NOT TO BE CITED WITHOUT
PERMISSION OF THE AUTHORS[1]

Canadian Atlantic Fisheries
Scientific Advisory Committee

CAFSAC Research Document 88/61

NE PAS CITER SANS
AUTORISATION DES AUTEURS[1]

Comité scientifique consultatif des
pêches canadiennes dans l'Atlantique

CSCPCA Document de recherche 88/61

Collected papers on stock assessment methods.

Edited by

D. Rivard,
Fisheries Research Branch,
Department of Fisheries and Oceans,
200 Kent Street,
Ottawa, Ontario.
K1A 0E6

---

[1] This series documents the scientific basis for fisheries management advice in Atlantic Canada. As such, it addresses the issues of the day in the time frames required and the Research Documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Research Documents are produced in the official language in which they are provided to the Secretariat by the author.

[1] Cette série documente les bases scientifiques des conseils de gestion des pêches sur la côte atlantique du Canada. Comme telle, elle couvre les problèmes actuels selon les échéanciers voulus et les Documents de recherche qu'elle contient ne doivent pas être considérés comme des énoncés finals sur les sujets traités mais plutôt comme des rapports d'étape sur les études en cours.

Les Documents de recherche sont publiés dans la langue officielle utilisée par les auteurs dans le manuscrit envoyé au secrétariat.

Abstract

This publication contains six papers outlining various methods for the assessment of fish stocks. These papers review the current practices for the assessment of marine fish stocks in the Northwest Atlantic. The first paper discusses the definition of abundance indices from commercial fishing data. In that context, the various techniques used for the standardization of commercial catch rates are exposed. The estimation of stock abundance from research surveys is presented in the second paper. Current survey designs are discussed, together with the transformation of data. Biological production is discussed in a third paper, together with the use of production models for stock assessment. The fourth paper presents an overview of sequential population methods and discusses their underlying assumptions. The procedures used for making projections of catches and yield per recruit analysis are outlined in the last two papers.

Résumé

Cette publication contient six articles résumant les méthodes utilisées pour l'évaluation des stocks de poissons. Ces articles font le point sur les pratiques courantes pour l'évaluation des stocks de poissons marins dans l'Atlantique Nord-Ouest. Le premier article discute la définition des indices d'abondance établis à partir des données sur la pêche commerciale. Dans ce contexte, les diverses techniques utilisées pour la normalisation des taux de capture commerciaux sont exposées. On présente dans le second article l'estimation de l'abondance des stocks au moyen des relevés scientifiques; on discute ici les patrons d'échantillonnage utilisés, ainsi que la transformation des données. La production biologique est discutée dans le troisième article, ainsi que l'utilisation des modèles de production pour l'évaluation des stocks. Le quatrième article donne une vue d'ensemble des méthodes d'analyse séquentielle des populations et discute les hypothèses sous-jacentes. Les algorithmes utilisés pour faire la projection des prises et les analyses de rendement par recrue sont présentés dans les deux derniers articles.

CONTENT

Abundance Indices from Commercial Fishing

Abundance Indices from Research Survey Data

Biological Production and Production Models

Catch Projections

Yield per Recruit Analysis

Foreword

These essays on stock assessment methods were prepared in recent years at the request of the Canadian Atlantic Fisheries Scientific Advisory Committee (CAFSAC). Their aim is to document current practices for the assessment of marine fish stocks in the Northwest Atlantic. For an account of recent developments in the calibration of virtual population analyses, the reader is referred to the July 1988 report of the Workshop on Methods of Fish Stock Assessment (ICES C.M. 1988/Assess:26) and to CAFSAC Res. Doc. 88/29 ("An adaptive framework for the estimation of population size", by S. Gavaris). The June 1987 Report of the Working Group on Methods of Fish Stock Assessments (ICES C.M. 1987/Assess:24) also contains an update on recent developments in surplus production models.

Collected Papers on stock assessment methods

Abundance Indices from Commercial Fishing

by **S. Gavaris**

Marine Fish Division,
Biological Sciences Branch,
Department of Fisheries and Oceans,
Biological Station,
St. Andrews, New Brunswick
E0G 2X0

Introduction

Many fish stock assessment techniques in use today require some measure of population abundance. Two common methods which employ abundance estimates are general production models (Rivard 1988) and the calibration of sequential population analysis (Winters 1988). Typically, as in the case for these two methods of assessments, a relative measure of abundance over time, i.e. an index, is sufficient. One source for an abundance index is catch and effort data from the commercial fishery (Ricker 1940). Gulland (1964) demonstrated that indices of abundance based on catch rate data from the commercial fishery, gave good agreement with trends in abundance derived from age data.

Abundance indices may also be calculated using data from research surveys (Smith 1988). The survey approach has been promoted because it does not suffer from two major difficulties associated with commercial fishery catch rate indices: a) potential bias arising from non-random distribution of fishing locations; and b) the amalgamation of information from all cohorts necessitating the use of indirect methods to obtain abundance at age. It is recognized that indices based on research survey data can potentially provide more detailed and more accurate results than indices derived from commercial catch rate data although the results may not necessarily be more precise. Notwithstanding their limitations, it is still often desirable, and sometimes necessary, to examine abundance indices based on catch and effort data from the commercial fishery, as exemplified by recent stock assessments reviewed in CAFSAC, NAFO and ICES. Commonly, a series of annual abundance is required and the concepts of this chapter will be developed for a 1-year time interval. Results can be extended for other time frames.

Basic Model

Definition

The use of catch rate, say U, as a measure of abundance is based on the mathematical model:

$$U = f(B) \quad , \qquad (1)$$

where B represents the biomass. Although more complex functional forms have been explored, the most common relationship used assumes that abundance is directly proportional to catch rate:

$$U = q\ B \quad , \qquad (2)$$

where q = catchability coefficient. This model implicitly relates effort to fishing mortality.

Fish populations are composed of individuals belonging to various age-classes. Age-specific differences in factors such as distribution, both vertical and areal, and size may result in variation of the catchability coefficients by age. Two general approaches have been used to accommodate age-specific differences in the catchability coefficient. In the first method, the problem is addressed directly by applying the basic model equation to each age separately. In order to accomplish this, sufficiently detailed information must be available to enable the partitioning of catch rate data by age.

Since the age composition of the catch by various components of the fishery may vary, it is preferable to perform the partitioning before summarizing the catch rates. Although this approach is attractive from a theoretical point of view, data limitations may preclude its use. An alternative involves the application of partial recruitment factors to the biomass at age in order to derive the exploitable biomass, that portion of the stock which is being exploited by the fishery represented in the catch rate, U. Exploitable biomass would then be the appropriate variable for use in the model. In subsequent references to the basic model, it will be assumed that appropriate measures have been taken into account for the effect of age structure on the catchability coefficient.

Limitations and underlying assumptions

The model has been criticized (Paloheimo and Dickie 1964) because it represents fishing as a simple process and does not account for phenomena such as searching time, gear saturation and gear competition. These and other possible departures from the basic model, such as dependence of the catchability coefficient on biomass or the expenditure of effort without the generation of fishing mortality, were reviewed by Rothschild (1977). Although some of these situations, particularly those classed as competition models, have received some attention in the literature from a theoretical point of view, available data are rarely sufficient to entertain models which are more complex than equation 2. The most notable exception has been the inclusion of an intercept term in the model. This modification has been based on empirical evidence. *Post-hoc* explanations for the existence of an intercept have been proposed but not established.

Considerable attention has been given to the ways in which effort can be measured for various gears (Anon. 1976). The Scientific Council of the Northwest Atlantic Fisheries Organization has endorsed the use of the list presented in Table 1. Experience has shown that hours fished by trawlers is related to fishing mortality; however, the adequacy of the effort measures for the other gears as a measure of fishing mortality has not been demonstrated universally. The application of the model should be limited to fishery components in which suitable effort measures were recorded.

The model implicitly requires that fish be uniformly dispersed within a region. For species which exhibit strong schooling tendencies or limited movement, catch rates may not reflect abundance. For example, the school size may shrink as abundance decreases but, if density remains constant and the schools are located by the fishery, the catch rates may not decrease. On the other hand, the localized depletion of a portion of the stock without replenishment from immigration may result in low catch rates which overestimate the decline in abundance. Clearly then, the model is more useful for species whose members are relatively well dispersed and for fisheries which do not experience high localized exploitation rates. Although quantitative guidelines for these criteria cannot be provided, experience has shown that the model is more reliable for demersal species than for pelagic species.

A further complication associated with the measurement of catch rate arises when a fishery exploits more than one species (Ketchen 1964). Although some attempts have been made to account for this situation by modifying the model (e.g. Chikuni 1976), it is more commonly dealt with by considering only that data where the proportion of the desired species in the catch exceeds some threshold level. The determination of an appropriate threshold level is generally based on the nature of the particular fishery. Other similar procedures have also been used to distinguish a "directed" fishery, i.e. the most prominent species in the catch. The utility of the model depends on the success in defining an appropriate threshold level.

Table 1. The following list of definitions of effort measures was endorsed by the Scientific Council of the Northwest Atlantic Fisheries Organization for use in reporting of effort data.

FISHING GEARS	EFFORT MEASURES	DEFINITION
Trawls	No. of hours fished	Number of hours during which the trawl was in the water (mid-water), or on the bottom (bottom trawl), and fishing.
Boat seines Danish seines, etc.	No. of hours fished	Number of hours during which the seine was on the bottom and fishing.
Beach seines	No. of sets	Number of times the gear has been set or shot, whether or not a catch was made.
Boat dredges	No. of hours fished	Number of hours during which the dredge was on the bottom and fishing.
Surrounding nets (purse seines)	No. of sets	Number of times the gear has been set or shot, whether or not a catch was made.
Gillnets (set or drift)	No. of effort units	Length of nets expressed in 100-meter units multiplied by the number of sets made (=accumulated total length, in meters, of nets used in a given time period divided by 100).
Gillnets (fixed)	No. of effort units	Length of net expressed in 100-meter units multiplied by the number of times the net was cleared.
Traps (uncovered pounds nets)	No. of effort units	Number of days fished times the number of units hauled.
Covered pots and fyke nets	No. of effort units	Number of lifts times the number of units (=total number of units fished in a given time period).
Longlines (set or or drift)	Thousands of hooks	Number of hooks fished in a given time period divided by 1000.
Handlines (pole, troll, jig, etc.)	No. of line-days	Total number of lines used in the given time period.
Harpoons	-	(Report effort in days only)

The possible improvement in fishing success which may result from technological advances has also raised concerns. The data generated from such a fishery may not be comparable from year to year. Halliday and Doubleday (1976) applied a 4% adjustment in each year of a series to compensate for such an effect. This approach is based on ancillary information regarding the advantage provided by the change. If the technological improvement only affects certain gear categories or if both the old and new technologies are present for some time period, an alternative may be to differentiate between the two technologies in applying some of the summarizing techniques described below. Use of the model is limited to situations where the effect of technological change can be effectively removed from the catch rate series.

Catch-and-effort data from a commercial fishery are generally of a cross-sectional nature, i.e. several gear types may be used in various regions during different seasons. As was discussed earlier, it may not be possible to derive catch rate by age. In such situations, it is necessary that recruitment of fish to the various components of the fishery being analyzed, the possible gear, region and season combinations be similar in order to facilitate interpretation and summarization. It should be pointed out that, in this paper, gear is used in a broader context than usual and may apply to differences in such things as size of vessel, method of gear deployment, or home port of a fleet. Also, season is used as a convenient word to refer to time of year and is not intended to be restrictive in the sense of four seasons. Regions here are used to denote any convenient division of area.

Summarizing Catch Rate Data

As was mentioned previously, catch-and-effort data from a commercial fishery are often of a cross-sectional nature. Having reconciled the limitations of the basic model, the major task in applying the model is the summarization of the raw data to obtain an annual series of an abundance index. The cross-sectional nature of the data presents two related problems: a) neither the absolute nor relative levels of effort applied by the various components of the fishery are likely to remain constant; and b) catch rates from the different components are generally not directly comparable, i.e. the coefficient of catchability will vary from one to the other. For a discussion of the causes of variation in the coefficient of catchability, refer to Gulland (1964) and Anon. (1976).

The problem of catch rate summarization is related to the broad class of statistical methods applied to comparative studies (Anderson et al. 1980). Generally, the aim is to determine the value of the outcome variable for various levels of a treatment factor while adjusting for confounding factors. In this context, estimates of catch rate (outcome variable) are required for each year (treatment factor) while taking into account the effect of gear, region and season (confounding factors). In this case, the outcome variable is a continuous variable and the treatment is considered to be discrete. Confounding factors may be continuous or discrete but, because less restrictive assumptions are required, the discrete case is generally chosen, as will be done here. Since the treatment and confounding factors do not necessarily have any meaningful order of magnitude, the values these variables take will be referred to as categories rather than levels. For example, January, February and April would be categories in the factor season. The confounding factors are not necessarily simple; they may be a compounding of classifications which affect the outcome. For example, a category in the gear factor could be something like Canada-otter trawl-tonnage class 5. The nature and number of confounding factors and the degree of resolution within each factor are determined by the available data and the type of fishery being studied. The possible cross-classifications of the confounding factors will be referred to as components, i.e.

the component for January, Canada-otter trawl-tonnage class 5 in region I. Cross-classifications of treatment (year) and components will be referred to as cells. Often, data for a cell are only available in aggregated form, i.e. only the mean catch rate for that cell is known. All methods presented below can be applied to aggregated data, while some are flexible enough to accommodate disaggregated data.

In the trivial case, when the amount of fishing effort for each component stays constant from year to year, the total catch in each year would be a meaningful measure of abundance. Similarly, when the relative amount of effort for each component stays constant from year to year, the mean catch rate, or some other measure of central tendency such as total catch divided by total effort, could be used to reflect abundance. In general, however, when these two situations do not apply, the effect of confounding factors must be removed. The penalty for overlooking confounding factors is a biased index of abundance. On the other hand, the risk for including unnecessary factors in an analysis is reduced precision.

Before proceeding with the discussion of techniques for summarization, consider briefly the theoretical relationship between annual catch rate series from individual components. Two cases are distinguished: A) *the individual catch rate series are proportional to each other*; and B) *the individual series are not related in a simple manner*. The concept embodied by case A has arisen through experience and is intuitively appealing. Say, for example, that there are two components denoted by I and II. If the catch rates for I and II were 1.0 t/h and 1.5 t/h, respectively, in year 1, then in year 2 under case A, the expected catch rate for II would be 3 t/h, given that the catch rate for I was 2 t/h. In the context of equation 2, this catch rate pattern would be due to a constant ratio in the catchability coefficient for the various components. In case B, no discernible relationship exists. As an example, for two regions within a stock boundary the trends in abundance may show no similarity, perhaps due to low migration rates between the areas. Case B is less likely to apply to gear or season categories. It is important to differentiate between the two cases because of the implications in interpreting the results. For Case A, a single global value may be calculated to reflect the change in catch rates. For Case B, however, summarization of the change in catch rates can only be made with reference to a specific weighting scheme for each component's catch rate series. The derived index is specific to the weighting scheme used and can only be interpreted in the context of the meaningfulness of the weights. This weighting of series is not to be confused with the situation where Case A applies but, due to heteroscedasticity, the data are weighted in the analysis.

Alternatives

Standardized Weighting

This technique can be applied when observations of catch rate are available for all cells. As a result of this requirement for a full data matrix, no assumptions are required about the underlying structure. It has been noted that catch rate indices calculated on the basis of actual levels of fishing activity reflect not only changes in catch rates within components but also changes in the relative levels of activity between years for each component. To demonstrate this phenomenon, consider the simple situation with two components over 2 years. The following table shows the catch, effort and catch rate:

	YEAR 1			YEAR 2		
Component	Catch	Effort	Catch rate	Catch	Effort	Catch rate
A	1000	1000	1.00	100	50	2.00
B	1000	100	10.00	2000	100	20.00
TOTAL	2000	1100	1.82	2100	150	14.00

For each component, the catch rate doubled between years 1 and 2. Taking the mean catch rate weighted by the actual level of effort in each year, which is equivalent to the total catch divided by the total effort each year, we see that the catch rate increased by almost 10 times. This is a reflection of both the increase in catch rate for each component and the increased relative effort exerted by component B.

An obvious solution to the problem is to avoid changing the weighting scheme from year to year. The strategy for this technique, then, is to determine a set of meaningful standard weights for the various components of the fishery and apply them uniformly over all years. This can be expressed mathematically as:

$$ {}_{sw}I_y = \sum_k w_k U_{ky} \;/\; \sum_k w_k \quad , \qquad (3)$$

where ${}_{sw}I_y$ is the index for year y and where w_k corresponds to the weight for component k to be applied in all years. It will be noted that equation 3 is simply a weighted average of catch rates with the weights for each component w remaining constant for all years. A special case would be $w = 1$ for all k, i.e. the unweighted mean. If a variance for U_{ky} can be calculated, then an estimate of the variance of ${}_{sw}I_y$ is given by:

$$ \mathrm{VAR}({}_{sw}I_y) = \sum_k w_k^2 \,\mathrm{VAR}(U_{ky}) \;/\; (\sum_k w_k)^2 \quad . \qquad (4)$$

When the relationship between components is proportional, i.e. case A, the choice of standard weights does not affect the expected value of the index however the variance of the index may vary depending on the particular choice. Kalton (1968) gives some guidelines for selection of weights when the overall variance is to be minimized, but the situation may dictate departures from such strategy. For example, it may be desirable to use a long-term average pattern of fishing activity or a current pattern of fishing activity. The former might give relatively even variances over years while the latter would tend to favor the recent years. When case B applies, the usefulness of the index will depend on the success in finding meaningful weights.

A specific situation where the weighting scheme is based on other considerations is when the components pertain to regions only. In such a case, it may be assumed that the catch rate for each region is a measure of the density in the region. An approximately unbiased estimate of the total abundance would be obtained by weighting each regional series by the area for the respective region (Quinn et al. 1982). Care must be taken to ensure that the catch rate is representative for the entire area of the respective region. In this instance, meaningful results are possible for case B.

