289, 294, 298, 354,
 356, 360, 362, 366, 371, 422, 424
 Francis, C.R. 23, 31
 Fréon, P. 344, 351
 Frontini, N. 312, 319
 Fulton, T.W. vii, 9, 31
 Garcia, S. 1, 147, 166, 336
 Gaschütz, G. 15, 33, 76, 87, 88, 100, 128, 142, 146,
 456, 462
 Gayanilo, F.C. 375, 386, 391
 George, M.J. 19, 34
 Giannetti, G.F. 312, 319
 Gill, P.E. 94, 99
 Gislason, H. 179, 187
 Gjøsaeter, J. 19, 31
 Goeden, G. 14, 31
 Gomez, E. 33
 Gomez, O. 21, 33
 Gompertz, B. 13, 31
 Grandperrin, R. 19, 31
 Gray, H.L. 345, 351
 Green, R.H. 45, 51
 Gulland, J.A. vii, 7, 22, 23, 27, 33-34, 54, 74, 137,
 166, 178, 187, 193, 197, 202, 299, 303, 335, 336,
 338, 339, 341, 342, 344, 349, 351, 353, 358, 362,
 368, 371, 380, 421
 Gupta, R.P. 371
 Hakim, S. 19, 31, 330, 334
 Halldorsson, D. 288, 298
 Hald, A. 76, 99
 Hampton, J. 17, 22, 25, 139, 159, 193, 194, 196, 202-
 205, 212, 215, 216, 284, 336, 363, 365, 367-369,
 371, 372, 375, 456
 Hanlon, R.T. 20, 32
 Harding, D. 171
 Harding, J.P. 8, 32, 76, 99, 366, 371
 Hasselblad, V. 76, 98, 99, 354, 362, 366, 371
 Hearn, W.S. 196, 202
 Heincke, F. 22, 32, 67, 74
 Helgason, T. 179, 187
 Heltshe, J.F. 345, 351
 Hinkley, D.V. 62, 74
 Hixon, R.F. 20, 32
 Hoenig, J.M. 22, 121, 123, 128, 244, 281, 305, 336,
 340, 343, 349, 351, 355
 Hoenig, N.A. 123, 128, 244, 281
 Holden, M.J. 321-323, 329, 331, 334
 Holt, S.J. vii, 23, 31, 54, 56, 74, 123, 128, 204, 216,
 232, 235, 238, 281, 300, 303, 350, 351, 358, 362
 Hosillos, L.V. 394
 Houghton, R.G. 351
 Hoydal, K. 23, 32, 76, 77, 90, 100
 Hubold, G. 288, 298
 Hulet, W.H. 20, 32
 Hunter, W.G. 308, 309
 Ibrahim, M.A. 151, 166
 Ikenoue, H. 162, 164
 Ingles, J. 8-10, 15, 20, 25, 26, 32, 33, 49, 51, 206, 216,
 312, 317, 319
 Jacob, T. 19, 34
 Jakobsson, J. 288, 298
 James, M.K. 216, 371
 Jenrich, R. 37, 43
 Jensen, A.L. 23, 32
 Johannes, R.E. 17, 32, 33
 Jones, R. vii, 1, 6, 27, 32, 60, 67, 74, 138, 139, 146,
 148, 151-155, 165, 167, 168, 176, 178, 180, 182,
 187, 193, 202, 204, 214, 216, 217, 225, 236, 238,
 240, 243, 253, 281, 285, 299, 303, 306, 336, 346,
 351, 363, 364, 367, 369, 371, 375, 418, 445, 446,
 453, 455
 Kanciruk, P. 122, 128
 Kendall, M.G. 76, 78, 97, 100
 Kirkwood, G.P. 74, 138, 146, 368, 371
 Kleijnen, J.P.C. 196, 202
 Kleiner, B. 347, 348, 351
 Klima, E.F. 165
 Knafl, G. 308, 309
 Knights, B.J. 182, 187
 Krüger, F. 13, 32
 Krutchkoff, R.G. 308, 309
 Kumar, D.K. 366, 371
 Lamboy, W.F. 308, 309
 Landry, A.M. 165
 Larkin, P.A. 57, 74
 Larson, G.L. 16, 32
 Lasker, R. 321, 333
 Lassen, H. 22
 Laurec, A. 139, 239, 277, 281, 306, 308, 336, 363,
 367-369
 Laurs, R.M. 74
 Lawing, W.D. 123, 128, 244, 281, 349, 351
 Lea, E. 9, 32
 Levi, D. 22, 311, 312, 319, 336, 340, 341, 343, 345,
 369, 370
 Liew, H.C. 376, 394
 Longhurst, A. 17, 32
 Lough, R.G. 45, 51
 Lwin, T. 308, 309
 Macdonald, P.D.M. 36, 43, 76, 78, 82, 100, 113, 119,
 167, 176, 284, 294, 298, 354, 362, 364, 366, 371,
 372, 422, 424
 Maclean, J.L. 394
 Majkowski, J. 17, 22, 25, 65, 74, 139, 159, 179, 187,
 193, 194, 203-205, 212, 215, 216, 245, 281, 284,
 336, 363, 365, 367-369, 371, 372, 375, 456
 Mandecki, W. 368, 372
 Marsch, J.A., Jr. 33
 Mathews, C.P. 8, 10, 19, 31, 32, 53, 147-149, 150, 151,
 155, 157, 163-166, 284, 321-323, 325, 327, 330,
 331, 333, 334, 336, 341, 345, 348, 349, 353, 401,
 402, 421, 423

