

# A manual of mathematical techniques for linefish assessment

---

D S Butterworth, A E Punt, D L Borchers, J B Pugh and G S Hughes

Incorporating a Report of the SANCOR Marine Linefish Programme's  
Workshop on Population Dynamics, 4 – 6 February 1987, Cape Town

**SOUTH AFRICAN NATIONAL SCIENTIFIC PROGRAMMES REPORT NO**

**160**

**1989**

Issued by

Foundation for Research Development (FRD)  
P O Box 395  
PRETORIA  
0001

from whom copies of reports in this series are available on request.

Printed in the Republic of South Africa by Scientia Printers  
of the Technical Support Services Group of the CSIR.

ISBN 0 7988 4513 9

Prepared by

D S Butterworth, A E Punt, J B Pugh and G S Hughes  
Department of Applied Mathematics  
University of Cape Town  
Private Bag  
RONDEBOSCH  
7700

and

D L Borchers  
IWC  
The Red House  
Station Road  
Histon  
CAMBRIDGE CB4 4NP  
United Kingdom

Cover: Linefish catch (Photograph by R P van der Elst)

## CONTENTS

	page
I. INTRODUCTION	1
II. FITTING CURVES TO DATA	3
Introduction	3
Which curve to fit and how to fit it	3
Underlying assumptions	4
Assessing the quality of the fit	7
III. PREDICTIVE LINEAR AND NONLINEAR REGRESSION ANALYSIS	9
Introduction	9
How many parameters to use?	11
Precision of parameter estimates	12
Accuracy of the model	13
IV. ANALYSIS OF CATCH AND EFFORT DATA	14
(1) Fishery Management in the Context of Surplus Production Models	14
Management objectives	14
What constitutes "overexploitation"?	14
Models for surplus production	16
Management methods	17
(2) Information Derivable from Catch-Effort Data	19
Data available for analysis	19
Detection of overexploitation	19
Problems associated with catch-effort analyses	22
Effort indices and beach patrols	25
V. ANALYSIS OF LENGTH COMPOSITION DATA	27
Introduction	27
Transformation of length distribution to age distribution	28
Virtual Population Analysis (VPA)	29
Modelling the catch-curve to find $F$ - an alternative analysis of age data	32
Yield-per-recruit analysis	36
The $F_{0.1}$ strategy	39

VI. STANDARD ANALYSES FOR LINEFISH DATA	41
CPUE Analysis	41
Yield-per-Recruit Analysis	42
VII. TAG-RECAPTURE METHODS	44
General aspects	44
Analysis of tag-recapture data for the Natal eel resource	45
VIII. HOOK SELECTIVITY IN A LINEFISHERY	48
IX. MULTI-SPECIES AND ENVIRONMENTAL ASPECTS	51
ACKNOWLEDGEMENTS	53
REFERENCES	54
APPENDICES	
APPENDIX I : Functional Regression Analysis	
II : Outliers and Influential Points	
III : The Runs Test for Serial Correlation	
IV : Predictive Linear Regression Formulae	
V : Schnute's Generalised Growth Model (Schnute 1981)	
VI : The Bootstrap Method of Variance Estimation	
VII : The Schaefer Model	
VIII : Proof of the Relationship between CPUE (CGD) and CND (Cooke 1984)	
IX : Using an Age-Length Key to Transform the Length-Frequency Distribution of the Catch into an Age-Frequency Distribution	
X : Derivation of the Beverton-Holt Yield-per-Recruit Equations	
XI : Derivation of the Relationship Between Mean Tag Return Time ( $t_P$ ) and Total Mortality, Z	
XII : Derivation of the Formula for N(gape)	

## ABSTRACT

This manual contains an expanded version of material presented at a Linefish Population Dynamics Workshop held in Cape Town in February 1987. First the statistical procedures for fitting curves to data are discussed. The standard single-species approaches to analysis of fisheries catch and effort data, and of length (or age) composition data, are then developed. Tag-recapture methods, hook selectivity, and multi-species aspects are discussed briefly. The topics are presented in the context of making scientifically based recommendations for management given limited data, as is usually the case for a recreational line fishery. A checklist of standard analysis is suggested, many of which may be implemented using the microcomputer package *PC-YIELD*. A number of the analysis techniques are illustrated by application to data for the Natal elf resource.

## UITTREKSEL

Hierdie handleiding bevat 'n uitgebreide weergawe van materiaal aangebied tydens 'n werksessie oor lynvisbevolkingsdinamika wat in Februarie 1987 in Kaapstad gehou is. Eerstens word die statistiese metodes vir krommepassing by data bespreek. Vervolgens word die standaard enkel-spesie benaderings vir die ontleding van vang-en-pogingdata, en van lengte-(of ouderdom-)samestelling ontwikkel. Merk-en-hervangmetodes, vishoekselektiwiteit, en multispesie-aspekte word kortliks bespreek. Die onderwerpe word aangebied in die konteks van die behoefte aan wetenskaplik gefundeerde bestuursaanbevelings gebaseer op beperkte inligting - gewoonlik die geval vir ontspanningslynvisserye. 'n Kontrolelys van standaardontledings word aanbeveel. Verskeie hiervan kan aangewend word deur gebruik te maak van die mikrorekenaarprogram *PC-YIELD*. 'n Aantal ontledingstegnieke word geïllustreer aan die hand van Natalse elfdata.

## 1. INTRODUCTION

The Linefish Population Dynamics Workshop held in February 1987 was convened chiefly to meet the need felt by biologists in the SANCOR Marine Linefish Programme for expert guidance in the analysis of catch-effort, length and age data. It also provided a useful forum for in-depth discussion of the practical problems in such analyses encountered by scientists in this field, both with mathematicians and with fellow biologists. Discussion was focused in the main on two important topics:

- (i) the theoretical background to fitting mathematical models to biological data sets for describing an underlying process, estimating the parameters of the models and using them to make predictions, and
- (ii) the application of particular mathematical models to catch statistics data with a view to fishery management for optimal resource utilisation.

This report is a formalisation of the discussions which it is hoped will provide a useful reference for the scientists concerned. To this end, it goes further than was possible within the time constraints of the Workshop itself, and includes aspects covered inadequately at the time or not at all, which are nonetheless relevant to the general problems of fisheries data analysis and of linefish management in particular.

A characteristic of linefisheries, particularly recreational fisheries, is that usually only limited scientific information is available. For example, this may be no more than a time series of catch rates (CPUE - catch per unit effort) and mean masses, for only a sector of the complete fishery (e.g. angling competitions). The suggestions for analysis that follow have been made with these limitations in mind. Partly for the same reason, the fish population dynamics models presented are all of the "single-species" type, ignoring (explicitly) that harvesting one species must in turn affect others with which it has biological interactions. This limitation is discussed further in Chapter IX, which deals with multi-species and environmental effects - potentially important (though as yet not well understood) aspects of linefishery management.

As a guideline intended to facilitate the interpretation and comparison of results, a checklist of standard analyses and useful methods for presentation of linefish data is provided in Chapter VI, based on recommendations agreed upon at

the Workshop. Many of these analyses (discussed in more depth in Chapters IV and V) may be accomplished using the program *PC-YIELD* (Hughes and Punt 1988), which has been specifically designed for fitting curves to age-length data and performing yield-per-recruit analysis on an IBM-compatible personal computer. For ease of reference, therefore, the analyses available in *PC-YIELD* are indicated in this checklist, as well as in the text, where relevant. Theoretical discussion of curve-fitting in general, and some associated practical suggestions, are given in Chapters II and III, while Chapters VII and VIII briefly cover the two special topics of tag-recapture methods and hook size analysis.

For the convenience of readers, the document incorporates most of the discussion contained in an earlier internal SANCOR document concerning the scientific information requirements for linefish management (Butterworth 1983).

## II. FITTING CURVES TO DATA

### Introduction

The fitting of a curve to data points is an attempt to describe mathematically an underlying biological process, and is a useful technique in the statistical analysis of biological data sets. Before attempting to fit a curve, however, two important questions should be asked. These are:

- (i) what type (shape) of curve can best be fitted to the particular data set, and
- (ii) what is the use to which the fitted curve is to be put, e.g. prediction, utilisation of the curve parameters estimated, or comparison with other curves.

Once these questions have been answered and the curve fitted, the adequacy of the fit should be evaluated and the model altered if necessary. A further important consideration is the precision of the resulting parameter estimates, as these are virtually meaningless without an associated measure of error. This is usually represented in the form of a standard error (s.e.) or coefficient of variation (c.v.), the latter being the ratio of the standard error to the mean, generally expressed as a percentage.

### Which curve to fit and how to fit it

It must be emphasized that the data points should always be plotted before attempting to fit a curve to them. Not only does this make for easier identification of the type of curve which should be fitted - it also facilitates identification of outliers and influential data points by visual inspection. In certain instances a particular type of curve may be chosen on the basis of prior knowledge of the underlying biological process. The Von Bertalanffy curve, for example, is derived from a differential equation which reflects the anabolic and catabolic contributions to the growth process in animals.

The actual fitting procedure depends on the use to which the curve is to be put. If, for example, one wishes to predict the dependent variable  $Y$  from a variable  $X$ , where  $X$  is known without error, a suitable technique would involve minimisation of the sum of the squared vertical distances along the  $Y$ -axis of the data points from the fitted curve. The quantity to be minimised is then:



$$\sum_{i=1}^n (y_i - \hat{y}_i)^2$$

where  $\hat{y}_i$  is the expected value of Y (i.e. the value given by the curve) for each value of X,  $x_i$ . This is known as predictive least squares regression. If, rather than prediction however, estimation of the parameters of the relationship between X and Y is required, where each variable has an associated observation error, a more appropriate technique would involve minimisation of some combination of both vertical and horizontal deviations of the data points from the curve. This is known as functional regression analysis (see Appendix I).

### Underlying assumptions

Any process of model-fitting implies certain assumptions regarding the distribution of the data points about the resulting curve. For the least squares regression procedure above, these are:

- (i) the errors or "residuals"  $e_i = (y_i - \hat{y}_i)$  are normally distributed about a mean of zero, with variance  $\sigma_i^2$ , written  $e_i \sim N(0, \sigma_i^2)$
- (ii)  $\sigma_i^2 = \sigma^2$ , i.e. the variance of the residuals is the same for all points  $i$  (homoscedasticity)
- (iii) the residuals  $e_i$  are independent of each other, i.e. they contain no systematic trends.

These assumptions are important and their validity should be confirmed before continuing, for the following reasons:

- (i) normality is required so that asymptotically (i.e. in the limit of a large number of data points), the parameter estimates are minimum-variance estimates and are themselves each normally distributed, so that normal distribution theory may be used to estimate confidence intervals
- (ii) homoscedasticity is required to obtain the "best" estimate  $\hat{y}_i$  for any given  $x_i$ , i.e. the estimate having the smallest variance
- (iii) the presence of systematic trends in the residuals indicates that the chosen model is a poor representation of the relationship between the two variables X and Y, i.e. that additional and unexplained non-random variation is still present - the fit is poor.

### - Checking for outliers and influential observations

Outliers are easily detected in a plot of  $e_i$  versus  $x_i$ . Strictly, these are observations which are situated three or more standard deviations ( $> 3\sigma$ ) from the fitted curve and are considered atypical, warranting further investigation. (Note that identification of outliers using this criterion is not entirely straightforward, as  $\sigma$  is not known but has to be estimated from the data set which includes the suspected outlier point.) While their existence may often be explained in terms of sampling or experimental error, they may alternatively represent genuine results which can provide additional information about the process being modelled. Rejection of an outlier is ultimately a subjective decision which should not be made without reasonable certainty that the suspect point is, in fact, an error.

Influential observations are data points whose presence or absence makes a critical difference to the parameter estimates for the fitted model; such observations are more difficult to identify than outliers. They are often not detectable from residual plots and may or may not be outliers. This is illustrated in Appendix II. A number of statistics (e.g. Cook's distance) have been devised to test for the existence of influential observations in the case of linear regression [see Draper and Smith (1966)]. A crude preliminary test can be made by deleting possible suspects in turn and noting the effect on estimated parameters and on conclusions drawn from the model. The reason it is important to identify influential points is that their existence indicates that the fitted model is open to question, and more data in the neighbourhood of the influential points are desirable.

### - Checking for normality of the residuals

A simple visual test for normality provided by many statistical computer packages is the normal probability plot; if the residuals are normally distributed, this plot should show a strongly linear trend. The same result can be achieved manually by plotting the residuals on normal probability paper. Deviations from normality of the residuals are, in fact, difficult to detect unless the data set is very large; in practice, however, these are generally less crucial to the quality of the fit than is the presence of a trend or of non-homoscedasticity in the residuals.

### - Checking for homoscedasticity of the residuals

Plotting  $e_i$  against  $\hat{y}_i$  is a useful technique here, as illustrated in Figure 1. [ $\hat{y}_i$  and not  $y_i$  should be used, as  $e_i$  and  $y_i$  are usually correlated (Draper and Smith 1966).] If the residuals have uniform variance, points on this plot will be distributed in a horizontal band, showing no trend. If a trend is evident, however, a transformation may be required to produce homoscedasticity, e.g. if the residual variance increases with  $\hat{y}_i$  as in Figure 1(a), a log transformation may be warranted, i.e. for a Von Bertalanffy growth curve, for example, fit:

$$\ln y_i \text{ to } \ln[L_{\infty}(1 - e^{-\kappa(t_i - t_0)})]$$

Instead of:

$$y_i \text{ to } L_{\infty}(1 - e^{-\kappa(t_i - t_0)}).$$

In Figure 1(b) the effect of this transformation on the residual plot can be seen. Also shown are curves representing  $\pm 2$  standard errors (s.e.'s) of the growth curve fit to the mean length-at-age for both the transformed and untransformed data. These standard errors were calculated using the Jack-knife procedure (see Chapter III) and correspond to approximate 95% confidence intervals for the fitted curve. Note that these confidence intervals are generally narrower for the fit to the transformed data, particularly for the smaller ages, illustrating the improvement in precision obtained for homoscedastic residuals.

This improvement is also evident for the estimates of the curve parameters. For the untransformed data of Figure 1(a), for example,  $\hat{\kappa} = 0.104$  with s.e. = 0.022, while for the transformed data of Figure 1(b),  $\hat{\kappa} = 0.112$  with s.e. = 0.016. Visual inspection is often sufficient to determine the type of transformation to use, which depends on the nature of the trend. After fitting the transformed data, the residuals should be plotted again to ensure that the trend has indeed been removed, and additional transformations made if necessary. A statistical test for the presence of a trend should be used - a linear trend, for example, can be detected by testing for a significantly nonzero slope in a regression of  $|e_i|$  against  $\hat{y}_i$ . A straight line fit to the moduli of the residuals shown in Figure 1(a) has a slope which is significantly different from zero at the 5% level ( $F_{48,2}^{0.05} = 9.549$ ;  $P > 0.01$ ), justifying the application of a transformation to the data. The slope of the fit to the moduli of the residuals after transformation [Fig. 1(b)] is no longer significant ( $F_{48,2}^{0.05} = 1.251$ ;  $P < 0.05$ ).

(a)

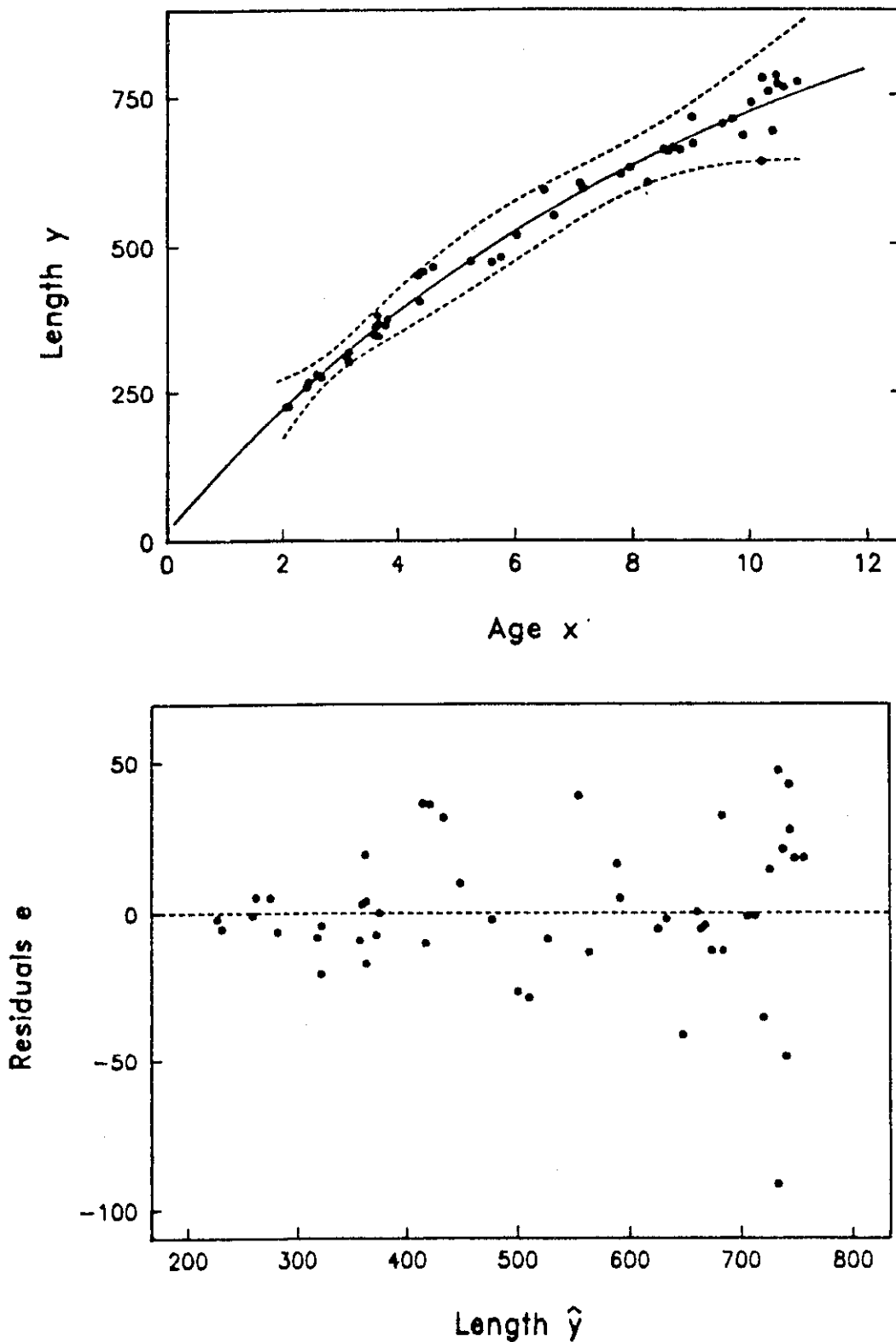
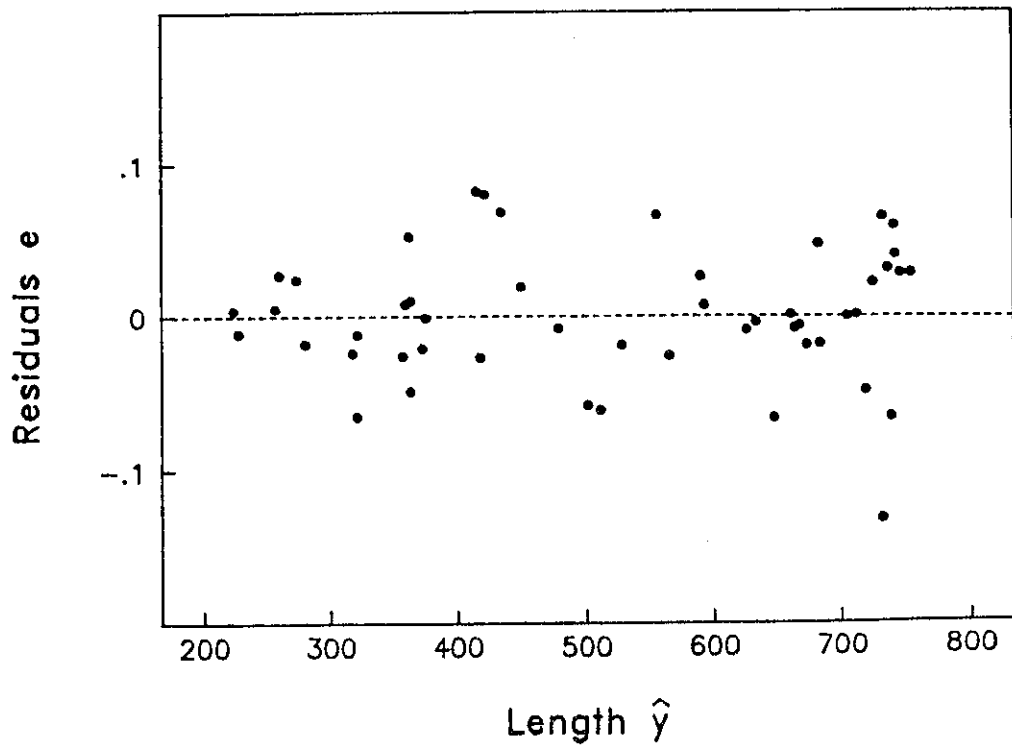
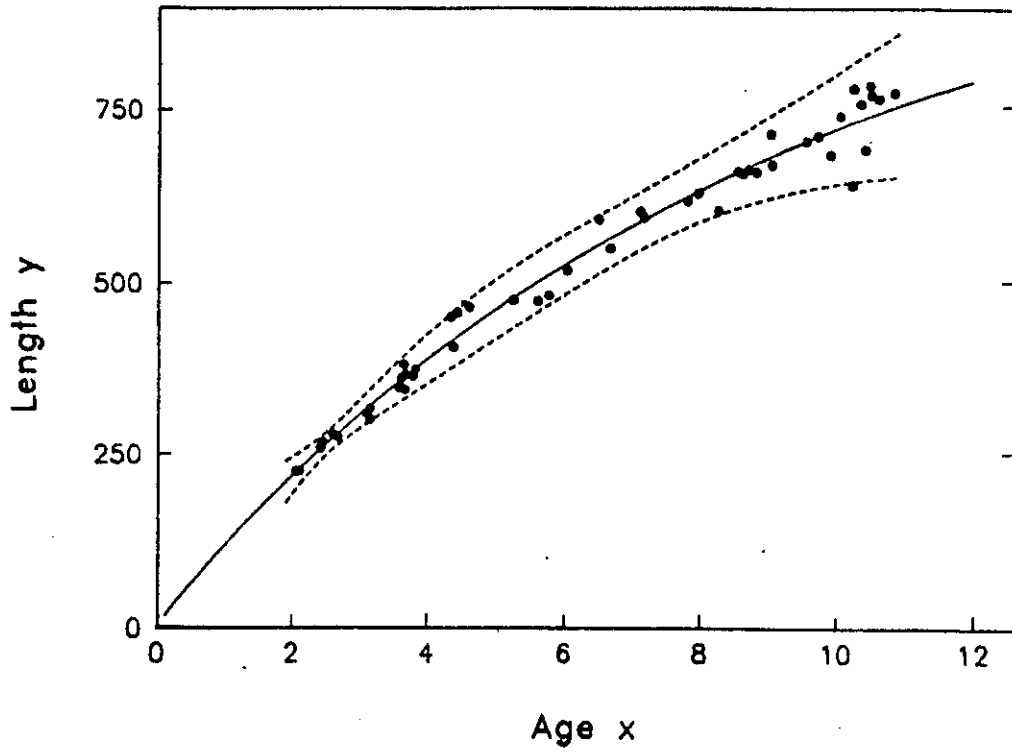


Figure 1:

A Von Bertalanffy growth curve fitted to age-length data, and its associated residual plot (a) before transformation of  $y$  and (b) after log transformation. Dashed lines on the curve plots indicate approximate 95% confidence limits about the curves. Note the absence of trend in the typical size of the residuals for case (b), indicating that homoscedasticity has been achieved.

(b)



(a)

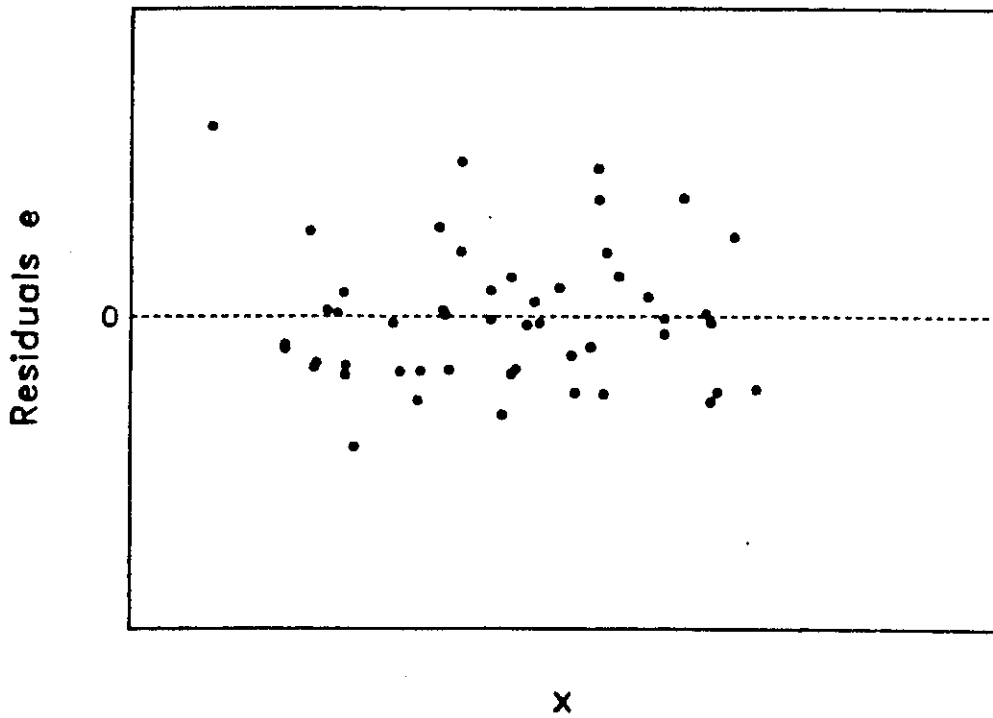
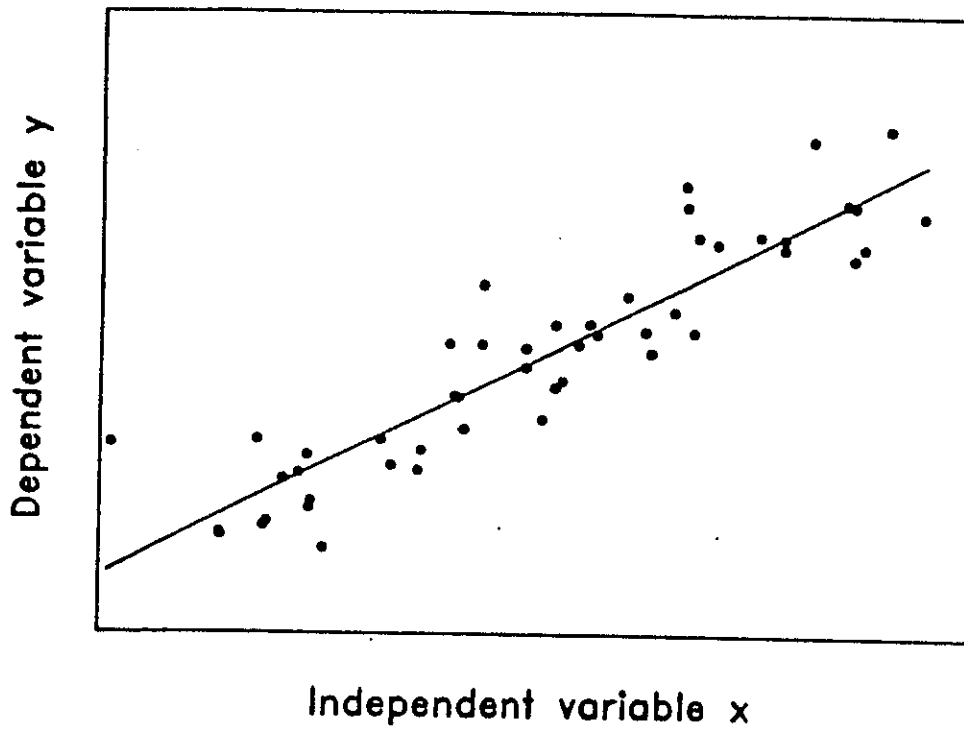
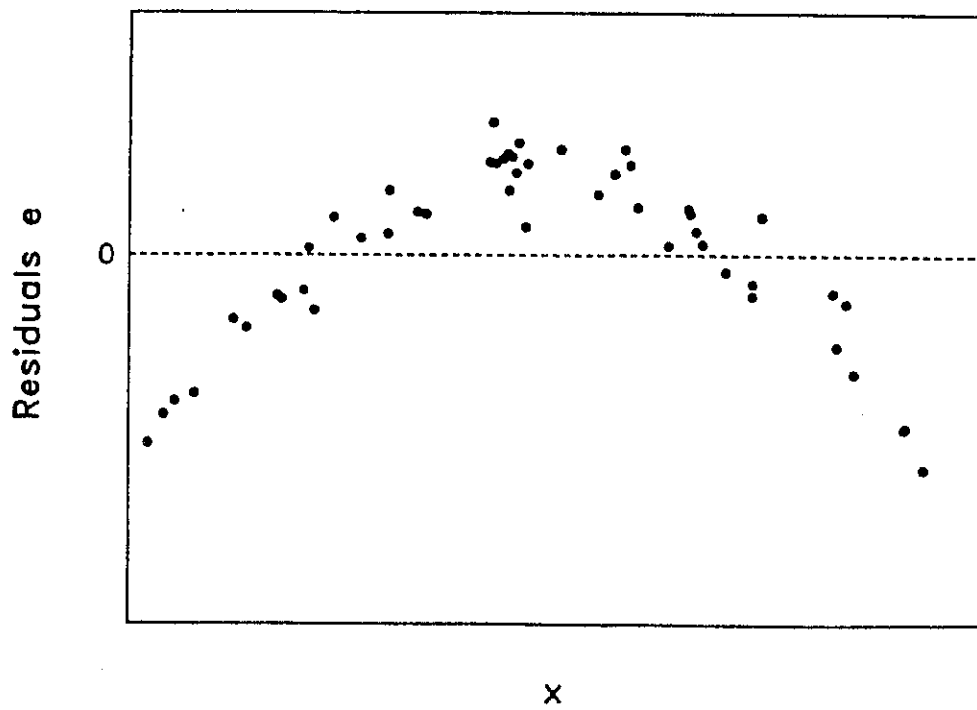
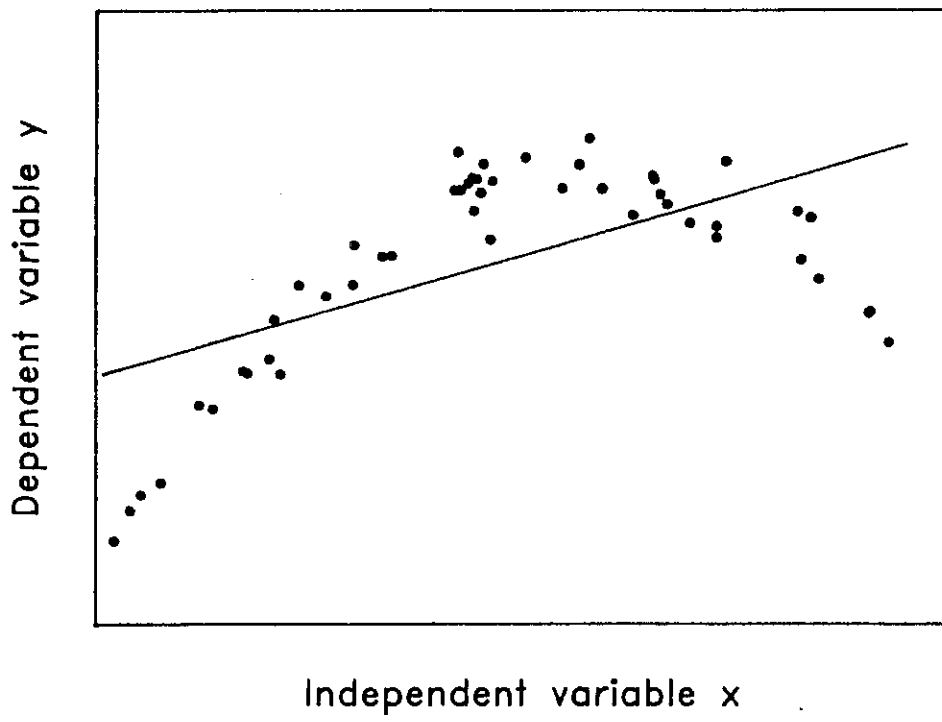


Figure 2: Linear model fits to data, and their associated residual plots. Note the clear systematic trend in the latter for case (b).