Scaled Response

In contrast to standardized weighting where the effect of differences in the level of fishing are accounted for by selecting a set of meaningful standard weights to be applied in all years, in the scaled response it is the catch rates which are manipulated. The catch rates for each classification are scaled to a common year, the reference year. Now, the catch rates can be transformed to "relative" numbers and they are all based on the same reference year. In principle, this means that for case A, these relative catch rates can be considered equivalent measures which can be averaged, the weighting being dependent on the precision of each point. Clearly it is necessary to have catch rate data from all components for at least 1 year, the reference year. This procedure has been recently described in Anon. (1981). Generally, then the catch rates are converted to relative numbers:

$$\gamma_{ky} = U_{ky} \;/\; U_{ky'} \qquad . \qquad (5)$$

where y' is the reference year. These are subsequently averaged:

$$_{sr}I_y = \sum_k w_{ky}\,\gamma_{ky} \;/\; \sum_k w_{ky} \qquad . \qquad (6)$$

Given the variance of γ_{ky}, an estimate of the variance of $_{sr}I_y$ is given by

$$\mathrm{VAR}(\,_{sr}I_y\,) = \sum_k w_{ky}{}^2\,\mathrm{VAR}(\,\gamma_{ky}\,) \;/\; (\,\sum_k w_{ky}\,)^2 \qquad . \qquad (7)$$

This method, although used mostly when data are available in all cells, is not restricted to those conditions. When case B applies, however, the results will not be meaningful unless data exist for all cells and, even then, the index will be specific to the choice of weights. Under case A the choice of weights will only affect the variance, as was the case for standardized weighting. In Anon. (1981), it was recommended that the catch for the year by that component be used for weighting. Note that the weight for each component may change from year to year. This form of the index is referred to as the inverse Paasche index (Allen 1975) by economists. Other indices commonly used by economists are the Laspeyres and Paasche forms (Allen 1975). The Paasche form is equivalent to a scaled response index where the weight for a component in the current year is the reference year's catch rate times the current year's effort, i.e. the catch which could have been taken in the reference year if the current year's effort were expended. The Laspeyres form is a mix of scaled response and standardized weighting indices. The catch rates are converted to relative numbers with respect to a reference year, similar to the scaled response approach, but the weights remain fixed for all years, as in the standardized weights method, those being the catch in the reference year for the component.

It has been noted (Anon. 1981) that converting the catch rates to relative numbers can distort the results since the catch rate in the reference year is also a random variable measured with error. This effect could be reduced by taking the average of several years as a reference; however, to implement such a technique, catch rate data must be available for all components in the years used for reference. Despite these difficulties, this method allows considerable flexibility in the choice of weights so that the precision of data within each cell can be reflected. Because the relative catch rate is the ratio of two random variables, the overall index will likely be biased; however, the extent of that bias has not been quantified.

Standardized Response

The third approach capitalizes on the proportional relationship between components and consequently only has meaning if case A applies. Typically, data from one component, the standard, is used as the explanatory variables in regressions where the other components are used as response variables. The relationships are forced through the origin to comply with the proportionality assumption of the model. Conventional scatterplots of paired observations or residual plots can be used to examine how closely the data follow this model. It is evident from the description that considerable overlap in time is required between the standard and each of the other components in order to establish these relationships. This approach has only been applied in cases with one confounding factor.

It is known that the data for the standard are not error free; therefore ordinary least squares does not provide unbiased results (Snedecor and Cochran 1967). The extent of the bias depends on the relative magnitude of the error in the explanatory variable with respect to the error in the response variable. Only if the errors in the catch rate for the standard are small with respect to the errors in the catch rate for other components will it be justifiable to ignore this bias.

The principle behind this approach is to determine the slope of the line passing through the origin for the pairs of observed catch rates between the standard and each component. The estimate of the slope can then be used to "standardize" the effort, i.e. convert the effort for each component to the effort which would have been expended had the catch been taken by the standard. Mathematically, the procedure can be expressed with the following formulas.

For each classification determine the slope using the model

$$U_k = P_k \, U_{k'} + \varepsilon \quad , \qquad (8)$$

where $U_{k'}$ is the catch rate for standard component k', P_k is the relative power of component k to the standard and ε is an independent random error. Using the relative power, then, the index can be calculated from

$$_{st}I_y = \sum_k C_k \, / \, \sum_k P_k \, f_k \quad . \qquad (9)$$

where C_k and f_k are, respectively, the catch and the effort for component k. Although this formula is not strictly equivalent to an effort weighted average standardized catch rate, it is similar, i.e. it tends to give more weight to catch rates associated with relatively greater effort. Estimators for the precision and accuracy of this index have not been derived. Application of this technique has generally been restricted to cases where the components refer only to the gear factor (e.g. Parsons et al. 1976).

An alternative approach, which could have provided the possibility of estimating the variance of this index in a relatively straightforward manner, involves the use of inverse regression (Draper and Smith 1981; Hunter and Lamboy 1981). Essentially, rather than using the slope from the line to standardize effort, standardized catch rates with associated variance could be obtained from the relationship. These could then be averaged and their variances could be used to obtain the precision of the resulting index.

Multiplicative Model

If it can be assumed that case A applies, the estimation of an index can be put in the context of a model which reflects the proportional relationships. The model should allow for any number of confounding factors and recognize the nature of the errors in the data. Further, data requirements may be relaxed such that the model can be applied even when several cells do not contain data, including the situation where all components do not overlap with the "standard".

The multiplicative model is an application of conventional linear modeling techniques, i.e. analysis of variance, to the log transformed catch rates. Due to the unbalanced nature of the data (cells without catch rate information), the analysis is more conveniently carried out using the regression approach with dummy variables as the explanatory variables. The basic model can be expressed as:

$$\widehat{\ln U} = \widehat{\ln U_{k'}} + \sum_{ij} ((\ln P_{ij})\, X_{ij})) + \varepsilon \quad , \qquad (10)$$

where $U_{k'}$ = catch rate for reference classification;
P_{ij} = relative power for category *j* in factor *i*;
X_{ij} = 1 when the observation is associated with category *j*;
= 0 otherwise.

Variations of this model appeared in Beverton and Holt (1957) and Gulland (1956). Recent papers by Gavaris (1980), Kimura (1981) and Robson (1966) give further discussion of the statistical properties of various estimators obtained from the application of this model. Robson (1966) was concerned primarily with estimation of the relative power's, the P_{ij}'s, while Gavaris (1980) and Kimura (1981) focused on obtaining catch rate indices. A catch rate index is obtained using the predicted value from the regression for a selected standard component in each year:

$${}_mI_y = \exp(\widehat{\ln U}) \exp[(\hat{\sigma}^2 - \hat{\sigma}^2_{\widehat{\ln U}})(\nu + 1) / 2\nu] \quad , \qquad (11)$$

where $\widehat{\ln U}$ = predicted value;
$\hat{\sigma}^2$ = mean square residuals;
$\hat{\sigma}^2_{\widehat{\ln U}}$ = estimate of the variance of ln U;
ν = degrees of freedom for the residuals.

An estimator for the variance of ${}_mI_y$ is given by:

$$\begin{aligned} \mathrm{VAR}({}_mI_y) = {} & \exp(2\widehat{\ln U}) \{ (\exp[(\hat{\sigma}^2 - \hat{\sigma}^2_{\widehat{\ln U}})(\nu+1)/2\nu])^2 \\ & - \exp[(\hat{\sigma}^2 - 2\,\hat{\sigma}^2_{\widehat{\ln U}})(\nu+1)/\nu] \} \quad . \qquad (12) \end{aligned}$$

These estimators are based on computational formulas given by Ebbeler (1973) for the bias correction when using log transformations with linear models (Bradu and Mundlak 1970). They are based on asymptotic properties and provide a close approximation for large v, i.e. greater than 50.

A computationally more involved method of obtaining a catch rate index would be to use the estimated relative power for the categories in the confounding factors in order to standardize effort as was done in the standardized response method. Kimura (1981) found that differences were small when he compared this approach to the former.

Discussion

The selection of a method will often be restricted by data limitations or the specific situation. Comparison of the methods when more than one approach is applicable has not been reported but results are likely to depend on the problem being analyzed. It may be deemed appropriate to use an approach for one phase of the analysis and another approach subsequently. For example, the multiplicative model may be applied to gears within each region followed by the scaled response method to combine regions.

The issue of the most suitable level of data aggregation has received little attention, perhaps due to data limitations. Further work on this subject is warranted since erroneous statistical models may affect parameter estimates. For the standardized weighting and scaled response methods, the use of a log transformation could be explored.

References

Allen, R.G.D. 1975. Index numbers in theory and practice. Macmillan, London, 278 p.

Anderson, S., A. Auquier, W.W. Hanck, D. Oakes, W. Vandaele, and H.I. Weisberg. 1980. Statistical methods for comparative studies. John Wiley & Sons, New York, 289 p.

Anon. 1976. ACMRR Working Party on Fishing Effort and Monitoring of Fish Stock Abundance. FAO Fish. Tech. Pap. 155: 101 p.

Anon. 1981. Report of the Working Group on Methods of Fish Stock Assessments. ICES Doc. C.M. 1981/G:8.

Beverton, R.J.H., and S.J. Holt. 1957. On the dynamics of exploited fish populations. U.K. Min. Agric. Fish. Invest. (Ser. 2) 19: 533 p.

Bradu, D., and Y. Mundlak. 1970. Estimation in log-normal linear models. J. Amer. Statist. Assoc. 65: 198-211.

Chikuni, S. 1976. Problems in monitoring abundance in multi-species, multi-year groundfish fisheries in the Bering Sea, p. 23-36. *In* Monitoring of fish stock abundance: the use of catch and effort data. FAO Fish. Tech. Pap. 155.

Draper, N.R., and H. Smith. 1981. Applied Regression Analysis, 2nd Ed. John Wiley & Sons, New York, 709 p.

Ebbeler, D.H. 1973. A note on estimation in log-normal linear models. J. Statist. Comput. Simul. 2: 225-231.

Gavaris, S. 1980. Use of a multiplicative model to estimate catch rate and effort from commercial data. Can. J. Fish. Aquat. Sci. 37: 2272-2275.

Gulland, J.A. 1956. On the fishing effort in English demersal fisheries. U.K. Min. Agric. Fish., Fish. Invest. (Ser. 2), 20 : #5.

Gulland, J.A. 1964. The reliability of catch per unit effort as a measure of abundance in North Sea trawl fisheries. Rapp. P.-v. Réun. Cons. Perm. Int. Expl. Mer 155: 99-102.

Halliday, R.G., and W.G. Doubleday. 1976. Catch and effort trends for the finfish resources of the Scotian Shelf and estimates of the maximum sustainable yield of groundfish (except silver hake). ICNAF Selected Pap. 1:117-128.

Hunter, W.G., and W.F. Lamboy. 1981. A Bayesian analysis of the linear calibration problem. Technometrics 23: 323-328.

Kalton, G. 1968. Standardization: a technique to control for extraneous variables. Appl. Stat. 17: 118-136.

Ketchen, K.S. 1964. Measures of abundance from fisheries for more than one species. Rapp. P.-v. Réun. Cons. Perm. Int. Expl. Mer 155: 113-116.

Kimura, D.K. 1981. Standardized measures of relative abundance based on modelling log (c.p.u.e.) and their application to Pacific ocean perch (*Sebastes alutus*). J. Cons. Int. Expl. Mer 39: 211-218.

Paloheimo, J.E., and L.M. Dickie. 1964. Abundance and fishing success. Rapp. P.-v. Réun. Cons. Perm. Int. Expl. Mer 155: 152-163.

Parsons, L.S., A.T. Pinhorn, and D.G. Parsons. 1976. An evaluation of the northern Newfoundland-Labrador and Flemish Cap redfish fisheries. ICNAF Res. Bull. 12: 37-48.

Quinn, T.J., S.H. Hoag, and G.M. Southward. 1982. Comparison of two methods of combining catch-per-unit-effort data from geographic regions. Can. J. Fish. Aquat. Sci. 39: 837-846.

Ricker, W.E. 1940. Relation of catch per unit effort to abundance and rate of exploitation. J. Fish. Res. Board Can. 5: 43-70.

Rivard, D. 1988. Biological production and production models. *In* Rivard D. [Ed], Collected papers on assessment methods. CAFSAC Res. Doc. 88/61: 44-110.

Robson, D.S. 1966. Estimation of the relative fishing power of individual ships. ICNAF Res. Bull. 3: 5-14.

Rothschild, B.J. 1977. Fishing effort. *In* J.A. Gulland (ed.) Fish Population Dynamics. John Wiley & Sons, New York, 372 p.

Smith, S.J. 1988. Abundance indices from research survey data. *In* Rivard D. [Ed], Collected papers on assessment methods. CAFSAC Res. Doc. 88/61: 16-43.

Snedecor, G.W., and W.G. Cochran. 1967. Statistical Methods, 6th ed. Iowa State University Press, Ames, Iowa, 593 p.

Winters, G.H. 1988. The development and utility of sequential population analyses in contemporary assessment science. *In* Rivard D. [Ed.], Collected papers on assessment methods. CAFSAC Res. Doc. 88/61: 111-136.

Abundance Indices from Research Survey Data

by **S.J. Smith**

Marine Fish Division,
Biological Sciences Branch,
Department of Fisheries and Oceans,
Bedford Institute of Oceanography,
Dartmouth, Nova Scotia
B2Y 4A2

Introduction

The provision of scientific advice for the management of groundfish stocks in areas under CAFSAC jurisdiction is dependent upon two major sources of information: the commercial fishery and bottom trawl surveys. Bottom trawl surveys have been used in the Northwest Atlantic according to various designs and purposes for more than 30 years. A history of surveys in the Northwest Atlantic areas by ICNAF/NAFO member countries has been given by Doubleday (1981). Detailed discussions of use in specific areas have been given by Pitt et al. (1981) for the Newfoundland area, Halliday and Koeller (1981) for the Scotian Shelf, Bay of Fundy and southern Gulf of St. Lawrence and by Azarovitz (1981) for the Atlantic coastline of the United States.

Attention in this chapter is confined to discussing bottom trawl surveys for groundfish only. Other surveys such as ichthyoplankton and juvenile surveys as well as hydroacoustic surveys have yet to have as much of an impact on the assessment of groundfish stock status in CAFSAC.

A series of standardized bottom trawl surveys were initiated in 1970 on the Scotian Shelf and later expanded to all areas currently under CAFSAC jurisdiction. The choice of a stratified-random sampling design as the standard over all of the areas was based on the U.S. experience with this design for their surveys which had begun in 1963. At the moment, most surveys are conducted annually and usually at the same time every year. On the Scotian Shelf, there have been three surveys a year during the period 1978 to 1984. This has resulted in there being as much as 17 years of data from bottom trawl surveys in some areas. Within the CAFSAC forum, use of the information collected from these surveys can be categorized as follows:

1) Estimation of absolute or relative abundance;
2) Direct or relative index of population trends;
3) Pre-recruit indices of fish younger than those included in Sequential Population Analysis (SPA);
4) Estimation of mortality rates;
5) Derivation of partial recruitment factors for SPA;
6) Determination of final year fishing mortality for SPA.

This paper will be directed toward the first two assessment uses. The sections entitled Survey Design and Estimation contain a basic review of well established techniques for stratified random surveys. Modifications specifically made for groundfish surveys are also included. The next two sections deal with modeling catch per tow and trends in population abundance over time. These sections are confined to reviewing current practise in the CAFSAC forum and much of the discussion of these topics remains speculative in nature.

Survey Design

The emphasis on the stratified-random design in this paper has been chosen here for two reasons. In the first place, detailed comparisons of different sampling designs are given in most standard sampling texts and do not bear repeating here. In addition, a number of papers dealing with such comparisons within the context of trawl surveys are already available (e.g. Saville 1977; Doubleday 1981; Grosslein and Laurec 1982; Sissenwine et al. 1983; Fogarty 1985).

Our experience with respect to the collection and use of information from an ongoing and standardized sampling design has been limited to the stratified-random design. Therefore it makes sense to confine our attention to this design in order to document procedures and evaluate performance.

The bases for the definition of the strata used in the CAFSAC management area are given in Doubleday (1981). The major stratifying variable used was depth with strata classified according to specific depth intervals. Further delineations were based on fish distribution patterns and as much as possible on ICNAF (NAFO) stock boundaries.

Within each strata, the sampling unit or trawl set was defined to the area swept by a trawl with a specified width for a distance of 1.75 nautical miles. The width of the trawl is usually defined to be the wingspread and therefore a constant net opening over the duration of the tow is assumed. The actual number of these trawl units, N_h, is determined by dividing the total area of the strata by the area of the standard trawl unit. The total number of sampling units is simply

$$N = \sum_h N_h \quad , \qquad (1)$$

where $h = 1,...,L$. Total area has been calculated according to water area and does not take into account the actual surface of the bottom. The actual sampling frame for each strata consists of 5' latitude by 10' longitude rectangles each subdivided into a number of trawling locations. The choice of sample units is done by first randomly selecting a rectangle and then randomly selecting a location within. This two-stage selection of samples may not be the same as a truly random selection of sample units. Consider for example five rectangles each with ten trawling locations. The probability of choosing a trawling location either by ignoring the rectangles (1/50) or by using the two-stage selection (1/5 x 1/10) is equivalent. However, this will not be so if there are unequal numbers of trawling locations in one or more of the rectangles. The standard estimators for the means, totals and associated variances given for the stratified random design have been derived assuming equal probabilities within strata. Unequal probabilities will impact on the properties of these estimators (e.g. unbiasedness). Fortunately, the number of trawling locations per rectangle for CAFSAC management areas have been set to ten (Doubleday, 1981) and therefore the practise of two-stage sample unit selection should not be a problem here.