- Mathisen, O.A. 16, 32
 Mattice, J.S. 286, 298
 McFarlane, G.A. 10, 31
 McNew, R.W. 76, 98, 100, 366, 372
 McNiff, M. 396
 Mead, R. 148, 165
 Mendelsohn, R. 28, 32
 Mendelhall, W. 344, 351
 Mendo, J. 19, 21, 28, 32, 33
 Menon, M.D. 7, 32
 Mercer, M.C. 177, 187
 Mesnil, B. 27, 32, 139, 239, 277, 281, 306, 308, 367-369
 Miller, B.S. 40-43, 48, 51
 Miller, D.M. 369, 372
 Miller, R.G. 345, 351
 Mines, A.N. 32
 Mohammed, K.H. 148, 151, 159, 165
 Mohn, R.K. 368, 372
 Molloy, J. 288, 298
 Moreau, J. 368, 372, 391
 Morgan, G.R. 1, 7, 8, 10, 19, 27, 30-32, 35, 38, 45, 49, 53, 74, 75, 77, 95, 103, 113, 116, 121, 129, 134, 137, 138, 141, 145-147, 149, 165-167, 177, 193, 203, 217, 239, 283, 296, 299, 300, 305, 311, 321-325, 327, 329-337, 341, 343, 345, 349, 353, 357, 359, 361, 362, 365, 373, 375, 392, 394, 421, 424
 Morizur, Y. 240
 Muck, P. 21, 33
 Munro, J.L. vii, 12, 22, 23, 29, 32, 163, 166, 348, 350, 351, 375, 391
 Murphy, G.I. 32, 33, 74, 166, 178, 187, 202, 216, 281, 298, 334, 368, 371, 372, 380
 Murray, W. 94, 99
 NAG Library 94, 100
 Navaluna, N.A. 8, 19, 25, 26, 32, 33, 48, 51
 Neal, R. 8, 9, 15, 20, 33, 380
 Nickelson, T.E. 16, 32
 Nielsen, L.A. 351, 352
 Nishimoto, R. 74
 Olsen, S. 16, 32
 Orth, D.J. 368, 372
 Ott, L. 344, 351
 O'Connor, M. 31
 Paciorkowski, A.J. 288, 298
 Pagdilao, C.R. 19, 32
 Pallares, P. 281
 Palmer, D. 116, 119
 Palomares, M.L. 21, 33, 386
 Pannella, G. 8, 33, 321, 334
 Paranjape, M.A. 182, 188
 Parker, K.R. 300, 303
 Pauly, D. 1, 6-10, 13-17, 19-23, 25-27, 29-33, 35, 36, 38, 42, 43, 45, 48, 49, 51, 53, 64, 67, 74-77, 87, 88, 95, 100, 103, 113, 116, 119, 121-126, 128, 129, 134, 137-142, 145-149, 157-160, 163, 166, 167, 176, 177, 193, 195, 202-209, 212, 214-217, 238, 239, 281, 283, 284, 288, 293, 298, 300, 303, 305, 311, 312, 318, 319, 321-323, 329-331, 334, 335, 336, 337, 343, 345, 346, 348, 350, 352-354, 357, 358, 362-365, 367, 371-373, 375-377, 380, 381, 383-386, 391, 392, 394, 414, 456, 462
