FAO
COMPUTERIZED INFORMATION SERIES
fisheries

8

FAO-ICLARM
stock assessment tools
Reference manual

F. C. Gayanilo, Jr.
International Center for Living Aquatic Resources Management (ICLARM)

D. Pauly
International Center for Living Aquatic Resources Management (ICLARM) and
Fisheries Centre, University of British Columbia
In 1990, the FAO of the United Nations and ICLARM (International Center for Living Aquatic Resources Management) signed an agreement to develop a joint software package for fish stock assessment for use with microcomputers, primarily aimed at helping scientists working on tropical fisheries resources to produce scientific advice for fisheries management. The resulting software, called FISAT (FAO–ICLARM Stock Assessment Tools), was released in 1996, along with a short user’s guide (Gayanilo et al. 1996). This document complements the user’s guide by presenting the theoretical background, assumptions and limitations of the models/methods incorporated in FISAT.

The materials included in this document were compiled from the following three sources, in order of importance:

(i) adaptation of previous publications of ICLARM or FAO, especially the manuals of Pauly (1984a) and Sparre and Venema (1992);

(ii) publications by other authors, specially the authors of original methods incorporated into FISAT;

(iii) new material written by the editors, required to fill in gaps or to document methods incorporated in FISAT, but not previously documented.

A footnote to the title of each section indicates which of these sources was used in cases (i) and (ii); case (iii) is indicated by the absence of an attribution. Permission to use published material was obtained whenever appropriate.

When applicable, models and methods are demonstrated with (a) derivations and assumptions, (b) an example with intermediary calculations, and (c) application of the model to a hypothetical data and/or "real" data using FISAT. The sample data used in this manual are provided in the sample data disk distributed with the original software package.

The disks distributed along with this document contain an update to Shepherd’s method (SLRCA) and a new Microsoft Windows® freeware, ABee.

* Microsoft Windows is a registered trademark of Microsoft Corporation, USA.
A SETUP routine is available in the ABee subdirectory to properly install the software. ABee includes routines to estimate the coefficients (a and b) of length–weight relationship from either pairs of length–weight data, weighted or unweighted by sample size (using either the linear regression or the non-linear least square method) and from length–frequency samples and their bulk weights; the latter is a new method.

Acknowledgments

The editors would like to express their gratitude to the authors who contributed to this manual and especially to Siebren Venema, project manager (projects CGP/INT/392/DEN and GCP/INT/575/DEN), for critically reading the draft of this manuscript.

Special thanks are also due to Per Sparre for his contribution to the initial phase of the project which led to this manual.

Also, we thank the scientific staff of DIFRES (Danish Institute of Fisheries Research) for their comments on various aspects of FISAT, and Ms. Rachel Atanacio, Mr. Rustico S. Espiritu and Mr. Alvin Laquian for their drafting and scanning of the figures used in the manual.

Finally, we thank the participants of successive FAO/DANIDA training courses in tropical fish stock assessment for inspiring many of the methods presented here, and for forcing us to learn how to present them effectively.
**ABSTRACT**

This manual presents key concepts and methods found useful for assessing the components of tropical and subtropical fish stocks, with emphasis on short- and medium-lived species of finfish and invertebrates (shrimps, squids, etc.) and on approaches based on the detailed analysis of length–frequency data. This account covers the preprocessing of length–frequency and ancillary data, growth estimation, mortality, gear selection, virtual population analysis and yield prediction.

The manual is illustrated by means of computational examples, graphics and tabular printouts generated by the FISAT software (FAO–ICLARM Stock Assessment Tools), and thus complements the FISAT User’s Manual.

---

<table>
<thead>
<tr>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparation of this document ................................................................. iii</td>
</tr>
<tr>
<td>Acknowledgments ...................................................................................... v</td>
</tr>
<tr>
<td>Abstract ..................................................................................................... vi</td>
</tr>
<tr>
<td>Contents ..................................................................................................... vii</td>
</tr>
<tr>
<td>Chapter 1. <em>Introduction</em> ................................................................. 1</td>
</tr>
<tr>
<td>What you will learn from this chapter ................................................ 1</td>
</tr>
<tr>
<td>Why stock assessment? ................................................................. 1</td>
</tr>
<tr>
<td>Basic concepts of fisheries science .................................................... 4</td>
</tr>
<tr>
<td>The stock concept ................................................................................. 4</td>
</tr>
<tr>
<td>Tropical multispecies stocks ............................................................. 9</td>
</tr>
<tr>
<td>Characteristics of fisheries arising from biological factors .................. 13</td>
</tr>
<tr>
<td>Problems related to statistical data ..................................................... 17</td>
</tr>
<tr>
<td>Fishery research problems ............................................................... 18</td>
</tr>
<tr>
<td>Chapter 2. <em>Overview of Fish Stock Assessment Models</em> .... 20</td>
</tr>
<tr>
<td>What you will learn from this chapter ................................................ 20</td>
</tr>
<tr>
<td>Introduction ............................................................................................ 20</td>
</tr>
<tr>
<td>Analytical models ............................................................................... 21</td>
</tr>
<tr>
<td>Holistic models .................................................................................... 24</td>
</tr>
<tr>
<td>Overview of surplus-production models ......................................... 27</td>
</tr>
<tr>
<td>Flow diagrams in fish stock assessment ............................................. 29</td>
</tr>
<tr>
<td>Estimating growth parameters directly from length compositions ........ 30</td>
</tr>
<tr>
<td>Converting length data to age information prior to estimation of growth parameters ........................................... 32</td>
</tr>
<tr>
<td>Length-based fish stock assessments ................................................. 32</td>
</tr>
<tr>
<td>Chapter 3. <em>Data Handling for Fish Stock Assessment</em> ...... 35</td>
</tr>
<tr>
<td>What you will learn from this chapter ................................................ 35</td>
</tr>
<tr>
<td>Introduction ............................................................................................. 35</td>
</tr>
<tr>
<td>Chapter 4. Analysis of Growth</td>
</tr>
<tr>
<td>-------------------------------</td>
</tr>
<tr>
<td>What you will learn from this chapter</td>
</tr>
<tr>
<td>Introduction</td>
</tr>
<tr>
<td>Using time series of L/F data</td>
</tr>
<tr>
<td>ELEFAN I</td>
</tr>
<tr>
<td>Shepherd's method</td>
</tr>
<tr>
<td>Original version</td>
</tr>
<tr>
<td>Improved version</td>
</tr>
<tr>
<td>Powell-Wetherall method</td>
</tr>
<tr>
<td>Using size-at-age data</td>
</tr>
<tr>
<td>Modal progression analysis (MPA)</td>
</tr>
<tr>
<td>Bhattacharya's method</td>
</tr>
<tr>
<td>Hasselblad's NORMSEP</td>
</tr>
<tr>
<td>Linking of mean lengths</td>
</tr>
<tr>
<td>Estimating the C.V. of L_{ct}</td>
</tr>
<tr>
<td>Estimating growth from length increments</td>
</tr>
<tr>
<td>Gulland &amp; Holt plot</td>
</tr>
<tr>
<td>Background and 1st example</td>
</tr>
<tr>
<td>Amplifications, 2nd and 3rd examples</td>
</tr>
<tr>
<td>Munro's method</td>
</tr>
<tr>
<td>Fabens' method</td>
</tr>
<tr>
<td>Modified Appeldoorn method</td>
</tr>
<tr>
<td>Comparing the growth of fishes</td>
</tr>
<tr>
<td>Growth comparison</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 5. Mortality, Gear Selection and Recruitment</th>
<th>104</th>
</tr>
</thead>
<tbody>
<tr>
<td>What you will learn from this chapter</td>
<td>104</td>
</tr>
<tr>
<td>Introduction</td>
<td>104</td>
</tr>
<tr>
<td>Total mortality from catch curves</td>
<td>106</td>
</tr>
<tr>
<td>Age-structured catch curve</td>
<td>109</td>
</tr>
<tr>
<td>Length-converted catch curves (without seasonality of growth)</td>
<td>110</td>
</tr>
<tr>
<td>Jones and van Zalinge cumulative plot</td>
<td>116</td>
</tr>
<tr>
<td>Length-converted catch curves accounting for seasonal growth</td>
<td>118</td>
</tr>
<tr>
<td>Selecting points for the regression</td>
<td>122</td>
</tr>
<tr>
<td>Z from Hoenig's models</td>
<td>124</td>
</tr>
<tr>
<td>Estimating L_{max}</td>
<td>130</td>
</tr>
<tr>
<td>Natural mortality</td>
<td>132</td>
</tr>
<tr>
<td>Pauly's M empirical equation</td>
<td>132</td>
</tr>
<tr>
<td>Rickhert and Efandov's method</td>
<td>134</td>
</tr>
<tr>
<td>Estimating M from selection data</td>
<td>135</td>
</tr>
<tr>
<td>The original approach</td>
<td>135</td>
</tr>
<tr>
<td>Improving Munro’s approach</td>
<td>138</td>
</tr>
<tr>
<td>Gear selectivity</td>
<td>141</td>
</tr>
<tr>
<td>Trawl mesh selection</td>
<td>141</td>
</tr>
<tr>
<td>Gill net selection</td>
<td>146</td>
</tr>
<tr>
<td>Baranov/Holt model</td>
<td>148</td>
</tr>
<tr>
<td>Multiple mesh sizes</td>
<td>152</td>
</tr>
<tr>
<td>Estimation of probabilities of capture from length-converted catch curves</td>
<td>154</td>
</tr>
<tr>
<td>Recruitment patterns</td>
<td>158</td>
</tr>
</tbody>
</table>
Chapter 1. Introduction

What you will learn from this chapter

In this chapter, you will be introduced to the rationale and fundamentals behind fish stock assessment, and to specific problems of tropical fish stock assessment.

Why stock assessment?\(^1\)

Capture fisheries exploit living and, hence, renewable resources. In the overwhelming majority of cases, fisheries resources are “common property”. In many cases they are also “open access”—i.e., fishing licenses are not required, or are available at nominal fees, or other limitations to entry are not effectively enforced. Open and free access generally implies competition for the resources which, in an unregulated fishery, leads ultimately to biological and economic overfishing.

Biological overfishing has different effects (growth and/or recruitment overfishing with regard to single species, ecosystem overfishing with regard to a multispecies resource), all discussed, if briefly, in this manual.

Economic overfishing is also multifaceted. When one considers the first boat operating in a fishery, “trouble” actually begins when a second unit starts fishing, thus reducing the biomass (and hence, catch/effort and economic returns) available to the first unit. Progressive reduction of the resource then occurs, until there are so many units in the fishery that the boats cannot manage on average to break even, and fishery

---

\(^1\) Adapted from Pauly et al. (1991) and Sparre and Venema (1992).
operators clamor loudly for subsidies (for example in the form of tax reductions on fuel or gear, subsidies for the renewal of boats, or even more research and management).

At the level of society, trouble begins when the difference between the total revenue in a fishery and the total cost of fishing—i.e. the net economic yield or resource rent—starts to decline. Fisheries resource rents can be considerable, but are dissipated in unmanaged fisheries.

Providing a basis for a rational management of exploited resources thus involves dealing with a wide variety of issues, ranging from political problems to socioeconomic and resource conservation issues. However, this manual emphasizes only the biological aspects, i.e. fish stock assessment. Needless to say, when used as the basis for management, the other components—economic and sociological—should also be considered (see Coppola et al. 1992, Cochet and Gilly 1990, Sparre and Willman 1992).

In essence, the ultimate goal of fish stock assessment is to provide technical advice to maintain the productivity of fish stocks.

To further illustrate this basic objective, consider Fig. 1.1. On the horizontal axis is fishing effort, \( f \) (measured here in number of boat days). On the other axis is yield, i.e. landings (usually in weight; however if the landings consist of different groups of animals, for example shrimp, finfish and squid, it may be more appropriate to express the yield in terms of value). The figure shows that up to a certain level we gain by increasing fishing effort, but beyond that level, the renewal of the resource (through reproduction and the body growth of individuals) cannot keep pace with the removal caused by fishing, at which point (MSY or “maximum sustainable yield”) any further increase in exploitation level causes yield to fall.

A few fisheries resources in the world may still be underexploited—particularly in inland water bodies in which a combination of factors (remoteness, lack of suitable gear and crafts, lack of markets, or insecure conditions) have hindered the build-up of fishing capacities. To account for such cases, fish stock assessment should also:

- evaluate the potential yields of underutilized water bodies, taking the need for resource conservation into account;
- consider adoption and improvement of suitable fishing gear and craft combinations (for example through fishing trials), and of processing methods and marketing;
- define technical assistance schemes that allow a phased growth of the fisheries, and the adjustment of effort to those levels leading to high, but sustainable yields.

Living aquatic resources have become a major source of food protein, especially in the tropics, providing around a quarter of the total animal protein in many developing
countries and up to three quarters in some others. Yet despite their importance, tropical and subtropical fisheries are most often badly managed (if at all) — the resources are generally overexploited and the fisheries overcapitalized.

Fisheries scientists alone cannot remedy this situation; however, they can provide the biological advice required to rehabilitate stocks that are mismanaged. The methodology for this exists. It consists of various refinements of standard assessment techniques, together with the judicious application of new methods developed for use in tropical waters. This manual is structured around a methodology of this sort.

Basic concepts of fisheries science

Fish are usually not seen, counted or weighed before they are caught, and hence fisheries research has, from the very start, depended on indirect inferences (mathematical and statistical) for quantifying the vital statistics of fish populations required for fisheries management.

Thus, some of the earliest applications of mathematics and statistics (outside of the physical sciences) are those of fisheries biologists. However, when describing the dynamics of an exploited aquatic resource, the first concept that needs defining is that of the stock.

The stock concept

A stock is a subset of a species, the species itself being the basic taxonomic unit. Thus, a prerequisite for the identification of stocks is the ability to distinguish between different species. Because of the great number of different, often similar species observed in tropical fisheries, their identification can be difficult. The fishery scientist, however, must master the techniques of species identification if any meaningful fish stock assessment is to come out of the data collected. Tools to solve species identification problems are provided by the FAO species identification sheets for fishery purposes (Fischer 1978, Fischer and Bianchi 1984, Fischer, Bianchi and Scott 1981, Fischer and Whitehead 1974), the FAO species catalogues (Allen 1985, Carpenter 1988, Carpenter and Allen 1989, Cohen et al. 1990, Colette and Nauen 1983, Compagno 1984 and 1984a, Holthuis 1980 and 1991, Marquez 1990, Nakamura 1985, Roper, Sweeney and Nauen 1984, Russel, 1990, Whitehead 1985, Whitehead, Nelson and Wongratana 1988) and most recently by the development of a computerized global fish database on fish biology, FishBase, also usable for species identification (Froese and Pauly 1996), available on CD-ROM.*

By a stock we mean a subset of one species having the same growth and mortality parameters, and inhabiting a particular geographical area.

A group of animals for which the geographical limits can be defined may be considered a stock for the purposes of fish stock research. Such a group of animals should however share a common gene pool. For species showing little migratory behavior (mainly demersal species), it is easier to identify a stock than for highly migratory species such as tunas.

Cushing (1968) defines a stock as a subset of a species with a single spawning group which the adults join year after year. Larkin (1972) defines a stock as "a population of organisms which, sharing a common gene pool, is sufficiently discrete to warrant consideration as a self-perpetuating system which can be managed", while Ihssen et al. (1981) define a stock as "an

* Contact ICLARM (or fishbase@cgnet.com) for information on this joint ICLARM/FAO product.
intraspecific group of randomly mating individuals with temporal or spatial integrity".

Ricker (1975) defines a fish stock as "the part of a fish population which is under consideration from the point of view of actual or potential utilization". This definition reflects a different view of the stock concept.

A definition of the stock concept acceptable to everyone with an interest in intraspecific grouping may be unattainable. For reviews of this concept see Booke (1981), Ihssen et al. (1981) and Maclean and Evans (1981).

Perhaps the most suitable definition in the context of fish stock assessment was given by Gulland (1983), who stated that for fisheries management purposes the definition of a unit stock is an operational matter, i.e. a subgroup of a species can be treated as a stock if differences within the group and interchanges with other groups can be ignored without making invalid the conclusions reached in the course of an assessment.

This implies that it will generally be preferable to initiate stock assessments over the entire area of distribution of a species, as long as there are no indications that separate unit stocks exist in that area. If it becomes clear that the growth and mortality parameters differ significantly in various parts of that area, it will be necessary to perform the assessment on a stock basis. (Note that the identification of separate stocks is a complex matter, usually requiring many years of data collection and analysis.)

When distinct stocks have been identified, assessments should refer separately to each stock. The results may (or may not) subsequently be pooled for assessment of a multispecies fishery. Therefore, in multispecies situations, the input data should be available for each stock of each species considered, i.e. they should remain available in their unaggregated form.

The stock concept is closely related to the growth and mortality characteristics displayed by a group of fish. The growth parameters are numerical values in an equation by which we can predict the body size of fish when they reach a certain age. The mortality parameters reflect the rate at which the animals die, i.e. the number of deaths per time unit. The mortality parameters considered in this manual are fishing mortality, which defines the "deaths" caused by fishing, and natural mortality, which accounts for all other causes of death (predation, disease, etc.).

Some practical difficulties associated with stock identification are:

(i) generally, only a part of the distribution area is studied;

(ii) often, several independent stocks coexist locally, such that their areas of distribution overlap at least some of the time;

(iii) several countries may exploit the same stock, as in the case of many highly migratory stocks such as tunas.

Fig. 1.2. illustrates cases (ii) and (iii). In the left panel, we consider a fish stock, the distribution area of which is indicated by a solid line; the broken lines delimit the EEZ (exclusive economic zone) of each country. The dotted area indicates the fishing area of country C and the hatched areas those of countries A and B. If country C were to base its assessment on the assumption that the stock was limited to its own fishing area, thus ignoring the fishing pattern exhibited by A and B, it would be likely to draw wrong conclusions. If, for example, countries A and B maintain intensive fisheries on the stock in question, such that it is overexploited (i.e. a reduction of the fishing intensity would increase the yield), there is little that country C can do on its own to improve the situation.
Thus, assessments based on the assumption of a stock limited to country C’s waters alone would lead that country to conclude that the stock is overexploited, and it might therefore introduce management measures to reduce fishing effort. However, the expected effect of the management measures might not be realized if countries A and B did not impose similar actions.

The right panel of Fig. 1.2 illustrates the case where one fishery exploits several stocks. In this scenario, the assessment can refer only to the “average stock”, since it may be difficult to separate the catches by stocks. If, however, the fishing effort expended is similar for each stock, the result of the assessment would still be correct. There are other complications for this scenario: suppose that the three stocks 1, 2 and 3 are heavily overfished, and that the fishery is expanded to include the unexploited stock 4; in this case the average catch rate would increase, which would lead to false conclusions about the status of stocks 1, 2 and 3.

The key to understanding stock structures is knowledge of migration routes. Nearly all exploited marine organisms undertake migrations, such as to and from spawning grounds. Information regarding migration can be obtained from tagging experiments and from data provided by the fishery itself.

Some general conclusions can be drawn from these facts. First, it is usually safer to assume that fish from neighboring fishing areas belong to one unit stock than to assume that each separate fishery exploits its own unit stock. Further, proper assessment can be carried out only when the biology of a resource species, including its migrations, spawning habits, etc., is well understood. Fish stocks are not bound by human borders and this means that proper assessments can be made only when the effects of such borders on the distribution of fishing effort is understood, or when the borders are overcome through international cooperation.

**Tropical multispecies stocks**

The first and most obvious feature of tropical multispecies stocks is the multitude of species occurring on the fishing grounds. For example, the following species numbers were reported from surveys on Southeast Asia:

- eastern peninsular Malaysia: 341 species (Pauly 1979b);
- Java Sea and southern tip of South China Sea: 230 species (Widodo 1976);

These figures are lower limits and depend on the numbers of stations covered. Current estimates for the total number of fish species in the Indo-Pacific area are as high as 6,000–7,000 (Carcasson 1977), of which a large proportion occurs in Southeast Asia.

---

3 Adapted from Pauly (1979b).
In general, single hauls with 50 species or more are quite frequent. For thoughts on the implications of such numbers of species for routine assessment see Marr (1976) and Pauly (1979b).

A second, very marked feature of tropical multispecies stocks is that in general, most of the component species are small-sized. In shallow waters in Southeast Asia the bulk of the catch generally comprises Leiognathidae, which have a mean maximum length of about 12 cm. (One species, *Leiognathus equalus*, reaches up to 30 cm. The figure of 12 cm refers to the rest of the leiognathid species, which are all small.) In deeper waters the bulk of the catch is often represented by Gerridae, with a similarly small length. Large fish are much less common, the whole picture being that of a typical "food pyramid".

A third, very important feature of the tropical stocks is that the peak occurrence of many of their constituent species is in shallow waters. The Leiognathidae for example have the maximum of their biomass at a depth of about 25 m, while the Trygonidae (rays) are most abundant at 10–20 m.

Migratory movements of demersal species have been barely studied in the tropics. However, tagging studies in the Gulf of Thailand suggest "that the demersal fishes do not make any extensive migrations" (Chomjurai and Buñag 1970). On the other hand, there is ample evidence that most species are represented by larger specimens in the offshore, deeper waters. This can be demonstrated for a large number of species, for example on the basis of the extensive length–frequency data presented by Martosubroto and Pauly (1976), which cover approximately 90 species (ca. 40,000 measured specimens) from the Java and South China seas. (This theme is further developed in Longhurst and Pauly 1987.)

As a whole, however, these data also suggest that there are no distinct gaps or discontinuities separating the young from the adults, or the reproductive stages from the reproductive stocks.

Finally, it appears that tropical species assemblages are in most cases climax communities, the outcome of a long, common, evolutionary history in an extremely stable environment (Eckman 1967). That such assemblages of fish species differ from those occurring, say, in the North Atlantic, is quite obvious.

On the other hand, it is similarly obvious that acknowledging the existence of these differences between high-latitude and tropical ecosystems has seldom prevented fishery biologists from applying principles derived from high-latitude marine ecosystems to the fundamentally different tropical marine ecosystems. Garrod (1973) wrote that for example in high latitude "multiple stock fisheries resources form a robust system" which "can tolerate wide variations in fishing mortality...without adverse effects". However, before applying this concept of a "robust system" to tropical marine ecosystems, the following questions should be answered:

- Is the statement correct as a whole, or does it exclude certain groups of species, such as the clupeoids (see Murphy 1977)?

- If the statement does apply, at least to predominantly demersal systems, why are high-latitude multiple-species systems robust? Is it because of their "system" property, or rather because high-latitude systems are composed of single species each of which can withstand high variations in fishing mortality?

Obviously, the answers to these last questions are crucial to the management of multispecies fisheries. A positive answer to the first question would, for example, imply that the knowledge derived from, say, the North Atlantic fisheries and the stock interactions observed there can be generalized and then applied
to a tropical situation. On the other hand, a positive answer to the second question would imply that tropical marine ecosystems may not be robust at all.

Ecological theory, as reviewed in various texts (e.g. Ricklefs 1973), does not seem to provide a clear-cut answer to these questions, at least when fish communities are considered. It seems generally accepted, however, that tropical fishes interact most strongly with the biotic components of their environment, while temperate fishes seem to be more strongly affected by the abiotic components of their environment (see for example Nurstall 1977). This seems to be confirmed by the fact that natural mortality (as caused mainly by predation), which in fishes is a function of both size and growth rate, is also a function of environmental temperature (Pauly 1980; and see the section on natural mortality in chapter 5). This relationship, demonstrated on the basis of literature data on 175 fish stocks, suggests that natural mortality (M) in tropical fishes is, other things being equal, twice as high in tropical as in temperate waters. Similarly, evidence is available showing that, other things being equal, fishes living in warm water eat more than their cold water counterparts (Palomares and Pauly 1989, Pauly 1989).

Some qualitative differences between tropical and temperate ecosystems (both terrestrial and marine) have been reviewed by Robinson (1978). With respect to marine ecosystems, coral reef communities can be used to show that major qualitative differences do occur between tropical and temperate ecosystems. Three phenomena where identified which have no obvious counterparts in temperate waters:

- cleaning symbioses (for example between cleaner wrasses and various reef fishes);
- mutualistic associations (damsel/fish sea anemone, goby/snapping shrimp);
- mimicry (for example by Aspidontus taenaiatus posing as a cleaner wrasse).

Robinson (1978) concluded his review of tropical biology by suggesting that in tropical communities, species interactions may be more intensive and/or numerous than in temperate communities, even after accounting for the number of species involved in the interactions (which is generally higher in the tropics and thus by itself would provide for more sites for interaction). Although Robinson's hypothesis is falsifiable in principle, it seems to have been rarely tested rigorously (at least as far as aquatic systems are concerned), despite its obvious relevance to fishery research.

**Characteristics of fisheries arising from biological factors**

The theme of difference between tropical and temperate systems and their relevance to fish stock assessment is illustrated in Fig. 1.3. Fig. 1.3A shows the yield-per-recruit obtained for a typical temperate fish with a low ratio of mortality to growth (M/K) value (how to construct such curves is shown later in this manual). As can be seen, the curve has a pronounced maximum corresponding to the relatively low fishing mortality (Fmax) generating maximum yield per recruit (Y/Rmax). Fig. 1.3B, on the other hand, presents the yield per recruit obtained from a typical tropical fish with a high M/K value. As might be seen, the clear maximum of Fig. 1.3A has disappeared, and is replaced by an ill-defined plateau at very high fishing mortalities.

\[\text{Adapted from Pauly (1979b).}\]
Clearly, attempts at maximizing yield-per-recruit in the case of Fig. 1.3B would lead to doom because maximization leads, in such cases, to extremely low stock biomasses, and hence to equally low catch per effort, and to potential recruitment failures.

This is an example where the yield-per-recruit concept can be used meaningfully only in conjunction with additional information or assumptions on the appropriate level of fishing mortality. One approach, based on the assumption that the optimum fishing mortality ($F_{opt}$=$F_{MSY}$; see Symbols and Definitions) should be approximately equal to the natural mortality prevailing in a stock, leads to the well-known equation

$$MSY \approx 0.5MB_o$$

where $B_o$ is the size of the unexploited stock. However, several studies have shown that this equation tends to overestimate yields by a factor of two or more (Francis 1974, Caddy and Csirke 1983, Beddington and Cooke 1983).

Another approach is the "$F_{0.1}$" concept, i.e., an arbitrary limit to fishing mortality set at a level such that the marginal increase in yield-per-recruit is 1/10 of its value near the origin of the curve (Gulland and Boerema 1973, Caddy and Garcia 1983; see Fig. 1.3B).

These considerations can be used to derive simple generalizations for use with most single-species stocks of small to medium-sized tropical fish, in which the yield-per-recruit model of Beverton and Holt (1957, 1959) is reduced to the four-options model in Table 1.1.

Such generalizations, when applied judiciously, can considerably reduce the costs of preliminary resources assessment and fishery management for tropical multispecies demersal stocks.

<table>
<thead>
<tr>
<th>Option</th>
<th>Relative size of fish caught</th>
<th>Fishing mortality</th>
<th>Results (in terms of Y/R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>large</td>
<td>low</td>
<td>Underfishing</td>
</tr>
<tr>
<td>B</td>
<td>large</td>
<td>high</td>
<td>eumorotic fishing</td>
</tr>
</tbody>
</table>
An effect of the multitude of species on the demersal fishing grounds is the occurrence of a multitude of species in the catch. This statement is not as trivial as it sounds, since it implies that no selective fishing is attempted for any given species or group of species. So, in many tropical multispecies stocks, the closest one gets to any single species fishery is by "shrimping", with subsequent discard of most of the (fish) catch (see Alverson et al. 1994).

Alternatively, the predominance of small-sized fishes on the fishing grounds also forces the fishery to use very fine-meshed gear, so as to catch the large valuable fishes as well as the less valuable small fishes which contribute to the value of the catch by sheer bulk.

Further, the occurrence of the largest part of the stock in shallow waters has two important consequences for the fishery. First, it is possible for a large number of small-scale fisheries, operating even with low efficiency in very shallow waters, to significantly reduce the stock, even if mainly by impairing recruitment to the deeper, offshore component of these stocks.

Second, the commercial fisheries are more or less forced to operate in shallow waters and thus to compete with the small-scale fisheries for the same resource. (It should be noted, however, with respect to points made in this and the preceding paragraphs that there is probably a substantial self-reinforcing component at work. As the trawl fisheries developed, the average size of individual fish decreased, as did their abundance. Thus, to maintain catch rates, the fishers decreased mesh sizes and moved into other fishing grounds, including the very shallow inshore areas.)

As the reproductive stages of most fishes are in reach of the commercial and especially the artisanal fisheries, and as both types of fishery will catch fish of any size from a few centimeters upward, there is a marked tendency for the catch in tropical demersal fisheries to consist to a significant extent of the juveniles for the large valuable fishes. This feature is likely to affect recruitment to the adult stock whenever the spawning stock has been significantly reduced. Therefore, in these demersal fisheries, there is the likely possibility that "recruitment overfishing" will occur, in addition to the "growth overfishing" induced by the small meshes in use. (These terms are defined further below.)

Finally, the fact that the stocks are composed of an assemblage of species with a very long, common evolutionary history has the grave implication that any fishery, by removing specific prey fishes, may disrupt and eventually destroy the original food web and lead to the emergence in the system of generalists, often of less value. The implications for biodiversity are obvious.

**Problems related to statistical data**

For statistics, the multitude of species is again the predominant problem. In the statistics of many countries this species multitude is summarily dealt with and reduced to its simplest expression, namely "various sea fishes". This greatly reduces the usefulness of these statistics for fisheries management. Some crude differentiation is often made, however, and it frequently pertains to value: we often find "good fish", marketed whole, iced, and used for human consumption, and "trash fish", used for duck or other animal

---

5 Adapted from Pauly (1979b).
feed and consisting of three different categories: the young of highly valuable fishes such as the Lutjanidae; smaller-sized fishes (such as the Leioognathidae), which in the unexploited stock form the bulk of the food of the large valuable fishes; and real trash fish—that is, fishes not used for direct human consumption and not forming a significant part of the food of the larger valuable fishes. Fishes in this last category, represented by such families as the Triacanthidae, Aluteridae and Ostracionidae, tend to increase as the biomass of the fishes of the first two categories is seriously reduced. Moreover, because of the simultaneous existence of two fisheries, one commercial and one artisanal, the latter using a multitude of different kinds of gear, each with a different "power factor", in most cases it is extremely difficult to obtain, for any given stock, a series of mutually compatible effort data against which the catch per effort could be plotted.

Fishery research problems

Fisheries research, which ideally should provide the basis for sound management, is faced in the case of tropical multispecies stocks with a series of practical, theoretical, and institutional problems which have greatly hampered its development and which, in most cases, have altogether prevented an understanding of the dynamics of the stocks that were being investigated.

There have been four main problems. First, perhaps, was that associated with properly identifying and naming the fishes which contributed to the fisheries in question. With the completion of several series of the FAO identification sheets and of revisions for some important families, and the release of databases such as FishBase 96 (Froese and Pauly 1996), this taxonomic problem seems now to have been largely overcome (see also p. 5). The problem remains, however, that many of the identification aids are not readily available in various local languages such as to permit their use at all levels in all countries with multispecies fisheries.

Second, poor species identification has been a major reason why fisheries scientists have been deprived of a body of data sufficient for their needs, gained from the fisheries statistics of their countries. This is similar to the third problem: in high-latitude countries, the statistical services linked with the commercial fisheries tend to generate, at little added cost, a tremendous body of data which is available to fisheries scientists; this additional source of information is absent in most tropical fisheries.

A final problem gravely affecting the development of fisheries biology as it relates to multispecies stocks is the continuing dependence of scientists in tropical countries on methods, concepts, theories, and even expertise from high-latitude countries, often with little or no attempt to really adapt the imported concept or theory to tropical situations.

Therefore, in the following chapters we shall emphasize the tropical context and its implications when we present the various stock assessment models.

Chapter 2. Overview of Fish Stock Assessment Models

What you will learn from this chapter

---

6 Adapted from Pauly (1979b).
In this chapter, you will be introduced to analytical and holistic models for fish stock assessment, and length-based methodologies incorporated in the FISAT software.

Introduction

The structure of most of the models for stock assessment discussed in this manual is defined by three components:

- the input (the fishing effort);
- the output (the fish landed);
- the processes which link input and output (the biological processes and the fishing operations).

The tools to describe the processes linking input and output are called models. Thus, a model is defined here as a series of instructions on how to perform calculations, and is constructed on the basis of what we can observe or measure, such as for example fishing effort and landings.

The actual processes, leading from a certain number of days fishing with a certain number of boats to a certain number of fish being landed, are quite complicated. However, the basic principles are usually well understood, so that by processing suitable input data in suitable models, we can predict the output.

A model is good if it can predict the output with a reasonable precision. However, since models are simplifications of an ever-changing reality, they will rarely generate exact predictions.

Fish stock assessment models may be classified as either holistic or analytical in nature. Usually, holistic models use fewer parameters than analytical models. They consider a fish stock as a homogeneous biomass and do not take into account, for example, its length or age composition. The analytical models, on the other hand, are based on a more detailed description of the stock and are more demanding in terms of quality and quantity of input data. However, they also allow a better definition of available options, and also tend to give more reliable estimates.

The type of model to be used in a given situation depends on the quality and quantity of the inputs. If data are available for an advanced analytical model, then such a model should be used; the simple models should be reserved for situations where data are limited. Often a complete set of input data suitable for an analytical approach is not available, but the available data exceed the demand of the simple models. As an alternative to using simple models in this case, the missing input data can be replaced by assumptions or qualified guesswork. (Often, the missing parameter for a particular stock can be replaced by known parameter values from another, similar stock.)

Analytical models

A basic feature of analytical models as developed by, among others, Baranov (1914), Thompson and Bell (1934) and Beverton and Holt (1957), is that they require the age composition of catches to be known. For example, the number of one-year-old fish caught, the number of two-year-old fish caught, etc., may form the input data.

The basic ideas behind the analytical models may be expressed as follows:

- If there are "too few old fish" the stock is overfished and the fishing pressure on the stock should be reduced.

---

7 Adapted from Pauly (1984a) and Sparre and Venema (1992).

8 Adapted from Sparre and Venema (1992).
If there are "very many old fish" the stock is underfished and more fish may be caught as the yield is increased.

Thus, analytical models are age-structured and require estimates of growth and mortality rates.

The basic concept in age-structured models is that of a cohort. For example, a cohort of the threadfin bream, *Nemipterus marginatus* could be all the fish of that species that hatched from June to August in 1976 near Tanjung Pinang in the South China Sea. Suppose there were one million specimens in that cohort. After August 1976 the original one million fish (Fig. 2.1A) would decrease in number because of deaths due to natural causes (predation, disease, etc.) or fishing. However, while the number of survivors of the cohort decreases with time, the average individual body length (Fig. 2.1B) and body weight (Fig. 2.1C) increase.

The curve resulting from combining the number of survivors with their mean body weight is the biomass curve (Fig. 2.1D), which displays a clear maximum. Thus, to realize the (hypothetical) maximum yield in weight from that cohort all fish should be caught exactly when the cohort has reached the age corresponding to that maximum. This, of course, is not possible in practice. However, the goal of fish stock assessment is to manage fisheries in such a way that catches come as close as possible to this theoretical maximum.

The implication is that the fish should be caught neither too young, nor too old. If the fish are caught too young there is growth overfishing of the stock; conversely growth underfishing occurs when the fish are caught too old.

There are thus two major elements describing the dynamics of a cohort:

- the increase of length and weight, due to growth;
- the decline of the number of survivors, due to fishing and/or natural processes.

When working with analytical models using length frequencies, another concept that needs a clear definition is that of age. As was noted earlier in connection with body length, we do not operate at the level of individual fish, so age implies the average age of a cohort. To define age, we must start with a
definition of day of birth. The obvious biological definition is the day larvae hatch from the egg (if applicable).

In the first part of their life, the larvae (or juveniles) are not affected by the fisheries. We say that the fish are then in the unexploited phase.

Let $t_r$ be the youngest age at which the young fish reach the fishing grounds. The fish of age $t_r$ are called recruits. By recruitment, we mean the number of fish that have attained age $t_r$ during a recruitment season. The recruitment intensity is the number of recruits per unit time. The recruitment pattern of a typical temperate species is as shown in Fig. 2.2A. In most tropical fish stocks, recruitment may occur throughout the year; however, seasonal peaks will occur, for example in conjunction with the monsoons (Pauly and Navaluna 1983; see Fig. 2.2B).

**Holistic models**

In data-sparse situations, for example when initiating the exploitation of a hitherto unexploited resource or in cases of limited sampling capability, one solution would be to start up the collection of the data types required for the analytical approach and then wait until a sufficient amount of data is available. This approach might of course solve the problem in the long run. However, it might take years for an analysis to emerge, while advice on an exploitation or development strategy might be required on short notice.

---

9 Adapted from Sparre and Venema (1992).
Other methods worth mentioning here are the surplus-production models briefly discussed in the next section. These use catch per unit of effort (for example kg of fish caught per hour trawling) as input. The data usually represent a time series of years and usually stem from commercial fishery samples. The models are based on the assumption that the biomass of fish in the sea is proportional to catch/effort. An estimate of the yield is obtained by multiplying effort by catch/effort, and maximum sustainable yield (MSY) is defined as the maximum of the resulting curve (Fig. 2.3B).

![Fig. 2.3. The simple Schaefer model: A: the logistic curve and its first derivative; B: the yield–biomass and the yield–effort relationship.](image)

**Overview of surplus-production models**

10 Adapted from Pauly (1984a).

Based on earlier work by Baranov (1927), Graham (1935) and others, Schaefer (1954, 1957) presented a model which, in its standard formulation (Ricker 1975) can be used for stock assessment when a minimum of data is available (only catch and effort data are required) and which has been applied, with varying success, to a number of fisheries throughout the world.