For trawl surveys, the main limit to the total number of sample units or sets ($n = \Sigma n_h$, where n_h are the number of sets in stratum h) is usually the time available for the cruise. The initial number of sets to be covered in a survey are determined according to past experience and available resources. Additional sets are usually chosen as alternates for those original locations which can not be successfully trawled due to bad bottom or other impediments. In practise, the allocation of sets to the strata is loosely defined. Basically, there is a minimum of two sets defined for all strata with additional sets allocated on a roughly proportional to strata area rule (i.e. larger strata - more sets). Adjustments from the proportional rule have been made to some strata according to fish distribution considerations.

The guidelines found in most sampling texts for allocating the number of sets to each stratum incorporate cost and variance criteria. The cost of a survey, as stated above, is usually fixed and defined according to time constraints. However, if cost is ignored, then the set to stratum allocation rule has implications for the level of precision of the estimates of mean catch per tow and total abundance. Judicious attention to strata construction and set allocation rules can result in a stratified-random design that will provide more precise estimates of the sample mean than a comparable simple random design. This aspect was investigated by Gavaris and Smith (1987) for trawl surveys of the Eastern Scotian Shelf (NAFO 4VsW). They reported that the present stratification-allocation scheme performed poorly relative to a simple random design with respect to the precision of the estimate of mean catch per tow for cod. The authors recommended that the number of strata be reduced so as to increase the flexibility for different allocation plans over the present plan. A strict proportional to area of stratum rule is preferred when the species being surveyed does not have a persistent and known spatial distribution or if many species with very different distributions are being surveyed. However, even slight deviations from this proportional rule can result in a loss of potential precision. Francis (1984) has suggested that the set allocation rule can be modified to be optimal with respect to precision while the survey is being done. This procedure requires that all strata are initially allocated two sets which are covered in the first phase of the survey. In the second phase, the remaining sets are allocated to the strata in proportion to the estimated variances measured from the first two sets. This method may be useful for single species surveys but may well be more expensive because the vessel will have to cover the survey area for each phase.

Estimation

The basic estimation formulae for means, totals and variances for stratified-random designs can be found in standard sampling texts. The notation given in Cochran (1977) will be adopted as the standard here. This includes the use of upper case letters to denote population values and lower case to denote sample estimates.

Numbers and Weight

Further to the notation given in the Survey Design Section, define y_{hi} to be the total number or total weight of a specific species caught in *i*th standard set ($i = 1, \ldots, n_h$) and *h*th stratum ($h = 1, \ldots, L$) . Recall that a standard set was defined as the area swept by a trawl when towed at a constant speed of 3.5 kt for 30 minutes. It is assumed that a distance of 1.75 nm over the bottom is covered in a 30 minute standard set for surveys in divisions 2J, 3KLNOP and 4RS. For sets in which the duration of the tow is less than 30 minutes, the observations are adjusted such that

$$y_{hi} = y'_{hi} (30 / t) \qquad , \qquad (2)$$

where y'_{hi} is the actual observation and t is the actual duration of the tow. This procedure makes the following assumptions:

1) the speed can be held constant at 3.5 kt for the duration of the tow;
2) progress through the water is unimpeded by tides, currents, etc.;
3) distance through the water is the same as distance over bottom; and
4) catch is linearly related to distance (and hence area) towed.

At present, the start and end positions of a tow are obtained from the ship's navigation instruments for surveys in the southern Gulf of St. Lawrence (4T) and Scotian Shelf (4VWX). Distance towed is calculated as the difference between these two points and catches are adjusted such that

$$y_{hi} = y'_{hi} (1.75 / d) \qquad , \qquad (3)$$

where again y'_{hi} is the actual observation (catch) and d is the actual distance towed. This procedure only requires the latter two assumptions above with the additional proviso that a straight line was travelled between the beginning and end points.

In the case of large catches where sub-sampling is carried out for numbers and sometimes for total weight as well, the observations used for the analysis are adjusted for the sampling fraction. That is, when the catch has been sub-sampled for numbers but the total weight (A) and sample weight (a) are known, the catch is adjusted as follows:

$$y_{hi} = y'_{hi} (A / a) \quad . \qquad (4)$$

In some cases, the total weight is not measured and instead is estimated from multiplying the total number of baskets of fish of the specific species caught by the average weight of a basket estimated from those baskets which were weighed. This estimated value for A is then substituted into the adjustment above.

Once all these adjustments have been made, the estimation of means and variances is straightforward. Within strata, the estimated mean and variance are given by

$$\bar{y}_h = \sum_{i=1}^{n_h} y_{hi} / n_h \quad ; \qquad (5)$$

$$s_h^2 = \sum_{i=1}^{n_h} \frac{(y_{hi} - \bar{y}_h)^2}{(n_h - 1)} \quad . \qquad (6)$$

The estimated mean over all strata is a weighted average of the strata means, i.e.,

$$\bar{y}_{st} = \sum_{h=1}^{L} \frac{N_h}{N} \bar{y}_h \quad , \qquad (7)$$

$$= \sum_{h=1}^{L} W_h \bar{y}_h \quad .$$

The variance of $\bar{y}_{st}$ is estimated by

$$\text{var}[\bar{y}_{st}] = \sum_{h=1}^{L} \frac{N_h (N_h - n_h)}{N^2} \frac{s_h^2}{n_h} \quad ;$$

$$= \sum_{h=1}^{L} W_h (W_h - \frac{n_h}{N}) \frac{s_h^2}{n_h} \quad . \qquad (8)$$

For trawl surveys, n_h/N is negligible and can be ignored, which reduces (8) to

$$\text{var}[\bar{y}_{st}] = \sum_{h=1}^{L} W_h^2 \frac{s_h^2}{n_h} \quad . \qquad (9)$$

The estimated total abundance with respect to numbers of fish or weight caught is simply estimated by

$$y_{st} = N \bar{y}_{st} \quad , \qquad (10)$$

with corresponding variance

$$\mathrm{var}[\ y_{st}\ ] = N^2\ \mathrm{var}[\ \bar{y}_{st}\ ] \quad . \qquad (11)$$

All of the above estimators are defined to be unbiased estimators of their respective finite population parameters. In sample survey methodology, it is assumed that we have N population units with each unit having associated with it a defined characteristic or vector of characteristics (e.g. numbers, weights, etc. for all the species being surveyed) which we want to measure. These characteristics are assumed to be fixed and it is the sampling plan which assigns a known probability of selection to the population units. Therefore, properties of estimators are a consequence of the sample design and not of any probability distribution assumptions for the measured characteristics. In fact, the measured characteristics are not assumed to follow any probability model. The estimation of confidence intervals for the mean or the total within strata relies on the distribution of $\bar{y}_h$ (or y_h) over all possible samples of size n_h in stratum h being well described by a probability distribution. Application of the central limit theorem to finite populations has been used to justify the choice of the Normal distribution for the mean (Little and Rubin, 1987). This does not require that the individual y_{hi} be normally distributed. The situation becomes more complicated when constructing confidence intervals for $\bar{y}_{st}$ or y_{st}. Typically, these intervals are expressed as,

$$\bar{y}_{st} \pm t\ \mathrm{se}[\ \bar{y}_{st}\ ] \quad ; \qquad (12)$$

$$y_{st} \pm t\ N\ \mathrm{se}[\ \bar{y}_{st}\ ] \quad , \qquad (13)$$

where

$$\mathrm{se}[\ \bar{y}_{st}\ ] = \sqrt{\mathrm{var}[\ \bar{y}_{st}\ ]} \quad .$$

The multiplier t is usually assumed to be the 1-(α/2) percentile of the Students-t distribution (Cochran, 1977) and therefore implies that the choice of the Normal distribution within strata can be extended over all strata. Because the variances within strata are assumed to be unequal, that is var[$\bar{y}_{st}$] is not a pooled estimate but a linear combination of potentially different variances, the degrees of freedom for the t variate is approximated by the following formula from Satterthwaite (1946).

$$\mathrm{EDF} = \frac{\left(\sum_{h=1}^{L} g_h\, s_h^2 \right)^2}{\sum_{h=1}^{L} g_h^2\, s_h^4 \,/\, (n_h - 1)} \quad , \qquad (14)$$

where,

$$g_h = N_h\, (\, N_h - n_h\,)\, /\, n_h \quad .$$

This approximation is referred to as the effective degrees of freedom (EDF). Since the above equation will usually give a real number as a solution, the answer should be truncated to an integer to assure that the confidence level is greater than or equal to the value that would have been obtained had the degrees of freedom been taken literally (R.F. Ling, Clemson University, pers. com.). Use of this approximation requires the stronger assumption that the y_{hi} are normally distributed. Discussion of this assumption is deferred to the section on Models for Catch per Tow.

Examples of these calculations are given in Tables 1 and 2 which are representative of printouts from the STRAP system (Smith and Somerton 1981). This system is currently in use in the Newfoundland and Quebec regions. A modified version has been implemented in the Scotia-Fundy region. The Gulf region currently uses RVAN, a system written for the microcomputer.

Numbers by Length Category

All commercial species are measured for length in order to provide length frequencies for species in the sets in which they were caught. Let y_{hij} be the number of fish of a specific species in the jth length category ($j = 1,..., J_i$, where J_i is the maximum length category observed for that species in that set) in the ith set and the hth stratum. Length categories are defined according to species, for example cod length categories are 3 cm groupings. A complete listing for species in the Scotia-Fundy region is given in Table 5 of Koeller (1981). Estimates of means, variances, etc. described in the subsection above are the same for the numbers in each of the length categories with y_{hij} substituted for y_{hi} in (5) and (6). In addition, the same adjustments for distance towed and sub-sampling are applied to the y_{hij} such that

$$y_{hi} = \sum_{j}^{J_i} y_{hij} \quad . \tag{15}$$

Note however that in general

$$s_h^2 \neq \sum_{j}^{J_i} s_{hj}^2 \quad . \tag{16}$$

An example of a STRAP printout for an analysis of numbers by length category is given in Tables 3 and 4.

Numbers at Age

Conversion of numbers of fish per length category to numbers at age requires information on the age-length relationship, more commonly referred to as the age-length key. During the survey, fish measured for length are sub-sampled for the removal of otoliths. These otoliths are assigned ages back in the respective laboratories and the derived ages along with the associated lengths are used to form the age-length key. Otolith sampling strategies vary between DFO regions, with the Scotia-Fundy and Gulf regions requiring a set maximum number of otoliths sub-sampled per length category per set. The Newfoundland region sets an overall quota for a cruise, area and species with sub-sampling to a set maximum per length category until the quota per length category is met. An example of an age-length key is given in Table 5.

Table 1. Numbers and weights listed for each set and adjusted by distance towed.

```
        POPULATION BIOMASS  4VWX - 5Y
                      NAFO: 4VN
                       WING: 41.0

                                 COD

           NUMBERS AND WEIGHTS PER STANDARD 30 MINUTE TOW

             STRATUM           SET             NUMBER           WT(KGS)
                40             58.                .00               .00
                40             59.                .00               .00
                40             67.               3.09              6.18
                40             68.                .00               .00
                41             60.              45.29            121.47
                41             61.              84.58            120.56
                41             62.              10.29             12.35
                41             65.              38.75             60.00
                41             66.             203.82            240.88
                42             63.              42.21             71.03
                42             64.            1460.37           1765.75
                42             70.               5.15              5.15
```

Table 2. Intermediate and final quantities used in calculating the stratified estimates for numbers caught.

```
        POPULATION BIOMASS  4VWX - 5Y
        ANALYSIS FOR CRUISE: N048,N049        JULY 1985
                       NAFO: 4VN
                       WING: 41.0

                                   COD

           NUMBERS

     STRATUM       NO. SETS      TOTAL       AV./SET      UNITS          TOTAL NO       VAR.
       40             4           3.09        0.77        78301.          60453.         2.38
       41             5         382.75       76.55        84741.        6486877.      5764.68
       42             3        1507.73      502.58       121773.       61200424.    688377.62

        LOWER CONFIDENCE LIMIT IS LESS THAN
         OR EQUAL TO ZERO
         ****-VARIANCE TOO LARGE FOR VALID
         CONFIDENCE INTERVAL AT THIS VALUE OF ALPHA-****

                         TOTAL                                         AVERAGE
           TOTAL         UPPER         LOWER          MEAN         UPPER          LOWER
         67747760      319050752    -183555264       237.86        1120.2        -644.47

        EFFECTIVE DEGREES OF FREEDOM=  2
        STUDENTS T-VALUE=  4.30 ALPHA= .05
```

Table 3. Set details for analysis of numbers caught by length group.

```
                                              LENGTH COMPOSITION-NUMBERS PER STANDARD TOW

         POPULATION ESTIMATES BY LENGTH JULY 1986 4VWX-5Y       FRI, AUG 29, 1986, 9:18 AM

         ANALYSIS FOR CRUISE: N065,N066         JULY 1986
                      STRATA: 440-442
                        WING:34.0

         SPECIES:     COD
         STRATUM:     442
         UNITS:     146844
         SEX:COMBINED

 LENGTH                                 SET DETAILS            SET STATISTICS      TOTAL
                                                                                 ABUNDANCE
 IN CMS.   58       59       61       62       63                 AVG.     VAR.   (1000'S)
  1.0     0.00     0.00     0.00     0.00     0.00                0.00     0.00     0.00
  4.0     0.00     0.00     0.00     0.00     0.00                0.00     0.00     0.00
  7.0     0.00     0.00     0.00     0.00     0.00                0.00     0.00     0.00
 10.0     0.00     0.00     0.00     0.00     0.00                0.00     0.00     0.00
 13.0     0.00     0.00     0.00     0.00     0.00                0.00     0.00     0.00
 16.0     0.00     0.00     0.00     0.00     0.00                0.00     0.00     0.00
 19.0     0.00     0.00     0.00     0.00     0.00                0.00     0.00     0.00
 22.0     0.00     0.00     0.00     0.00     0.00                0.00     0.00     0.00
 25.0     0.00     0.00     0.00     0.00     0.00                0.00     0.00     0.00
 28.0     0.00     0.00     0.00     0.00     0.00                0.00     0.00     0.00
 31.0     0.00     0.00     0.00     1.25     0.00                0.25     0.31    36.71
 34.0     0.00     1.03     0.00     0.00     0.00                0.21     0.21    30.23
 37.0     2.06     5.15     0.00     0.00     0.97                1.64     4.57   240.18
 40.0    10.29     4.12     2.69     3.75     1.94                4.56    11.02   669.57
 43.0     7.21     6.18     2.69     1.25     2.92                4.05     6.36   594.47
 46.0     4.12     6.18     1.35     2.50     4.86                3.80     3.65   558.05
 49.0     2.06     3.09     1.35     0.00     0.00                1.30     1.79   190.70
 52.0     5.15     3.09     1.35     0.00     3.89                2.69     4.17   395.61
 55.0     1.03     3.09     0.00     0.00     0.00                0.82     1.80   120.93
 58.0     0.00     4.12     1.35     0.00     2.92                1.68     3.31   246.13
 61.0     1.03     3.09     0.00     0.00     0.00                0.82     1.80   120.93
 64.0     3.09     2.06     0.00     0.00     0.00                1.03     2.12   151.16
 67.0     2.06     1.03     1.35     0.00     0.00                0.89     0.79   130.23
 70.0     2.06     1.03     0.00     1.25     0.00                0.87     0.77   127.41
 73.0     3.09     0.00     0.00     1.25     0.00                0.87     1.83   127.41
 76.0     4.12     1.03     0.00     0.00     0.00                1.03     3.18   151.16
 79.0    11.32     0.00     0.00     0.00     0.00                2.26    25.64   332.56
 82.0     5.15     0.00     0.00     0.00     0.00                1.03     5.30   151.16
 85.0     3.09     0.00     0.00     0.00     0.00                0.62     1.91    90.70
 88.0     2.06     0.00     0.00     0.00     0.00                0.41     0.85    60.47
 91.0     3.09     0.00     0.00     0.00     0.00                0.62     1.91    90.70
 94.0     0.00     0.00     0.00     0.00     0.00                0.00     0.00     0.00
 97.0     2.06     1.03     0.00     0.00     0.00                0.62     0.85    90.70
100.0     0.00     0.00     0.00     0.00     0.00                0.00     0.00     0.00
103.0     0.00     0.00     0.00     0.00     0.00                0.00     0.00     0.00
106.0     1.03     0.00     0.00     0.00     0.00                0.21     0.21    30.23
TOTAL    75.15    45.29    12.12    11.25    17.50               32.26   768.57  4737.40

         ESTIMATION TYPE:STANDARD    TRANSFORMATION TYPE:NONE
```

Table 4. Stratified estimates of numbers by length group in population. Note recent change in STRAP to giving standard errors instead of confidence intervals.

```
         POPULATION ESTIMATES BY LENGTH JULY 1986 4VWX-5Y
         ANALYSIS FOR CRUISE: N065,N066     JULY 1986
                      STRATA: 440-442
                        WING: 34.0

                                   LENGTH COMPOSITION-NUMBERS PER STANDARD TOW
      SUMMARY TABLE
      SPECIES:    COD
      SEX:COMBINED

       LENGTH          STRATIFIED        STANDARD            STRATIFIED         STANDARD
      IN CMS.          TOTAL NUMBERS     ERROR               MEAN               ERROR
           1.0                    0              0.00              0.00              0.00
           4.0                    0              0.00              0.00              0.00
           7.0                    0              0.00              0.00              0.00
          10.0                    0              0.00              0.00              0.00
          13.0                    0              0.00              0.00              0.00
          16.0                    0              0.00              0.00              0.00
          19.0                    0              0.00              0.00              0.00
          22.0                64431          26330.67              0.19              0.08
          25.0               232741         180139.94              0.68              0.52
          28.0               320841         200347.16              0.93              0.58
          31.0               740195         379373.75              2.16              1.10
          34.0               940159         652949.50              2.74              1.90
          37.0              1724729         702269.87              5.02              2.04
          40.0              2272458         599908.00              6.62              1.75
          43.0              2093477         493302.19              6.10              1.44
          46.0              1646805         272659.56              4.79              0.79
          49.0              1744936         334887.19              5.08              0.98
          52.0              1781537         288899.69              5.19              0.84
          55.0              1568660         355785.19              4.57              1.04
          58.0              1367754         443524.62              3.98              1.29
          61.0               901711         312028.19              2.63              0.91
          64.0               680793         140447.03              1.98              0.41
          67.0               424492          93728.55              1.24              0.27
          70.0               341458         134438.91              0.99              0.39
          73.0               295435         105535.09              0.86              0.31
          76.0               258703         135358.09              0.75              0.39
          79.0               416715         336523.06              1.21              0.98
          82.0               215595         153437.03              0.63              0.45
          85.0               173254         101853.70              0.50              0.30
          88.0               102543          65725.58              0.30              0.19
          91.0                90698          90696.55              0.26              0.26
          94.0                    0              0.00              0.00              0.00
          97.0               134091          66053.56              0.39              0.19
         100.0                19440          19439.29              0.06              0.06
         103.0                    0              0.00              0.00              0.00
         106.0                30233          30232.19              0.09              0.09
      TOTAL                20583884        2911599.00             59.93              8.48

      ESTIMATION TYPE:STANDARD    TRANSFORMATION TYPE:NONE
```