 Pearson, K. 76, 100
 Peloni, E. 312, 319
 Pennington, R.H. 37, 43
 Pereiro, J.A. 252, 281, 367, 372
 Perez, A. 151, 166
 Petersen, C.G.J. vii, 8, 9, 33, 76, 100
 Piernaar, L.V. 36, 43, 351, 352
 Pinder, J.E. 41, 43
 Pitcher, T.J. 36, 43, 76, 78, 82, 100, 284, 294, 298, 354, 362, 364, 366, 371, 372, 422, 424
 Polovina, J.J. 53, 123, 244, 355, 357, 366, 380, 384, 401
 Pomeroy, A.B. 216, 371
 Poole, L. 396
 Pope, J.A. 197, 202
 Pope, J.G. 17, 27, 30, 31, 33, 103, 167, 177-179, 181, 182, 187, 193, 204, 216, 238, 249, 253, 281, 299, 305, 338, 339, 354, 355, 357, 358, 360, 361, 366, 367, 372, 375, 386, 425
 Powell, D.G. 57, 74, 364, 372
 Prince, E.D. 365, 372
 Puga, R. 151, 166
 Pulos, L.M. 365, 372
 Qasim, S.Z. 7, 17, 33
 Ragonese, S. 22, 336, 340, 343
 Raith, D.F.S. 321, 323, 329, 331, 334
 Ralston, M. 37, 43
 Ralson, S. 53, 63, 123, 244, 355, 357, 366, 380, 384, 401
 Rao, C.R. 78, 100
 Reck, G.K. 19, 34
 Rheinhard, E.G. 159, 166
 Richards, F.J. 16, 34, 36, 43, 365, 372
 Ricker, W.E. 22, 34, 67, 74, 300, 303, 351, 352, 358, 362, 368, 372
 Ridgeway, J. 369, 372
 Rijnsdorp, A.D. 288, 298
 Robson, D.S. 204, 216, 297, 298
 Rodrigues, J. 151, 166
 Rodriguez, A. 18, 19, 34
 Roedel, P.M. 166, 351
 Rohde, J. 19, 34
 Rohlfs, J. 312, 318, 319, 345, 352
 Rørvik, C.J. 23, 32, 76, 77, 90, 100
 Rosenberg, A.A. 13, 17, 25, 30, 139, 145, 193, 283, 299, 305, 336, 338, 363-365, 369
 Rothschild, B.J. 33, 34, 166, 380
 Sacks, J. 308, 309
 Saeger, J. 377, 394
 Saif, J. 330, 334
 Saini, S.B. 45, 51, 166, 347, 351, 352
 Sainsbury, K.J. 13, 206, 208, 216, 269, 281, 284, 286, 298, 306, 309, 364, 372
 Sallares, S. 367, 372