(b)



If the relationship between  $X$  and  $Y$  appears to be linear, however, homoscedasticity is often more easily achieved by the use of a weighted least-squares method of fitting the curve, rather than by means of a transformation which will introduce non-linearity. In this case, each residual is weighted by the inverse of the estimated variance  $\hat{\sigma}_i^2$  at the corresponding point  $x_i$ , assuming that there are enough replicate observations at each  $x_i$  to provide estimates of  $\sigma_i^2$  which are reasonably precise.

#### - Checking for Independence of the residuals

The presence of a systematic trend in the residuals can be identified quantitatively using a statistical test for randomness such as the One-Sample Runs Test (Appendix III) or the Durbin-Watson Test [see Draper and Smith (1966)], although visual inspection of a plot of  $e_i$  against  $x_i$  is often sufficient (Fig. 2). The nature of the trend should give an indication of how the model may be improved to incorporate the non-random variation - by including higher-order terms, for example, or additional predictor variables. The obvious trend in the residuals plotted in Figure 2(b), for instance, suggests that a convex function may be more appropriate than a linear one in this case; in contrast, the lack of trend in the residuals of Figure 2(a) indicates that the linear form is adequate there.

#### Assessing the quality of the fit

A "better" fit in terms of a smaller sum of squared errors does not necessarily indicate that the process (e.g. growth) is being more accurately described. For example, it is always possible to fit a polynomial of  $(n-1)$ th degree to  $n$  data points (having distinct  $x$  values), producing an error sum of squares of zero. This does not imply perfect characterisation of the process, as any new data point will be extremely unlikely to fall exactly on the fitted curve. This is a case of under-determination, i.e. the polynomial model has as many parameters as there are data points, and the data will not be sufficiently precise to warrant fitting a model with that degree of complexity.

Depending on the reason for fitting the curve, one may choose to fit it to a certain subset of the data only. If, for example, inferences about fish growth are required after a certain age  $t$  and not before, as in yield-per-recruit analyses (see Chapter V) for age-at-first-capture,  $t_c$ , there may be some justification for fitting only the relevant section of the data. In this



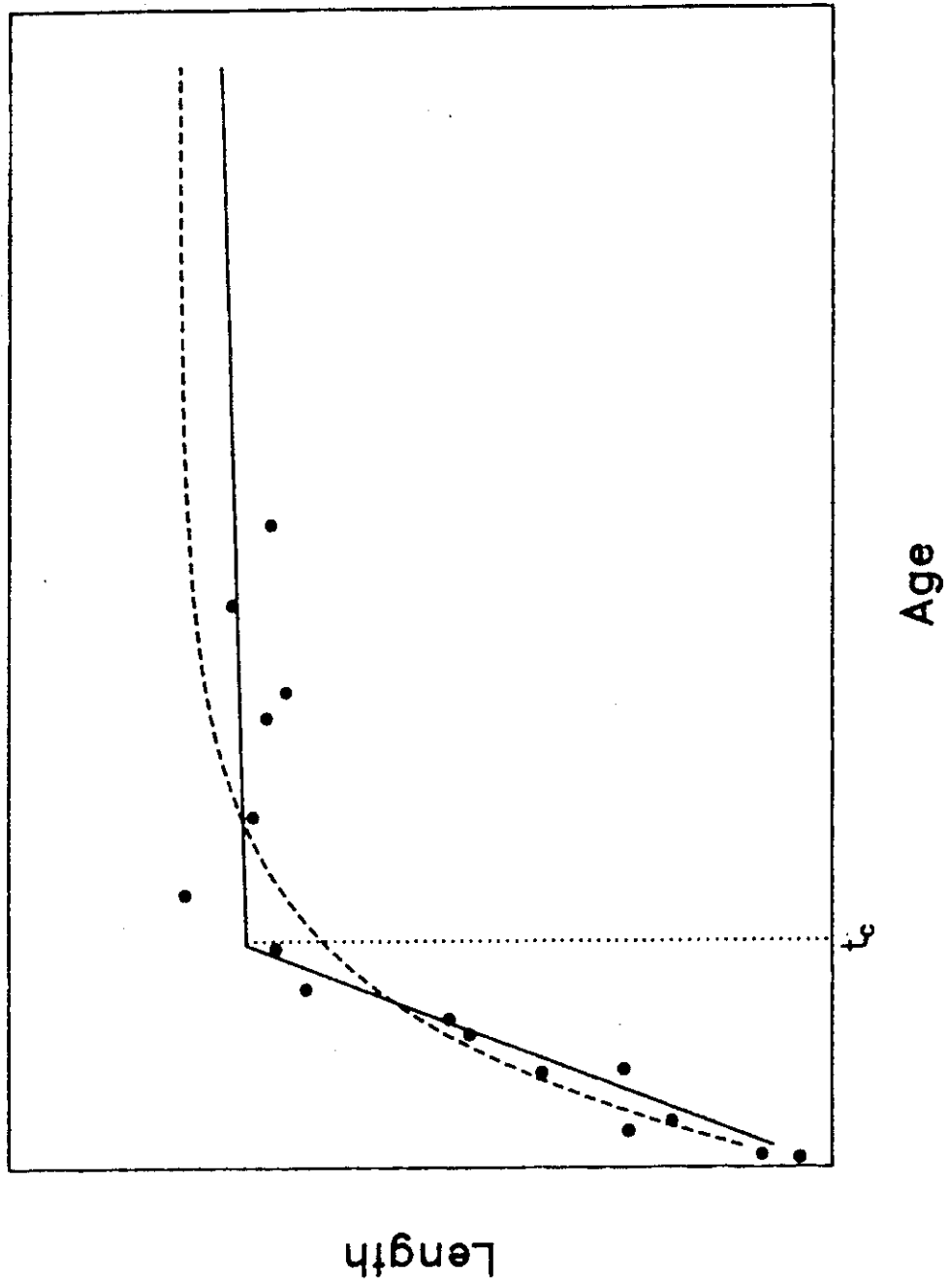
situation, however, useful information about the growth process may be discarded, resulting in parameter estimates with higher variances. On the other hand, if distinct processes appear *a priori* to be operating over different ranges of the data set, fitting a different curve to each is sensible; the points at which such discontinuities occur should be specified beforehand, however. Figure 3 illustrates a situation in which age-length data produce a poor fit to a Von Bertalanffy curve, while the straight lines shown may legitimately describe two distinct processes operating at different stages of the growth of the animal, provided there is reason to believe *a priori* that age  $X$  represents the discontinuity between them. Such a discontinuity could be related, for example, to the animal reaching maturity, after which some energy intake has to be allocated to reproduction, leaving less available for growth.

In summary, the following procedure is suggested when testing for adequacy of fit, the order of the various tests corresponding to their relative importance in most cases:

- (1) plot the data,  $(x_i, y_i)$ , the fitted curve, and the residuals - both  $(x_i, e_i)$  and  $(\hat{y}_i, e_i)$
- (2) check for systematic trends in the residuals; if these exist, the model requires modification
- (3) check for outliers and influential observations and decide whether or not there is a valid basis for exclusion of any of the former
- (4) check for homoscedasticity of the residuals and, if necessary, transform or weight the data accordingly before re-fitting the curve
- (5) check for normality of the residuals; be wary of estimating confidence intervals for the parameter estimates and of conducting hypothesis tests based on normal distribution theory, if this normality assumption is rejected.

The *PC-YIELD* program (Hughes and Punt 1988) provides useful aids in choosing and fitting an appropriate growth model, incorporating the procedure described above with the exception of the test for outliers and influential observations in (3). Practical examples of curve-fitting, with particular application to the Von Bertalanffy growth model, are also discussed by Hughes (1986).

Figure 3: A Von Bertalanffy curve (dotted line) fitted to age-length data, with the straight lines representing distinct growth processes operating over different age ranges - above and below age  $t_c$ .



### III. PREDICTIVE LINEAR AND NONLINEAR REGRESSION ANALYSIS

#### Introduction

The linear ("straight line") model is represented by the equation:

$$y_i = a + bx_i + e_i \quad (3.1)$$

where the residuals  $e_i$  meet the assumptions described in Chapter II (i.e.  $e_i \sim N(0, \sigma^2)$ ), and the least squares method of fitting this model requires minimisation of the quantity:

$$\sum_{i=1}^n [y_i - (a + bx_i)]^2 \quad (3.2)$$

which is often referred to as the "sum of squares" (or SS) function. Formulae for the estimation of the parameters and calculation of their associated variances and confidence intervals are given in Appendix IV. While data transformations may produce a linear model from an essentially nonlinear relationship, they do not always yield an error structure which meets the required assumptions. In such cases, and others where no simple transformation to a linear form is possible, a nonlinear minimisation procedure must be used to fit a curve to the data. The nonlinear model is represented by the equation:

$$y_i = f(x_i; p_1, \dots, p_m) + e_i \quad i = 1, \dots, n \quad (3.3)$$

where  $e_i \sim N(0, \sigma^2)$  as before, and the function  $f(\cdot)$  is nonlinear in the parameters  $p_1, \dots, p_m$ . The quantity to be minimised is then:

$$SS = \sum_{i=1}^n [y_i - f(x_i; p_1, \dots, p_m)]^2. \quad (3.4)$$

Generally, no simple formulae exist for the estimation of the parameters of such a model, in contrast to the linear case (Appendix IV). However, computer packages are available which perform the minimisation iteratively [using, for example, the method of Nelder and Mead (1965) or Powell (1964)], starting from an initial guess at the position of the minimum and successively improving on this until (hopefully) the true overall (global) minimum is reached. The disadvantages of these methods include:

- (i) the amount of computing time required, which can be prohibitive, depending on the closeness of the starting point to the final minimum (if it is found), the number of data points and the complexity of the model,
- (ii) the existence of local minima towards which the procedure may converge, and
- (iii) cases of SS surfaces which are "flat" in the vicinity of the minimum SS value, for which the iterative search procedure implemented by the computer package may terminate at parameter values some distance from those corresponding to this actual minimum (this occurs when the available data contain inadequate information (or "contrast") to provide precise estimates of all the model parameters).

The importance of the initial guess is thus clear, and avoidance of these problems may demand *a priori* knowledge of the vicinity of the global minimum. [For certain models it may be possible to rewrite or transform the equation to a form which is linear in the parameters. The parameter estimates corresponding to this linear form are then readily obtained using the equations of Appendix IV, appropriately expanded if more than two parameters are involved (Draper and Smith 1966, pp. 85-96); these estimates provide the initial guesses for the full non-linear minimisation process. In the special case of a Von Bertalanffy curve, the traditional "eye-fitting" method (Beverton 1954) can be used to fit the data roughly and obtain an initial guess for the parameter estimates.]

Identification of a minimum as global or local is often difficult and in most cases intuitive. It is advisable to repeat the minimisation process several times from different starting points, noting whether convergence to the same point occurs. If so, this can be regarded as the global minimum with reasonable confidence. If no satisfactory minimum point is found at all, the curve may have more parameters than can be estimated reliably from the data. Note that this is not the same as under-determination; rather it means that there is insufficient "contrast" in the data set for the effects of different parameters to be distinguished. Nonconvergence due to too many indistinguishable parameters is a common problem in fitting nonlinear curves to growth data. When the data lie essentially along a straight line, for example, the iteration procedure may have difficulty estimating a parameter describing the curvature. A case in point here is the Von Bertalanffy growth curve:

$$l_t = l_{\infty}[1 - e^{-\kappa(t - t_0)}] \quad (3.5)$$

where  $l_t$  = length at age  $t$   
 $l_{\infty}$  = asymptotic length  
 $\kappa$  = growth rate parameter  
 $t_0$  = age at zero length .

Now, for  $[\kappa(t - t_0)] \ll 1$  this equation can be approximated as follows:

$$\begin{aligned} l_t &= l_{\infty}[1 - (1 - \kappa(t - t_0))] \\ &= l_{\infty}\kappa t - l_{\infty}\kappa t_0 \\ &= \alpha t - \beta \end{aligned}$$

where  $\alpha = l_{\infty}\kappa$  and  $\beta = l_{\infty}\kappa t_0$  - i.e. by a straight line. For growth data which show little deviation from a straight line,  $\alpha$  and  $\beta$  and hence  $t_0$  will be readily determined. However, a nonlinear minimisation procedure using the exact Von Bertalanffy equation will have difficulty distinguishing between  $l_{\infty}$  and  $\kappa$ , as the data only provide information about their product ( $l_{\infty}\kappa$ ). It may therefore have trouble finding a minimum at all, or may even converge to a somewhat arbitrary point. These last two problems also occur in situations where the computational truncation errors "swamp" the small differences in the sum of squares function as parameter values are changed in the minimisation process.

How many parameters to use?

The task of choosing an appropriate model to describe length-at-age data from the extensive menu documented in the literature is not a simple one. Even once a particular model has been chosen, numerous methods of parameter estimation exist and problems such as nonconvergence may also occur. Schnute (1981) provides a new, comprehensive growth model which incorporates most of the curves commonly used in growth studies as special cases (Appendix V), and which avoids many of the problems of fitting the standard curves in their familiar parameterisations.

Schnute's set of curves has four statistically stable parameters:  $a$ ,  $b$ ,  $l(t_1)$  and  $l(t_2)$ . (In contrast, the parameter  $l_{\infty}$  in the Von Bertalanffy equation is statistically unstable, as it is not well determined and may drift to very large values in a minimisation procedure for growth data showing little curvature.) Parameters  $l(t_1)$  and  $l(t_2)$  are well-defined, representing the length of a fish at

two different, prespecified (and preferably well separated) times (ages) for which data are available, and good initial guesses are thus not a problem. The parameters  $a$  and  $b$  determine the basic shape of the curve [see Figure A5.1(a)], and the values obtained by fitting Schnute's equations provide insight into the type of curve (perhaps one of the standard curves) that best fits the data, and thus the number of parameters required. The confidence intervals for these last two parameters may also give an indication of the type of curve to use. These can be calculated from the variance estimates for  $a$  and  $b$  obtained by using the "bootstrap" or "jack-knife" re-sampling techniques, discussed below.

### Precision of parameter estimates

In the case of linear regression, and where the assumptions of normality, homoscedasticity and independence of errors are met, the parameter estimates have the Student's  $t$  distribution, so that if their variances are known, exact confidence intervals for them can be calculated and significance tests performed based on this distribution [see Zar (1974), pp. 205-213]. In the case of nonlinear models, on the other hand, it is seldom possible to calculate exact variance estimates for the parameters analytically. These can be acquired quickly and easily, however, using the jack-knife re-sampling method, which gives the standard error (s.e.) of a parameter  $p$  as follows:

$$(\text{s.e.})^2 = [(n-1)/n] \sum_{j=1}^n [p(j) - p(\cdot)]^2 \quad (3.6)$$

where  $p(j)$  is the estimate of  $p$  obtained when the  $j$ th data point is omitted, and  $p(\cdot)$  is the mean value of the  $p(j)$ 's. If it is assumed that the estimates of  $p$  have (asymptotically) a normal distribution, an approximate confidence interval for  $p$  can be calculated and tests of significance performed, based on normal distribution theory. (The validity of the assumption of normality should strictly be tested by simulation; for reasonably large sample sizes, however, the results should be adequate for most biological purposes.) Although easy to use, it has been noted (Efron 1981) that the jack-knife procedure tends to produce s.e. estimates which are larger than their "true" values in certain situations. The bootstrap method (see Appendix VI) is likely to provide more precise and less biased estimates. Both of these methods of variance estimation are available in *PC-YIELD* (Hughes and Punt 1988).

### Accuracy of the model

The quality of the fit of a model is determined by the amount of systematic error it contains; in general, the "discrepancy due to approximation", i.e. approximation of the real process by a simple equation, decreases as the number of parameters increases. On the other hand, as the number of parameters describing a given data set increases, so too do the sizes of the standard errors of the estimates of these parameters, i.e. the "discrepancy due to estimation". The "total discrepancy" is the "sum" of the discrepancies due to estimation and approximation and is a measure of the overall accuracy of a model - how well it represents, and can be used to predict, reality (Linhart and Zucchini 1986). The "best" model is therefore the one having the smallest "total discrepancy", as illustrated graphically in Figure 4.

As a general rule in marine fisheries studies, the smallest "total discrepancy" is usually found in models with two to four parameters only [see comments by Schnute (1985)]. Increasing the sample size tends to shift the "discrepancy due to estimation" to the right in Figure 4, allowing more parameters to be estimated for a smaller "total discrepancy".

#### IV. ANALYSIS OF CATCH AND EFFORT DATA

##### (1) Fishery Management in the Context of Surplus Production Models

###### Management objectives

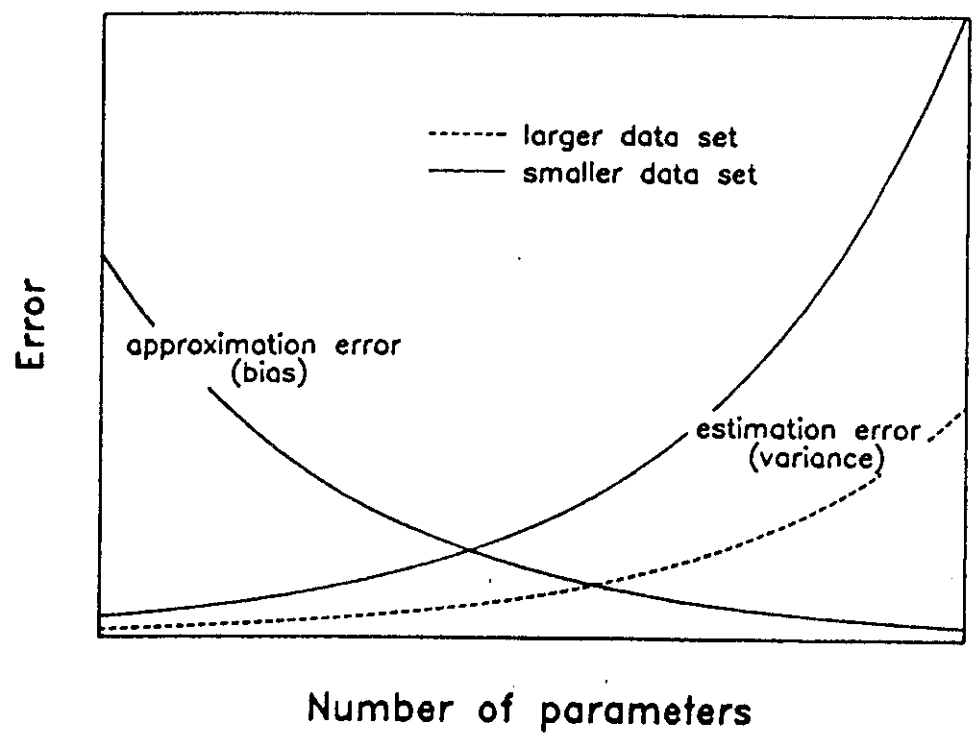
Management of a line fishery is generally directed towards the avoidance of overexploitation of the stock by fishermen with otherwise unlimited access to the fishery, and at the same time towards ensuring sustained maximal utilisation of the resource. The interpretation of "optimal" resource management, however, depends on the fishery concerned and the different groups exploiting it, as their objectives may be in conflict. For example, while the goal of a commercial fishery may be to maximise yield in terms of the mass of fish landed, a recreational fisherman is usually more concerned to ensure high catch rates and the continued existence of the larger size classes of a species (even if his fishing is restricted to a short season to achieve this).

###### What constitutes "overexploitation"?

The term "overexploited" is often used rather loosely, without a clear definition of what is meant. A stock is usually considered to be biologically overexploited if its biomass drops below that level (MSYL) at which the maximum sustainable yield (MSY) is achieved, where the sustainable annual yield at any given biomass level is the catch which will allow the biomass to return to that level after one year; this is equated in most modelling calculations to the increase in the population from that level over a one year period in the absence of fishing. The specification of MSYL for a given population, however, is rather arbitrary in the sense that it depends on the particular model used to describe its dynamics, and available data are usually inadequate to distinguish between models with quite different MSYL's. The Schaefer model (Schaefer 1954, 1957), for example, specifies the MSYL as a population size of 50% of the environmental carrying capacity,  $K$ , which is the average unexploited stock size. (Note that the word "average" is used because stock levels will tend to fluctuate even in the absence of human exploitation, due to environmental variability.) Although arguments exist which suggest that MSYL may be lower than this for the smaller sized pelagic



(a)



(b)

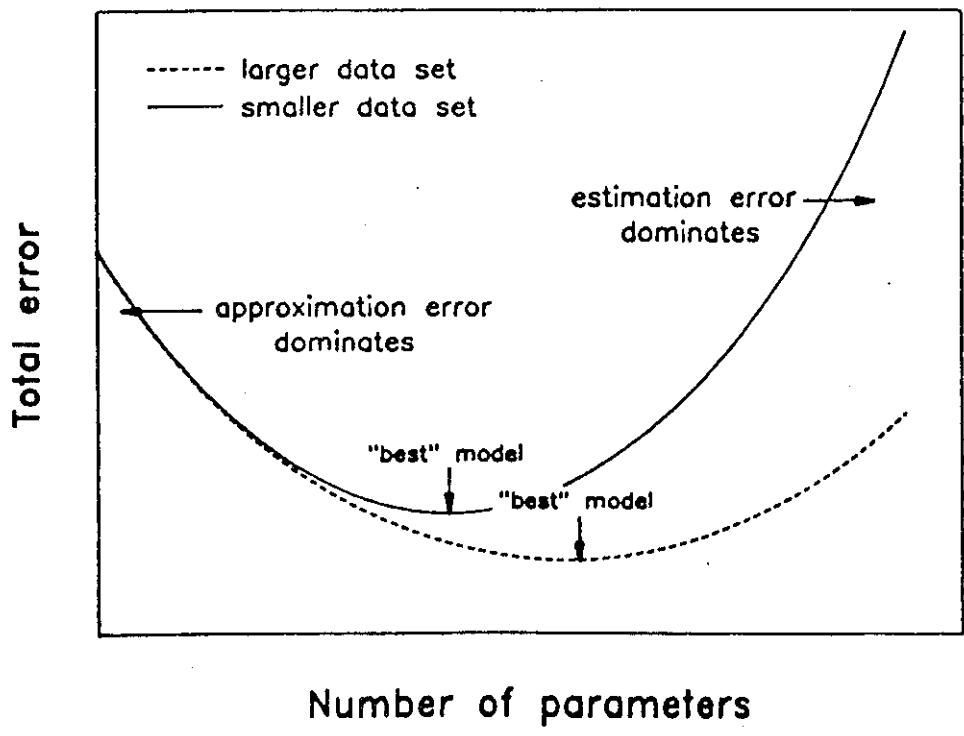


Figure 4: Behaviour of the curves representing (a) the error due to estimation and approximation and (b) the total amount of error, as the number of model parameters increases, for different sized data sets.

species (MacCall 1980), in general it is probably wiser to act conservatively and adopt an MSYL of  $0.5K$  as the minimum safe biomass level.

In a mixed species fishery one ideally would like each species harvested to be maintained at its own MSYL, but (even ignoring multi-species interaction complications) if fishing effort cannot be precisely species directed, problems arise. The level of effort required to harvest (at MSYL) a species that is difficult to catch (low catchability coefficient  $q$  - see equation 4.2.1), may constitute substantial overexploitation of the more catchable species. However, managing the overall fishery on the basis of keeping the species with the highest catchability coefficient at MSYL may be unrealistically restrictive, and wasteful of the other resources. It may therefore be necessary to adopt a management strategy whereby some species will be overexploited. Other conservation measures may be possible in such circumstances; for example, the creation of reserves could be considered if these highly catchable species are non-migratory.

A different approach towards preventing biological overexploitation is the use of the fishing mortality  $F$  [the ratio of the (instantaneous) catch rate to the population size] as a measure of the degree of exploitation. In terms of the yield-per-recruit model discussed in Chapter V, for example,  $F_{MSYR}$  (sometimes denoted  $F_{max}$ ) represents the fishing mortality at which the Maximum Sustainable Yield per Recruit is obtained.  $F_{0.1}$  is used alternatively as a target fishing mortality level, particularly if  $F_{MSYR}$  is infinite, and is defined in Chapter V. Biological overexploitation is considered to have occurred if  $F$  exceeds whichever of  $F_{MSYR}$  or  $F_{0.1}$  has been chosen as the target level. Maintaining  $F$  at a level less than the natural mortality rate  $M$  of the stock has also been suggested as a rule of thumb for managing fisheries (Gulland 1971), although a value of  $F$  less than  $0.3M$  has been proposed as being more appropriate following the analyses of Beddington and Cooke (1983), which take recruitment fluctuations into account. In contrast, however, Caddy and Csirke (1983) list assessments of a number of stocks for which the value of  $F$  corresponding to MSY is estimated to be larger than  $M$ .

This discussion has thus far referred to "biological overexploitation" to draw a distinction from "economic overexploitation". A resource is economically overexploited if it is depleted to a level below which it provides no economic rent - essentially when catch rates have dropped to the extent that the costs of fishing (including normal salaries) balance the revenue to be derived from sale of

the catch, so that the fishermen no longer make a profit. The biomass level at which economic overexploitation occurs depends on economic parameters (prices and costs); it may be either above or below the biologically preferred MSYL, and it will also change with time. Economic overexploitation is a concept more pertinent to a commercial fishery, however, than to a line fishery with a predominantly recreational character, and therefore further references to overexploitation in the following will imply biological overexploitation only.

### Models for surplus production

Models which attempt to provide a mathematical formulation of the net growth rate of a population as a function of its biomass are termed surplus production models. Note that these models assume that the single quantity biomass is the principal determinant of surplus production, and that the effects of other factors such as the population's age- and sex-structure can be ignored. All such models must sensibly respect three conditions:

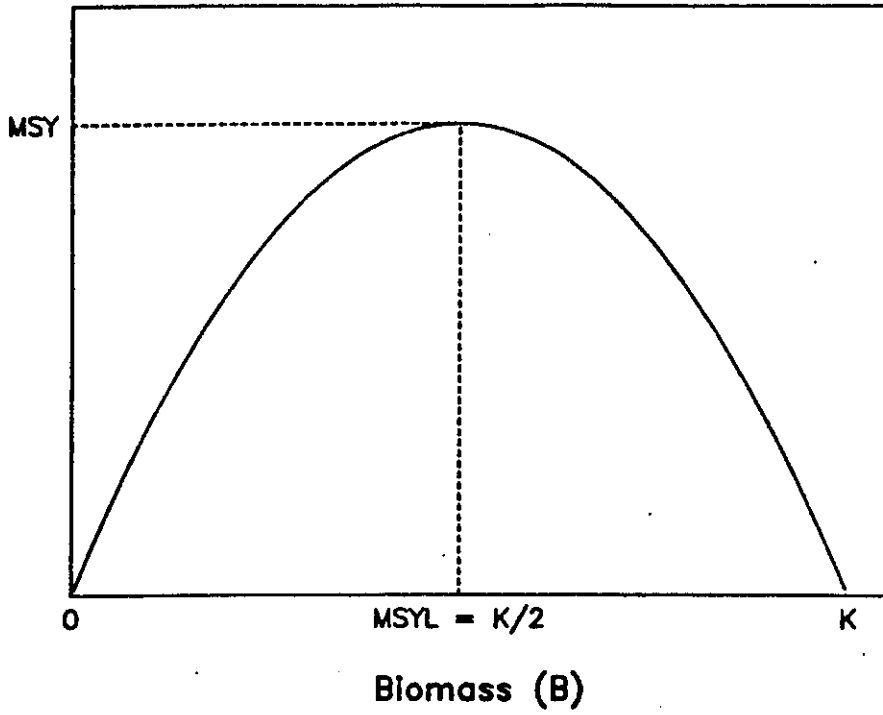
- i) the growth rate at zero biomass is zero,
- ii) the growth rate at carrying capacity ( $K$ ) is zero, and
- iii) the growth rate is positive for some part of the  $(0, K)$  biomass domain.

An immediate implication of this is that MSY (corresponding to the maximal growth rate) must occur at an MSYL less than  $K$ . In fact, any sustainable harvest regime requires a biomass level less than  $K$ , and hence also leads to a catch rate lower than in the pristine (unexploited) situation. Under harvesting, a certain level of stock decline (as far as, say, MSYL in terms of the discussion above) is inevitable and not undesirable.