Table 5. Example of an age-length key.

```
ANALYSIS FOR CRUISE: N049                    JULY 1985
                NAFO: 4VN
                WING: 41.0

AGE/LENGTH KEY

SPECIES:    COD
SEX:COMBINED
                                 AGE IN YEARS
 LENGTH   0   1   2   3   4   5   6   7   8   9  10  11  12  13  14      SUM
  25.0 .......... 1 ...............................................        1
  28.0 .......... 1   6 ...........................................        7
  31.0 .............. 8   1 .......................................        9
  34.0 .............. 6   4   1 ...................................       11
  37.0 .............. 1   8 .......................................        9
  40.0 .................. 3   3 ...................................        6
  43.0 .................. 1   4   1 ...............................        6
  46.0 ...................... 7   1 ...............................        8
  49.0 ...................... 6   3   2 ...........................       11
  52.0 ...................... 6   4   1 ...........................       11
  55.0 ...................... 5   3 ...............................        8
  58.0 ...................... 1   4   3   1 .......................        9
  61.0 .......................... 2   4   2 ..... 1 ...............        9
  64.0 .......................... 3   2   4   3 ...................       12
  67.0 .............................. 4   2 .......................        6
  70.0 ............................................................        0
  73.0 ............................................................        0
  76.0 .................................. 2 ......... 1 ...........        3
  79.0 ............................................................        0
  82.0 .................................. 1 ................. 1 ...        2
  85.0 ............................................................        0
  88.0 ............................................................        0
  91.0 ............................................................        0
  94.0 ............................................................        0
  97.0 .............................................. 1 ...........        1
 100.0 .............................................. 1 ...........        1
 103.0 ............................................................        0
 106.0 .......................................................... 1        1
 109.0 ............................................................        0
 112.0 ...................................................... 1 ...        1
                                                                         132
          0   0   2  21  17  33  21  16  12   3   1   3   0   2   1
```

For the conversion of numbers per length category to numbers at age, the entries of the age-length key are expressed as proportions such that

P_{jm} = proportion of animals in the jth length category at age m ($m = 0, 1, \ldots, M$, where M is the maximum age observed).

Note that $0 \leq P_{jm} \leq 1$ and

$$\sum_{m=0}^{M} P_{jm} = 1 \quad .$$

The number of fish age m, y_{him}, in the ith set is estimated by

$$y_{him} = \sum_{j}^{J} P_{jm}\, y_{hij} \quad . \tag{17}$$

The same corrections for distance towed and sub-sampling, i.e. those given in the Numbers and Weight Subsection, are applied here such that

$$y_{hi} = \sum_{m}^{M} y_{him} \quad . \tag{18}$$

The same estimation procedures for means and variances, i.e. those given in the Numbers and Weight subsection, are applied to these data with y_{him} substituted for y_{hi}. Note that, as was the case for numbers at length,

$$s_h^2 \neq \sum_{m}^{M} s_{hm}^2 \quad . \tag{19}$$

Examples of analyses from the STRAP program for these kind of data are given in Tables 6 and 7. Note that for species where there are differences between sexes with respect to the age-length relationship, each sex is analyzed separately with its own age-length key.

Models for Catch per Tow

The finite population randomization basis for sample survey methods requires that the y_{hi} are fixed and the sample selection method be the random process (Cochran 1977; Little and Rubin 1987). No probability distributions need to be assumed for the y_{hi} to derive means and variances. However, the estimation of confidence intervals for $\bar{y}_{st}$ (and y_{st}) via the Normal assumption with the EDF approximation introduced earlier, does require that the number (or weight) caught per tow be normally distributed. The EDF approximation depends on the theory for the Normal distribution through the assumption that the variance of s_h^2 is given by $2\sigma_h^4\,(n_h - 1)$ (Cochran 1977, pg. 96).

Application of the t-distribution with the EDF approximation to mean catch per tow from survey data often gives unrealistic results. The analyses presented in Tables 2 and 7 resulted in negative lower bounds for mean catch per tow. The Normal distribution includes negative values in its range, however, they have no meaning for our purposes. For this reason, the most recent Scotia-Fundy version of STRAP provides standard errors and does not give confidence intervals as standard output (see Table 4).

Table 6. Set details for analysis of numbers by age group.

```
ANALYSIS FOR CRUISE: N049                  JULY 1985
                NAFO: 4VN
                WING: 41.0
SPECIES:    COD
STRATUM: 42
UNITS:    121773.
SEX:COMBINED

                                            SET DETAILS          SET STATISTICS        TOTAL
    AGE                                                                              ABUNDANCE
  IN YEARS      63        64        70                           AVG.        VAR.    (1000'S)
      0.0     0.00      0.00      0.00                           0.00        0.00        0.00
      1.0     0.00      0.00      0.00                           0.00        0.00        0.00
      2.0     0.29      0.00      0.00                           0.10        0.03       11.94
      3.0     4.36      5.46      1.48                           3.77        4.23      458.61
      4.0     2.61    116.57      1.00                          40.06     4390.71     4878.30
      5.0    15.30    807.44      1.81                         274.85   212784.44    33469.59
      6.0     7.55    326.19      0.76                         111.50    34581.01    13577.75
      7.0     5.26    134.19      0.09                          46.51     5771.99     5664.05
      8.0     3.83     38.05      0.00                          13.96      438.78     1699.82
      9.0     0.77      7.61      0.00                           2.79       17.54      340.16
     10.0     0.34      3.80      0.00                           1.38        4.43      168.35
     11.0     0.34      2.54      0.00                           0.96        1.90      116.93
     12.0     0.00      0.00      0.00                           0.00        0.00        0.00
     13.0     1.54      0.00      0.00                           0.51        0.79       62.68
UNKNOWN       0.00     19.03      0.00                           6.34      120.73      772.50
TOTAL        42.21   1460.97      5.15                         502.74   688855.75    61220.67

ESTIMATION TYPE:STANDARD   TRANSFORMATION TYPE:NONE
```

Table 7. Stratified estimates of numbers by age group in population.

```
      ANALYSIS FOR CRUISE: N049                    JULY 1985
                     NAFO: 4VN
                     WING: 41.0

                                                   AGE COMPOSITION-NUMBERS PER STANDARD TOW

      SUMMARY TABLE
      SPECIES:    COD
      SEX:COMBINED

   AGE
IN YEARS   TOTAL NUMBERS  UPPER LIMIT  LOWER LIMIT   MEAN PER TOW   UPPER LIMIT   LOWER LIMIT  D.F.
   0.0               0            0            0           0.00          0.00          0.00    0
   1.0               0            0            0           0.00          0.00          0.00    0
   2.0          136281       467087      -194524           0.48          1.64         -0.68    4
   3.0         1069278      2093054        45502           3.75          7.35          0.16    5
   4.0         5441037     25498668    -14616598          19.10         89.53        -51.32    2
   5.0        35871784    175506400   -103762832         125.95        616.21       -364.32    2
   6.0        14847152     71168496    -41474200          52.13        249.88       -145.62    2
   7.0         6373084     29402424    -16656258          22.38        103.23        -58.48    2
   8.0         2067293      8428020     -4293433           7.26         29.59        -15.07    2
   9.0          409655      1681437      -862127           1.44          5.90         -3.03    2
  10.0          220551       867945      -426842           0.77          3.05         -1.50    2
  11.0          192162       622784      -238460           0.67          2.19         -0.84    2
  12.0               0            0            0           0.00          0.00          0.00    0
  13.0           88848       368798      -191102           0.31          1.29         -0.67    2
  14.0           16478        62227       -29272           0.06          0.22         -0.10    4
UNKNOWN        1034414      4412302     -2343474           3.63         15.49         -8.23    2
TOTAL         67768016    319157952   -183621920         237.94       1120.58       -644.70    2

      ESTIMATION TYPE:STANDARD         TRANSFORMATION TYPE:NONE

      CONFIDENCE LEVEL:   95%

      ****-ONE OR MORE OF THE LOWER LIMITS IN THE ABOVE
      TABLE IS LESS THAN OR EQUAL TO ZERO. VARIANCE IS TOO LARGE FOR
      VALID CONFIDENCE LIMITS ****
```

A number of authors (e.g. Taylor 1953, Forest and Minet 1981, Koeller 1981) have presented empirical evidence for the unsuitability of the Normal assumption by showing that the mean and variance by stratum appear to be functionally related. The pattern in Figure 1 with standard deviation plotted against the mean is typical of their findings. The relationship appears to be proportional, which is often taken to imply that catch per tow can be modeled by one of the "contagious" probability distributions. The presence of the mean-standard deviation relationship precludes, of course, the y_{hi} being adequately described by a Normal distribution.

Attempts to model catch per tow by a statistical distribution have mainly concentrated on postulating a random distribution (i.e. Poisson) as the null hypothesis and testing this hypothesis via goodness-of-fit tests (e.g. Taylor 1953). Rejection of this null hypothesis by these tests and the consideration that fish tend to congregate or school for a number of possible reasons usually leads to the negative binomial or a related distribution being chosen to approximate the distribution of catch per tow (Taylor 1953, Houser and Dunn 1967 and Pennington and Grosslein 1978). Taylor (1953) proposed that if the number of a species caught in any tow behaved as a Poisson random variable but the mean varied from tow to tow as a Gamma variate due to aggregation, then the compound distribution would be negative binomial. Further, the observed proportional relationship between the mean and standard deviation implies that the variance is in excess of that expected for the Poisson model which has equal mean and variance. Whatever the underlying mechanism, the point made by Taylor (1953 pg. 57) that ..."high variance associated with the mean catch per tow is shown (i.e. by his analyses) to be a necessary condition of the basic heterogeneous distribution of the numbers of fish..." is well taken.

Within the CAFSAC forum, no use has been made of any distribution other than the Normal to construct confidence intervals. At present, any such intervals are not of direct use in stock assessments. However, the findings discussed above are often used to justify some action being taken with the data for even simple presentation of the mean catch per tow over time. The two main actions that are usually taken are either to apply a transformation to all of the data or to modify only a few of the observations.

Transformation of the Data

The presence of a mean to standard deviation relationship and/or an hypothesis of a negative binomial distribution for catch per tow has been used to justify a logarithmic transformation of the data (e.g. Pennington and Grosslein 1978). It is standard textbook policy (e.g. Snedecor and Cochran 1967, pg. 329) to use a logarithmic transformation to stabilize the variance or at least weaken the relationship between mean and variance. Within CAFSAC, the common practise is to substitute $x_{ih} = \log_e (y_{hi} + C)$ for y_{hi} in the formulae in the section on Estimation. The constant C is added to make the logarithm valid for sets where no fish of that particular species was caught. This constant is often set equal to one.

The log transformation or any other transformation for that matter had not been used to any great degree in CAFSAC assessment documents previous to 1985. This may have been due to software limitations and problems with data access and ease of data manipulation in the past. The STRAP systems and microcomputer-based software have removed such limitations since at least 1981. Halliday and Koeller (1981) reviewed the sue of trawl survey data in stock assessments for stocks in areas 4VWXT from 1970 to 1980 and reported that log transformations had been used for some herring and mackerel assessments only. Prior to 1985, only three documents (Sinclair and Maguire 1981; Zwanenburg 1983; Campana and Simon 1984) have reported using log transformations for groundfish stocks in these areas. Pitt *et al.* (1981) give examples of using log transformations for data from the redfish stock in the Gulf of St. Lawrence and the cod stock on Flemish Cap.

Two things are common to those analyses of groundfish stocks which have reported the use of the log transformation. First, the estimates are never left in the log scale. Instead, they are back-transformed to the arithmetic scale. Secondly, the back-transformation is confined to estimates of mean or total abundance. Estimates are back-transformed either at the stratum mean level (Zwanenburg 1983; Campana and Simon 1984) or at the stratified mean level (e.g. Sinclair and Maguire 1981).

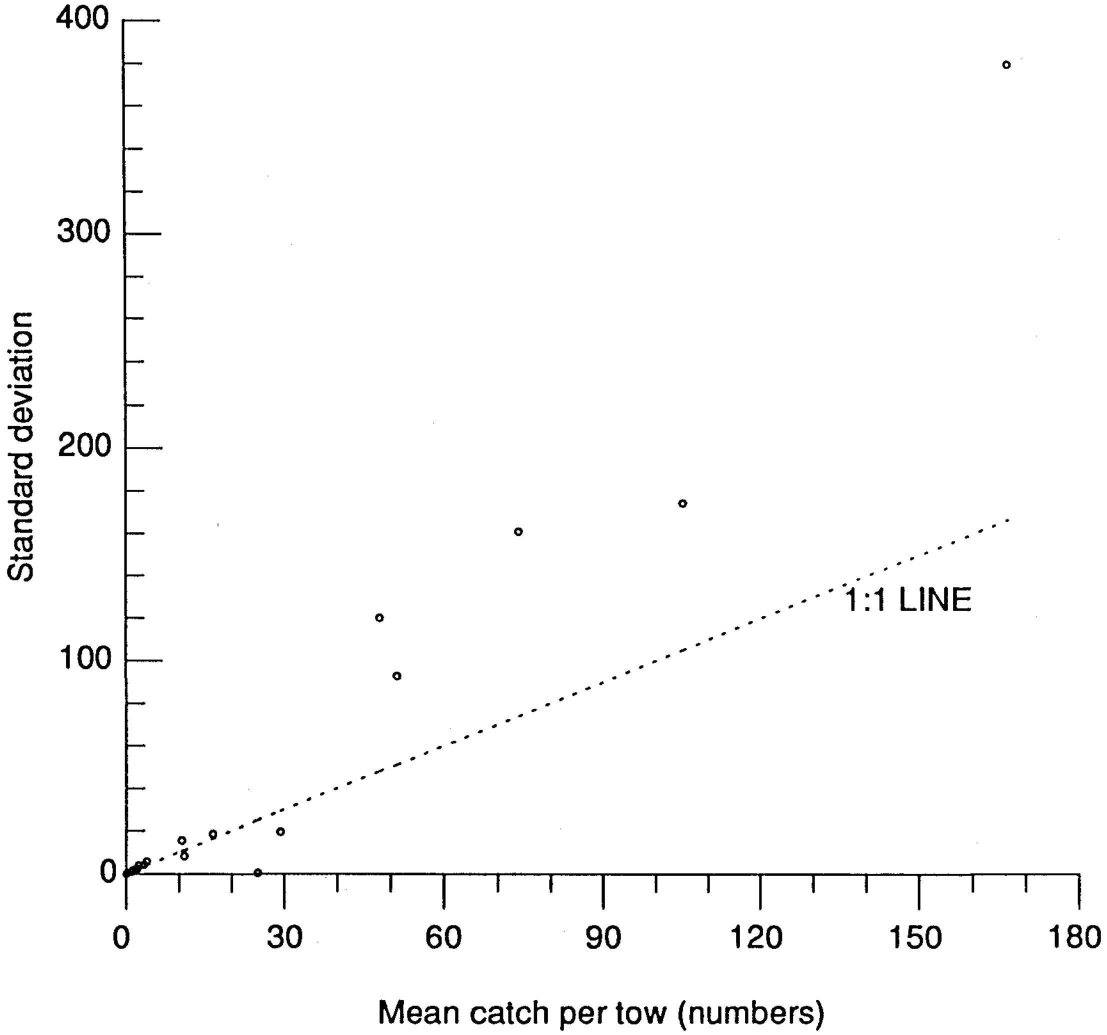


FIGURE 1. Comparison of mean catch per tow (numbers) of cod (4VsW) with standard deviation by strata. Research cruises N060, N078 (March 1986 and 1987).

The method of back-transformation also varies. Estimates made in the log scale may simply be exponentiated (i.e. $\bar{y}_h' = \exp(\bar{x}_h)$), or exponentiated and then have C subtracted to account for adding the constant in the first place. The sample form of the first moment of the Log-Normal distribution has also been used (Maguire 1979; Clarke 1981),

$$\bar{y}'' = \exp [\bar{x} + (\mathrm{var}[\bar{x}]/2)] - 1$$

where var($\bar{x}$) is the estimated variance of the mean in the log scale.

The Log transformation was proposed to make the mean independent of the variance, a property of symmetric distributions such as the Normal. By leaving the observations in the log scale, one could use the Normal-distribution theory to full advantage in estimating the necessary quantities. However, trawl survey data is often characterized by a large number of sets where no fish of a particular species were caught. Addition of 1 or any C to the zero observations and then applying a Log transformation only replaces a large number of zeroes in the arithmetic scale with a large number of values equal to log C on the log scale. This approach will not necessarily result in a symmetric distribution.