The assumptions made for the derivation of this model are as follows:

- Any fish population newly colonizing a given, finite ecosystem grows in weight until it approaches the maximum carrying capacity (most often in terms of available food) of this ecosystem, after which its increase in total weight gradually ceases as the stock size comes closer (asymptotically) to the carrying capacity of the environment (B$_{m}$).
- B$_{m}$ more or less corresponds to the unfished biomass (B$_{0}$).
- The growth, in time, of the fish biomass toward B$_{m}$ may be described by a logistic curve (Fig. 2.3A), the first derivative of which, dB/dt, has a maximum at B$_{m}$/2 and zero values at B$_{m}$ and B = 0 (Fig. 2.3B).
- The fishing effort which reduces B$_{m}$ to half its original value will produce the highest net growth of the stock, that is the maximum *surplus yield* available to a fishery (Fig. 2.3B).
- This maximum surplus yield can be sustained indefinitely (hence, the term *maximum sustainable yield*), as long as the biomass of the exploited stock is maintained at B$_{m}$/2.
There is biological evidence to make these assumptions appear reasonable (Odum 1971; Silliman and Gutsell 1958). Some reasons for the low surplus production at stock size \( B > B_x/2 \) are given here (from Ricker 1975):

- **Near maximum stock density, efficiency of reproduction, and often the actual number of recruits, is less than at smaller densities. In the latter event, reducing the stock will increase recruitment.**

- **When food supply is limited, food is less efficiently converted to fish flesh by a large stock than by a smaller one. Each fish of the larger stock gets less food individually; hence, a larger fraction is used merely to maintain life, and a smaller fraction for growth.**

- **An unfished stock tends to contain older individuals, relatively, than a fished stock. This makes for decreased production, in at least two ways; a) Larger fish tend to eat larger foods, so an extra step may be inserted in the food pyramid, with consequent loss of efficiency of utilization of the basic food production, and b) Older fish convert a smaller fraction of the food they eat into new flesh—partly, at least, because mature fish annually divert much substance to maturing eggs and milt.**

From the assumptions listed above, two very important features of the Schaefer and related models follow—namely, that the growth of a stock is a function of its size and of its size only, and that, therefore, a stock should respond by changes in its growth rate \( (dB/dt) \) instantaneously to any change of its size (e.g. by fishing). This may be expressed as:

\[
(dB/dt) = \left( r_m B (B_x - B) \right) / B_x ...2.1)
\]

where \( B \) is the stock size, \( B_x \) is the carrying capacity of the environment, \( r_m \) is the intrinsic rate of growth of the stock in question.

However, in demersal fisheries, especially in the tropics, the catch tends to consist of a multitude of species for which individual assessments are often impossible or inappropriate. It has been a common practice to treat the various fish of tropical and other multispecies stocks as a single entity, applying a model to the total multispecies catch of these fisheries (FAO 1978, Pope 1979, Pauly 1979b).

Crucial to surplus production models is the availability of long time-series of catch-and-effort data; it is worthwhile to derive such series in an ongoing fishery. Where possible, one should also attempt to reconstruct time-series of total mortality (for example from length–frequency data) for use with available time-series of catch.

Ricker (1975) gives a good account of the historical development of surplus yield models, but it is best to read also some of the original papers on the topic, notably those by Graham (1949), Schaefer (1954, 1957), Silliman and Gutsell (1958), Schaefer and Beverton (1963), Gulland (1969) and Schnute (1977). These models are not further discussed—except for passing references—in this manual.

### Flow diagrams in fish stock assessment

Fish stock assessments require a measure of biological time. Though the measure of biological time has traditionally been (absolute) age, size (in terms of length and weight) is also often used instead (see contributions in Pauly and Morgan 1987). Since management advice, as the ultimate goal of any assessment, should be framed in terms of a time unit (e.g. in years, seasons, etc.), age-structured models are most
appropriate for management. However, as discussed earlier, ages may be difficult and expensive to estimate in tropical fishes.

Thus, for tropical situations, it has become necessary to refine approaches based on the analysis of length–frequency data, which are cheap and relatively easy to obtain and through which growth parameters can be obtained which bridge the gap between sizes and ages. With regard to the estimation of growth parameters, two school of “thoughts” have evolved. One of these (see Fig. 2.4) proposes that growth parameter estimations should be based directly on the length composition of the stock, without first translating the lengths into ages. Another school suggests that lengths should be translated to (relative) ages (see Fig. 2.5). These two approaches are further detailed below.

**Estimating growth parameters directly from length compositions**

Fig. 2.4 summarizes the routines which, within FISAT, embody what was previously known as the "ELEFAN approach" implemented in various version of the ELEFAN package (Brey and Pauly 1986, Gayanilo et al. 1989). The contributions in Pauly and Morgan (1987), Basson et al. (1988) and Isaac (1990) may be consulted for reviews of the precision and accuracy of the methods included in Fig. 2.4.

A few points need to be emphasized here:

- The analytic path taken [(a), (b), (c) or (d), with or without (e) or (f)] will depend largely on the nature of the data and the assumptions made in the derivation of the model (e.g. path (d) is taken only when growth parameter estimates exist).
Path (h) should be considered when the available data do not enable reliable estimates or do not fulfill the assumptions of the models in the flow diagram.

Most length–frequency data sets are biased by selection, and should be corrected (see (i) and (j)) to obtain more accurate parameter estimates.

Converting length data to age information prior to estimation of growth parameters

It is often possible to convert length composition data to age composition information then to proceed using age-structured methods. Fig. 2.5, based on what may be described as “the LFSA approach”, represents a sequence of steps which include the prior decomposition of normal distributions using Bhattacharya’s method (or Hasselblad’s NORMSEP or a combination of both). The means generated by these methods are then linked to form either a cohort (with growth parameter estimation by non-linear fitting of the von Bertalanffy function) or growth increment data, which may then be analyzed using the Gulland and Holt plot and/or other methods for the analysis of such data (see Chapter 4).

Length-based fish stock assessments

Fig. 2.6 summarizes a suggested flow of analyses for performing length-based fish stock assessment.

Following the estimation of growth parameters (for which the approaches in both Fig. 2.4 and Fig. 2.5 should be used), we suggest that natural mortality should be estimated. Then, either of two routes may be followed leading (a) to a Beverton and Holt type of yield per recruit model, or (b) to a Thompson and Bell type of analysis. Both of these lead to biologically-based long-term management advice, which may be modified in the light of socioeconomic considerations, not discussed in this manual.
Chapter 3. Data Handling for Fish Stock Assessment

What you will learn from this chapter

In this chapter, you will be introduced to the basic concepts of sampling data for fish stock assessment and some techniques for handling fishery data.

Introduction

The quality of information derived from a given amount of data can be quite variable. It is a function of the biology of the species in question, the place and time of study, the models used by the investigator, his or her skills, and—last but not least—the purpose of the assessment. Conversely, the data “fed” to a model influence the quality of the output and hence the quality of management advice. It is therefore necessary to examine how samples are best handled prior to their being used as inputs to the models presented in this manual.

Length measurements

The term “length measurement”, in this manual, refers to the mean length of some body part required for growth studies; it is not important what is measured as long as length increases can be interpreted as increases of age. However, it is common practice to use total length unless some anatomical features makes this impractical (see Fig. 3.1). Fork length may be used for fish with stiff caudal fins (e.g. tunas) or caudal filaments

---

11 Adapted from Sparre and Venema (1992).
Standard length, much used by museum workers, is not recommended for length–frequency sampling due to the excessive time required for its measurement in the field.

In fisheries research, it is important to always specify what kind of length measurement has been used, as one may otherwise run into difficulties when comparing results with those of other investigations.

Other examples given in Fig. 3.1 are squid, abalone, scallop and sea cucumber. For animals with a hard shell or skeleton, it is not a problem to define a suitable length measure. It may, in certain cases, be preferable to work with body weight rather than length, as the former is obviously measurable with greater accuracy; but these data will need to be converted to a linear measure as most of the models included in this manual, as well as in the FISAT software, are length-based. (One approach is to treat the cube root of the weight, i.e. \( W^{1/3} \), as equivalent to length.)

Note that for rapid field measurements, length is measured to the unit below, i.e. is defined by the lower limit of the length class to which it belongs; in this case, half a class interval must be subsequently added to obtain the midpoint of a class interval, here called midlength (ML).

**Age readings vs. length measurements**

As stated earlier, age reading is a simple technique in the case of species from temperate or subtropical waters, because their otoliths or scales often show seasonal markings, one for the summer and one for the winter, which together form an annual ring (Panella 1971, Bagenal 1974, Beamish and McFarlane 1983). The annual rings give sufficient information for most stock assessment purposes.

Tropical fish, on the other hand, rarely show clear rings in their otoliths or scales, because the strong seasonality which characterizes colder areas is lacking. Recent discoveries,

---

12 Adapted from Sparre and Venema (1992).
however, have created opportunities to also read ages in such fish, albeit within limited ranges and at a high cost in terms of personnel and capital investment. In-depth study into the formation of the rings in otoliths and scales has revealed that daily increments can be detected by means of a strong microscope. However, the findings indicate that the daily rings are sometimes so thin that they defeat ordinary microscopes, whose detection power is limited by the wavelength of light; such rings can be read only by a scanning electron microscope (Panella 1971, Williams 1986, Brother 1980, Gjøsæter et al., 1984, Morales-Nin 1988a, 1988b).

In tropical fish stock assessment, it is therefore necessary to emphasize length measurements and put less emphasis on age data. However, it may be important to deal with age data for two reasons. First, it may be possible to carry out a small number of age readings to calibrate the findings obtained from length measurements. Second, it is often easier to explain stock assessment concepts and theories on the basis of age and length data than on the basis of length data only.

Time-series of length frequencies are the most common data type collected for tropical fish stock assessment. Data files of this form can either be grouped with a constant class size (i.e. class interval; see Table 3.1), grouped with variable class sizes (common for data from commercial sampling of shrimp) or collected as individual measurements, requiring subsequent grouping. In the FiSAT software, all of these data types can be converted to length frequencies grouped with constant class size, which is the input to most of the routines in the software.

### Table 3.1. Sample length frequencies measured in cm for *Upeneus moluccensis* (goat fish; mean annual catch of 25,000 t) from Ragay Gulf, Philippines in 1980.

<table>
<thead>
<tr>
<th>Midlength (TL)</th>
<th>20/01</th>
<th>02/03</th>
<th>21/05</th>
<th>23/07</th>
<th>27/09</th>
<th>20/11</th>
<th>11/12</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.5</td>
<td>4</td>
<td></td>
<td></td>
<td>9</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.5</td>
<td>22</td>
<td>3</td>
<td></td>
<td>9</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.5</td>
<td>13</td>
<td>10</td>
<td>6</td>
<td>19</td>
<td>4</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>10.5</td>
<td>105</td>
<td>77</td>
<td>26</td>
<td>26</td>
<td>13</td>
<td>108</td>
<td>33</td>
</tr>
<tr>
<td>11.5</td>
<td>103</td>
<td>198</td>
<td>423</td>
<td>61</td>
<td>37</td>
<td>245</td>
<td>42</td>
</tr>
<tr>
<td>12.5</td>
<td>85</td>
<td>62</td>
<td>415</td>
<td>148</td>
<td>44</td>
<td>235</td>
<td>101</td>
</tr>
<tr>
<td>13.5</td>
<td>68</td>
<td>58</td>
<td>289</td>
<td>130</td>
<td>115</td>
<td>133</td>
<td>97</td>
</tr>
<tr>
<td>14.5</td>
<td>53</td>
<td>24</td>
<td>185</td>
<td>75</td>
<td>80</td>
<td>107</td>
<td>42</td>
</tr>
<tr>
<td>15.5</td>
<td>15</td>
<td>40</td>
<td>32</td>
<td>38</td>
<td>30</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>16.5</td>
<td>3</td>
<td>36</td>
<td>21</td>
<td>8</td>
<td>10</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>17.5</td>
<td>3</td>
<td></td>
<td>3</td>
<td></td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>18.5</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td>473</td>
<td>419</td>
<td>1,430</td>
<td>499</td>
<td>354</td>
<td>897</td>
<td>407</td>
</tr>
<tr>
<td><strong>Sample weight (g)</strong></td>
<td>1,050</td>
<td>710</td>
<td>3,100</td>
<td>1,150</td>
<td>850</td>
<td>1,600</td>
<td>1,020</td>
</tr>
<tr>
<td><strong>Total catch (t)</strong></td>
<td>2,700</td>
<td>1,660</td>
<td>2,900</td>
<td>1,950</td>
<td>1,500</td>
<td>2,400</td>
<td>2,510</td>
</tr>
</tbody>
</table>

An age/length key is a table showing, for each length class of fish in a particular stock, its percentage (or fractional) age frequency distribution. Once such a key is available, samples of fish which were measured for length only can be distributed over age groups according to the key. Moreover, such a key can also be used to estimate the age composition of the entire catch from a stock, based on a small sample; then we can restrict further sampling to the collection of length–frequency data. The same key may be used in consecutive years if the assumption that no major changes in the age composition of the stock is met. However, in periods of strongly increasing effort, when the old fish may disappear from catches, new age/length keys must be frequently prepared (see Sparre and Venema 1992 for details).
Length–weight relationships

Introduction

The relationship between the length (L) and the weight (W) of fish can generally be expressed by the equation:

\[ W = aL^b \]  

where “a” is a factor discussed below and the exponent “b” lies between 2.5 and 3.5 (usually close to 3). Carlander (1969, 1977) has demonstrated from an extraordinarily large number of length–weight data, derived from a wide variety of fishes, that values of \( b < 2.5 \) or \( b > 3.5 \) are generally based on a very small range of sizes and/or that such values of \( b \) are most likely erroneous. When \( b=3 \), weight growth is called isometric, meaning that it proceeds in the “same” dimension as the cube of length (\( L^3 \)). When \( b \neq 3 \), weight growth is allometric, meaning that it proceeds in a “different” dimension (differing from \( L^3 \)). Allometric growth can be either positive (\( b>3 \)) or negative (\( b<3 \)). Another way of relating length and weight is to define a condition factor (c.f.) such that

\[ \text{c.f.} = \frac{W}{100} \frac{1}{L^3} \]  

When weight growth is isometric (\( b=3 \)), we also have

\[ \frac{\text{c.f.}}{100} = a \]  

where “a” is the multiplicative factor in Equation 3.1. The reason for the multiplication by 100 in Equation 3.2 is to bring the value of the condition factor of fishes with a “normal” shape close to unity when grams are used to express the weight, and centimeters to express the length. It must be emphasized, however, that the condition factor in a given fish species or

---

13 Adapted from Pauly (1984a).
stock can be compared to that of another species or stock only if the same units and definitions have been used (e.g. total length in cm and live or gutted weight in g). The units and definitions must always be stated.

In addition many variables, such as sex, time of year, stage of maturity and stomach contents, influence the numerical magnitude of the condition factor. Comparisons should be made only when these variables are roughly equivalent among samples to be compared.

The values of a in Equation 3.1, on the other hand, cannot be used for interspecies or interstock comparisons, even when the same units and definitions are used, unless the values of b are exactly the same. The values of b, finally, are not affected by the units or definitions used.

**Parameter estimation**

The a and b coefficients of the length–weight relationships can be estimated either from pairs of length–weight data, from the mean weights of small class intervals, or from length–frequency samples and their corresponding sample weights.

**From individual pairs of length–weight data**

When pairs of length–weight data are available (as in Table 3.2), the values of a and b in Equation 3.1 can be computed by means of “linearized” form of the equation, namely

\[ \log W = \log a + b \log L \]  

...3.4

that is by taking (base 10) logarithms (natural logarithms may also be used) on both sides and by estimating the values of \( \log a \) and \( b \) by means of a linear regression.

This procedure of using ordinary least-square regression to estimate a and b only approximates these parameters, and results in estimates of the standard errors that are not very reliable; alternative procedures, such as the use of non-linear least squares estimations, should be considered where possible. The disk distributed with this manual contains a software package, ABee, which contain a routine to fit Equation 3.1 using a least squares estimation approach (see Fig. 3.3) as well as the linearized approach, also available in FiSAT.

<table>
<thead>
<tr>
<th>Obs. No.</th>
<th>TL (cm)</th>
<th>W (g)</th>
<th>Obs. No.</th>
<th>TL (cm)</th>
<th>W (g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.1</td>
<td>6.3</td>
<td>9</td>
<td>16.6</td>
<td>65.5</td>
</tr>
<tr>
<td>2</td>
<td>8.1</td>
<td>9.6</td>
<td>10</td>
<td>17.7</td>
<td>69.4</td>
</tr>
<tr>
<td>3</td>
<td>10.2</td>
<td>11.6</td>
<td>11</td>
<td>18.7</td>
<td>76.4</td>
</tr>
<tr>
<td>4</td>
<td>11.9</td>
<td>18.5</td>
<td>12</td>
<td>19.0</td>
<td>82.5</td>
</tr>
<tr>
<td>5</td>
<td>12.2</td>
<td>26.2</td>
<td>13</td>
<td>20.6</td>
<td>106.6</td>
</tr>
<tr>
<td>6</td>
<td>13.3</td>
<td>36.1</td>
<td>14</td>
<td>21.9</td>
<td>119.8</td>
</tr>
<tr>
<td>7</td>
<td>14.8</td>
<td>40.1</td>
<td>15</td>
<td>22.9</td>
<td>169.8</td>
</tr>
<tr>
<td>8</td>
<td>15.7</td>
<td>47.3</td>
<td>16</td>
<td>23.5</td>
<td>173.3</td>
</tr>
</tbody>
</table>

This table (Table 3.2) gives individual length–weight data pairs used for establishing a length–weight relationship for the threadfin bream (Nemipterus marginatus) from the southern tip of the South China Sea (live weight in grams).
When large numbers of fish have been measured, entering L/W data pairs can become tedious. In such cases, a common practice is to arrange the data by (relatively narrow) length groups, and to calculate the mean weight for each class (see Table 3.3). (Grouped data of this sort may also result from field sampling, where individual fish are measured, grouped and weighted by length class, and the mean weights plotted against the class midlengths.)

Table 3.3. Data for establishing the length–weight relationship of splendid pony fish (*Leiognathus splendens*) from the Eastern Java Sea (total length in cm, live weight in g).

<table>
<thead>
<tr>
<th>Class limits Low</th>
<th>Class limits High</th>
<th>Class midlength</th>
<th>Mean weight</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.00</td>
<td>6.49...</td>
<td>6.25</td>
<td>5.28</td>
<td>1</td>
</tr>
<tr>
<td>6.50</td>
<td>6.99...</td>
<td>6.75</td>
<td>4.07</td>
<td>1</td>
</tr>
<tr>
<td>7.00</td>
<td>7.49...</td>
<td>7.25</td>
<td>6.91</td>
<td>11</td>
</tr>
<tr>
<td>7.50</td>
<td>7.99...</td>
<td>7.75</td>
<td>8.46</td>
<td>26</td>
</tr>
<tr>
<td>8.00</td>
<td>8.49...</td>
<td>8.25</td>
<td>10.15</td>
<td>26</td>
</tr>
<tr>
<td>8.50</td>
<td>8.99...</td>
<td>8.75</td>
<td>11.68</td>
<td>23</td>
</tr>
<tr>
<td>9.00</td>
<td>9.49...</td>
<td>9.25</td>
<td>13.77</td>
<td>16</td>
</tr>
<tr>
<td>9.50</td>
<td>9.99...</td>
<td>9.75</td>
<td>17.13</td>
<td>2</td>
</tr>
<tr>
<td>10.00</td>
<td>10.49...</td>
<td>10.25</td>
<td>19.29</td>
<td>7</td>
</tr>
<tr>
<td>10.50</td>
<td>10.99...</td>
<td>10.75</td>
<td>22.57</td>
<td>9</td>
</tr>
<tr>
<td>11.00</td>
<td>11.49...</td>
<td>11.25</td>
<td>25.84</td>
<td>11</td>
</tr>
<tr>
<td>11.50</td>
<td>11.99...</td>
<td>11.75</td>
<td>28.66</td>
<td>3</td>
</tr>
<tr>
<td>12.00</td>
<td>12.49...</td>
<td>12.25</td>
<td>34.02</td>
<td>7</td>
</tr>
<tr>
<td>12.50</td>
<td>12.99...</td>
<td>12.75</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>13.00</td>
<td>13.49...</td>
<td>13.25</td>
<td>46.73</td>
<td>1</td>
</tr>
<tr>
<td>13.50</td>
<td>13.99...</td>
<td>13.75</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>14.00</td>
<td>14.49...</td>
<td>14.25</td>
<td>55.91</td>
<td>1</td>
</tr>
<tr>
<td>14.50</td>
<td>14.99...</td>
<td>14.75</td>
<td>65.63</td>
<td>1</td>
</tr>
<tr>
<td>15.00</td>
<td>15.49...</td>
<td>15.25</td>
<td>61.72</td>
<td>1</td>
</tr>
</tbody>
</table>

Using Equation 3.2, the length–weight relationship and/or the mean condition factor may be calculated with or without weighting the L/W data pairs by the sample size.

During demersal trawl and other sea-going surveys (or even during port-sampling), there is often not enough time to fully analyze the catch of one station before the catch of the next station is hauled in. This results in aggregated data, i.e. samples of fish that have been measured individually, leading to length–frequency samples (L/F), but not weighted individually. Thus only the bulk weights of the L/F are available (accurate shipborn weighing of small fishes is usually not possible anyway). Table 3.1 gives an example of such data.

“Pseudoweights”, i.e. sample weights obtained using estimates of the parameters a and b of the length–weight relationship, are first estimated for such data. Estimating the pseudoweights of samples requires an accurate estimation of the mean weight ($\overline{w}_i$) of the fish within a given length class (i), which is not equal to the weight corresponding to the midpoint of that length class, or midlength (remember: the relation between weight and length is nonlinear; see also Lenarz 1994). For this, we use the equation of Beyer (1987):

$$\overline{w}_i = \left( \frac{1}{L_{i+1}-L_i} \right) \left( a + \frac{b}{b+1} \right) \left( L_{i+1} - L_i \right)$$

where $a$ and $b$ are as defined in Equation 3.1, and which provides unbiased estimates of $\overline{w}_i$.

The pseudoweight ($W'_j$) of a given sample (j) can then be estimated from

$$W'_j = \sum_{i=1}^{n_j} \left( \overline{w}_i \cdot f_i \right)$$

where

- $\overline{w}_i$ is the mean weight of class i,
- $f_i$ is the frequency of class i in sample j, and
- $n_j$ is the number of classes in sample j

---

14 Adapted from Pauly and Gayanilo (1996).
When a number (≥3) of length–frequency samples and their bulk weight are available, a and b can be estimated iteratively, starting with arbitrary seed values of each (e.g. a = 0.01 and b = 3), and using a non-linear least square procedure (here: Marquardt’s compromise algorithm (Marquardt 1963) which minimizes the sum of the squared differences (SSE) between the sample weights (W_j) and the pseudoweights (W'_j), both previously log-transformed to stabilize the variance, or

\[ SSE = \sum [\log(W_j) - \log(W'_j)]^2 \]  

...3.7

The results of the final iteration can be shown by plotting the sample pseudoweights against the observed sample weights; this leads to a graph such as in Fig. 3.4 which can be used to identify outliers, and tag them for deletion.

![Fig. 3.4. ABeef plot, showing estimation of parameters of a length–weight relationship from length–frequency samples of Upeneus moluccensis and their bulk weight (based on the data in Table 3.1).](image)

These steps are all quickly performed by the ABeef software, which also outputs the standard errors for all parameters estimated.

### Sampling from commercial catches15

Data for estimation of population parameters may also be obtained by sampling commercial catches—or more precisely, commercial landings. The basic principles for analyzing such samples are the same as for research survey data. The major difference lies in the bias problems: commercial boats never attempt to collect a random sample of the stock, because they always go for the marketable sizes and try to find the areas with the highest concentrations of fish. However, if one keeps in mind the sources of bias and tries to stratify sampling to minimize the bias, one can use data from commercial fisheries to estimate population parameters.

The major advantage of sampling commercial landings is that samples are much cheaper to collect, and thus sampling can be much more frequent than is possible with a single research vessel.

One basic problem in sampling is that in principle we want a random sample. The landing of an individual fishing vessel is rarely a random sample of the population because it reflects factors such as the locality fished and the properties of the vessel. Moreover, landings are often sorted into market categories such as fish for direct human consumption and fish for industrial purposes. Therefore, fish of the same species may occur in more than one market category, and it requires careful sampling and analyses to be able to estimate the total amount landed and to reconstruct the length composition of a given species.

It is not possible to give a complete set of general rules for the best way to collect data. Each fishery has its own characteristics, which must be taken into account when planning a sampling scheme. However, some generalizations can be stated here. First, a decision must be made as to which

---

15 Adapted from Sparre and Venema (1992).
species should be investigated. This decision, of course, must be based on the economic or ecological importance of the various species (or species groups) and the resources (funds and personnel) available for the sampling program. There may be other important considerations worth exploring; for example, we may want to investigate the potential of a species which we believe is underexploited, and which could perhaps be developed into a fishery of economic importance. However, it is suggested that data collection be limited to the economically/ecologically most important species (though it must also include investigation of their juveniles, even if they are landed as by-catch or “trash” fish, or thrown overboard rather than landed).

To obtain an overview of the complex system from which landing data are generally obtained, a flow diagram should be constructed which takes into account the possibility that:

- several boats exploit a common stock;
- several landing sites exist;
- not all catches are landed at the sampling site;
- "trash" fish is either landed or thrown overboard.

However, given the many possibilities in commercial fishing operations, care must be taken when raising length frequencies to the catches. For instance, in cases where the size and capacity of the boats are very dissimilar or the gears used are different and not all catches are known, the number of boats in a fleet will not be a good measure of overall effort, and it becomes imperative to obtain complete catch data for all fleets (see Sparre and Venema 1992 for details).

**Underlying assumptions of random samples**

As mentioned earlier, the basic versions of all the methods presented in this manual assume the input data to be derived from random samples. A sample of fish, for example a length–frequency sample representing the stock, is a random sample if any fish in the entire stock has the same probability of being drawn as any other. Usually it is difficult or even impossible to obtain samples that are really random. If, for example, the juvenile fish are located in certain nursery areas which do not coincide with the fishing grounds from which most samples originate, the juvenile fish will be under-represented in the samples. A similar problem is created by the selectivity of fishing gears. Often the small fish are under-represented because they are not vulnerable to the gear, whereas the larger samples are. Samples which are not random are here referred to as biased samples.

The feature of fish behavior which is believed to create the most serious bias is migration. Almost all marine animals perform some sort of migration. Pelagic fish such as mackerels, scads and tunas undertake long migrations between their feeding and spawning grounds. Most penaeid shrimps start their life cycle as larvae in the open sea, drift to shallow waters, e.g. lagoons and mangroves, then migrate as juveniles back to the open sea, where the adults reproduce, thus closing the cycle.

The implication of such migratory behavior is that a large area must be covered in order to obtain random samples from the entire population. Often samples can be obtained only from the fishery which concentrates on the grounds where the resources are easiest to catch in large quantities. Thus, we are often in the situation that random samples of the population are not available, and biases must be accounted for in the analyses. Some types of bias are easier to deal with than others, but bias created by migration can be handled properly only when the migration routes are known (see Chapter 8).

---

16 Adapted from Sparre and Venema (1992).
Chapter 4. Analysis of Growth

What you will learn from this chapter

In this chapter, you will learn about the models that can be used to estimate directly or indirectly the parameters of the von Bertalanffy growth function from length frequencies, size-at-age data or growth increment data. The growth performance index $\varphi$ and related concepts will also be introduced.

Introduction

Growth may be defined as the change over time of the body mass (= body weight) of a fish, being the net result of two processes with opposite tendencies, one the building-up of body substances (anabolism) and the other the breaking-down of these substances (catabolism), or

$$\frac{dW}{dt} = HW^\frac{d}{3} - kW$$ (4.1)

where $dW/dt$ is the change in body weight per unit time, $H$ is the coefficient of anabolism and $k$ is the coefficient of catabolism. The process of anabolism is here viewed as being proportional to a power ($d$) of fish weight ($W$), while catabolism is proportional to weight itself (Bertalanffy 1934, Pauly 1981).

Details on the integration of Equation 4.1 to a growth curve have been presented in Taylor (1962) and Pauly (1979a). It suffices to mention here that, in the course of this integration, it is assumed that

$$W \propto pL^a$$ (4.2)

and

$$W \propto qL^b$$ (4.3)

where $b = 3$, $a = 2$, and hence $d = 2/3$. This, following integration gives the equation defining the length ($L_t$) at age $t$

$$L_t = L_\infty \left[1 - e^{-K(t-t_0)}\right]$$ (4.4)

where

$L_\infty$ is the asymptotic length, that is the mean length the fish of a given stock would reach if they were to grow indefinitely;

$K$ is the rate (of dimension time$^{-1}$) at which $L_\infty$ is approached; and

$t_0$ is the "age of the fish at zero length" if they had always grown in the manner described by the equation (note that $t_0$ is generally negative).

Equation 4.4 is the VBGF (von Bertalanffy growth function; Bertalanffy 1934). The equation can also be expressed in terms of weight, as

$$W_t = W_\infty \left[1 - e^{-K(t-t_0)}\right]^3$$ (4.5)

where $W_\infty$ is the asymptotic weight and all other parameters are as defined above.

When, as in Equation 4.5, the exponent of the (implied) length–weight relationship set is equal to 3, fitting the equation to weight–growth data is the same as fitting it to length–growth data, except that the cubic root of all weight values is taken prior to all calculations, these cube root values being then

$^{17}$ Adapted from Pauly (1984a).
treated exactly as if they were length values. This is justified because, when weight growth is isometric:

\[ L \propto (W)^{1/3} \quad \text{...4.6} \]

More generally, weights can be rendered proportional to length by raising them to the inverse of the power of the length–weight relationship \( b \), or

\[ L \propto W^{1/b} \quad \text{...4.7} \]

which can be used when weight growth is either isometric or allometric to obtain growth parameters \((K, t_0)\) generally equivalent to those that would be obtained by fitting the corresponding length data. Note, however, that when empirical data are used, slight differences might still occur between values of \( K \) and \( t_0 \) computed from weight and length data, depending on sample sizes and method of fitting.

This is illustrated by the set of constructed data in Table 4.1, in which values of \( W_{\text{er}} \), \( L_{\text{er}} \), \( K \) and \( t_0 \) were obtained for data with the length–weight relationships \( W = a L^{2.5} \), \( W = a L^{3.0} \) and \( W = a L^{3.5} \). These data demonstrate that this property can be used when using FiSAT to fit weight data with the programs that fit length data, the only prior requirement being that the weight data must be converted to "length" using the appropriate routine.

### Table 4.1: Data (A) and results (B) of the comparison of growth parameters obtained from length-at-age and weight-at-age data.*

<table>
<thead>
<tr>
<th>Error</th>
<th>Age (year)</th>
<th>Length (cm)</th>
<th>Weight (in arbitrary units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( W_{\text{er}} )</td>
<td>1</td>
<td>15</td>
<td>871.4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>18</td>
<td>1,375</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>20</td>
<td>1,789</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>21</td>
<td>2,021</td>
</tr>
<tr>
<td>( L_{\text{er}} )</td>
<td>22.68</td>
<td>22.68</td>
<td>22.68</td>
</tr>
<tr>
<td>( K )</td>
<td>0.511</td>
<td>0.511</td>
<td>0.511</td>
</tr>
<tr>
<td>( t_0 )</td>
<td>-1.116</td>
<td>-1.116</td>
<td>-1.116</td>
</tr>
</tbody>
</table>

\*The length–weight relationship used here was of the form \( W = a \cdot L^b \), with "a" set equal to unity.

While the version of the VBGF presented above assumes constant environmental conditions, it has been demonstrated by various authors that fish live in an environment with seasonally oscillating features (temperature, food availability, etc.) and that their growth may also oscillate seasonally (Ursin 1963, Pitcher and Macdonald 1973, Lockwood 1974, Daget and Ecoutin 1976, Cloern and Nichols 1978, Pauly and Gaschütz 1979, Pauly et al. 1992). The model incorporated in FiSAT for expressing the seasonally oscillating length growth of fishes was presented by Somers (1988) and has the form

\[
L_t = L_0 \left[ 1 - c^{(K)(t-t_0)} \right] \sin(2\pi(t-t_0)) \quad \text{...4.8} \]

where \( L_0 \), \( K \), \( t_0 \), \( t \) and \( L_t \) are as defined above, while the new parameters \( C \) and \( t_s \) refer to the intensity of the (sinusoid) growth oscillations of the growth curve and the onset of the first oscillation relative to \( t = 0 \), respectively. It is a property of the parameter \( C \) that it takes a value of 1 when the growth rate \((dL/dt)\) has exactly one zero value per year (in which case Equation 4.8 reduces to Equation 4.4), and that it takes intermediate values when the seasonal oscillations are sufficient to reduce, but not to halt, growth in length (Fig. 4.1).

Various authors (Pauly 1984a, Gaschütz et al. 1980, Longhurst and Pauly 1987) have shown a strong correlation between \( C \) and \( \Delta T \) (= difference between the highest and lowest average mean temperature to which the fish are exposed in the course of a year, in °C), which suggests that:

- seasonal length growth oscillations are caused by temperature fluctuations or by another environmental parameter which itself correlates with temperature (e.g. food availability or feeding activity);
- slight seasonal fluctuations of temperature (or of correlated parameters) such as occur in the tropics are sufficient to generate seasonally oscillating growth.
curves, and a seasonally oscillating growth model should be used to fit growth data pertaining to intervals of less than one year (e.g., when fitting growth data obtained by reading daily otolith structures or when analyzing length–frequency data sampled at less than yearly intervals).

Using time series of L/F data

The methods conventionally used for the analysis of length–frequency data were introduced by Petersen (1892), and can be reduced to two basic techniques:

The “Petersen method” (sensu stricto) involves attribution of relative ages to the peaks of a length–frequency sample. Note that the time separating various peaks must be assumed, a difficult task in animals which may or may not spawn several times a year (Fig. 4.2A).

"Modal progression analysis" consists of linking up the peaks of length–frequency data sampled at known times \(t_1, t_2\) by means of growth increments (Fig. 4.2B).

With the Petersen method, the problem consists of identifying those peaks representing cohorts spawned at known or assumed time intervals. The method generally involves the separation of the length–frequency samples into normally or otherwise distributed subsets by graphical methods, such as those initially proposed by Harding (1949), Cassie (1954), Tanaka (1956) and Bhattacharya (1967), or by means of computer programs such as NORMSEP (Abrahamson 1971) or ENORMSEP (Yong and Skillman 1971).

---

18 Adapted from Pauly (1987).
"Modal progression analysis", on the other hand, has its major problems in the identification of those peaks which should be connected (by growth increments) with each other.

Both methods are subjective and in the following we shall discuss some of the techniques devised to make them less subjective, and their results reproducible.

**ELEFAN I**

ELEFAN I (Electronic Length Frequency Analysis) is a program for the estimation of growth parameters from sequential length–frequency data. The ELEFAN I approach attempts to combine the logic of the Petersen method and that of modal progression analysis with a minimum of subjective inputs.

Within ELEFAN I, Equation 4.8 is used with two of the original parameters replaced by others. First, $t_s$ is replaced by the "winter point" (WP; Equation 4.9), which designates the period of the year (expressed as a fraction of a year) when growth is slowest. The WP is related to $t_s$ through

$$t_s + 0.5 = \text{WP} \quad \ldots(4.9)$$

The other parameter of Equation 4.8 not used explicitly in ELEFAN I is $t_o$. (ELEFAN I uses another parameter to position a growth curve along the time axis.) One way this could be done is by assuming that the curve in question also describes the growth of very small fishes. In such cases, the parameter $t_o$ could be replaced by the parameter $t_o$, expressing the origin of the growth curve as fraction of a year. Alternatively, a growth curve can be positioned by identifying a point, here called starting point, through which the growth curve passes.

The use of starting points associated with middle-sized fishes avoids having to make the assumption that very small fishes grow as described by the VBGF (which they usually don’t). In ELEFAN I, starting points are defined by a sample number and a length, with both coordinates defining a size at which some non-zero frequencies are observed. Starting points may be fixed or variable. When a starting point is fixed, different growth curves (as defined by different growth parameter values) will be forced to pass through that same point. In the case of a variable starting point, growth curves are tested for several possible choices of starting points, and only that with the best fit is retained. Variable starting points require more computational time but can be expected to yield better fitting curves, and should be used as a general approach—except in those cases where, for specific reason(s), a curve is to be forced through a specific set of coordinates.

The key assumptions of ELEFAN I are:

- samples used are representative of the population;
- all length differences can be attributed to differences in age;
- growth is similar from one year to the next, i.e. there are no factors inducing any strong changes in growth over the years;
- the seasonally oscillating version of the VBGF (Equation 4.8) provides an appropriate description of the growth of fish and aquatic invertebrates.

In ELEFAN I, the goal is to search for a growth curve that will fit the peaks of the length–frequency data such that under given conditions (i.e. based on the data used), an optimum fit is achieved, much the same as in the computation of the parameters of a linear regression. However, while the criterion...
for "best fit" in linear regression is defined in rigorous mathematical terms, the method used in ELEFAN I to identify peaks and troughs in length–frequency histograms is a simple highpass filter, i.e. the application of a running average; this leads to identification of "peaks" as those parts of a length–frequency distribution that are above the corresponding running average and conversely of "troughs" separating peaks. The running average used is taken over five classes. Analysis suggests that the results of ELEFAN I are sensitive, but not strongly so, to changes in the number of classes over which the running average is computed.

Fitting of 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); "available points" are points which can possibly be accumulated by one single growth curve (see Fig. 4.3C). This sum is called the "available sum of peaks" (ASP).

- "Trace" through the (set of) length–frequency sample(s) sequentially arranged in time, for any arbitrary seed input of \(L_\infty\) and \(K\), a series of growth curves started from the base of each of the peaks, and projected backward and forward in time so as to meet all other samples of the sample set and/or the same sample, repeated as often as necessary (see below).

- Accumulate the points obtained by each growth curve when passing through peaks (positive points) or through the troughs separating peaks (negative points) (see Fig. 4.3C).

- 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 the "explained sum of peaks" (ESP).

- Decrement or increment the "seeded" values of \(L_\infty\) and \(K\) (and \(C\) and \(WP\) when growth is seasonal) until the

![Fig. 4.3. 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 are 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 troughs. 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 text). 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 text).](image)
goodness of fit index, \( R_n (=10^{\text{ESP/ASP}/10}) \), reaches an optimum (where \( 0 < R_n < 1 \)), and output the growth parameters corresponding to this optimum ratio.

Table 4.2 is a detailed explanation, using an example, of the data restructuring applied within ELEFAN I.