- Sambilay, V., Jr. 377, 394
 Samuel, M. 149, 155, 165, 166, 323, 325, 327, 330,
 331, 333, 334
 Sanders, M.J. 22, 129, 134-136, 299, 305, 336, 338,
 340, 343, 356, 358, 367, 375, 427
 Sandland, R.L. 196, 202
 Sann, Aung 19, 33, 209, 216
 Scheaffer, R.L. 344, 351
 Scheffé, H. 308, 309
 Schnute, J. 43, 53, 74, 76, 78, 82, 95, 97, 100, 140,
 146, 193, 202, 284, 285, 294, 298, 354, 356, 360-
 362, 366, 372, 422, 424
 Schoch, W.F. 351, 352
 Schucany, W.R. 345, 351
 Scorcetelli, R. 312, 319
 Sharp, G.D. 33, 34, 51, 216
 Sharp, G.V. 17, 31
 Sheldon, R.W. 182, 188
 Shepherd, J.G. 104, 113, 116, 167, 168, 171, 176, 193,
 336, 337, 339, 341, 349, 353-357, 359, 365, 375,
 414, 418, 421, 424, 426
 Shoushani, M. 148, 166
 Silas, E.G. 19, 34
 Sims, S.E. 253, 281
 Sivasubramaniam, K. 151, 166
 Skillman, R.A. 8, 13, 34
 Smith, H. 78, 82, 99
 Smith, M.H. 41, 43
 Smith, S.J. 345, 352
 Sokal, R.R. 312, 318, 319, 345, 352
 Soriano, M. 375, 385, 391, 392, 394
 Sousa, M.I. 19, 31
 Sparre, P. 13, 17, 23, 30, 33, 75-77, 90, 100, 180, 188,
 284, 285, 291, 293, 294, 337, 338, 354, 357, 358,
 360, 361, 366, 375, 408, 410, 412, 418, 420, 422
 Spiegelman, C. 308, 309
 Ssentongo, G.W. 57, 74
 Stark, K.P. 216, 371
 Stuart, A. 76, 78, 97, 100
 Summerfelt, R.C. 76, 98, 100, 366, 372
 Sutcliffe, J.R. 182, 188
 Tanaka, S. 8, 34, 76, 100
 Tandog, D.D. 19, 34
 Thompson, D.W. 9, 34
 Thompson, J.A. 36, 43
 Tiews, K. 34
 Tomlinson, P.K. 76, 100
 Tsuda, R.T. 33
 Tsukayama, I. 29, 32, 33, 204, 216, 386
 Tukey, J.W. 347, 352
 Tukey, P.A. 347, 348, 351
 Uchiyama, J.H. 74
 Ursin, E. 10, 134, 179, 181, 187, 188
 Van Winkle, W. 371
 van Zalinge, N.P. 148, 151-155, 157, 164-166
 Vaughan, D.S. 122, 128
 Venta, G. 166
 Wadhwa, L.C. 216, 371
 Waring, G. 288, 298
 Weber, W. 17, 34
 Wellings, S.R. 40-43, 48, 51
 Went, A.E.J. 10, 34
 Wetherall, J.A. 53, 74, 123, 244, 355, 357, 364, 367,
 380, 384, 394, 401
 White, T.F. 15, 19, 34
 Wiener, J.G. 41, 43
 Williams, E.J. 308, 309
 Williams, T. 321, 323, 326, 331, 334
 Wolsteholmen, G.E.W. 31
 Woolner, L. 167
 Yahiaoui, M. 24
 Yang (*see* Yang, Jiming)
 Yang, Jiming 30, 177, 339, 425
 Yivisaker, D. 308, 309
 Yong, M.Y. 8, 13, 34