An example of a surplus production model is the Schaefer model (Appendix VII and Fig. 5), for which MSY occurs at 50% of the carrying capacity. In this particular model, the relative (or *per capita*) net growth rate is always positive and a decreasing function of the biomass. This is referred to as "pure compensation", as the population always responds in a stabilising manner, increasing *per capita* net growth rate if the biomass declines and *vice versa*. This may not be a realistic model for all populations. Figures 6 and 7 represent two surplus production models which incorporate different compensatory effects over some part of the  $(0, K)$  domain. In Figure 6 the phenomenon of "critical depensation" is illustrated, in which the stock, once reduced below a critical

(a)

Equilibrium yield = net growth rate ( $\frac{dB}{dt}$ )



(b)

Per capita net growth rate ( $\frac{1}{B} \frac{dB}{dt}$ )

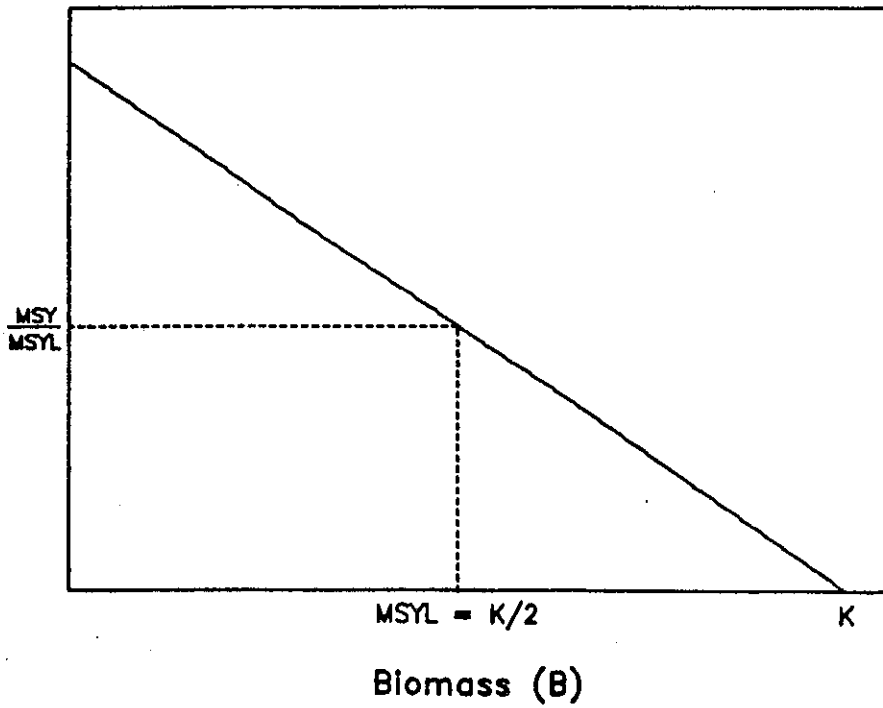
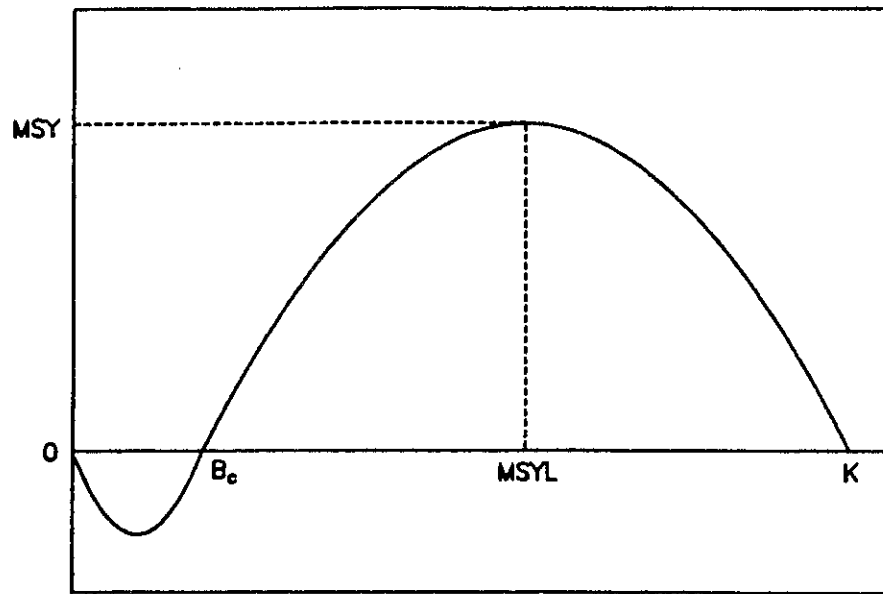


Figure 5: The Schaefer model for surplus production. Note that the relative net growth rate (b) is always a decreasing function of the biomass ("pure compensation").

(a)

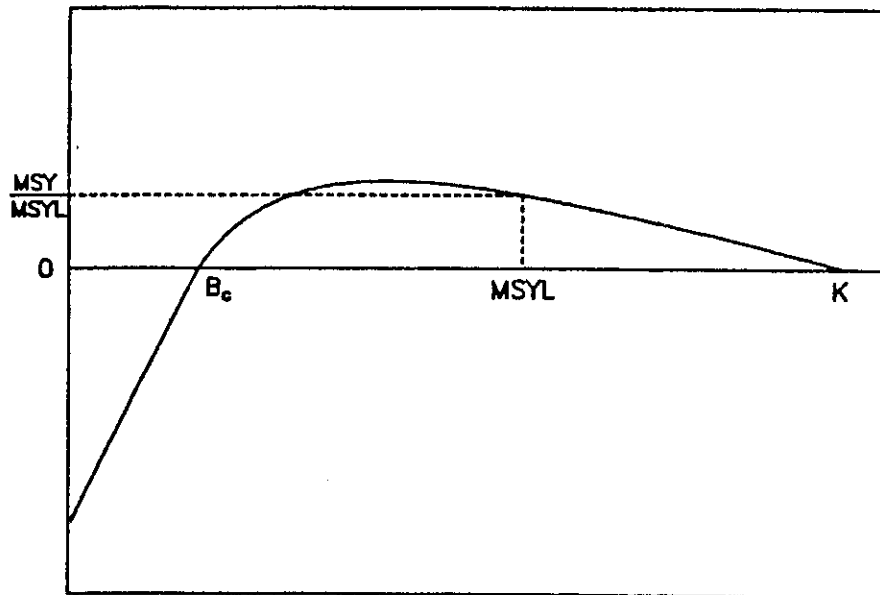
Equilibrium yield = net growth rate ( $dB/dt$ )



Biomass (B)

(b)

Per capita net growth rate ( $\frac{1}{B} \frac{dB}{dt}$ )

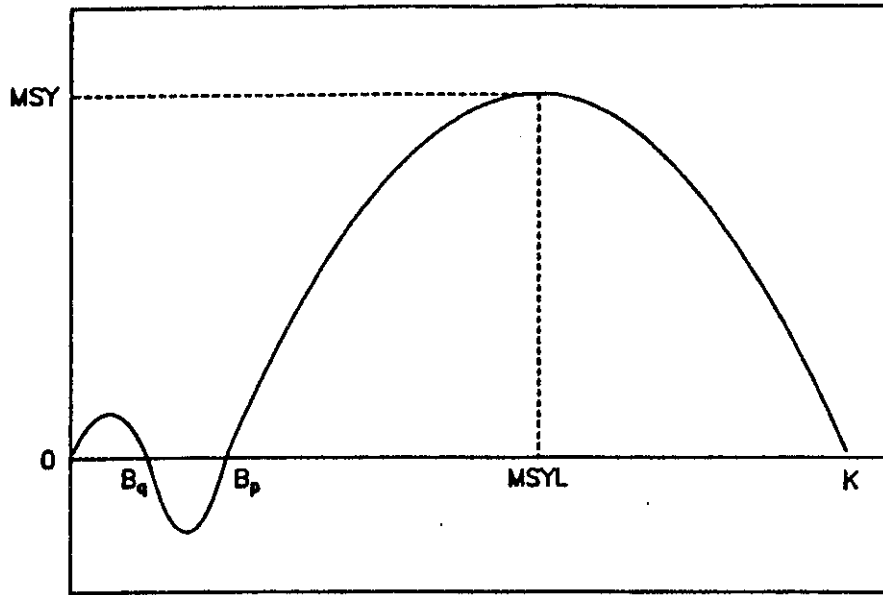


Biomass (B)

Figure 6: A surplus production curve showing (a) critical depensation and (b) the corresponding relative net growth rate as a function of biomass. Once the biomass decreases below  $B_c$ , the population can never recover and goes to extinction.

(a)

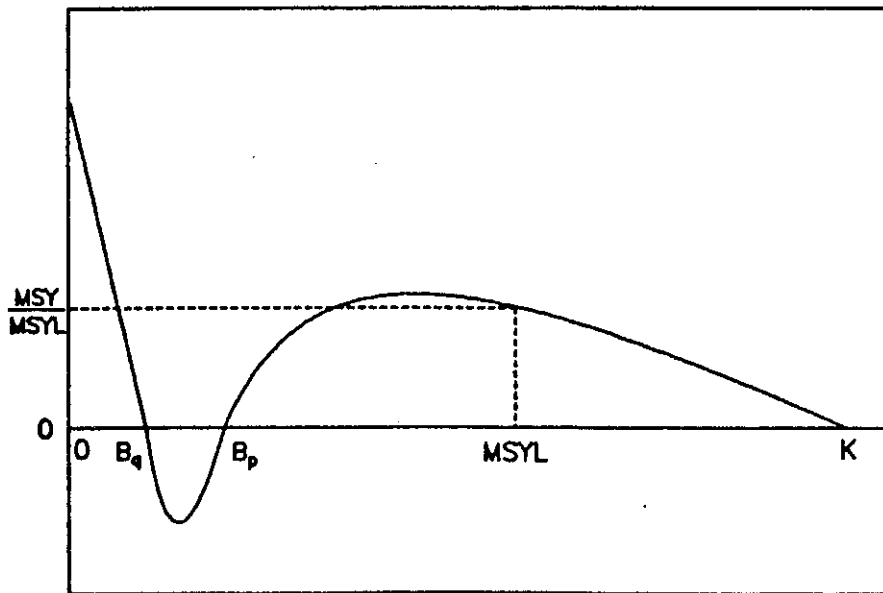
Equilibrium yield = net growth rate ( $\frac{dB}{dt}$ )



Biomass (B)

(b)

Per capita net growth rate ( $\frac{1}{B} \frac{dB}{dt}$ )



Biomass (B)

Figure 7: A depensatory surplus production curve (a) for a prey population with a "predator pit", and (b) the corresponding relative net growth rate as a function of biomass. Once the stock is reduced below  $B_p$ , the predators prevent it increasing beyond the lower of the two unexploited stable equilibrium levels,  $B_q$ .

level  $B_c$ , cannot recover and becomes extinct. This may occur in populations of widely spaced individuals with a low reproductive rate, such as whales. The model in Figure 7 exhibits a "predator pit" in which the prey population becomes "trapped" by its predators. Under continued exploitation the population may be driven below biomass  $B_p$ , where its defence mechanisms against natural predators (e.g. forming large shoals) become impaired, and the attentions of these predators alone are sufficient to deplete the resource still further. However, once the population has dropped below  $B_q$ , it becomes too small to satisfy the requirements of its natural predators, which then switch to other prey. If the population starts to recover, the predators will switch back, so that it becomes maintained by its predators at a new, lower unexploited equilibrium level,  $B_q$ .

In applying surplus production models to fisheries data, it is usual to ignore phenomena such as those discussed above. Essentially, it is assumed that maintaining the population above MSYL will prevent these possible effects from becoming a concern.

#### Management methods

Although in theory a stock may be maintained at MSYL by continually harvesting at a rate corresponding to the MSY, this is not true in practice because its dynamics cannot be described accurately in terms of an exact (deterministic) functional relationship but contain, in addition, a random component due to environmental fluctuations. Further, these fluctuations, together with the sampling errors in the statistics derived from catch data, mean that even the average MSY value cannot be estimated exactly, but will have an associated estimation error. Management strategies need to take into account the stochastic (randomly fluctuating) nature of a resource's dynamics and the imprecise estimates of its harvesting potential by continually monitoring its status.

The choice of a management strategy involves a trade-off between the rewards, in terms of the number and size of fish landed, and the risk of reducing the population size, if not to extinction, to a level at which the surplus production (and therefore the sustainable yield) is substantially reduced. The simple strategy of a constant smallish quota is one with a relatively low risk; however, it also leads to smaller rewards than would a more adventurous policy of varying the quota depending on the assessed status of the resource.

An example of this trade-off is provided by simulation model calculations of the consequences of different harvesting strategies for the South African anchovy (*Engraulis japonicus*) resource by Bergh and Butterworth (1987). Figure 8 compares the constant-catch strategy for this resource with one in which a fixed proportion of the estimated recruitment is caught each year (conditional on annual catches not changing by more than 15% from one year to the next, for reasons of industrial stability). The latter strategy shows higher rewards at all levels of risk and is therefore superior, but requires annual monitoring of the recruitment, which presents some practical difficulties.

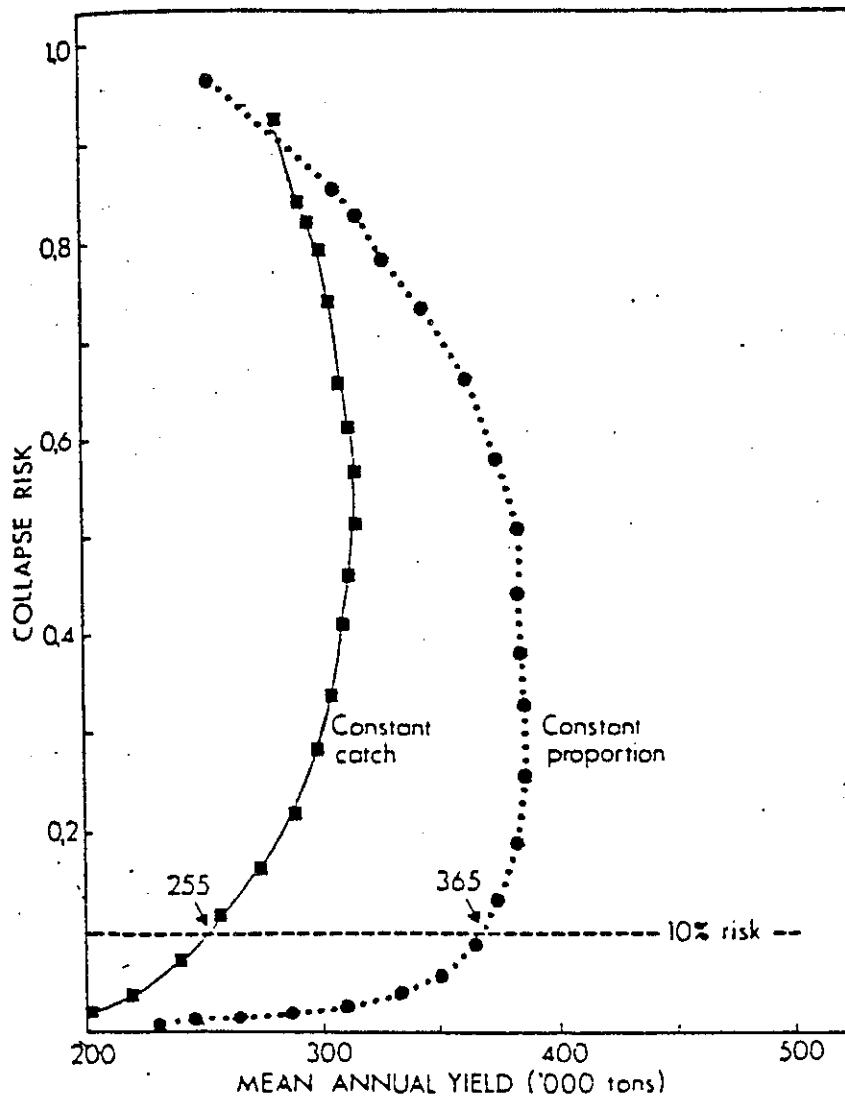
A commonly used management approach is illustrated in Figure 9. Here, a target biomass level which is considered to be the optimum, is specified ( $K/2$  in this case, i.e. MSYL for the Schaefer model), and a fishing quota is set according to the estimated current biomass level, with the objective of eventually stabilizing the resource at the optimum level. The recommended quota,  $Q$ , (also referred to as the TAC - Total Allowable Catch) is given by the equation:

$$Q = (B^*/MSYL).MSY$$

where MSYL is considered here to be the optimum biomass and  $B^*$  is the current biomass. When  $B^*$  is below the optimum level, a catch smaller than the corresponding optimum catch (MSY) is taken and *vice versa*. The advantage here is that, natural fluctuations aside, this strategy will ensure that the resource eventually stabilises at the target biomass level. This procedure is equivalent to taking out a fixed fraction of the current biomass each year, and is known as a "Constant F" (constant fishing mortality) strategy, or as the " $F_{MSY}$ " strategy when the target biomass level corresponds to MSYL. Under the additional assumption that CPUE (catch-per-unit-effort) is proportional to biomass, this is also a constant fishing effort policy, customarily denoted as an " $f_{MSY}$ " strategy, if MSYL is the target biomass level.

As a means of reducing the effort in an overexploited fishery, closed seasons are often imposed. Before such a strategy is implemented, however, an evaluation of the reduction in effort this is likely to achieve should be made. Methods of estimating effort in a linefishery are discussed in the following section.

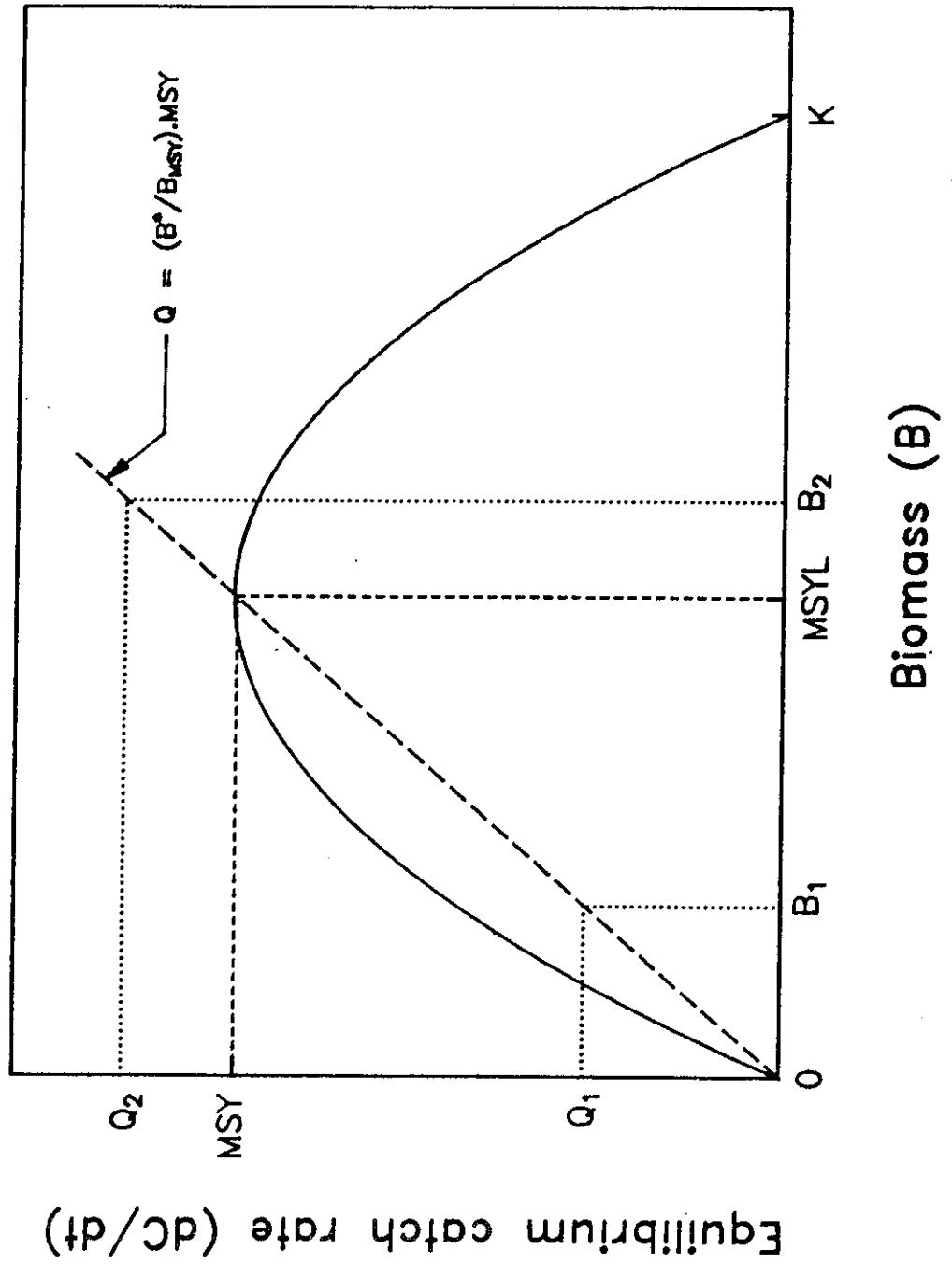




**Figure 8:**

Comparison of risk and reward for two different harvesting strategies for management of the South African anchovy stock: a constant-catch strategy and the more adventurous strategy of a constant proportion of estimated recruitment (after Bergh and Butterworth 1987, Fig. 8). The latter shows higher rewards at all levels of risk. (The results were obtained by integrating over the joint probability distributions of model parameters as estimated in 1986, and respect a 15% limitation on inter-annual quota variation.)

Figure 9: The  $F_{MSY}$  strategy (constant fishing mortality) for determining a series of quotas which will eventually stabilise the stock at MSYL is shown for the Schaefer model.  $Q_1$  and  $Q_2$  are the allowable catches corresponding to biomass levels  $B_1$  (below MSYL) and  $B_2$  (above MSYL) respectively, in terms of this strategy.



## (2) Information Derivable from Catch-Effort Data

### Data available for analysis

Data collected in line fisheries for catch-effort analysis usually comprise only relative indices of the catch and effort, i.e. statistics (assumed to be linearly) proportional to the absolute values. Thus:

$$\begin{aligned} E &= k_1 E_R \\ C &= k_2 C_R \\ \text{and } (C/E) &= k_3 (C/E)_R, \end{aligned}$$

where  $E$  = absolute effort  
 $k_{1/2/3}$  = constants of proportionality  
 $E_R$  = relative index of effort  
 $C$  = absolute catch by mass  
 $C_R$  = relative index of catch  
 $C/E$  = absolute value of CPUE  
 $(C/E)_R$  = relative index of CPUE .

The above equations are only valid if the  $k_i$  are constant over the time period spanned by the data series. It is, however, possible that effects such as changes in fishing legislation, equipment, etc. may result in the  $k_i$  varying with time.

### Detection of overexploitation

A linear relationship is frequently assumed to exist between CPUE and biomass:

$$\widehat{C/E} = qB \quad (4.2.1)$$

where  $\widehat{C/E}$  is the expected (or average) catch-per-unit-effort,  $q$  the catchability coefficient and  $B$  the biomass. For the same biomass level,  $\widehat{C/E}$  will differ from the observed catch rate  $(C/E)$  due to sampling variability and fluctuations in catchability caused by changing environmental conditions. The equations:

$$\begin{aligned} \sqrt{C} &= \sqrt{qEB} + \epsilon & \epsilon &\sim N(0, \sigma^2) \\ \text{and } C/E &= qBe^\epsilon & \epsilon &\sim N(0, \sigma^2) \end{aligned} \quad (4.2.2)$$

represent two models which are often used to incorporate a stochastic component into the relationship between CPUE and biomass to account for such random sampling and catchability fluctuations respectively. The second equation above has the advantage that observed CPUE values are necessarily positive.

In reality CPUE is proportional, not necessarily to overall abundance, but rather to local stock density, and even this relationship may not hold for certain types of fishing operation (see, for example, Cooke 1985). Extrapolation to overall abundance involves the more stringent assumption of random fishing over the entire range of the resource. Fishing is often a decidedly non-random operation, but the need for the latter assumption falls away if the fish can be considered sufficiently mobile to compensate for "gaps" caused by heavy localised effort. Alternatively, if fishing is restricted to only a part of the range of the stock, one needs to assume that the overall relative spatial density pattern remains the same from year to year. However, certain species may extend or decrease their range in response to abundance increases or decreases, thereby maintaining a near constant core area density; in such cases, CPUE trends from a fishery localised in the central high density region (or, similarly, one localised near the periphery of the range) would present a false reflection of true changes in abundance. A knowledge of the total distribution of the stock in relation to that portion fished is therefore important.

Given a time series of relative CPUE data  $(C/E)_R$ , then, on the basis of the Schaefer model and assuming CPUE proportional to biomass, a reduction of more than 50% in this index over the period covered by the series is sufficient to lead to the conclusion that the stock is biologically overexploited. For example, Figure 10(a) shows the catch-per-hour of elf (*Pomatomus saltatrix*) in angling competitions off the Natal coast from 1956 to 1976. A linear regression fit to the downward trend of this CPUE index reveals that the index has fallen over the 20 year period by 58% of its value at the start of the period, indicating that biological overexploitation has occurred. Such a conclusion fails, however, to take into account factors such as sampling variability and changes in the value of  $q$  due to environmental fluctuations [see equations (4.2.2)]. Estimates of stock declines should always be quoted with an associated standard error in the light of

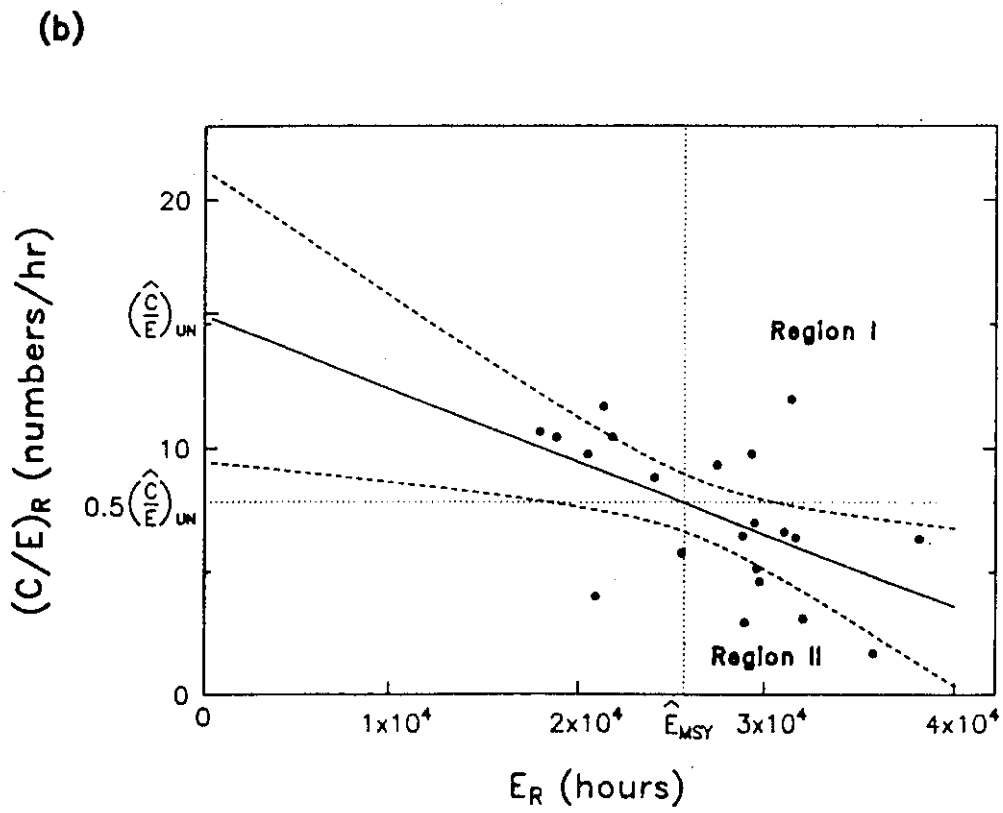
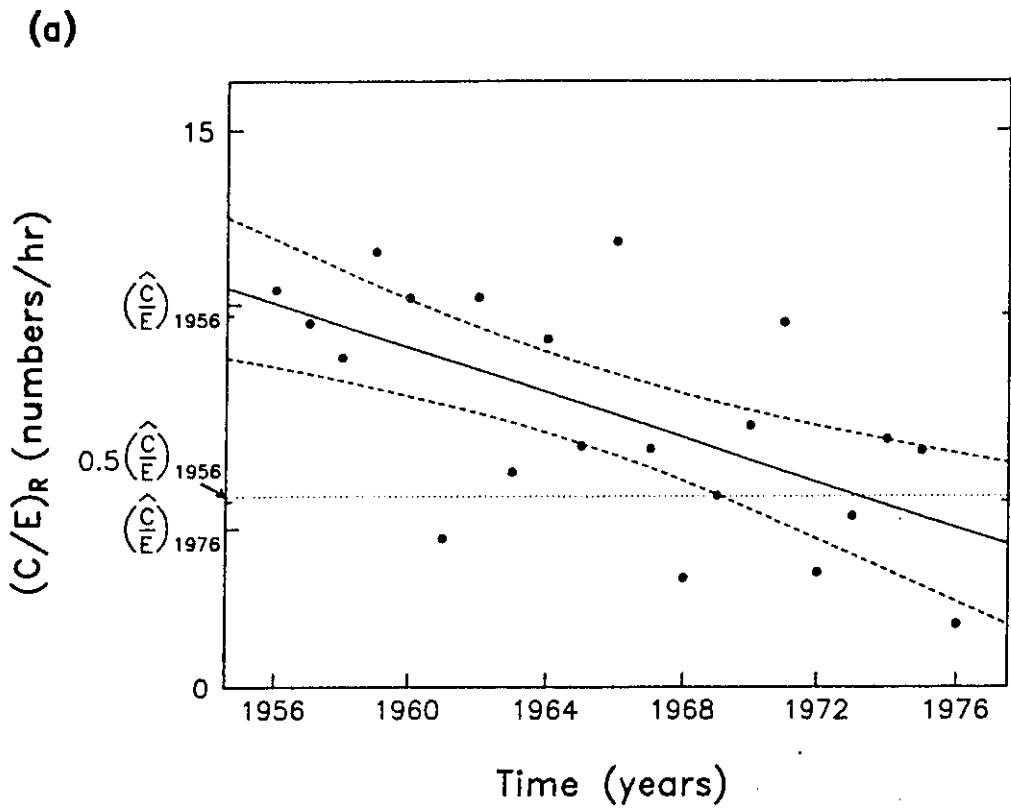


Figure 10: Relative CPUE trends with (a) time and (b) relative effort for the Natal eif fishery. Approximate 95% confidence intervals ( $\pm 2$  s.e.'s) on the curves are represented by dashed lines in each case.

these factors. Figure 10(a) also shows approximate 95% confidence intervals about the linear trend fitted to the data; these have been calculated using the jack-knife technique (see Chapter III). The standard error of the percentage decline is readily calculated by this procedure, yielding a result of 12% for the elf example considered. With the approximate 95% confidence interval for the decline from the 1956 estimate thus spanning the range [34%, 82%], the conclusion of biological overexploitation becomes less certain.