Pennington (1983) suggested that the data should be modeled as a two stage process with the first stage being a model of the event of fish of a certain species being absent or present in the set. The number of fish caught conditional on there being fish present in the set would be modeled by another probability distribution at the second stage. Pennington (1983) proposed the zero-augmented Log-Normal or the so-called Δ-distribution as a useful model for trawl catches. This distribution has three parameters, with δ representing the proportion of zeroes at the first stage. The number of fish caught is then assumed to follow a Log-Normal distribution with parameters μ and σ^2 representing the population mean and variance of the observations on the log scale. Estimators for the mean (k) of the Δ-distribution in the arithmetic scale and its variance are given in Pennington (1983) and in Smith (1988). If the Log-Normal model is appropriate for the distribution of catch per tow, then the estimate of the mean ($\hat{k}$) for the Δ-distribution is more precise than the arithmetic mean ($\bar{x}$). How much more precise depends on δ, n and σ^2 . The relative precision of the two estimators for various n, σ^2=0.4 and δ=0.1 is given in Figure 2. There, the relative precision is expressed as percent efficiency which is given by

$$\mathrm{eff}[\bar{x}] \;=\; (\mathrm{var}[\hat{k}] \;/\; \mathrm{var}[\bar{x}]) \times 100 \qquad .$$

The line labelled "Approximation" refers to the large sample approximation used in Pennington (1986), while the other lines are exact results for each value of n calculated by the formula given in Smith (1988). At n=2, neither estimator is to be preferred over the other, while large benefits using $\hat{k}$ require sample sizes within strata larger than usual for trawl surveys and large variances in the log scale.

A number of stock assessments used $\hat{k}$ and its variance in 1986 (e.g. Zwanenburg *et al.*, 1986; O'Boyle and Wallace, 1986) but the Statistics, Sampling and Surveys Subcommittee of CAFSAC recommended that this practise be discontinued. There have been questions raised about the robustness of $\hat{k}$ relative to the arithmetic mean when the Log-Normal is not appropriate. Sample sizes at present are too small to evaluate this distribution and there are no obvious biological bases for assuming a Log-Normal distribution for trawl catches. This question of robustness applies to using estimators specific to any distribution. This problem is avoided by finite population theory where no distribution is postulated for the quantities being measured (see Smith and Mohn, 1987, and references cited therein for more discussion on this topic).

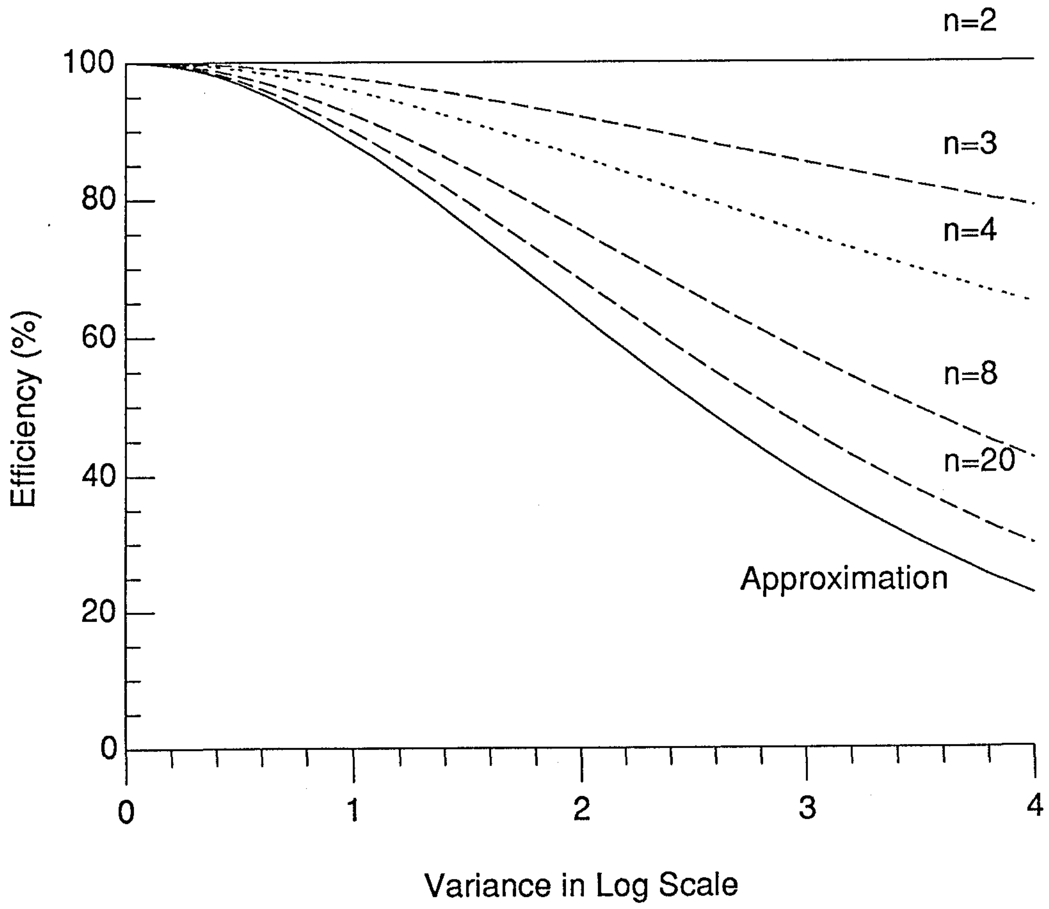


FIGURE 2. Comparison of the exact and approximate forms for the efficiency of the estimate of the mean of the Δ–distribution ($\hat{k}$) with respect to the arithmetic mean ($\bar{x}$). (δ=0.1)

Treatment of Large Observations

In contrast to the use of transformations which are applied to all of the observations, the methods presented in this section confine their attention to the large catches only. Given that between 2 to 8 sets are made in a stratum and that strata are quite variable in size, a few large sets in moderate to large strata can have a great deal of influence on the estimate of total abundance. Their effect can be quite marked, for example O'Boyle and White (1982) report that for haddock, five sets from two strata in the 1981 4X survey contributed 74.8% to the stratified mean and 98.0% to the variance. A more extreme case is from the 1973 4VSW survey where two sets from two strata, combined with their strata weights, made up 86.4% of stratified mean number of cod (Gray 1979).

The process of dealing with these sets has two steps; identification and treatment. The methods employed in these steps are basically *ad hoc* in nature and no methods have gained general acceptance as standard.

The identification phase begins with a comparison of the stratified mean or estimated total abundance from a particular year's survey with estimates for adjacent years. If the estimate for that year appears to be unusually large then a search is carried out through the catch per tow information for large sets. The estimate for the total abundance of cod in 1973 survey referred to above was found to be between 4 to 5 times that of adjacent years. The two largest sets mentioned above were then identified as the culprits (Gray 1979).

The effect that a large set can have on the final estimate is a function of the size of the strata it belonged to, since the final estimates are weighted means. In recognition of this fact, Maguire and Waiwood (1982) weighted each set by the stratum size before searching for possible "anomalous" sets. Comparisons were made by plotting the n_h weighted observations for the L strata over time. Suspect sets were then identified by comparing weighted catch per tows between and within years for each stratum. Clay (1980) also took stratum area into account in his method which identified anomalous sets as those which contributed 90% to the stratum mean and 20% to the stratified mean estimate.

These methods are obviously time consuming and extremely subjective. McKone et al. (1980) circumvented the time problem by simply identifying as anomalous the largest set from each survey in this series. However, how large is large in order to be identified as anomalous set(s) appears to be defined by each assessment scientist based on experience and the importance of this data to their stock assessment. Once these sets have been identified and determined to be anomalous, they are then treated in one of three ways.

The first method involves simply deleting these sets from the analysis (e.g. Gray (1979), Maguire and Waiwood (1982) and Atkinson (1983)). This approach assumes that the identified sets are not indicative of general abundance in their strata and that, instead, they reflect a very local event. There may also be other reasons for suspecting these sets. An error may have been made when the data was entered on the computer. Further, the set could have been of less than full duration and the assumption of linearity between amount caught and duration of tow (see subsection on Numbers and Weight) may be inappropriate. A minimum of 3 sets in the strata so modified is required for the strata to be included in the calculations.

The next two methods use some sort of replacement rule for the anomalous sets.

Clay (1980), in conjunction with his 90-20 rejection rule, replaced the identified sets with the mean of the observations from 3 or more adjacent years in that same stratum. The assumptions made in the deletion method hold here with the addition of stationarity of mean catch over time within the strata.

Recently, the method of "winsorizing" (see Smith 1981) was introduced as a substitution rule (Gagné et al. 1983; Maguire et al. 1983). Basically, the catches identified as anomalous are replaced by the next largest catches in their respective strata. The assumption made here is that these large catches may be indicative of large concentrations of fish in the area and the next largest catches are used to confirm or deny this possibility. For example, set 6 and set 7 with catches of 1347.5 and 8425.9 cod fish (numbers) were identified as anomalous in the 1973 4VSW survey. The decimal values are due to the correction for distance towed discussed in the Numbers and Weight subsection. Winsorizing involves replacing the catches for these two sets with the next largest in their respective strata, i.e. with 11.05 and 7.0 fish, respectively. The implication of this example is that the concentrations are local and not indicative of conditions in general in these strata.

All of the above methods implicitly assume that events are contained by the strata boundaries and cannot be transboundary. This assumption is the strongest for the winsorizing method since the next largest sets in the same strata are used to judge the validity of the sets identified as anomalous. Set 6 and 7 in the example above are in two adjacent strata. The sets were made consecutively and both had large numbers of juveniles in them. Therefore these sets may very well indicate a concentration which was both widespread and transboundary but the information contained in them is artificially reduced by confining attention to within strata comparisons.

A final important point must be made with respect to applying any of the methods presented in this section. The fact that an error may have been made in transcription or coding of the data for analysis was alluded to above. This possibility should be explored when trying to identify sets as anomalous by some criteria. This point has become extremely poignant lately with the recent discovery of an error (S. Gavaris, Marine Fish Division, pers. com.) in the 4VSW cod data we have been using to illustrate the above methods. It appears that the duration of the tow for set 6 was incorrectly coded such that the recorded catch of 8425.9 may be three times too large. This error is being rectified. In addition, the duration of the tow in set 7 was 20 minutes, which resulted in the actual catch being increased by 32% (speed of 4 knots). The adjusted catch then is dependent upon the linearity assumption between catch and tow duration.

Therefore, identification of sets as unusual should not automatically lead to implemention of deletion or substitution. The possibility of errors or other reasons must be exhausted first. Details of such investigations and any other aspects of these sets should be included in the final report of the analyses.

Estimates of Population Size and Trends

Once the estimates of total number or weight ($\hat{y}_{st}$) and stratified mean catch ($\bar{y}_{st}$) have been obtained from a particular trawl survey, the next step is to relate these estimates in some way to the actual population size. Assuming that a relationship does exist, the most general model that can be proposed is $E[\hat{y}_{st}] = f(P)$, where P is the form of population size being estimated, i.e. numerical abundance or biomass.

The simplest and probably the most unrealistic form for the right side of this model is $f(P) = P$. That is, the survey results are used to estimate absolute or actual abundance. This form is unrealistic because of the restrictive assumptions required for it to be applicable. However, Halliday and Koeller (1981) report this to be a very common usage of the survey estimates. The assumptions required for this form are as follows:

a) All size groups of the fish of a particular species are equally catchable. This implies that the size composition in the survey catch is an unbiased estimator of the size composition of the target population as a whole.
b) No escapement
    i) all fish in area swept are caught
    ii) no diel movement and/or vertical distribution above headline of net
    iii) negligible horizontal movement (e.g. migration) during survey period.
c) Area swept constant from tow to tow and/or linear relationship between distance towed and amount caught holds. This also implies that the net behaves the same under all conditions (bottom type, sea state, etc.).

All of the above factors are associated with the unbiasedness property of the model. The possible contagious distribution of the fish on the bottom would affect the precision of the estimates.

Echo soundings taken during research cruises have provided evidence that groundfish do distribute themselves in the water column above the headline of the net (Byrne et al. 1981 and Forest and Minet 1981). Therefore assumption (b.ii) could be modified to state that "there are no diel movements and that a vertical distribution exists but is static". The form of the model would then be changed to $f(P) = KP$ where $0<K<1$ with $(1-K)$ of the population uncatchable by virtue of their location in the water column. All of the other assumptions remain in force and the estimate $\hat{y}_{st}$ is now a measure of relative abundance or trawlable abundance with respect to location of the fish in the water column.

In situ observations of the behavior of fish when a trawl is moving over the bottom have shown that not all of the fish in the area swept are caught. Escapement occurs around the net and through the mesh. Assumption (b.i) needs to be modified to allow for this. Let $P' = KP$ represent that proportion of the population accessible to the trawl gear by virtue of location in the water column. Further, let

$$P' = \sum_{j}^{J} P_j' \quad ,$$

where J as before represents the maximum size (length) category in the accessible portion. Then we may express

$$f(P) = \sum_{j}^{J} S_j P_j' \quad ,$$

where $0 \leq S_j \leq 1$, to take into account the active escapement of the fish from the trawl in the area swept. The factors S_j are a form of selectivity coefficients for the species sought and gear used in the survey. Assumption "a" can also be weakened for this form of the model. However, it must be assumed that, for any species, the S_j's are constant from tow to tow which requires that assumption "c" must hold.

Age classes are used more often than size classes in the above model. Doubleday's (1981) "survivor method" is an example of this, which proposes that $E[\hat{R}_i] = K_{t-i} P_i$

where $\hat{R}_i$ is the survey estimate of abundance at year class i and t is the year. The index $(t-i)$ is equal to the age of the ith year class in year t .

A variant of the age based model is the practice of using only a range of ages in the estimate. Commonly referred to as *plus groups*, this method restricts the estimates to include fish age l and older (i.e. age l^+, hence 'plus groups'). The reasoning behind this method appears to be that the above assumptions are better met for those ages which are fully recruited to the research trawl.

It is more probable that selection to the gear is more related to the size of the fish than age. Therefore using age as a proxy for size requires the additional assumption that the age-length relationship be constant over time.

Violations of assumptions b.ii, b.iii and c have been discussed by a number of authors (e.g. Byrne *et al.* 1981, Forest and Minet 1981 and Pitt *et al.* 1981). In particular, Carrothers (1981) and Foster et al. (1981) discuss problems with using research trawls as quantitative tools.

Authors rarely state what assumptions are being made when abundance estimates from surveys are used in assessments. Basically, the estimates are called population estimates, relative estimates (or population indices) or estimates of trawlable biomass. Use of the term *minimum trawlable biomass* (e.g. Pitt *et al.* 1981) appears to refer to the third form of $f(P)$, with *minimum* indicating that the assumptions may be violated in some way.

The survey estimates of abundance are rarely considered by themselves for any one year but as a series of estimates from surveys carried out over a number of years. Trends over time in these estimates are expected to reflect trends in the population. The general form of the model linking estimates and population size becomes $E[\hat{y}_{st}(t)] = f(\ P(t)\ )$ where t indexes time in years. The addition of the temporal dimension to the general model makes it necessary to make additional assumptions. The use of the total estimate whether for absolute or relative abundance requires that all of the same strata are sampled each year. If there are missing strata then, if it can be assumed that the trend over time is the same in all strata, one would use the mean $\bar{y}_{st}(t)$ for relative indices. Otherwise a model to estimate the mean for the missing strata would need to be postulated in order to use the estimate of the total. The surveys should be carried out at the same time every year in order to minimize seasonal movements, etc. This of course assumes that conditions are the same, year to year, during that period.

For the form $K_tP(t)$, it must be assumed that $K_t = \hat{K}$, for any t in order for the estimate to be useful. This constancy over time must also be true for the size selection model, i.e. for $S_{jt}P_j(t)$, $S_{jt} = S_j$ for any t, and further, for age-based models, the age-length relationship must also be constant over time.

Obviously, the form used for f depends on knowing a great deal about the gear and the population being modeled. Models for relating the survey to predicted population sizes from sequential population analysis are beyond the scope of this chapter. The chapter on Sequential Population Analysis by G.H. Winters in this manual should be referred to for more information on these models.

The rest of this section will briefly review methods for presenting information on trends over time in the survey indices. Commonly, plots of mean catch per tow or total abundance versus years are characterized by large changes in abundance levels over short periods of time. The trend given by the solid line in Figure 3 is a good example of this. Large changes are often due to a few large catches in the survey and because of this, the estimate may not reflect true changes in abundance or at least overemphasize any real change (see section on Treatment of Large Values, this chapter).

Data from time trends can be expressed as having two components,

$$\text{DATA} = \text{SMOOTH} + \text{ROUGH} \qquad ,$$

where the SMOOTH is the signal we are interested in extracting and the ROUGH is the extraneous noise we wish to remove. To date, two methods have been used to filter out the SMOOTH component. The first method filtered or smoothed the data by using running means to extract the SMOOTH component. If we define our observations as y_t where t indexes time, then the running mean method replaces the y_t value with the average of y_{t-a}, ..., y_t, ..., y_{t+a}, where a is the number of years on either side of the point to be included in the average. The trend given by the running mean method with a=1 (referred to as a 3-year mean) is presented as the dotted line in Figure 3. Note that the large changes, albeit reduced somewhat in magnitude, are still evident. In addition, the effect of the large changes have spread out over adjacent years, making the changes appear to be smoother. Velleman (1980) has suggested that large peaks such as are present in the raw index in Figure 3 may be considered to be outliers if unsupported by information provided by the next point or points in the series. If we define these large peaks to be outliers, then the method of running means allows these points to contaminate the sequences they are included in.

Tukey (1977) and others have suggested that nonlinear smoothers based on medians would provide protection against these unsupported spikes in the data. The method of smoothing by medians proceeds by replacing y_t with the median of y_{t-a}, ..., y_t, ..., y_{t+a}, where a is defined as above. The process is repeated on the smoothed series until no change occurs between the last smoothed series and the current series. The dashed line gives the median smooth line for a=1 in Figure 3. In this case, large changes have been eliminated if unsupported by adjacent points. The methods differ therefore on how rapid changes in mean catch per tow are treated. The method of running means appears to require the assumption that changes in one year were actually occurring over a number of years. On the other hand, large spikes (negative or positive) are to be ignored for the method of medians if only observed in one year.