**Table 4.2.** Steps performed in restructuring length–frequency data (adapted from Saeger and Gayanilo 1986). See text for explanations.

<table>
<thead>
<tr>
<th>ML</th>
<th>Obs.</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>6</td>
<td>7.800</td>
<td>0.769</td>
<td>0.385</td>
<td>(2)</td>
<td>-0.385</td>
<td>-0.079</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>19</td>
<td>9.800</td>
<td>1.939</td>
<td>0.550</td>
<td>1</td>
<td>0.275</td>
<td>0.275</td>
<td>0.275</td>
</tr>
<tr>
<td>17</td>
<td>14</td>
<td>11.000</td>
<td>1.273</td>
<td>0.017</td>
<td>0</td>
<td>0.017</td>
<td>0.017</td>
<td>0.017</td>
</tr>
<tr>
<td>19</td>
<td>10</td>
<td>9.800</td>
<td>1.020</td>
<td>-0.184</td>
<td>(1)</td>
<td>-0.184</td>
<td>-0.038</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>6</td>
<td>6.000</td>
<td>1.000</td>
<td>-0.200</td>
<td>(2)</td>
<td>-0.200</td>
<td>-0.041</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>4.400</td>
<td>0.000</td>
<td>-1.000</td>
<td>(1)</td>
<td>-1.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>2.600</td>
<td>0.000</td>
<td>-1.000</td>
<td>(1)</td>
<td>-1.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>27</td>
<td>6</td>
<td>1.400</td>
<td>4.286</td>
<td>2.426</td>
<td>3</td>
<td>0.303</td>
<td>0.303</td>
<td>0.303</td>
</tr>
<tr>
<td>29</td>
<td>1</td>
<td>1.800</td>
<td>0.556</td>
<td>-0.556</td>
<td>(2)</td>
<td>-0.556</td>
<td>-0.114</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>0</td>
<td>2.000</td>
<td>0.000</td>
<td>-1.000</td>
<td>(0)</td>
<td>-1.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>33</td>
<td>2</td>
<td>1.000</td>
<td>2.500</td>
<td>0.999</td>
<td>2</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
</tr>
<tr>
<td>35</td>
<td>1</td>
<td>0.600</td>
<td>1.667</td>
<td>0.332</td>
<td>3</td>
<td>0.042</td>
<td>0.042</td>
<td>0.042</td>
</tr>
</tbody>
</table>

The successive steps for each sample are as follows (the capital letters A - G refer to the column heads in the table):

**A:** Computation of the moving average (ma) over 5 class intervals; the computation also considers 2 class intervals above and below the highest and lowest recorded length respectively, e.g.:

\[
\text{ma}_{13 \text{ cm}} = (0 + 0 + 6 + 19 + 14)/5 = 7.800
\]

\[
\text{ma}_{15 \text{ cm}} = (0 + 6 + 19 + 14 + 10)/5 = 9.800
\]

\[
\text{ma}_{33 \text{ cm}} = (1 + 1 + 2 + 1 + 0)/5 = 1.000
\]

\[
\text{ma}_{35 \text{ cm}} = (0 + 2 + 1 + 0 + 0)/5 = 0.600
\]

**B:** Computation of quotients \( N_i/ma \). Also, the mean quotient \((m')\) is computed using

\[
m' = (1/n) \sum(N_i/ma) ...
\]

where \( n \) = number of class intervals. In our example, \( m' \) is given by

\[
m' = 1/12((6/7.800) + (19/9.800) + ...+ (1/0.600))
\]

\[
= 15.010 / 12 = 1.251
\]

**C:** Division of \( (N_i/ma) \) by \( m' \) and subtraction of 1 (one) from quotient, e.g. for 13 and 15 cm midlengths:

\[
(0.769 / 1.251) – 1.0 = -0.385
\]

\[
(1.939 / 1.251) – 1.0 = 0.550
\]

**D:** Identifying "isolated peaks", i.e. counting in column C the occurrences of zero-frequencies in the two class intervals above and below a given positive value and defining \( nz \) = (the number of neighboring classes with zero-frequencies). (If a given frequency is equal to zero (e.g. the 23, 25 and 31 cm classes of our example) or if the value in column C is negative, this routine will not become effective.)

**E:** De-emphasizing positive values in column C by multiplying them by \( 0.5^{nz} \) (rather than \( 0.2^{nz} \) in Brey and Pauly 1986). Thus

\[
N_{r} = [(N_i/ma)/m'] \cdot 0.5^{nz} ...
\]

where \( N_{r} = \) restructured value and \( nz = 0, 1, 2, 3 \) or 4. This gives, for example:

\[
N_{r15 \text{ cm}} = 0.550 \cdot 0.5^1 = 0.275
\]

\[
N_{r27 \text{ cm}} = 2.426 \cdot 0.5^3 = 0.303
\]

\[
N_{r33 \text{ cm}} = 0.999 \cdot 0.5^4 = 0.250
\]
Experience has shown (Pauly 1986) that the negative restructured value(s) corresponding to \( L_{\max} \), the largest recorded length(s), had too strong an influence on the shape of the estimated growth curve. To compensate for this, the following routine is applied, which uses \( L_{\max-1} \), the length class immediately preceding that containing \( L_{\max} \):

<table>
<thead>
<tr>
<th>Point value</th>
<th>( L_{\max-1} )</th>
<th>( L_{\max} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>point value not modified</td>
<td>point value not modified</td>
</tr>
<tr>
<td></td>
<td>(but counted only once)</td>
<td>(but counted only once)</td>
</tr>
<tr>
<td>Zero</td>
<td>no change</td>
<td>no change</td>
</tr>
<tr>
<td>Negative</td>
<td>point value multiplied by 0.5</td>
<td>point value set equal to zero</td>
</tr>
</tbody>
</table>

F: All values in column E that are equal to -1.000 are set to zero. All values that are smaller than zero but not equal to -1.000 are multiplied by \( SPV/-SNV \), where \( SPV = \) sum of positive values in column E, and \( SNV = \) sum of negative values in column E.

In Table 4.2, we have \( SPV = 0.275 + 0.017 + 0.303 + 0.250 + 0.042 = 0.887 \) and \( SNV = -0.385 - 0.184 - 1 - 1 - 0.556 = -4.125 \). Consequently, \( SPV/-SNV = (0.887/4.125) = 0.215 \). Values in column E that have to be multiplied by \( SPV/-SNV \) are: 0.385; -0.184; -0.200; and -0.556.

G: Computing the available sum of peaks (ASP) by identifying the highest points in a run of positive point values, bordered on either side by zero or by negative values (troughs). Such run may consist of one length class only. In our example, these values (in column F) are 0.275; 0.303 and 0.250. Their sum is \( ASP = 0.828 \).

During the computation of ESP, any "peak" (i.e. run of positive values) being hit once by a growth curve is flagged so that it will not be counted again. However, negative points, representing troughs, still are counted every time they are hit (Pauly 1985).

In the FISAT software, steps A–G are performed automatically every time length–frequency data are restructured. Fig. 4.4 illustrates a set of L/F data restructured following these seven steps.

![Fig. 4.4. Restructured version of the L/F data in Table 3.1, pertaining to Upeneus moluccensis (Ragay Gulf, Philippines, 1980).](image)

Different approaches can be used to identify the parameter values (\( L_o \), and \( K \), and \( C \) and \( WP \) if seasonality is considered) which best fit a data set such as in Fig. 4.7; in FISAT, the following are provided:

1. **Curve-fitting by eye**, i.e. entering growth parameter values thought to be appropriate, and using the resulting value of \( R_n \) to evaluate goodness of fit, relative to a previous attempt; an approach that is extremely slow and inefficient, but which is nevertheless recommended as first step, if only to get a feel for the data at hand;

2. **Response surface analysis**, i.e. plotting values of \( R_n \) along axes representing values of two of the parameters (usually, \( L_o \) and \( K \), with \( C \) and \( WP \) set at fixed values, based on a seasonal cycle of temperature); the approach is efficient, but can be confusing;
3. **Scan of K-value**, the recommended approach: it requires that \( L_{\infty} \) be estimated first, using a Powell-Wetherall Plot (or assuming \( L_{\infty} = L_{\text{max}} \)); then K (and its approximate variance) can be estimated through a plot of \( R_n \) vs. K, with K ranging from 0.10 to 10year\(^{-1} \), the range of acceptable values acceptable by **FiSAT** (Fig. 4.5);

4. **Automatic search routine**, which can also be used, especially on fast computers, as it allows varying all four parameters simultaneously; its use requires an understanding of how to conduct searches in multidimensional space, in the presence of multiple local optima (Fig. 4.6);

5. **Output routine**, which can be used for fitting a secondary growth curve (i.e. a curve forced through a given peak in the L/F data), as in the case of approach 1; this is not a very efficient approach, however, as the true function of this routine is to present results (Fig. 4.7).
Shepherd's method

Original version

Shepherd's method 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 non-feasible interpretations of the data and yielding direct estimates of growth parameters.

The goodness-of-fit criterion used in the original version of Shepherd’s method 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 for example Bloomfield 1976). The criterion is designed to take account of data in proportion to their quantity and likely reliability without pre-processing, and not to be unduly sensitive to possible spurious modes created by sampling noise.

The approach interprets one or more length compositions (arising from data collected in different years, seasons or months) in terms of a single growth curve. For simplicity, this is assumed to be of the standard VBGF, i.e. without seasonal modulation of growth rate.

Given current estimates of the parameters, \( L_0 \), \( K \) and \( t_o \) of the growth curve, the lengths where modes are to be expected in a season which is a fraction \( t \) of the annual cycle are given by

\[
L = L_0 \left( e^{K(t+1-t_o)} \right)
\]

for \( t = 0, 1, 2, \ldots \) 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 \( t = 0.5, 1.5, 2.5, \ldots \)) constitute evidence against them.

For any reason, one may construct a test function which is positively near the expected modal lengths and negatively near the intermodal lengths. The sum of products of the observations with this test function then constitutes a criterion of goodness-of-fit measure (or “score”), constructed by summing the 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 \( L_0 \) 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 \( \cos(2\pi(t-t_o)) \) where \( t \) is estimated age-at-length (measured from 1 January), i.e.

\[
t = \left( \frac{1}{K} \right) \cdot \ln \left( \frac{(L_0 - L_o)/(L_0 - L_i)}{L_0} \right) \quad \ldots 4.13)
\]

or

\[
t = t_o - \left( \frac{1}{K} \right) \cdot \ln \left[ 1 - \left( \frac{L_i/L_o}{1} \right) \right] \quad \ldots 4.14)
\]

where \( L_o = L_0 \left( 1 - e^{-Kt_o} \right) \)

---

20 Adapted from Shepherd (1987).
For a length interval whose upper and lower bounds correspond to ages-at-length of $t_{\text{max}}$ and $t_{\text{min}}$, it is easily shown that

$$T_i = \frac{\sin\pi(t_{\text{max}} - t_{\text{min}})}{\pi(t_{\text{max}} - t_{\text{min}})} \cos 2\pi(\bar{t} - t_i) \quad ...4.15$$

where $T_i$ is the average of this test function over the interval, and $\bar{t}$ is the average of $t_{\text{max}}$ and $t_{\text{min}}$. The first term of this expression is the 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 goodness-of-fit (or score) function is then

$$S = \sum_l \sum_i T_{li} \cdot N_{li}^{1/2} \quad ...4.16$$

where $l$ indexes the length groups and $i$ indexes the various distributions available.

Noting that, because of Equation 4.14, $t_{\text{max}}$ and $t_{\text{min}}$ and, therefore, also $\bar{t}$ and $T_i$ are periodic in $t_z$ ($t_z$ is introduced here to denote a calendar date in terms of years as opposed to $t_0$ which is the age at length zero, usually not known), it is clear that $S$ is similarly periodic (with a period of unity). It is therefore possible to determine very easily the value of $t_z$ leading to a maximum of $S$ for any choice of values for $K$ and $L_n$. If we denote by $A$ the value of $S$ obtained with $t_z$ set to zero, and $B$ that obtained with $t_z$ set to 0.25, it can be determined that

$$S_{\text{max}} = (A^2 + B^2)^{1/2} \quad ...4.17$$

and

$$t_z = \left(\frac{1}{2\pi}\right) \tan^{-1} \left(\frac{B}{A}\right) \quad ...4.18$$

For a number of practical reasons this scoring function is, in FiSAT, re-expressed in relative terms, i.e. for each run the values of $S$ are divided by $S_{\text{max}}$, the maximum value of $S$ in that run (which is also shown, however). The relative score value resulting from this, although ranging from 0 to 1, must not be seen as similar to a correlation coefficient or other measure of association. There is no "test" to tell whether the fit is "significant" or not.

Fig. 4.8 shows an application of this method to *Upeneus moluccensis* from Ragay Gulf, Philippines, with scores plotted against $K$ values ranging, in log steps, from 0.1 to 10 years$^{-1}$.

![Fig. 4.8. K-scan routine for Upeneus moluccensis from Ragay Gulf, Philippines (L/F data of Table 3.1; note differences from Fig. 4.5).](image)

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 the contours of constant score, output by FiSAT, 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 (or local optimum) is associated with something like a 95 percent confidence level (since they are likely to be observations in several length groups contributing to the total).

Improved version

While Shepherd’s Length Composition Analysis (SLCA) is rather ingenious in its design, and runs fast on any computer, it has some disadvantages:

(i) it does not incorporate seasonal oscillations;

(ii) its score function does not have obvious maxima and minima, for best and worst possible fits, respectively;

(iii) it usually does not converge, i.e. the highest value of the score function does not correspond to the best combination of \( L_i \) and \( K \) values (which are usually identified by a local optimum).

The problems in (i) and (ii) will not be dealt here. However, the problem in (iii) is straightforward to resolve. Fig. 4.9 illustrates this problem: the scoring function of SLCA tends to increase without bounds as \( K \) increases, and the optimum value of \( K \) is identified—at best—by a local optimum (see also Fig. 4.8). Moreau et al. (1995) present applications of SLCA to 56 length–frequency datasets pertaining to African freshwater fishes (and included in the report in question); the scoring function had its maximum at \( K = 10 \) year\(^{-1} \) in 29 of these cases, and \( K \) was overestimated in nearly all cases.

The reason for these features of SLCA is that its score function gives the same weight to length classes pertaining to small (young) fishes, of which several contribute to an age group, as to length classes pertaining to large (old) fishes, each of which consists of the fish of several cohorts.

Countering this is surprisingly simple: to give adequate weights to the different frequencies to be analyzed (see Fig. 4.10), it is sufficient to divide each frequency associated with a given length class by \( \Delta t_i \), the time needed by the fishes to grow from the lower limit (\( L_i \)) to the upper limit (\( L_{i+1} \)) of length class \( i \); this should be done right after tracing the oscillatory curve in Fig. 4.10, and before counting the scores. The \( \Delta t_i \) values are computed from

\[
\Delta t_i = \frac{1}{K} \ln \left( \frac{L_{i+1} - L_i}{L_i - L_{i-1}} \right)
\]

\[\text{...4.19}\]
where $L_\infty$ and $K$ are the growth parameters being evaluated by the scoring function. It will be noted that in a length-converted catch curve (see Chapter 5), the same equation is used to correct for the piling-up of old, large fish at the tail end of the length–frequency distributions, which otherwise leads to underestimation of $Z$ (Pauly et al. 1995).

**Fig. 4.10.** The SLCA method estimates $L_\infty$ and $K$ by re-expressing the VBGF as an oscillating line whose limit is $L_\infty$ and whose frequency of oscillation is determined by $K$. The length frequencies comprised between the positive parts of oscillations are multiplied with the value of the corresponding section of the oscillatory curve, then added to a positive counter ($N+$); while the corresponding product of the negative part is added to a negative counter ($N-$). The scoring function is then $S = (N+) + (N-)$. 

**Fig. 4.11.** Three examples of the differences and similarities between the old version of SLCA (top row), the new version of SLCA (NSLCA; middle row) and the ELEFAN I routine (bottom row):

- Goldband goat fish (*Upeneus moluccensis*): ($L_\infty = 20$ cm) the old SLCA overestimates $K$ (see also Fig. 4.5); NSLCA and ELEFAN I estimate $K$ as 0.84 and 0.87 year$^{-1}$, respectively.
- Mango tilapia (*Sarotherodon galilaeus*): ($L_\infty = 30.1$ cm) the old SLCA does not allow estimation of $K$; NSLCA leads to $K = 0.58$ year$^{-1}$; for ELEFAN I, $K = 0.51$ year$^{-1}$.
- Shortfin tilapia (*Tilapia brevimanus*): ($L_\infty = 19.6$ cm) the old SLCA does not allow estimation of $K$; NSLCA leads to $K = 0.65$ year$^{-1}$; for ELEFAN I, $K = 0.69$ year$^{-1}$.

The application of this simple modification has the effect of suppressing the increase of the score with increasing $K$ (see Fig. 4.11), and to shift the optima to the left, largely countering the tendency of SLCA to overestimate $K$. Indeed, when seasonal growth oscillations are not considered, the results obtained with the improved SLCA are virtually the same as those obtained with ELEFAN I, as should be expected if both methods work as they should. The new version of SLCA was not included in the first release of **FiSAT** but an update of this routine can be found in the update disk distributed with this manual.
Powell–Wetherall method\textsuperscript{22}

Wetherall (1986), following up on work by Powell (1979), suggested a simple method to estimate the asymptotic length ($L_\infty$) and the ratio of the coefficients of mortality and growth ($Z/K$), using only length–frequency data from a fish catch. This is based on the well-known equation of Beverton and Holt (1956):

$$Z = K \left[ \frac{(L_\infty - L) / (L - L')}{L} \right]$$ \quad ...4.20

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 ($\bar{L}$) of a random sample of fish above length $L'$. When using Equation 4.20, it is assumed that $L'$ and the VBGF parameters $K$ and $L_\infty$ are given. If only $L'$ and $L_\infty$ are known (or estimated from other information), slight manipulation of Equation 4.20 yields an estimate of the $Z/K$ ratio, a component of many standard yield models and stock assessment procedures (see Chapter 7).

It can be shown that the mean length of the $n$ selected fish ($\bar{L}$) is a linear function of the knife-edge selection length ($L'$):

$$\bar{L} = L_\infty \left\{ \frac{1}{1 + (Z/K)} \right\} + L' \left\{ \frac{1}{1 + (Z/K)} \right\}$$ \quad ...4.21

where $Z$ is the instantaneous mortality rate, and $K$ and $L_\infty$ are the growth coefficient and asymptotic length of the VBGF.

For a series of arbitrary cutoff lengths – within the size range of the $n$ selected fish, we can thus construct a corresponding series of partially overlapping sub-samples. The $i$-th sub-sample consists of those $n_i$ fish whose lengths exceed $L_i$ ($i = 1, \ldots, n$). If the mean lengths for the sub-samples are plotted against the cutoff lengths, a positive linear relationship results, as predicted by Equation 4.21. Thus, if we label the intercept of the straight line as “$a$”, and its slope as “$b$”, we have

$$a = L_\infty \left[ \frac{1}{1 + (Z/K)} \right]$$ \quad ...4.22a

and

$$b = \frac{Z/K}{1 + (Z/K)}$$ \quad ...4.22b

The parameters of interest, $L_\infty$ and $Z/K$, can be estimated in two simple steps:

Step 1. Estimate $a$ and $b$ from a linear regression of $L_1$ on $L_i'$. Statistical weights should be used, but it will usually suffice to weight each sub-sample mean length by the corresponding sub-sample size.

Step 2. Given the estimates of $a$ and $b$, compute $L_\infty$ and $Z/K$ as the solution of Equations 4.22a and 4.22b:

$$L_\infty = a / (1 - b)$$ \quad ...4.23

and

$$Z/K = b / (1 - b)$$ \quad ...4.24

This approach improves on the Beverton-Holt method, allowing the estimation of $L_\infty$ in addition to $Z/K$. It does this by making full use of information in the sample besides the overall mean length. Specifically, it takes advantage of the linear relationship between sample mean length and the selection length.

Wetherall et al (1987) and Isaac (1990) discussed various properties of this model and should be consulted for details. In FiSAT, it is incorporated in the form proposed in Pauly (1986), wherein Equation 4.21 has been modified to

$$L_i - L_i' = a + bL_i$$ \quad ...4.25

\textsuperscript{22} Adapted from Wetherall et al. (1987).
Fig. 4.12. Estimation of \( L_{1/2} \) and \( Z/K \) using the modified Powell-Whetherall method, based on length–frequency data of *Upeneus moluccensis* from Ragay Gulf, Philippines (Table 3.1).

where \( L_{1/2} = a/b \), and \( Z/K = (1+b)/-b \), the advantage being that \( L_{1/2} \) is estimated as intercept with the abscissa, which facilitates interpretation of the plot (see Fig. 4.12).

**Using size-at-age data**

When size-at-age data are available, the non-linear least-squares method is the best approach for estimating growth parameters.

When estimating the parameters of a given function using least squares, the point is to minimize the sum of squared errors (SSE):

\[
SSE = \sum (y - \hat{y})^2 \tag{4.26}
\]

where \( y \) is the observed value of the dependent variable and \( \hat{y} \) is the predicted value of the dependent variable (current estimates). The nonlinear least squares method uses the following assumptions (Saila et al. 1988): correct model specification, independence of observations, homoscedasticity, normally distributed errors with zero mean, and knowledge of the independent variable(s) without error. Please consult statistics texts if you have reasons to think that these assumptions are not met by the data available to you.

Assume a series of pairs of observations \((L_i, t_i)\), as in ordinary linear least squares where \( L_i \) refers to the length of fish no. \( i \), \( t_i \) is the age of fish no. \( i \) and \( i = 1, 2, 3, \ldots, n \).

The sum of squared errors (Equation 4.26) can be expressed, using the non-seasonalized version of the VBGF, as

\[
SSE = \sum_{i=1}^{n} \left( L_i - L_{0} \left[ 1 - e^{-K(t_i - t_0)} \right] \right)^2 \tag{4.27}
\]

and for the seasonalized version,

\[
SSE = \sum_{i=1}^{n} \left( L_i - L_{0} \left[ 1 - e^{-K[t_i - t_0] + iCK(t_i - t_0)} \right] \right)^2 \tag{4.28}
\]

where

\[ t_i = \sin 2\pi(t_i - t_0), \text{ and } t_0 = \sin 2\pi(t_0 - t_0). \]

There are a number of established algorithms to minimize SSEs (e.g. the Taylor series or linearization, steepest descent; see also Soriano and Jarre 1988). Marquardt's method (Marquardt 1963) appears to work well as far as the VBGF is concerned (Saila et al. 1988), and is thus a sensible choice. The algorithm itself will not be discussed here, and we refer to the literature. It is, however, important to note here the importance of initial inputs, or “seed values”.

\[
L - L_{1/2} = 4.05 - 0.208 L^{1/2}
\]

Estimate of \( L_{1/2} = 19.5 \text{ cm} \),

Estimate of \( Z/K = 3.81 \).
Procedures that estimate parameters iteratively require initial guesses as close to the true value as possible. Not only do these inputs determine the time of convergence toward a solution, but wrong initial values may result in convergence to biologically impossible "local" minima. It is therefore suggested that some effort should be made to obtain good starting values, from the literature or from other methods.

In the section that follows, a technique is introduced which makes it possible to drive series of sizes-at-(relative)age data from length–frequency data; Fig. 4.13 illustrates an example of the VBGF fitted to such data.

![VBGF fitted using the Marquardt algorithm and superimposed onto the set of lengths-at-(relative)age data, derived from length frequencies of Upeneus moluccensis (Ragay Gulf, Philippines 1980) using modal progression analysis (MPA) and linking of the mean lengths (see next section).](image)

With regards to the fit of such functions, it was recommended by Kvalseth (1985; see also Saila et al. 1988) that \( r^2 \) be adjusted for the positive bias of a nonlinear fitting, i.e.,

\[
 r^2 = 1 - \frac{\sum(y - \hat{y})^2}{\sum(y - \bar{y})^2} \quad \ldots \ldots 4.29
\]

and

\[
 r^2_{adj} = 1 - \left[ \frac{(n - 1)}{(n - k)} \right] (1 - r^2) \quad \ldots \ldots 4.30
\]

where \( y, \hat{y} \) and \( n \) are as defined earlier, \( \bar{y} \) is the mean of the dependent variables and \( k \) is the number of parameters estimated.

Modal progression analysis (MPA)\(^{23}\)

Prior to the discovery of daily and seasonal rings in fish otoliths and scales, the detailed analysis of length–frequency data was, besides mark–recapture studies and direct observation of captive fish, the only method available to draw inferences on the growth of tropical fishes, and indeed the only method that could be applied routinely (Pauly 1987).

A great deal of confused terminology surrounds the use of length–frequency data in growth analysis, and before discussing this approach it is necessary to clarify the terms we shall use.

Here, we use the term modal progression analysis (MPA) for the method in which several length–frequency samples are plotted sequentially, and where the apparent shift of modes is used to infer growth (see Fig. 4.2B). In MPA the first step, "linking", is the identification of the means or modes representing cohorts, after which subjective identification of the means or modes perceived to belong to the same cohort of fish enables the tracing of growth curves, or the computation of growth increments between modes. Thus, in MPA, the critical issue is not the attribution of "ages" to the various groups—as

---

\(^{23}\) Adapted from Gayanilo et al. (1989).
in the Petersen method—but the linking of means perceived to belong to a same cohort.

There are a number of methodologies that can be applied to decomposed distributions into their components as the initial step in MPA, but only two of these are introduced here—Bhattacharya’s method and Hasselblad’s NORMSEP, as incorporated in FiSAT.

Bhattacharya’s method

Bhattacharya (1967) developed a methodology useful for splitting a composite distribution into separate normal distributions when several age groups (cohorts) of fish are contained in the same sample. They key advantage of this method is that it is straightforward to understand (Pauly and Caddy 1985).

The Bhattacharya method basically consists of removing normal distributions, each representing a cohort of fish, from a mixture of distributions, starting on the left-hand side of the overall distribution. Once the first normal distribution has been identified, it is subtracted from the total distribution and the procedure is repeated as long as it is possible to identify distinct normal distributions.

Fig. 4.14 summarizes this method via a series of steps; Fig. 4.15 presents an application to the data in Table 4.3, which also presents intermediate results.

---

24 Adapted from Sparre and Venema (1992).
There exist a number of graphical and mathematical techniques other than Bhattacharya's method which can be used for cases where mixtures of distribution need to be separated into their component distributions (see Everitt and Hand 1981, for a review). Most of them assume that the distributions to be separated from the mixture are normal distributions (Pearson 1894, Harding 1949, Cassie 1954, Tanaka 1962, Hasselblad 1966, Yong and Skillman 1971, Macdonald and Pitcher 1978). The NORMSEP program, of which an algorithm is briefly presented here, requires inputs as to the expected number of component distributions in the mixture.

Although NORMSEP may not be the most sophisticated tool available to analyze size–frequency distributions, it is useful in situations where the data to be separated do not greatly overlap, or where previous knowledge exists regarding the number and probable positions of the means of the component distributions, as obtained for example through previous analysis with Bhattacharya's method.

Hasselblad (1966) presented the problems for a mixture of normal distributions by assuming that the elements of each sub-population, \( N_i \), normally distributed with mean and variance, are defined by a mixed population;

The major modifications made to the original version are:

- An algorithm was incorporated in FiSAT which automatically computes, using the (guessed) means entered, a very large range of possible values for the standard deviation corresponding to each mean, by using a very small (0.01) and a very large (0.50) coefficient of variation (C.V. = s.d./\( \bar{X} \)) to generate the range. This follows Abrahamson's suggestion that "anyone desiring to maintain

\[\text{Hasselblad's NORMSEP}^25\]

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.5</td>
<td>4</td>
<td>1.366</td>
<td>1.366*</td>
<td>-</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>13.5</td>
<td>20</td>
<td>3.186</td>
<td>0.786*</td>
<td>-</td>
<td>20</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>14.5</td>
<td>38</td>
<td>3.638</td>
<td>0.460*</td>
<td>-</td>
<td>38</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>15.5</td>
<td>42</td>
<td>3.738</td>
<td>0.100*</td>
<td>-</td>
<td>42</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>16.5</td>
<td>33</td>
<td>3.497</td>
<td>-0.241*</td>
<td>-</td>
<td>33</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>18.5</td>
<td>53</td>
<td>3.83</td>
<td>0.501*</td>
<td>-</td>
<td>53</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>19.5</td>
<td>7</td>
<td>1.946</td>
<td>-1.050</td>
<td>-</td>
<td>7</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>20.5</td>
<td>3</td>
<td>1.099</td>
<td>-0.847</td>
<td>-</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>22.5</td>
<td>2</td>
<td>1.099</td>
<td>0.000</td>
<td>-</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>23.5</td>
<td>1</td>
<td>1.099</td>
<td>-1.050</td>
<td>-</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>24.5</td>
<td>2</td>
<td>2.079</td>
<td>0.470</td>
<td>-</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>25.5</td>
<td>11</td>
<td>2.398</td>
<td>0.316</td>
<td>-</td>
<td>11</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>26.5</td>
<td>14</td>
<td>2.639</td>
<td>0.241</td>
<td>-</td>
<td>14</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>27.5</td>
<td>17</td>
<td>2.833</td>
<td>0.194</td>
<td>-</td>
<td>17</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

25 Adapted from Pauly et al. (1986).
the rigor of Hasselblad's original procedure only needs to place extremely wide bounds on the standard deviations.

- An algorithm was added which internally computes the number and approximate position of the "cutoff" points.

**Linking of mean lengths**

Modal progression analysis (MPA) does not stop where the decomposition of normal distributions ends, but rather the key results of the decomposition (i.e. the means) should be plotted to form either growth increments (see below) or a time series likely to represent the progression, through time, of the mean length of a cohort.

In Fig. 4.16, the mean lengths of the components are plotted against the sampling dates and those mean lengths which are believed to belong to the same cohort are linked.

**Fig. 4.16.** Mean lengths computed using Bhattacharya's method and Hasselblad's NORMSEP from length frequencies of Upeneus moluccensis (Ragay Gulf, Philippines; see Table 3.1), plotted against the sampling dates and linked in a fashion likely to represent the growth of assumed cohorts. Note that preliminary estimates of growth parameters can be obtained, when sufficient links are available, by using the Gulland and Holt plot (see next section).

It should be emphasized here that the linking process is highly subjective and that the reliability of the links thus rests entirely on the experience of the person performing the analysis. Note that for the data presented here, the tracing of the growth increments appears easy. However, in practice this process will generally be more problematic (see Fig. 4.17).

**Estimating the C.V. of L<sub>∞</sub>**

When the standard deviations associated with each mean length (derived through Bhattacharya's method or NORMSEP) are available, the coefficient of variation (C.V.) of L<sub>∞</sub> can be estimated. Moreover, a plot of the average C.V.s associated with pairs of connected mean lengths vs. the corresponding growth increments (ΔL/Δt) can be used to identify the most likely of two alternative hypotheses concerning the distribution of length about age (see Fig. 4.18):

(i) Constant C.V.—i.e. the standard deviation of length around age—increases in proportion of length (see Fig. 4.18A).

(ii) From intermediate lengths on, the C.V.s decline with length, and hence with ΔL/Δt (Fig. 4.18B).
Fig. 4.18. Effects of the variability of von Bertalanffy growth parameters (\(L\) and \(K\)) about mean growth curves (solid lines). **A**: effect of a variable \(L\); **B**: effects of a variable \(K\) (modified from Sainsbury 1980). In most cases, both parameters will vary (as will the \(t_o\) values, not discussed here), leading to a mixture of the effects in **A** and **B**.

In the case of (i), a plot of C.V. vs. \(\Delta L/\Delta t\) will result in a vertically arranged series of dots, and a non-significant correlation between the C.V. and \(\Delta L/\Delta t\) values. Alternatively, for hypothesis (ii) to be most likely, the correlation coefficient (\(r\)) must be significant and positive and the x-intercept must be positive. In case (i) the C.V. of \(L_o\) is estimated as the mean of the C.V. values; in case (ii) it is estimated as the x-intercept, i.e. where \(\Delta L/\Delta t\) is equal to zero.

A significantly negative x-intercept or a positive slope is indicative of the following:

(i) a flaw in either the decomposition of the normal distribution, particularly in identifying "older" groups (due to overlaps), or in the linking procedure itself, or a cumulative effect of biases in both procedures; or

(ii) under-representation of "older" groups in the samples.

Fig. 4.19. Illustration of the two alternative hypotheses: (i) assuming constant C.V. (broken line) and (ii) assuming C.V. declining with \(\Delta L/\Delta t\) (solid regression line). Based on the linked mean lengths in Fig. 4.16.

These considerations, inspired from Sainsbury (1980) are incorporated, in **FISAT**, in the form of a plot (such as in Fig. 4.19) based on the linked means and their C.V., and which can be seen and printed as an intermediate step every time an MPA is being performed.

**Estimating growth from length increments**

When growth increment data are available (e.g. from mark-recapture information, or from increments resulting from connecting mean lengths), the parameters describing a growth curve can be estimated. In the following, three approaches are outlined which may be applied when non-seasonalized growth is assumed: the Gulland and Holt plot, Munro’s method and Fabens’ method. For seasonally oscillating growth, a modification of the method of Appeldoorn (1987) is presented.
Gulland & Holt plot

Background and 1st example

The VBGF, presented earlier implies that growth rate \( (dL/dt) \) declines linearly with length. Since growth increments \( (\Delta L/\Delta t) \) derived from successive lengths \( (L_i; L_{i+1}) \) approximate \( dL/dt \) values (Gulland and Holt 1959), the VBGF can be linearized by

\[ (\Delta L/\Delta t) = a + b \bar{L} \]

...4.31) where \( \bar{L} = L_i + \frac{1}{2} \Delta L \), \( K = -b \) and \( L_\infty = -a/b \).

Table 4.4 contains an example of such data and Fig. 4.20 shows the corresponding plot.

<table>
<thead>
<tr>
<th>Length ( (L_i) )</th>
<th>Length ( (L_{i+1}) )</th>
<th>Date ( (L_i) )</th>
<th>Date ( (L_{i+1}) )</th>
<th>( \Delta L ) (cm)</th>
<th>( \Delta t ) (days)</th>
<th>( \Delta L/\Delta t )</th>
<th>( \bar{L} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.60</td>
<td>10.39</td>
<td>20/11/80</td>
<td>20/01/81</td>
<td>1.79</td>
<td>61</td>
<td>0.029</td>
<td>9.50</td>
</tr>
<tr>
<td>10.39</td>
<td>12.72</td>
<td>20/01/81</td>
<td>21/05/81</td>
<td>2.33</td>
<td>121</td>
<td>0.019</td>
<td>11.56</td>
</tr>
<tr>
<td>12.02</td>
<td>13.06</td>
<td>11/12/80</td>
<td>20/01/81</td>
<td>1.04</td>
<td>40</td>
<td>0.026</td>
<td>12.54</td>
</tr>
<tr>
<td>12.72</td>
<td>13.64</td>
<td>21/05/81</td>
<td>23/07/81</td>
<td>0.92</td>
<td>63</td>
<td>0.015</td>
<td>13.18</td>
</tr>
<tr>
<td>13.64</td>
<td>14.68</td>
<td>23/07/81</td>
<td>27/09/81</td>
<td>1.04</td>
<td>66</td>
<td>0.016</td>
<td>14.16</td>
</tr>
<tr>
<td>14.68</td>
<td>15.35</td>
<td>27/09/81</td>
<td>11/12/81</td>
<td>0.67</td>
<td>75</td>
<td>0.009</td>
<td>15.02</td>
</tr>
</tbody>
</table>

From the estimate of the slope in Table 4.4 we get:

\[ K = -b = 1.22 \text{ year}^{-1} \]

and

\[ L_\infty = -a/b = 18.4 \text{ cm} \]

The 95 percent confidence limits for \( K \) can be taken as the same as those for \( b \) but with the sign changed. The confidence limits of \( L_\infty \) are more complicated to obtain. However, we may say that for a given value of \( K \), the confidence interval of \( L_\infty \) can be calculated from the confidence interval of \( a \). The confidence limits of \( a \) are not the confidence limits of \( L_\infty \), but only the limits conditioned on one fixed value of \( K \).

Mathematically it can be proven that Equation 4.31 is equivalent to the VBGF. However, this is correct only if the time interval \( (\Delta t) \) is very small. Thus, the Gulland and Holt plot, which is based on Equation 4.31, is an approximation which is reasonable only for small values of \( \Delta t \).

Amplifications, 2nd and 3rd examples

The Gulland and Holt plot, being approximative, should be used only for visualizing the data at hands and for obtaining seed values for more rigorous iterative methods. To help with
this, the Gulland and Holt plot is complemented in **FiSAT** with two routines.

The first of these routines is an option to fix the value of $L_\infty$, and with the value, force a line (of slope $\approx -K$) through a relatively narrow cloud of $\Delta L/\Delta t$, $\bar{L}$ values (Fig. 4.21). This routine is called a “forced” Gulland and Holt plot (Pauly 1984a).

![Fig. 4.21. Forced Gulland and Holt plot for mark–recapture data of bonito (Sarda chiliensis), off California, forcing the $L_\infty$ to be equal to 90 cm (see also Fig. 4.22).](image)

Fig. 4.21. Forced Gulland and Holt plot for mark–recapture data of bonito (Sarda chiliensis), off California, forcing the $L_\infty$ to be equal to 90 cm (see also Fig. 4.22).

The second routine to be presented here analyzes the residuals of Gulland and Holt plots (both standard and forced) in terms of their distribution within the year (see Fig. 4.22). This routine, by systematically varying the cutoff point between six-month periods, identifies the times of the year when the difference between two groups of residuals is greatest, usually representing a winter/spring and a summer/fall period.

![Fig. 4.22. Residuals of Fig. 4.21, plotted against mean days at large. Note the difference in their distribution over time, resulting in an estimate of $C \approx 0.2$ (see text).](image)

Fig. 4.22. Residuals of Fig. 4.21, plotted against mean days at large. Note the difference in their distribution over time, resulting in an estimate of $C \approx 0.2$ (see text).

These two groups of residuals can be tested for significant difference in their mean $\Delta L/\Delta t$ (with one degree of freedom lost for the estimation of the cutoff point). If significant, this difference can be used to obtain a minimum estimate of $C$ in the seasonally oscillating version of the VBGF, using

$$ C \geq \frac{\pi}{2} \cdot \bar{m} $$

where $\bar{m}$ is the mean difference between the first and second half of the year and relates a square wave to a sinus function (H. Lassen, pers. comm.).