Geographic Index

- Adriatic 311
 Africa, West vii, 388
 Arabian Gulf 333
 Asia, Southeast 10
 Atlantic 2, 7, 9
 Australia 14
 Bahamas 19
 Bubiyan 151, *see also* Kuwait
 Burma 19, 26
 Cadiz, *see* Spain
 California, Gulf of 163
 Campeche Bank, *see* Mexico
 Cesenatico 312, 318, *see also* Adriatic
 Cochin, *see* India
 East Sound, *see* Washington
 Europe 10, 321
 Galapagos Island 19
 Great Barrier Reef, *see* Australia
 Heron Island, *see* Australia
 India vii, 19
 Indonesia 20
 Irish Sea 117, 118
 Italy 373
 Jamaica vii
 Kuwait 10, 19, 137-140, 143-152, 154, 155, 157-164,
 322-333, 358, 359, 401, 402, 418, 421

- Manila Bay, *see* Philippines
 Mazara del Vallo, *see* Italy
 Mexico 322
 Mexico, Gulf of 19, 20, 26
 Morocco 387
 Mozambique 19
 North Sea 7, 10, 19, 104-107, 111, 170-172, 176, 177, 180, 182, 186, 187, 219, 220, 418
 Orcas Island, *see* Washington
 Pacific 19, 74
 Papua New Guinea 19
 Peru 19, 28, 386
 Philippines 10, 19, 20, 48, 26
 Qatar 151
 Ragay Gulf 377, *see also* Philippines
 Rixa 151, *see also* Kuwait
 Saudi Arabia 151
 Shatt Al-Arab 151, *see also* Kuwait
 Shuaiba 156, 159, *see also* Kuwait
 Sicily vii, 19
 Spain 19
 Thailand, Gulf of 163
 Visayan Sea, *see* Philippines
 Washington 40

Species Index

- Abalone 99, 146, 371
Acanthopagrus cuvieri 324
Acanthopagrus latus 324
 Albacore, *see* *Thunnus alalunga*
 Anchoveta 19, 21, 28, 32, 33, 389, 390, *see also*
 Engraulis ringens
 Anchovy 12, 319
Arius thalassinus 324
Asaphis deflorata 19
 Billfish 371, 372
 Bonito, *see* *Sarda chiliensis*
 Bopyrid, *see* *Epipenaeon elegans*
 Butterfly fish, *see* *Chaetodon miliaris*
 Caesionidae 31
Carcharhinus obscurus 122, 123, *see also* shark
 Catfish, sea, *see* *Arius thalassinus*
Chaetodon miliaris 122, 123
 Chim 327, *see also* *Arius thalassinus*
Clupea harengus 19, 287, 288
 Clupeidae 31, 312
 Cod 34, 99, 103-105, 107, 109-111, 171-173, 180, 182, 187
Codakia orbicularis 19
 Croaker 32, 322, 324, *see also* Newaiby: *Otolithes ruber*
 Crustacea 19, 166
 Cutthroat trout 32
Decapterus russelli 19, 31
Engraulis encrasicolus 319
Engraulis ringens 287, 389, 390, *see also* Anchoveta
Epinephelus morio 322
Epinephelus sexfasciatus 26, 123, *see also* Grouper
Epinephelus spp. 328, *see also* Grouper
Epinephelus tauvina 19, 322, 324, 333, *see also* Hamoor
Epipenaeon elegans 159, 165
Gadus morhua 19, 103, 104, 187, 287, 288, *see also* Cod
 Grouper 25
 Grouper, banded, *see* *Epinephelus sexfasciatus*
 Grouper, orange-spotted; *see* Hamoor
 Grunt, silvery, *see* *Pomadasys argenteus*
 Haddock 182, 219, 220, 359, 418
 Haddock, faroe 223, 224
 Hake 28, 281, 388, 394
 Halibut, *see* *Hippoglossus hippoglossus*
 Hamoor 165, 322-325, 328, 329, 333, 334
 Hamra 137-141, 143-145, 322-325, 328, 329, 334
 Herring 298
Hilsa kelee 19, 31
Hippoglossoides elassodon 35, 40-43, 51, *see also* Sole,
 flathead
Hippoglossus hippoglossus 32
Katsuwonus pelamis 19, 31, 34
 Kelee shad, *see* *Hilsa kelee*
 Krill 336
 Leiognathidae 32
Leiognathus spp. 19
Limanda limanda 19
Loligo pealei 18, 20, 26, 32, *see also* Squid
Lutjanus coccineus 137, 139-141, 143-145, 322, 324
 Mackerel 298
 Mackerel, horse, *see* *Trachurus trachurus*
 Makrelen, *see* *Scomber japonicus*
Melanogrammus aeglefinus 19, 359, *see also* Haddock
Merlangius merlangus 19
Merluccius cadenati 394, *see also* Hake
 merluccius 28, 388, 394, *see also* Hake
 senegalensis 394, *see also* Hake
Metapenaeus affinis 147-158, 160-165
 dobsoni 19