While data are seldom available for the initial stages of exploitation, it is clearly important to obtain an estimate of  $(C/E)_R$  for the unexploited resource, as, in cases where a stock has been heavily depleted prior to the beginning of the data series, to regard a reduction of less than 50% in the CPUE index as an indication that the stock is not overexploited, could constitute a serious error. A simple visual procedure which can be used in such circumstances to decide whether or not the stock is overexploited, provided a time series of (relative) total effort data ( $E_p$ ) is also available, is illustrated in Figure 10(b) for the Natal elf data for 1956 to 1976. Assuming a linear relationship between CPUE and effort (Appendix VII), the vertical axis intercept of the regression line in this Figure provides an estimate of the CPUE for this resource when exploitation commenced. In Region I of the Figure, although the biomass index is still greater than half its estimated unexploited level,  $E_{(R)}$  is greater than  $E_{MSY}$ , so that more effort is being used than is required to maintain the stock at or above MSYL. As the effort remains above  $E_{MSY}$ , the stock eventually declines to below MSYL (Region II) and the resource becomes biologically overexploited. Quantitatively, the estimate of  $(C/E)_{UN}$ , together with that for the "current" (i.e. 1976) CPUE obtained from the regression line, indicate a CPUE decline of 72% (s.e. 11%) since the beginning of the fishery, i.e. the biomass is now estimated to be at  $(0.28 \pm 0.11)K$ , which is significantly lower than  $K/2$ . This result is far more clearly suggestive of biological overexploitation than what is apparent from Figure 10(a).

Even this procedure is open to question, however, in the main because the linear relation between CPUE and effort holds only for equilibrium situations (see Appendix VII), while the resource abundance is changing with time. In cases of a decreasing CPUE trend with time [Fig. 10(a)], the procedure will produce a positively biased estimate of  $E_{MSY}$ . Techniques exist to correct for this bias (Butterworth and Andrew 1984, Walters 1986), but in general the quality of

linefishery data renders it questionable whether these would be worth implementing.

$(C/E)_R$  data for a recreational line fishery can be obtained from angling competitions and beach patrols, while  $E_R$  can be estimated from beach patrols, telephone surveys, licence issues and equipment sales.

#### Problems associated with catch-effort analyses

The methods above rely on the assumption of a linear relationship between  $(C/E)_R$  and stock size. However, various factors may lead to bias in estimates of population reduction based on this assumption, and it is important to attempt to identify and quantify such biases so that more reliable inferences can be drawn. Some such factors are:

##### - Species-direction of effort

In order to obtain a reasonable estimate of the fishing effort applied separately to a number of different species, some account needs to be taken of the fact that effort is not entirely random, but may be preferentially directed at certain species, for instance by choice of hook size, location, bait etc. One rule of thumb to define directed effort is to assume that if species A constitutes more than 50% of the catch for any appropriate sampling unit (e.g. vessel-day), the associated effort is considered to be directed at A. The total effort on A over the time period of interest is then calculated as follows:

$$E_A^{TOT} = (E_A^D C_A) / C_A^D \quad (4.2.3)$$

where  $E_A^D$  = total directed effort on A

$C_A$  = total catch of A

$C_A^D$  = catch of A made during time of directed effort.

This is only one method of taking species-direction of effort into account, however, and although it should provide reasonable results for the more abundant species, it is not particularly reliable for the less abundant ones. Very little effort will tend to be allocated as "directed" at the latter species in terms of this rule of thumb, so that the "directed" catch for each of them will constitute only a very small fraction of the corresponding total catch.

Another approach is to decide which species were definitely targeted, and which were possibly targeted, for each unit of effort; then, for any species, calculations involving effort should be performed twice, assuming (i) directed effort = "definite effort" and (ii) directed effort = "definite effort" + "possible effort", to see whether the two methods give substantially different results. In general, it is advisable to use more than one method of taking species-direction effects into account when calculating the time series of effort directed at a species, in order to gauge whether or not the choice of method makes a substantial difference to estimated trends. Two possibilities (corresponding to what may be near-extreme opposites) could be to compare the series obtained using equation (4.2.3) for determining the directed effort with that obtained assuming no species-direction of effort at all; if the resulting two series reveal very similar trends, species-direction effects probably need not be of much concern.

In a species-directed fishery where few individuals are caught (e.g. marlin), it may be possible to draw inferences about CPUE even if complete effort data have not been recorded. All that is required is information on the dates on which each individual fish was caught, and the vessel making the catch. If catch data for each fishing trip are assumed to have identical and independent Poisson distributions, then the Catch-per-Gross-Vessel-Day (CPUE) is represented by the Poisson distribution parameter,  $\lambda$ . From the available data, the only index that can be calculated is the Catch-per-Net-Vessel-Catchday (CND), i.e. the average catch per vessel per day, counting only day-vessel combinations on which a catch was made (so that CND is necessarily  $\geq 1$ ). CPUE can then be estimated from this statistic using the relation:

$$\text{CPUE} = \text{CND} - 1.$$

The derivation (Cooke 1984) for this formula is given in Appendix VIII.

#### - "Power factors"

The assumption that CPUE is proportional to biomass over a period of time implicitly requires that consistent units are used to measure effort over this period, so that the catchability  $q$  remains constant in this context. Where this is not the case, the introduction of "power factors" in the calculation of effort is an attempt to quantify changes in the basic units of fishing effort as gear



becomes more efficient and fish-finding aids (e.g. sonar) become available. A larger vessel, for example, will have a greater power factor than a smaller one, and contribute more effort over the same period of time. Procedures to evaluate power factors are available, usually based on multi-linear regression models (Robson 1966).

#### - Rod level saturation

As the number of anglers on a stretch of beach increases beyond a certain limit, the probability per unit time of one of them catching a fish (to which the catchability coefficient,  $q$ , is proportional) decreases, resulting in smaller CPUE values, which therefore provide a negatively biased index of abundance. This effect may be insubstantial, however, when averaged over larger time and spatial scales.

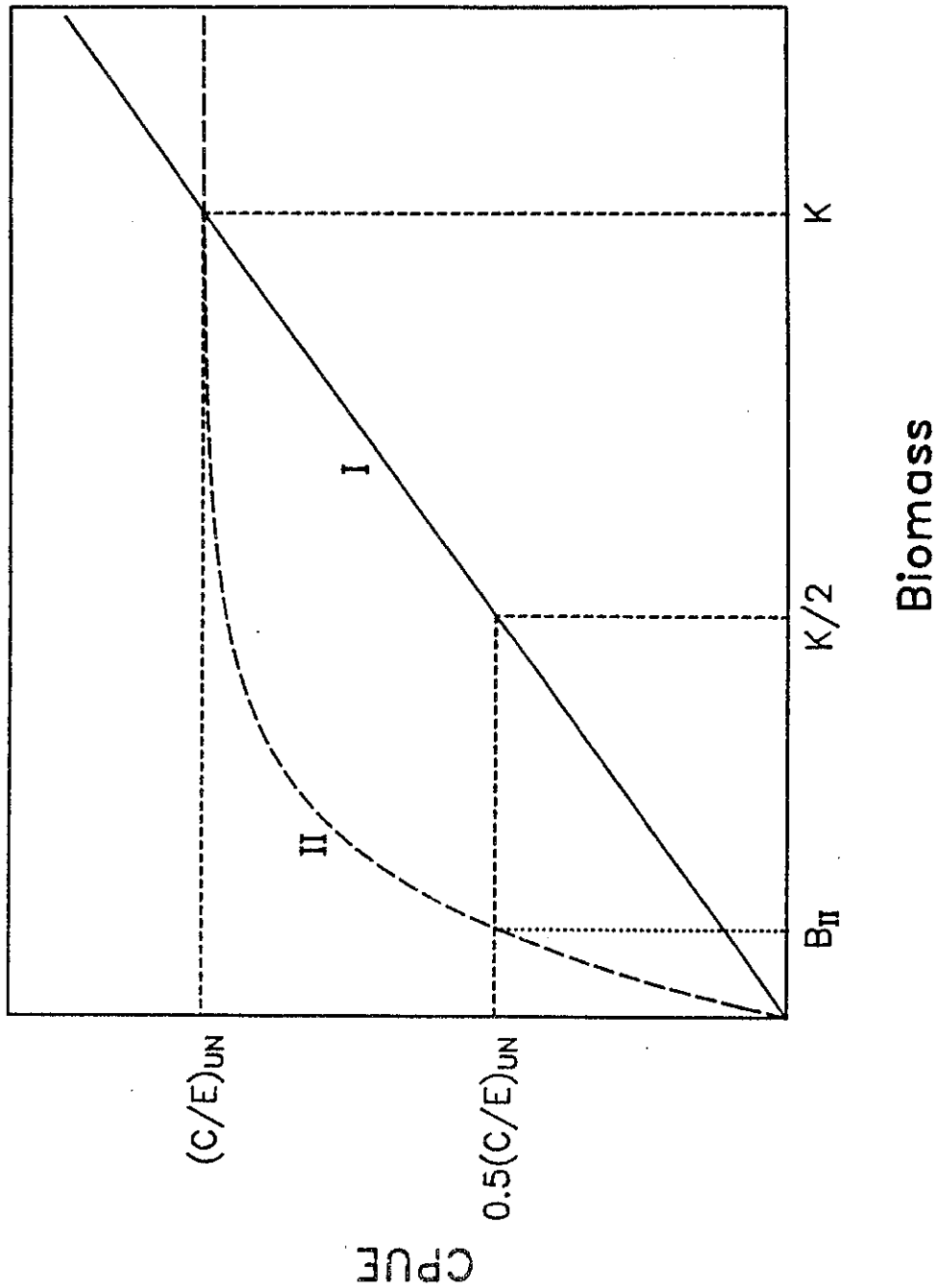
#### - Changes in spatial fishing patterns

Trends in spatial fishing patterns may also produce apparent changes in  $q$ . Once catch-rates in prime fishing spots decrease, fishermen tend to move to previously unexploited areas. This may lead to average catch rates being maintained even though total resource abundance is declining. This situation is illustrated in Figure 11, in which the relationship between CPUE and biomass is not a straight line but a convex curve, and the assumption of a linear relationship for drawing inferences about the state of the resource could have disastrous consequences. Ideally, data should be stratified by area with an index of abundance in each region provided by the relation  $C/E \propto \text{density}$ . The total biomass is given by  $\sum(d_i A_i)$ , where  $d_i$  is the fish density (per unit area) and  $A_i$  the area of region  $i$ , and this is indexed by  $\sum[A_i (C/E)_i]$ . This approach, however, involves the estimation of many regional catch rates [the  $(C/E)_i$ 's], so that the associated increase in the "discrepancy due to estimation" may more than offset the reduction in bias.

#### - Changes in angler quality profile

The basic unit of effort for anglers is the rod-hour. While rod efficiency is probably nearing its limit, variations in the prowess of different anglers certainly exist. These will not result in bias, however, if the proportions of anglers with different degrees of skill remain constant. If the skill profile

Figure 11: A linear (I) and nonlinear (II) relationship between CPUE and biomass. Most effects which bias the linear relationship tend to render it convex, as illustrated by (II). Note that for (II), a decline in abundance is not reflected as markedly by the drop in CPUE values. To base management recommendations on the assumption of linearity, should (II) in fact apply, could have serious consequences for the stock; a 50% drop in CPUE from its unexploited level corresponds to a stock level ( $B_{II}$ ) well below  $K/2$  in this case.



changes with time though, artificial trends in CPUE as an index of biomass could result, e.g. an influx of inexperienced anglers may lead to a decrease in the recorded CPUE although the stock abundance is unaltered. It may therefore be helpful to monitor angler quality profile so that allowance for changes can be made.

#### - Fish distributional factors

A reduction in population size may result either in a reduced stock density over an unchanged area, or the same density over a smaller area. In the former case, this will be reflected in the CPUE index. In the latter situation, no decline may be detected if fishing has been concentrated in the smaller area throughout the period; however, if the fishery has been operating on the periphery of the stock's range, even a small reduction in biomass will result in a substantial decrease in CPUE. This emphasises the fact that monitoring CPUE alone may be insufficient to detect overexploitation of a resource, as illustrated in Figure 11.

#### - Migration

Migration of some percentage of the population may be misinterpreted as a change in resource abundance, because of the resultant effect on the CPUE. It may be necessary to remove short-term (e.g. seasonal) trends from CPUE time series data to avoid false conclusions due to this effect. Knowledge of the behaviour patterns of a species is clearly important in interpreting catch-effort data, and an attempt should be made to incorporate such information (suitably quantified) into analyses before conclusions are drawn concerning the state of a resource.

#### Effort indices and beach patrols

Beach patrols provide a practical means of monitoring effort in a linefishery. An increase in relative effort,  $E_R$ , of 30% (s.e. 10%) for Natal beach fishing between 1978 and 1982 has been estimated (Hughes 1985) from data obtained from the Natal Parks Board beach patrols. The 95% confidence limits on this figure encompass a large increase in relative effort of 50% (which would certainly be a cause for concern and is possibly indicative of the need for restrictive management action), and a 10% increase which is of little consequence in a

management context. This example illustrates the importance of the degree of precision obtainable when using data sets such as those acquired from beach patrols to estimate fishing effort trends. For meaningful results, the degree of uncertainty should always be quantified, and an attempt made to improve the precision of the estimates. Precision may be increased by stratification of the data to take account of patterns due to seasons, weather, tides, holidays and weekends, for example. Other factors which may influence the amount of fishing effort expended, and hence affect estimation precision, should also be investigated and wherever appropriate incorporated into the analysis.

Special problems relating to beach patrols encountered during the Natal beach fishing study (Hughes 1985), serve to illustrate the need for careful consideration when choosing an index of effort or CPUE. The index of relative effort ( $E_R$ ) used in prior analyses of these data was based on the number of rods counted during a single patrol. This index proved to be inconsistent from zone to zone, however, as distances covered per unit time differed substantially both within and between zones. A more accurate index in this situation is the number of rods per length of patrolled beach, which is then multiplied by the total length of the zone.

CPUE indices obtained from beach patrols during this study also proved to be problematical, as illustrated by the following example. If 10 anglers fish for 10 hours each, or 100 anglers fish for one hour each during the same 10 hours, the average effort level reported will be (quite correctly) 10 units. The catch reported to the patrol by the 10 fishermen in the former case, however, will be 10 times that of the "average" 10 in the latter case. Obtaining an accurate measure of CPUE therefore also requires knowledge of the time spent fishing by each angler, either acquired *in situ* by direct questioning (usually impractical) or (preferably) from an independent survey of the distribution of fishing times on different beaches.

## V. ANALYSIS OF LENGTH COMPOSITION DATA

### Introduction

Models and management methods based only on catch and effort data are of an empirical nature. A different type of model frequently used in fisheries assessment and management is the "analytic model", which attempts to represent some of the underlying biological processes occurring, such as the birth, growth and death of an individual fish. For many linefisheries, effort data sets may not be available, but even if they are, implications for management may be difficult to infer from them due to uncertainty about the magnitudes of the possible biasing factors discussed in the previous chapter. It is therefore important to evaluate the available biological data and interpret these, using analytical models, in terms of their implications for management.

The principal underlying concept of analytic models in their application to management is that the greater the fishing pressure on a resource, the shorter the period any single fish is likely to live, and so the smaller and lighter the average fish caught will become. Frequently, some information on the mass or length distributions of the fish caught will be available; most methods of analysis are based on the age distribution of the catch, so that it is necessary first to convert the mass distribution to length (via a mass-length relationship) and second, the length distribution to age, using one of the methods discussed in the following section.

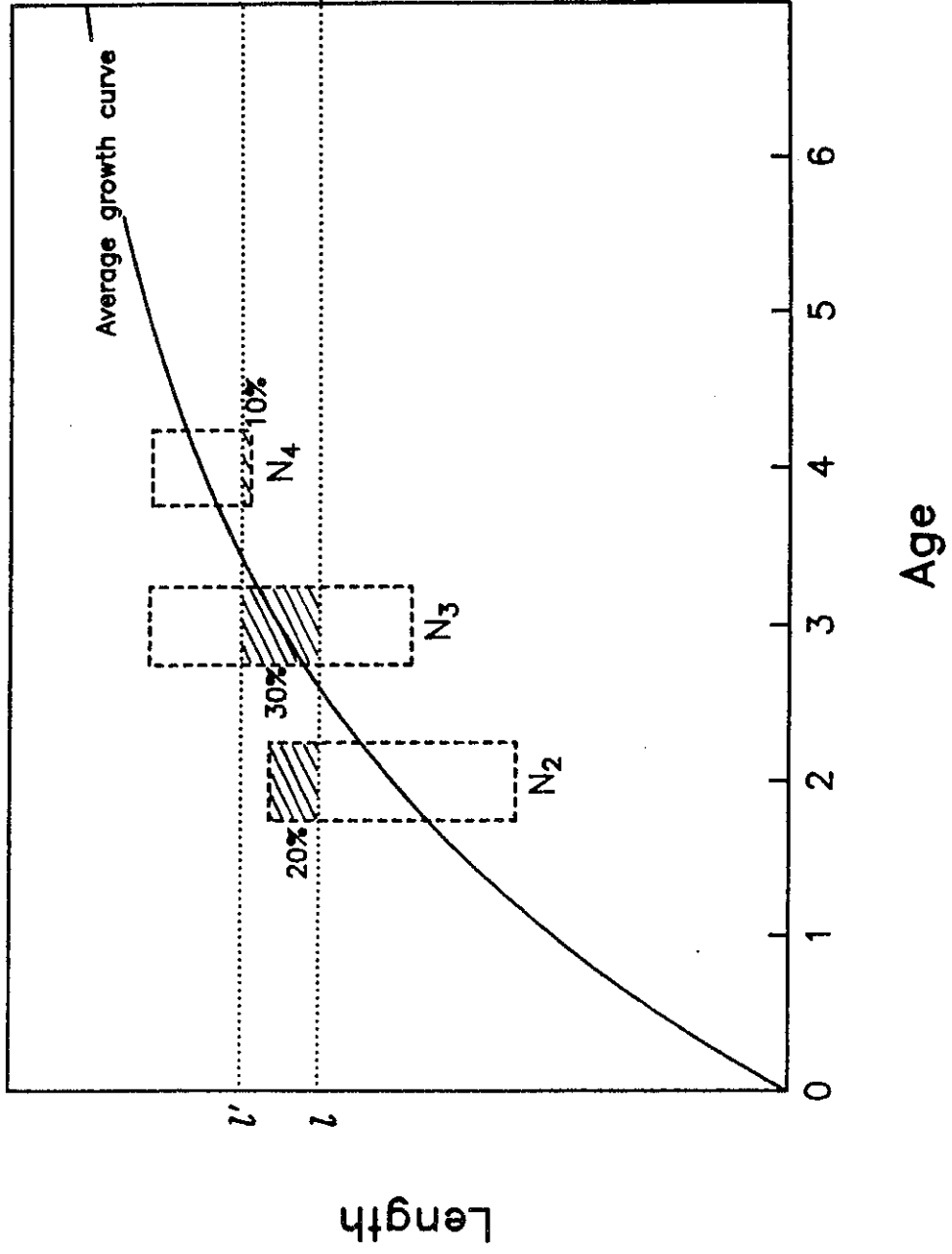
Most fish species have distinct spawning seasons, usually one per year. It is therefore customary to divide the catch (and also the population) into "age-classes", which may be fairly broad in terms of the range of true birth-dates included. For simplicity, however, all fish are assumed to have been born on the same day of the year, so that a fish, once captured, is assigned to the 2-year-old class if it has passed its second but not its third birthday, for example. A "cohort" is a group of fish born in the same spawning season, and therefore regarded as having the same birth date. Thus, at a given time, a population consists of the 1979-cohort (or 1979-year-class), the 1980-cohort, and so on. As the years pass, a particular cohort moves from one age-class to the next, with its numbers being depleted continuously by fishing and by natural mortality due to factors such as disease and consumption by predators.

### Transformation of length distribution to age distribution

Length distribution data can be converted to estimates of the age distribution of the catch by a number of methods which include the following:

- i) Allocate a single age to each length-class in the length distribution histogram, using the length-at-age given by, say, the Von Bertalanffy equation, and then sum for each age. (This is the crudest approach, and care should be taken when fitting a simple form such as the Von Bertalanffy growth curve to a wide range of ages, as this may produce unreliable estimates for certain ages-at-length, particularly for the younger age classes.)
- ii) Polymodal analysis. This technique is based on the interpretation of "bumps" in the length distribution as different cohorts, and various more or less complicated methods of implementing it are available (see, for example, MacDonald and Pitcher 1979, Schnute and Fournier 1980). A set of normal distribution curves, for example, may be fitted to the length distribution, or information from a fitted growth curve (e.g. Von Bertalanffy) incorporated.
- iii) The use of an age-length key, which provides an estimate of the age distribution for each length-class (see Appendix IX). As the key depends on relative cohort strengths as well as on the growth curve, an age-length key for one year may not be applicable to other years. Figure 12 shows the (hypothetical) percentage of fish in each of three age-classes which fall into length-class [ $L, L'$ ] on a Von Bertalanffy growth curve in a particular year. Even if the distribution of length-at-age is assumed to be constant over time, relative cohort strength factors will change the key from one year to the next. For example, if in year 1 the sizes of the cohorts  $N_2$ ,  $N_3$  and  $N_4$  are 100, 30 and 10 respectively, the age-length key entries for this length-class are 67% (age 2), 30% (age 3) and 3% (age 4). If, due to good recruitment, these cohort sizes change the following year to 200, 30 and 10 respectively, the corresponding age-length key entries then become 80% (age 2), 18% (age 3) and 2% (age 4). Note also that two factors contribute to the distribution of length at age: firstly, all individuals do not necessarily follow exactly the same growth curve and secondly, individuals from the same cohort are not all born simultaneously (as

Figure 12: Age-length keys: the shaded areas superimposed on the growth curve represent the percentage number of fish falling into each of three age-classes of sizes  $N_2$ ,  $N_3$  and  $N_4$ , which comprise the length class  $[L, L']$ .



sometimes has to be assumed when plotting growth data); instead their births are spread over a spawning season whose time and duration may vary from year to year. This second factor may result in a length-at-age distribution which is not constant over time. Thus, it is quite possible that (as in Figure 12) the largest fish in the 4-year-old cohort, for example, are hardly bigger than the largest in the next youngest cohort.

### Virtual Population Analysis (VPA)

VPA is a technique used to estimate cohort strengths and fishing mortality  $F$ , and requires that a matrix of catch-at-age data (in numbers) is available. The original size of a recruiting cohort ( $N_0$ ) must have been greater than the total catch (or "virtual population") taken subsequently from that cohort, i.e.:

$$N_0 > \sum_{a=0}^m C_a$$

where

$N_0$  = original number of fish in cohort

$C_a$  = catch of fish of age  $a$  (in numbers) from that cohort.

Assuming further that, in the absence of fishing, a cohort declines exponentially with time due to natural mortality, then:

$$N(t) = N_0 e^{-Mt} \quad \left( \text{which follows from } \frac{dN}{dt} = -MN \right) \quad (5.1)$$

where

$N(t)$  = number of fish at time  $t$

$N_0$  = number of fish at time  $t = 0$  (original size of cohort)

$M$  = natural mortality rate.

For the pulse-fishing approximation, in which it is assumed that all fishing takes place instantaneously at mid-year, and that all fish are born at the beginning of the year, the number of fish remaining 6 years (for example) after birth,  $N_6$ , is given by:

$$N_6 = (N_5 e^{-M/2} - C_5) e^{-M/2}$$

where  $N_5 e^{-M/2}$  is the population size on which fishing takes place, and the final  $e^{-M/2}$  factor represents the decline due to natural mortality for the balance of the year.



Alternatively, for fishing throughout the year, if it is assumed that the catch rate is proportional to population size:

$$dC/dt = FN$$

the governing differential equation becomes:

$$\frac{dN}{dt} = -MN - FN = -(M + F)N . \quad (5.2)$$

Generalising to all years ( $y$ ) and ages ( $a$ ), where the fishing mortality may vary for the different age-groups in any particular year, the solution to these differential equations is:

$$(i) \quad N_{y+1,a+1} = N_{y,a} e^{-(F_{y,a} + M)} \quad (5.3)$$

where  $N_{y,a}$  = number of fish of age  $a$  at the start of year  $y$   
 $F_{y,a}$  = fishing mortality on fish aged  $a$  in year  $y$ ,

$$(ii) \quad C_{y,a} = F_{y,a} N_{y,a} \frac{1 - e^{-(F_{y,a} + M)}}{F_{y,a} + M} \quad (5.4)$$

where  $C_{y,a}$  = catch of fish aged  $a$  in year  $y$ .

Given estimates either of terminal cohort numbers or of terminal fishing mortalities (numbers of, or fishing mortalities applied to, fish in the oldest age-class sampled, for each cohort), these equations can be used, working backwards, to calculate all earlier  $F$ 's and  $N$ 's. [The results from VPA are crucially dependent on these terminal estimates. For commercial fisheries, they are most often determined by a technique known as *ad hoc* tuning, which relies on the availability of (usually) a time series of (relative) effort values, to which the fishing mortalities are assumed to be related (customarily linearly proportional). For linefisheries, even this procedure may not be viable, and a cruder set of assumptions may be required, as in the example discussed in the following section. (It should always be remembered, however, that conclusions drawn, particularly as regards trends, remain dependent on this cruder set of assumptions.)]

The solution of the VPA equations (5.3) and (5.4) above is not straightforward, so that approximations are often used to make the calculations easier. Two such

approximations are those of Pope (1972) (the pulse-fishing situation discussed earlier), and a slightly more accurate version proposed by MacCall (1986):

$$N_{y,a} = N_{y+1,a+1}e^M + C_{y,a}e^{M/2} \quad \dots \text{Pope} \quad (5.5)$$

$$N_{y,a} = N_{y+1,a+1}e^M + C_{y,a}(1 - e^{-M})/M \quad \dots \text{MacCall} .$$

The accuracy of these approximations should be quite adequate for linefishery applications. The equation chosen is used, again working backwards, to obtain  $N_{y,a}$  values, given estimates of terminal cohort numbers. Fishing mortality estimates then follow from the equation:

$$F_{y,a} = \ln(N_{y,a}/N_{y+1,a+1}) - M \quad . \quad (5.6)$$

#### - Application of VPA to the species *Pomatomus saltatrix* (elf)

VPA requires data for the total catch-at-age of a species over a period. Such data are not available for the elf off Natal. What is available, however, are catch-at-age figures for elf taken in Natal angling competitions, which are given in Table 1 (Van der Elst, pers. comm.). It is still possible to apply VPA to these data, provided the assumption can be made (within reason) that the ratio of angling competition fishing effort to total fishing effort remained constant over the period considered. If this assumption holds, the VPA will provide correct estimates for fishing mortality and relative year-to-year recruitment and biomass levels. The absolute recruitment and biomass estimates, however, will be incorrect; they need to be scaled upwards by the inverse of the effort ratio mentioned above, a value not yet well known. Thus, the values in Tables 3 and 4 following are reliable; the "numbers of fish" in Table 2 need to be adjusted by an unknown multiplicative factor.

The following further assumptions have been made to carry out the VPA:

$$M = 0.4$$

$$N_{y,11} = 0$$

$$N_{1977,a} = 0 \quad \text{for } a \geq 7.$$

The number of fish in each age-class for the cohorts recruiting in the years 1956-1972, calculated using Pope's approximation, is given in Table 2. The reason

that the post-1972 cohorts have been excluded from this Table is the absence of a basis on which to set the corresponding terminal cohort numbers; the assumption of  $N_{1977,a} = 0$  for  $a < 7$  would be unrealistic. These figures are then used to calculate the corresponding values for  $F_{y,a}$  (Table 3) from equation (5.6).

What pertinent conclusions can be drawn from these calculations? The coefficient of variation of the annual recruitment values is 24%; this information can be used in simulation models to check how low the resource is likely to be driven under a particular target fishing mortality in the presence of natural recruitment fluctuations. Greater care needs to be taken with populations showing larger recruitment fluctuation because of the danger of the spawning biomass dropping to a very low level following periods of poor recruitment. Knowledge of the extent of this variability is also required to assess the usefulness of the average mass of fish caught ( $m$ ) as an index of the level of fishing mortality (see discussion on yield-per-recruit analysis following).

Also of interest are trends in recruitment and in fishing mortality with time. Two series of fishing mortality values for elf, averaged over different sets of ages, are shown in Table 4. Note that the first of these data sets suggests a decreasing trend in fishing mortality over the 1959 to 1972 period, although no such trend is evident when the mortalities for ages 3 and 4 only are considered. The two average values of  $F$ , however, are 0.74 and 0.92, which are high compared with the recommended guideline of  $F < M$  (here  $M = 0.4$ ). Nevertheless, the recruitment series shows no marked trend, suggesting that the high fishing mortalities have not impaired recruitment success over the period considered. It is important to remember, however, that these conclusions depend on the appropriateness of the assumptions made for the terminal cohort numbers in initiating the VPA calculations.

#### Modelling the catch-curve to find $F$ - an alternative analysis of age data

When only limited information (or even none at all) is available on the age structure of the catch for certain years, the application of VPA may not be possible and an approach is required which can extract pertinent information (particularly an estimate of the fishing mortality) from the available age data accumulated over a period of several years (or perhaps a number of such periods). Further, VPA requires not only age-distribution data, but estimates for the total

TABLE 1: Observed catch-at-age in numbers for *P. saltatrix* (Van der Elst pers. comm.).

YEAR	CATCH-AT-AGE								
	$C_{y,2}$	$C_{y,3}$	$C_{y,4}$	$C_{y,5}$	$C_{y,6}$	$C_{y,7}$	$C_{y,8}$	$C_{y,9}$	$C_{y,10}$
1956	756	869	177	84	29	3			
1957	269	649	582	325	153	17	6	1	
1958	327	871	468	318	117	12	3	1	
1959	367	1283	515	299	22	1			
1960	394	962	487	88	37	2	1		
1961	141	375	246	46	16	7	4		
1962	323	931	888	107	18	4	0	2	
1963	276	682	406	32	9	3	2	1	
1964	389	1637	471	39	93				
1965	460	1137	232	25	5	1			
1966	1057	2428	212	4	152				
1967	458	1129	385	24	17	6	2		
1968	185	475	176	9	4	2			
1969	415	875	186	19	6	6	3		
1970	488	1319	197	24	16	5	1		
1971	539	1661	508	85	35	8	4		
1972	65	503	291	65	33	12	8	3	10
1973	146	936	248	20	4	1	4	5	7
1974	285	1140	457	72	44	28	9	6	20
1975	318	1230	714	81	31	10	16	7	17
1976	74	269	177	30	19	14	12	8	6

TABLE 2: Estimated numbers-at-age for *P. saltatrix* from VPA (Pope's approximation).