Fig. 4.21, based on the **SAMPLE.GIN** file distributed with the **FiSAT** software, illustrates this procedure. This value of $C$ can then be used as a seed in the modified Appeldoorn method (see below), which explicitly considers seasonal growth oscillations.
Munro's method\textsuperscript{27}

If a fish is tagged at time $t$ and recovered at time $t + \Delta t$, from the von Bertalanffy equation

$$\ln(L_{\infty} - L_t) = \ln (L_{\infty}) + K \cdot t - K \cdot t_t$$  \hspace{1cm} (4.33)

which may be rewritten as

$$\ln(L_{\infty} - L_{t+\Delta t}) = \ln L_{\infty} + K \cdot t - K \cdot t_{t+\Delta t}$$  \hspace{1cm} (4.34)

where $L_t$ is length at tagging and $L_{t+\Delta t}$ the length at recapture, from which $K$ can be estimated given a tentative estimate of $L_{\infty}$ and a set of $L_t, L_{t+\Delta t}$ data pairs.

These definitions (Equations 4.33 and 4.34) lead, after some rearrangements, to

$$K = \frac{\ln(L_{\infty} - L_t) - \ln(L_{\infty} - L_{t+\Delta t})}{(t_{t+\Delta t} - t_t)} = Y/X$$  \hspace{1cm} (4.35)

The value of $L_{\infty}$ which yields the lowest coefficient of variation of the $Y:X$ ratios is the "best" estimate of asymptotic length. A set of data which conformed perfectly to the von Bertalanffy growth function would have a coefficient of variation of zero when the correct asymptotic length is inserted in Equation 4.35. The 95 percent confidence limits for the estimate of $K$ are given by $K \pm (1.96) \frac{\text{s.d.}}{\sqrt{n}}$, where $n$ is the number of $L_t, L_{t+\Delta t}$ data pairs.

Values of $L_{t+\Delta t}$ or $L_t$ which exceed the trial values of $L_{\infty}$ introduce negative values into the logarithmic expression and render the method inoperable. It is otherwise independent of the sizes of the fish marked and is also not dependent upon a wide variety of values for the time interval, $t_{t+\Delta t} - t_t$. The method, like the forced Gulland and Holt plot, can give estimates of growth parameters when the standard Gulland and Holt plot fails because of an insufficient range of lengths at marking and recovery. This applies particularly to $K$, which can be estimated from a single pair of $L_t, L_{t+\Delta t}$ values when an estimate of $L_{\infty}$ is available.

Fabens' method\textsuperscript{28}

We can use the von Bertalanffy equation to express the size, $L_{t+\Delta t}$, of the fish at time $t + \Delta t$ in terms of its size $L_t$ at time $t$:

$$L_{t+\Delta t} = L_{\infty} \left[1 - e^{-K(1+\Delta t)}\right]$$  \hspace{1cm} (4.36)

which may be decomposed into

$$L_{t+\Delta t} = L_{\infty} \left[1 - \left[e^{-Kt} \cdot e^{-K\Delta t}\right]\right]$$  \hspace{1cm} (4.37)

and

$$e^{-Kt} = 1 - \left(\frac{L_t}{L_{\infty}}\right)$$  \hspace{1cm} (4.38)

Thus, we can rewrite as follows:

$$L_{t+\Delta t} = L_{\infty} \left[1 - \left(1 - \left(\frac{L_t}{L_{\infty}}\right) \cdot e^{-K\Delta t}\right)\right]$$  \hspace{1cm} (4.39)

and

$$L_{t+\Delta t} = L_t + \left(L_{\infty} - L_t\right) \left(1 - e^{-K\Delta t}\right)$$  \hspace{1cm} (4.40)

Equation 4.40 represents a relationship that does not require the input of $t_t$ and in which the new size depends only on the old size, the time elapsed and the parameters $K$ and $L_{\infty}$.

To fit the curve, we let $n$ be the number of observations with paired linear sizes, $L_t$ and $L_{t+\Delta t}$, and with time difference

\textsuperscript{27} Adapted from Munro (1982).

\textsuperscript{28} Adapted from Fabens (1965).
\(\Delta t\). Using Equation 4.40, given \(L_i\) and \(\Delta t\), we can predict \(L_{i+\Delta t}\) as

\[L_{i+\Delta t} = L_i + \left(L_{\infty} - L_i\right)\left(1 - e^{-K\Delta t}\right)\]  
...4.41

If we apply the least squares principles to \(L_{i+\Delta t}\), we should minimize SSE:

\[SSE = \sum_{i=1}^{n}(L_{i+\Delta t} - L_i)^2\]  
...4.42

and

\[SSE = \sum_{i=1}^{n}\left[L_{i+\Delta t} - L_i - (L_{\infty} - L_i)(1 - e^{-K\Delta t})\right]^2\]  
...4.43

To minimize SSE (Equations 4.42 and 4.43), we set partial derivatives with respect to \(K\) and \(L_{\infty}\) equal to zero. Therefore:

\[L_{\infty} = \frac{\sum_{i=1}^{n}(1 - e^{-K\Delta t})\left(L_{i+\Delta t} - L_i\cdot e^{-K\Delta t}\right)}{\sum_{i=1}^{n}(1 - e^{-K\Delta t})^2}\]  
...4.44

Differentiating SSE with respect to \(K\) gives

\[\frac{d(SSE)}{dK} = -2\sum_{i=1}^{n}\Delta t\cdot e^{-K\Delta t}\left(L_{\infty} - L_i\right)\left(L_{i+\Delta t} - L_i\right)\]  
...4.45

If we introduce simplifying substitutions:

\[p_i = e^{-K\Delta t}\]
\[r_i = L_{\infty} - L_i\]
\[S_i = L_{i+\Delta t} - L_i\]
\[d_i = \Delta t\]

and make the left side of Equation 4.45 equal to zero, we can state that

\[F(K) = \sum d_i p_i r_i (r_i + S_i) = 0\]  
...4.46

Equation 4.46 can be solved by using the Newton-Raphson iteration method, wherein successive values of \(K\) are improved by using

\[K_{i+1} = K_i - \frac{F(K_i)}{F'(K_i)}\]  
...4.47

where the derivative \(F'(K)\) in Equation 4.47 can be expressed as

\[F'(K) = \sum d_i p_i \left\{\left(dL_n/dK\right)\left[S_i + r_i(1 - 2p_i)\right] + \ldots + r_i p_i \left(2p_i r_i - S_i\right)\right\}\]  
...4.48

and where

\[\frac{dL_n}{dK} = \sum d_i p_i \left[L_{i+\Delta t} + L_i(1 - 2p_i) - 2L_n(1 - p_i)\right] / \sum (1 - p_i)^2\]  
...4.49

Fabens' method, although rigorous mathematically and much used by fisheries scientists, shares with Munro's method the disadvantage that it cannot be used to fit seasonally oscillating growth increment data. Yet seasonal growth oscillations are ubiquitous, and occur even in coral reef fishes (Pauly and Ingles 1981). Thus, any analysis performed with Fabens' or Munro's method should be complemented by an analysis using the method presented below—at least in those cases where sufficient data pairs are available.

Modified Appeldoorn method

Somers (1988) showed that the seasonalyzed VBGF (see Equation 4.8) can be adapted for use with seasonally oscillating growth increments by letting
\[
S(t) = \frac{CK}{2} \sin 2\pi(t - t_0) \quad ...4.50
\]

and
\[
S(t_0) = \frac{CK}{2} \sin 2\pi(t_0 - t_0) \quad ...4.51
\]

which makes it possible to rewrite the Equation 4.8 as
\[
L_i = L_\infty \left\{ 1 - \exp \left[ \frac{K}{2} \sin 2\pi(t - t_i) \right] \right\} \quad ...4.52
\]

After time of liberty (\(\Delta t\)), the length attained, \(L_{t + \Delta t}\), after release at length \(L_t\) will be expressed as
\[
L_{t + \Delta t} = L_\infty \left\{ 1 - \exp \left[ \frac{K}{2} \sin 2\pi(t - t_0) - \sin 2\pi(t_0 - t_i) \right] \right\} \quad ...4.53
\]

Substituting \(S(t + \Delta t)\) for \(S(t + \Delta t) + S(t) - S(t)\) in Equation 4.53, we can write
\[
L_{t + \Delta t} = L_\infty \left\{ 1 - \exp \left[ \frac{K}{2} \sin 2\pi(t - t_0) - \sin 2\pi(t_0 - t_i) \right] \right\} \quad ...4.54
\]

From Equation 4.51,
\[
L_\infty - L_i = L_\infty \cdot \exp \left[ \frac{K}{2} \sin 2\pi(t - t_0) - \sin 2\pi(t_0 - t_i) \right] \quad ...4.55
\]

Therefore Equation 4.54 can be simplified to
\[
L_{t + \Delta t} = L_\infty \left\{ 1 - \exp \left[ \frac{K}{2} \sin 2\pi(t - t_0) - \sin 2\pi(t_0 - t_i) \right] \right\} \quad ...4.56
\]

or alternatively, growth increments can be described by
\[
(L_{t + \Delta t} - L_t) = (L_\infty - L_i) \left\{ 1 - \exp \left[ \frac{K}{2} \sin 2\pi(t - t_0) - \sin 2\pi(t_0 - t_i) \right] \right\} \quad ...4.57
\]

Soriano and Jarre (1988) presented a linearized version of Equation 4.57 for fitting growth increment data using multiple linear regression. In FISAT, Equation 4.57 is implemented by fitting the function using a non-linear fitting technique, as also suggested by Appeldoorn (1987).

### Comparing the growth of fishes

#### Growth comparison\(^{29}\)

When studying the growth of tropical fishes and invertebrates, the question of validating growth parameter estimates often arises, due to the low reliability of some of the methods used to obtain such estimates.

A comparative approach can be used for this, i.e. comparison of the growth performance among organisms whose growth is described by the same mathematical model.

In fishes and invertebrates whose growth can be described by the von Bertalanffy growth function (VBGF), the comparison of growth performance is facilitated by the feature, demonstrated empirically and on theoretical grounds by Pauly (1979a), that for fish of the same species, double logarithmic plots of the coefficient \(K\) vs. the asymptotic weight, \(W_\infty\), tends to be linear with a slope of 2/3. Thus,

\[
\log_{10}(K) = \phi - 0.67 \log_{10}(W_\infty) \quad ...4.58
\]

and hence
\[
\phi = \log_{10}(K) + 0.67 \log_{10}(W_\infty) \quad ...4.59
\]

The implication of the constant slope of 0.67 is that the differences in the growth performance of different organisms are reflected solely in the value of the y-axis intercept, \(\phi\), which may be viewed as the (theoretical) value of \(K\) that would occur in fishes and invertebrates with a \(W_\infty\) value of 1 unit of mass.

---

\(^{29}\) Adapted from Munro and Pauly (1983).
Fig. 4.23 shows the frequency distribution of values of $\phi$ for different stocks of a number of species of fishes (for which $W_{\infty}$ is expressed in grams). As can be seen, the use of Equation 4.59 to estimate $\phi$ from published growth parameters for various stocks of the species produces distributions of $\phi$ values that are essentially normal (bell-shaped), and rather sharply peaked, suggesting that Equation 4.59 indeed describes the interrelationship of $K$ and $W_{\infty}$ in those fishes whose growth can be described by the VBGF.

The relationship expressed in Equations 4.58 and 4.59 can also be used to characterize the animals of a given family. Fig. 4.24 shows examples of the frequency distributions of $\phi$ for species of mesopelagic myctophids, penaeid shrimps, reef-dwelling serranids and tropical scombrids. The distributions are quite sharply peaked and cover a relatively narrow range, which immediately suggests that the growth parameter estimates which produce the values of $\phi$ at the lower or upper limits of the distributions should be checked (for example the single estimate of $\phi = 3.3$ in the Scombridae).

It also appears possible to obtain estimates of $K$ in those species whose range of $\phi$ values is known, if the asymptotic weight can be estimated. Another way of estimating $W_{\infty}$ is to use the relationship (Pauly 1979a)

$$W_{\text{max}}/0.86 \approx W_{\infty}$$

in which $W_{\text{max}}$ is the largest animal observed in a given stock and $W_{\infty}$ distinguishes a preliminary estimate of asymptotic length from one obtained by fitting the VBGF to growth data.

For example, applying Equation 4.58 to tropical scombrids, which have an overall $\phi$ range 2 to 3, the median value of $\phi = 2.5$ in conjunction with Equation 4.57 will provide a value of $K = 1.08 \text{ year}^{-1}$ for an asymptotic weight of 5,000 g and of $K = 0.23 \text{ year}^{-1}$ for an asymptotic weight of 50,000 g.
The ø' concept

If we assume that \( W_e = aL_e^3 \), the length equivalent of Equation 4.60 is

\[
ø' = \log_{10}(K) + \frac{2}{3}\log_{10}(a) + 2\log_{10}\left(\frac{L_e}{g_{102}}\right) \quad ...4.61)
\]

or

\[
ø' = \log_{10}(K) + 2\log_{10}(L_e) \quad ...4.62)
\]

in which

\[
ø' = ø - \frac{2}{3}\log(a) \quad ...4.63)
\]

Thus ø', which will have values different from ø, is an index for comparing the growth performance of fish in terms of their growth in length. Table 4.5 illustrates a case where the use of ø' values allowed the identification of a biased growth parameter estimate in the short mackerel (\( Rastrelliger brachysoma \)).

<table>
<thead>
<tr>
<th>Area</th>
<th>( L_e ) (FL, cm)</th>
<th>( K ) (yea (^{-1}))</th>
<th>ø'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inner Gulf of Thailand</td>
<td>20.9</td>
<td>3.38</td>
<td>3.17</td>
</tr>
<tr>
<td>Inner Gulf of Thailand</td>
<td>20.9</td>
<td>4.20</td>
<td>3.36</td>
</tr>
<tr>
<td>Gulf of Thailand (10°N, 100°E)</td>
<td>20.0</td>
<td>3.53</td>
<td>3.15</td>
</tr>
<tr>
<td>Gulf of Thailand (10°N, 100°E)</td>
<td>19.6</td>
<td>4.14</td>
<td>3.20</td>
</tr>
<tr>
<td>Indonesia (Tajung Satat)</td>
<td>22.9</td>
<td>2.28</td>
<td>3.08</td>
</tr>
<tr>
<td>Burma coast, uncorrected(^d)</td>
<td>27.0</td>
<td>0.96</td>
<td>2.84</td>
</tr>
<tr>
<td>Burma coast, corrected(^d)</td>
<td>27.0</td>
<td>1.60</td>
<td>3.07</td>
</tr>
</tbody>
</table>

\(^a\) From Pauly and Sann Aung (1984).
\(^b\) All growth parameter estimates based on length–frequency data, with growth curves fitted by eye by various authors, except in the case of data from Burma, which were fitted with the ELEFAN I program.
\(^c\) Raw length–frequency data, growth parameter estimated with ELEFAN I.
\(^d\) Length–frequency data corrected for gear selection, then growth parameters estimated with ELEFAN I.

It must be realized however, that ø' can be used only to compare the growth performance of fish with similar shapes; in this it differs from ø which, being based on weight, can also be used to compare the growth performance of fish of different shapes.

More on ø'

A preliminary knowledge of the range of ø' that can be expected in a given species of fish can be useful in the context of length-based analyses. Thus, abscissa scales with ø' values are provided for the K-scan routines associated with ELEFAN I and Shepherd's length composition analysis (see above), enabling users to distinguish between likely and less likely solutions. Preliminary analyses (Moreau et al. 1986) suggest that the C.V. of ø' for several stocks of the same species should not exceed 5 percent, which may provide some guideline as to which values of ø' are credible or not.

A more rigorous approach for dealing with ø' values, proposed by J. Hoening (pers. comm.) is, when possible (i.e. when size-at-age data are available), to estimate the standard error of ø', then to perform t-tests to assess the significance of whatever differences might have been observed. The standard error of ø' is derived using the delta method (see for example Seber 1982), which allows approximating the variance of functions with random variables.

Here, ø' is a function of the two random variables K and \( L_e \), and the standard error can be expressed as

\[
s.e.(ø') = \frac{1}{\log_{10}(a)^2} \left\{ \left[1/K^2\right](s.e.(K))^2 + \ldots \right. \\
\left. \ldots \left[4/(L_e)^2\right] \cdot s.e.(L_e) \right\} + \frac{4}{4} \cdot \left[4/(K \cdot L_e)^2 \cdot \text{cov}(K, L_e) \right] \quad ...4.64)
\]

Similarly, when \( W_e \) is given,
s.e.(\(\varphi\)) = \left(\frac{1}{\log_{10}}\right)^2 \left\{ \frac{1}{K^2} \cdot \text{s.e.}(K) + \ldots \right\} 
\ldots \left[ \frac{1}{W^2} \cdot \text{s.e.}(W) \right] + \left\{ \frac{1}{KW} \cdot \text{cov}(K, W) \right\} \ldots \text{4.65}

The approximate 95 percent confidence intervals of the estimated \(\varphi'\) can then be obtained though
\[ \varphi' \pm 2 \cdot \text{s.e.}(\varphi') \ldots \text{4.66} \]

with the assumption that the sample size is large enough to justify a normal distribution for the estimators.

**FiSAT** outputs \(\varphi'\) and its standard error, based on these equations, when it fits the VBGF to length-at-age data.

---

**Chapter 5. Mortality, Gear Selection and Recruitment**

**What you will learn from this chapter**

In this chapter, you will learn about length-based models that can be used to derive mortality estimates, about a methodology for deriving seasonal recruitment patterns, and procedures for deriving trawl mesh or gillnet selection curves.

**Introduction**

In fishery biology, the most useful way to express the decay (= decrease) through time of a group of fish born at the same time (a cohort) is by means of "instantaneous" rates. These rates, of which there are three (\(Z\), \(M\) and \(F\)), are defined by two equations, of which the first is

\[ N_{i+1} = N_i \cdot e^{Z(t_{i+1} - t_i)} \ldots 5.1 \]

where \(N_i\) is the (initial) number of fish at time \(t_i\), and \(N_{i+1}\) is the number of remaining fish at time \(t_{i+1}\); \(Z\) is the instantaneous rate of total mortality. An advantage of such decay rates is that they can be added or subtracted. Thus, we have the second equation:

\[ Z = M + F \ldots 5.2 \]

where \(M\) is the instantaneous rate of natural mortality and \(F\) the instantaneous rate of fishing mortality. Obviously, when \(F = 0\), \(Z = M\), which means that natural and total mortality have the

---

31 Adapted from Pauly (1984a).
same value when there is no fishing, i.e. in an unexploited stock (Fig. 5.1).

![Image: Decrease of a cohort starting with 100 fish, subjected to different levels of natural mortality (M), and to fishing mortality (F) from age $t_0$.](image)

Fig. 5.1. Decrease of a cohort starting with 100 fish, subjected to different levels of natural mortality (M), and to fishing mortality (F) from age $t_0$. on.

Instantaneous rates (i.e. "exponential" rates) of mortality can be converted to the fraction surviving through equations such as

$$S = \frac{N_{i+1}}{N_i} \quad \text{...5.3) }$$

where $S$ is the fraction surviving at time $t_{i+1}$, while

$$A = 1 - S \quad \text{...5.4) }$$

is the fraction of the stock dead at time $t_{i+1}$. Although used by a few authors, fractional (and percentage) mortalities are not further discussed in this manual, because they are too cumbersome to handle in comparison with instantaneous rates (see Beverton and Holt 1956, p. 68).

Mortalities, whether expressed as instantaneous rates or as fractions, always refer to a certain period of time. Throughout this manual, the year is used as the conventional unit, unless mentioned otherwise.

Fisheries biologists have two main jobs as far as mortalities are concerned:

- to estimate total mortality;
- to split their estimates of total mortality, where appropriate, into separate estimates of natural and fishing mortalities.

A number of methods are proposed here by which these aims can be achieved, given suitable inputs.

**Total mortality from catch curves**

The most common approach in estimating total mortality is through a so-called catch curve, i.e. a plot of the logarithms of the numbers of survivors vs. their age. This method is based on the assumption of a constant parameter system.

Consider, as an example, a fish stock during the period 1971–1975 with a life span of five years. Let the numbers of survivors be those given in Table 5.1A. For simplicity, we assume only one cohort per year, recruiting on January 1. Note that a specific cohort can be followed diagonally across the table. Every year, the survivors of that cohort advance to a new age group. Each column, on the other hand, contains the numbers of survivors of five different cohorts on 1 January.

---

32 Adapted from Sparre and Venema (1992).
Table 5.1. Illustration of the “variable parameter system” (A) and the “constant parameter system” (B; see text).

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>1971</th>
<th>1972</th>
<th>1973</th>
<th>1974</th>
<th>1975</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: VARIABLE PARAMETER SYSTEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2105</td>
<td>1111</td>
<td>9560</td>
<td>1869</td>
<td>1236</td>
</tr>
<tr>
<td>1</td>
<td>2575</td>
<td>736</td>
<td>405</td>
<td>3817</td>
<td>618</td>
</tr>
<tr>
<td>2</td>
<td>155</td>
<td>1097</td>
<td>281</td>
<td>142</td>
<td>1193</td>
</tr>
<tr>
<td>3</td>
<td>102</td>
<td>58</td>
<td>298</td>
<td>109</td>
<td>59</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>38</td>
<td>18</td>
<td>138</td>
<td>29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B: CONSTANT PARAMETER SYSTEM</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2560</td>
<td>2560</td>
<td>2560</td>
<td>2560</td>
<td>2560</td>
</tr>
<tr>
<td>1</td>
<td>942</td>
<td>942</td>
<td>942</td>
<td>942</td>
<td>942</td>
</tr>
<tr>
<td>2</td>
<td>346</td>
<td>346</td>
<td>346</td>
<td>346</td>
<td>346</td>
</tr>
<tr>
<td>3</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
</tr>
<tr>
<td>4</td>
<td>47</td>
<td>47</td>
<td>47</td>
<td>47</td>
<td>47</td>
</tr>
</tbody>
</table>

Altogether, Table 5.1A contains parts of nine different cohorts; of these, only one is represented in all years, i.e. from 1971 to 1975 (the 1971 cohort). As might be seen, recruitment \((N_r)\) has varied from 1971 to 1975; the 1971 cohort and the 1973 cohort were strong, whereas the 1972 cohort was a weak one (“strong” means that \(N_r\) is considerably greater than average recruitment).

Now suppose that recruitment remains constant every year, as shown in Table 5.1B; assume further that \(F\) and \(M\) also remain constant. In this case, the number of survivors along the diagonals is the same as along the columns. Thus, in the case of constant recruitment, considering one cohort over a longer period is the same as considering all different cohorts (the entire stock) in one particular year.

We know that the assumptions of a constant parameter system are never strictly fulfilled in reality. However, it is only by making such assumptions that we are able to carry out an analysis of available data; also it appears that the data points used to define catch curves (as defined below) usually fall along straight lines, suggesting that recruitment varies little, at least past the early juvenile stage.

If steady-state is assumed and \(Z\) is assumed to be constant throughout the life of a cohort, the catch can be derived from Equation 5.1 as

\[
C_{i+1} = \left(\frac{F}{Z}\right) \cdot \left(N_i - N_{i+1}\right)
\]

Equation 5.5 is the "catch equation" (sometimes called "Baranov's equation"—Baranov 1918), where \(F/Z\) expresses the fraction of the mortality caused by fishing (also called "exploitation rate", \(E\)).

Equation 5.6 is not linear in \(t_i\) (nor in \(t_{i+1}\)), but linearization can be achieved by replacing \(N_i\) using Equation 5.5 which results in

\[
C_{i+1} = N_i \cdot e^{-Z \Delta t} \cdot \left(\frac{F}{Z}\right) \cdot \left(1 - e^{-Z \Delta t}\right)
\]

where \(\Delta t = t_{i+1} - t_i\).

If we take the logarithm of Equation 5.7 and rearrange the terms, we obtain

\[
\ln(C_{i+1}) = \ln\left[N_i \cdot \left(\frac{F}{Z}\right)\right] + Z \Delta t - Z t_i + \ln\left(1 - e^{-Z \Delta t}\right)
\]

Here, \(t_i\) appears only in one linear term, \(-Zt_i\), while the term \(\ln\left[N_i \cdot \left(\frac{F}{Z}\right)\right] + Z \Delta t\) can be treated as constant (a) since \(N_i, t_i, F\)
and Z are also assumed to remain constant. Thus, we can simplify Equation 5.8 by replacing these terms by a, i.e.:

\[
\ln(C_{i+1}) = a - Zt_i \ln(1 - e^{-Zt_i}) \quad \ldots 5.9
\]

A linear expression in \(t_i\) (and \(t_{i+1}\)), has been obtained, except for the last term, \(\ln(1 - e^{-Zt_i})\). There are several ways to deal with this term (Edser 1908; Heincke 1913; Baranov 1918; Ricker 1975); two of these are presented below.

**Age-structured catch curve**

The non-linear term (Equation 5.9) can be included in the intercept term. Thus, if the constant a is redefined to include the term \(\ln(1 - e^{-Zt_i})\), Equation 5.9 reads

\[
\ln(C_{i+1}) = a - Zt_i \quad \ldots 5.10
\]

Equation 5.10 is the catch curve equation for constant time intervals, also known as "age-structured catch curve".

Fig. 5.2 illustrates the estimation of Z based on an age-structured catch curve. Data for the three first age groups have been excluded from the regression analysis, since they do not fit on the straight line corresponding to a constant Z. The most probable reason is that these age groups were not under full exploitation. Their fishing mortality is lower because they escape through the meshes of the trawl-net (most of the North Sea whiting are caught by trawls with codend mesh sizes from 70 mm to 110 mm). Another reason for expecting a reduced fishing mortality on the small fish is that most are still in their nursery, and are not yet recruited to the fishing grounds.

**Length-converted catch curves (without seasonality of growth)**

Length-converted catch curves are functionally equivalent to age-structured catch curves and rest on the same assumptions, but are commonly used where age data (for example otolith or scale readings) are difficult to obtain.

Here again, the basic model consists of a plot of the natural logarithm of the numbers of fish caught in various age groups \(C_i\) against their corresponding age or age group \(t_i\), i.e.:

\[
\ln(C_i) = a + b \cdot t_i \quad \ldots 5.11
\]

and Z is estimated from the slope b, with sign changed, of the descending right arm of the plot.

---

33 Adapted from Sparre and Venema (1992).

34 Adapted from Pauly (1984a).
The following assumptions are involved here:

- Z is the same in all age groups used in the plot.
- All age groups used in the plot were recruited with the same abundance (or the recruitment fluctuations were small and of random character).
- All age groups used in the plot are equally vulnerable to the gear used for sampling.
- The sample used is large enough and covers enough age groups to effectively represent the average population structure over the period considered.

In practical terms, the estimation of Z from a length-converted catch curve then involves the following steps (assuming that the sample used is large enough and covers enough age groups to effectively represent the average population structure over the period of time considered):

(i) pooling of catch data to obtain a single, large sample representative of the population for the period under consideration;

(ii) construction of the catch curve proper, using the large sample in (i) and a set of growth parameters;

(iii) estimation of Z from the descending right arm of the catch curve.

Pooling of samples (such as monthly catches) over a relatively long period (at least one year) is particularly important in short-lived fish and shrimps, as their whole population structure is affected by seasonal "pulses" of recruitment, generally one or two per year (Pauly and Navaluna 1983). Also, to prevent a single, larger (monthly) sample from unduly affecting the total (annual) sample, the samples may all be given the same weight, by conversion to percentages prior to adding to obtain a single overall sample.

There are many alternatives to a scheme where each sample is given the same weight. For example, it might be more appropriate to weight the samples by the square root of their size when the fishery catch is not known, or by the catch when it is. However, empirical studies into appropriate sample sizes and weighting factors for length-converted catch curves are still lacking. Thus, it will be best to always compare the results obtained using the three weighting options incorporated in FISAT, and to use their correspondence (or the lack thereof) as a basis for further decisions or studies.

There are also several ways to construct a length-converted catch curve. However, they must all account for the fact that fish growth in length is not linear, but slows down as length and age increase. This slowing down has the effect that older size groups contain more age groups than do younger size groups. In other words, it takes larger fishes longer to "leave" a certain size group—they "pile-up" (Baranov 1918), or "stack-up" (van Sickle 1977) in the size classes pertaining to old, large, slow-growing fish. Correcting for this effect is rather straightforward, and three methods will be discussed here.

The first approach, analogous to but improved upon those discussed in Ricker (1975, p. 33 and pp. 60–64) and van Sickle (1977), consists of multiplying the number in each length class by the growth rate of fish in that class. This results in a catch curve equation of the form

\[
\ln(C_i) \cdot \left(\frac{\Delta L_i}{\Delta t}\right) = a + b \cdot t_i'
\]

...5.12)

where \(\Delta L_i/\Delta t\) is the growth rate and \(t_i'\) the relative age corresponding to length class i. In practice, \(\Delta L_i/\Delta t\) can be estimated from the VBGF as the growth rate pertaining to the midlength of length class i, while \(t_i'\) can be estimated as the relative age corresponding to the median of class i as estimated, using the appropriate growth parameters, through conversion using the VBGF. Relative ages are used here because using \(t_i\)
(which leads to absolute ages) is not necessary in conjunction with catch curves, where Z is estimated from a slope.

Equation 5.12 allows ready estimation of the bias caused by not accounting for the "pile-up" effect mentioned above. This is done by first rewriting it as

\[ \ln(C) - \ln(\Delta L/\Delta t) = a + b \cdot t' \]  

or

\[ \ln(C) = a + b \cdot t' - \ln(\Delta L/\Delta t) \]  

In terms of the VBGF, growth rate (\(\Delta L/\Delta t\)) can be computed from

\[ (\Delta L/\Delta t) = \ln(K \cdot L_\infty) + K \cdot (t' - t_o) \]  

where K, L_\infty and t_o are parameters of the VBGF, and t', the relative age corresponding to a given midlength. Inserting Equation 5.15 in Equation 5.14 gives

\[ \ln(C) = a + b \cdot t' - \ln(K \cdot L_\infty) - K \cdot (t' - t_o) \]  

or

\[ \ln(C) = a + b \cdot t' - \ln(K \cdot L_\infty) - K \cdot t' + K \cdot t_o \]  

Equation 5.17, it will be noted, has two constant terms with regard to the variables N and t', namely, \(\ln(K \cdot L_\infty)\) and K \cdot t_o.

Since Z in Equation 5.12 is estimated as a slope, these two constant terms can be grouped into one single new term (a), which becomes the intercept of a new equation of the form

\[ \ln(C) = a + b \cdot t' - K \cdot t' \]  

which gives, rearranged

\[ \ln(C) = a + (b - K) \cdot t' \]  

as a new equation for a length-converted catch curve, wherein \((b + K) = Z\).

It follows from this that the bias resulting from not considering the "pile-up" effect (i.e. from using ln(N) instead of ln(N \cdot \Delta L/\Delta t) as ordinate of a length-converted catch curve) is equal to K.

Two practical applications of this finding come to mind:

- It becomes possible to correct biased values of Z obtained by authors who did not account for the "pile-up" effect (by simply adding K to their (biased) estimate of Z; see for example Berry 1970, Nzioka 1983).
- The estimation of Z from a length-converted catch curve becomes simpler, since one can first ignore the "pile-up" effect then compensate for it by adding K to the absolute value of the curve's slope.

Another, now widely used, type of length-converted catch curve is defined by the equation

\[ \ln(C_i/\Delta t_i) = a + b \cdot t'_i \]  

where C_i and t'_i are as used in Equation 5.12, and where \(\Delta t_i\) is the time needed, on average, by the fish to grow through length class i; \(\Delta t_i\) is estimated from

\[ \Delta t_i = \left(1/K\right) \cdot \ln \left[ \frac{L_{i+1} - L_i}{L_\infty - L_i} \right] \]  

and t'_i is estimated from

\[ t'_i = \left(1/K\right) \ln \left[ 1 - \left(L_i/L_\infty\right) \right] \]
where $L_i$ is the mid-point of length class $i$, i.e., its "midlength". An application example of this model is given in Fig. 5.3.

![Fig. 5.3. A length-converted catch curve for *Upeneus moluccensis*, Ragay Gulf, Philippines ($L_{102} = 19.5$ cm and $K = 0.9$ year$^{-1}$). The first point to be included in the estimation of $Z(P_1)$ is clearly defined (see text).](image)

Equation 5.20 accounts for the "piling-up" effect through division of the $C_i$ values by $\Delta t_i$, the inverse of the growth rates by which the $C_i$ values are multiplied in Equation 5.11. Hence, Equation 5.20 is a slightly modified version of Equation 5.12, and its properties, for example with regard to not accounting for the "piling-up" effect, are the same.

When $K$ is not known, Equations 5.12 and 5.20 can still be used; in such cases, a value of unity (1) should be used instead of $K$ when computing the relative ages. The slope of the catch curve, with sign changed, will then be equal to $Z/K$.

**Jones and van Zalinge cumulative plot**

The Jones and van Zalinge cumulative plot is a variant of the length-converted catch curve approach, but rests on slightly different assumptions.

We recall the basic equation of age-structured catch curves (Equation 5.9). If we let $t_{i+1}$ take a very high value (i.e. $t_{i+1} \to \infty$), the term $e^{-Z(L_{i+1}-L_i)}$ will approach zero and $\ln(1-e^{-ZK})$ will therefore approach zero as well. Thus, if $C_{i,\infty}$ designates all fish caught at age $t_i$ and older, there is a linear relationship (Jones and van Zalinge 1981):

$$
\ln(C_{i,\infty}) = d - Z \cdot t_i \quad \ldots \ldots 5.23
$$

If $t_i$, the age in Equation 5.23, is replaced by the inverse of the VBGF, we obtain

$$
\ln(C_{i,\infty}) = d - Z \left( \frac{1}{K} - \ln(1 - \frac{L_i}{L_{\infty}}) \right) \quad \ldots \ldots 5.24
$$

which can be reduced to the form

$$
\ln(C_{i,\infty}) = a + (Z/K) \cdot \ln(L_{\infty} - L_i) \quad \ldots \ldots 5.25
$$

where $C_{i,\infty}$ is the cumulative catch (computed from the highest length class with non-zero catch) corresponding to a given length class, $i$ is the lower limit of that length class, and the $\infty$ symbol indicates that the catch refers to a range from $L_i$ to all larger sizes. Fig. 5.4 and Table 5.2 show an application of Equation 5.25.

---

35 Adapted from Sparre and Venema (1992).
As in catch curves, it is necessary when applying this method to first draw a scattergram of the computed values, then to select visually the points belonging to the straight segment of the plot, representing fully selected and recruited fish.

This method can be applied also when K is unknown. In this case a value of one (1) is entered instead of K; this results in the relative ages being defined as

$$t' = (t - t_0) \cdot K$$  \hspace{1cm} ...5.26

which leads to a slope equal to Z/K.

The Jones and van Zalinge method is extremely sensitive to the values of the catches in the largest size groups, even when they are not included in the linear regression. Thus, it should not be used when the catch composition data used were obtained from gears that markedly select for or against very large fish. Also, the statistics of the plot (correlation coefficient, confidence intervals, etc.) must be taken with a grain of salt, since they are based on data points that are not independent of each other. Finally, the method does not take account of seasonal growth oscillations, and thus should not be used when such oscillations are known to occur (see below).

Length-converted catch curves accounting for seasonal growth

The methods presented above for constructing length-converted catch curves (and the Jones and van Zalinge method as well) all overestimate Z when growth is not seasonal (Sparre 1990).

---

36 Adapted from Pauly (1990).
The reason for this is their use of the standard VBGF for estimating $\Delta t$ and $t'$, the time needed to grow through length class $i$ and the relative age corresponding to $i$, respectively:

- There is only one value of $\Delta t$ for any length class, i.e. the $L_i$ and $L_{i+1}$ values are completely determined by $L_c$ and $K$.
- There is only one value of $t'$ for any length class, i.e. $L_t$ is completely determined by $L_c$ and $K$.
- Hence, class-specific $N$ values can be added across samples (i.e. time) without effect on the values of $\Delta t$ and $t'$.

These features do not apply in the case of seasonal growth, where $\Delta t$ and $t'$ depend not only on length (as implied by the standard VBGF) but also on the time of the year—for example $\Delta t$ is larger during the season of slow growth ("winter") and smaller in "summer".

Seasonal growth can however be represented by the curve of Hoenig and Chaudary Hanumara (1982) and Somers (1988), i.e.:

$$L_t = L_c \left[ 1 - e^{-(\frac{K}{2\pi}) \sin(t-t_c)} \right]$$  \hspace{1cm} \text{(5.27)}

where $S(t) = (\frac{CK}{2\pi}) \sin(t-t_c)$, and $S(t_c) = (\frac{CK}{2\pi}) \sin(t_c-t_c)$.

and all other parameters are as defined in Equation 4.8.

Seasonal growth oscillations imply that in a given sample, $\Delta t$ depends not only on $L_i$, $L_{i+1}$, $L_c$ and $K$, but also on $C$ and, more importantly, on the difference between WP and the date the sample in question was obtained. Hence, $N$ values pertaining to the same value of $L_i$ but different samples cannot be added across time (as is usually done with length-converted catch curves), because they are all associated with different values of $\Delta t$. Similarly, there is no one-on-one correspondence between $L_t$ and $t'$ (as implied by Equation 5.14), because (for example) one-year-old fishes will have very different sizes depending on whether they hatched before or after the winter period of reduced growth.

The following steps may be applied to resolve this (and are implemented in FISAT; see Fig. 5.5):

---

**Fig. 5.5. Schematic representation of the method for constructing length-converted catch curves accounting for seasonality.**

**A:** The first operation is using the parameters of a seasonally oscillating growth curve to identify a number of (pseudo) cohorts, i.e. fish between two successive growth curves; the next step is adding fish belonging to different samples, but to the same (pseudo) cohort to obtain successive values of $C_i$.

**B:** Construction of catch curve as a plot of $\ln(C_i)$ vs. relative ages, and estimation of $Z$ from straight descending arm.

**C:** Standard length-converted catch curve, also based on data in A, but not accounting for seasonal growth. Note overestimation of $Z$. 