One series against which the trawl survey indices may be compared is the commercial catch rate series for the same stock. However, the differences in scale due to the use of mean catch per tow in kg from the survey and thousands of tonnes for the commercial catch as well as the presence of large spikes in the former series will complicate direct comparison. These differences may be accounted for by scaling and standardizing the two series. A robust method of doing this consists of centering each series by their respective medians. The resultant series are deviations from the median and the next step involves dividing each series by their respective median absolute deviation (MAD). This method is analogous to centering by the mean and scaling by the standard deviation. The standardized catch series for the 4Vn (May-December) cod stock is presented in Figure 4 along with the standardized indices from the July surveys for stock. Changes in each series can now be compared on the same basis as so many MAD's from zero. The smoothed (median, a=1) standardized trawl series is also presented. All series tend to follow the same pattern with the standardized survey appearing to exaggerate changes seen in the catch series. The smoothed series is remarkably similar to the catch series.

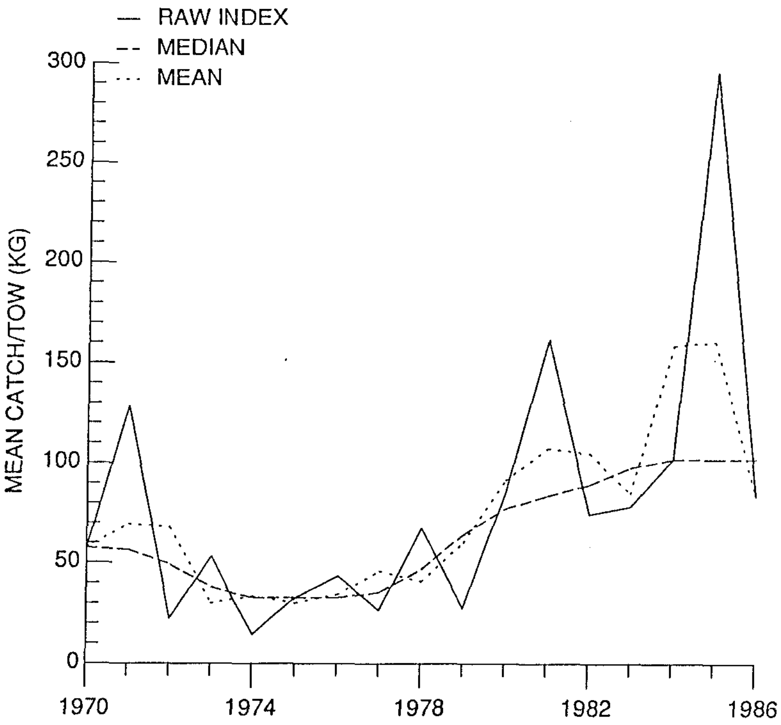


FIGURE 3. Effect of smoothing mean catch per tow (kg) using MEDIAN (3–year median smoothing) and MEAN (3–year running average).

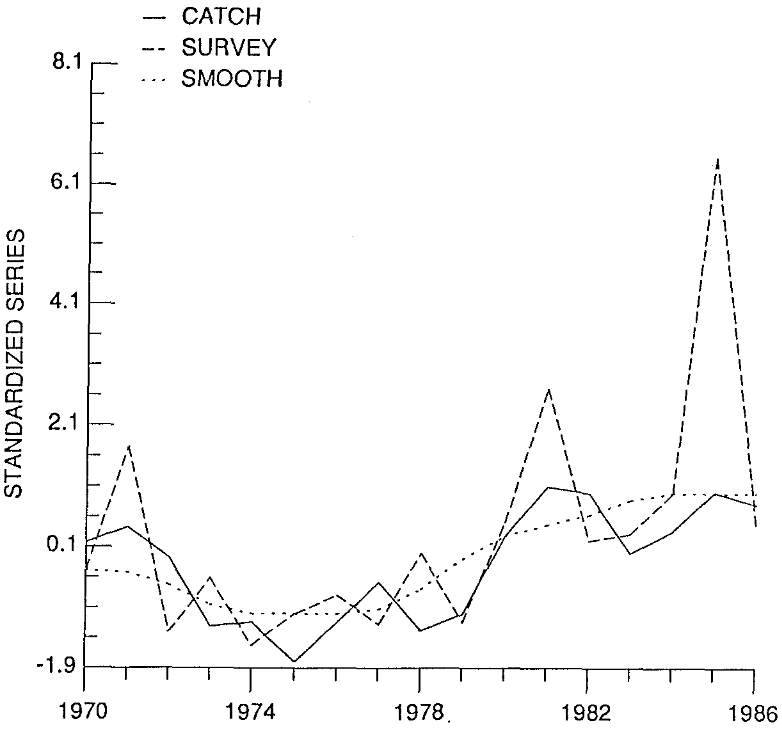


FIGURE 4. Robust standardization of catch ('000 t), survey (kg) and 3–year median smooth of the standardized survey index.

References

Atkinson, D.B. 1983. 2+3k Redfish. Can. Atl. Fish. Sci. Adv. Comm. Res. Doc. 83/23. 19 p.

Azarovitz, T.R. 1981. A brief historical review of the Woods Hole Laboratory trawl survey time series. *In* Doubleday, W.G. and D. Rivard (ed.), Bottom Trawl Surveys. Can. Spec. Publ. Fish. Aquat. Sci. 58:62-67.

Byrne, C.J., T.R. Azarovitz and M.P. Sissenwine. 1981. Factors affecting variability of research vessel trawl surveys. *In* Doubleday, W.G. and D. Rivard (ed.). Bottom Trawl Surveys. Can. Spec. Publ. Fish. Aquat. Sci. 58:258-273.

Campana, S. and J. Simon. 1984. The 4X cod fishery: a biological update. Can. Atl. Fish. Sci. Adv. Comm. Res. Doc. 84/43. 40 p.

Carrothers, P.J.G. 1981. Catch variability due to variations in groundfish otter trawl behaviour and possibilities to reduce it through instrumented fishing gear studies and improved fishing procedures. *In* Doubleday, W.G. and D. Rivard (ed.). Bottom Trawl Surveys. Can. Spec. Publ. Fish. Aquat. Sci. 58:247-257.

Clark, S.H. 1981. Use of trawl survey data in assessments. *In* Doubleday, W.G. and D. Rivard (ed.). Bottom Trawl Surveys. Can Spec. Publ. Fish. Aquat. Sci. 58:82-92.

Clay, D. 1980. Variability in abundance of Atlantic redfish derived from Canadian summer groundfish surveys on the Scotian Shelf (1970-1979). Can. Atl. Fish. Sci. Adv. Comm. Res. Doc. 80/31. 41 p.

Cochran, W.G. 1977. Sampling Techniques. Third Edition. John Wiley & Sons. New York. 428 p.

Doubleday, W.G. (ed.) 1981. Manual on Groundfish Surveys in the Northwest Atlantic. NAFO Scientific Council Studies No. 2. 55 p.

Doubleday, W.G. 1981. A method of estimating the abundance of survivors of an exploited fish population using commercial fishing catch at age and research vessel abundance indices. *In* Doubleday, W.G., and D. Rivard (ed.). Bottom Trawl Surveys. Can. Spec. Publ. Fish. Aquat. Sci. 58:164-178.

Fogarty, M.J. 1985. Statistical considerations in the design of trawl surveys. FAO Fish. Circ. (786):21p.

Forest, A. and J.P. Minet. 1981. Abundance estimates of the trawlable resources around the islands of St.Pierre and Miquelon (NAFO Subdiv. 3Ps): Methods used during the French research surveys and discussion of some results. *In* Doubleday, W.G. and D. Rivard (ed.). Bottom Trawl Surveys. Can. Spec. Publ. Fish. Aquat. Sci. 58:68-81.

Foster, J.J., C.M. Campbell and G.C.W. Sabin. 1981. The fish catching process relevant to trawls. *In* Doubleday, W.G. and D. Rivard (ed.). Bottom Trawl Surveys. Can. Spec. Publ. Fish. Aquat. Sci. 58:229-246.

Francis, R.I.I.C. 1984. An adaptive strategy for stratified random trawl surveys. New Zealand Journal of Marine and Freshwater Research. 18:59-71.

Gagné, J.A., A.F. Sinclair and L. Currie. 1983. Status of the 4VsW cod stock complex. Can. Atl. Fish. Sci. Adv. Comm.. 83/56. 56 p.

Gavaris, S. and S.J. Smith. 1987. Effect of allocation and stratification strategies on precision of survey abundance estimates for Atlantic cod (*Gadus morhua*) on the Eastern Scotian Shelf. J. Northw. Atl. Fish. Sci., Vol. 7: 137-144.

Gray, D.F., B.A. Fowler and O.V. Maessen. 1979. 1979 review of inshore cod stock in subdivision 4Vn (May-Dec.). Can. Atl. Fish. Sci. Adv. Comm. Res. Doc. 79/49. 19p.

Grosslein, M.D. and A. Laurec. 1982. Bottom trawl surveys design, operation and analysis. FAO, CECAF/ECAF Series 81/22 (En). 25p.

Halliday, R.G. and P.A. Koeller. 1981. A history of Canadian groundfish trawling surveys and data usage in ICNAF divisions 4TVWX. *In* Doubleday, W.G. and D. Rivard (ed.). Bottom Trawl Surveys. Can. Spec. Publ. Fish. Aquat. Sci. 58:27-41.

Houser, A. and J.E. Dunn. 1967. Estimating the size of the thread-fin shad population in Bull Shoal reservoir from midwater trawl catches. Trans. Am. Fish. Soc. 96:176-184.

Koeller, P.A. 1981. Manual for groundfish survey personnel - cruise preparation, conduct and standing orders. Marine Fish Division, Department of Fisheries and Oceans, Dartmouth, Nova Scotia, Canada. Internal Laboratory Reference Series No. 81/3.

Koeller, P.A. 1981. Distribution and sampling variability in the Southern Gulf of St. Lawrence groundfish surveys. *In* Doubleday, W.G. and D. Rivard (ed.). Bottom Trawl Surveys. Can. Spec. Publ. Fish. Aquat. Sci. 58:194-217.

Little, R.J.A. and D.B. Rubin. 1987. Statistical analysis with missing data. John Wiley & Sons, New York. 278 p.

Maguire, J.J. 1979. An analytical assessment of SA 3-6 mackerel with information from egg and larval survey. Can. Atl. Fish. Sci. Adv. Comm. Res. Doc. 79/46. 42 p.

Maguire, J.J. and K.G. Waiwood. 1982. The 1982 assessment of cod in NAFO Division 4T and Subdivision 4Vn (Jan-Apr.). Can. Atl. Fish. Sci. Adv. Comm. Res. Doc. 82/22. 29 p.

Maguire, J.J., J.-P. Lussiaa-Berdou and P. Rubec. 1983. The 1982 stock status and 1983-84 yield projections for 4RST redfish. Can. Atl. Fish. Sci. Adv. Comm. 83/50. 39 p.

McKone, W.D., D.B. Atkinson and W.E. Leggett. 1980. Gulf of St. Lawrence redfish assessment. Can. Atl. Fish. Sci. Adv. Comm. Res. Doc. 80/60. 43 p.

O'Boyle, R.N. and G.N. White III. 1982. An evaluation of the 4X haddock population characteristics during 1962-1981 with yield projected to 1983. Can. Atl. Fish. Sci. Adv. Comm. Res. Doc. 82/53. 51 p.

O'Boyle, R.N. and D. Wallace. 1986. An evaluation of the population dynamics of 4X haddock during 1962-1985 with yield projected to 1987. CAFSAC Res. Doc. 86/98. 71 p.

Pennington, M. 1983. Efficient estimators of abundance, for fish and plankton surveys. Biometrics 39:281-286.

Pennington, M. 1985. Estimating the relative abundance of fish from a series of trawl surveys. Biometrics 41:197-202.

Pennington, M. 1986. Some statistical techniques for estimating abundance indices from trawl surveys. Fish. Bull. 84:519-526.

Pennington, M.R. and M.D. Grosslein 1978. Accuracy of abundance indices based on stratified-random trawl surveys. ICNAF Res. Doc. 78/VI/77, Ser. No. 5264, 42 p.

Pitt, T.K., R. Wells and W.D. McKone. 1981. A critique of research vessel otter trawl surveys by the St. John's Research and Resource Services. *In* Doubleday, W.G. and D. Rivard (ed.). Bottom Trawl Surveys. Can. Spec. Publ. Fish. Aquat. Sci. 58:42-61.

Satterthwaite, F.E. 1946. An approximate distribution of estimates of variance components. Biometrics 2:110-114.

Saville, A. (ed.) 1977. Survey methods of appraising fishery resources. FAO Fish. Tech. Pap. (171). 76 p.

Sissenwine, M.P., T.R. Azarovitz and J.B. Suomala. 1983. Determining the abundance of fish. In A.G. Macdonald and L.G. Pringle (ed.). Experimental Biology at Sea. Academic Press Inc. (London). p. 51-101.

Snedecor, G.W. and W.G. Cochran. 1967. Statistical Methods. Sixth Edition. The Iowa State University Press. Ames, Iowa. 593 p.

Smith, S.J. 1981. A comparison of estimators of location for skewed populations - with application to groundfish trawl surveys. *In* Doubleday, W.G. and D. Rivard (ed.). Bottom Trawl Surveys. Can. Spec. Publ. Fish. Aquat. Sci. 58:154-163.

Smith, S.J. 1988. Evaluating the efficiency of the Δ-distribution mean estimator. Biometrics 44: 485-493.

Smith, S.J. and R.K. Mohn. 1987. Considerations on the representation and analysis of a spatially aggregated resource: Georges Bank scallops. ICES C.M. 1987/K:26. 19 p.

Smith, S.J. and G.D. Somerton. 1981. STRAP: A user-oriented computer analysis system for groundfish trawl survey data. Can. Tech. Rep. Fish. Aquat. Sci. 1030: iv+66 p.

Taylor, C.C. 1953. Nature of variability in trawl catches. Fish. Bull. 54:145-166.

Tukey, J.W. 1977. Exploratory Data Analysis. Addison-Wesley, Reading, Mass. 688 p.

Velleman, P.F. 1980. Definitions and comparison of robust nonlinear data smoothing algorithms. Journal of the American Statistical Association, 75: 609-615.

Zwanenburg, K.C.T. 1983. 1983 status report on the 4VWX redfish. Can. Atl. Fish. Sci. Adv. Comm. Res. Doc. 83/17. 25 p.

Zwanenburg, K.C.T., P. Fanning, R. Mahon, D. Waldron and P. Simpson. 1986. Haddock in management unit 4TVW: an assessment of present resource status - 1986. CAFSAC Res. Doc. 86/117. 51 p.

Biological Production
and
Surplus Production Models.

by **D. Rivard**

Fisheries Research Branch,
200 Kent, 12th floor,
Ottawa, Ontario, Canada.
KlA 0E6

Introduction

Certain population dynamics models depend on fairly detailed knowledge of growth, mortality rates, catch composition and recruitment. For many commercially exploited fish populations, such information is simply not available, or unobtainable. Often, historical records on catch and effort constitute the only information on a fishery. In this case, it is possible to relate yield to fishing effort in order to determine how, on the average, fishing effort affects yield. This approach gave rise to a family of models called surplus production models. In fisheries literature, surplus production models are also referred to as "stock production models", "surplus yield models" or simply as "production models".

It is particularly important to discuss at the beginning of this chapter the definition of production in the context of a fish stock. The terms "production" and "surplus production" have often been employed without a clear understanding of what production is in the context of the dynamics of a fish stock under exploitation. In many instances, it is not clear whether the term "production" refers to total production, surplus production or net production. What are the main components of production in an exploited fish population? Which age-groups are the main contributors to biological production? What are the basic mechanisms regulating recruitment and growth in a fish population? We will thus divert momentarily, in the next pages, from the main purpose of this chapter - i. e. the use of surplus production models for management - to gain insight into processes governing biological production for the fish stocks of the Northwest Atlantic.

Production from a single cohort

As in Ricker (1975), we will assume that the biomass of a cohort is governed through time by the relationship

$$B(t) = B_0 \, e^{(G-F-M)\,t} \quad , \qquad (1)$$

where B_0 is the initial biomass of the cohort, G is the mean instantaneous rate of growth, F is the instantaneous rate of fishing mortality and M is the instantaneous rate of natural mortality. The parameters G, F and M are assumed to be constant. The average biomass of the cohort over a unit time period, say one year, can be found from

$$\bar{B} = \int_0^1 B_0 \, e^{(G-F-M)\,t} \, dt \quad ,$$

$$= B_0 \, (e^{G-F-M} -1) \, / \, G-F-M \quad . \qquad (2)$$

The instantaneous net production at a given time t, $P^*(t)$, can be found from the instantaneous rate of change of population biomass, $B(t)$. That is

$$P^*(t) = d\,B(t) \,/\, dt = B_0 \, (G-F-M) \, e^{(G-F-M)\,t} \quad . \qquad (3)$$

Over a unit period, the net production (or overall change in biomass) is given by

$$P^* = \int_0^1 (d\,B(t) \,/\, dt) \, dt$$

$$= \int_0^1 B_0 \, (G-F-M) \, e^{(G-F-M)\, t} \, dt \quad ,$$

$$= (G-F-M) \, B_0 \, (e^{(G-F-M)} - 1) \, / \, G-F-M \quad . \qquad (4)$$

By combining equations (2) and (4), we can express the net production as

$$P^* = (G-F-M) \, \bar{B}$$

$$= G\bar{B} - F\bar{B} - M\bar{B} \quad . \qquad (5)$$

Equation (5) represents an important result since it serves to identify three components for net production: the increase of biomass through somatic growth, the loss of biomass through fishing and the loss of biomass through natural death. Ricker (1975) indicates that $F\bar{B}$ represents the weight of the catch (or yield); similarly, MB represents the weight of fish that die from natural causes. Total production, i.e. the total amount of tissue-weight gained by the cohort through growth over the year, is given by $G\bar{B}$: that is,

$$P = G\bar{B} \quad .$$

This definition can be shown to be equivalent to the definition employed by Eggers *et al.* (1978):

$$P = \int_0^1 N(t) \, (d \, W(t) \, / \, dt) \, dt \quad ,$$

where $W(t)$ is a time-dependent function describing somatic growth and $N(t)$ represents the number of fish in the population at time t. In our case, we can assume that

$$N(t) = N_0 \, \mathrm{e}^{-(F+M)\, t} \quad , \qquad (6)$$

and that

$$W(t) = W_0 \, \mathrm{e}^{Gt} \quad , \qquad (7)$$

and finally we find

$$P = N_0 \, W_0 \int_0^1 \mathrm{e}^{-(F+M)\, t} \, G \, \mathrm{e}^{Gt} \, dt \quad , \qquad (8)$$

$$= B_0 \, G \, (e^{G-F-M} - 1) \, / \, G-F-M \quad ,$$

$$= G \, \bar{B} \quad .$$

Finally, the excess of growth over natural mortality, say P^{**} can be expressed as

$$P^{**} = (G-M) \, \bar{B} \qquad (9)$$

The quantity P^{**} represents the weight of fish which has to be removed from the cohort in order to leave the stock at its current level (assuming constant recruitment). It thus represents the portion of total production which is available for exploitation if the number of recruits entering the fishery is constant.