YEAR	COHORT NUMBERS-AT-AGE									
	$N_{y,2}$	$N_{y,3}$	$N_{y,4}$	$N_{y,5}$	$N_{y,6}$	$N_{y,7}$	$N_{y,8}$	$N_{y,9}$	$N_{y,10}$	$N_{y,11}$
1956	4972	3097	1130	414	75	17	2	-	-	-
1957	3825	2714	1364	613	209	26	9	1	-	-
1958	4410	2344	1288	438	145	15	4	1	0	-
1959	3385	2689	858	480	33	1	0	0	0	-
1960	3772	1969	752	154	77	4	0	-	-	-
1961	3206	2206	532	105	31	21	1	-	-	-
1962	3164	2033	1171	155	33	8	9	0	-	-
1963	4128	1856	601	58	16	7	2	2	-	-
1964	3612	2541	686	70	13	4	2	1	0	-
1965	6474	2103	363	74	15	1	0	0	0	-
1966	4880	3963	479	53	29	6	0	-	-	-
1967	3032	2406	669	147	33	7	2	-	-	-
1968	3351	1657	688	133	79	8	0	0	-	-
1969	5208	2095	722	317	82	50	4	-	-	-
1970	5013	3151	688	332	197	50	28	0	-	-
1971	2855	2961	1032	300	203	119	29	18	-	-
1972	3471	1472	625	276	131	107	73	16	12	-
1973		2273	575	180	132	61	62	43	9	0
1974			757	182	105	85	40	38	24	0
1975				134	63	34	34	19	21	0
1976					23	17	15	10	7	0
1977						0	0	0	0	0

Note: The "0"'s shown above correspond to terminal cohort numbers assumed to enable calculations using equation (5.5) to be carried out.

TABLE 3: Estimated fishing mortality-at-age for *P. saltatrix* from VPA (Pope's approximation).

FISHING MORTALITY-AT-AGE									
YEAR	F <sub>y,2</sub>	F <sub>y,3</sub>	F <sub>y,4</sub>	F <sub>y,5</sub>	F <sub>y,6</sub>	F <sub>y,7</sub>	F <sub>y,8</sub>	F <sub>y,9</sub>	F <sub>y,10</sub>
1956	0.21	0.42	0.21	0.28	0.64	0.24	-	-	-
1957	0.09	0.35	0.74	1.04	2.26	1.57	1.61	-	-
1958	0.09	0.60	0.59	2.18	4.37	-	-	-	-
1959	0.14	0.87	1.32	1.43	1.65	-	-	-	-
1960	0.14	0.91	1.57	1.20	0.89	0.85	-	-	-
1961	0.06	0.23	0.83	0.76	1.00	0.51	-	-	-
1962	0.13	0.82	2.60	1.84	1.10	1.03	0.85	-	-
1963	0.09	0.60	1.75	1.11	1.10	0.70	-	-	-
1964	0.14	1.55	1.82	1.13	1.95	-	-	-	-
1965	0.09	1.08	1.52	0.53	0.51	-	-	-	-
1966	0.31	1.38	0.78	0.10	0.98	0.51	-	-	-
1967	0.20	0.85	1.21	0.22	1.02	-	-	-	-
1968	0.07	0.43	0.37	0.09	0.06	0.37	-	-	-
1969	0.10	0.71	0.38	0.08	0.09	0.16	-	-	-
1970	0.13	0.72	0.43	0.09	0.10	0.13	0.04	-	-
1971	0.26	1.16	0.92	0.43	0.24	0.09	0.18	0.00	-
1972	0.02	0.54	0.84	0.34	0.37	0.15	0.14	0.25	-
1973		0.70	0.75	0.15	0.04	0.02	0.08	0.15	-
1974			1.34	0.66	0.72	0.51	0.32	0.21	-
1975				1.35	0.91	0.44	0.85	0.58	-
1976					-	-	-	-	-

Note: " - "s indicate that F values cannot be determined from the data available in Table 2, or are infinite as a result of the zero terminal cohort size assumption.

TABLE 4: Summary of mean F values for *P. saltatrix* obtained from Table 3.

YEAR	MEAN F-AT-AGE	
	A	B
1956	0.33	0.32
1957	1.09	0.54
1958	1.57	0.60
1959	1.08	1.10
1960	0.93	1.24
1961	0.57	0.53
1962	1.20	1.71
1963	0.89	1.17
1964	1.32	1.68
1965	0.75	1.30
1966	0.68	1.08
1967	0.70	1.03
1968	0.23	0.40
1969	0.25	0.55
1970	0.23	0.57
1971	0.47	1.04
1972	0.33	0.73
average:	0.742	0.917

A - average over all defined F values  
 B - average over  $F_{y,3}$  and  $F_{y,4}$

catch every year, which are seldom available for a recreational fishery, so that an alternative method becomes a necessity. Such methods model the "catch-curve", which is a plot of catch in numbers ( $C_a$ ), usually summed over a period of several years, against age ( $a$ ).

Two processes contribute to the shape of the catch curve: the first is the age structure of the resource under exploitation, and the second is the pattern of fishing mortality with age, which is called "selectivity". A changing selectivity with age represents a situation in which the age composition of the catch does not match that of the population. This can occur, for example, if the net mesh-size used allows the smaller (younger) fish to escape, if the hook size is too large to catch the smaller fish, or if the population tends to be geographically stratified by age with the fishery concentrated on certain strata only.

For simplicity, assume that the fishing pattern does not change with time. Selectivity is then defined by the equation:

$$F_{y,a} = F_y S_a$$

where  $S_a$  = age-specific selectivity (i.e. partial recruitment) at age  $a$   
 $F_y$  = fishing mortality on fully recruited ( $S_a = 1$ ) cohorts in year  $y$ .

Assuming further that fishing mortality is proportional to fishing effort, this equation may be written:

$$F_{y,a} = qE_y S_a$$

where  $q$  = catchability coefficient  
 $E_y$  = effort in year  $y$ .

In cases where the dominant selection effect is the avoidance of young fish, the selectivity curve will have an ogive shape as illustrated by curves (b) and (c) in Figure 13. Combining this with the generally decreasing number of fish with age in the population as a whole, produces a catch-curve showing an initial peak followed by an exponential-like decline, as exhibited by the accumulated data for Natal elf in Table 1 (Fig. 14).

Modelling the catch-curve requires a model for the selectivity function,  $S_a$ . The simplest assumption is that of "knife-edge" selectivity: no fish are caught below



the age-at-first-capture ( $t_c$ ), whilst above this age they are equally available to fishermen, independent of their age (see Figure 13). This assumption is often made for yield-per-recruit analyses (see following section), and is written:

$$s_a = \begin{cases} 0 & \text{for } a < t_c \\ 1 & \text{for } a \geq t_c \end{cases} \quad (5.7)$$

Under these circumstances (and conditional on the further assumption that recruitment remains constant with time, discussed at the end of this section), the catch-curve will show an exponential decline whose slope is determined by the total mortality rate  $Z$ , where  $Z = (F + M)$  and  $F$  is some average of the  $F_y$ 's for the period under consideration. Even if the knife-edge selectivity assumption is not satisfied, this relation between the total mortality and the catch-curve slope still holds provided that only those ages above the age at full recruitment ( $a_f$ ) are considered, where  $a_f$  is defined by the relation:

$$s_a = 1 \quad \text{for } a \geq a_f .$$

The total mortality rate ( $Z$ ) is usually estimated from the slope of the linear regression of  $\ln(C_a)$  against age (for  $a \geq a_f$ ), and an estimate of  $F$  is then obtained by subtracting  $M$ . However, this procedure tends to produce substantially biased estimates of  $Z$  if the numbers caught at any age included in the regression are low. A statistically superior estimator for  $Z$  (Chapman and Robson 1960, Cooke and Beddington 1981) which overcomes this problem is:

$$\hat{Z} = \ln[1 + 1/(\bar{a} - a_f)] \quad (5.8)$$

with approximate 95% confidence limits:

$$\hat{Z}[1 \pm 2/\sqrt{N}]$$

where  $a_f$  is the age at full recruitment defined above  
 $\bar{a}$  is the mean age of all fully recruited fish sampled (i.e. with age  $a \geq a_f$ )  
 $N$  is the number of fully recruited fish sampled (i.e. excluding any fish caught with ages  $a < a_f$ ).

[This formula for confidence limits may indicate greater levels of precision than actually pertain, because it depends on the assumptions of the underlying model

Figure 13: Three selectivity curves, (a) representing knife-edge selection with constant asusceptibility to fishing gear after the age-at-first-capture,  $t_c$ , while (b) and (c) show more general selectivity functions of the logistic form which depend on the parameter  $\delta$  [see equation (5.7)], and for which  $a_r$  represents the age at 50% selectivity.

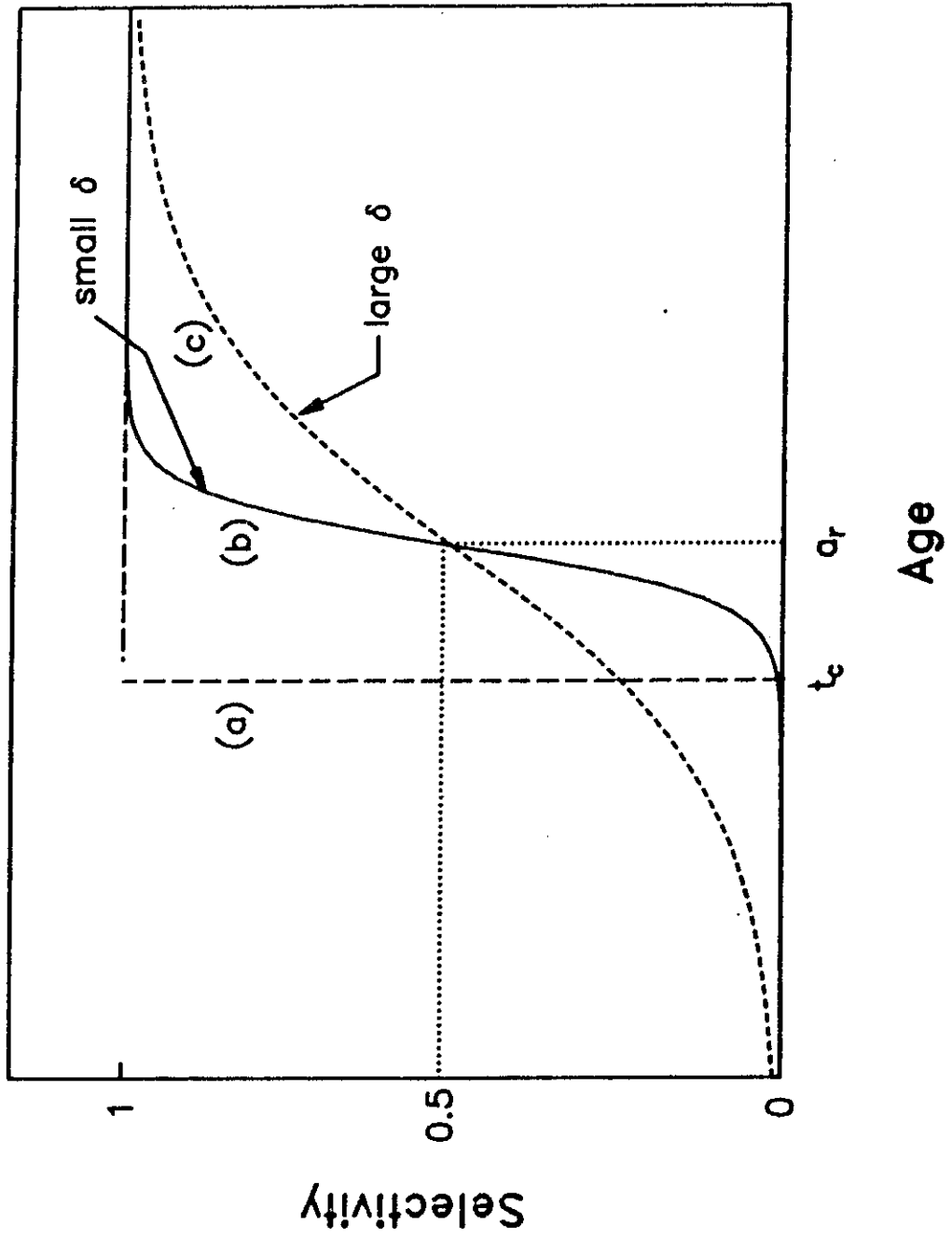
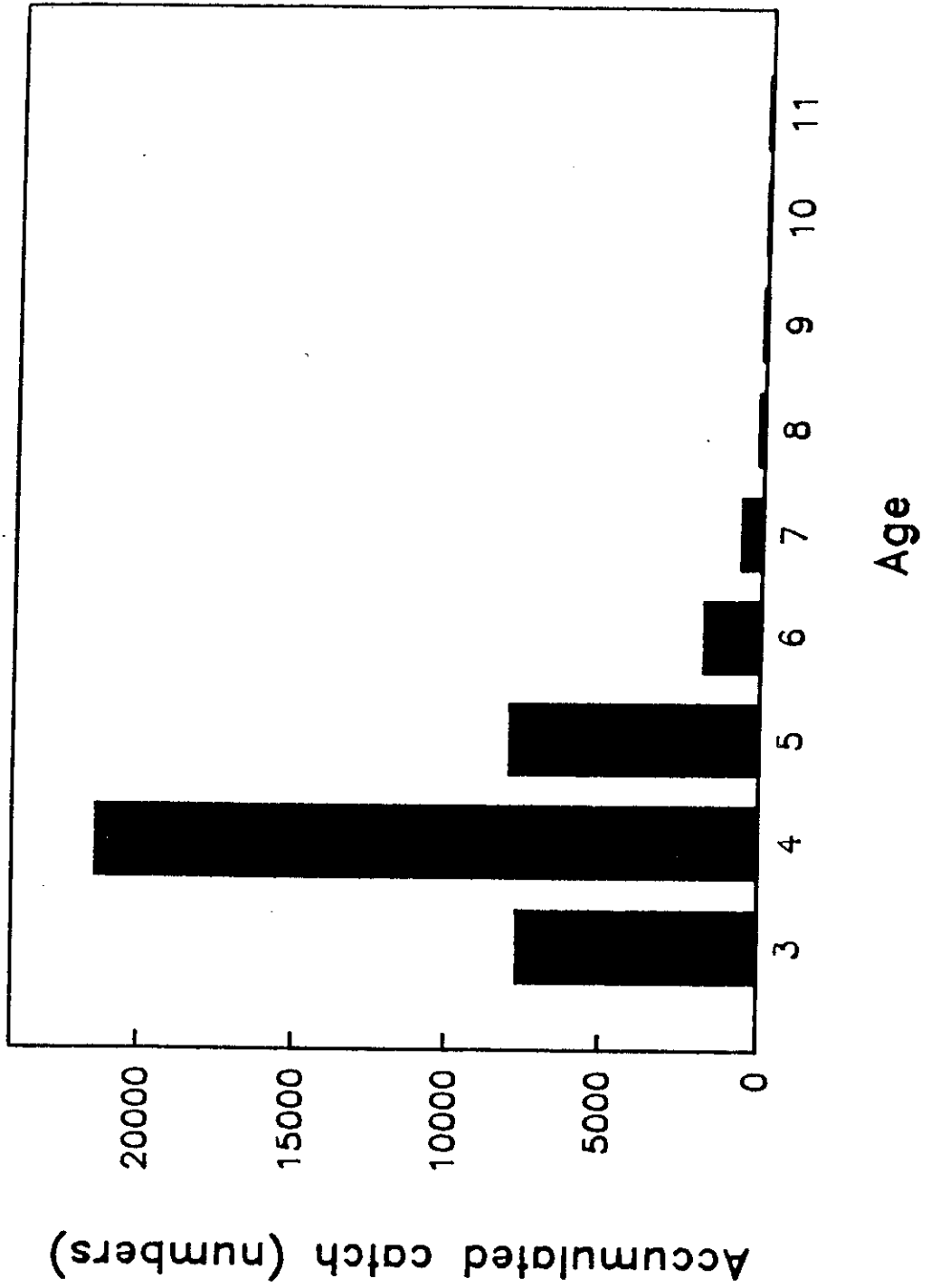


Figure 14: Accumulated elf catch-by-age in Natal angling competitions from 1956 to 1976 - the "catch-curve".



being met exactly. In reality, the annual variations in recruitment, fishing mortality and natural mortality may not be negligibly small. A more realistic variance estimate may be provided by the Jack-knife approach (Chapter III), using equation (3.6), for which the complete set of age data from each year is treated as the sampling unit.]

Table 5 gives the results of sample calculations of  $\hat{F}$  for the Natal elf data, using the two methods described above. Two different values of  $t_c$  and  $a_f$  are used, to provide an indication of the sensitivity of the methods to the values chosen for these parameters. The standard errors of  $\hat{F}$  for the second method [equation (5.8)] are noticeably small and are, in fact, considerably smaller than the difference in the estimates themselves produced by changing the value of  $a_f$ . In such a case the comments of the preceding paragraph would appear to be relevant: the assumptions of the underlying model are being violated and the error estimates are therefore not reliable. The estimates of  $F$  itself using equation (5.8) are similar to those deduced from VPA (see Table 4), suggesting that this estimation procedure may nevertheless be superior to the  $\ln(C)$  vs age regression, which yields much lower estimates of  $F$ . It should also be noted that estimates of  $F$  increase when either  $t_c$  or  $a_f$  is changed from 3 to 4, suggesting that 3-year-old elf are not fully recruited (available) to the fishery, i.e.  $S_3 < 1$ .

A potential problem with the above approaches is that a considerable fraction of the data may be discarded as a result of restriction to ages  $a \geq a_f$ , so that the resultant estimate of  $F$  may have poor precision. Improved precision may be achieved by taking the full range of available age data into account, but this requires a more realistic model of selectivity than the knife-edge approximation of equation (5.7) above. A more flexible functional form which has the ogive shape is the logistic:

$$S_a = 1/[1 + e^{-(a - a_r)/\delta}] \quad (5.9)$$

where  $a_r$  is the age corresponding to 50% selectivity (50% recruitment)  
 $\delta$  is a parameter related to the age range over which the selectivity changes from values near 0 to values near 1 .

As  $\delta$  tends to zero, this function approaches the knife-edge selection case [curve (a) in Figure 13].

The catch-at-age vector ( $C_a$ ) can now be expressed approximately as follows:

$$C_a \propto S_a F e^{-[M(a-1) + \sum_{a'=0}^{a-1} S_{a'} F]} \quad (5.10)$$

i.e.  $C_a$  is related to the selectivity for fish of age  $a$  ( $S_a$ ), the fishing mortality  $F$  and the number of fish surviving to age  $a$ . [Note that for this model of selectivity, as age increases,  $S_a F$  approaches  $F$ , so that:

$$C_{a+1}/C_a = e^{-(M + F)}$$

i.e. the plot of  $\ln(C)$  against age has a slope equal to  $-(M + F)$  for large age.] The complete catch-curve is then fitted to equation (5.10), where  $S_a$  is given by equation (5.9). Now, however, three parameters ( $F$ ,  $a_r$  and  $\delta$ ) must be estimated from the data instead of only one ( $Z$  in equation 5.8); thus, even though more data are available, the ultimate precision of the estimate of  $F$  may not improve because of the additional parameters requiring simultaneous estimation from the data.

All of the procedures described above for estimating  $F$  are based on the additional assumption that there is no temporal trend in recruitment. If there is an increasing recruitment trend over the period considered, the estimate of  $F$  will be positively biased. More seriously, a low estimate of  $F$  (suggesting that the stock is not in danger of overexploitation) could be only a reflection of a decreasing trend in recruitment.

#### Yield-per-recruit-analysis

Determination of the appropriate age-at-first-capture for a particular stock involves a trade-off between two factors: the increase in mass of an individual fish as it gets older, and the parallel increase in the probability that it will be eaten by a predator. The Beverton-Holt model describes this trade-off in terms of the yield-per-recruit, which can be translated into total yield under the assumption that recruitment is constant from year to year. (This is a limitation of the model, in that effects such as possible recruitment failure due to depletion of the spawning biomass under heavy fishing mortality are not incorporated.)

Assuming knife-edge selection operating at the age-at-first-capture,  $t_c$ , it follows that the number  $N$  remaining in a cohort at a time  $t$  after birth is:

$$N(t) = \begin{cases} Re^{-Mt} & (0 < t < t_c) \\ Re^{-Mt_c} e^{-(M+F)(t-t_c)} & (t \geq t_c) \end{cases}$$

where  $R$  = recruitment (at birth), i.e. 0-year-class strength  
 $t_c$  = age-at-first-capture .

Using the standard Von Bertalanffy growth model with the assumption that the mass of a fish,  $w(t)$ , is proportional to the cube of its length,  $l(t)$ , gives:

$$w(t) = w_\infty [1 - e^{-\kappa(t-t_0)}]^3 .$$

Now, under steady-state assumptions, the total yield ( $Y$ ) in any year from all cohorts is the same as that from a single cohort over its whole lifetime, i.e.:

$$Y = \int_{t_c}^{\infty} FN(t).w(t)dt .$$

Evaluation of this expression (Appendix X) produces the Beverton-Holt equation for the yield-per-recruit as a function of fishing mortality and age-at-first-capture:

$$Y/R = Fw_\infty e^{-Mt_c} \left[ \frac{1}{(F+M)} - \frac{3e^{-\kappa(t_c-t_0)}}{(F+M+\kappa)} + \frac{3e^{-2\kappa(t_c-t_0)}}{(F+M+2\kappa)} - \frac{e^{-3\kappa(t_c-t_0)}}{(F+M+3\kappa)} \right]$$

and the average mass of each fish caught:

$$\bar{m} = w_\infty \left[ 1 - \frac{3e^{-\kappa(t_c-t_0)} \cdot (F+M)}{(F+M+\kappa)} + \frac{3e^{-2\kappa(t_c-t_0)} \cdot (F+M)}{(F+M+2\kappa)} - \frac{e^{-3\kappa(t_c-t_0)} \cdot (F+M)}{(F+M+3\kappa)} \right] .$$

The combined assumptions of a Von Bertalanffy growth curve and mass proportional to length cubed are usually used because they produce a simple formula for yield-per-recruit in terms of  $F$  and  $t_c$ . The analysis may be conducted for alternative growth curves and mass-length relationships (Hughes and Punt 1988), and also for non-knife-edge selectivity, but in these cases numerical integration procedures (e.g. Simpson's Rule) may have to be utilised.

Plotting  $Y/R$  against  $F$  for a given value of  $t_c$  facilitates determination of the maximum yield-per-recruit and the corresponding fishing mortality  $F_{MSYR}$ , alternatively termed  $F_{max}$ . This is not necessarily, however, the most economically practical level at which to stabilise fishing effort. In Figure 15(a), for example, in order to achieve MSYR when  $t_c = 4$ , an infinite  $F$  and hence an infinite fishing effort would be required. In this case  $F_{0.1}$  (see following section) may be a more appropriate target fishing mortality level.

The value of  $t_c$  should be chosen carefully. If  $t_c \leq t_m$  (age at 50% maturity), the spawning biomass-per-recruit (SB/R) approaches zero as  $F$  becomes large [see Figure 15(b)]. Recruitment may be impaired once SB/R drops below a certain (imprecisely known) critical level, probably somewhere between 20% and 50% of its average unexploited level. On the other hand, if  $t_c > t_m$ , SB/R approaches a non-zero limit for large  $F$ ; this does not alone, however, necessarily constitute sufficient protection against overexploitation, as:

- i) this non-zero limit may nevertheless be below the critical level necessary to maintain pre-exploitation recruitment levels of the stock,
- ii) heavy fishing reduces the effective number of year-classes in the population, making it less stable in the face of recruitment fluctuations and
- iii) when  $F$  is large, "growth overfishing" (a poorer yield as a result of catching too many fish before they have had a chance to grow sufficiently) may still occur in species for which  $Y/R$  is a decreasing function of  $F$  for large  $F$ .

It should thus be clear that a management strategy based entirely on the setting of minimum size limits above the age at maturity will not necessarily provide sufficient protection for the population; further, a yet smaller minimum size limit than this may be more appropriate in certain circumstances. What is important in making management decisions based on yield-per-recruit analyses, is

TABLE 5: Estimates of  $F$  for *P. saltatrix* using catch-at-age data from Table 1 and a catch-curve analysis approach (see Chapter V). Standard errors on the estimates are given in brackets.

			$\hat{F}$
$\ln(C)$ vs age regression	$t_c = 3$	slope = -0.95	0.44 (.17)
	$t_c = 4$	slope = -0.88	0.48 (.23)
$Z = \ln[1 + 1/(\bar{a} - a_f)]$	$a_f = 3$	$\bar{a} = 3.47$	0.74 (.0056)
	$a_f = 4$	$\bar{a} = 4.40$	0.85 (.0096)



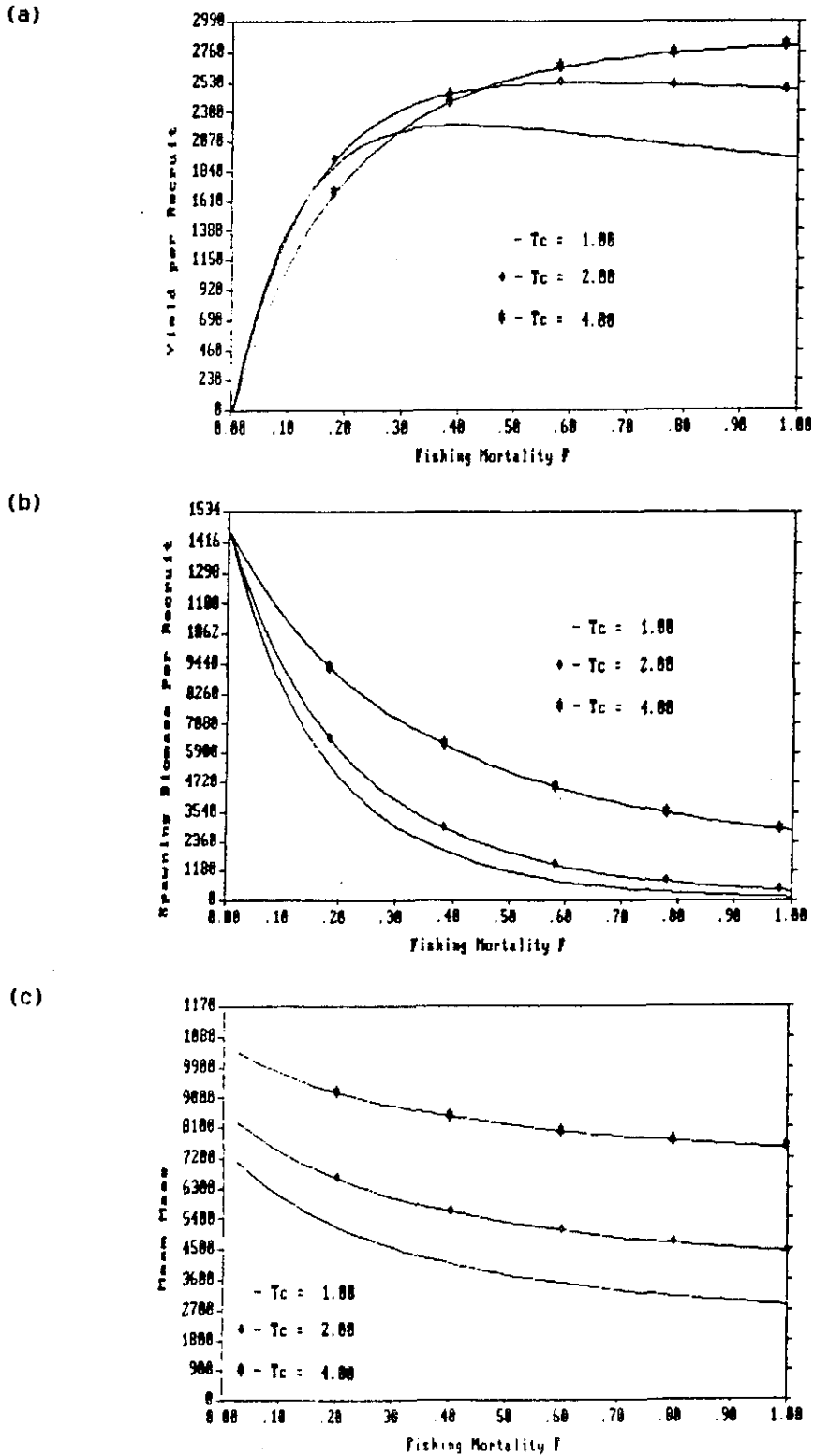


Figure 15:

Yield-per-recruit (a), spawning biomass-per-recruit (b) and average mass of fish caught (c) for a hypothetical population with  $t_m = 4$ , plotted against  $F$  for various values of  $t_c$  (age-at-first-capture) [scaled-down examples of output obtainable using the computer program PC-YIELD (Hughes and Punt 1988).]

to consider the behaviour of the spawning biomass-per-recruit [Fig. 15(b)] in addition to the yield-per-recruit, with  $F$  (for various values of  $t_c$ ). (The formula for spawning biomass-per-recruit is given in Appendix X.)

By analogy with the rule-of-thumb deduced in the previous Chapter from the Schaefer model, viz. that stock biomass should not be allowed to fall to less than half its average unexploited level, it would seem undesirable to select a target fishing mortality value that corresponds to the spawning biomass-per-recruit falling to less than half its pristine level (i.e.  $F = 0$ ). The heavier the exploitation becomes (i.e. the larger the value of  $F$ ), the earlier in its life a given fish is likely to be captured. Thus as  $F$  increases, a decrease in the average length and hence the average mass of fish landed is to be expected [Fig. 16(c)]. This average mass,  $\bar{m}(F)$ , can therefore be used to indicate whether overfishing ( $F > F^*$ , where  $F^*$  is the target fishing mortality level, e.g.  $F_{MSYR}$ ) is occurring, but this method is fairly crude as  $\bar{m}$  is relatively insensitive to changes in  $F$ . The observed  $\bar{m}$  may also vary substantially due to fluctuations in year-class size, so that care must be taken in the interpretation of such an index; for example, an increase in the value of  $\bar{m}$  may be a reflection of poor recruitment rather than a reduction in fishing mortality.