---
• Create a length–frequency file in which all fishes are assumed to have been caught within the same period of one year (since no inter-year differences of growth or mortality are assumed to occur).

• Estimate, by solving Equation 5.20 iteratively, the (relative) age of the youngest and oldest fish in the file ($t'_{\text{min}}$ and $t'_{\text{max}}$ respectively).

• Divide the time difference, $t'_{\text{max}} - t'_{\text{min}}$, by the number of length classes in the file, such as to obtain a number of time intervals ($n$) equal to the number of data points that would have been obtained from the corresponding standard length-converted catch curve (some value of $n$ has to be used, and the proposed value has the advantage of facilitating comparisons between different types of catch curves).

• For each interval, starting from $t'_{\text{min}}$, and moving backward along the time axis, draw successive growth curves at regular (time) intervals, and add up across samples all fishes ($C_i$) between the two growth curves defining an interval. (This step, which is illustrated in Fig. 5.5A, is equivalent to adding up, across samples, the fish within defined length class limits.)

• Plot the $\ln(C_i)$ values thus obtained against the midpoint of the relative age intervals ($t'_i$), and estimate $Z$ from the slope of the right, descending arm of the curve.

It should be noted that this approach for constructing catch curves from length–frequency data gives exactly the same results as the standard method when $C = 0$, i.e. when growth is not seasonal, and it can handle any number of cohorts per year. This would make the new method universally applicable were it not for the fact that the left (ascending) arm of the new catch curve cannot readily be used to assess the impact of size-specific gear selection or recruitment (see below, Fig. 5.17). Hence, standard length-converted catch curves will continue to be helpful.

Selecting points for the regression

The selection of those data points that can be used for estimating $Z$ after a catch curve has been constructed is crucial, whatever the type of catch curve. Two important selection criteria are:

• The points belonging to the left, ascending arm of the curve must not be included because they represent incompletely selected and/or incompletely recruited fish.

• Generally, points must not be included which do not align themselves with the rest of the points, especially if they represent age (size) groups for which sample sizes are relatively small.

The second criterion must not be misunderstood as providing an excuse for the whole deletion of points until one's preconceived notion of linearity is achieved; rather it allows deletion of one point or two at most. When the curve as a whole seems to deviate from linearity, the appropriate next step is to test whether this deviation is significant or not, using any of the statistical tests available for this purpose (e.g. Guilford and Fruchter 1978, pp. 277–280).

In addition to two general criteria above, it should be realized that in length-converted catch curves (and in the Jones and van Zalinge method), the conversion of length to (relative) ages by means of the VBGF, when involving fish whose length is very close to $L_v$, often generates unrealistically high "ages"

37 Adapted from Pauly (1990).
which cannot be included either. Such points usually occur when lengths are within 5 percent of \( L_c \).

To select the points to be used when estimating \( Z \) from a catch curve, a routine was implemented in **FISAT** which automatically selects a set of appropriate points. These points are *default* values, and need not be used in every case—they can easily be overwritten and other points used. They are identified using the following subsidiary rules:

1. Take only points which lead to a regression with a slope significantly less than zero (this ensures that only the descending, right-arm of the curve is considered).

2. Use as first point (\( P_1 \)) the one immediately to the right of the highest point (\( P_{\text{max}} \)), because \( P_{\text{max}} \) is likely to be contaminated by incomplete selection/recruitment (Robson and Chapman 1961).

3. If \( P_{\text{max}} \) is above the leftward projection of an initial regression line that did not include \( P_{\text{max}} \), it was not contaminated and should be included (this ensures that \( P_{\text{max}} \) is deleted only if it was contaminated).

4. Include the rightmost point (\( P_{\text{last}} \)) of the curve only if it lies below the rightward projection on an initial regression line that did not include \( P_{\text{last}} \) (this ensures that \( P_{\text{last}} \) is not included when it was derived from a length close to \( L_c \), and hence shifted to the right).

5. Estimate \( Z \) from the sequence of points \( P_1, P_2, \ldots P_n \) which maximizes:

\[
F_{\text{ratio}} = \left[ \frac{\sum (y_i - \bar{y})^2}{\sum y_i^2 - (\sum y_i)^2/n} \right]/n - 2 \quad \ldots (5.28)
\]

where \( x, y \) are the coordinates of the points included (Sokal and Rohlf 1981). This can be done by scanning all possible combinations of points for which rules (1) to (4) apply.

6. Rules (1) to (5) may still fail in cases where catch curves display a wide, flat plateau, or two (or more) descending sequences of points separated by an ascending sequence, or other pathological cases.

Note that the necessary (though not sufficient) condition for a catch curve estimate of \( Z \) to be reliable is that it be derived from a regression line as straight as possible (implying a constant value of \( Z \)), and including as many age groups as possible (such as to dampen the effects of fluctuations of mortality and of recruitment). In statistical terms, this means that a large fraction of the variance of the points \( P_1, P_2, \ldots P_n \) should be explained by the regression, and also that the value of \( n \) should be as large as possible. A high \( F_{\text{ratio}} \), estimated as per Equation 5.28, ensures this.

For cases where the above approach fails, **FISAT** makes it possible to overwrite the defaults and to estimate \( Z \) through another set of points.

### \( Z \) from Hoenig's models\(^{38} \)

Following a number of works from authors who had demonstrated the existence of a strong relationship between the longevity of fish (in the wild) and their mortality, Hoenig (1982) assembled data on a large number of aquatic animals (molluscs, fish and cetaceans) from which he derived the empirical relationship

\[
\ln(Z) = 1.44 - 0.984 \ln(t_{\text{max}}) \quad \ldots (5.29)
\]

\(^{38}\) Adapted from Hoenig (1982).
where \( t_{\text{max}} \) is the maximum age (in years) observed in a given stock, and \( Z \) is defined as above.

Although the "fit" of Equation 5.29 is rather good \( (r^2 = 0.82 \) for 130 data pairs), it should be realized when using this model that the estimates of \( Z \) thus obtained are very approximate, and should therefore be revised as additional information becomes available. Table 5.3 gives some examples of the application of the equation.

Table 5.3. Maximum observed size \((L_{\text{max}}, W_{\text{max}})\), maximum observed age \((t_{\text{max}})\) and estimated mortality \((Z)\) for 12 coral reef fish of New Caledonia.\(^a\)

<table>
<thead>
<tr>
<th>Family</th>
<th>Species</th>
<th>( L_{\text{max}} ) ((\text{SL, cm}))</th>
<th>( W_{\text{max}} ) ((\text{live weight, g}))</th>
<th>( t_{\text{max}} ) ((\text{years}))</th>
<th>( Z ) ((\text{year}^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holocentridae</td>
<td>Adioryx spinifer</td>
<td>25.8</td>
<td>572</td>
<td>13</td>
<td>0.34</td>
</tr>
<tr>
<td>Serranidae</td>
<td>Epinephelus summa</td>
<td>20.8</td>
<td>263</td>
<td>16</td>
<td>0.28</td>
</tr>
<tr>
<td>Carangidae</td>
<td>Caranx ignobilis</td>
<td>76.4</td>
<td>10,765</td>
<td>9</td>
<td>0.49</td>
</tr>
<tr>
<td>Lutjanidae</td>
<td>Lutjanus argentimaculatus</td>
<td>60.7</td>
<td>5,780</td>
<td>18</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>Lutjanus gibbus</td>
<td>37.0</td>
<td>1,735</td>
<td>18</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>Lutjanus sebae</td>
<td>69.5</td>
<td>13,810</td>
<td>35</td>
<td>0.13</td>
</tr>
<tr>
<td>Pomadasyidae</td>
<td>Plectorhynchus chaetodonoides</td>
<td>43.1</td>
<td>2,715</td>
<td>21</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>Plectorhynchus pictus</td>
<td>39.2</td>
<td>1,970</td>
<td>11</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>Pomadasys hasta</td>
<td>31.8</td>
<td>87.3</td>
<td>12</td>
<td>0.37</td>
</tr>
<tr>
<td>Lethrinidae</td>
<td>Lethrinus harak</td>
<td>24.3</td>
<td>450</td>
<td>15</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>Lethrinus obsolatus</td>
<td>25.0</td>
<td>501</td>
<td>14</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>Monotaxis grandoculis</td>
<td>39.2</td>
<td>2,730</td>
<td>11</td>
<td>0.40</td>
</tr>
</tbody>
</table>

\(^a\)Size and age data adapted from Loubens (1980, Table VI); the values of \( t_{\text{max}} \) are based on limited samples (sample sizes not given) which, however, contained large-sized adults.

\(^b\)Estimated from Equation 5.29.

When, in addition to \( t_{\text{max}} \) and \( t_c \), the size of the sample from which \( t_{\text{max}} \) was determined is also known, it becomes possible to estimate \( Z \) and its standard error \((\text{s.e.}(Z))\) from the relationships derived by Hoenig and Lawing (1982):

\[
Z = \frac{1}{c_1(t_{\text{max}} - t_c)}
\]

and

\[
\text{s.e.}(Z) = \sqrt{c_2 \cdot Z^2}
\]

where \( c_1 \) and \( c_2 \) are coefficients whose values depend on \( n \) (see Table 5.4).

Table 5.4. Table of coefficients for estimating \( Z \) and its standard error using Equations 5.30 and 5.31 (from Hoenig and Lawing 1982, as incorporated in FISAT).

<table>
<thead>
<tr>
<th>( n )</th>
<th>( c_1 )</th>
<th>( c_2 )</th>
<th>( n )</th>
<th>( c_1 )</th>
<th>( c_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.583</td>
<td>0.416</td>
<td>10</td>
<td>0.405</td>
<td>0.196</td>
</tr>
<tr>
<td>15</td>
<td>0.344</td>
<td>0.142</td>
<td>20</td>
<td>0.311</td>
<td>0.117</td>
</tr>
<tr>
<td>25</td>
<td>0.290</td>
<td>0.102</td>
<td>30</td>
<td>0.274</td>
<td>0.091</td>
</tr>
<tr>
<td>35</td>
<td>0.263</td>
<td>0.084</td>
<td>40</td>
<td>0.253</td>
<td>0.078</td>
</tr>
<tr>
<td>45</td>
<td>0.245</td>
<td>0.074</td>
<td>50</td>
<td>0.239</td>
<td>0.070</td>
</tr>
<tr>
<td>55</td>
<td>0.233</td>
<td>0.067</td>
<td>60</td>
<td>0.228</td>
<td>0.064</td>
</tr>
<tr>
<td>65</td>
<td>0.224</td>
<td>0.062</td>
<td>70</td>
<td>0.220</td>
<td>0.060</td>
</tr>
<tr>
<td>75</td>
<td>0.217</td>
<td>0.058</td>
<td>80</td>
<td>0.214</td>
<td>0.057</td>
</tr>
<tr>
<td>90</td>
<td>0.208</td>
<td>0.054</td>
<td>100</td>
<td>0.204</td>
<td>0.052</td>
</tr>
</tbody>
</table>

\(^a\)Interpolate for intermediate values of \( n \).

Hoenig and Lawing (1982), whose paper should be consulted for the derivation of Equations 5.29 and 5.30 and Table 5.4, stressed that "fast growing, short-lived species with minimal variability in length about age are best suited for this method\(^b\). This is so because in such cases \( n \), the sample size, is not the number of fish actually aged, but the number of fish from which a subsample consisting of the largest fish was taken. Thus if, say, 200 fish have been inspected, from which the 20 largest were selected for aging, the value of \( n \) will be
Table 5.5. Maximum reported age and estimated total mortality of selected Brazilian freshwater and marine fish.a

<table>
<thead>
<tr>
<th>Species</th>
<th>$t_{	ext{max}}$ (year)</th>
<th>n</th>
<th>Location, sampling date(s)</th>
<th>Author(s) (year)</th>
<th>Estimated $Z$ (year$^{-1}$)</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Trachychorystes galeatus</em></td>
<td>3.5</td>
<td>83</td>
<td>Banabuiú Reservoir</td>
<td>Nomura et al. (1976)</td>
<td>1.35</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>3.5</td>
<td>99</td>
<td>Caera State, 1971</td>
<td></td>
<td>1.40</td>
<td>0.32</td>
</tr>
<tr>
<td><em>Prochilodus scrofa</em></td>
<td>13</td>
<td>451</td>
<td>Mossi Guassu River, Sao Paulo State, 1947</td>
<td></td>
<td>0.50</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>485</td>
<td></td>
<td>Godoy (1959)</td>
<td>0.73</td>
<td>0.13</td>
</tr>
<tr>
<td><em>Plagioscion squamosissimus</em></td>
<td>6</td>
<td>103</td>
<td>Amanari Reservoir, Caera State, 1960–62</td>
<td>Nomura and Olivera (1960–62)</td>
<td>0.82</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>134</td>
<td></td>
<td></td>
<td>0.74</td>
<td>0.16</td>
</tr>
<tr>
<td><em>Micropogon furnieri</em></td>
<td>6</td>
<td>229</td>
<td>Off Iguape, Caera State, 1966–67</td>
<td>Rodrigues (1968)</td>
<td>0.96</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>115</td>
<td></td>
<td></td>
<td>0.72</td>
<td>0.16</td>
</tr>
<tr>
<td><em>Micropogon ancyodon</em></td>
<td>11</td>
<td>9,947</td>
<td>Off Sao Paulo, 1975–76</td>
<td>Lara (1951)</td>
<td>0.66</td>
<td>0.11</td>
</tr>
</tbody>
</table>

a Total mortality and its standard error estimated from equations 5.30 and 5.31, with $t_c$ set at zero because very small fish were included in the catch samples.

Table 5.5 gives values of $Z$ and its standard error as obtained by application of Equations 5.30 and 5.31.

**Total mortality from mean length**

**Beverton and Holt model**

The model proposed by Beverton and Holt (1956) is based on a deterministic expression for the mean length of fish in the catch above a length at first capture, $L_c$. Assuming a population in equilibrium, the mean length ($\bar{L}$) in the catch is given by

$$\bar{L} = \frac{F}{N_t} \int L N(L) dL = \frac{L'}{N_t} \int L N(L) dL$$

...5.32)

where population size can be expressed by an exponential mortality model, such that

$$N_t = N_t' e^{-(F+M)(1-t')}$$

...5.33)

where $t'$ is the smallest age, $L'$ the corresponding length at which fish are fully represented in catch samples, and $N_t'$ the corresponding number of fishes. Substituting Equation 5.33 into Equation 5.32 and integrating from $t'$ to $\infty$ result in an explicit expression for average length:

$$\bar{L} = L_c \left(1 - \frac{F + M}{F + M + K} \right) e^{-\frac{K}{L_c} (t' - L')\int e^{-\frac{K}{L_c} (t' - L')}}$$

...5.34)

$L'$ may be expressed as a function of $t'$:

$$e^{-\frac{K}{L_c} (t' - L')} = \frac{(L_c - L')}{L_c}$$

...5.35)

Substitution of Equation 5.35 into Equation 5.34 and rearrangement of terms result in the Beverton and Holt (1956) formula

$$Z = K \cdot \frac{(L_c - \bar{L})}{(L_c - L')}$$

...5.36)

where $L'$ is the smallest length of fish fully represented in the length–frequency data at hand (corresponding to the lower class limit of $P_1$ in length-converted catch curves), and $\bar{L}$ is the mean length of all fish $\geq L'$.

**Ault and Ehrhardt method**

Adapted from Ault and Ehrhardt (1991).

Adapted from Ault and Ehrhardt (1991).
Although this is not usually not made explicit, Equation 5.36 assumes that the species studied exhibit infinite life spans, i.e. that they are extremely long-lived. In the tropics, many species have high rates of growth and natural mortality and are therefore relatively short-lived. Tropical fisheries are usually characterized, moreover, by numerous and heterogeneous artisanal fleets which have restricted operational ranges and highly selective gears. These conditions often result in length–frequency distributions excluding the largest lengths. This results in biases for estimates of Z based on the Beverton and Holt estimators because the maximum age in the catch is very small, much less than the high ages for which the assumption of "infinite" age makes no differences.

This can be accounted for by replacing $\infty$ by $t_{max}$ in Equation 5.32, which leads to

$$\bar{L} = \frac{\int_{t}^{t_{max}} L_{0} \, dL}{\int_{t}^{t_{max}} N_{0} \, dt} \quad \ldots 5.37$$

Use of the VBGF and of Equation 5.33 for expressing $L_{t}$ and $N_{t}$ and their incorporation into Equation 5.37, along with integration and some simplification, lead to a new relationship:

$$\left[ \frac{L_{x} - L_{max}}{L_{x} - L'} \right]^{Z} = \left[ \frac{Z(L' - \bar{L}) + K(L_{x} - \bar{L})}{Z(L_{max} - \bar{L}) + K(L_{x} - \bar{L})} \right] \quad \ldots 5.38$$

If independent estimates of the growth parameters $L_{x}$ and $K$ are available, and if $L'$, $L_{max}$ and $\bar{L}$ can be estimated from length–frequency distributions, Equation 5.38 contains only one unknown, the instantaneous total mortality rate $Z$. However, Equation 5.38 does not generate an explicit solution for $Z$; hence, total mortality rate must be obtained by successive approximations, i.e. iteratively.

Estimating $L_{max}$

To estimate unobserved $L_{max}$ values such as required for the method of Ault and Ehrhardt (see above), the theory of extreme values (Gumbel 1954) may be applied (Fig. 5.6).

Fig. 5.6. Flow diagram of the procedure to estimate the theoretical $L_{max}$ using the extreme value theorem.

The theory assumes the following conditions:

- Lengths are continuous.

---

41 Adapted from Formacion et al. (1991).
• Samples from which the extreme lengths are drawn have a constant distribution with fixed parameters.

• The extreme lengths are taken from independent samples.

Fig. 5.7. Observed $L_{\text{max}}$ in individual samples of Upeneus moluccensis (Table 3.1), confidence interval and predicted $L_{\text{max}}$ values at the 95 percent level of confidence.

This method (see Fig. 5.7 for an example) is independent of any deterministic growth model (such as the VBGF). Rather it is based on the observation that the largest fish in any given sampling process varies randomly according to some probabilistic law.

**Natural mortality**

Natural mortality ($M$), as defined in the literature, is caused by all possible causes of death except fishing. Direct estimates of $M$ can therefore be obtained only from completely unfished stocks. Yet values of this parameter are needed, from exploited stocks, for most of the models currently used in fish population dynamics.

In exploited fish stocks, values of $M$ may be obtained from estimates of total mortality ($Z$) from which fishing mortality ($F$) is subtracted, or by a plot of $Z$ against fishing effort ($f$), $M$ then representing the intercept (at $f = 0$) (see Beverton and Holt 1956, Ricker 1975, Pauly 1984a). These two approaches obviously have their limitations, the first in the fact that most exploitable fish stocks are currently exploited, the second in that total mortality and contemporary fishing effort data are very often unavailable.

This has led to various empirical models for the estimation of $M$ from values of correlates, i.e. of parameters related to $M$.

**Pauly's M empirical equation**

It has been shown by various authors that the parameter $K$ of the VBGF is closely linked with longevity in fish (see for example Beverton and Holt 1959). This can be demonstrated on the basis of the rule that in nature the oldest fish of a stock grow to about 95 percent of their asymptotic length (Taylor 1958, Beverton 1963). This rule, which was derived from growth data used in conjunction with the special VBGF, does not strictly apply to large fish, such as tuna (see Pauly 1981), but seems to work for medium-sized and small fish growing according to the VBGF. Thus, if we have

$$t_{\text{max}} = \left\{ \ln\left[1 - \left(\frac{L_{\text{max}}}{L_o}\right)\right] - K \right\} - t_o \quad \ldots5.39$$

we can insert $L_{\text{max}} = 0.95 \cdot L_o$, which gives

$$t_{\text{max}} = \left(2.9957/K\right) + t_o \quad \ldots5.40$$

---

Adapted from Pauly (1984a).
Ignoring \( t_i \) and simplifying gives

\[
t_{\text{max}} \approx \frac{3}{K}
\]

...5.41

where \( t_{\text{max}} \) is the approximate longevity of the fish in question.

That natural mortality should, in fishes, be inversely correlated with longevity and hence be correlated with \( K \) thus seems obvious. Natural mortality should also inversely correlate with size since large fish should have, as a rule, fewer predators than small fish. Natural mortality can also be demonstrated to be correlated with mean environmental temperature (Pauly 1980). These interrelationships can be expressed for length growth data by the multiple regression

\[
\log_{10} M = -0.0066 - 0.279 \log_{10} L_{\infty} + \ldots
\]

... 0.6543 \log_{10} K + 0.463 \log_{10} T

...5.42

and for weight growth data by

\[
\log_{10} M = -0.2107 - 0.0824 \log_{10} W_{\infty} + \ldots
\]

... 0.6757 \log_{10} K + 0.4687 \log_{10} T

...5.43

where \( M \) is the natural mortality in a given stock, \( L_{\infty} \) (total length, in cm) and \( W_{\infty} \) (live weight, in g) being the asymptotic sizes of that stock; \( K \) refers to the curvature parameters of the VBGF and is expressed on an annual basis; the value of \( T \) is the annual mean habitat temperature (°C) of the water in which the stock in question lives.

Note that negative temperature values for polar fishes down to -2°C may be used as input in FISAT because an “effective physiological temperature” (Pauly 1980), which happens to be always positive, is computed internally for all values of \( T < 3.5^\circ \) and \( T > -2.0^\circ \).C.

In general, the estimates of \( M \) provided by Equations 5.42 and 5.43 are quite reasonable, especially because a very large number (175) of independent estimates of \( M \) have been used for their derivation. Also the fish considered covered an extremely wide range of sizes, taxa and habitats.

However, estimates of \( M \) obtained from these expressions may be biased upward in the case of strongly schooling fishes, such as the sardine-like fishes, and downward in the case of polar fishes. Correction factors and a further discussion of Equations 5.42 and 5.43 are given in Pauly (1980 and 1985), along with all data used in the derivation. Also, note that this model, which may also give reasonable results for shrimps and cephalopods (where \( TL \) must be used for \( L_{\infty} \), will not work for bivalves, for which \( M \approx K \) is a better approximation.

Rickhter and Efanov’s method

Beverton and Holt (1959) investigated the relationship between longevity (\( t_{\text{max}} \)) and the \( L_{\infty}/L_{m} \) ratio, where \( L_{m} \) is the mean length at first maturation, and this was continued by Holt (1962) and Rickhter and Efanov (1976), who showed a close relationship between \( M \) and the age at which most of the fish of a stock reach maturity (age at massive maturation or \( t_{\text{mass}} \)). This relationship is expressed by

\[
M = [(1.52/t_{\text{mass}}) \cdot 0.72] - 0.16
\]

...5.44

where \( M \) is expressed in years.

Note that this model was based on only few \( M \) and \( t_{\text{mass}} \) data pairs, covering a relatively small range of \( M \) values pertaining exclusively to boreal/temperate stocks; hence it should not be used for tropical stocks without prior comparison with previous models.

Adapted from Rickhter and Efanov (1976).
Estimating M from selection data

The original approach

The method of Munro (1984) is based around two premises: (i) that for any fishing gear, the number of fish caught (N) in each successive length class is a function of the probability of capture (P) of that length group by the fishing gear and the relative abundance (A) of that length classes in the population; and (ii) that the total mortality coefficient (Z) observable between successive length groups in an exploited stock is the sum of the natural mortality coefficient plus the product of the probability of capture, P, and the prevailing fishing mortality coefficient (F).

Thus, for the i-th length class

\[ Z_i = M + P_{i+\frac{1}{2}} \cdot F \]  ...5.45)

Here, a regression of estimates of \( Z_i \) against \( P_{i+\frac{1}{2}} \) (interpolated probability of capture at the upper limit of the length group i) should yield a regression of slope F and a y-axis intercept equal to M.

The requirements of this methodology can be summarized as:

- catch length composition data averaged over at least one, and preferably several years;
- estimates of parameters of the von Bertalanffy growth function \( L_\infty \) and \( K \);
- estimates of the probability of capture (\( P_i \)) at successive lengths (\( L_i \)), acquired independently of the estimates of the mean annual length composition.

The procedure is illustrated in Table 5.6 and Fig. 5.8 by a hypothetical example from a gill net fishery. The steps are as follows:

### Table 5.6. Hypothetical example (\( K = 0.5 \) year\(^{-1} \), \( L_\infty = 30 \) cm, \( M = 1 \) year\(^{-1} \), \( F = 2 \) year\(^{-1} \)) showing steps in obtaining estimates of total mortality rate (\( Z_i \)) from catches (\( C_i \)). The ages, \( t_i \), at successive lengths are the relative ages which are calculated by assigning a zero value to the parameter, \( t_{\infty} \), of the VBGF. See text for details.

<table>
<thead>
<tr>
<th>( L_i ) (cm)</th>
<th>( P_i )</th>
<th>( C_i )</th>
<th>( S_i )</th>
<th>( S_{i+1} )</th>
<th>( t_{\infty} ) (year)</th>
<th>( S_{i+1}/S_i )</th>
<th>( (S_{i+1}-S_i)/(t_{\infty} - t_i) )</th>
<th>( Z_i ) (year(^{-1} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.5</td>
<td>0.12</td>
<td>89</td>
<td>742</td>
<td>0.148</td>
<td>5.004</td>
<td>0.790</td>
<td>0.154</td>
<td>1.53</td>
</tr>
<tr>
<td>17.5</td>
<td>0.43</td>
<td>272</td>
<td>633</td>
<td>0.160</td>
<td>3.951</td>
<td>0.694</td>
<td>0.167</td>
<td>2.27</td>
</tr>
<tr>
<td>18.5</td>
<td>0.85</td>
<td>400</td>
<td>471</td>
<td>0.174</td>
<td>2.704</td>
<td>0.598</td>
<td>0.182</td>
<td>2.83</td>
</tr>
<tr>
<td>19.5</td>
<td>0.97</td>
<td>299</td>
<td>308</td>
<td>0.191</td>
<td>1.617</td>
<td>0.551</td>
<td>0.190</td>
<td>2.98</td>
</tr>
<tr>
<td>20.5</td>
<td>0.97</td>
<td>182</td>
<td>188</td>
<td>0.211</td>
<td>0.890</td>
<td>0.541</td>
<td>0.222</td>
<td>2.76</td>
</tr>
<tr>
<td>21.5</td>
<td>0.82</td>
<td>93</td>
<td>113</td>
<td>0.236</td>
<td>0.481</td>
<td>0.544</td>
<td>0.250</td>
<td>2.43</td>
</tr>
<tr>
<td>22.5</td>
<td>0.60</td>
<td>42</td>
<td>70</td>
<td>0.267</td>
<td>0.262</td>
<td>0.557</td>
<td>0.286</td>
<td>2.05</td>
</tr>
<tr>
<td>23.5</td>
<td>0.40</td>
<td>18</td>
<td>45</td>
<td>0.308</td>
<td>0.146</td>
<td>0.521</td>
<td>0.334</td>
<td>2.00</td>
</tr>
<tr>
<td>24.5</td>
<td>0.22</td>
<td>6</td>
<td>27</td>
<td>0.365</td>
<td>0.075</td>
<td>0.666</td>
<td>0.401</td>
<td>1.01</td>
</tr>
<tr>
<td>25.5</td>
<td>0.09</td>
<td>2</td>
<td>22</td>
<td>0.446</td>
<td>0.040</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Fig. 5.8. Regression of coefficients of mortality, \( Z_i \), between successive length groups against the probabilities of capture (\( P_i \)). (Data from Table 5.6.)
• Tabulate the average catch in numbers ($C_i$) and the probabilities of capture ($P_i$) of successive length groups.

• Calculate the apparent relative abundance ($S'_i$) of successive length groups within the selection curve as
  \[ S'_i = \frac{C_i}{P_i} \]

• Calculate the ages (in years) at the mid-point ($t_i$) and the time required ($\Delta t_i$) for a fish to grow through each length class.

• Divide $S'_i$ by $\Delta t_i$ to get the true relative abundance ($S_i$) of the fish in each length class.

• Calculate the annual coefficients of total mortality, $Z_i = \frac{\ln(S_{i+1}/S_i)}{(\Delta t_{i+1} - \Delta t_i)}$, between successive length classes.

• Plot $Z_i$ against interpolated probability of capture computed at the upper limit of the class $i$ ($P_{i+\frac{1}{2}}$) to obtain estimates of $F$ and also of $M$, using the equation
  \[ Z_i = (P_{i+\frac{1}{2}} \cdot F) + M. \]

**Improving Munro’s approach**

A variation of the above methodology assumes that the values of $C_i$, $S_i$, and $R_i$ are derived from a sampling gear (here: gear 1) with a retention range smaller than that of a commercial gear (here: gear 2), and that values of $Z_i$ are computed from the above values and plotted against the probabilities of retention ($P_i$) of gear 2. In such cases, estimates of $Z_i$ pertaining to lengths below the retention range of the commercial gear (where $P_i = 0$) are direct estimates of natural mortality ($M$); or put differently: the probabilities of capture for gear 1 are used to compute $Z_i$ values, which are plotted against the probabilities of capture of gear 2 to estimate $M$ as the y-intercept of the regression.

The method as originally presented by Munro has two disadvantages:

• The observed $P_i$ values are not those that are used for Equation 5.45, which relies on interpolated values of $P_i$ (see Fig. 5.9).

• The variability of the estimated $Z_i$ values is unnecessarily high.

Both these disadvantages can be remedied quite straightforwardly by using interpolated values of $Z_i$ (which thereby become less variable) and by plotting these against their corresponding values of $P_i$ (i.e. using the observed values of $P_i$; see Fig. 5.9B).

Table 5.7, illustrating the approach of Moreau (1988) and its difference from Munro’s original method, is based on data pertaining to brown trout from the Vèbre River, France. As might be seen from Fig. 5.10, the points resulting from the modification proposed here have a higher correlation than the original version of the method ($r = 0.944$ vs. $r = 0.887$).

---

45 Adapted from Moreau (1988).
Fig. 5.9. Schematic representation of differences in computations of Z and P values in Munro’s original method and Moreau’s modification (see also Table 5.6 and text).

Fig. 5.10. Plots of Z vs. P for estimation of M using Munro’s original method and the modification proposed by Moreau (1988). The higher estimate of M appears more realistic, especially when compared to an estimate derived using Equation 5.42 and estimates of $L_\infty$, K and T in Table 5.7.

We must note, however, that whatever the merits of the modified approach, it has crucial assumptions which must be met if the method is to work. Of these, the most important is probably that the size frequency sample that underlies the analysis (the $C_i$ values in Table 5.7) must indeed represent a steady-state situation (i.e. be the average of a multi-year series of seasonal or monthly samples).
Gear selectivity

Generally fishing gears, whether used by fishers or by fishery biologists are “selective”, i.e. they catch fish only within a certain range of sizes. Thus, if one wishes to know the true size structure of a fish population (for example to assess whether there has been a reduction of mean size over time), it is necessary to account for the effect of gear selection.

This can be achieved by estimating, for each size class of fish sampled, the probability of capture by the gear in question, then dividing, for each length class, the numbers actually caught by the gear’s specific probability of capture. Methods are presented below to estimate the probability of capture (= fraction retained) of different size groups of fish caught by fishing gears. The first of these methods pertains to trawl selection, the second is a simple method applied to gillnets (but also applicable to fishing hooks and some other gears with similar characteristic) and the last pertain to estimates of probabilities of capture extrapolated from length-converted catch curves.

Trawl mesh selection

The selectivity of trawl meshes is generally determined through trawl selection experiments. Such experiments consist of covering the cod end whose selectivity is to be assessed with a fine-mesh cover (see Fig. 5.11). After fishing, in each length group, a certain fraction of the total number of fish caught will be retained in the cod end, and this fraction (the probability of capture) will tend to increase with increasing fish length (Table 5.8).

From such data, the probability of capture can be obtained via a plot of the fractions retained against the corresponding

---

46 Adapted from Sparre and Venema (1992).

---

Table 5.8. Trawl selection data for slipmouth (*Leiognathus equulus*) obtained with 7.8 cm meshed nets in Mombasa Harbor, Kenya. * From Pauly (1984a), based on selection experiments conducted during the FAO/DANIDA Training Course on the Methodology of Fisheries Sciences (Biology), held in Mombasa, Kenya, 19 May–14 June 1980.

<table>
<thead>
<tr>
<th>Lower class limit (cm)</th>
<th>Fishes in cover (no.)</th>
<th>Fishes in cod end (no.)</th>
<th>Total fish caught</th>
<th>P = fishes in codend as fraction of total (= fraction retained)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 = Lmin</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>0.000</td>
</tr>
<tr>
<td>9</td>
<td>35</td>
<td>2</td>
<td>37</td>
<td>0.054</td>
</tr>
<tr>
<td>10</td>
<td>198</td>
<td>22</td>
<td>220</td>
<td>0.100</td>
</tr>
<tr>
<td>11</td>
<td>170</td>
<td>56</td>
<td>226</td>
<td>0.248</td>
</tr>
<tr>
<td>12</td>
<td>76</td>
<td>42</td>
<td>118</td>
<td>0.356</td>
</tr>
<tr>
<td>13</td>
<td>45</td>
<td>34</td>
<td>79</td>
<td>0.430</td>
</tr>
<tr>
<td>14</td>
<td>25</td>
<td>19</td>
<td>44</td>
<td>0.432</td>
</tr>
<tr>
<td>15</td>
<td>7</td>
<td>21</td>
<td>28</td>
<td>0.750</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>12</td>
<td>12</td>
<td>1.000</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>0.750</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td>1.000</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td>1.000</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>1.000</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1.000</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1.000</td>
</tr>
<tr>
<td>23 = Ln1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1.000</td>
</tr>
<tr>
<td>24 = Ln+1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3P = 10.120</td>
</tr>
</tbody>
</table>

From Pauly (1984a).
lengths. A smooth curve can then be drawn from which the probability of capture can be read for each length class (Fig. 5.12).

Several of the models discussed in this manual require estimates of mean size at first capture, that is the length at which 50 percent of the fish entering a trawl net are retained by the gear \((L_{50}, \text{ or } L_c)\).

The parameter \(L_c\) is particularly interesting in that, given symmetry of the selection curve, it is the length at which the numbers of smaller fish retained by the codend compensate for the number of larger fish not retained by the codend (see shaded areas in Fig. 5.12).

While \(L_c\) can be estimated graphically from plots such as Fig. 5.12, a more precise method is to order the catch data as in Table 5.8 and to estimate \(L_c\) from

\[
L_c = L_{n+1} - \Sigma P_i \quad \ldots5.46
\]

where \(L_n\) is the lower limit of the highest length class considered (when this equation is used the fish must be grouped in classes of width equal to unity, e.g. 1 cm), while \(\Sigma P_i\) is the sum of the fractions retained, as shown in Table 5.8.

Another method to estimate \(L_c\) is to fit the retention data with a logistic curve of the form

\[
P_i = \frac{1}{[1 + e^{r_m(l_c-L_i)}]} \quad \ldots5.47
\]

where \(P_i\) is the probability of capture at the midpoint of a length class \(i\) and \(r_m\) a constant whose value increases with the steepness of the selection curve; both Equations 5.46 and 5.47 assume the observed selection pattern to be symmetrical (or nearly so). Equation 5.47 may also be rewritten

\[
P_i = \frac{1}{[1 + e^{(S_1-S_2)\cdot L_i}]} \quad \ldots5.48
\]

where

\[
P_i = \left(\frac{\text{no. of fish of length } L_i \text{ in the codend}}{\text{no. of fish of length } L_i \text{ in the codend and in the cover}}\right) \quad \ldots5.49
\]

and \(L_i\) is the length interval midpoint, \(S_1\) and \(S_2\) being constant (Paloheimo and Cadima 1964, Kimura 1977 and Hoydal, Rørvik and Sparre 1982). Equation 5.48 can be re-expressed as

\[
\ln\left[\frac{1}{P_i} - 1\right] = S_1 - S_2 \cdot L_i \quad \ldots5.50
\]

which can be identified with a regression line, where \(S_1 = a\) and \(S_2 = b\).

There is a one-to-one correspondence between \(S_1\) and \(S_2\) and \(L_{25}, L_{50}\) and \(L_{75}\), the lengths at which respectively 25, 50 and 75 percent of the fish are retained in the codend. The
length range from $L_{25}$ to $L_{75}$, which is symmetrical around $L_{50}$, is called the selection range.

The formulae for calculating $L_{25}$, $L_{50}$ and $L_{75}$ are

$$L_{25} = \left[ \ln(3) - S_1 \right] / S_2 \quad \ldots 5.51)$$

$$L_{50} = S_1 / S_2 \quad \ldots 5.52)$$

$$L_{75} = \left[ \ln(3) + S_1 \right] / S_2 \quad \ldots 5.53)$$

$S_1$ and $S_2$ can be derived from $L_{75}$ and $L_{50}$ using:

$$S_1 = L_{50} \cdot \ln(3) / (L_{75} - L_{50}) \quad \ldots 5.54)$$

$$S_2 = \ln(3) / (L_{75} - L_{50}) = S_1 / L_{50} \quad \ldots 5.55)$$

Note that Equation 5.50 is not defined for $P_i = 0$ or $P_i = 1$.

The above considerations assume more or less symmetrical selection ogives. There are cases where a trawl selection ogive is not symmetrical; in such cases, an approach to estimating $L_{25}$, $L_{50}$ and $L_{75}$ and obtaining a smooth selection curve from observed probabilities of capture is through a running average, and subsequent graphical estimation of the selection range and of $L_{50}$.

In trawls, $L_c$ can be considered proportional to the mesh size of the codend; the proportionally constant is called the selection factor (SF). When known, it can be used to estimate $L_c$ from the relationship

$$L_c = SF \cdot \text{mesh size} \quad \ldots 5.56)$$

It has been demonstrated by several authors that the selection of fishes is generally related to their overall shape: slender fishes have high selection factors while bulky ones have low selection factors. This property has been used by Pauly (1984a) to derive a nomogram (Fig. 5.13) based on a large number of published results of selection experiments, which can be used to estimate approximate values of selection factors of fishes, given their "girth factor" (maximum girth/total length) or their "depth ratio" (standard length/maximum body depth).

![Fig. 5.13. Nomogram for the estimation of selection factors of fishes from their body proportion (from Pauly 1984a).](image)

**Gill net selection**

Whereas trawl selection is generally a one-sided affair (with only smaller fish having a reduced probability of capture), gill nets tend to select against both small and large fish. The former simply go through the mesh without getting caught, while the latter are too big to insert themselves into a mesh.