The reader should note that the change in biomass over the year is equal to the net production, P^* . In fact,

$$\begin{aligned} \Delta B &= B_1 - B_0 , \quad (10)\\ &= B_0\, e^{(G-F-M)} - B_0 ,\\ &= B_0\, (G-F-M)\, (e^{(G-F-M)\,t} - 1) / G-F-M ,\\ &= (G-F-M)\, \bar{B} = P^* . \end{aligned}$$

Therefore, we can express total production over the year as (from equations 8 and 10)

$$P = \Delta B + F\bar{B} + M\bar{B} , \quad (11)$$

where ΔB is the biomass change over the year, $F\bar{B}$ is the catch expressed in weight and $M\bar{B}$ is the loss of biomass through natural death.

Production from a stock over a given interval

From a stock management viewpoint, production can be described as being composed of two components: an exogenous component, recruitment, and an endogenous component, total amount of tissue weight gained by all cohorts through somatic growth. In other words, total production from a given fish stock over a period t can be expressed as

$$P_{\bullet,t} = B_{b,t} + \sum_i P_{i,t} , \quad (12)$$

where $B_{b,t}$ is the biomass of the youngest age-group entering the exploited population at the beginning of the period t, and $P_{i,t}$ is the total production from age group i over t due to somatic growth. The quantity $B_{b,t}$ is thus a measure of the net contribution of prerecruits to production over t . The ratio $B_{b,t}/P_{\bullet,t}$ would also provide an indication of the relative contribution of recruitment to total production. The higher this ratio, the more dependent the fishery is upon new recruits.

For each age-group represented in the stock in a given year, the observed change in biomass over t, say $\Delta B_{i,t}$, can be defined as the difference between the growth component of production and the biomass loss due to fishing and natural mortalities. From equation (10), we write

$$\Delta B_{i,t} = P_{i,t} - D_{i,t} - Y_{i,t} , \quad (13)$$

where $Y_{i,t}$ refers to the loss of biomass for age-group i over t through fishing mortality, and where $D_{i,t}$ represents the loss of biomass for age-group i over t through natural mortality. The quantity $\Delta B_{i,t}$ can be either positive or negative, depending on the relative importance of the growth component with respect to the losses through fishing and natural deaths. If $\Delta B_{i,t}$, $D_{i,t}$ and $Y_{i,t}$ are known, the growth component of production over t can be calculated for age-group i by rearranging equation (13). That is,

$$P_{i,t} = \Delta B_{i,t} + D_{i,t} + Y_{i,t} , \quad (14)$$

Consequently, total production over *t* appears as

$$P_{\bullet,t} = B_{b,t} + \sum \Delta B_{i,t} + \sum D_{i,t} + \sum Y_{i,t} \quad , \tag{15}$$

where the summations are taken over all age-groups *i*. Annual net production can then be defined as the total production diminished by the losses of biomass due to fishing and natural death. That is,

$$\begin{aligned} P_{\bullet,t}^{*} &= P_{\bullet,t} - \sum_{i} D_{i,t} - \sum_{i} Y_{i,t} \\ &= B_{b,t} + \sum_{i} \Delta B_{i,t} \quad . \end{aligned} \tag{16}$$

From equation (16), annual net production appears as being composed of two components: an exogenous component, recruitment, and an endogenous component, the actual change in population biomass over *t*. The manager has direct control on an annual basis over the endogenous component of net production through the management of fishing effort. However, the exogenous component can only be weakly controlled through long-term management practices for "optimum" recruitment.

Finally, the "surplus production", i.e. the excess of recruitment and growth over natural deaths, is defined as

$$P_{\bullet,t}^{**} = P_{\bullet,t} - \sum_{i} D_{i,t} \quad . \tag{17}$$

The quantity $P_{\bullet,t}^{**}$ represents the portion of total production which is available for exploitation if we want to keep the stock at its current level. In fact, we can rewrite equation (16) as

$$P_{\bullet,t}^{*} = P_{\bullet,t}^{**} - \sum_{i} Y_{i,t} \quad ,$$

which indicates that the sign of net production (and consequently the sign of the overall change in population biomass over *t*) is governed by the quantity of fish removed.

The term "production" is not used here in the eco-physiological sense and does not refer to the total amount of energy needed for maintaining life, for somatic growth and for reproductive functions. Total production has been defined here as an *annual change of biomass resulting from somatic growth and recruitment in the exploitable portion of a fish population.* Total production, in an eco-physiological context, would also consider food ingested, metabolic functions and excretion. For studying the dynamics of a stock, it will be sufficient to define total production as the net change in biomass resulting from food ingestion, metabolic functions (e.g. respiration, reproduction, growth, etc.) and excretion for a stock as a whole.

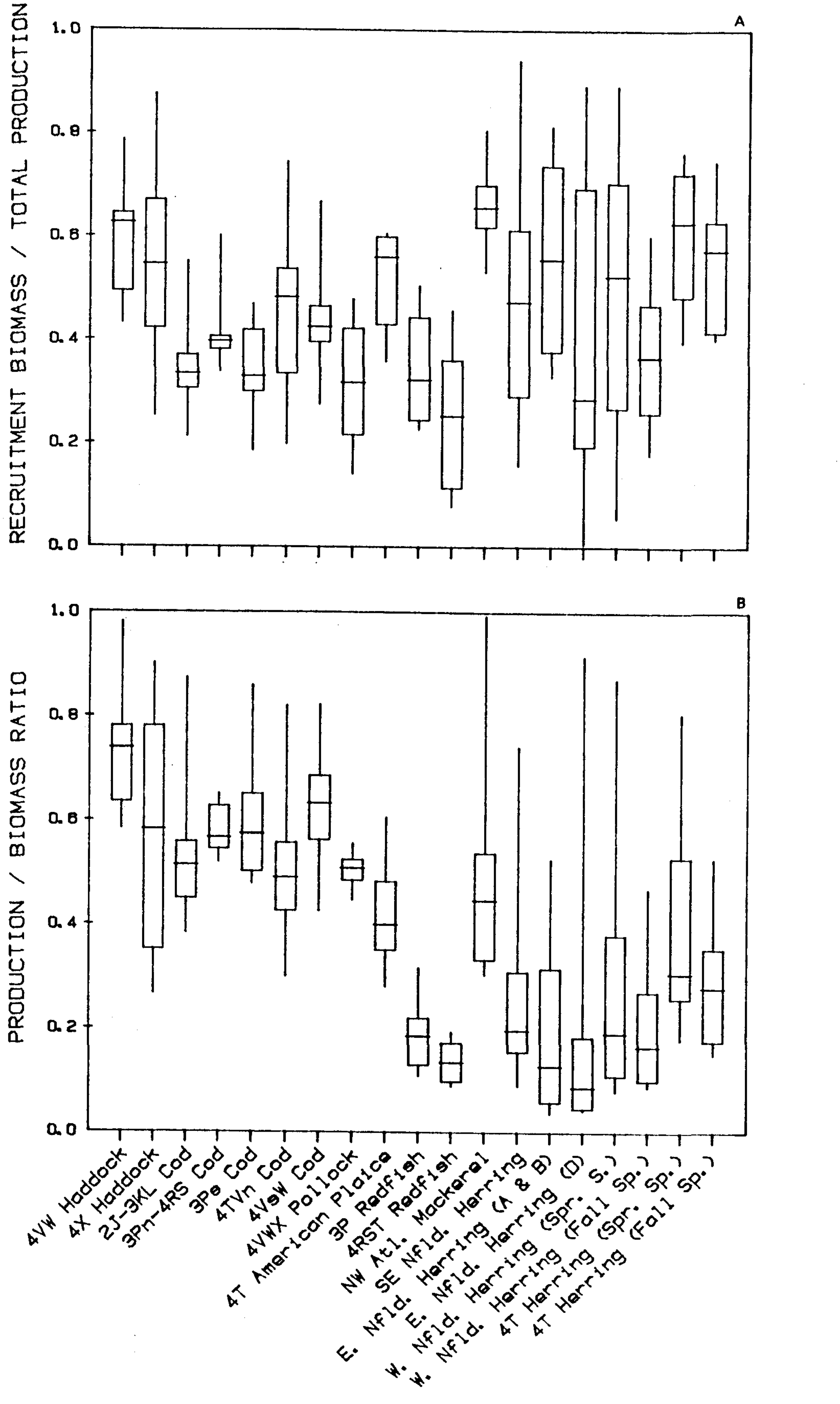
Examples from the Northwest Atlantic.

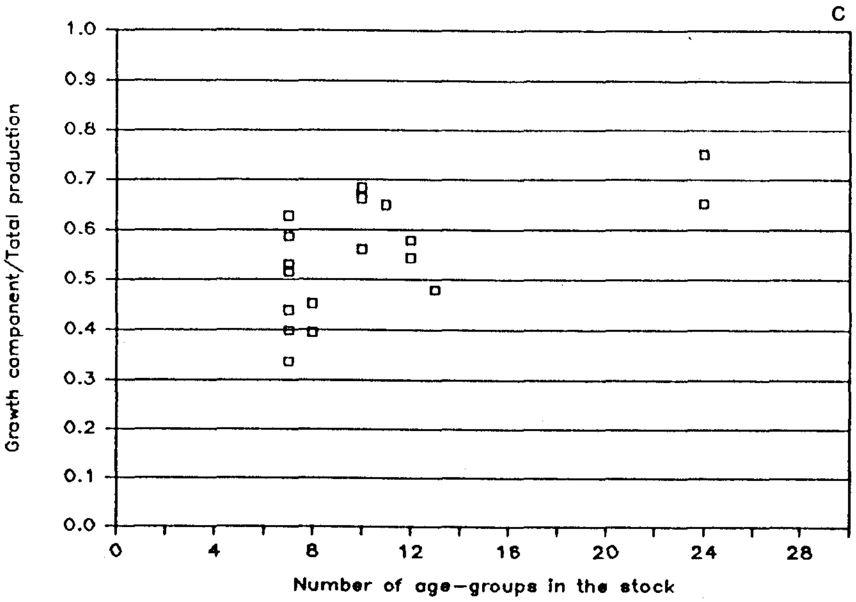
The equations derived above can be used to back-calculate production from the results of cohort analysis (see Rivard, 1980). In particular, Figure 1 summarizes the results of such calculations for a number of stocks of the Northwest Atlantic for haddock, cod, plaice, redfish, mackerel and herring. For this analysis, each stock is defined as all age-groups having a partial recruitment coefficient higher than 0.5 . For each stock, the relative contribution of recruitment and growth to total production is evaluated, as well as the *production/biomass* ratio.

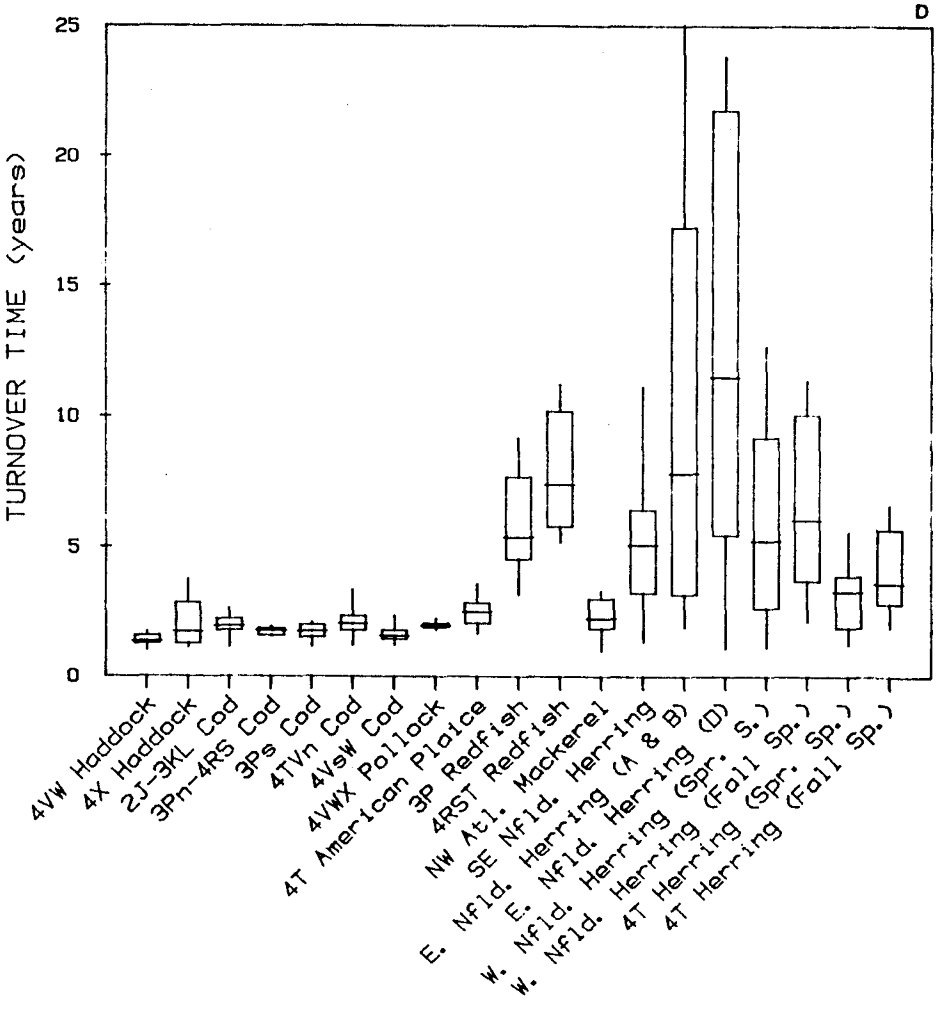
FIGURE 1.

A. Relative contribution of recruitment biomass to total production;
B. Production/biomass ratio;
C. Dependence of total production upon growth in recruited age groups;
D. Turnover times;
E. Relative contribution of each age group to growth; the age groups represented in each histogram correspond to the age span of the stock considered.

Box and whiskers plots identify, for each variable, the median (center bar), hinges (upper and lower edges of each box), as well as the minimum and maximum observed values (the end points of whiskers).


A
RECRUITMENT BIOMASS / TOTAL PRODUCTION
1.0
0.8
0.6
0.4
0.2
0.0
B
PRODUCTION / BIOMASS RATIO
1.0
0.8
0.6
0.4
0.2
0.0
4VW Haddock
4X Haddock
2J-3KL Cod
3Pn-4RS Cod
3Ps Cod
4TVn Cod
4VsW Cod
4VWX Pollock
4T American Plaice
3P Redfish
4RST Redfish
NW Atl. Mackerel
SE Nfld. Herring
E. Nfld. Herring (A & B)
E. Nfld. Herring (D)
W. Nfld. Herring (Spr. S.)
W. Nfld. Herring (Fall Sp.)
4T Herring (Spr. Sp.)
4T Herring (Fall Sp.)


C
Growth component/Total production
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
4
8
12
16
20
24
28
Number of age-groups in the stock


D
TURNOVER TIME (years)
25
20
15
10
5
0
4VW Haddock
4X Haddock
2J-3KL Cod
3Pn-4RS Cod
3Ps Cod
4TVn Cod
4VsW Cod
4VWX Pollock
4T American Plaice
3P Redfish
4RST Redfish
NW Atl. Mackerel
SE Nfld. Herring
E. Nfld. Herring (A & B)
E. Nfld. Herring (D)
W. Nfld. Herring (Spr. S.)
W. Nfld. Herring (Fall Sp.)
4T Herring (Spr. Sp.)
4T Herring (Fall Sp.)