- Metapenaeus* spp. 151
stebbingi 151
- Micropterus salmoides* 99
- Mollusc 19, 31
- Nakroor 326-328, 331, *see also Pomadasys argenteus*
- Nephrops* 146, 187, 216, 241, 281
- Newaiby 145, 166, 322-325, 328, 330, 331, 334, 360, 418, *see also Otolithes argenteus*
- Otolithes argenteus* 145, 166, 322, 324, 358, 359, 360, *see also Newaiby*
- Otolithes ruber* 19, 32, *see also Croaker*
- Oyster, pearl 19
- Pagophilus groenlandicus* 372
- Pampus argenteus* 19, 32, 146, 165, 322, 324, *see also Zobaidy*
- Panulirus* spp. 19
- Parapenaeopsis stylifera* 147-150, 152, 153, 155-158, 160-163, 165
- Penaeus kerathurus* 18, 19, 34
notialis 151
schmitti 151
semisulcatus 19, 147-165, 401-403
vannamei 163
- Perca flavescens* 187
- Perch, Yellow, *see Perca flavescens*
- Plaice 31, 32, 74, 109, 298
- Plectropomus leopardus* 14
- Pleuronectes platessa* 287-288
- Pomadasys argenteus* 324, 327
- Pomadasys argyreus* 25
- Pomfret, *see Pampus argenteus*
- Porgy 34
yellow-finned black, *see Acanthopagrus latus*
silvery black, *see Acanthopagrus cuvieri*
- Pterocaesio pisang* 19, 31
- Raja clavata* 116-119
- Rastrelliger brachysoma* 19, 26
- Russel's scad, *see Decapterus russelli*
- Sandeel 182
- Sarda chiliensis* 300, 303
- Sardina pilchardus*, *see Sardine*
- Sardine 311-314, 319, 387, 394
- Sardinella longiceps* 25
- Scomber japonicus* 19, 32
- Scomber scombrus* 287, 288
- Seal, harp, *see Pagophilus groenlandicus*
- Shark 122, 123, 128, 371, 372
- Sheim 326, 328, *see also Acanthopagrus latus*
- Shrimp 4, 10, 12, 15, 18-20, 31, 33, 147-152, 154, 155, 157-166, 322, 333, 334, 336, 380, 401, 423
- Slipmouth, *see Leiognathidae*
- Snapper, crimson, *see Hamra*
- Snapper, red, *see Hamra*
- Sobaity 326, 328, 334, *see also Acanthopagrus cuvieri*
- Sole 298, 333
- Sole, flathead 35, 40-43, 48, 51
- Solea solea* 288
- Squid 4, 15, 18, 19, 32, 33, 147, 166
- Stizostedion vitreum vitreum* 187
- Stolephorus* spp. 19
- Stolephorus zollingeri* 26
- Taius tumifrons*, *see porgy*
- Teleost 18, 20
- Thunnus alalunga* 74
- Thunnus albacares* 19, 31
- Thunnus maccoyii* 371, *see also Tuna, bluefin*
- Trachurus trachurus* 287, 288, 298
- Trout, coral *see Plectropomus leopardus*
- Tuna 368, 371, 372
bluefin 74, 368, 371, 372
skipjack, *see Katsuwonus pelamis*
- Upeneus moluccensis* 377-379
- Walleye, *see Stizostedion vitreum vitreum*
- Whiting 182
- Zoarces viviparus* 76
- Zobaidy 323, 324, 330, 331, 334, *see also Pampus argenteus*