Gill nets are usually long rectangular nets where the upper edge (the head rope) has floats while the foot rope has sinkers. Often gill nets (drifting and set nets) are deployed in the form

47 Adapted from Sparre and Venema (1992).
of "gangs" of nets with different mesh sizes, to avoid the effects of selection, but this is a feature not discussed here.

Gill nets are *passive gear*, i.e. the fish have to swim into the net to get caught. This implies that fish which move fast have a larger probability of encounter with the gear than slow moving fish.

Rudstam et al. (1984) included swimming speed into a model for gill net selection. They considered the selection as the product of two probabilities, probability of encounter and probability of being caught given an encounter.

The picture becomes somewhat more complicated when various ways by which the fish can get stuck in a gill net are also considered. Baranov (1914) recognized three of these (see Fig. 5.14):

- *Gilled*, where the mesh is around the fish just behind the gill cover;
- *Wedged*, where the mesh is around the body as far as the dorsal fin; and
- *Entangled*, where the fish is held in the net by its teeth, jaws, fins or other projections.

Fig. 5.14. Illustration of the different ways fish can be caught by gill nets (from Karlsen and Bjarnasson 1986).

We shall deal here only with the probability of being caught given an encounter with a gill net, and with fish caught by being "gilled".

**Baranov/Holt model**

One model which has been widely used for estimating gill net selectivity parameters is the Baranov/Holt model, originally developed by Baranov (1914), subsequently improved by Holt (1963), and presented in various stock assessment texts (Gulland 1983, Pauly 1984a, Sparre and Venema 1992). This model uses two sets of catch data obtained by two gill nets with different mesh sizes \( m_A \) and \( m_B \) (note that henceforth \( m_A < m_B \)), and assumes that:

- Both sets of catch-at-length data were obtained with a known level of effort (i.e. data must be standardized by effort).

---

\[\text{Adapted from Pauly (1984a).}\]
• The selection curves for both mesh sizes \( m_A \) and \( m_B \) can be represented by normal (i.e. bell-shaped) distributions with the same standard deviations (s.d.s).

• The means of both normal distributions, i.e. their values of \( L_{opt} \), \( L_A \) and \( L_B \), are related such that \( m_A/m_B = L_A/L_B \).

• The catches at length obtained with the gill nets overlap over several length classes.

Given that these assumptions are met and with given catches obtained by the smaller mesh \( m_A \) and the larger mesh \( m_B \), the optimum length corresponding to \( A \) \( (L_A) \) and the optimum length corresponding to \( B \) \( (L_B) \) can be estimated from the catch by length class of each mesh \( (i.e. \ C_A \ and \ C_B) \) through a linear regression of the form

\[
\ln(C_B/C_A) = a + b \cdot L_i \quad \ldots 5.57
\]

where the ratio \( C_A/C_B \) is called the "catch ratio".

The intercept and slope of this regression can then be used to estimate the optimum lengths from

\[
L_A = \frac{(-2a \cdot m_A)}{b \cdot (m_A + m_B)} \quad \ldots 5.58
\]

and

\[
L_B = \frac{(-2a \cdot m_B)}{b \cdot (m_A + m_B)} \quad \ldots 5.59
\]

while the standard deviation (s.d.) of both selection curves is estimated from

\[
s.d. = \frac{\sqrt{2a(m_A - m_B)}}{b(m_A + m_B)} \quad \ldots 5.60
\]

and the selection factor from

\[
SF = -\frac{2a}{b \cdot (m_A + m_B)} \quad \ldots 5.61
\]

Note also that

\[
L_A = SF \cdot m_A \text{ and } L_B = SF \cdot m_B \quad \ldots 5.62
\]

Once \( L_A \), \( L_B \) and s.d. have been estimated, the probability of capture \( (P) \) at a given length \( (L) \) is given for mesh \( m_A \) by

\[
P_{A,L} = e^{\left[\frac{(L-L_A)^2}{2(s.d.)^2}\right]} \quad \ldots 5.63
\]

and for mesh \( m_B \) by

\[
P_{B,L} = e^{\left[\frac{(L-L_B)^2}{2(s.d.)^2}\right]} \quad \ldots 5.64
\]

Although the method gives reasonable results in the case of the example provided here (Table 5.9, Fig. 5.15 and Fig. 5.16), various authors have shown that gill net selection curves frequently have shapes other than normal. This applies especially to large, spiny fishes, which, in addition to being gilled, often entangle themselves, which results in asymmetrical selection curves. In such cases, it may be necessary to use more elaborate methods to estimate the selectivity of the net(s) under investigation, e.g. those of Gulland and Harding (1961), or Hamley (1975).
Table 5.9. Estimation of gill net selection curves for *Tilapia esculenta*, Lake Victoria (data from Garrod 1961).

<table>
<thead>
<tr>
<th>Length interval midpoint (L) (x-axis)</th>
<th>Numbers caught (mesh 8.1 cm)</th>
<th>Numbers caught (mesh 9.1 cm)</th>
<th>ln(CB/CA) (catch ratios) (y-axis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.5</td>
<td>7</td>
<td>0</td>
<td>not used</td>
</tr>
<tr>
<td>19.5</td>
<td>90</td>
<td>1</td>
<td>used</td>
</tr>
<tr>
<td>20.5</td>
<td>199</td>
<td>9</td>
<td>-3.096</td>
</tr>
<tr>
<td>21.5</td>
<td>182</td>
<td>53</td>
<td>-1.234</td>
</tr>
<tr>
<td>22.5</td>
<td>119</td>
<td>290</td>
<td>0.891</td>
</tr>
<tr>
<td>23.5</td>
<td>29</td>
<td>357</td>
<td>2.510</td>
</tr>
<tr>
<td>24.5</td>
<td>17</td>
<td>225</td>
<td>not used</td>
</tr>
<tr>
<td>25.5</td>
<td>3</td>
<td>82</td>
<td>used</td>
</tr>
<tr>
<td>26.5</td>
<td>0</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>27.5</td>
<td>0</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

intercept = -41.907, slope = 1.894

\[
L_A = \frac{-2 \cdot (-41.907 \cdot 8.1)}{1.894 \cdot (8.1 + 9.1)} = 20.8 \text{ cm}
\]

\[
L_B = \frac{(20.8 \cdot 9.1)/8.1}{8.1} = 23.4 \text{ cm}
\]

\[
s.d. = \left[\frac{-2(-41.907)(9.1 - 8.1)}{1.894^2(8.1 + 9.1)}\right]^{1/2} = 1.17
\]

When the selection curves for a given fish species are only slightly asymmetrical and drawn to the right, it is still possible to apply the Baranov/Holt method outlined above using the logarithm of the lengths instead of the lengths in all computations. Indeed, FISAT allows “flipping” between the linear and non-linear versions of the Baranov/Holt model.

Multiple mesh sizes

In situations where more than two mesh sizes are used, an overall selection factor and a common standard deviation may be estimated from the results of an analysis of each pair of successive mesh sizes, if the same assumptions apply as in the case of two mesh sizes.

---

49 Adapted from Sparre and Venema (1992).
Assume that there are \( n \) mesh sizes. There will then be \( n - 1 \) estimates of the intercept (\( a \)) and the slope (\( b \))
\[
[a_1, b_1], [a_2, b_2], \ldots, [a_{n-1}, b_{n-1}]
\]
corresponding to the mesh sizes:
\[
[m_1, m_2], [m_2, m_3], \ldots, [m_{n-1}, m_n]
\]
Thus, one can estimate a common selection factor by performing a regression analysis, the line being forced through the origin \( y_i = b \cdot x_i \), with \( y_i = -2 \cdot a_i/b_i \) as dependent variable and \( x_i = m_i + m_{i+1} \) as independent variable. The slope (\( b \)) then becomes the selection factor (SF), according to a rearranged version of Equation 5.61:
\[
-2 \cdot a_i/b_i = SF(m_i + m_{i+1}), \quad i = 1, 2, \ldots, n-1 \quad \ldots 5.65
\]
or simply,
\[
b = \frac{\sum x_i y_i}{\sum x_i^2} \quad \ldots 5.66
\]
Therefore,
\[
SF = \frac{-2 \sum_{i=1}^{n-1} (m_i + m_{i+1}) \cdot (a_i / b_i)}{\sum_{i=1}^{n-1} (m_i + m_{i+1})^2} \quad \ldots 5.67
\]
Then, standard deviations, as the mean values of the individual estimates for each consecutive pair of mesh sizes, can be estimated from
\[
s.d. = \sqrt{\left(1 / n - 1\right) \sum_{i=1}^{n-1} \left[-2 a_i (m_{i+1} - m_i) / b_i^2 (m_i + m_{i+1})\right]} \quad \ldots 5.68
\]
The optimum length for mesh size \( i \) is finally obtained by:
\[
L_i = SF \cdot m_i \quad \ldots 5.69
\]
There will be cases where this approach cannot be used, because the underlying relationships are not linear (see for example Cortez-Zaragoza et al. 1983) and it is thus important to check that Equation 5.65 leads to a plot that is linear, or nearly so.

**Estimation of probabilities of capture from length-converted catch curves**

Probabilities of capture for trawl-type selection can also be estimated from the left, ascending arm of length-converted catch curves. Essentially, the method (see Fig. 5.17) 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 "expected" numbers to those that are actually caught being used to estimate probabilities of capture. This approach, which is related to work done by Jensen (1982) and Hoydal et al. (1982), has been found (Anon. 1982) to provide reasonable estimates of mean size at first capture (\( L_c \)). Thus, it has the potential to make some mesh selection experiments superfluous, especially when length–frequency data are available that include very small fish, i.e. covering the month(s) in which recruitment occurs.

The probabilities that are estimated here may be used, in the absence of appropriate selection/recruitment data, to correct the length frequencies for the effects of selection (if any) and/or incomplete recruitment (if applicable).

\[50\] Adapted from Pauly (1984a).
Fig. 5.17. Schematic representation of method to derive probabilities of capture from the left, ascending arm of a length-converted catch curve for *Upeneus moluccensis*, Ragay Gulf, Philippines (L<sub>50</sub> = 19.5 cm and K = 0.9 year<sup>-1</sup>). P<sub>1</sub> identifies the first point from which 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<sub>0</sub> represents the first point (when going from right to left) where the probability of capture is zero (0); the mortality corresponding to this point is M, by definition. With M at P<sub>0</sub> and Z at P<sub>1</sub>, intermediate mortalities can be interpolated, and population sizes reconstructed by backward extrapolation, starting from point P<sub>1</sub>. 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. (see Table 5.10).

Table 5.10. Illustrating how approximate probabilities of capture can be estimated from a length-converted catch for *Upeneus moluccensis*, Ragay Gulf, Philippines.

<table>
<thead>
<tr>
<th>Freq.</th>
<th>ln(N/Δt)</th>
<th>Rel. age (year&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>F (year&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>M (year&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>Z (year&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>No available</th>
<th>Prob. ( (C_{ai}/C_{at}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.50</td>
<td>0.38</td>
<td>0.37</td>
<td>1.74</td>
<td>2.11</td>
<td>2,1032</td>
<td>12.458</td>
</tr>
<tr>
<td>4</td>
<td>3.82</td>
<td>0.47</td>
<td>0.74</td>
<td>1.74</td>
<td>2.48</td>
<td>1,7439</td>
<td>12.203</td>
</tr>
<tr>
<td>17</td>
<td>5.19</td>
<td>0.56</td>
<td>1.11</td>
<td>1.74</td>
<td>2.85</td>
<td>1,3818</td>
<td>11.898</td>
</tr>
<tr>
<td>37</td>
<td>5.89</td>
<td>0.65</td>
<td>1.48</td>
<td>1.74</td>
<td>3.22</td>
<td>1,0391</td>
<td>11.535</td>
</tr>
<tr>
<td>254</td>
<td>7.73</td>
<td>0.76</td>
<td>1.85</td>
<td>1.74</td>
<td>3.59</td>
<td>7353</td>
<td>11.104</td>
</tr>
<tr>
<td>388</td>
<td>8.06</td>
<td>0.87</td>
<td>2.22</td>
<td>1.74</td>
<td>3.96</td>
<td>4844</td>
<td>10.594</td>
</tr>
<tr>
<td>1,109</td>
<td>9.02</td>
<td>1.01</td>
<td>2.59</td>
<td>1.74</td>
<td>4.33</td>
<td>2930</td>
<td>9.989</td>
</tr>
<tr>
<td>1,090</td>
<td>8.88</td>
<td>1.15</td>
<td>2.96</td>
<td>1.74</td>
<td>4.70</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>566</td>
<td>7.95</td>
<td>1.49</td>
<td>2.96</td>
<td>1.74</td>
<td>4.70</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>186</td>
<td>6.67</td>
<td>1.71</td>
<td>2.96</td>
<td>1.74</td>
<td>4.70</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>99</td>
<td>5.83</td>
<td>1.97</td>
<td>2.96</td>
<td>1.74</td>
<td>4.70</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>3.45</td>
<td>2.30</td>
<td>2.96</td>
<td>1.74</td>
<td>4.70</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>3.37</td>
<td>2.75</td>
<td>2.96</td>
<td>1.74</td>
<td>4.70</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

To illustrate the approach, consider the data presented in Table 5.10 (columns 1 to 3). If we assume that Z = 4.7 year<sup>-1</sup> and M = 1.74 year<sup>-1</sup> (from Equation 5.42), F can be estimated to be 2.96 year<sup>-1</sup>. However, this estimate of F would refer to lengths from P<sub>1</sub> onward only (see Fig. 5.17). If we assume that F decreases gradually with length, with F at P<sub>0</sub> as zero (0), we can interpolate intermediate values of R by subtracting a value X defined by

\[
X = F/\text{(no. of classes below P<sub>1</sub> + 1)} \quad \text{...5.70)
\]

Thus, in our example, if P<sub>1</sub> is catch curve point 8,

\[
X = 2.96/8 = 0.37
\]

and extrapolating backward (column 4),

\[
F_{i-1} = F_i - X \quad \text{...5.71)
\]

Similarly, the numbers of fish that ought to have been caught \( (C_{ai}) \) is given by

\[
C_{ai-1} = C_i \cdot e^{Zt} \quad \text{...5.72)
\]

where Z is interpolated between two classes: \( Z = (Z_i + Z_{i+1})/2 \), with \( Z_i = M + F_i \) (from Equation 5.71).

The probability of capture can thus be computed as the ratio of the numbers observed over the numbers available \( (N_i) \). From Equation 5.72,

\[
P_i = \ln(C_i/\Delta t) - \ln (C_{ai}/\Delta t) \quad \text{...5.73)}
\]
The accuracy of the method depends on the following assumptions being met (Pauly 1984a):

- The gear in question is a trawl or has a selection curve similar to that of a trawl (where it is only the smaller fish that are selected against).
- The smallest fish caught ($L_{\text{min}}$) are fully recruited.
- The value of $M$ used for the fish between $P_0$ and $P_1$ and the mortalities generated by interpolation between $M$ and the $Z$ value for the fully selected fish are accurate.

The first of these assumptions can be easily verified. The second, when violated, implies that the computed probabilities will not strictly refer to a selection curve, but to a resultant curve, i.e. to the product of a selection and a recruitment curve (Gulland 1983). Whether the second assumption is met or not will thus affect the interpretation of the results, but not their computation.

The importance of the third assumption can be assessed using a simple sensitivity analysis. As might be seen from Fig. 5.18, $L_c$ shows little sensitivity to changes of $M$. Estimated probabilities of capture are, on the other hand, sensitive to the value of $M$ used, especially at smaller lengths.

Note that this method can be used only in conjunction with catch curves that do not take account of seasonal growth oscillations. Also note that FiSAT treats probability values estimated through this method as if they had been estimated externally, i.e. they can be used to estimate $L_{50}$ and a selection range using either a logistic curve or running averages (see above section on selection).

---

**Recruitment patterns**

Since it is the pulsed nature of (seasonal) recruitment into a population that determines the structure of a set of length–frequency data, the converse also applies that one should be able to recover some information on the seasonality of recruitment from length frequencies (Pauly 1982). This inverse approach is implemented in FiSAT in the form of “recruitment patterns”. These allow identifying the number of recruitment pulses per year and evaluating the relative importance of these pulses when compared to each other. However, recruitment patterns are based on two assumptions usually not met in reality:

- All fish in a given data set grow as described by a single set of growth parameters.
- One month out of twelve always has zero recruitment.

---

51 Adapted from Gayanilo et al. (1989).
The first of these two assumptions is common to other models described earlier; its validity and overall impacts are discussed in Hampton and Majkowski (1987), Rosenberg and Beddington (1987), and Isaac (1990). As far as recruitment patterns are concerned, this assumption has limited impact on the results because the computation of recruitment patterns assigns very strong weight to smaller fish, whose deviations from the growth curve typical of the whole population are much smaller than in larger, older fish (see Fig. 5.19).

Recruitment patterns are obtained by backward projection, onto the length axis, of a set of length–frequency data (Fig. 5.20). The steps involved are:

1) projection onto the time axis of the frequencies after they have been divided by Δt, the time needed to growth through the length class. (Alternatively, FISAT allows, following Moreau and Cuende (1991) the use of restructured frequencies (as used for ELEFAN I); this leads to recruitment patterns with peaks much narrower than when untransformed L/F data are used, thus partly compensating for the "spread" illustrated in Fig. 5.19);

2) summation for each month (and irrespective of year) of the adjusted frequencies projected onto each month;

3) subtraction, from each monthly sum, of the lowest monthly sum to obtain a zero value where apparent recruitment is lowest;

The second of the above two assumptions will probably never be strictly met, because (i) there may be more than one month per year in which no recruits enter a stock (especially in temperate fishes and invertebrates) or because (ii) some recruits may be entering a stock every month (especially in tropical fishes and invertebrates).
(4) output of monthly relative recruitment as a percentage of annual recruitment.

Concerning points (3) and (4), it must be noted that the monthly values pertain to months of the year only when a precise estimate of $t_o$ is available.

Note that the time lag between peak spawning and the peak(s) of a recruitment pattern can be used to estimate $t_o$; this, however, is not explored here.

Fig. 5.21. Above: seasonal pattern of spawning condition in gonads for six-banded grouper ($Epinephelus seafasciatus$) 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. 5.21 shows the match between a recruitment pattern 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, Moreau and Cuende 1991, Moreau et al. 1995) suggest that recruitment patterns obtained through the process outlined above do contain useful information, from which legitimate inferences on the seasonal dynamics of fish and invertebrate recruitment can be drawn.
Chapter 6. Virtual Population Analysis

What you will learn from this chapter

In this chapter, you will be introduced to the models leading to virtual population analysis (VPA), used to reconstruct fish population structure by age or length.

Introduction

Beverton and Holt (1957) showed that the catch \( (C_i) \) from a population during a unit time period, \( i \), is equal to the product of the population size at the beginning of the time period \( (N_i) \) times the fraction of the deaths caused by fishing, times the fraction of total deaths, which can be written

\[
C_i = (F_i/Z_i) \cdot (e^{Z_i} - 1)
\]

This is the equation for the virtual population analysis (VPA) of Gulland (1965); it can be derived from Equation 5.5 by substituting for \( N_i \) by \( N_{i+1} \cdot e^{Z_i} \).

Given values of \( C_i \) and an estimate of \( M \), Equation 6.1 can be used to estimate (retroactively) the size of past cohorts (i.e. of groups of fish born at the same time and exposed to the same mortalities throughout their lives), if an estimate of \( N_{i+1} \) is available from which to start the computation. Estimates of \( N_{i+1} \) (expressing the last population size a cohort had before it became extinct) are called “terminal populations” \( (N_t) \).

Values of \( N_t \) can be obtained from

\[
N_t = Z_t \cdot C_t/F_t
\]

where \( C_t \) is the terminal catch (i.e. the last catch taken from a cohort before it went extinct) and \( F_t \) is the terminal fishing mortality, i.e. the fishing pressure that generated \( C_t \) (Mesnil 1980).

In principle, VPA is thus a method that uses a quantity that we can be straightforwardly estimate—the catch—to draw inferences about the fishing mortality or a quantity that cannot be readily estimated—the population that generated that catch.

The feature of VPA that is most important for practical use is that, given a high fishing pressure, estimates of population size obtained by repeated application of Equation 6.1 tend to converge rapidly toward their true value, and hence usually provide, given a reasonable estimate of \( M \), reliable estimates of recruitment (Pope 1972).

Equation 6.1 does not have a direct solution, and to deal with the "cohort analysis" of Pope (1972), reviewed in Jones (1984) and Pauly (1984a), is often used. Cohort analysis is based on an approximation, i.e. on the assumption that all fish are caught on a single day, exactly in the middle of the time period considered (conventionally one year—see Fig. 6.1).

Consequently, in the first half year the fish suffer only natural mortality, so the number of survivors after that period becomes

\[
N_{i+1} = N_i \cdot e^{-M/2}
\]

\[52\] Adapted from Pauly (1984a).
Then, all the catch is taken and the number of survivors becomes

\[ N_{i+\frac{1}{2}} = N_i \cdot e^{-M/2} - C_i \] ...6.4)

These survivors only suffer natural mortality in the second half year and thus their number at the end of the year is

\[ N_{i+1} = S_{i+\frac{1}{2}} \cdot e^{-M/2} \] ...6.5)

Considering all three steps and going backward in time, i.e. from \( N_{i+1} \) to \( N_i \), leads, with some rearrangements, to

\[ N_i = \left( N_{i+1} \cdot e^{-M} + C_i \right) \cdot e^{M/2} \] ...6.6)

Note that fishing mortality, for which Equation 6.1 cannot be solved directly, does not occur here as an explicit parameter.

Thus, by repeatedly invoking Equation 6.6, we work backwards in time, estimating a new population size (\( N \)) at each step. Fishing mortality estimates can be then obtained from the successive \( N \) values, using:

\[ F_{i+1} = \ln\left(\frac{N_i}{N_{i+1}}\right) - M \] ...6.7)

Length-structured VPA

Generalizing Equation 6.1 for any time interval (\( \Delta t \)) gives

\[ \frac{N_{i+\Delta t}}{C_i} = \frac{Z_i \cdot e^{-Z_i \cdot \Delta t}}{F_i \left(1 - e^{-Z_i \cdot \Delta t}\right)} \] ...6.8)

which may also be written as

\[ C_i = N_{i+\Delta t} \left( F_i / Z_i \right) \left( e^{Z_i \cdot \Delta t} - 1 \right) \] ...6.9)

with all other parameters defined as in Equation 6.1; these equations allow for structuring catch data in terms of length, rather than time intervals.

Converting length to age requires the use of a mathematical expression of fish growth, here the VBGF. Thus, any age \( t_i \) pertaining to a length \( L_i \) can be obtained from

\[ t_i = \left(1/K\right) \cdot \ln \left[1 - \left( L_i / L_o \right)^\frac{1}{\alpha}\right] + t_o \] ...6.10)

and similarly for age \( t_{i+1} \), pertaining to \( L_{i+1} \). From the length–age relationships for \( L_i \) and \( L_{i+1} \), \( \Delta t \) is obtained as the difference between \( t_{i+1} \) and \( t_i \), or after some rearrangement

---

\(^{53}\) Adapted from Pauly (1984a).
\[ \Delta t = \left( \frac{1}{K} \right) \cdot \ln \left( \frac{L^\alpha - L_i}{L^\alpha - L_{i+1}} \right) \]  

...6.11)

which can be substituted for \( \Delta t \) in Equation 6.9. (It will be noted that the estimation of \( \Delta t \) does not require \( t_0 \).)

Thus, given catch-at-length data from a stock with stable age distribution, Equation 6.9 can be used in a fashion similar to Equation 6.1 to estimate, starting from a (guessed) terminal fishing mortality (affecting the largest length group), the number of fish in the smaller size classes and the fishing mortalities affecting them.

When Equation 6.9 is used in conjunction with values of \( \Delta t \) that are not constant (i.e. when the \( \Delta t \) values are computed from length-converted ages), the results obtained will not apply to a specific cohort of fish, but rather pertain (for a given value of \( M \)) to the population sizes (per length class) that must have existed, on the average, for the observed catch to have been produced by the estimated values of \( F \). The method is thus analogous to Jones' length cohort analysis (Jones 1974, 1979, 1981) which can be expressed by

\[ N_i \approx \left( N_{i+1} \cdot X_i + C_i \right) \cdot X_i \]  

...6.12)

where

\[ X_i = \left( \frac{L^\alpha - L_i}{L^\alpha - L_{i+1}} \right)^{M/K} \]  

...6.13)

where \( C_i \) is the number of fish caught in a given time period with stable age distribution and with length between \( L_i \) and \( L_{i+1} \), and where \( N_i \) and \( N_{i+1} \) represent the population size (in number) with length \( L_i \) and \( L_{i+1} \) respectively.

Jones' length cohort analysis is particularly helpful in that it requires a knowledge of only two parameters, \( L_\alpha \) and the ratio \( M/K \); the latter, as shown by Beverton and Holt (1959) tends to vary less between different groups of fish than either \( K \) or \( M \) alone. However, a problem with Jones' method is that it is derived from Pope's approximate "cohort analysis" (see above) through generalizing for any time interval:

\[ N_i \approx N_{i+1} \cdot e^{M \Delta t} + C_i \cdot e^{M \frac{\Delta t}{2}} \]  

...6.14)

Since Equation 6.9, which gives precise results, and the approximation in Equation 6.12, can both be used to obtain estimates of population size and fishing mortality from the same set of catch-at-length data, Equation 6.9 can be used to assess the closeness of the approximation involved in Equation 6.12. This is done in the example in Table 6.1.

<table>
<thead>
<tr>
<th>Length groups (cm)</th>
<th>Annual catch ('000)</th>
<th>Population ('000) A</th>
<th>Population ('000) B</th>
<th>Population ('000) C</th>
<th>Fishing mortality (year(^{-1})) A</th>
<th>Fishing mortality (year(^{-1})) B</th>
<th>Fishing mortality (year(^{-1})) C</th>
<th>(% diff.) A</th>
<th>(% diff.) B</th>
<th>(% diff.) C</th>
</tr>
</thead>
<tbody>
<tr>
<td>6–12</td>
<td>1,823</td>
<td>98,919</td>
<td>98,238</td>
<td>-0.7</td>
<td>0.040</td>
<td>0.040</td>
<td>0.040</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>12–18</td>
<td>14,463</td>
<td>64,393</td>
<td>63,801</td>
<td>-0.7</td>
<td>0.386</td>
<td>0.392</td>
<td>1.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>18–24</td>
<td>25,227</td>
<td>59,476</td>
<td>59,010</td>
<td>-0.8</td>
<td>1.066</td>
<td>1.111</td>
<td>4.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>24–30</td>
<td>8,134</td>
<td>27,632</td>
<td>27,428</td>
<td>-0.8</td>
<td>0.647</td>
<td>0.661</td>
<td>2.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>30–36</td>
<td>3,889</td>
<td>15,968</td>
<td>15,849</td>
<td>-0.8</td>
<td>0.491</td>
<td>0.500</td>
<td>1.8</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>36–42</td>
<td>2,959</td>
<td>9,961</td>
<td>9,872</td>
<td>-0.8</td>
<td>0.647</td>
<td>0.666</td>
<td>3.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>42–48</td>
<td>1,781</td>
<td>5,501</td>
<td>5,455</td>
<td>-0.8</td>
<td>0.647</td>
<td>0.666</td>
<td>3.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>48–54</td>
<td>653</td>
<td>2,619</td>
<td>2,797</td>
<td>-0.8</td>
<td>0.385</td>
<td>0.392</td>
<td>1.8</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>54–60</td>
<td>322</td>
<td>1,691</td>
<td>1,678</td>
<td>-0.8</td>
<td>0.286</td>
<td>0.293</td>
<td>2.7</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>60–66</td>
<td>228</td>
<td>1,057</td>
<td>1,048</td>
<td>-0.9</td>
<td>0.307</td>
<td>0.313</td>
<td>1.6</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>66–72</td>
<td>181</td>
<td>621</td>
<td>616</td>
<td>-0.8</td>
<td>0.401</td>
<td>0.412</td>
<td>2.7</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>72–78</td>
<td>96</td>
<td>314</td>
<td>312</td>
<td>-0.6</td>
<td>0.389</td>
<td>0.399</td>
<td>2.6</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>78–84</td>
<td>16</td>
<td>149</td>
<td>148</td>
<td>-0.7</td>
<td>0.110</td>
<td>0.111</td>
<td>0.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>84–\infty</td>
<td>45(C_i)</td>
<td>92(N_i)</td>
<td>92(N_i)</td>
<td>-</td>
<td>0.280(F_i)</td>
<td>0.280(F_i)</td>
<td>-</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

A = Jones' length cohort analysis.
B = Length-structured VPA.
C = (B/A -1) \cdot 100 = C (% difference).

As might be seen in Table 6.1, the combination of parameter values used generates a mean difference between the
results obtained with Jones’ method and those obtained using Equation 6.9 of only 0.7 percent for the population estimates and 2.2 percent for the fishing mortality estimates.

However, regrouping the catch data in Table 6.1 into larger and larger length class intervals produces increasing differences between the fishing mortality and population estimates obtained by the two methods (Table 6.2), suggesting that Jones’ length cohort analysis may indeed be quite sensitive to coarse groupings of the catch data (see Jones and van Zalinge 1981).

Table 6.2. Comparison of results using (A) Jones’ length cohort analysis and (B) length-structured VPA (24 cm classes)(see also Table 6.1).

<table>
<thead>
<tr>
<th>Length groups (cm)</th>
<th>Annual catch ('000)</th>
<th>Population ('000)</th>
<th>Fishing mortality (year⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C (%) diff.</td>
</tr>
<tr>
<td>12–36</td>
<td>51,713</td>
<td>93,010</td>
<td>84,379</td>
</tr>
<tr>
<td>36–60</td>
<td>5,805</td>
<td>11,592</td>
<td>10,265</td>
</tr>
<tr>
<td>60–84</td>
<td>521</td>
<td>1,236</td>
<td>1,087</td>
</tr>
<tr>
<td>84–inf</td>
<td>45 (C₀)</td>
<td>92 (N₀)</td>
<td>92 (N₀)</td>
</tr>
</tbody>
</table>

Varying the value of natural mortality used for the analysis produces, on the other hand, virtually no additional differences between the results of the two methods, i.e. the difference remained close to 2 percent for M = 0.1 year⁻¹ to M = 1.0 year⁻¹.

If the coefficients (a and b) of the length–weight relationship are known, the steady-state biomass can also be computed from the relationship

\[ B_i = \bar{w}_i \cdot C_i \cdot \Delta t \]  \hspace{1cm} \text{...6.15) \}

where \( \bar{w}_i \) is the mean weight of the fish in length class i and \( \Delta t \) is the time needed for the fish to grow through length class i. Here, \( \bar{w}_i \) is obtained, based on Beyer (1987), from

\[ \bar{w}_i = \left( \frac{1}{L_{i+1} - L_i} \right) \left( \frac{a}{b+1} \right) \left( L_{i+1}^{b+1} - L_i^{b+1} \right) \]  \hspace{1cm} \text{...6.16) \}

Equation 6.15 may also be written as

\[ B_i = \bar{w}_i \cdot \left( N_i - N_{i-1} \right) / Z_i \]  \hspace{1cm} \text{...6.17) \}

where \( Z_i \) is the total mortality (year⁻¹) as estimated given computed \( F_i \) and a fixed value of M, i.e. \( Z_i = M + F_i \).

Similarly, the yield \( (Y_i) \) (catch in weight) can also be computed using the relationship

\[ Y_i = \bar{w}_i \cdot C_i \]  \hspace{1cm} \text{...6.18) \}

where \( C_i \) is the catch (in numbers) for length class i.

The main drawback of the length-structured VPA form presented as Equation 6.9 and as implemented in \texttt{FiSAT} (see Fig. 6.2) (and of length–cohort analysis; see Equation 6.12), is the necessary assumption of a stable age distribution, which is not required in the case of age-structured VPA. However, a number of methods resting on the same assumption of a stable age distribution have become widely accepted and used for stock assessment, such as the estimation of total mortality from catch curves or from the mean length of fish in catch samples. As in the case of the procedures recommended for use with these other methods, a stable age distribution can be simulated in the case of length-structured VPA or length–cohort analysis by averaging catch data for a length of time during which recruitment and fishing mortality can be assumed to have been constant, or varying randomly without trend.
Jones' length–cohort analysis has the following advantages over its VPA equivalent (Equation 6.8):

- It does not require separate estimates of $K$ and $M$, but only of the ratio $M/K$.
- It provides direct solutions, i.e. the solution does not need to be obtained iteratively, as in the case of solutions to Equation 6.9.

However, to obtain population sizes and fishing mortalities based on length-structured VPA (or cohort analysis methods), it is necessary to have good catch-at-length data.

Converting catch in weight to catch-at-length data is rather straightforward, given length–frequency data representative of the catch and the parameters of the length–weight relationship in the stock in question. Since the mortalities are estimated in terms of years, the catch data should also be raised to represent the mean annual catch.

Unfortunately, the catch and landing data-collection systems of most tropical countries are currently not geared toward collecting length–frequency data representative of catches and landings, with the result that VPA and cohort analysis cannot be applied to those fisheries without a preliminary phase of data acquisition and standardization.

Another important property of VPA and related methods is that the resulting population estimates of young (small) fish are estimates of absolute recruitment. Recruitment is generally extremely difficult to estimate, although it is an important parameter.

It thus seems appropriate to stress here the need for fishery biologists working in tropical countries to help their fisheries department set up a catch reporting system which—at least for major fisheries—will allow for catch-at-length and later catch-at-age data to emerge.

When growth oscillates seasonally, as will generally be the case, the length VPA and length cohort analysis models presented here cannot be used, and must be replaced by an approach that explicitly takes account of such oscillations. The length/age VPA method presented below is such a method, recommend for all cases where seasonal growth oscillations cannot be neglected.

**Length/age VPA**

Length/age VPA is a version of age-structured VPA performed on "cohorts" (actually pseudocohorts) obtained by superimposing growth curves, drawn at monthly intervals, onto a set of catch-at-length data, the catch pertaining to each "cohort" and month being simply that part of the monthly catches contained between two adjacent growth curves (see Fig. 6.3). Using the computed catches of a pseudo-cohort, an age-structured VPA is performed to estimate the population, fishing mortality and biomass.

---

54 Adapted from Pauly et al. (1987).
Fig. 6.3. A: monthly catch-at-length data (not to scale; Peruvian anchoveta (*Engraulis ringens*), Peru) with one of the many monthly cohorts which can be superimposed onto these data. B: population and fishing mortality estimates pertaining to this cohort. C: biomass estimates. The population estimate pertaining to month 1 is an expression of recruitment of fish of length $L_r$ (here, 3.75 cm, i.e. the lower limit of the class with midlength 4.25 cm).

For such cohorts to really consist of fish recruited at the same time, the growth curves used for ‘slicing up’ a cohort must be as close as possible to the true growth curve of that cohort. This, among other things, makes it imperative that a seasonally oscillating growth curve be used since, as shown in Pauly and Ingles (1981) and Pauly (1982, 1985, 1989), virtually all natural fish stocks, including those occurring in tropical waters, display seasonally oscillating growth.

Not all fish of a given cohort have the same growth parameters (Isaac 1990) and it can be expected that some fish will "leave their cohort" because they grow either faster or slower than predicted by the mean cohort growth curve. Such differences in growth rate have the effect of increasing the autocorrelation between estimates of recruitment (Mendelsohn and Mendo 1987, Pauly 1987).

The growth curves used here to "slice" cohorts are based on the seasonally oscillating version of the von Bertalanffy growth function (VBGF) presented by Somers (1988; Equation 4.8); estimates of $t_0$ are not required, as in the case of length-structured VPA. It is assumed here that growth does not change between years, although such changes can be accommodated (see Palomares et al. 1987).

**Interpolation of missing sets of monthly size-composition data**

Although available monthly catch data (in weight) must be complete, the monthly size composition data used to derive monthly catch-at-length data need not be complete. Rather, one may linearly interpolate size–frequency distributions where gaps occurred in the original data (see Fig. 6.4). This approach may not have a major detrimental impact on results for three reasons:

- Linearly interpolated length–frequency samples are very similar to "real" samples (see Fig. 6.4) if only because their shape is determined mainly by continuous, rather "smooth" processes (i.e. growth and mortality).
Estimation of monthly recruitment and biomass

Monthly population and biomass estimates per length class are obtained by regrouping monthly values obtained on a cohort basis into regular class intervals (the same intervals as those in which the catch data were originally grouped).

Fig. 6.5 shows how this is accomplished.

- VPA estimates of $F$ and population size in a given time interval are determined not only by the catch composition in that interval, but by the catch in that interval plus the catches and catch composition in previous intervals (VPA runs backwards!). Hence, whatever error is introduced by interpolating samples will be spread over several intervals, and its absolute impact on mortality and population estimates pertaining to a given interval thus reduced.

- Catch-at-length data as used in this approach are estimated from both length composition data and catch data in weight. While the former may be partly interpolated, the latter are always "real"; hence, the overall number of fish caught in a given month tends to be reasonably approximated, even though their size composition may have been interpolated.

This powerful technique, initially developed by J. G. Pope and D. Pauly in the early 1980s, has been applied so far to only few stocks:

- the northern/central stock of the Peruvian anchoveta (*Engraulis ringens*) (Pauly and Tsukayama 1983, Pauly et al. 1987, Pauly and Palomares 1989);
• a Bolinao (northwestern Philippines) reef-flat population of mottled spinefoot (Siganus fuscescens) (Del Norte and Pauly, 1990);

• a stock of round sardinella (Sardinella aurita) off eastern Venezuela (Mendoza et al. 1994).

Also, Christensen (1995) incorporated this approach as an element of his generic multispecies VPA package.

We recommend the application of this method to various stocks of short-lived species, for which it was conceived, and about which it provides insights not obtainable using other approaches.

Chapter 7: Yield per Recruit and Prediction

What you will learn from this chapter

In this chapter, you will be introduced to the concept and application of the Beverton and Holt yield-per-recruit and Thompson and Bell yield/stock prediction models, which allow prediction of long term yields under different fishing regimes.

Introduction

The previous chapters described the methodologies and models used to estimate the basic parameters of a stock. However, these estimates by themselves are an insufficient basis for deriving the technical advice needed to manage a fishery. To do this, predictions must be made and evaluated, and presented to managers in the form of achievable options.