### The $F_{0.1}$ strategy

The discussion thus far has touched on the importance of economic factors only with reference to the specification of optimum fishing mortality levels. The  $F_{0.1}$  strategy attempts to take into account the decrease in marginal yield per recruit,  $d(Y/R)/dF$ , as effort (assumed proportional to fishing mortality,  $F$ ), and hence fishing costs, increases.  $F_{0.1}$  (or  $F_{max}$ ) is defined as the effort level at which the marginal yield per recruit drops to 10% of its value for the unexploited stock (Gulland 1968, Gulland and Boerema 1973):

$$\left. \frac{d(Y/R)}{dF} \right|_{F = F_{0.1}} = 0.1 \left. \frac{d(Y/R)}{dF} \right|_{F = 0}$$

Increasing  $F$  above  $F_{0.1}$  tends to provide very small additional returns in terms of  $Y/R$ , although costs may increase considerably. For  $Y/R$  curves that are monotonic with respect to  $F$  (i.e.  $F_{MSYR} = \infty$ ),  $F_{0.1}$  is a far more realistic management target at which to aim.

Application of yield-per-recruit analysis as described here (i.e. using the Von Bertalanffy growth model) to South African linefish, is discussed by Hughes (1986). A more general procedure is now available, however, in *PC-YIELD* (Hughes and Punt 1988), which provides a wide choice of growth curves as a basis for the analysis.

## VI. STANDARD ANALYSES FOR LINEFISH DATA

The Workshop agreed to recommend the following checklist as a guideline to be followed in the analysis of catch-effort data and in the presentation of yield-per-recruit analyses, for linefish species. The reason for suggesting this form of standardisation is to facilitate the interpretation and comparison of results; it is not intended to discourage the reporting of alternative calculations as well. The symbol "\*" in the margin indicates the procedures provided by *PC-YIELD* (Hughes and Punt 1988), while the symbol "\*\*\*" denotes those previously available in *PC-VONBERT* (Punt 1987), which was circulated to Workshop participants in July 1987.

### CPUE Analysis

- (1) Alternative bases for taking into account species-direction of effort:
  - (a) ignore possible species-direction of effort
  - (b) species-direction of effort is proportional to the catch composition over a "suitable" time period (STP)
  - (c) use the "Penney Method", i.e. allocate all the effort to a species, provided that the species was caught in the STP
  - (d) effort is considered to be directed at a species if that species makes up more than 50% of the catch in the STP (see equation 4.2.3).
- (2) Regress CPUE against time and calculate the proportionate decline over the period spanned by the data together with its associated standard error, to give a measure of the precision of the estimate obtained.
- (3) Plot CPUE against effort to estimate the catch rate for the unexploited stock and the present relative depletion level, with their associated standard errors.
- (4) Select and discuss the two factors considered to be the most serious potential causes of bias in the analyses above for the stock under investigation.

## Yield-per-Recruit Analysis

\* (1) Fit a growth curve (e.g. Von Bertalanffy) to the age-length data. Investigate the quality of the fit, ensuring that the requisite assumptions have been met. [PC-YIELD (Hughes and Punt 1988) provides alternative growth curves if the Von Bertalanffy is inappropriate; in such a case, however, the formula of Pauly in (2) below cannot be used to estimate M.] Use a re-sampling technique such as "jack-knifing" to calculate standard errors for the parameter estimates and the lengths-at-age.

(2) Use a formula of Pauly's (1980) or that of Rikhter and Efanov (1977) to estimate M:

$$\left. \begin{aligned} \text{(a)} \quad \ln(M) &= -0.0066 - 0.279\ln(L_{\infty}) + 0.6543\ln(\kappa) + 0.4634\ln(T_0) \\ \ln(M) &= -0.2107 - 0.0824\ln(w_{\infty}) + 0.6757\ln(\kappa) + 0.4627\ln(T_0) \end{aligned} \right\} \dots \text{Pauly}$$

where  $T_0$  is mean sea temperature in  $^{\circ}\text{C}$ , with  $w_{\infty}$  in grams,  $L_{\infty}$  in cm and  $\kappa$  and  $M$  in  $\text{yrs}^{-1}$

$$\text{(b)} \quad M = 1.521/(0.7t_m) - 0.155 \quad \dots \text{Rikhter}$$

where  $t_m$  = age (in years) at which 50% of the stock is sexually mature.

[Note that the precision of estimates provided by these formulae is low; for example, 95% confidence intervals for the Pauly estimates span a range of about a third to three times the value obtained. Thus, while the formulae do at least provide estimates, the sensitivity of the results to alternative values should always also be checked - see 8(a) below.]

(3) Estimate  $Z$  from the catch-curve by using equation (5.8) and obtain  $F$  by subtracting  $M$ . (Omit low ages where curve appears convex.) Calculate the 95% confidence interval for the estimate of  $F$ . If data are available for a long period over which a relative effort index is also available, obtain an estimate of  $Z$  for each year or group of years using the regression procedure above, and plot this against the associated (average) effort index; the vertical axis intercept of the  $Z$  vs Effort regression fit provides an estimate of  $M$ .

- (4) Find  $t_c$  (current), i.e. age-at-first-capture (for initial analyses, assume that it corresponds to the maximum of the catch-curve).
- (5) Find  $t_m$  (the age at which 50% of individuals are sexually mature).
- (6) Plot the following as functions of  $F$ :
- \*(a)  $Y/R$  (using the Beverton-Holt equations)
  - \*(b)  $\bar{m}$  (average mass of individuals caught)
  - \*(c) spawning biomass per recruit.
- (7) Calculate the following and compare them with the current value of  $F$ :
- (a)  $F_{MSYR}$  (if it exists)
  - \*(b)  $F_{0.1}$
  - (c)  $F_{S'=0.5}$ , where  $S' = \frac{\text{spawning biomass per recruit}}{\text{unexploited spawning biomass per recruit}}$ .
- (8) Repeat steps (6) and (7) for different values of  $M$  and  $t_c$ , including:
- \*(a)  $M = M_{\text{estimated}} \pm 0.1$ #
  - \*(b)  $t_c = t_m$ , and  $t_c = t_c(\text{current}) \pm 1$ .
- # A more appropriate value than 0.1 may be required if the estimate of  $M_{\text{estimated}}$  is very large or very small. In such a case a variation factor equal to (perhaps) 20% of the estimate is recommended.
- \*\* Alternatively, plot isopleths of  $Y/R$  (for specific values of  $F$ ),  $\bar{m}$  and spawning biomass-per-recruit on the  $(M, t_c)$  plane.

An updated version of *PC-YIELD* (Hughes and Punt 1988) is planned, which (hopefully) will:

- (a) incorporate a selectivity function into the yield-per-recruit analysis and
- (b) plot an ellipsoidal joint probability density distribution for parameters  $a$  and  $b$  of the Schnute growth equations, in order to provide a visual basis to aid the process of rejecting possible alternative growth models for the comparative analysis described in Appendix V.

## VII. TAG-RECAPTURE METHODS

## General aspects

The greatest value of tag-recapture exercises to a linefishery probably lies in the information these give about migration patterns and stock boundaries. The data collected may also provide a direct estimate of the size of a population. If it can be assumed that:

- i) the population is closed, i.e. its size remains constant (no recruitment, immigration or emmigration occurs in the time between tagging and the recapture sample),
- ii) tagging does not affect the catchability or natural mortality of an animal,
- iii) no tag-loss occurs in the time between marking and the recapture sample,
- iv) the tagged and recapture samples are taken at random from the population (i.e. all animals are equally likely to be caught) and
- v) all tags recovered are reported,

then the following equation (Seber 1973) gives an estimate (the "Petersen Estimate") of the population size:

$$\hat{N} \cong nT/r \quad (7.1)$$

$$\text{c.v.}(\hat{N}) \cong 1/\sqrt{r}$$

where  $\hat{N}$  = estimated population size  
 $T$  = number of fish tagged  
 $r$  = number of tags recovered  
 $n$  = recapture sample size .

The recapture sample is usually provided by normal fishing operations, e.g. a commercial catch or angling competition. Equation (7.1) gives a biased estimate of population size if  $n$ ,  $r$  or  $T$  are small ( $< 10$ ); some improvement may be achieved by using the formula:

$$\hat{N} = \frac{(n+1)(T+1)}{r+1} - 1 \quad (7.2)$$

[see Seber (*op. cit.*), p.60]. Under the above assumptions, an approximate estimate for  $F$  is provided by the tag recovery rate:

$$\begin{aligned} \hat{F} &\cong r/T & (7.3) \\ \text{c.v.}(\hat{F}) &\cong 1/\sqrt{F} \end{aligned}$$

In practice, however, these assumptions are unlikely to be met and the resulting estimates will tend to be biased. Potential sources of bias should always be carefully considered; it may be possible to assign a direction to some of these, even if their magnitudes cannot be quantified. For example, initial tagging mortality is hardly measurable in practice, but the population estimate obtained by assuming that this mortality is zero may nevertheless provide an informative "upper limit" in some sense.

#### Analysis of tag-recapture data for the Natal elf resource

Tag-recapture data for the Natal elf (shad) population may be used to illustrate the application of these techniques. If it is assumed that fishing and natural mortality are the only important effects operating, the following equations allow calculation of the catch within one year of tagging (see equation 5.4):

$$\begin{aligned} C &= FN_0[1 - e^{-(F+M)}]/(F+M) & (7.4) \\ \hat{C}/N_0 &= r/T \end{aligned}$$

[The exact form of the catch equation (from which  $F$  is calculated) is used here, rather than the approximate form given by equation (7.3), which is accurate only for small values of  $F$  and  $M$ .]

Given that a tag-recapture experiment for this species yields  $T = 5111$  and  $r = 402$ , and taking  $M$  to be 0.4 (Van der Elst pers. comm.), these equations provide a value for  $\hat{F}$  of 0.1. The average return time [ $t_R = 1/Z$  (see Appendix XI), where  $Z = F + M$ ] for these tags, however, is 36 days, giving estimates for  $Z$  and  $F$  of 10.1 and 9.7  $\text{yr}^{-1}$  respectively. Thus two substantially different estimates of  $F$  are obtained. In an attempt to account for this discrepancy, two further effects - tag loss ( $L \text{ yr}^{-1}$ ) and emigration ( $E \text{ yr}^{-1}$ ) may be included, so that:



$$\frac{dN}{dt} = -(F + L + M + E)N = -ZN$$

and

$$C/N_0 = F \left[ \frac{1 - e^{-(F + M + L + E)t}}{(F + M + L + E)} \right]$$

$$\cong F/(F + M + L + E) \quad [\text{for } (F + M + L + E) \text{ large}].$$

Now experiments have shown that the average tag-loss time ( $\bar{t}_L$ ) for captive shad is 145 days, which yields an estimate for  $L = 1/\bar{t}_L$  of  $(145/365)^{-1} = 2.5 \text{ yr}^{-1}$ . Thus the equations to be solved for the remaining unknowns, F and E are:

$$F/(F + M + L + E) = t/T = 402/5111$$

$$t_R = 1/(F + M + L + E) = 36/365$$

$$L = 2.5$$

$$M = 0.4$$

which give:

$$F = 0.8$$

$$E = 6.4$$

For the untagged fish then:

$$\frac{dN}{dt} = -(F + M + E)N$$

so that:

$$C/N_0 = \frac{F}{(F + M + E)} [1 - e^{-(F + M + E)t}]$$

$$= 0.11$$

which suggests that fishing removes approximately 11% of the population per year. Substitution into equation (7.4) above shows that this rate of exploitation corresponds to an equivalent average annual F value of  $\sim 0.14 \text{ yr}^{-1}$ . This figure provides an appropriate reference value for comparison with M, for example, in assessing whether the stock is overexploited. In this case, although the "true" value of  $F = 0.8 \text{ yr}^{-1}$  may seem large, it is applicable to a section of the tagged population only (in which the fish have neither lost their tags nor emigrated); hence the calculation of an equivalent average annual F to provide a value more representative of the impact on the population as a whole. A similar analysis for

Cape shad tag-recapture results yields an estimate for the equivalent average annual  $F$  of  $0.05 \text{ yr}^{-1}$ , which corresponds to approximately 4% of the population being caught each year.

In summary then, three different approaches to assessing the status of the Natal elf resource have provided three very divergent results:

- i) tag-recapture analysis indicates that  $F \approx 0.14$ , suggesting that the stock is underexploited ( $F \ll M$ )
- ii) the C/E vs Time regression (Chapter IV) suggests that the stock is (only just) overexploited ( $B = 0.42K$ ), although this conclusion becomes stronger when the C/E vs E regression is also taken into account
- iii) catch-age-composition analysis (VPA - Chapter V) yields a value for  $F$  of  $0.7 - 1.0 \text{ yr}^{-1}$  (see Table 4), suggesting that the stock is heavily overexploited ( $F \gg M$ ).

These three conclusions are incompatible. The next steps in the modelling process are to re-examine the reliability of the assumptions on which each conclusion is based, to develop hypotheses which could reconcile the differences, and to design experiments which would test such hypotheses. Elf tagged in the field, for example, possibly suffer large initial tagging mortality, lose their tags more easily than in the laboratory study, or are easier prey for natural predators; all of these effects would result in the equivalent annual average  $F$  estimate from the tag-recapture analysis being negatively biased. On the other hand, the emigration effect used to account for the tag-recapture results could also be influencing the catch-age analysis;  $Z$  would then include an emigration component, so that  $F$  would be smaller than the value of  $0.7-1.0 \text{ yr}^{-1}$  estimated. If this is the case, however, where is the older component of the population - why is it not available to beach anglers? Experimental off-shore fishing could perhaps help to resolve some of these questions.

## VIII. HOOK SELECTIVITY IN A LINEFISHERY

This section describes the initial steps in the formulation of a model of hook selectivity, using experimental data, which is currently under development for several commercially important species of linefish. The model is to be used to establish the potential of hook size limitation regulations for preventing the overexploitation of certain species. It incorporates the following two assumptions:

- (i) the total number of fish caught is directly proportional to the fishing effort applied and
- (ii) for each species the abundance within any cohort declines exponentially with age (a), which may be simply related to a fish's size-class expressed in terms of gape size, g.

Assumption (ii) gives:

$$N_s(\text{gape}) \propto e^{-M_s t} \quad (8.1)$$

$$\propto e^{-M'_s g}$$

where  $M_s$  is the natural mortality for species s (see Appendix XII). The equation:

$$S \propto e^{-(g - g_h)^2 / (2\sigma_h^2)} \quad (8.2)$$

defines the selectivity function for a given hook size, h; this hook is most effective in catching fish with gape size  $g_h$ , its effectiveness decreasing for both smaller and larger fish. Note that implicit in this is the important assumption that selectivity depends only on the hook size, and not on the species, i.e. individuals of two different species with the same gape size and abundance are equally likely to be caught by a given hook. A global model for the catch of all species in the experiment then follows from the general relation:

$$C = qEN \quad (8.3)$$

which, for a particular age group (a), may be written:

$$C_a = q_0 S_a E N_a \quad (8.4)$$

$$= E(q_0 N_a) S_a$$

or alternatively, in terms of gape-size ( $g$ ):

$$C_g = E(q_0 N_g) S_g \quad (8.5)$$

Substituting from equations (8.1) and (8.2) above and generalising, this becomes:

$$C_{hsgd} = E_{hd} \lambda_s e^{-M'_s g} e^{-(g - g_h)^2 / (2\sigma_h^2)} \quad (8.6)$$

where

$C_{hsgd}$  = the number of individuals of species  $s$  and gape size  $g$  caught with hook  $h$  on day  $d$

$E_{hd}$  = the time spent fishing using hook  $h$  on day  $d$

$\lambda_s$  = a constant for each species  $s$ , related to its relative abundance.

This model has ( $4 \times s$ ) parameters (assuming that the  $g_h$ 's and  $\sigma_h$ 's are species-independent); it may be simplified, however, by noting that the optimal gape size of a fish caught with hook  $h$ ,  $g_h$ , must necessarily increase with  $h$  (big hooks catch bigger fish), so that the number of parameters can be reduced by assuming the linear relation:

$$g_h = g_1 + hg_2$$

where  $g_1$  and  $g_2$  are constants. In addition, as the size range of fish caught by a hook is also believed to increase with hook size, a further possible assumption is:

$$\sigma_h = \sigma_1 + h\sigma_2$$

so that the simplified model now becomes:

$$C_{hsgd} = E_{hd} \lambda_s e^{-M'_s g} e^{-(g - g_1 - hg_2)^2 / [2(\sigma_1 + h\sigma_2)^2]} \quad (8.7)$$

which has at most ( $2s + 4$ ) parameters.

The data set obtained from the experiment has been trimmed to contain 9 species, each represented by at least 50 individuals. A further simplification may be achieved by noting that the expected mortality rate  $M'_s$  is similar for many species and these may therefore be allocated a single, common value, leaving only

three values of  $M'$  to be estimated. Thus, instead of the ( $4s = 36$ ) parameters for the original model, a total of only 16 parameters remains:

$$\lambda_s: s = 1, \dots, 9 \quad (9)$$

$$M'_s \quad (3)$$

$$g_1, g_2 \quad (2)$$

$$\sigma_1, \sigma_2 \quad (2) .$$

It is important always to start with the simplest possible model, for two reasons:

- (i) to reduce technical problems in computing the parameter values that best fit the data and
- (ii) to ensure that the parameter estimates are as precise (low variance) as possible.

Further complexity (more parameters) should only be introduced if the simpler model provides statistically unsatisfactory fits to the data.

## IX. MULTI-SPECIES AND ENVIRONMENTAL ASPECTS

As remarked in the Introduction, the fish population dynamics models described in the preceding sections have all been of the "single-species" type, treating the stock under consideration in isolation. Harvesting a given species necessarily has an effect on its predators, however, as well as on its prey species, and thus also has an effect on harvesting of these species. (The single species approach conveniently dismisses these effects by assuming they may be regarded as random fluctuations about the single species model predictions.) Multi-species models, however, while in principle both desirable and conceivable, have associated with them serious practical difficulties and have not yet reached the stage of development at which they can be used to provide quantitative management advice. Gulland (1981) summarises the situation as follows: "The simpler models tend to produce what is intuitively obvious, or directly follows from the assumptions made and are little help in specific management situations. The more realistic and complex models ..... at present make such demands on data, and the estimation of numerous parameters that they only provide some guidelines on the qualitative effects that might be expected - and in particular some warnings that the effects may be quite different from those expected from simple single species models. They cannot as yet be used to give quantitative predictions. A small change in some of the parameters, e.g., of the predation rate of one species on another, which is well within the range of values consistent with available observations, can make a big difference to the expected results. Nevertheless, the careful use of multi-species models can provide useful guidance in managing these fisheries, and result in better decisions than those based on a single species approach". Although some advances in multi-species fisheries modelling have been made since 1981, they are not as yet sufficient to negate these general conclusions.

Pollution (or habitat degradation) may also be considered as a generalised form of multi-species effect. Perhaps the simplest way of treating such an effect mathematically is to regard it (in a single-species context) as something which decreases the carrying capacity,  $K$ . To be useful for management purposes, however, the amount of effective decrease would need to be quantified and it is difficult to see how this could be achieved.

Conflict of interests between different fisheries is another aspect of the multi-species problem. Recreational fishermen argue that purse-seining adversely

affects their catches; on the other hand, purse-seining scatters the bait-fish shoals and, as shoaling is considered to be a mechanism by which to secure protection from predators, purse-seining could therefore be argued to be increasing the attractiveness of the area for game-fish. Perhaps the predators, feeding more easily, are less interested in the fishermen's bait. The point, as before, is that a number of effects are operative and it is impossible to tell *a priori* which will dominate and what direction the net effect will have - one has to model the effects quantitatively and/or design specific experiments in order to ascertain this.

The usually substantial "noise" in fisheries data (i.e. the wide scatter of points in, say, a CPUE vs Effort or Recruits vs Parent Stock plot) is conveniently ascribed to "environmental factors". The objective of "environmental" research is to relate these deviations (at least in part) to environmental measurements - mean temperature, salinity, rainfall, upwelling, etc. Knowledge of such relationships would improve the ability to predict (which is fundamental to the ability to manage). If, for example, it were known that in years of below average temperature, recruitment values for a certain species tend to fall below the average indicated by the stock-recruitment curve, it would be appropriate to institute conservative management measures for that species, should lower than average temperatures be observed.

Unfortunately, little success has been achieved world-wide in discovering such relationships, despite long and intensive research. This suggests that they are unlikely to be simple, and identifying them will probably take many more years.

**ACKNOWLEDGEMENTS**

The authors and other Workshop participants are indebted to the staff of the Foundation for Research Development (FRD) at the University of Cape Town for their co-operation in helping to organise the Meeting and for the use of their Committee Room. The manuscript was typed by Ms Lee Goddard - to her our grateful thanks. Financial assistance for the Workshop was provided by SANCOR (CSIR).



## REFERENCES

1. Beddington, J.R. and Cooke, J.G. 1983. The potential yield of fish stocks. *FAO Fish. Tech. Pap.* 242 (47 pp).
2. Bergh, M.O. and Butterworth, D.S. 1987. Towards rational harvesting of the South African anchovy considering survey imprecision and recruitment variability. In *The Benguela and Comparable Ecosystems*. A.I.L. Payne, J.A. Gulland and K.H. Brink (Eds). *S. Afr. J. mar. Sci.* 5: 937-951.
3. Beverton, R.J.H. 1954. *Notes on the use of theoretical models in the study of the dynamics of exploited fish populations*. U.S. Fish. Lab., Beaufort, N.C., Misc. Contrib. 2 (159pp).
4. Beverton, R.J.H. and Holt, S.J. 1957. On the dynamics of exploited fish populations. *Fish Invest., Ser. II*, 19 (533 pp).
5. Butterworth, D.S. 1983. *Some aspects of the scientific information requirements for line-fish management decisions*. Internal SANCOR Line-Fish Programme Document (15 pp).
6. Butterworth, D.S. and Andrew, P.A. 1984. Dynamic catch-effort models for the hake stocks in ICSEAF Divisions 1.3 to 2.2. *ICSEAF 1984 Colln scient. Pap. int. Commn SE. Atl. Fish.* 11: 29-58.
7. Caddy, J.F. and Csirke, J. 1983. Approximations to sustainable yield for exploited and unexploited stocks. *Océanogr. trop.* 18(1): 3-15.
8. Chapman, D.G. and Robson, D.S. 1960. The analysis of a catch curve. *Biometrics* 16: 354-368.
9. Cooke, J.G. and Beddington, J.R. 1981. Biases in estimates of total mortality and how to avoid them. Paper SC/33/05, presented to the IWC Scientific Committee, June 1981 (unpublished).

10. Cooke, J.G. 1984. The relationship between the net catcher day and the gross catcher day as a unit of effort, with reference to the Norwegian minke whale fishery. *Rep. int. Whal. Commn* 34, 288-290.
11. Cooke, J.G. 1985. On the relationship between catch per unit effort and whale abundance. *Rep. int. Whal. Commn* 35: 511-519.
12. Draper, N.R. and Smith H. 1966. *Applied Regression Analysis*. John Wiley and Sons Inc., New York (709 pp).
13. Efron, B. 1981. Nonparametric estimates of standard error: the jack-knife, the bootstrap and other methods. *Biometrika* 68(3): 589-599.
14. Gulland, J.A. 1968. The concept of the marginal yield from exploited fish stocks. *J. Cons. perm. int. Explor. Mer.* 32: 256-261.
15. Gulland, J.A. 1971. The fish resources of the oceans. Revised edition of *FAO Fish. Tech. Pap.* (97): 45.
16. Gulland, J.A. 1981. An overview of applications of operations research in fishery management. In *Applied Operations Research in Fishing*. K.B. Haley (Ed). Plenum Press, New York: 125-135.
17. Gulland, J.A. and Boerema, L.K. 1973. Scientific advice on catch levels. *Fish. Bull.* 71: 325-335.
18. Hughes, G.S. 1985. *Interim Report on the Catch Statistics Sampling Strategy Evaluation Project*. Internal SANCOR Line-Fish Programme Document (200 pp).
19. Hughes, G.S. 1986. *Examining methods of fitting age/length data to the Von Bertalanffy growth curve with a view to applying a simplified version of the Beverton and Holt yield-per-recruit model*. Internal SANCOR Line-Fish Programme Document (98pp).
20. Hughes, G.S. and Punt, A.E. 1988. *PC-YIELD User Manual*. Internal SANCOR Line-Fish Programme Document (50pp).
21. ICES 1985. Report of the Working Group on Methods of Fish Stock Assessments. *ICES Cooperative Research Report No. 133* (57 pp).

22. Linhart, H. and Zucchini, W. 1986. *Model Selection*. John Wiley and Sons, New York (301pp).
23. MacCall, A.D. 1980. The consequences of cannibalism in the stock-recruitment relationship of planktivorous pelagic fishes such as *Engraulis*. IOC Workshop Report 28: 201-220.
24. MacCall, A.D. 1986. Virtual population analysis (VPA) equations for nonhomogeneous populations, and a family of approximations including improvements on Pope cohort analysis. *Can. J. Fish. Aquat. Sci.* 43(12): 2406-2409.
25. MacDonald, P.D.M. and Pitcher, T.J. 1979. Age groups from size-frequency data: a versatile and efficient method of analysing distribution mixtures. *J. Fish. Res. Board Can.* 36: 987-1001.
26. Nelder, J.A. and Mead, R. 1965. A simplex method for function minimisation. *Computer J.* 7: 308-313.
27. Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters and mean environmental temperature in 175 fish stocks. *J. Cons. perm. int. Explor. Mer.* 39(2): 175-192.
28. Pope, J.G. 1972. An investigation of the accuracy of Virtual Population Analysis using cohort analysis. *Int. Commn Northwest Atl. Fish. Res. Bull.* 9: 65-74.
29. Powel, M.J.D. 1964. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Computer J.* 7: 155-162.
30. Punt, A.E. 1987. *PC-VONBERT*. Computer software for plotting yield-per-recruit curves using a personal computer. Department of Applied Mathematics, University of Cape Town.
31. Punt, A.E. 1988. Bias and variance of catch-effort model estimation procedures used in the management of the ICSEAF hake stocks - a preliminary comparative investigation. Submitted to *ICSEAF 1988 Colln scient. Pap. int. Commn SE. Atl. Fish.* 15.

32. Ratkowsky, D.A. 1983. *Nonlinear Regression Modelling*. Marcel Dekker Inc., New York (276 pp).
33. Ricker, W.E. 1973. Linear regressions in fisheries research. *J. Fish. Res. Board Can.* 30: 409-434.
34. Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. *J. Fish. Res. Board Can. Bull.* 191 (382 pp).
35. Rikhter, V.A. and Efanov, V.A. 1977. On one of the approaches of estimating natural mortality of fish populations. *Trudy Atlant. NIRO* 73: 77-85.
36. Robson, D.S. 1966. Estimation of the relative fishing power of individual ships. *Int. Commn Northwest Atl. Fish. Res. Bull.* 3: 5-14.
37. Schaefer, M.B. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. *Inter-Amer. Trop. Tuna Commn, Bull.* 1(2): 26-56.
38. Schaefer, M.B. 1957. A study of the dynamics of a fishery for yellowfin tuna in the eastern tropical Pacific Ocean. *Inter-Amer. Trop. Tuna Commn, Bull.* 2(6): 247-285.
39. Schnute, J. 1981. A versatile growth model with statistically stable parameters. *Can. J. Fish. Aquat. Sci.* 38: 1120-1140.
40. Schnute, J. 1985. A general theory for analysis of catch and effort data. *Can. J. Fish. Aquat. Sci.* 42(3): 414-429.
41. Schnute, J. and Fournier, D. 1980. A new approach to length-frequency analysis: growth structure. *Can. J. Fish. Aquat. Sci.* 37: 1337-1351.
42. Seber, G.A.F. 1973. *The Estimation of Animal Abundance and Related Parameters*. Charles Griffin & Company Ltd., London (506 pp).
43. Von Bertalanffy, L. 1934. Untersuchungen über die Gesetzmäßigkeit des Wachstums. *I. Roux Arch. Entw Mech.*, 131: 613-652.

44. Walters, C. 1986. *Adaptive Management of Renewable Resources*. MacMillan Publishing Company, New York (374pp).
45. Zar, J.H. 1974. *Biostatistical Analysis*. Prentice-Hall Inc., New Jersey (620 pp).

## APPENDIX I

## Functional Regression Analysis

A number of articles in the scientific literature address the question of which type of linear regression analysis is appropriate for application to a particular problem [see Ricker (1973)]. The most recent, and also perhaps the most authoritative and useful comments on the subject in the context of fisheries science, are to be found in ICES (1985). While a few pertinent points have been extracted for discussion here, readers are referred to the original Report for more details. In the first instance, it should be emphasised that, where the data already show a well-defined linear relationship, the particular type of regression analysis chosen is likely to make little practical difference to the parameter estimates obtained. It is only where there is considerable fluctuation in the data (and hence a small correlation coefficient) that the choice of procedure becomes important.

In functional regression analysis, the quantities of interest are the values of the parameters of the (linear) fit to the data. For the case where the underlying relationship is exact, but both  $x$  and  $y$  variables are subject to independent measurement error, estimates of these parameters can be determined using the "maximum likelihood" estimation principle. If the underlying exact model is:

$$Y = \alpha + \beta X$$

with the observed variables:

$$\begin{aligned} x_i &= X_i + \eta_i & \eta_i &\sim N(0, \sigma_x^2) \\ y_i &= Y_i + \varepsilon_i & \varepsilon_i &\sim N(0, \sigma_y^2) \end{aligned}$$

then, if  $a$  is an estimate of  $\alpha$  and  $b$  an estimate of  $\beta$ , and the data are  $\{(x_i, y_i): i = 1, \dots, n\}$ :

$$\begin{aligned} a &= \bar{y} - b\bar{x} \\ b &= p + \sqrt{p^2 + \lambda} \end{aligned}$$

where

$$p = \frac{\sum_{i=1}^n (y_i - \bar{y})^2 - \lambda \sum_{i=1}^n (x_i - \bar{x})^2}{2 \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}$$

$$\lambda = \sigma_y^2 / \sigma_x^2$$

$\bar{x}$ ,  $\bar{y}$  are the means of the  $x_i$ 's and  $y_i$ 's respectively.

Note that the value of  $b$  always lies between two extremes:

$\lambda = \infty$ , corresponding to negligible error in the  $x$  measurements, for which the formula for  $b$  is the same as for the predictive regression of  $y$  on  $x$  (see Appendix IV) and

$\lambda = 0$ , corresponding to negligible error in the  $y$  measurements, for which the formula for  $b$  is the same as for the predictive regression of  $x$  on  $y$ .