ICLARM PUBLICATIONS ON STOCK ASSESSMENT

- Theory and management of tropical fisheries. D. Pauly and G.I. Murphy, Editors. 1982.
- Theory and management of tropical multispecies stocks: a review, with emphasis on the Southeast Asian demersal fisheries. D. Pauly. 1979. Reprinted 1983.
- Caribbean coral reef fishery resources. J.L. Munro, Editor. 1983.
- Fish population dynamics in tropical waters: a manual for use with programmable calculators. D. Pauly. 1984.
- Indonesian marine capture fisheries. C. Bailey, A. Dwiponggo and F. Marahudin. 1987.
- The Peruvian anchoveta and its upwelling ecosystem: three decades of change. D. Pauly and I. Tsukayama, Editors. 1987.
- Small-scale fisheries of San Miguel Bay, Philippines: biology and stock assessment. D. Pauly and A.N. Mines, Editors. 1982.
- An atlas of the growth, mortality and recruitment of Philippine fishes. J. Ingles and D. Pauly. 1984.
- Growth, mortality and recruitment of commercially important fishes and penaeid shrimps in Indonesian waters. A. Dwiponggo, T. Hariati, S. Banon and D. Pauly. 1986.

For further information, write to The Editor, ICLARM, MC P.O. Box 1501, Makati, Metro Manila, Philippines.



INTERNATIONAL CENTER FOR LIVING AQUATIC RESOURCES MANAGEMENT

The International Center for Living Aquatic Resources Management (ICLARM) is an autonomous, nonprofit, international scientific and technical center which has been organized to conduct, stimulate and accelerate research on all aspects of fisheries and other living aquatic resources.

The Center was incorporated in Manila on 20 January 1977 and its operational base was established in Manila in March 1977.

ICLARM is an operational organization, not a granting entity. Its program of work is aimed to resolve critical technical and socioeconomic constraints to increased production, improved resource management and equitable distribution of benefits in economically developing countries. It pursues these objectives in the fields of aquaculture, resource assessment and management, education and training, and information through cooperative research with institutions in developing and developed countries. ICLARM's work is international.

Policies are set by a Board of Trustees with members drawn from the international community. Direction of ICLARM, under the policies set by the Board, is the responsibility of the Director General. Advice on programs is received by the Director General from *ad-hoc* Program Advisory Meetings composed of scientists drawn from the international community.

The ICLARM core staff consists of internationally recruited scientists with expertise in aquaculture, fishery biology, population dynamics and socioeconomics. In addition, provision is made for interns, consultants and visiting fellows, contributing to breadth of competence and flexibility. The programs of ICLARM are supported by a number of private foundations and governments.



KUWAIT INSTITUTE FOR SCIENTIFIC RESEARCH

The Kuwait Institute for Scientific Research (KISR) is the principal government research establishment in Kuwait and provides research and development support to both Kuwaiti industry and other government organizations.

In addition to aquaculture and fisheries, its staff are involved in research areas related to the special needs of Kuwait such as petroleum technology, improving agricultural and industrial production, solar energy research and environmental protection. Training of Kuwaiti nationals in these fields is also a major responsibility of KISR.

The Mariculture and Fisheries Department was one of the first research departments to be established within KISR. Its work on marine fish culture and management of the country's fish resources has become widely known both locally and internationally and is contributing to Kuwait's move towards greater self-sufficiency in fisheries products.