In this manual, two approaches are provided for predictions and the evaluation of management options: (i) the Beverton and Holt and (ii) the Thompson and Bell models. Both approaches are based on yield-per-recruit consideration, i.e. they estimate the yield that would be obtained from a given number of recruits and a given fishing regime.

As will be discussed later, the model of Thompson and Bell is more versatile than that of Beverton and Holt in that it allows any fishing pattern to be used, while Beverton and Holt's model assumes the same fishing mortality for all (fully recruited and selected) size groups. Both models allow "relative value" (e.g. size-dependent prices, or egg produced per gram of female fish as index of biological value, etc.) to be considered when
computing yields, and this leads to far more realistic management options.

Paradoxically, it is easier, in fishery science at least, to make a long-term prediction than to make a short-term prediction. Both prediction models that will be discussed in the following sections make long-term predictions, i.e. they predict expected average yields assuming an unchanged fishing regime.

**Beverton and Holt yield-per-recruit model**

Beverton and Holt (1957) developed a yield-per-recruit model describing the state of the stock and the expected yield in a situation where a given fishing pattern has been operating for a long time, i.e. under steady-state conditions.

The definition of recruits may vary between authors; here we may visualize recruits as (i) fully metamorphosed young fish, (ii) whose growth is described adequately by the VBGF, (iii) whose instantaneous rate of natural mortality is similar to that of the adults, and (iv) which occur at (or swim into) the fishing ground(s).

Fig. 7.1 describes the fate of a cohort as implied by this model. At age $t_r$ all members of the cohort recruit to the fishing ground; this is called *knife-edge recruitment*, and the number of fish at age $t_r$ is the recruitment ($R$) to the fishery. From age $t_r$ to age $t_c$ (age at first capture), the members of the cohort are not exposed to fishing mortality. (It is thus assumed that all fish of ages between $t_r$ and $t_c$ escape through the meshes if they happen to enter the gear.) Thus, in that period they suffer only from natural mortality ($M$), which is assumed to remain constant throughout the life of the cohort. At age $t_c$ the cohort is assumed to be suddenly exposed to full fishing mortality, $F$, which is also assumed to remain constant for the rest of the cohort's life. Because of this "knife-edge" selection, the catch from the cohort is assumed to be zero before the cohort has attained the age $t_c$. Recruitment ($R$) is defined

\[ R = N_t \quad \ldots 7.1 \]

From Equation 5.1, the number of survivors at age $t_c$ can be expressed as

\[ N_c = R \cdot e^{-M(t_c-t_r)} \quad \ldots 7.2 \]

Thus, at any age $t_i$ ($> t_c$) the number of survivors can be written as

\[ N_i = N_c \cdot e^{-(M+F)(t_i-t_c)} \quad \ldots 7.3 \]

and substituting Equation 7.2,

---

55 Adapted from Pauly (1984a) and Sparre and Venema (1992).
The fraction of \( R \) surviving until age \( t_i \) therefore is
\[
N_i = R \cdot e^{-M(t_i-t_i)(M+F)(t_i-t_i)} \quad \ldots 7.4)
\]
The fraction of \( R \) surviving until age \( t_i \) therefore is
\[
N_i/R = e^{-M(t_i-t_i)(M+F)(t_i-t_i)} \quad \ldots 7.5)
\]

Since the Beverton and Holt model expresses yields on a per recruit basis, the yields are relative, i.e. relative to the recruitment. If, say, a recruitment of 1 million fish gives a yield of 100, then 200 million recruits would yield 20,000. (This assumption may appear trivial, but is not, as one could well imagine that the larger the biomass of a stock becomes, the worse certain factors—such as competition for food—would become for the individuals in that stock.) The results of the model are expressed in units of yield per recruit per time (e.g. grams per recruit per year).

To derive the mathematical expression for yield per recruit, we can take a starting point in the catch equation
\[
C_i = \Delta t \cdot F \cdot N_i \quad \ldots 7.6)
\]
Equation 7.6 gives the number of fish caught \( (C_i) \) during the time period from \( i \) to \( i+\Delta t \) in the life of a cohort. To obtain the corresponding yield in weight, this number should be multiplied by the mean individual weight of the fish. If \( \Delta t \) is small, the body weight of the fish will remain approximately constant during the time period from \( i \) to \( i+\Delta t \), and the yield becomes
\[
Y_i = \Delta t \cdot F \cdot N_i \cdot w_i \quad \ldots 7.7)
\]
where \( w_i \) is the body weight of a fish aged \( i \) years, as defined by the VBGF equation for growth in weight. To obtain the yield per recruit for the period from \( i \) to \( i+\Delta t \), Equation 7.7 is divided by the number of recruits, \( R \):
\[
Y_i/R = \Delta t \cdot F \cdot \left( N_i/R \right) \cdot w_i \quad \ldots 7.8)
\]
where \( N_i/R \) is defined by Equation 7.5.

Equation 7.8 is the Beverton and Holt model for a short period of time where \( \Delta t \) is relatively small. To obtain the total yield-per-recruit \( (Y/R) \) for the entire life span of the cohort, all the small contributions defined by Equation 7.8 must be summed. Beverton and Holt (1957) proposed for this purpose
\[
Y/R = F \cdot e^{3M}W_\infty \left[ \frac{1-e^{-Zr_1}}{Z} \right] - \left[ \frac{3e^{-Kr_1} \left( 1 - e^{-Z+K}r_1 \right)}{Z+K} \right] + 
\]
\[
\ldots \left[ \frac{3e^{-2Kr_1} \left( 1 - e^{-Z+2Kr_1}r_1 \right)}{Z+2K} \right] - \left[ \frac{e^{-3Kr_1} \left( 1 - e^{-Z+3Kr_1}r_1 \right)}{Z+3K} \right] \quad 7.9)
\]
where \( Z = F+M \),
\[
r_1 = t_c-t_\infty,
\]
\[
r_2 = t_c-t_\infty, \text{ and}
\]
\[
r_3 = t_{\text{max}}-t_c
\]
with \( W_\infty, K \) and \( t_\infty \) being the VBGF growth parameters (assuming \( L_\infty \propto W_\infty \)), \( t_c \) the mean age at first capture, \( t_\infty \) the mean age at recruitment, and \( t_{\text{max}} \) the “maximum age of significant contribution to the fishery” or more simply, the longevity of the fish in question (see Ricker 1975).

The effect of \( t_{\text{max}} \) is generally very small, especially when \( Z \) is high, and Equation 7.9 can be considerably simplified by setting \( t_{\text{max}} = \infty \), in which case it becomes
\[
Y/R = F \cdot e^{3M}W_\infty \left[ \frac{1-e^{-Zr_1}}{Z} \right] - \left[ \frac{3e^{-Kr_1} \left( 1 - e^{-Z+K}r_1 \right)}{Z+K} \right] + 
\]
\[
\ldots \left[ \frac{3e^{-2Kr_1} \left( 1 - e^{-Z+2Kr_1}r_1 \right)}{Z+2K} \right] - \left[ \frac{e^{-3Kr_1} \left( 1 - e^{-Z+3Kr_1}r_1 \right)}{Z+3K} \right] \quad 7.10)
\]
Equation 7.10 is the Beverton and Holt yield-per-recruit model (1957), in the form suggested by Gulland (1969).
The two parameters $F$ and $t_c$ are those which may be influenced by fishery managers, because (i) $F$ is usually proportional to fishing effort, and (ii) $t_c$ is a function of gear selectivity.

Applying Equation 7.10 to a range of $F$ values enables us to produce the graph similar to that shown in Fig. 7.2. The $F$ value that provides the maximum value of $Y/R$ (the maximum yield per recruit — $Y/R_{\text{max}}$), is referred to as the optimum fishing mortality, $F_{\text{max}}$.

When $F$ is altered, it takes some time before $Y/R$ becomes equivalent to that predicted by the curve. How long it takes depends, among other factors, on the longevity of the species in question.

Fig. 7.3 presents two yield-per-recruit curves, differing only in their values of $M$. From the figure one may derive the following:

(i) A lower $M$ value produces a lower $F_{\text{max}}$ and a higher $Y/R_{\text{max}}$. Fishing effort levels above $F_{\text{max}}$ lead to a big reduction in total yield.

In some cases the $Y/R$ curve does not have a maximum (see Fig. 7.3B); in such cases one must not conclude that effort should be increased indefinitely. Rather, it is the biomass per recruit ($B/R$; see below) which should be examined (see also Fig. 7.4), and/or another version of the $Y/R$ model should be used, incorporating a realistic selection ogive and/or values (see below).

Because the $Y/R$ model assumes a constant parameter system, the results which may be read from the curve will be true only if the system has had constant parameters for a while.
(ii) When M is high, the Y/R curve will tend to be flat and it may be difficult to estimate $F_{\text{max}}$ on the curve (see Fig. 7.3B).

In the case of (ii), if M is high one must catch the fish before too many die from natural causes, i.e. a high M implies a high F for optimum exploitation. Conversely, if M is low it pays to let the fish grow large before they are caught, because the gain in biomass from individual growth exceeds the loss caused by natural mortality for a large part of the cohort's life.

Equation 7.10 can also be used to assess the effect of different values of $t_c$ (such as those corresponding to a given mesh size) and values of F (corresponding to a certain amount of fishing effort) on yield per recruit. The results of such computations are generally presented in the form of "yield curves", as in Fig. 7.4, from which the effect of increasing mesh size (for example, from a size generating $t_c = 0.2$ year to a size generating $t_c = 0.3$ year) can be assessed.

Further, Equation 7.10 leads to a definition of growth overfishing similar to that form which occurs when F is higher than required to maximize Y/R, i.e. $F_{\text{max}}$ (see Fig. 7.5).

**Biomass-per-recruit**

Mean annual biomass ($\bar{B}$) can be expressed as

$$\bar{B} = N \cdot \bar{W} \quad ...7.11$$

where $\bar{W}$ is the mean weight of the fish in the stock. Substituting Equation 7.11 into Equation 7.7 leads to

$$\bar{Y} = F \cdot \bar{B} \quad ...7.12$$

Thus,

$$\frac{B}{R} = \frac{\bar{Y}}{F} \quad ...7.13$$

which can be inverted

---

56 Adapted from Sparre and Venema (1992).
From Equation 7.10 we may also derive a relationship to estimate the biomass of the exploited part of the cohort:

\[
\frac{B}{R} = e^{\frac{-M}{R} - \frac{1}{2}W} \left( \frac{1}{Z} - \frac{3S}{Z + K} + \frac{3S^2}{Z + 2K} + \frac{S^3}{Z + 3K} \right) \quad ...7.15
\]

where

\[
S = e^{Kt}
\]

Fig. 7.5 shows a B/R curve which illustrates that, for the threadfin bream, *Nemipterus marginatus*, the biomass corresponding to \( F_{\text{max}} \) is only 15 percent of the unexploited stock (\( B_0 \)).

Note that MSY is generated when relative biomass is 50 percent or 37 percent of unexploited stock, at least in terms of the Schaefer (1957) and Fox (1970) models, respectively.

One should ask whether such reductions in (spawning) stock biomass influence the productivity of the stock. Obviously, there must be a level of stock biomass below which recruitment is reduced because of the reduced number of adult fishes. These thoughts lead to a second overfishing concept, that of recruitment overfishing. Here, a stock is “recruitment overfished” if the parent stock biomass is reduced to a level where subsequent recruitment is affected. However, the relationship between recruitment and spawning stock biomass (see Fig. 7.6) is usually not well defined. The only fact known for sure is that with no parents there can be no offspring. As the parent stock starts to increase in size from (close to) zero, recruitment also increases, and then stabilizes, or varies randomly (see Allen 1971; Goodyear 1980).

Another, and sometimes more appropriate, way of using the B/R curve is to exploit the fact that it is (in principle at least) proportional to a catch per unit effort (C/\( f \)) curve. When managing a fishery considerations on economic return per boat are essential, and this quantity is closely related to \( C/f \) and hence to biomass per recruit. This very important aspect relates to fisheries economics, not further dealt with in this manual.

The relative Y/R concept

Beverton and Holt (1964) developed a more general version of their Y/R model (Equation 7.10) based on the realization that, in any case, the absolute value of \( Y/R \), expressed for example in grams per recruit per year, has no direct relation to fisheries management. What matters are the relative differences of \( Y/R \) for different values of \( F \). An example of the kind of conclusions we can draw from the Y/R curve is: if \( F \) increases by 20 percent yield is reduced by 15 percent. The “relative”

---

57 Adapted from Pauly (1984a).
Y/R model resulting from these considerations has the advantage of requiring fewer parameters and is especially suitable for assessing the effect of mesh size regulations.

Let \( L_{c} \) be the 50 percent retention length (\( L_{50} \)), i.e. the length at which 50 percent of the fish are retained by the gear if they enter it, and define the auxiliary variables:

\[
U = 1 - \left( \frac{L_{c}}{L_{w}} \right) \quad \text{...7.16}
\]

and

\[
m = (1 - E)/ \left( \frac{M}{K} \right) = K/Z \quad \text{...7.17}
\]

where \( E (= F/Z) \) is the exploitation rate, or the fraction of deaths caused by fishing, \( M \) the natural mortality and \( K \) the VBGF curvature parameter. Then we have

\[
Y'/R = EU^{m} \left[ 1 - \left( \frac{3U}{1 + m} \right) + \left( \frac{3U^{2}}{1 + 2m} \right) - \left( \frac{U^{3}}{1 + 3m} \right) \right] \quad \text{...7.18}
\]

for computation of "relative" yield-per-recruit (\( Y'/R \)), as shown by Beverton and Holt (1964).

\( Y'/R \) is a function of \( U \) and \( E \) and the only parameter required here is \( M/K \). The equation gives a quantity which is proportional to \( Y/R \) (Equation 7.10); the transformation of \( Y'/R \) into \( Y/R \) is given by

\[
Y'/R = Y/R \cdot e^{m(1-t_{c})/w_{c}} \quad \text{...7.19}
\]

"Relative" biomass per recruit (\( B'/R \)) can be derived from Equation 7.12, which can readily be converted to absolute values, in analogy to Equation 7.18, through the relationship

\[
B'/R = (B'/R) \cdot W_{c} \cdot e^{-M(t_{c}-1)} \quad \text{...7.20}
\]

Note, however, with regard to fishery management, that fishers are not interested in an imaginary "yield per recruit"; they are interested in a physical "yield of fish", and this yield is the product of the yield per recruit times the absolute number of recruits produced in the stock. It is directly proportional to yield per recruit over a wide range of fishing mortalities only if it can be assumed that there is no relationship, over a wide range of \( F \) values, between the size of the parental stock of fish and its progeny.

When the values of \( F \) go beyond that range, "maximum yield per recruit" may be associated with an abysmally low yield (\( Y \)), because the "best" value of \( F \) (the one maximizing yield per recruit) could also reduce the parental stock to a level at which virtually no recruits are produced.

Moreover, it must be realized that the results of yield-per-recruit analyses apply to long-term or equilibrium situations only. In the short term, an increase of fishing mortality or a decrease in size at first capture always results in higher yields, even when yield-per-recruit analysis predicts lower yields. (Conversely, a decrease of fishing mortality or an increase in size at first capture always results in lower yields in the short term, although in the long term higher yields may be attained.)

The duration of the transition period can be of several years in fish which have a high longevity and are subjected to exploitation over a number of years, as in a number of temperate stocks such as cod or halibut. In short-lived animals the transition period will be much shorter; in the case of very short-lived animals, such as most penaeid shrimps, the distinction between "immediate" and "long-term" effect does not even apply, because the stocks are never in equilibrium. This and related problems are reviewed in Garcia and Le Reste (1981) who present a number of methods for the quantification of short- and long-term effects of changes in fishing mortality and mesh size (see also Jones 1981).
Another important feature of the yield-per-recruit model is that yield per recruit is maximized at low values of F only in the case of large, long-lived, low mortality fish (see Fig. 7.3A). In small tropical fish, the values of F which maximize yield per recruit are generally extremely high (see Fig. 7.3B and Fig. 7.4). Thus, managing a tropical fishery based on a species of small fish (let alone a multispecies fishery based on such fish) using only yield-per-recruit analyses can be very misleading (see Pauly 1979b, Pauly and Martosubroto 1980).

It may be mentioned, finally, that in temperate waters, an (arbitrary) agreement has emerged to generally limit F (for assessment of stocks whose stock–recruitment relationships are unknown) to the value which corresponds to 1/10th of the rate of increase of yield per recruit that can be obtained by increasing F at low levels of F (Gulland and Boerema 1973). This concept, called F\textsubscript{0.1}, is illustrated in Fig. 7.7. The F\textsubscript{0.1} concept may be viewed as a surrogate for MEY (maximum economic yield), applicable in situations where economic data on the performance of a fishery are lacking.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
F & Y/R & Diff./10^{-4} & \hline
0.00 & 0.000 & 0.00 & \\
0.01 & 0.030 & 3.00 & \\
0.1 & 0.270 & 2.40 & \\
0.2 & 0.485 & 2.15 & \\
0.3 & 0.656 & 1.71 & \\
0.4 & 0.794 & 1.38 & \\
0.5 & 0.905 & 1.11 & \\
0.6 & 0.995 & 0.90 & \\
0.7 & 1.068 & 0.73 & \\
0.8 & 1.127 & 0.59 & \\
0.9 & 1.175 & 0.48 & \\
\hline
\end{tabular}
\caption{Data for the computation of F\textsubscript{0.1} for Nemipterus marginatus from the South China Sea (see also Fig. 7.7).}
\end{table}

\[ \text{Eq. 7.21} \]

In analogy to F\textsubscript{0.1}, defined above, one can also use, E\textsubscript{0.1}, defined as the exploitation rate at which the marginal increase of relative yield per recruit is 1/10th of its value at E = 0.

The first derivative of Equation 7.18 is

\[ \frac{d(Y'/R)}{dE} = \left[ U^{M/K} (Y'/R) \right] + \left[ \frac{3E \cdot U^{1-M/K}}{(M/K)} \right] \cdot A \quad \text{...7.21} \]

where

\[ A = 2U \cdot \left( 1 + \left[ \frac{2(1-E)}{M/K} \right] - U^2 \right) \]

\[ \text{...7.22} \]

Equation 7.21 can be solved for any value of E, including E = 0 and E = 1. Using Equation 7.21 it is a simple matter, using an appropriate search algorithm and for a given pair of M/K and c values, to identify the value of E generating a value of \( d(Y'/R)/dE \) equal to 1/10th of the value of \( d(Y'/R)/dE \) at E = 0.
Equation 7.20 can also be used to estimate the value of E at which yield per recruit is maximized, i.e. the value of E at which \( \frac{d(Y'/R)}{dE} \) is equal to zero.

**Compensating for the effects of a wide selection range**

In large, long-lived fish, the selection process usually takes place over a relatively narrow range of sizes, such that the assumption of "knife-edge" selection, presented above, is acceptable. In some small animals such as shrimps caught by trawls, the selection range (\( L_{25} \) to \( L_{75} \)) may cover most size classes represented in the population. In such cases, yield-per-recruit computations involving the assumption of knife-edge selection will involve a large bias. A simple method is used here to show the extent of and to help overcome this bias.

Selection curves provide a probability of capture (\( P_i \)) for catch-length class \( i \) between \( L_{\text{min}} \) (the smallest) and \( L_{\text{max}} \) (the maximum length) represented in the available catch samples. In Equation 7.10, it is assumed that \( P_i = 0 \) when \( L < L_c \) and \( P_i = 1 \) when \( L > L_c \) (hence also, \( L' = L_c \)). The implicit assumption here is that, if selection is not knife-edged, the yield from the fish caught below \( L_c \) will compensate for the yield losses due to the fact that not all fish larger than \( L_c \) are caught.

However, although some compensation may occur, the assumption of knife-edge selection does generate a large bias, especially for high values of \( E \), as can be shown by reformulating Beverton and Holt's method for computation of yield per recruit for different \( E \) values over the life span of a fish (Beverton and Holt 1964) such that \( E \) is assumed constant, but \( P \) variable. This gives

\[
\frac{Y'}{R} = \sum_{i=L_{\text{min}}}^{L_{\text{max}}} P_i \left[ \left( \frac{Y'}{R} \right)_i \cdot G_i \right] - \left[ \left( \frac{Y'}{R} \right)_{i+1} \cdot G_i \right] \quad \text{...7.23}
\]

in which \( (Y'/R)_i \) and \( (Y'/R)_{i+1} \) refer to relative yield per recruit as computed from the lower limit of length class \( i \), and \( P_i \) refers to the probability of capture between \( L_i \) and \( L_{i+1} \). Finally, \( G_i \) is defined by

\[
G_i = \prod_{j=1}^{i} r_j \quad \text{...7.24}
\]

where \( r_j \) is a factor expressing the proportion of recruits of length \( L_i \) which survive, grow and reach length \( L_{i+1} \). The value of \( r_j \) is computed, for \( 0 < E < 1 \), from

\[
r_j = \left\{ \left[ \frac{U_i^{(M/K)(E/(1-E)/P_i)}}{U_{i+1}^{(M/K)(E/(1-E)/P_i)}} \right] \right\} \quad \text{...7.25}
\]

Equation 7.24 and 7.25 are analogous to Beverton and Holt's "reduction factor", but \( P \) is considered as a variable and the exponent \( (M/K)(E/(1-E)) \) is used instead of \( F/K \).

Replacing the \( Y'/R \) terms in Equation 7.23 by \( B'/R \) as given by Equation 7.20 is straightforward and will lead to estimates of biomass per recruit independent of the knife-edge assumption (see Fig. 7.8).

The concept presented above can be applied to represent yield per recruit as a function of mesh size or age (size) at first capture and of fishing effort or fishing mortality through a yield isopleth diagram.

---

58 Adapted from Pauly and Soriano (1986).
Table 7.2. Probabilities of capture for a simulation of effects on relative yield per recruit of increasing the selection range (for $L_{\text{c}}/g_{102} = 10$ cm, $M/K = 2$).

<table>
<thead>
<tr>
<th>Length class (cm)</th>
<th>1.0–2.0</th>
<th>2.0–3.0</th>
<th>3.0–4.0</th>
<th>4.0–5.0</th>
<th>5.0–6.0</th>
<th>6.0–7.0</th>
<th>7.0–8.0</th>
<th>8.0–9.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Case 2</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.3</td>
<td>0.7</td>
<td>0.9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Case 3</td>
<td>0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
<td>0.9</td>
<td>1</td>
</tr>
</tbody>
</table>

*aKnife-edge selection.

Table 7.2 gives probabilities of capture for selection ogives with increasingly large ranges, from knife-edge selection in Case 1 to a selection range spanning most of the range between 0 and $L_{\text{c}}$ in Case 3. In Fig. 7.8A, departure from knife-edge selection has a profound impact on yield-per-recruit estimation, particularly at high values of $E$; similar results are obtained for relative biomass per recruit (Fig. 7.8B).

The implications are obvious: one should use the FISAT routine that incorporates knife-edge selection only for reference purposes, while the routine incorporating realistic selection ogives should be used for all assessments.

Thompson and Bell yield–stock prediction

In virtual population analysis (VPA; Chapter 6), estimates of the fishing patterns are based on historical (past) data. The method of Thompson and Bell (1934) is the converse — a prediction of the development of a fishery given assumptions on future recruitment and values of fishing effort (expressed in terms of size-(age-)specific fishing mortality) and gear selectivity.

Such predictions can be made assuming a decrease or increase in $F$ over the entire $F$-array, or changes in $F$ for certain size (age) groups only. The latter is applied, for example, in the case of changes in mesh size, or to separate the effect of different fleets (for example artisanal and industrial fisheries) on the stock.

The method of Thompson and Bell (1934) consists of two main parts:

- an analysis, based on fishing mortalities per size (age) group (a so-called $F$-array), size (age) group-specific catches, deaths, yields, biomasses and values;
- a prediction of the effect of changes in the $F$-array on the catches, deaths, yields, biomasses and values.

The first of these two parts can be achieved through VPA or a slight modification of the catch curve routine, where the fishing mortalities are estimated for each age or length group. If we recall Eqs. 5.1 and 6.9, and assume that fishing mortality is given, the numbers caught for a given period ($\Delta t$) can be predicted through the relationship

$$C_i = (N_i - N_{i+1})(F_i/Z_i)$$

... 7.26)
where \( Z_i = M + F_i \) and the population \( (N) \) at age \( i+1 \) is

\[
N_{i+1} = N_i \cdot e^{-Z_{i+1}}. \quad \ldots 7.27
\]

From Equation 7.7, yield (i.e. the catch in weight in each period) can be computed through the relationship

\[
Y_i = C_i \cdot w_i \quad \ldots 7.28
\]

where \( w_i \) is the mean weight of the fish of that age group. The total yield can be computed by summing all estimated yields for all age groups. It also follows, as with Equation 7.13, that the predicted biomass \( (B_i) \) can be computed from

\[
B_i = \frac{Y_i}{F_t + M} \quad \ldots 7.29
\]

The approximate mean biomass \( (\bar{B}) \) is

\[
\bar{B} = \frac{\sum(B_i \Delta t_i)}{\sum \Delta t_i} \quad \ldots 7.30
\]

When prices are available for each age group, these can be multiplied with the yield to obtain the value of the age group \( (V_i) \), i.e.

\[
V_i = Y_i \cdot v_i \quad \ldots 7.31
\]

where \( v_i \) is the price.

**Single-fleet example**

Consider the data from the Kuwait shrimp fishery (from Garcia and van Zalinge 1982). Columns A to E in Table 7.3 contain the input data, with the fishing mortalities estimated using the swept-area method and catch data. (They may equally have been approximated by cohort analysis or VPA). In Column F we start with an assumed number of 1,000 recruits of one month of age \( (t_r) \) at the beginning of the period. (All subsequent calculations are relative to 1,000 recruits).

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.7</td>
<td>0.73</td>
<td>1.20</td>
<td>4.20</td>
<td>1,000.0</td>
<td>295.3</td>
<td>84.4</td>
<td>481</td>
<td>4,809</td>
<td>351</td>
</tr>
<tr>
<td>2</td>
<td>9.3</td>
<td>0.93</td>
<td>1.32</td>
<td>4.32</td>
<td>704.7</td>
<td>213.0</td>
<td>65.1</td>
<td>605</td>
<td>5,504</td>
<td>563</td>
</tr>
<tr>
<td>3</td>
<td>13.0</td>
<td>1.20</td>
<td>1.32</td>
<td>4.32</td>
<td>491.6</td>
<td>148.6</td>
<td>45.4</td>
<td>590</td>
<td>5,367</td>
<td>708</td>
</tr>
<tr>
<td>4</td>
<td>17.6</td>
<td>1.45</td>
<td>1.44</td>
<td>4.44</td>
<td>343.0</td>
<td>106.1</td>
<td>34.4</td>
<td>606</td>
<td>5,046</td>
<td>876</td>
</tr>
<tr>
<td>5</td>
<td>22.0</td>
<td>1.70</td>
<td>1.92</td>
<td>4.20</td>
<td>236.9</td>
<td>79.7</td>
<td>31.1</td>
<td>684</td>
<td>4,276</td>
<td>1,163</td>
</tr>
<tr>
<td>6</td>
<td>26.1</td>
<td>1.90</td>
<td>1.20</td>
<td>4.20</td>
<td>157.2</td>
<td>46.4</td>
<td>13.3</td>
<td>346</td>
<td>3,463</td>
<td>658</td>
</tr>
<tr>
<td>7</td>
<td>30.3</td>
<td>2.08</td>
<td>1.56</td>
<td>4.56</td>
<td>110.8</td>
<td>35.0</td>
<td>12.0</td>
<td>363</td>
<td>2,793</td>
<td>795</td>
</tr>
<tr>
<td>8</td>
<td>33.8</td>
<td>2.14</td>
<td>1.20</td>
<td>4.20</td>
<td>75.8</td>
<td>22.4</td>
<td>6.4</td>
<td>216</td>
<td>1,261</td>
<td>462</td>
</tr>
<tr>
<td>9</td>
<td>37.0</td>
<td>2.18</td>
<td>1.20</td>
<td>4.20</td>
<td>53.4</td>
<td>15.8</td>
<td>4.5</td>
<td>167</td>
<td>1,667</td>
<td>363</td>
</tr>
<tr>
<td>10</td>
<td>40.3</td>
<td>2.23</td>
<td>1.60</td>
<td>4.80</td>
<td>37.6</td>
<td>12.4</td>
<td>4.7</td>
<td>187</td>
<td>1,250</td>
<td>418</td>
</tr>
<tr>
<td>11</td>
<td>43.1</td>
<td>2.24</td>
<td>2.76</td>
<td>5.76</td>
<td>25.2</td>
<td>9.6</td>
<td>4.6</td>
<td>199</td>
<td>863</td>
<td>445</td>
</tr>
<tr>
<td>12</td>
<td>44.7</td>
<td>2.27</td>
<td>2.52</td>
<td>5.52</td>
<td>15.6</td>
<td>5.8</td>
<td>2.6</td>
<td>117</td>
<td>559</td>
<td>267</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>9.9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

On the basis of the input figures in columns A to E and the number of recruits at age one month, the population per age group, expressed in numbers present at the beginning of each month, can be calculated (column F). Subsequently, the following can be calculated: number of deaths per month (column G); catch in numbers, equivalent to the number of deaths due to fishing (column H); yield in grams (column I); mean biomass in grams (column J); and value of the catch in monetary units (column K), using Equations 7.26 to 7.31.

The last age group (age = 13 months), contains only the number of survivors and none of the other entries. That is because in this example, the number of survivors older than 12 months is small and has therefore been ignored.
In cases where the number of survivors is too large to be ignored, there is a way to account for it, even when basing the computations on the 12 age groups. This is done by treating age group 12 as a so-called plus group, i.e. by replacing the number of deaths between ages 12 and 13 \((N_{12} - N_{13})\) by the total number of deaths occurring beyond age 12. Since all shrimps will eventually die, this number is \(N_{12} = 15.6\) in the example (Table 7.3).

If it is further assumed that the older age groups have the same mortalities as age group 12, the number of shrimp caught from the plus group becomes

\[
C_{12+} = \frac{(F_{12}/Z_{12})}{N_{12}}
\]

or

\[
C_{12+} = (2.52/5.52) \cdot 15.6 = 7.1
\]

Thus, by leaving out the plus groups in Table 7.3 a catch of 7.1 – 2.6 = 4.5 has been ignored. As growth is considered to have halted at age 12 \((w_{12}\) is the maximum body weight), the catch corresponding to \(C_{12+}\) becomes 44.7 \(\cdot\) 4.5 = 201 g, which in turn corresponds to a value of 201 \(\cdot\) 2.27 = 456 monetary units, which constitutes some 6 percent of the total. Therefore, it is suggested that the last group be treated always as a plus group.

With this analysis, all the basic data are available to predict the effect of increases and decreases in fishing effort or fishing mortality. New values for total yield, total value and mean biomass can be computed by raising the fishing mortalities in column D of Table 7.3 by a certain percentage. The F-array presented in Table 7.3 is then replaced by a new one, by multiplying by some factor \((= \text{new } F/\text{reference } F)\).

If, for example, the effort is increased by 20 percent (corresponding to an \(f\)-factor of 1.2), the new fishing mortalities in column D would become

\[1.20 \cdot 1.20 = 1.44, \quad 1.32 \cdot 1.20 = 1.58, \quad \text{etc.}\]

Total yield, total value and mean biomass are obtained by going through the entire procedure for all groups. The results of such calculations are presented in Table 7.4. The reference F-array gives a total yield of 4,636, a total catch value of 7,200 and a mean biomass of 3,176 g. These results were obtained with the same input data as used in Table 7.3; however, there are slight differences from the results presented in that table due to the fact that the computer program used to calculate Table 7.4 had a high precision for its calculations. Such small (and, from the stock assessment point of view, unimportant) differences can also be found in other parts of this manual.

### Table 7.4. Yield, value of yield and biomass for various levels of \(f\). The reference F-array is given in column D of Table 7.3.

<table>
<thead>
<tr>
<th>(f)-factor</th>
<th>Total yield</th>
<th>Total value</th>
<th>Mean biomass</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5,382</td>
</tr>
<tr>
<td>0.4</td>
<td>2,549</td>
<td>4,209</td>
<td>4,271</td>
</tr>
<tr>
<td>0.8</td>
<td>4,055</td>
<td>6,396</td>
<td>3,466</td>
</tr>
<tr>
<td>1.0</td>
<td>4,636</td>
<td>7,200</td>
<td>3,176*</td>
</tr>
<tr>
<td>1.2</td>
<td>4,954</td>
<td>7,485</td>
<td>2,870</td>
</tr>
<tr>
<td>1.5</td>
<td>5,383</td>
<td>7,842</td>
<td>2,522</td>
</tr>
<tr>
<td>2.0</td>
<td>5,814</td>
<td>8,025</td>
<td>2,075</td>
</tr>
<tr>
<td>3.0</td>
<td>6,138</td>
<td>7,683</td>
<td>1,497</td>
</tr>
</tbody>
</table>

*The reference point which describes the present fishing regime, at \(f\)-factor = 1.

Note that the value estimates in Table 7.4 have a maximum whereas the yield estimates had none for the range of \(f\)-factors (effort factors) considered here. This is a common feature of analyses involving prices (e.g. $/kg) that increase with individual weights, and a good reason in itself to always include price in Thompson and Bell analyses and other forms of yield-per-recruit analyses.

### Multi-fleet concept
The shrimp fishery in Kuwait waters is composed of an *artisanal fishery* and an *industrial fishery*. Table 7.5 shows the split of fishing mortality (in Table 7.3) into an artisanal component and an industrial component (from Garcia and van Zalinge 1982). In general, such partitioning of fishing mortality into fleet components is made in proportion to the numbers caught.

Table 7.5. Partitioning of fishing mortality (F) in Kuwait shrimp fishery into an artisanal and an industrial component (from Garcia and van Zalinge 1982; see Table 7.3).

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Artisanal fleet (F1; year⁻¹)</th>
<th>Industrial fleet (F2; year⁻¹)</th>
<th>Total fleet (F total; year⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.720</td>
<td>0.480</td>
<td>1.20</td>
</tr>
<tr>
<td>2</td>
<td>0.960</td>
<td>0.360</td>
<td>1.32</td>
</tr>
<tr>
<td>3</td>
<td>0.840</td>
<td>0.480</td>
<td>1.32</td>
</tr>
<tr>
<td>4</td>
<td>0.480</td>
<td>0.960</td>
<td>1.44</td>
</tr>
<tr>
<td>5</td>
<td>0.600</td>
<td>1.320</td>
<td>1.92</td>
</tr>
<tr>
<td>6</td>
<td>0.480</td>
<td>0.720</td>
<td>1.20</td>
</tr>
<tr>
<td>7</td>
<td>1.080</td>
<td>0.480</td>
<td>1.56</td>
</tr>
<tr>
<td>8</td>
<td>0.480</td>
<td>0.720</td>
<td>1.20</td>
</tr>
<tr>
<td>9</td>
<td>0.084</td>
<td>1.116</td>
<td>1.20</td>
</tr>
<tr>
<td>10</td>
<td>0.120</td>
<td>1.680</td>
<td>1.80</td>
</tr>
<tr>
<td>11</td>
<td>0.240</td>
<td>2.520</td>
<td>2.76</td>
</tr>
<tr>
<td>12</td>
<td>0.240</td>
<td>2.280</td>
<td>2.52</td>
</tr>
</tbody>
</table>

The fishing mortality exerted by fleet fl is

\[
F_{fl} = F_{total} \left( \frac{C_{fl}}{C_{total}} \right)
\]

...7.32)

where \(C_{fl}\) is the number of shrimps (or fish) caught by fleet fl, and \(F_{total}\) and \(C_{total}\) are the fishing mortality and the numbers caught by all fleets. \(F_{total}\) may be derived by virtual population analysis (VPA; see Chapter 6). The split of the predicted catch (column H in Table 7.3) into fleet components is obtained by

\[
C_{fl} = C_{total} \left( \frac{F_{fl}}{F_{total}} \right)
\]

...7.33)

Thus, yield and value of yield are easily separated into fleet components. Table 7.6 shows this split for the total yields and corresponding values of yields given in Table 7.4 (see Fig. 7.9).

Table 7.6. Partitioning of yields and values of yields from Table 7.4 into an artisanal and an industrial component (see also Fig. 7.9).

<table>
<thead>
<tr>
<th>Total yield (000s $)</th>
<th>Total value (000s $)</th>
<th>Artisanal fleet yield (000s $)</th>
<th>Artisanal fleet value (000s $)</th>
<th>Industrial fleet yield (000s $)</th>
<th>Industrial fleet value (000s $)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2,549</td>
<td>4,209</td>
<td>0.4</td>
<td>1,048</td>
<td>1,531</td>
<td>0.4</td>
</tr>
<tr>
<td>4,059</td>
<td>6,396</td>
<td>0.8</td>
<td>1,774</td>
<td>2,486</td>
<td>0.8</td>
</tr>
<tr>
<td>4,580</td>
<td>7,029</td>
<td>1.0</td>
<td>2,048</td>
<td>2,814</td>
<td>1.0</td>
</tr>
<tr>
<td>4,854</td>
<td>7,465</td>
<td>1.2</td>
<td>2,281</td>
<td>3,073</td>
<td>1.2</td>
</tr>
<tr>
<td>5,383</td>
<td>7,842</td>
<td>1.5</td>
<td>2,564</td>
<td>3,354</td>
<td>1.5</td>
</tr>
<tr>
<td>5,814</td>
<td>8,025</td>
<td>2.0</td>
<td>2,904</td>
<td>3,627</td>
<td>2.0</td>
</tr>
<tr>
<td>6,138</td>
<td>7,683</td>
<td>3.0</td>
<td>3,291</td>
<td>3,783</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Fig. 7.9. Total yield (A) and total value of yield (B) from Table 7.5 and their separation into artisanal and industrial components.

In this case the same factors (i.e. \(f\)-factor1 = \(f\)-factor2) are applied to the \(F\)-values of both fleets, since in the exercise demonstrated in Table 7.6 it was assumed that the effort of the artisanal fleet is always the same proportion of the total effort.