A value of:

$$\lambda = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

provides an estimate of the slope,  $\beta$ , which is the geometric mean of the estimates for  $\lambda = 0$  and  $\lambda = \infty$ ; this is the "geometric mean functional regression" recommended by Ricker (1973) for certain situations. An approximate coefficient of variation for  $b$  is given by:

$$\text{c.v.}(b) \approx \sqrt{(1 - r^2)/(nr^2)}.$$

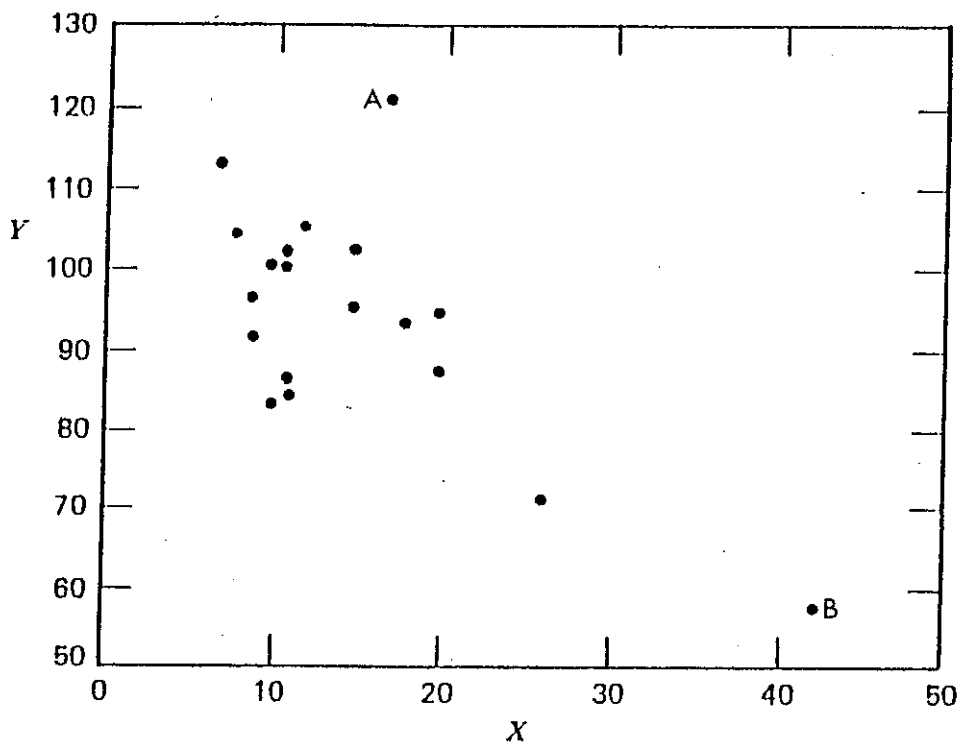
Application of these formulae requires a value for  $\lambda$ , the ratio of the error variances in the two variables. (A similar technique may be applied if the absolute value of either of these error variances, rather than their ratio, is available.) ICES (1985) emphasises the point that " $\lambda$  should be estimated from whatever knowledge is available on the nature and source of the data"; further it concludes that if no information in regard to the value of  $\lambda$  is available, the geometric mean functional regression "may be the best last resort of the desperate ... such a last resort should, however, be avoided if at all possible, and no further justification can be given for it than expedience", except that it is "a 'central' estimate, and therefore unlikely to be far wrong".

## APPENDIX II

## Outliers and Influential Points

Figure A2.1 below shows the scatter in the data obtained when a variable  $X$  is plotted against a second variable,  $Y$ . Both the points  $A$  and  $B$  indicated are clearly outliers. Point  $B$  is also an influential point, as, due to its relative isolation, it will have a substantial effect on the position of a curve fitted to the data. Point  $A$ , on the other hand, is not an influential point, as its effect is likely to be swamped by that of the neighbouring observations.

Figure A2.1: A linear regression showing an outlier which is not an influential point ( $A$ ) and one ( $B$ ) which may well be influential [after Draper and Smith (1966), p.170].





## APPENDIX III

## The Runs Test for Serial Correlation

A "run" is defined here as an unbroken sequence of identical elements within a series containing two different types of element. The series:

+ + - - - + + + + - - + + +

for example, contains 5 runs, the elements of which represent the signs of consecutive residuals from a fitted curve. To test the null hypothesis that the residuals show no positive serial correlation (too few runs), let  $n_1$  and  $n_2$  be the number of negative and positive residuals respectively, and  $u$  the total number of runs in the series:  $n_1 = 5$ ,  $n_2 = 9$ ,  $u = 5$ .

Table A3.1 gives the cumulative probability (under the null hypothesis) of obtaining the observed number of runs,  $u$ , for any  $(n_1, n_2)$  pair with  $n_1 \leq n_2 \leq 10$ . If this probability is less than a chosen critical probability,  $\alpha$ , the null hypothesis is rejected at the  $(1 - \alpha) \times 100\%$  level of significance. From Table A3.1 then, we have  $P(u = 5) = 0.119$ , which is larger than 0.05. We therefore accept the null hypothesis that the residuals show no positive serial correlation at the 5% level of significance. To test the opposite case of too many runs, the null hypothesis of no negative serial correlation is rejected if 1 minus the appropriate probability given in the Table is less than  $\alpha$ .

For either  $n_1$  or  $n_2 > 10$ ,  $u$  is distributed approximately normally with mean:

$$\mu = \frac{2n_1n_2}{n_1 + n_2} + 1$$

and variance:

$$\sigma^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}$$

and the standard normal deviate,  $Z$ , is given by:

$$Z = \begin{cases} \frac{(u - \mu + 0.5)}{\sigma} & \text{for the lower-tailed test} \\ & \text{(positive serial correlation)} \\ \frac{(u - \mu - 0.5)}{\sigma} & \text{for the upper-tailed test} \\ & \text{(negative serial correlation)} \end{cases}$$

(Draper and Smith 1966). In this case Table A3.2 gives the probability of obtaining the observed  $Z$  value under  $H_0$  for the lower- and upper-tailed tests. As before, if this probability is less than  $\alpha$ , we reject  $H_0$  and accept the alternative hypothesis that the residuals are correlated.

Table A3.1: Cumulative distribution of the total number of runs,  $u$ , in samples of sizes  $(n_1, n_2)$  [after Draper and Smith (1966), Table 3.1].

| $(n_1, n_2) u$ | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    | 11    | 12    | 13    | 14    | 15    | 16    | 17    | 18    | 19    | 20    |
|----------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| (2, 3)         | 0.200 | 0.500 | 0.900 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (2, 4)         | 0.133 | 0.400 | 0.800 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (2, 5)         | 0.095 | 0.333 | 0.714 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (2, 6)         | 0.071 | 0.286 | 0.643 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (2, 7)         | 0.056 | 0.250 | 0.583 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (2, 8)         | 0.044 | 0.222 | 0.533 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (2, 9)         | 0.036 | 0.200 | 0.491 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (2, 10)        | 0.030 | 0.182 | 0.455 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (3, 3)         | 0.100 | 0.300 | 0.700 | 0.900 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (3, 4)         | 0.057 | 0.200 | 0.543 | 0.800 | 0.971 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (3, 5)         | 0.036 | 0.143 | 0.429 | 0.714 | 0.929 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (3, 6)         | 0.024 | 0.107 | 0.345 | 0.643 | 0.881 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (3, 7)         | 0.017 | 0.083 | 0.283 | 0.583 | 0.833 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (3, 8)         | 0.012 | 0.067 | 0.236 | 0.533 | 0.788 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (3, 9)         | 0.009 | 0.055 | 0.200 | 0.491 | 0.745 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (3, 10)        | 0.007 | 0.045 | 0.171 | 0.455 | 0.706 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |       |
| (4, 4)         | 0.029 | 0.114 | 0.371 | 0.629 | 0.886 | 0.971 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |       |
| (4, 5)         | 0.016 | 0.071 | 0.262 | 0.500 | 0.786 | 0.929 | 0.992 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |
| (4, 6)         | 0.010 | 0.048 | 0.190 | 0.405 | 0.690 | 0.881 | 0.976 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |
| (4, 7)         | 0.006 | 0.033 | 0.142 | 0.333 | 0.606 | 0.833 | 0.954 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |
| (4, 8)         | 0.004 | 0.024 | 0.109 | 0.279 | 0.533 | 0.788 | 0.929 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |
| (4, 9)         | 0.003 | 0.018 | 0.085 | 0.236 | 0.471 | 0.745 | 0.902 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |
| (4, 10)        | 0.002 | 0.014 | 0.068 | 0.203 | 0.419 | 0.706 | 0.874 | 1.000 |       |       |       |       |       |       |       |       |       |       |       |
| (5, 5)         | 0.008 | 0.040 | 0.167 | 0.357 | 0.643 | 0.833 | 0.960 | 0.992 | 1.000 |       |       |       |       |       |       |       |       |       |       |
| (5, 6)         | 0.004 | 0.024 | 0.110 | 0.262 | 0.522 | 0.738 | 0.911 | 0.976 | 0.998 | 1.000 |       |       |       |       |       |       |       |       |       |
| (5, 7)         | 0.003 | 0.015 | 0.076 | 0.197 | 0.424 | 0.652 | 0.854 | 0.955 | 0.992 | 1.000 |       |       |       |       |       |       |       |       |       |
| (5, 8)         | 0.002 | 0.010 | 0.054 | 0.152 | 0.347 | 0.576 | 0.793 | 0.929 | 0.984 | 1.000 |       |       |       |       |       |       |       |       |       |
| (5, 9)         | 0.001 | 0.007 | 0.039 | 0.119 | 0.287 | 0.510 | 0.734 | 0.902 | 0.972 | 1.000 |       |       |       |       |       |       |       |       |       |
| (5, 10)        | 0.001 | 0.005 | 0.029 | 0.095 | 0.239 | 0.455 | 0.678 | 0.874 | 0.958 | 1.000 |       |       |       |       |       |       |       |       |       |
| (6, 6)         | 0.002 | 0.013 | 0.067 | 0.175 | 0.392 | 0.608 | 0.825 | 0.933 | 0.987 | 0.998 | 1.000 |       |       |       |       |       |       |       |       |
| (6, 7)         | 0.001 | 0.008 | 0.043 | 0.121 | 0.296 | 0.500 | 0.733 | 0.879 | 0.966 | 0.992 | 0.999 | 1.000 |       |       |       |       |       |       |       |
| (6, 8)         | 0.001 | 0.005 | 0.028 | 0.086 | 0.226 | 0.413 | 0.646 | 0.821 | 0.937 | 0.984 | 0.998 | 1.000 |       |       |       |       |       |       |       |
| (6, 9)         | 0.000 | 0.003 | 0.019 | 0.063 | 0.175 | 0.343 | 0.566 | 0.762 | 0.902 | 0.972 | 0.994 | 1.000 |       |       |       |       |       |       |       |
| (6, 10)        | 0.000 | 0.002 | 0.013 | 0.047 | 0.137 | 0.288 | 0.497 | 0.706 | 0.864 | 0.958 | 0.990 | 1.000 |       |       |       |       |       |       |       |
| (7, 7)         | 0.001 | 0.004 | 0.025 | 0.078 | 0.209 | 0.383 | 0.617 | 0.791 | 0.922 | 0.975 | 0.996 | 0.999 | 1.000 |       |       |       |       |       |       |
| (7, 8)         | 0.000 | 0.002 | 0.015 | 0.051 | 0.149 | 0.296 | 0.514 | 0.704 | 0.867 | 0.949 | 0.988 | 0.998 | 1.000 | 1.000 |       |       |       |       |       |
| (7, 9)         | 0.000 | 0.001 | 0.010 | 0.035 | 0.108 | 0.231 | 0.427 | 0.622 | 0.806 | 0.916 | 0.975 | 0.994 | 0.999 | 1.000 |       |       |       |       |       |
| (7, 10)        | 0.000 | 0.001 | 0.006 | 0.024 | 0.080 | 0.182 | 0.355 | 0.549 | 0.743 | 0.879 | 0.957 | 0.990 | 0.998 | 1.000 |       |       |       |       |       |
| (8, 8)         | 0.000 | 0.001 | 0.009 | 0.032 | 0.100 | 0.214 | 0.405 | 0.595 | 0.786 | 0.900 | 0.968 | 0.991 | 0.999 | 1.000 | 1.000 |       |       |       |       |
| (8, 9)         | 0.000 | 0.001 | 0.005 | 0.020 | 0.069 | 0.157 | 0.319 | 0.500 | 0.702 | 0.843 | 0.939 | 0.980 | 0.996 | 0.999 | 1.000 | 1.000 |       |       |       |
| (8, 10)        | 0.000 | 0.000 | 0.003 | 0.013 | 0.048 | 0.117 | 0.251 | 0.419 | 0.621 | 0.782 | 0.903 | 0.964 | 0.990 | 0.998 | 1.000 | 1.000 |       |       |       |
| (9, 9)         | 0.000 | 0.000 | 0.003 | 0.012 | 0.044 | 0.109 | 0.238 | 0.399 | 0.601 | 0.762 | 0.891 | 0.956 | 0.988 | 0.997 | 1.000 | 1.000 | 1.000 |       |       |
| (9, 10)        | 0.000 | 0.000 | 0.002 | 0.008 | 0.029 | 0.077 | 0.179 | 0.319 | 0.510 | 0.681 | 0.834 | 0.923 | 0.974 | 0.992 | 0.999 | 1.000 | 1.000 | 1.000 |       |
| (10, 10)       | 0.000 | 0.000 | 0.001 | 0.004 | 0.019 | 0.051 | 0.128 | 0.242 | 0.414 | 0.586 | 0.758 | 0.872 | 0.949 | 0.981 | 0.996 | 0.999 | 1.000 | 1.000 | 1.000 |

Source: Adapted from Frieda S. Swed and C. Eisenhart, "Tables for testing randomness of grouping in a sequence of alternatives," *Annals of Mathematical Statistics*, 14, 1943, 66-87.



## APPENDIX IV

## Predictive Linear Regression Formulae

Data:  $\{(x_i, y_i) : i = 1, \dots, n\}$

Model:  $y_i = \alpha + \beta x_i + e_i \quad e_i \sim N(0, \sigma^2)$

If  $a$  is an estimate of  $\alpha$ ,  $b$  an estimate of  $\beta$  and  $s^2$  an estimate of  $\sigma^2$ , then:

$$a = \bar{y} - b\bar{x}$$

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$s^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{(n - 2)}$$

where  $\bar{x}$  and  $\bar{y}$  are the means of the  $x$  and  $y$  values respectively .

Variances of the parameter estimates are estimated by:

$$\text{var}(b) = s^2 / \sum_{i=1}^n (x_i - \bar{x})^2$$

$$\text{var}(a) = s^2 \sum_{i=1}^n x_i^2 / [n \sum_{i=1}^n (x_i - \bar{x})^2] .$$

These calculations may be simplified by use of the following identities:

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2$$

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - n\bar{y}^2$$

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} .$$

## APPENDIX V

## Schnute's Generalised Growth Model (Schnute 1981)

Case 1:  $a \neq 0$ ;  $b \neq 0$

$$y(t) = \left[ y_1^b + (y_2^b - y_1^b) \frac{1 - e^{-a(t - t_1)}}{1 - e^{-a(t_2 - t_1)}} \right]^{1/b} \quad (\text{A5.1})$$

$y_1$ ,  $y_2$ ,  $a$  and  $b$  are the parameters of the model, with  $y_1$  the value of  $y(t)$  at time  $t = t_1$  and  $y_2$  the value of  $y(t)$  at  $t = t_2$  for  $y_2 > y_1 > 0$  and  $t_2 > t_1$ .

(a)  $a > 0$ ;  $b > 0$  (Generalised Von Bertalanffy curve)

$$y(t) = y_{\infty} [1 - e^{-a(t - t_0)}]^p$$

(i)  $p = 1$  ( $b = 1$ ) : Putter No. 1 (Specialised Von Bertalanffy growth)

(ii)  $p = 3$  ( $b = \frac{1}{3}$ ) : Putter No. 2

(b)  $a > 0$ ;  $b < 0$  (Richards' growth)

$$y(t) = y_{\infty} [1 + \frac{1}{p} e^{-a(t - t_0)}]^{-p}$$

(i)  $p = 1$  ( $b = -1$ ) : Logistic growth

(c)  $a < 0$ ;  $b > 0$

(i)  $b = 1$  : Exponential growth

$$y(t) = \alpha + \beta e^{at}$$

Case 2:  $a \neq 0$  ;  $b = 0$

$$y(t) = y_1 \exp[\ln(y_2/y_1) \frac{1 - e^{-a(t - t_1)}}{1 - e^{-a(t_2 - t_1)}}] \quad (\text{A5.2})$$

(a)  $a > 0$ ;  $b = 0$  (Gompertz curve)

$$y(t) = y_{\infty} \exp[-e^{-ag(t - t^*)}]$$

Case 3:  $a = 0$ ;  $b \neq 0$

$$y(t) = [y_1^b + (y_2^b - y_1^b) \frac{(t - t_1)}{(t_2 - t_1)}]^{1/b} \quad (\text{A5.3})$$

(I)  $b = 1$  : Linear form

$$y(t) = \alpha(t - t_0)$$

(II)  $b = \frac{1}{2}$  : Quadratic form

$$y(t) = (\alpha + \beta t)^2$$

Case 4:  $a = 0$  ;  $b = 0$

( $t^{\text{th}}$  Power curve)

$$y(t) = y_1 \exp[\ln(y_2/y_1) \frac{(t - t_1)}{(t_2 - t_1)}] \quad (\text{A5.4})$$

$$= \alpha \beta^t$$

Using the Schnute equations to determine alternative curve fits to size-at-age data

Schnute's generalised model [equations (A5.1 - A5.4)] encompasses many of the standard growth curves characteristically used in the analysis of size-at-age data (e.g. Von Bertalanffy, Gompertz, Logistic), each of which corresponds to specific (sets of the) values of the two parameters,  $a$  and  $b$ . These models are defined within the 8 different regions of the  $(a,b)$  plane by the lines:  $a = 0$ ,  $b = 0$ ,  $b = 1$  and  $b = a(t_2 - t_1)/\ln(y_2/y_1)$ , each region giving rise to a curve having a particular shape. Figures A5.1(a) and (b) depict these regions and the shapes of the corresponding growth curves. [Note that on these curves [Fig. A5.1(a)] is indicated which of the parameters  $t_0$ ,  $y_{\infty}$ ,  $t^*$  and  $y^*$  are defined in each case.

### - Procedure

The observed size-at-age data are fitted to the Schnute model using the appropriate equation (A5.1, 2, 3 or 4) and assuming a particular error model; errors are most commonly either additive or multiplicative:

$$y_i = \hat{y}_i + \varepsilon_i \quad (\text{additive})$$

$$y_i = \hat{y}_i e^{\varepsilon_i} \quad (\text{multiplicative})$$

where  $\hat{y}_i$  is the predicted size  
 $y_i$  is the observed size  
 $e_i \sim N(0, \sigma^2)$  are the errors.

Estimation of the parameters  $y_1$ ,  $y_2$ ,  $a$  and  $b$  involves minimisation of the sum of squares function:

$$SS(y_1, y_2, a, b) = \sum_{i=1}^n [y_i(\text{obs}) - \hat{y}_i(y_1, y_2, a, b)]^2$$

for the case of additive errors, and

$$SS(y_1, y_2, a, b) = \sum_{i=1}^n [\ln\{y_i^{(\text{obs})}\} - \ln\{\hat{y}_i(y_1, y_2, a, b)\}]^2$$

for the case of multiplicative errors. Comparison of the estimated values of  $a$  and  $b$  with those characteristic of the standard growth curves allows the identification of alternative, simpler models which may fit the data more parsimoniously. This procedure is illustrated in the following example [from Schnute (1981)].

#### Example:

Weight-at-age data for the species *Coregonus artedii* fitted to the 4-parameter Schnute equation (A5.1), provide estimates for  $a$  and  $b$  of 0.58 and -0.19 respectively. As the estimated value of  $b$  is close to zero, the Gompertz model ( $b = 0$ ) is a possible alternative. The difference in the fit of these two models may be compared using Analysis of Variance under the null hypothesis that  $b = 0$ , i.e. that the Gompertz model provides an adequate fit to the data. The  $F$  statistic:

$$F_{(1,6)}^{0.05} = \frac{(4.231 - 398.0)/1}{398.0/6}$$
$$= 0.378$$

is computed for a one-tailed test at the 5% significance level, where the number of degrees of freedom ( $p-q$ ,  $n-p$ ) depends on the number of parameters ( $p$  and  $q$  for the Schnute and Gompertz models respectively) and the number of data points ( $n = 10$  in this case). As the calculated value of  $F$  is much smaller than the critical value of 5.99, the null hypothesis cannot be rejected, i.e. the Gompertz model provides an adequate fit to the data set.

Table A5.1 gives the parameter estimates obtained by fitting the Schnute and Gompertz models as well as a third model, the Putter No. 2, chosen because the data are in the form of weight-at-age. Visual comparison of the sums of squared errors, SS, for this model and for the Gompertz (this is possible as both have the same number of parameters) reveals that the Gompertz model is superior (smaller sum of squares).



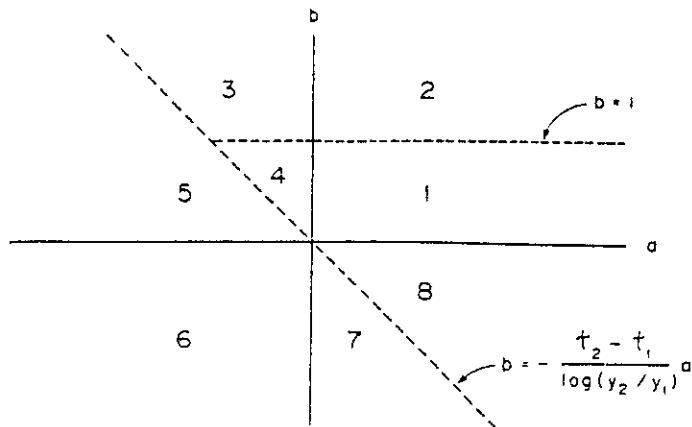
Table A5.1: Results of fitting three growth curves to the weight-at-age data for the species *Coregonus artedii*. Parameters not shown are undefined [after Schnute (1981)].

|                          | Schnute | Gompertz | Putter No. 2 |
|--------------------------|---------|----------|--------------|
| Sum of squares(SS)       | 398.0   | 423.1    | 615.3        |
| $y_1$ ( $t_1 = 2$ )      | 97.8    | 95.1     | 90.3         |
| $y_2$ ( $t_2 = 11$ )     | 538.2   | 539.9    | 543.3        |
| a                        | 0.58    | 0.54     | 0.46         |
| b                        | -0.19   | 0        | 0.33         |
| $t_0$                    | -       | -        | 0.27         |
| $t^*$                    | 3.22    | 3.05     | 2.68         |
| $y_{\infty}$             | 544.1   | 547.6    | 555.5        |
| $y^*$                    | 217.7   | 201.4    | 164.6        |
| Region In Figure A5.1(a) | 8       | 8        | 1            |

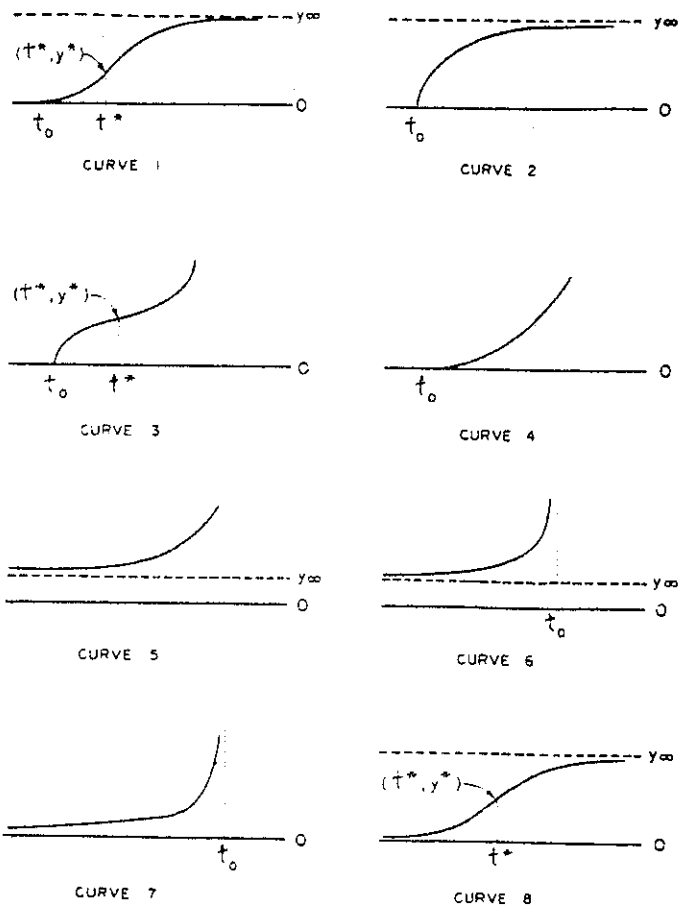
Note: *PC-YIELD* (Hughes and Punt 1988) provides the facility for determining the most appropriate curve-fit for any data set from a wide selection of growth curves, using the procedure described above.

Figure A5.1: (a) Set of 8 regions in the (a,b) plane defined by four lines [after Schnute (1981), Figure 1], (b) the 8 characteristic growth curves appropriate to a parameter pair (a,b) in each of the 8 regions shown in (a) [after Schnute (1981), Figure 2]. All curves represent size (y) against age (t).

a)



b)



## APPENDIX VI

## The Bootstrap Method of Variance Estimation

In order to illustrate the basic principles underlying this method, the example of estimating the standard error (s.e.) of a sample mean is used. A sample (the "original sample") of size  $n$  is available from some distribution  $F$ , the mean of which is estimated by the sample mean,  $\bar{x}$ . The s.e. of  $\bar{x}$  may be estimated by selecting, with replacement, a large number of random samples

$\{(X_1^j, X_2^j, \dots, X_n^j); j = 1, \dots, N\}$  (referred to here as "bootstrap samples") from the original sample and computing an estimate of the sample mean,  $\bar{x}^j$  for each. The bootstrap estimate of the s.e. of the sample mean is then given by:

$$\text{s.e.} = \sqrt{\frac{\sum_{j=1}^N (\bar{x}^j - \bar{x})^2}{N - 1}} \quad (\text{A6.1})$$

$$\text{where } \bar{x} = \frac{1}{N} \sum_{j=1}^N \bar{x}^j .$$

A more complex situation involving, for example, estimation of s.e.'s on the Von Bertalanffy growth curve parameter estimates ( $\hat{x}_\infty, \hat{k}$  and  $\hat{t}_0$ ) requires some alterations to the above method in order to take the regression-like nature of the estimation procedure into account. If a parameter for which an estimate of precision is required is denoted by  $p$  and its estimate by  $\hat{p}$ , the "conditioned" bootstrap method estimates the s.e. of  $\hat{p}$  by replacing  $\bar{x}^j$  by  $\hat{p}^j$  and  $\bar{x}$  by  $\bar{p}$  in equation (A6.1) to give:

$$\text{s.e.} = \sqrt{\frac{\sum_{j=1}^N (\hat{p}^j - \bar{p})^2}{N - 1}} \quad (\text{A6.2})$$

$$\text{where } \bar{p} = \frac{1}{N} \sum_{j=1}^N \hat{p}^j .$$

$\hat{p}^j$  is the estimate of  $p$  obtained from the  $j$ 'th bootstrap sample.

In this case, however, the method described above for generating the bootstrap samples cannot be used, as these samples must be conditioned upon the distribution of the values of the independent variable (in the case of the Von Bertalanffy growth curve these are the age values) in the original sample. A simple, though *ad hoc* method of conditioning is to ensure that each sample consists of the same number of observations at each realisation of the independent variable as there are in the original sample. [If conditioning is not used, serious biases may result - see Punt (1988) for the outcome of using an unconditioned bootstrap method (the "naive" bootstrap) in a regression problem.]

Table A6.1 provides fictitious examples of two different sample distributions of 20 length-at-age pairs obtained by such conditioned sampling. The growth curve is then fitted to each bootstrap sample in turn, providing each time a new parameter estimate:  $\hat{p}^1, \hat{p}^2, \dots, \hat{p}^N$ . This procedure should be performed several times using an increasing number of samples,  $N$ , until the variability in the estimates of the variance for each of the parameter estimates is sufficiently small (this depends on the degree of precision required); the most appropriate choice for  $N$  therefore depends on the amount of variation in the original sample.

**Table A6.1:** Distribution of 20 fictitious values of length-at-age for an original sample and two bootstrap samples obtained from it by conditioned replacement sampling. Note that each bootstrap sample has the same number of observations of length at each age as had the original sample ("conditioned sampling").

|                       |                                   | LENGTH-AT-AGE |     |     |     |     |     |
|-----------------------|-----------------------------------|---------------|-----|-----|-----|-----|-----|
| AGE                   |                                   | 1             | 2   | 3   | 4   | 5   | 6   |
| Original Sample       |                                   | 200           | 300 | 380 | 400 | 510 | 550 |
|                       |                                   | 210           | 270 | 350 | 430 | 480 |     |
|                       |                                   | 220           | 310 | 360 | 420 |     |     |
|                       | $(X_1, X_2, \dots, X_{20})$       | 190           |     | 370 | 440 |     |     |
|                       |                                   | 200           |     | 385 |     |     |     |
| <hr/>                 |                                   |               |     |     |     |     |     |
| Bootstrap<br>Sample 1 |                                   | 200           | 300 | 385 | 430 | 510 | 550 |
|                       |                                   | 190           | 270 | 360 | 420 | 510 |     |
|                       |                                   | 220           | 300 | 380 | 400 |     |     |
|                       | $(X_1^1, X_2^1, \dots, X_{20}^1)$ | 190           |     | 360 | 400 |     |     |
|                       |                                   | 220           |     | 350 |     |     |     |
| <hr/>                 |                                   |               |     |     |     |     |     |
| Bootstrap<br>Sample 2 |                                   | 210           | 300 | 370 | 400 | 510 | 550 |
|                       |                                   | 190           | 310 | 385 | 420 | 480 |     |
|                       |                                   | 190           | 270 | 385 | 430 |     |     |
|                       | $(X_1^2, X_2^2, \dots, X_{20}^2)$ | 200           |     | 350 | 430 |     |     |
|                       |                                   | 200           |     | 370 |     |     |     |

## APPENDIX VII

## The Schaefer Model

The Schaefer model is based on the assumption that the growth of a population may be described by a logistic curve, represented by the equation:

$$\frac{dB}{dt} = rB(1 - B/K) - C$$

where

|     |   |                       |
|-----|---|-----------------------|
| $r$ | = | intrinsic growth rate |
| $K$ | = | carrying capacity     |
| $B$ | = | biomass               |
| $C$ | = | catch rate .          |

(Note: It is frequently forgotten that  $r$ ,  $C$ ,  $E$  and  $MSY$  are rates; typically they have units  $yr^{-1}$ .)