Table 7.7 and Fig. 7.10 illustrate an example where \(f\)-factor1 (for the artisanal fleet) is kept constant whereas \(f\)-factor2 (the industrial fleet) is varied. This corresponds to a
situation where the industrial fishery is reduced, or increased, whereas the artisanal fishery remains at the same level. Note that the artisanal fleet gets a smaller share of the total catch the more effort the industrial fleet deploys: the effort increase of the industrial fleet reduces the stock so that a smaller share is left for the artisanal fleet.

Table 7.7. Assessment of the effect of varying industrial effort while artisanal effort is kept constant (see also Fig. 7.10).

<table>
<thead>
<tr>
<th>Total yield</th>
<th>Total value</th>
<th>Artisanal fleet</th>
<th>Industrial fleet</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>f-factor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>yield</td>
<td>value</td>
</tr>
<tr>
<td>2,479</td>
<td>3,603</td>
<td>1.0</td>
<td>2,479</td>
</tr>
<tr>
<td>3,524</td>
<td>5,404</td>
<td>1.0</td>
<td>2,289</td>
</tr>
<tr>
<td>4,270</td>
<td>6,598</td>
<td>1.0</td>
<td>2,124</td>
</tr>
<tr>
<td>4,560</td>
<td>7,029</td>
<td>1.0</td>
<td>2,048</td>
</tr>
<tr>
<td>4,811</td>
<td>7,383</td>
<td>1.0</td>
<td>1,979</td>
</tr>
<tr>
<td>5,120</td>
<td>7,783</td>
<td>1.0</td>
<td>1,883</td>
</tr>
<tr>
<td>5,501</td>
<td>8,203</td>
<td>1.0</td>
<td>1,740</td>
</tr>
<tr>
<td>5,951</td>
<td>8,499</td>
<td>1.0</td>
<td>1,510</td>
</tr>
</tbody>
</table>

Fig. 7.10. Assessment of the effect of regulating the industrial fishery while the artisanal fishery is kept at a constant level.

It is possible to assess in similar manner the effect of any regulatory measure for each fleet component as long as the effort regulation can be re-expressed in form of fishing mortalities. For example, Garcia and van Zalinge (1982) used the Thompson and Bell model to assess the effect of a closed season.

### Length-based version of the Thompson and Bell model

The length-based version of the Thompson and Bell model is, algebraically, the inverse of length-structured VPA (Chapter 6) and carries the same constraints and assumptions as the age-based Thompson and Bell model (see above). The algebraic inverse of Jones’ length cohort analysis (JLCA) could be applied; however, VPA is recommended over JLCA, which will not be presented in this manual.

While the same procedural flow is followed as with the age-structured Thompson and Bell model, some equations need to be altered to account for the conversion of length to age (or to relative age when $t_i$ is not known).

In Equation 7.26, the populations $N_i$ and $N_{i+1}$ are computed as the number of fish that would attain lengths $L_i$ (class lower limit) and $L_{i+1}$ (class upper limit) respectively. Equation 7.27 can be applied; however, $\Delta t_i$ is computed here based on Equation 6.11.

Given the length–weight relationship and the population per length group, the corresponding yield ($Y_i$) and value can be estimated through Equation 7.28. However, estimating the mean biomass by simply summing the predicted populations would be erroneous because a change in class interval would result in a different sum. In this case, the predicted population ($N_i$) refers only to the numbers alive at length $L_i$.

The proper procedure is to compute the mean number of fish in each length class and divide it by $\Delta t_i$. 

---
From Equation 5.1 the annual mean number in each length class can be expressed as,

\[ N_t' \cdot \Delta t_t = \left( N_i - N_{i+1} \right) / Z \quad ...7.34 \]

and correspondingly, the mean annual mean biomass in each length group can be computed as

\[ B_t' \cdot \Delta t_t = N_t' \cdot \Delta t_t \cdot \bar{w}_i \quad ...7.35 \]

where the mean weight of the length group (\( \bar{w}_i \)) can be estimated through Equation 6.16. The average biomass during the life span of the cohort, which is now independent of the length class interval, can be expressed as

\[ \bar{B} = \sum (B_i \cdot \Delta t_i) \quad ...7.36 \]

<table>
<thead>
<tr>
<th>Length group</th>
<th>N</th>
<th>Biomass</th>
<th>Yield</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>6–12</td>
<td>98,919.3</td>
<td>330.7</td>
<td>13.3</td>
<td>13.3</td>
</tr>
<tr>
<td>12–18</td>
<td>84,392.7</td>
<td>1,260.1</td>
<td>488.1</td>
<td>488.1</td>
</tr>
<tr>
<td>18–24</td>
<td>59,475.8</td>
<td>2,191.5</td>
<td>2,336.3</td>
<td>2,336.3</td>
</tr>
<tr>
<td>24–30</td>
<td>27,623.0</td>
<td>2,475.2</td>
<td>1,401.0</td>
<td>2,401.0</td>
</tr>
<tr>
<td>30–36</td>
<td>15,967.8</td>
<td>2,845.9</td>
<td>1,397.6</td>
<td>2,096.4</td>
</tr>
<tr>
<td>36–42</td>
<td>9,861.5</td>
<td>2,970.1</td>
<td>1,755.2</td>
<td>3,510.5</td>
</tr>
<tr>
<td>42–48</td>
<td>5,800.5</td>
<td>2,636.4</td>
<td>1,704.9</td>
<td>3,409.9</td>
</tr>
<tr>
<td>48–54</td>
<td>2,818.8</td>
<td>2,247.1</td>
<td>866.2</td>
<td>2,165.5</td>
</tr>
<tr>
<td>54–60</td>
<td>1,691.5</td>
<td>2,069.4</td>
<td>596.3</td>
<td>1,490.8</td>
</tr>
<tr>
<td>60–66</td>
<td>1,056.6</td>
<td>1,853.8</td>
<td>570.1</td>
<td>1,710.3</td>
</tr>
<tr>
<td>66–72</td>
<td>621.0</td>
<td>1,481.9</td>
<td>594.6</td>
<td>1,783.8</td>
</tr>
<tr>
<td>72–78</td>
<td>313.7</td>
<td>1,040.1</td>
<td>405.0</td>
<td>1,215.0</td>
</tr>
<tr>
<td>78–84</td>
<td>148.7</td>
<td>772.0</td>
<td>85.0</td>
<td>255.1</td>
</tr>
<tr>
<td>84–∞</td>
<td>92.0</td>
<td>2,012.6</td>
<td>5,635.1</td>
<td>6,690.6</td>
</tr>
<tr>
<td>Total</td>
<td>308,482.9</td>
<td>26,188.8</td>
<td>12,977.1</td>
<td>24,567.1</td>
</tr>
</tbody>
</table>

By multiplying the F-array by a factor (the f-factor) to simulate an increase or decrease of fishing effort, outputs similar to those in Fig. 7.11 can be derived. Thus, managers can straightforwardly estimate the amount of effort that should be added to or removed from a fleet.
The parameters used in this model assume, among others, a constant natural mortality. The procedure should therefore be run several times with different values of $M$, to evaluate the impact of varying natural mortality on the assessments.

**Multispecies/multi-fleet Thompson and Bell model**

Yield-per-recruit models such as Beverton and Holt's or Thompson and Bell's can be used straightforwardly in multispecies situations if two crucial assumptions are met:

(i) The fishing pattern has no influence on recruitment.

(ii) Biological interactions among species can be neglected.

Assumption (i) implies here not only that over a wide range, recruitment is not affected by changes in the effort level — as is also assumed for single-species $Y/R$ analyses — but also that the relative strength of recruitment between species remains unaffected by fishing. Thus, it is assumed that if three species A, B and C recruit to the fishing ground with relative strengths of 0.1, 0.6 and 0.3 respectively, species B will remain dominant even if its adults are targeted by the fishery.

This assumption is not likely to be met in reality — at least not strictly. However, radical changes of the relative species composition of a multispecies stock take a while to manifest themselves, even when they are induced by a fishery. Also, there are configurations that are more stable than others, with certain species remaining dominant over decades. Finally, it must be recalled that yield per recruit analyses usually lead to advice that, when implemented, may be conducive to stabilizing recruitment to the stock, especially when these analyses consider spawning biomass per recruit.

Assumption (ii), that species do not interact biologically means, in terms of the multi-species version of Thompson and Bell proposed here, that the species-specific $M$ values do not change as a function of fishing mortality; thus, it is assumed among other things that the natural mortality of small fish remains constant irrespective of the biomass of large fish, i.e. of actual and/or potential predators.

This assumption is evidently not likely to be met in any real stock. Models exist (e.g. multispecies VPA) in which $M$ is explicitly made to vary with predator biomass and size (age) structure (Christensen 1995). However, even without variable natural mortalities, the multispecies version of the Thompson and Bell model represents an improvement over the single-species approach. Further, there is always the possibility of running the model several times, with different values of $M$ such as to be able to assess the effects of changes of $M$ on yields.

With this in mind, we might for example use the multispecies version of the Thompson and Bell model to study a fishery with three dominant species A, B and C, the catch of which is determined not only by the fishing operation, but also by the availability of the fish (the species in question inhabit the same fishing grounds and are caught together). We might further assume length–frequency data from these three major species to be available; and that the price per kg differs between species and between size categories of the same species.

In this case, one cannot treat each species separately and subsequently add the results in terms of yield. Before adding makes sense, the yield must be converted into comparable units, e.g. "common value". Moreover, even if yield is converted into value, it is still not possible to add the results of single species assessments. This is because the effort level which gives the maximum economic return for species A ($V_{max}$), will usually not generate $V_{max}$ for species B and C.
Rather, in a multi-fleet environment, we assume that the following relationship is true:

\[ Z = M + \sum F_{fl} \]  

...7.37)

where \( F_{fl} \) is the fishing mortality caused by fleet \( fl \) and that the numbers caught by fleet \( fl \) for a single species are

\[ C_{fl} = \left( F_{fl}/Z \right) \cdot N_0 \cdot (1 - e^{-Z}) \]  

...7.38)

where \( N_0 \) is the standing stock in numbers.

The flow diagram that follows (Fig. 7.12) summarizes the procedure of the multispecies/multi-fleet Thompson and Bell model, i.e. the implementation of Equations 7.37 and 7.38.

As will be noted, the main concern with the multispecies/multi-fleet methodology of Thompson and Bell is how the fishing mortality is treated: the \( F \)-arrays derived from species-specific independent VPAs must incorporate the catches from all fleets. This allows the results to be redistributed, based on the catches made by each fleet.
Chapter 8. Monte Carlo Simulation

What you will learn from this chapter

In this chapter, you will learn how FiSAT simulates length–frequency distributions under several scenarios, including some pertaining to migratory stocks.

Introduction

In recent years length-based fish stock assessment has been applied to many tropical fish stocks. Several of the methods used for these applications, however, are relatively new and yet to be fully tested.

One way of testing length-based methods for growth parameter estimation and ancillary data is to use data for which the age distribution is known (from reading of hard parts, such as otoliths or scales), and for which accurate estimates of growth parameters can therefore be obtained. If the length-based method is capable of producing the same results as those derived from the age distribution, the method has passed the test. However, age composition data are rare for tropical fish stocks.

An alternative approach is "Monte Carlo" simulation. This computer-based technique can be used to simulate the dynamics of a fish stock and samples taken from it while taking into account various stochastic (randomly varying) elements of the biological system being simulated, and of the sampling procedure.

Adapted from Sparre and Venema (1992).
Essentially, simulation attempts to reconstruct or model certain parts of reality; and the performance of the process depends largely on our understanding of the real system.

The advantage of a simulated sample compared with a real sample is that the true parameters of the simulated sample (e.g. the VBGF parameters, the mortality rates, the gear selection parameters) are known.

However, the disadvantage is that the simulated population (apart from the superimposed “stochastic noise”) behaves exactly as believed by the simulator, though it is by no means obvious that our knowledge of fish populations is sufficient to ensure that such simulated behavior is in fact “realistic”. Thus, for example, the simulated samples may behave too “nicely”, as they do not incorporate the complexity of real sampling situations (see also Isaac 1990).

Erzini (1990) presents an example of a simulation useful for length-based assessment: an investigation of the effect of sample size and length class intervals on parameter estimation.

**Random number generators**

The number of fish in each length group is simulated as a quantity drawn at random from a population. A tool to simulate random sampling is therefore required: a “random number generator”.

The drawing of numbers at random from a population could be done with a roulette, such as used in the casinos of Monaco (hence “Monte Carlo”). There is, however, a much speedier way of simulating random numbers: the random number generators available in almost all computer languages (for example RND in the BASIC programming language).

Most random number generators produce, each time a program is executed, the same sequence of (pseudo) random numbers, drawn with equal probability between 0 and 1. Hence, it is possible to duplicate their results. However, it is also possible to create different sequences of random numbers by providing different "seeds" to the random number generator.

The mean value of RND is 0.5 and the variance is 1/12. Thus if RND₁, RND₂, ..., RNDₙ designate n calls of RND, each of which gives a random number, the sum

\[
X = \frac{1}{\sqrt{n/12}} \sum_{i=1}^{n} (\text{RND}_i - 0.5)
\]

will have mean value 0 and variance 1, for any value of n.

According to the central limit theorem (Rao 1965), X is approximately normally distributed. The larger n is, the better the approximation. If a random program produces many random numbers, we should expect the mean value and variance of these numbers to move closer to 0 and 1 respectively.

The distribution is a bell-shaped curve, symmetric about the mean, X. However, in the simulation procedure described below, where values of growth parameters (L₀, K, C and WP) are drawn from distribution, the possibility exists that the values thus generated have biologically impossible (e.g. negative) values. To avoid this, FiSAT uses normal distributions that are truncated on both side by the 95 percent probability limit.

**Simulation of length–frequency samples**
There are several ways to simulate length–frequency (L/F) samples using the basic Monte Carlo approach. In FiSAT, simulation of L/F data starts by assigning a fish of a randomly drawn age to a sample, after drawing the life history of that fish from a series of probability distributions which includes time of birth, recruitment, selection and random variability of growth. The procedure is repeated as many times as may be required to obtain the desired sample size and may be summarized as follows (see also Table 8.1):

**Step 1.**

This initial step is basically to determine if the randomly drawn age of a fish (or of the fish in a sample) lies within a specified age bracket (an input). This can be accomplished by drawing from probability distributions pertaining to the birth date, i.e. year and month (the day is here assumed to be the 15th of the month) and an individual set of growth parameters, generated from the mean growth parameters and their standard deviations.

Let $t_{\text{max}}$ be the maximum age a fish can reach (an input). The age, $t$, of the fish is drawn from the distribution:

$$P_t = \begin{cases} 
\frac{1}{t_{\text{max}}} & \text{if } t = 0,1,2,\ldots, t_{\text{max}} \\
0 & \text{otherwise}
\end{cases}$$

Table 8.1. Summary and sequence of procedures which define a Monte Carlo simulation to generate length–frequency samples.
This approach assumes that recruitment remains constant from year to year, i.e. that in each year there is the same probability that a fish will be hatched. If year-to-year differences in recruitment must be taken into account, this can be done by drawing ages \( t \) from a set of pre-specified distributions, one for each year of recruitment.

Thus, for pre-recruits there are only two possible options: to be recruited or not to be recruited. Let \( t \) be the year the 0-group was recruited, \( t-1 \) the year the 1-group was recruited, etc., where each distribution is assigned a relative strength of occurrence. We may call this distribution the relative year class strength. This probability may therefore be generalized as

\[
P_t = \frac{C_y}{\sum C_y} \quad \text{...8.2)}
\]

where \( C_y \) is the year class strength for class \( i \).

**Step 2.**

It is assumed that the fish of the stock in question spawn all year round and that recruitment thus takes place all year round. This is approximated by allowing for a number of recruits to enter the stock on the first day of each month. The drawing of the month of birth is similar to the drawing of age, where the probability of occurrence is drawn from the uniform distribution

\[
P_t = \frac{1}{12} \quad \text{(for 12 months/year)}
\]

This, however, may not be the case when the seasonality of recruitment is pronounced, in which case relative monthly recruitment strength \( (R') \) must be considered. The treatment of these inputs is similar to that of the previous step, where the probability distribution can be expressed as

\[
P_t = \frac{R'_t}{\sum R'_t} \quad \text{...8.3)}
\]

where \( R' \) is the relative monthly recruitment strength.

If we add the mathematical constraint that both probability distributions must have 1 as their maximum value, we can convert these relative quantities into absolute number of recruits. (Note that the constraint has no impact on the probability distributions.)

If we let \( N_{\text{max}} \) be the maximum number of recruits in any month of the period considered,

\[
\text{monthly recruitment} = \left(\frac{1}{12}\right) \cdot \left(\frac{R'_t}{\sum R'_t}\right) \cdot N_{\text{max}} \quad \text{...8.4)}
\]

whose value is expressed as absolute number of recruits.

**Step 3.**

The growth parameters \( L_\infty \) (asymptotic length), \( K \) (curvature parameter), \( C \) (amplitude of seasonal growth oscillations) and WP (“winter point”) are similarly drawn, from normal distributions, with pre-specified mean values and standard deviations as inputs.
Step 4.

At this stage of the simulation, the effects of migration can be simulated if so desired; otherwise proceed to Step 5.

The simulation of samples from a migratory stock has elements similar to those described above, but occurring at a number of scales, in a number of areas. Thus, when simulating migrations we must consider a number of areas, and simulate the movement of the stock among these areas.

Here, only large-scale horizontal migration is considered. Diurnal migrations (e.g. at the bottom during the day and in the water column at night) are not considered, although they might also affect length–frequency data, depending on the sampling gear and strategy.

The total area occupied by the stock in question is conceived here to consist of several "boxes" (Fig. 8.1). The fish stay in a box for a certain time period, and move to an adjacent box at the end of that period (note that the transition from one box to the next is assumed to be instantaneous).

Using only few boxes makes the model a crude approximation to reality. To approximate a continuous, smooth migration, a large number of boxes must be used. However, this requires more computer time and may result in an output that is difficult to analyze.

The migration route and the time spent in each area are deterministic quantities, i.e. inputs to the system. With regard to migrations, the random number generator is utilized only to determine the time when the fish leave the nursery ground (box 1 in Fig. 8.1): the model used is thus mathematically equivalent to that used for gear selection. In line with much biological evidence, but also to simplify the simulation, we assume size-dependent migration, i.e. the large and/or fast growing specimens are the first to leave the nursery ground, and generally the first to arrive at, and to leave boxes during their migration.

The probability that a fish of length \( L_i \) will leave its nursery ground at time \( t_i \) can be expressed as

\[
P_t = \frac{1}{1 + e^{S_5 + S_4 t_i}}
\]

where

\[
S_5 = \ln(3) \cdot \left[ L_i \left/ \left( R_{75} - R_{50} \right) \right. \right]
\]

and

\[
S_4 = -S_5 / R_{50}
\]
and \( R_{50} \) and \( R_{75} \) are lengths at which 50 and 75 percent of the fish recruit from the nursery ground into a “fishable” box (boxes 2 to 5 in Fig. 8.1), as determined by their size (in Step 1).

**Step 5.**

The probability of survival is calculated from the age of the fish and the cumulative total mortality \((Z)\) over all months.

Let \( F_{\text{max}} \) be the maximum fishing mortality a fish is exposed to during its life; \( M \) the natural mortality; and \( P \) the product of "left-hand selection" and "right-hand selection" (or de-selection and/or de-recruitment). \( F_{\text{max}}, M \) and \( P \) are given as input.

The cumulative total mortality then can be calculated as

\[
Z = \sum F_{\text{max}} \cdot P_{L_i} + M_i \quad \ldots8.8)
\]

Left-hand selection accounts for gear selection for small fish (e.g. small fish escape through the meshes of a trawl) and incomplete recruitment to the fishing ground. Right-hand selection accounts for gear selection against large fish (such as large fish not gilled in a gill net) and/or migration out of the fishing grounds, as applied by Equation 8.5. Thus,

\[
P_L = \left[ 1 / \left( 1 + e^{S_1 + S_2 + L_i} \right) \right] \cdot \left[ 1 / \left( 1 / 1 + e^{S_3 + S_4 - L_i} \right) \right] \quad \ldots8.9)
\]

where

\[
S_1 = \ln(3) \cdot L_{50} / (L_{75} - L_{50})
\]

\[
S_2 = -S_1 / L_{50}
\]

\[
S_3 = \ln(3) \cdot R_{50} / (R_{75} - R_{50})
\]

\[
S_4 = -S_3 / R_{50}
\]

and where \( L_{50} \) and \( L_{75} \) are the lengths at which 50 and 75 percent of the fish are available for left-hand selection and \( R_{50} \) and \( R_{75} \) are the lengths at which 50 and 75 percent respectively of the fish are not available, due to right-hand selection (or de-selection or de-recruitment).

The probability of survival is \( e^{-Z} \). Thus, survival has the probability distribution:

\[
P_t = e^{-Z} \quad \ldots8.10)
\]

By drawing from this distribution, it can be decided whether to include or exclude a fish in the sample.

**Step 6.**

The probability of being caught is given by the probability that a fish is available to the fishing gear (i.e. that it occurs on the fishing ground) and that it is large enough to be retained if it enters the gear. The probability distribution of being caught is thus expressed by a "resultant" curve, i.e. the product of the curve expressing the recruitment to the fishing ground and the gear selection curve proper (see the discussion of catch curves in Chapter 5).

Let \( P_L \) be the value of the resultant curve for length \( L \). \( P_L \) is the fraction of the stock of length \( L \) that is available to the fishery. \( P_L \) takes account of the fact that some fish are not available to the fishery because they have not yet been recruited.
to the fishing grounds or they have left the fishing grounds, and that other fish are not caught because of gear selection.

Whether to include or exclude a fish from the sample is thus determined by the probability distribution

\[ P_t = P_L \]

The samples not included in the length frequency file generated by FiSAT are also recorded to allow estimation of the total population.

**Step 7.**

This stage of the simulation is to determine if the desired number of samples has been reached. This can be achieved rather straightforwardly by adding all samples from the previous runs. If this does not produce the desired number of samples, repeat the process from Step 1. Otherwise, the array that records the length–frequency samples is saved as a file or printed as output of the simulation process.

The simulated L/F data generated by Steps 1–7, which may or may not consider migrations (Step 4), can be used for a variety of purposes, notably:

- to evaluate the precision and/or accuracy of a given method or set of methods for parameter estimation;
- to generate samples for testing students’ ability to analyze L/F data using, for example, FiSAT;
- to help identify, by comparison and backward inference, the process(es) which may have generated a given set of observed L/F data.

**Conclusion**
References


* This and subsequent articles by F. I. Baranov are available in English translation in “Selected Works on Fishing Gears: Theory of Fishing”. Israel Program for Scientific Translation, Jerusalem, Vol. 3. 234p.


Fischer, W., G. Bianchi and W. B. Scott (eds), 1981. FAO species identification sheets for fishery purposes. 7—Eastern Central Atlantic (Fishing Areas 34, 47). Canada Funds-in-Trust, Dept. of Fisheries and Oceans, Ottawa, Canada: by arrangement with FAO. Pag. var.


Garcia, S. and L. Le Reste, 1981. Life cycles, dynamics, exploitation and management of coastal penaeid shrimp...
stocks. FAO Fish Tech Pap. (203). 215p. (Issued also in French and Spanish)


Symbols and Definitions*

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>multiplicative factor generally used in exponential relationships linking L and W ($W = aL^b$); Y-intercept in a Type I, or AM linear regression</td>
</tr>
<tr>
<td>a'</td>
<td>Y-intercept in a Type II, or GM linear regression</td>
</tr>
<tr>
<td>A</td>
<td>number of fish dead after a given time; i.e. $A = 1 - S$, also relative abundance</td>
</tr>
<tr>
<td>AM</td>
<td>arithmetic mean; used to characterize Type I regressions</td>
</tr>
<tr>
<td>ASP</td>
<td>&quot;available sum of peaks&quot;; the sum of available &quot;points&quot; in a file restructured for analysis with the ELEFAN I routine</td>
</tr>
<tr>
<td>b</td>
<td>exponent of a length–weight relationship; slope of a Type I (AM) linear regression</td>
</tr>
<tr>
<td>b'</td>
<td>slope of a Type II (GM) linear regression</td>
</tr>
<tr>
<td>B</td>
<td>biomass, or stock size in weight of a population of fish</td>
</tr>
<tr>
<td>B̄</td>
<td>mean annual biomass of a stock</td>
</tr>
<tr>
<td>B₀</td>
<td>biomass prior to any fishing; unfished biomass</td>
</tr>
<tr>
<td>Bₓ</td>
<td>environmental carrying capacity for a given stock, in weight</td>
</tr>
<tr>
<td>B/R</td>
<td>biomass per recruit</td>
</tr>
<tr>
<td>B'/R</td>
<td>relative biomass per recruit</td>
</tr>
<tr>
<td>c</td>
<td>the fraction $L_c/L_x$</td>
</tr>
<tr>
<td>c₁, c₂</td>
<td>multipliers for estimating $Z$ and its standard error using one of Hoenig's methods</td>
</tr>
<tr>
<td>c.f.</td>
<td>condition factor</td>
</tr>
<tr>
<td>cov</td>
<td>covariance</td>
</tr>
</tbody>
</table>

* Only the most commonly used combinations of symbols and subscripts are listed here.
C catch in numbers; parameter expressing the amplitude of seasonal growth oscillations in the VBGF; a constant
C.V. coefficient of variation, i.e. C.V. = s.d./\bar{X}
(\text{also expressed in percent, i.e. C.V. = s.d.} \cdot \frac{100}{\bar{X}})
C/f catch per unit of effort (also: CPUE)
Ci,A cumulative catch in numbers for mesh size mA
Ci,B cumulative catch in numbers for mesh size mB
Ci,e cumulative catch in numbers from length i to L
Ci terminal catch, as used in VPA
\Delta denotes a difference between two lengths, times, etc.
\Delta \bar{L} difference between two successive mean lengths
\Delta L length increment; width of length class in grouped data
\Delta L/\Delta t growth rate expressed as difference
a temperature difference, such as the difference between warmest and coldest mean monthly temperature at a given place
\Delta t time difference, such as the time needed by an average fish to grow from the lower to the upper limit of a length class
d power of weight to which anabolism is proportional; also, a constant term
d.f. degrees of freedom, i.e. the "real" number of cases available for testing a statistical hypothesis
dt used to represent a rate, such as of growth
dB/dt growth rate (in weight) of a fish population, in numbers
dL/dt
dW/dt growth rate (in length); first derivative of the von Bertalanffy equation
e base of the natural (or Naperian) logarithms; e = 2.71828...
E exploitation rate; E = F/Z
E0.1 level of exploitation at which the marginal increase in yield per recruit reaches 1/10th of the marginal increase computed at a very low value of E
E0.5 exploitation level which results in a reduction of the unexploited biomass by 50 percent
E_{\text{max}} exploitation level which maximizes Y/R or Y'/R
ESP "explained sum of peaks"; the points "explained" (i.e. hit) by a growth curve traced by the ELEFAN I routine
f fishing effort
f-factor factor used as a multiplier to simulate a change in effort level, for a defined fishing regime
fl index for fleet in Thompson and Bell prediction model
f_{\text{MSY}} fishing effort generating MSY
F instantaneous rate of fishing mortality, i.e. rate of mortality, of dimension time^{-1}, due to fishing gears, whether or not the killed fish are caught and landed
F_{0.1} an arbitrary limit to fishing mortality, set at a level such that the marginal increase in yield
per recruit is one-tenth of its value at the origin of the curve; often used as a proxy for $F_{MEY}$.

- $F_{max}$: fishing mortality generating maximum yield per recruit
- $F_{opt}$: optimum fishing mortality
- $F_t$: terminal fishing mortality, as used in VPA and cohort analysis
- $F_{MSY}$: fishing mortality generating MSY; note that $F_{MSY} > F_{MEY}$
- $FL$: fork length

- $GM$: geometric mean, used for Type II regression
- $H$: coefficient of anabolism, used in the derivation of the VBGF
- $\infty$: symbol of infinity, as in $L_{\infty}$, i.e. the length associated with an infinitely high age
- $i$: symbol or subscript used in some equations for counting items (samples, means, etc.)
- $j$: symbol or subscript used in some equations for counting items (samples, means, etc.)
- $JLCA$: Jones' length cohort analysis
- $k$: coefficient of catabolism; the number of parameters estimated by a given procedure
- $K$: curvature parameter of the VBGF; rate of dimension time$^{-1}$ at which $L_{\infty}$ is approach

- $L$: any linear measurement; length
- $\bar{L}$: mean length of fish, computed from $L'$ upward; mean of two or more lengths, such as mean of length at tagging and at recapture
- $L'$: a length not smaller than the smallest length of fish fully represented in catch samples; used to compute $\bar{L}$
- $L_{\infty}$: parameter of the VBGF, expressing the asymptotic length, i.e. the mean length the fish in a population or stock would reach if they were to grow indefinitely; often close to the length of a very old fish
- $L^*$: largest observed specimen in a sample
- $L/F$: length–frequency or length–frequency samples
- $L_{25}$: length at which 25 percent of the fish will be vulnerable to the gear (left-hand selection)
- $L_{50}$: length at which 50 percent of the fish will be vulnerable to the gear (left-hand selection)
- $L_{75}$: length at which 75 percent of the fish will be vulnerable to the gear (left-hand selection)
- $L_c$: mean length of fish at first capture; equivalent to $L_{50}$
- $L_{m}$: length at tagging or marking
- $L_{mass}$: mean length at first maturity (or "massive maturation")
- $L_{max}$: upper class limit of the largest class in a sample; maximum length reached by the fish of a given stock; may also be predicted from the largest specimens of several samples using the extreme value theorem
- $L_{min}$: smallest length represented in one or several samples
- $L_{r}$: length, in Baranov/Holt selection model, at which $P$ is equal to 1
- $L_{rec}$: length at recapture or at recruitment
- $L_{r'}$: computed length at recapture given growth parameters ($L_c$ and $K$) and length at marking
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_t$</td>
<td>(mean) length at age $t$</td>
</tr>
<tr>
<td>$L_A$</td>
<td>optimum length for mesh size $m_A$</td>
</tr>
<tr>
<td>$L_B$</td>
<td>optimum length for mesh size $m_B$</td>
</tr>
<tr>
<td>$m$</td>
<td>mesh size</td>
</tr>
<tr>
<td>$m'$</td>
<td>a quantity used when restructuring length–frequency data</td>
</tr>
<tr>
<td>$m$''</td>
<td>mean growth rate difference between winter and summer</td>
</tr>
<tr>
<td>$m_A$</td>
<td>gillnet mesh size</td>
</tr>
<tr>
<td>$m_B$</td>
<td>another gillnet mesh size, with $m_B &gt; m_A$</td>
</tr>
<tr>
<td>$M$</td>
<td>instantaneous rate of natural mortality, i.e. due to all causes except fishing</td>
</tr>
<tr>
<td>$MEY$</td>
<td>maximum economic yield</td>
</tr>
<tr>
<td>$ML$</td>
<td>&quot;midlength&quot; or length class midpoint; midpoint of a class interval</td>
</tr>
<tr>
<td>$MPA$</td>
<td>modal progression analysis</td>
</tr>
<tr>
<td>$MSY$</td>
<td>maximum sustainable yield, a frequently criticized but still useful concept</td>
</tr>
<tr>
<td>$n$</td>
<td>number of items in a sample, number of cases investigated, etc.</td>
</tr>
<tr>
<td>$n_z$</td>
<td>number of classes with a frequency of zero adjacent to a given length class</td>
</tr>
<tr>
<td>$N$</td>
<td>number of fish in a length class</td>
</tr>
<tr>
<td>$N_o$</td>
<td>initial number of individuals in a population exposed to some source of mortality</td>
</tr>
<tr>
<td>$N_t$</td>
<td>initial population (in number) of a cohort, i.e. at age $t$</td>
</tr>
<tr>
<td>$N_{t_i}$</td>
<td>number of fish in length class $i$ following restructuring of a length–frequency sample</td>
</tr>
<tr>
<td>$N_{t_i}$</td>
<td>number of fish in the oldest age group of a cohort or population (&quot;terminal population&quot;)</td>
</tr>
<tr>
<td>$\pi$</td>
<td>equal to 3.1415...</td>
</tr>
<tr>
<td>$\phi$</td>
<td>phi, i.e. a weight-based index of growth performance ($\phi = \log_{10}K + \frac{2}{3}\log_{10}W$)</td>
</tr>
<tr>
<td>$\phi'$</td>
<td>phi-prime, i.e. a length-based index of growth performance ($\phi' = \log_{10}K + 2\log_{10}L$)</td>
</tr>
<tr>
<td>$\propto$</td>
<td>proportionality symbol</td>
</tr>
<tr>
<td>$p$</td>
<td>multiplicative factor in Equation 4.2</td>
</tr>
<tr>
<td>$P$</td>
<td>probability of capture or occurrence on the fishing ground; also position in extreme value plot</td>
</tr>
<tr>
<td>$P_1$</td>
<td>first point of a length-converted catch curve included in the computation of Z; this point is by definition the first where the probability of capture is 1</td>
</tr>
<tr>
<td>$P_{i+\frac{1}{2}}$</td>
<td>interpolated probability of capture for the upper limit of class $i$</td>
</tr>
<tr>
<td>$P_{\text{max}}$</td>
<td>highest point in a catch curve, usually a point to the left of $P_1$</td>
</tr>
<tr>
<td>$P_{\text{last}}$</td>
<td>last point to be included in a catch curve analysis</td>
</tr>
<tr>
<td>$q$</td>
<td>multiplicative factor in Equation 4.2</td>
</tr>
<tr>
<td>$r$</td>
<td>product–moment correlation coefficient</td>
</tr>
<tr>
<td>$r^2$</td>
<td>coefficient of determination</td>
</tr>
<tr>
<td>$r_m$</td>
<td>intrinsic rate of growth of a stock</td>
</tr>
<tr>
<td>$R$</td>
<td>recruitment; also true relative abundance of an age group</td>
</tr>
<tr>
<td>$R'$</td>
<td>relative strength of monthly recruitment</td>
</tr>
<tr>
<td>$R_{25}$</td>
<td>length at which 25 percent of the fish will no longer be vulnerable to the gear (right-hand selection, or deselection)</td>
</tr>
<tr>
<td>$R_{50}$</td>
<td>length at which 50 percent of the fish will no longer be vulnerable to the gear (right-hand selection, or deselection)</td>
</tr>
</tbody>
</table>
R75
length at which 75 percent of the fish will no longer be vulnerable to the gear (right-hand selection, or deselection)

Rn
'goodness-of-fit' index of the ELEFAN I routine (=10^{ESP/ASP}/10)

Σ
summation sign

s.d.
standard deviation of variates, used as a measure of their dispersion

s.e.
standard error of a statistic

S
score function in Shepherd's method; also number of surviving fish after a certain time period

S1, S2
variables used for estimating the probability of capture under the logistic model

S_{max}
maximum score in Shepherd's method for a range of L, and K parameters

SF
selection factor

SI
separation index

SL
starting length; one of the two coordinates used to locate a growth curve in the ELEFAN I routine

SS
starting sample; the other coordinate used to locate a growth curve in the ELEFAN I routine. Jointly, SL and SS define the location of a pre-selected point of a growth curve

SSE
sum of squared errors; a measure of dispersal from the mean

t
given time or age (normally expressed in years); absolute age of a fish, for example as estimated from daily otolith rings; age corresponding to L

\( t' \)
relative age of a fish, defined as \( t' = t - t_0 \)

\( \bar{t} \)
mean of \( t_{\text{min}} \) and \( t_{\text{max}} \), used in Shepherd's method

\( t_c \)
mean age at first capture, corresponding to \( L_c \)

\( t_i \)
mean age at length i

\( t_m \)
age at marking, corresponding to \( L_m \)

\( t_{\text{max}} \)
mean age at which most fish reach maturity

\( t_{\text{max}} \)
longevity (in the wild); also oldest (calculated) age in a sample

\( t_{\text{min}} \)
youngest calculated age in a sample

\( t_{\text{tr}} \)
parameter of the VBGF, expressing the theoretical age at which the fish would have a length zero if they had always grown according to that equation \( (t_{\text{tr}} \text{ almost always takes non-zero, negative values, because small fish usually grow faster than predicted by the VBGF}) \)

\( t_{\text{r}} \)
mean age at which fish "recruit" to an exploited population, i.e. reach the fishing ground(s)

\( t_s \)
parameter of the seasonally oscillating version of the VBGF

\( t_{\text{Z}} \)
in Shepherd's method: origin of the VBGF in calendar time (expressed as fraction of a year); here replaced (without affecting other results) by a starting point, defined by SS and SL

\( T \)
mean annual habitat temperature, in °C

\( T_1 \)
average of test function in Shepherd's method for length

\( \text{TL} \)
total length

\( U \)
a term used in yield-per-recruit models

\( v_i \)
estimated value (e.g. $ \cdot \text{kg}^{-1} \) at length i

VBGF
von Bertalanffy growth function, in either original or seasonally oscillating form

VPA
virtual population analysis
\( \bar{w}_i \) mean weight of fish within a given length class \( i \)

\( W \) live weight of an individual fish (if not explicitly stated otherwise); or of a sample; mean weight of fish in catch samples, computed from \( W' \) upward

\( \bar{W} \) mean annual weight of fish in a population

\( W' \) pseudoweight of length–frequency samples

\( W_\infty \) asymptotic weight, i.e. the mean weight the fish of a given stock would reach if they were to grow forever (see \( L_\infty \))

\( W_{(e)} \) rough estimate of asymptotic weight, derived from \( W_{\text{max}} \)

\( W_{\text{max}} \) maximum weight reached by the fish of a given stock

\( WC \) total weight of the catch

\( WF \) weighting factor assigned to an observation

\( WP \) “winter point”; in the seasonalized VBGF, the time of the year when growth rate is slowest; equivalent to \( t_e + 0.5 \) year

\( WS \) total weight of the sample

\( W_i \) mean weight of fish at age \( t \)

\( x \) any variable (often used for the abscissa in 2-dimensional plots)

\( x^2 \) chi-square statistics

\( y \) any variable (often used for the ordinate in 2-dimensional plots)

\( \hat{y} \) predicted value of \( y \)

\( \bar{y} \) mean value of \( y \)

\( Y \) yield, catch in weight

\( Y/R \) yield per recruit

\( Y'/R \) relative yield per recruit

\( Y/R_{\text{max}} \) maximum yield per recruit achievable under a given fishing regime

\( Y_i \) yield at length \( i \)

\( Z \) instantaneous rate of total mortality