Under equilibrium conditions (i.e.  $dB/dt = 0$ ) it follows that:

$$C = rB(1 - B/K)$$

with  $MSY = rK/4$ . Assuming that CPUE is linearly related to the biomass  $B$ , i.e.:

$$C/E = qB$$

where

|     |   |                          |
|-----|---|--------------------------|
| $q$ | = | catchability coefficient |
|-----|---|--------------------------|

and substituting for  $C$ , it follows that:

$$C/E = \alpha - \beta E$$

with  $\alpha = qK$  and  $\beta = q^2K/r$ . This gives  $MSY = \alpha^2/4\beta$ .

Values for  $\alpha$  and  $\beta$  can be obtained from a linear regression of  $C/E$  against  $E$ . Note, however, that this method of estimating  $\alpha$  and  $\beta$  may produce very poor estimates if the equilibrium assumption is invalid.

## APPENDIX VIII

## Proof of the Relationship between CPUE (CGD) and CND (Cooke 1984)

Let CND be the catch-per-net-vessel-day (the catch per vessel, counting only those vessel-days on which fish are actually caught), CGD the catch-per-gross-vessel-day (the catch rate over all days fished), ND the net-catch-days (the number of vessel-days on which fish are caught) and GD the total number of vessel-fishing days. Assuming that the catches are independent and Poisson-distributed, with  $\lambda = \text{CGD}$ , then:

$$\begin{aligned} \text{ND} &= \text{GD}[1 - \text{Prob}(\text{no catches})] \\ &= \text{GD}(1 - e^{-\lambda}) \end{aligned}$$

and therefore  $\text{ND}/\text{GD} = 1 - e^{-\lambda}$ . Also:

$$\begin{aligned} \text{CGD} &= \text{C}/\text{GD} \\ &= (\text{C}/\text{ND}) \cdot (\text{ND}/\text{GD}) \\ &= \text{CND}(1 - e^{-\lambda}) \end{aligned}$$

But:

$$\text{CGD} = \lambda \quad (\text{by assumption})$$

i.e.  $\lambda = \text{CND}(1 - e^{-\lambda})$

and  $\text{CND} = \lambda/(1 - e^{-\lambda})$   
 $\approx \lambda/(\lambda - \lambda^2/2)$   
 $\approx 1 + \lambda/2$ .

Therefore:

$$\begin{aligned} \lambda &\approx 2(\text{CND} - 1) \\ &\propto (\text{CND} - 1) \end{aligned}$$

It follows that  $\text{CGD} \propto (\text{CND} - 1)$ . For  $\text{CND} \leq 1.5$ , the error of the approximation is limited to 15%.

## APPENDIX IX

Using an Age-Length Key to Transform the Length-Frequency Distribution  
of the Catch Into an Age-Frequency Distribution

An age-length key is an  $(n \times m)$  matrix in which the entries,  $a_{ij}$ , represent the number of fish in each of  $n$  length-classes and  $m$  age-classes.

$$\begin{array}{c}
 \text{LENGTH-} \\
 \text{CLASSES} \\
 \left[ \begin{array}{cccc}
 a_{11} & a_{12} & a_{13} & \dots \dots \dots a_{1m} \\
 a_{21} & a_{22} & & \\
 a_{31} & & & \\
 \vdots & & & \\
 \vdots & & & \\
 \vdots & & & \\
 a_{n1} & \dots \dots \dots & & a_{nm}
 \end{array} \right]
 \end{array}$$

AGE-CLASSES

(Note that many of the entries in this matrix will be zero.) Normalising with respect to each length-class,  $n$ :

$$A_{ij} = a_{ij} / (a_{i1} + a_{i2} + \dots + a_{im})$$

e.g.  $A_{12} = a_{12} / (a_{11} + a_{12} + \dots + a_{1m})$  etc.

yields the normalised matrix,  $A_{nm}(\ell, t)$ . Now the length-frequency distribution,  $C(\ell)$ , of the catch may be expressed as a vector with  $n$  columns:

$$C_n(\ell) = (C_1, C_2, C_3, \dots, C_n)$$

where  $C_i$  represents the number of fish caught in each length class,  $i$ . Simple matrix multiplication then provides an estimate of the age-frequency distribution of the catch:



$$\begin{aligned}
\hat{C}_m(t) &= \sum_n C_n(x) \cdot A_{nm}(x, t) \\
&= (C_1, C_2, \dots, C_n) \cdot \begin{bmatrix} A_{11} & A_{12} & A_{13} & \dots & \dots & A_{1m} \\ A_{21} & A_{22} & & & & \\ \vdots & & & & & \\ A_{n1} & & & & & A_{nm} \end{bmatrix} \\
&= \begin{bmatrix} C_1 A_{11} + C_2 A_{21} + \dots + C_n A_{n1} \\ \vdots \\ C_1 A_{1m} + C_2 A_{2m} + \dots + C_n A_{nm} \end{bmatrix} \\
&= \begin{bmatrix} C'_1 \\ C'_2 \\ \vdots \\ C'_m \end{bmatrix}
\end{aligned}$$

where  $C'_j = C_1 A_{1j} + C_2 A_{2j} + \dots + C_n A_{nj}$ .

## APPENDIX X

## Derivation of the Beverton-Holt Yield-Per-Recruit Equations

Yield in terms of mass (rather than numbers) of fish is given by:

$$Y_m = \int_{t_c}^{\infty} FN(t) \cdot w(t) dt$$

where

$$N(t) = \begin{cases} Re^{-Mt} & (0 < t < t_c) \\ Re^{Ft_c - (M+F)t} & (t \geq t_c) \end{cases} \quad (\text{A10.1})$$

$$w(t) = w_{\infty} [1 - e^{-\kappa(t - t_0)}]^3 \quad (\text{A10.2})$$

and  $R$  is the recruitment. Substitution gives:

$$\begin{aligned} Y_m &= \int_{t_c}^{\infty} FRe^{Ft_c - (M+F)t} \cdot w_{\infty} [1 - e^{-\kappa(t - t_0)}]^3 dt \\ &= FRw_{\infty} e^{Ft_c} \int_{t_c}^{\infty} e^{-(M+F)t} \cdot [1 - e^{-\kappa(t - t_0)}]^3 dt . \end{aligned}$$

Integrating:

$$Y_m = RFw_{\infty} e^{-Mt_c} \left[ \frac{1}{(M+F)} - \frac{e^{-\kappa(t_c - t_0)}}{(M+F+\kappa)} + \frac{e^{-2\kappa(t_c - t_0)}}{(M+F+2\kappa)} - \frac{e^{-3\kappa(t_c - t_0)}}{(M+F+3\kappa)} \right] \quad (\text{A10.3})$$

and dividing  $Y_m$  by  $R$  provides the desired result.

Yield in terms of the number of fish caught is given by:

$$\begin{aligned}
 Y_n &= \int_{t_c}^{\infty} FN(t)dt \\
 &= \int_{t_c}^{\infty} FRe^{-Mt_c} \cdot e^{-(M+F)(t-t_c)} \cdot dt \\
 &= FRe^{-Mt_c} / (M+F) .
 \end{aligned} \tag{A10.4}$$

Now the average mass,  $\bar{m}$ , of a fish caught is given by  $Y_m/Y_n$  so that, from equations (A10.3) and (A10.4), it follows that:

$$\begin{aligned}
 \bar{m} &= w_{\infty} \left[ 1 - \frac{e^{-\kappa(t_c - t_0)} \cdot (M+F)}{(M+F+\kappa)} \right. \\
 &\quad \left. + \frac{e^{-2\kappa(t_c - t_0)} \cdot (M+F)}{(M+F+2\kappa)} - \frac{e^{-3\kappa(t_c - t_0)} \cdot (M+F)}{(M+F+3\kappa)} \right] . \tag{A10.5}
 \end{aligned}$$

The spawning biomass is defined as the total biomass of fish older than the age-at-maturity,  $t_m$ :

$$\text{SB} = \begin{cases} \int_{t_m}^{\infty} N(t) \cdot w(t) dt & (t_m \geq t_c) \\ \int_{t_m}^{t_c} N(t) \cdot w(t) dt + \int_{t_c}^{\infty} N(t) \cdot w(t) dt & (t_m < t_c) . \end{cases} \tag{A10.6}$$

Integrating and dividing by R gives the spawning biomass-per-recruit:

$$\begin{aligned}
 \text{SB/R} = & \left\{ \begin{aligned} & w_{\infty} e^{-Mt_m - F(t_m - t_c)} \left[ \frac{1}{(M + F)} - \frac{e^{-\kappa(t_m - t_0)}}{(M + F + \kappa)} \right. \\ & \left. + \frac{e^{-2\kappa(t_m - t_0)}}{(M + F + 2\kappa)} - \frac{e^{-3\kappa(t_m - t_0)}}{(M + F + 3\kappa)} \right] \quad (t_m \geq t_c) \\ & w_{\infty} \left[ \frac{e^{-Mt_m} - e^{-Mt_c}}{M} - \frac{e^{\kappa t_0} [e^{-(M + \kappa)t_m} - e^{-(M + \kappa)t_c}]}{(M + \kappa)} \right. \\ & \left. + \frac{e^{-2\kappa t_0} [e^{-(M + 2\kappa)t_m} - e^{-(M + 2\kappa)t_c}]}{(M + 2\kappa)} \right. \\ & \left. - \frac{e^{-3\kappa t_0} [e^{-(M + 3\kappa)t_m} - e^{-(M + 3\kappa)t_c}]}{(M + 3\kappa)} \right] + \frac{Y_m}{FR} \quad (t_m < t_c) \end{aligned} \right. \quad (\text{A10.7})
 \end{aligned}$$

## APPENDIX XI

Derivation of the Relationship Between Mean Tag Return  
Time ( $t_R$ ) and Total Mortality, Z

If, initially, a number  $N_0$  of fish is tagged, that number,  $N(t)$ , which remains tagged and potentially catchable by fishermen at a time  $t$  thereafter decreases exponentially at the total mortality rate, Z:

$$\frac{dN}{dt} = -ZN$$

$$N(t) = N_0 e^{-Zt} .$$

Note that the "mortality" of tagged and potentially catchable fish includes all those effects which may reduce this number - not just the factors which potentially cause actual death, such as fishing and natural mortality. Thus Z may in general also include the effects of tag-loss and emigration, both of which may contribute towards a decline in N. If the tagged fish are captured at a rate:

$$\frac{dC}{dt} = FN$$

then, between the times  $t$  and  $(t + dt)$  after tagging, the number of tagged fish captured (with tag return times,  $t$ ) is equal to  $FN(t)dt$ . The average tag return time,  $t_R$ , is given by:

$$t_R = \frac{\sum_{\text{all return times, } t} (\text{Return time, } t) \times (\text{Number of tags returned at time } t)}{(\text{Total number of tags returned})}$$

which becomes:

$$\begin{aligned} t_R &= \left[ \int_0^{\infty} t \cdot FN(t) dt \right] / \left[ \int_0^{\infty} FN(t) dt \right] \\ &= \int_0^{\infty} t \cdot e^{-Zt} \cdot dt / \left[ \int_0^{\infty} e^{-Zt} \cdot dt \right] \\ &= 1/Z . \end{aligned}$$

(Note that this relationship does not require that all fish are tagged simultaneously - only that the return time for any tagged fish should measure the time between its release and its recapture.)

## RECENT TITLES IN THIS SERIES

- 48 A bibliography of seabirds in the waters of southern Africa, the Prince Edward and Tristan Groups. J Cooper and R K Brooke. December 1981. 297 pp.
- 49 \*National Geoscience Programme. The evolution of earth resource systems. SACUGS. June 1981. 42 pp.
- 50 South African Antarctic Biological Research Programme. SASCAR. July 1981. 54 pp.
- 51 South African Marine Pollution Monitoring Programme 1979-1982. R J Watling and C E Cloete (editors). July 1981. 52 pp.
- 52 \*Structural characterization of vegetation in the Fynbos Biome. B M Campbell, R M Cowling, W J Bond and F J Kruger in collaboration with D P Bands, C Boucher, E J Moll, H C Taylor and B W van Wilgen. August 1981. 19 pp.
- 53 \*A bibliography of fynbos ecology. M L Jarman, R M Cowling, R Haynes, F J Kruger, S M Pierce and G Moll. August 1981. 73 pp.
- 54 \*A description of the Benguela Ecology Programme 1982-1986. SANCOR: W R Siegfried and J G Field (editors). March 1982. 39 pp.
- 55 Trophic ecology of Lepidoptera larvae associated with woody vegetation in a savanna ecosystem. C H Scholtz. June 1982. 29 pp.
- 56 Man and the Pongola floodplain. J Heeg and C M Breen. June 1982. 117 pp.
- 57 \*An inventory of plant communities recorded in the western, southern and eastern Cape Province, South Africa up to the end of 1980. C Boucher and A E McDonald. September 1982. 58 pp.
- 58 A bibliography of African inland water invertebrates (to 1980). B R Davies, T Davies, J Frazer and F M Chutter. September 1982. 418 pp.
- 59 An annotated checklist of dung-associated beetles of the Savanna Ecosystem Project study area, Nylsvley. S Endrödy-Younga. September 1982. 34 pp.
- 60 The termites of the Savanna Ecosystem Project study area, Nylsvley. P Ferrar. September 1982. 41 pp.
- 61 \*Conservation of Ecosystems: Theory and Practice. A report on a workshop meeting held at Tsitsikamma, South Africa, September 1980. W R Siegfried and B R Davies (editors). September 1982. 97 pp.
- 62 \*A description of the Grassland Biome Project. M T Mentis and B J Huntley (editors). October 1982. 29 pp.
- 63 Description of a fire and its effects in the Nylsvlei Nature Reserve: A synthesis report. M V Gandar. October 1982. 39 pp.
- 64 Terrestrial ecology in South Africa - project abstracts for 1980-1981. December 1982. 148 pp.
- 65 \*Alien invasive vascular plants in South African natural and semi-natural environments: bibliography from 1830. V C Moran and P M Moran. December 1982. 42 pp.

## APPENDIX XII

## Derivation of the Formula for N(gape)

If it is assumed that:

$$N(a) \propto e^{-Ma}$$

where  $N$  = number of fish present  
 $M$  = natural mortality  
 $a$  = age

it follows that, for some reference age  $a_0$ :

$$N(a)/N(a_0) = e^{-M(a - a_0)} \quad . \quad (A12.1)$$

Using the Taylor Series approximation gives:

$$l(a) = l(a_0) + \frac{dl}{da} \cdot (a - a_0) \quad (A12.2)$$

and  $g(l) = g(l_0) + \frac{dg}{dl} \cdot (l - l_0)$

where  $l(a)$  = length at age  $a$   
 $g(l)$  = gape size at length  $l$   
 $\frac{dl}{da}$  = rate of increase in length with age  
 $\frac{dg}{dl}$  = rate of increase in gape with length .

Manipulating equations (A12.1) and (A12.2) to solve for  $(a - a_0)$  gives:

$$(a - a_0) = (g - g_0) \cdot \frac{dl}{dg} \cdot \frac{da}{dl} \quad .$$

Substituting into equation (A12.1), it follows that:

$$N(a)/N(a_0) = e^{-M(g - g_0) \cdot \frac{dl}{dg} \cdot \frac{da}{dl}}$$

and simplification then gives:

$$N(g) \propto e^{-M'g}$$

where  $M' = M \cdot \frac{dl}{dg} \cdot \frac{da}{dl} \quad .$

## RECENT TITLES IN THIS SERIES

- 48 A bibliography of seabirds in the waters of southern Africa, the Prince Edward and Tristan Groups. J Cooper and R K Brooke. December 1981. 297 pp.
- 49 \*National Geoscience Programme. The evolution of earth resource systems. SACUGS. June 1981. 42 pp.
- 50 South African Antarctic Biological Research Programme. SASCAR. July 1981. 54 pp.
- 51 South African Marine Pollution Monitoring Programme 1979-1982. R J Watling and C E Cloete (editors). July 1981. 52 pp.
- 52 \*Structural characterization of vegetation in the Fynbos Biome. B M Campbell, R M Cowling, W J Bond and F J Kruger in collaboration with D P Bands, C Boucher, E J Moll, H C Taylor and B W van Wilgen. August 1981. 19 pp.
- 53 \*A bibliography of fynbos ecology. M L Jarman, R M Cowling, R Haynes, F J Kruger, S M Pierce and G Moll. August 1981. 73 pp.
- 54 \*A description of the Benguela Ecology Programme 1982-1986. SANCOR: W R Siegfried and J G Field (editors). March 1982. 39 pp.
- 55 Trophic ecology of Lepidoptera larvae associated with woody vegetation in a savanna ecosystem. C H Scholtz. June 1982. 29 pp.
- 56 Man and the Pongola floodplain. J Heeg and C M Breen. June 1982. 117 pp.
- 57 \*An inventory of plant communities recorded in the western, southern and eastern Cape Province, South Africa up to the end of 1980. C Boucher and A E McDonald. September 1982. 58 pp.
- 58 A bibliography of African inland water invertebrates (to 1980). B R Davies, T Davies, J Frazer and F M Chutter. September 1982. 418 pp.
- 59 An annotated checklist of dung-associated beetles of the Savanna Ecosystem Project study area, Nylsvley. S Endrödy-Younga. September 1982. 34 pp.
- 60 The termites of the Savanna Ecosystem Project study area, Nylsvley. P Ferrar. September 1982. 41 pp.
- 61 \*Conservation of Ecosystems: Theory and Practice. A report on a workshop meeting held at Tsitsikamma, South Africa, September 1980. W R Siegfried and B R Davies (editors). September 1982. 97 pp.
- 62 \*A description of the Grassland Biome Project. M T Mentis and B J Huntley (editors). October 1982. 29 pp.
- 63 Description of a fire and its effects in the Nylsvlei Nature Reserve: A synthesis report. M V Gandar. October 1982. 39 pp.
- 64 Terrestrial ecology in South Africa - project abstracts for 1980-1981. December 1982. 148 pp.
- 65 \*Alien invasive vascular plants in South African natural and semi-natural environments: bibliography from 1830. V C Moran and P M Moran. December 1982. 42 pp.



- 66 Environmental research perspectives in South Africa. December 1982. 39 pp.
- 67 The SANCOR Estuaries Programme 1982-1986. February 1983. 43 pp.
- 68 The SANCOR Programme on Coastal Processes April 1982 - March 1988. D H Swart (editor). February 1983. 30 pp.
- 69 \*Guidelines for the management of large mammals in African conservation areas. The proceedings of an international workshop held at Olifants Camp, Kruger National Park, South Africa. A A Ferrar (editor). May 1983. 95 pp.
- 70 \*Marine linefish programme - priority species list. SANCOR. J H Wallace and R P van der Elst (editors). May 1983. 113 pp.
- 71 \*Mineral nutrients in mediterranean ecosystems. J A Day (editor). June 1983. 165 pp.
- 72 South African programme for the SCOPE project on the ecology of biological invasions. A description and research framework produced by the Task Group for Invasive Biota of the National Programme for Environmental Sciences. July 1983. 25 pp.
- 73 \*South African marine pollution survey report 1976-1979. B D Gardner, A D Connell, G A Eagle, A G S Moldan, W D Oliff, M J Orren and R J Watling. September 1983. 105 pp.
- 74 Ecological notes and annotated checklist of the grasshoppers (Orthoptera: Acridoidea) of the Savanna Ecosystem Project study area, Nylsvley. M V Gandar. November 1983. 42 pp.
- 75 \*Fynbos palaeoecology: a preliminary synthesis. H J Deacon, Q B Hendey and J J N Lambrechts (editors). December 1983. 216 pp.
- 76 \*A South African perspective on conservation behaviour - a programme description. A A Ferrar (compiler). December 1983. 34 pp.
- 77 Limnology and fisheries potential of Lake Le Roux. B R Allanson and P B N Jackson (editors). December 1983. 182 pp.
- 78 Limnology of Lake Midmar. C M Breen (editor). December 1983. 140 pp.
- 79 The limnology of the Touw River floodplain. B R Allanson and A K Whitfield. December 1983. 35 pp.
- 80 \*SANCOR: Summary report on marine research 1983. SANCOR. February 1984. 52 pp.
- 81 \*South African Antarctic Earth Science Research Programme. SASCAR. February 1984. 53 pp.
- 82 \*The SANCOR Marine Sedimentology Programme January 1984 - December 1988. I C Rust (editor). March 1984. 15 pp.
- 83 \*A description of major vegetation categories in and adjacent to the fynbos biome. E J Moll, B M Campbell, R M Cowling, L Bossi, M L Jarman and C Boucher. March 1984. 29 pp.
- 84 Environmental research perspectives in South Africa. February 1984. 77 pp.

- 85 Invasive alien organisms in the terrestrial ecosystems of the fynbos biome, South Africa. I A W Macdonald and M L Jarman (editors). April 1984. 72 pp.
- 86 \*Terrestrial ecology in South Africa - project abstracts for 1982-1983. May 1984. 198 pp.
- 87 Conservation priorities in lowland fynbos. M L Jarman. 1986. 55 pp.
- 88 A synthesis of plant phenology in the fynbos biome. Shirley M Pierce. July 1984. 57 pp.
- 89 \*Aquaculture in South Africa: A cooperative research programme. O Safriel and M N Bruton. June 1984. 79 pp.
- 90 Pipeline discharges of effluents to sea. D A Lord, F P Anderson and J K Basson (editors). October 1984. 108 pp.
- 91 \*Monitoring in South African grasslands. M T Mentis. September 1984. 55 pp.
- 92 Conservation of threatened natural habitats. A V Hall (editor). November 1984. 185 pp.
- 93 Limnological criteria for management of water quality in the Southern Hemisphere. R C Hart and B R Allanson (editors). December 1984. 181 pp.
- 94 Water quality criteria for the South African coastal zone. J A Lusher (editor). December 1984. 42 pp.
- 95 \*National Programme for Weather, Climate and Atmosphere Research Annual Report 1984/85. C W Louw (compiler). December 1984. 28 pp.
- 96 A guide to the literature on research in the grassland biome of South Africa. N M Tainton. December 1984. 77 pp.
- 97 \*South African Red Data Book - Birds. R K Brooke. December 1984. 213 pp.
- 98 Directory of southern African conservation areas. T Greyling and B J Huntley (editors). December 1984. 311 pp.
- 99 The effects of crude oil pollution on marine organisms. A C Brown. February 1985. 33 pp.
- 100 \*"SANKON: Opsommingsverslag oor mariene navorsing 1984." SANCOR. February 1985. 51 pp.
- 101A \*"Verslag van die Hoofprogram vir Navorsingsondersteuning." February 1985. 30 pp.
- 101E \*Report of the Main Research Support Programme. February 1985. 30 pp.
- 102 National Programme for Remote Sensing. Report: 1984. P J van der Westhuizen. February 1985. 50 pp.
- 103 Bibliography of marine biology in South Africa. A supplement to the 1980 edition. A C Brown. March 1985. 83 pp.
- 104 The plant communities of Swartboschkloof, Jonkershoek. D J McDonald. March 1985. 54 pp.

- 105 Simulation modelling of fynbos ecosystems: systems analysis and conceptual models. F J Kruger, P M Miller, J Miller and W C Oechel (editors). March 1985. 101 pp.
- 106 The Kuiseb environment: the development of a monitoring baseline. B J Huntley (editor). March 1985. 138 pp.
- 107 Annotated bibliography of South African indigenous evergreen forest ecology. C J Geldenhuys. May 1984. 125 pp.
- 108 \*Review of metal concentrations in southern African coastal waters, sediments and organisms. H F-K O Hennig. August 1985. 140 pp.
- 109 \*Coastal dunes of South Africa. K L Tinley. September 1985. 293 pp.
- 110 The limnology of Hartbeespoort Dam. NIWR. September 1985. 269 pp.
- 111 Management of invasive alien plants in the fynbos biome. I A W Macdonald, M L Jarman and P Beeston. (editors). October 1985. 140 pp.
- 112 \*The SANCOR Marine Pollution Research Programme 1986-1990. October 1985. 16 pp.
- 113 Alien and translocated aquatic animals in southern Africa: a general introduction, checklist and bibliography. M N Bruton and S V Merron. October 1985. 59 pp.
- 114 A synthesis of field experiments concerning the grasslayer in the savanna regions of southern Africa. T G O'Connor. October 1985. 126 pp.
- 115 \*South African marine pollution survey report 1979-1982. B D Gardner, A D Connell, G A Eagle, A G S Moldan and R J Watling. December 1985. 81 pp.
- 116 \*Basic needs in rural areas. A report on a seminar held in Cape Town on 19 February 1985. December 1985. 103 pp.
- 117 \*South African Red Data Book: Plants - fynbos and karoo biomes. A V Hall and H A Veldhuis. 1985. 144 pp.
- 118 Invasive alien plants in the terrestrial ecosystems of Natal, South Africa. I A W Macdonald and M L Jarman (editors). 1985. 88 pp.
- 119 Invasive alien organisms in South West Africa/Namibia. C J Brown, I A W Macdonald and S E Brown. 1985. 74 pp.
- 120 The impact of climate and weather on the activities of the building and construction industry in South Africa. G du Toit de Villiers (compiler). 1986. 40 pp.
- 121 Ecological research on South African rivers - a preliminary synthesis. J H O'Keefe. 1986. 121 pp.
- 122 A description of the Karoo Biome Project. R M Cowling. 1986. 42 pp.
- 123 \*SANCOR: Summary report on marine research 1985. SANCOR. 1986. 57 pp.
- 124 The karoo biome: a preliminary synthesis. Part I - Physical environment. R M Cowling, P W Roux and A J H Pieterse (editors). 1986. 114 pp.

- 125 South African Red Data Book - Terrestrial Mammals. R H N Smithers. 1986. 216 pp.
- 126 A bibliography of sandy beaches and sandy beach organisms on the African continent. R Bally. 1986. 179 pp.
- 127 Activities of the National Programmes for Ecosystems and Aquaculture Research, 1983-1985. E W Auret. 1986. 68 pp.
- 128 Historical sites at the Prince Edward Islands. J Cooper and G Avery. 1986. 80 pp.
- 129 Richards Bay effluent pipeline. D A Lord and N D Geldenhuys. 1986. 30 pp.
- 130 An assessment of the state of the estuaries of the Cape and Natal 1985/86. A E F Heydorn (editor). 1986. 39 pp.
- 131 The conservation of South African rivers. J H O'Keeffe (editor). 1986. 117 pp.
- 132 SIBEX - II: Report of the South African study in the sector (48-64°E) of the Southern Ocean. D G M Miller (editor). 1986. 47 pp.
- 133 The regional landscape: Nylsvley in perspective. P G H Frost. 1987. 26 pp.
- 134 South African Southern Ocean research programme. SASCAR. 1987. 58 pp.
- 135 Disturbance and the dynamics of fynbos communities. R M Cowling, C D le Maitre, B McKenzie, R P Prys-Jones and B W van Wilgen (editors). 1987. 70 pp.
- 136 "SANKON: Opsommingsverslag oor mariene navorsing 1986." SANCOR. 1987. 45 pp.
- 137 South African Red Data Book - Fishes. P H Skelton. 1987. 199 pp.
- 138A "Verslag van die Hoofprogram vir Navorsingsondersteuning - September 1984 - Junie 1987". 1987. 23 pp.
- 138E \*Report of the Main Research Support Programme - September 1984 - June 1987. 1987. 23 pp.
- 139 Marine Research in Natal. Proceedings of a symposium and workshop held at Durban, South Africa, 10 and 11 February 1986. A P Bowmaker, D van der Zee and J H Ridder (editors). 1987. 184 pp.
- 140 An environmental impact assessment of a proposed emergency landing facility on Marion Island - 1987. Report to the Minister of Environment Affairs. G Heymann, T Erasmus, B J Huntley, A C Liebenberg, G de F Retief, P R Condy and O A van der Westhuysen. 1987. 209 pp. (Available only from Department of Environment Affairs)
- 141 A preliminary synthesis of pollination biology in the Cape flora. A G Rebelo (editor). 1987. 254 pp.
- 142 The karoo biome: a preliminary synthesis. Part II - vegetation and history. R M Cowling and P W Roux (editors). 1987. 133 pp.
- 143 The Vaal River catchment: Problems and research needs. E Braune and K H Rogers. 1987. 36 pp.

- 144 Atlas of alien and translocated aquatic animals in southern Africa. 1988. I de Moor and M N Bruton. 320 pp.
- 145 A description of the Research Programme for Wetlands. R D Walmsley. 1988. 26 pp.
- 146 The River Research Programme. A A Ferrar, J H O'Keeffe and B R Davies. 1988. 28 pp.
- 147 \*Dictionary of forest structural terminology. C J Geldenhuys, R S Knight, S Russell and M L Jarman (editors). 1988.
- 148 SANCOR: Summary report on marine research 1987. 1988. 60 pp.
- 149 Final Report of the Bagasse Programme. J C Patterson-Jones (editor). 1988. 250 pp.
- 150 Atmospheric pollution and its implications in the Eastern Transvaal Highveld. P D Tyson, F J Kruger and C W Louw. 1988. 150 pp.
- 151 South African Red Data Book - Reptiles and Amphibians. W R Branch (editor). 1988. 235 pp.
- 152 Remote sensing in the marine environment. L V Shannon and L Y Shackleton. 1988. 54 pp.
- 153 The culture of sharptooth catfish, *Clarias gariepinus* in South Africa. T Hecht, W Uys and P J Britz (editors). 1988. 106 pp.
- 154 A forest map of southern Africa with the aid of LANDSAT imagery. D W van der Zel. 1988. 79 pp.
- 155 Research needs in the Transkei and Ciskei coastal zone. Report on a workshop held at Mpekweni, Ciskei on 3 March 1988. G M Branch and L Y Shackleton (editors). 1988. 56 pp.
- 156 The Swartkops Estuary. Proceedings of a symposium held on 14-15 September 1987 at the University of Port Elizabeth. D Baird, J F K Marais and A P Martin (editors). 1988. 107 pp.
- 157 Long-term data series relating to southern Africa's renewable natural resources. I A W Macdonald and R J M Crawford (editors). 1988.
- 158 South African Red Data Book - Butterflies. S F Henning and G A Henning. 1989. 176 pp.
- 159 Data Report of the First Cruise of the Marion Off-shore Ecological Study (MOES-I). C M Duncombe Rae. 1989. 384 pp.
- 160 A manual of mathematical techniques for linefish assessment. D S Butterworth, A E Punt, D L Borchers, J B Pugh and G S Hughes. 1989. 89 pp.

\* Out of print.