E

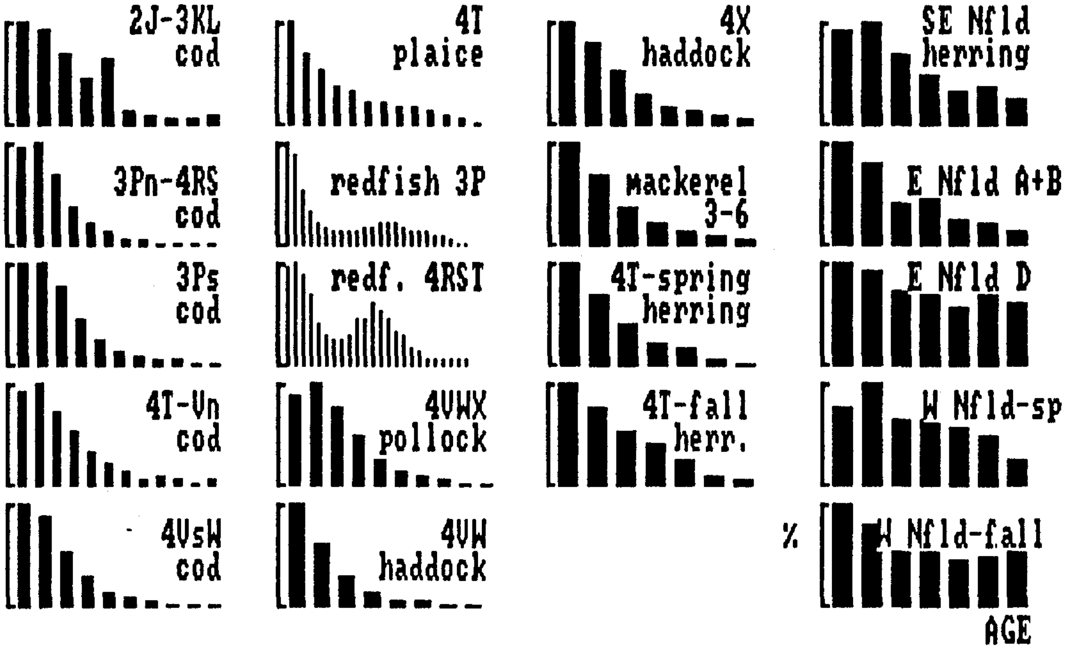

2J-3KL cod
4T plaice
4X haddock
SE Nfld herring
3Pn-4RS cod
redfish 3P
mackerel 3-6
E Nfld A+B
3Ps cod
redf. 4RST
4T-spring herring
E Nfld D
4T-Vn cod
4VWX pollock
4T-fall herr.
W Nfld-sp
4VsW cod
4VW haddock
%
W Nfld-fall
AGE

Figure 1A suggests that for haddock, plaice and mackerel in the Northwest Atlantic, total production is heavily dependent upon recruitment. For cod, pollock and redfish, total production is more dependent upon growth of the recruited age-groups. The relative contribution of growth and recruitment to total production is highly variable for most herring stocks of the Northwest Atlantic, as illustrated in Figure 1A. The Newfoundland stocks are dominated, over the time period considered, by one or only a few year classes. As a result of this, total production is mainly due to the growth of fish of the dominant year-class(es) in years of low recruitment. For the year(s) in which the dominant year class(es) is (are) recruited to the stock, total production is almost entirely the result of recruitment. For herring in Division 4T, recruitment is less variable and the dependence of total production upon growth from the recruited age-groups is reduced. In general, the growth component of production increases as the number of age-groups increases, as illustrated in Figure 1C. The dependence of total production upon recruitment and upon growth is also affected by the status of the stock, as illustrated by the herring stocks in Newfoundland: the dependence of total production upon growth in certain years for these stocks is the result of low recruitment levels in these years.

The ratio of production to mean "standing stock", ($P/\bar{B}$), often referred to as the turnover ratio (Waters, 1969; Bagenal, 1978), is summarized in Figure 1B. These results are consistent with the results of Mathew (1970) suggesting that a species with 7 or more age-groups commonly have $P/\bar{B}$ smaller than unity. The inverse of the turnover ratio provides an indication of the time it takes for a stock to replace its biomass: the turnover times are presented in Figure 1D. On the average, haddock stocks have very short turnover times - less than 2 years - while the two redfish stocks studied have rather long turnover times - 5 to 7 years. In years of poor recruitment, herring stocks show a low production and, consequently, long turnover times but in years of high recruitment, these stocks become more dynamic and are able of replacing their biomass in less than two years. The extreme variability demonstrated by the Newfoundland herring stocks is, here again, a result of the dependence of these stocks on only a few year classes.

The distribution of the growth component of production over ages is summarized in Figure 1E. For most groundfish stocks studied, as well as for mackerel in subarea 5-6 and for herring in 4T, the first few age-groups which are recruited to the fishery are the main contributors to production. As these age-groups are not yet fully recruited to the fishery for most species and stocks, production is mainly due to recruitment and to growth from partially recruited age-groups. Herring in Newfoundland again deviates from the general picture because of its reliance on a limited number of year classes. Consequently, the contribution of growth to production for herring in Newfoundland is, on the average, distributed almost uniformly over ages; this pattern is the result of growth of the dominant year class as it passes through the stock (particularly evident for herring at West Newfoundland). For redfish in 4RST, growth distribution over ages reveals a secondary mode at age 17, a result of the passage of the strong year classes of 1956 and 1958 through the stock (Gavaris and Atkinson, 1982). The dependence of production upon growth for this redfish stock is thus accentuated when a strong year class becomes fully recruited to the fishery. Similarly, a secondary mode appears at age 8 for cod in 2J-3KL. Following removals averaging about 650 000 mt in the period 1965-1969 for this stock, the catch rates declined drastically in the late sixties and early seventies, implying a decrease in the biomass of the stock. This, coupled with below average recruitment levels to the stock during that period, led to a higher dependence of production upon growth of fully recruited ages during the early- to mid-seventies.

Response of an age—structured fish population to fishing

As a means of illustrating the response of a fish population to fishing, an age-structured population having the characteristics of a cod stock was simulated and fished at various levels of fishing mortality for a period of 35 years. The projection algorithm described in Rivard (1982) was used for the simulation.

Equilibrium state: a simple definition.

If recruitment and fishing mortality (or effective fishing effort) are constant, then the response of the stock to fishing is described by Figure 2. When fishing is introduced, the stock decreases gradually from its initial level and "stabilizes" at a level which has often been referred to as an equilibrium level (Fig. 2B). The stock gives, at equilibrium, a constant yield: the equilibrium yield. For each effort level, a different equilibrium yield is achieved, as illustrated in Figure 3A. The curve thereby obtained is called the equilibrium curve. If recruitment is constant, the shape of the equilibrium curve is similar to the shape of the yield per recruit model. Similarly, the stock achieves a different equilibrium level for each effort level (Fig. 3B).

The transient state and transient paths: examples.

It is particularly interesting to analyse the response of the simulated stock when fishing effort changes. Figure 4, for example, shows the response of yield, catch rates and the *production/biomass* ratio to two sudden increases of effort. Note that the stock in year 1 was at equilibrium. After each change in effort level, yield rises sharply in the first year but decreases gradually in subsequent years to achieve a new equilibrium level. The various yield levels realized during the transient period - i.e. the time it takes for yield to move from one equilibrium level to a new one - are called "transient levels" and a series of transient levels define a "transient path". The transient yield levels are plotted against effort in Figure 7. It is important to understand the notions of "equilibrium levels", "transient levels" and "transient paths" early in this chapter, as these will play a major role in our analysis of production models.

The introduction of a higher effort level leads to declining catch rates during the transition period (Fig. 4C); in contrast, the *production/biomass* ratio increases rapidly in response to the introduction of higher effort levels (Fig. 4D). A higher fishing pressure removes older fish from the stock and reduces the growth component of production, as illustrated in Figure 6A. Fishing then becomes more dependent upon growth and upon recruitment from younger age-groups. For comparison, the response of yield, catch rates and the *production/biomass* ratio to a sudden drop in effort is illustrated in Figure 5. A lower effort leads to an immediate decline in yield but yield levels increase rapidly during the transition period to achieve a new equilibrium. Lowering effort also means an immediate pay off in terms of better catch rates (Fig. 5C). Finally, lower effort means that more age-groups are present in the stock: as a result of this, production is more dependent upon growth from these age-groups. As production is less dependent upon recruitment, a lower effort also leads to more stable stock levels.

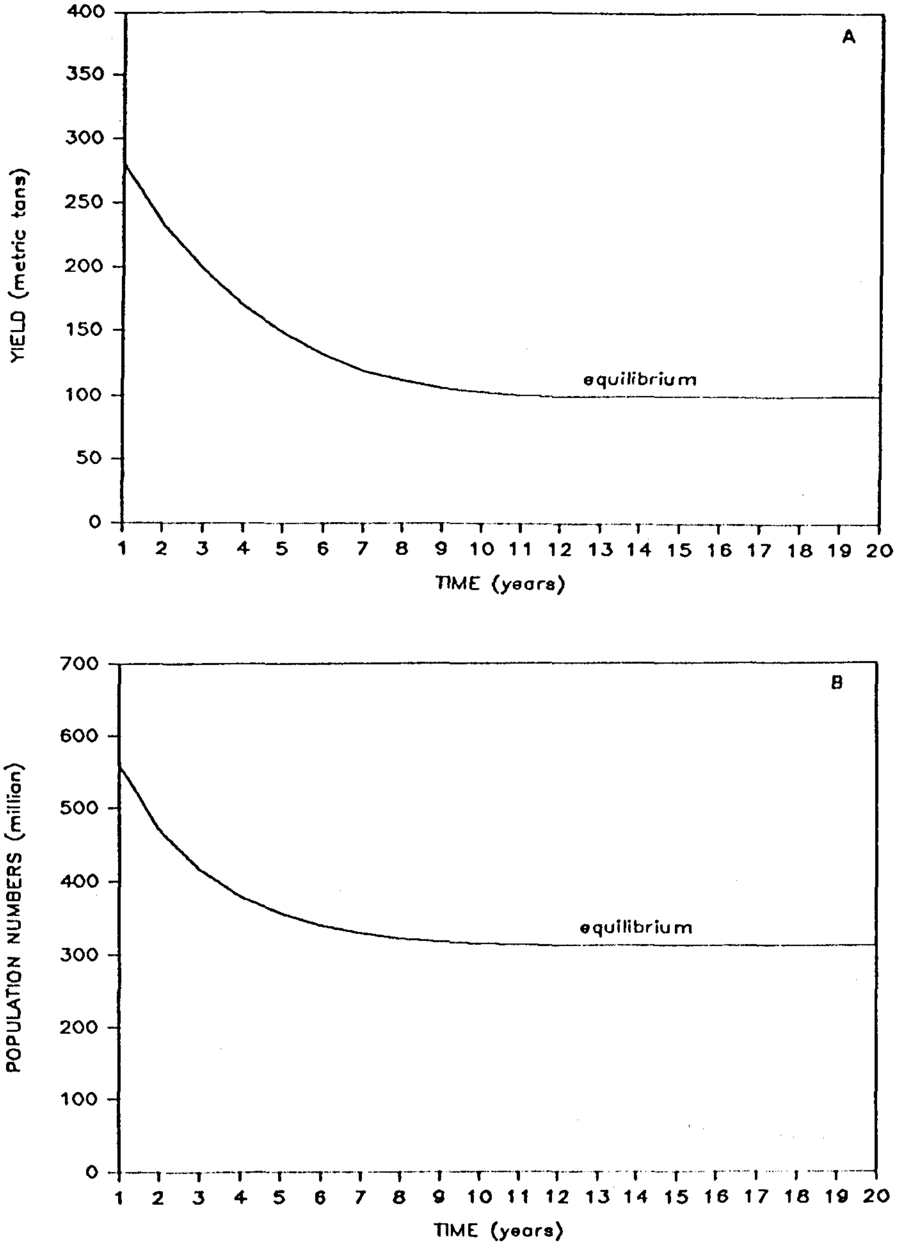


FIGURE 2. Response of a stock to fishing when recruitment and fishing mortality are constant: A) yield; B) stock size in numbers. When fishing is introduced, the stock, and subsequently the yield, decrease gradually from their initial level and stabilize at a level which has often been referred to as an *equilibrium level*.

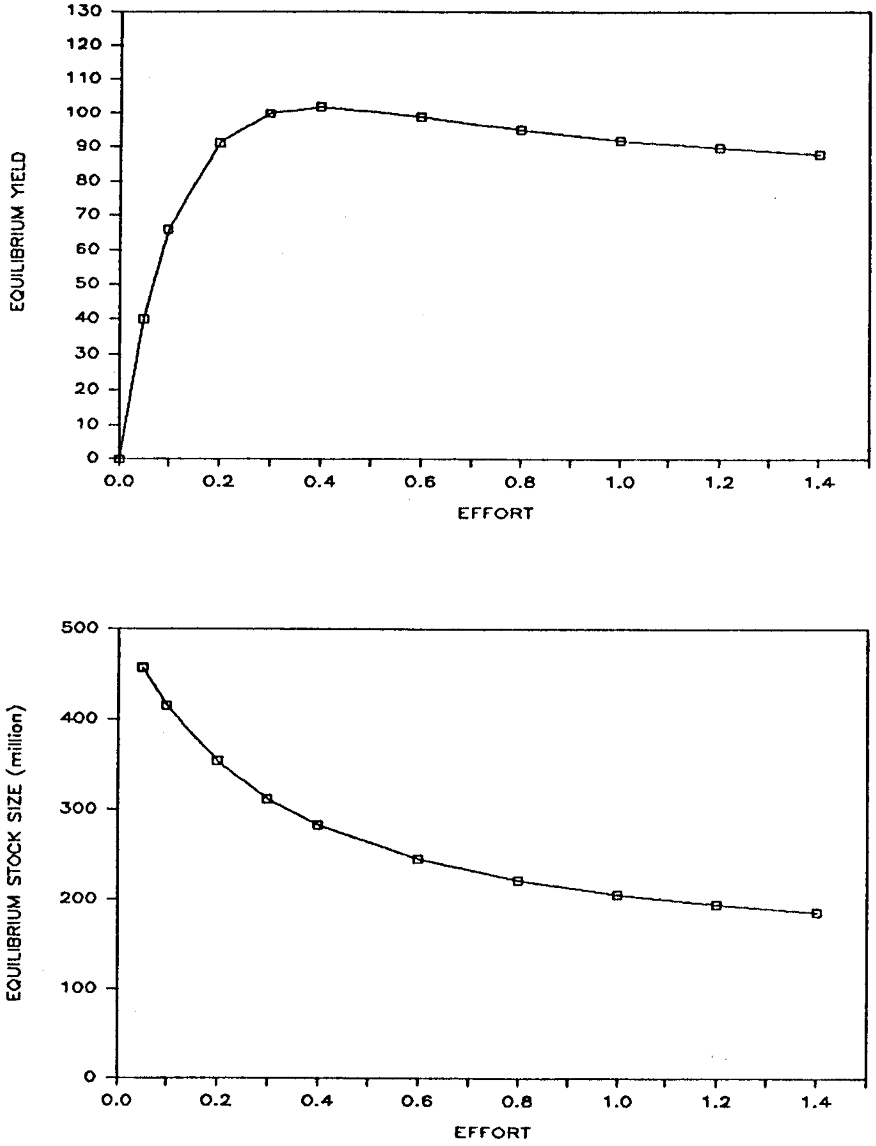


FIGURE 3. A) Equilibrium yield against effort. B) Stock size at equilibrium against effort (derived from a simulated stock having constant recruitment). Note that when recruitment is constant, the shape of the equilibrium curve is similar to the shape of the yield per recruit model.

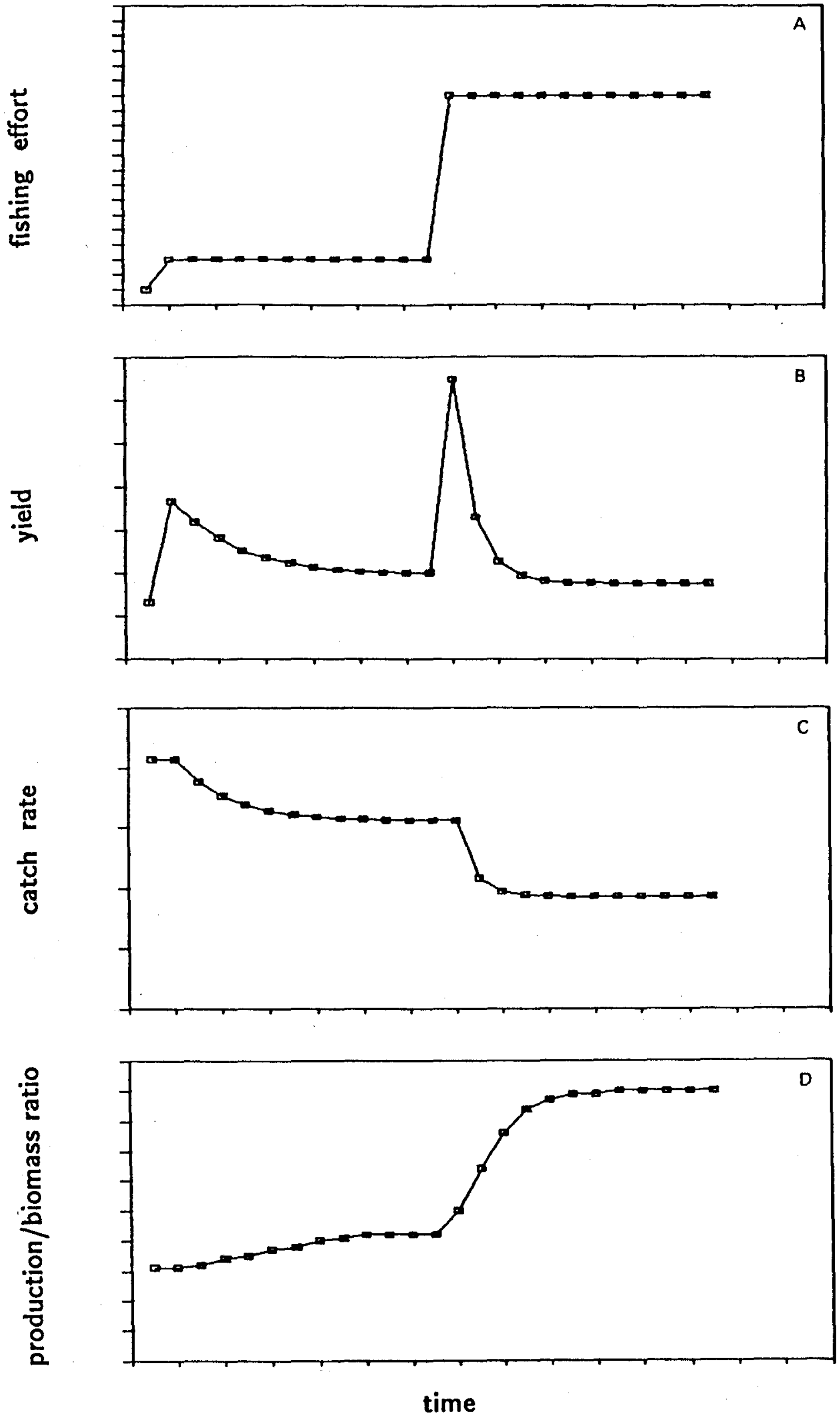


FIGURE 4. The transient state and transient paths. Response of the simulated stock when fishing effort increases: A) fishing effort; B) yield; C) catch rate; and D) production/biomass ratio.

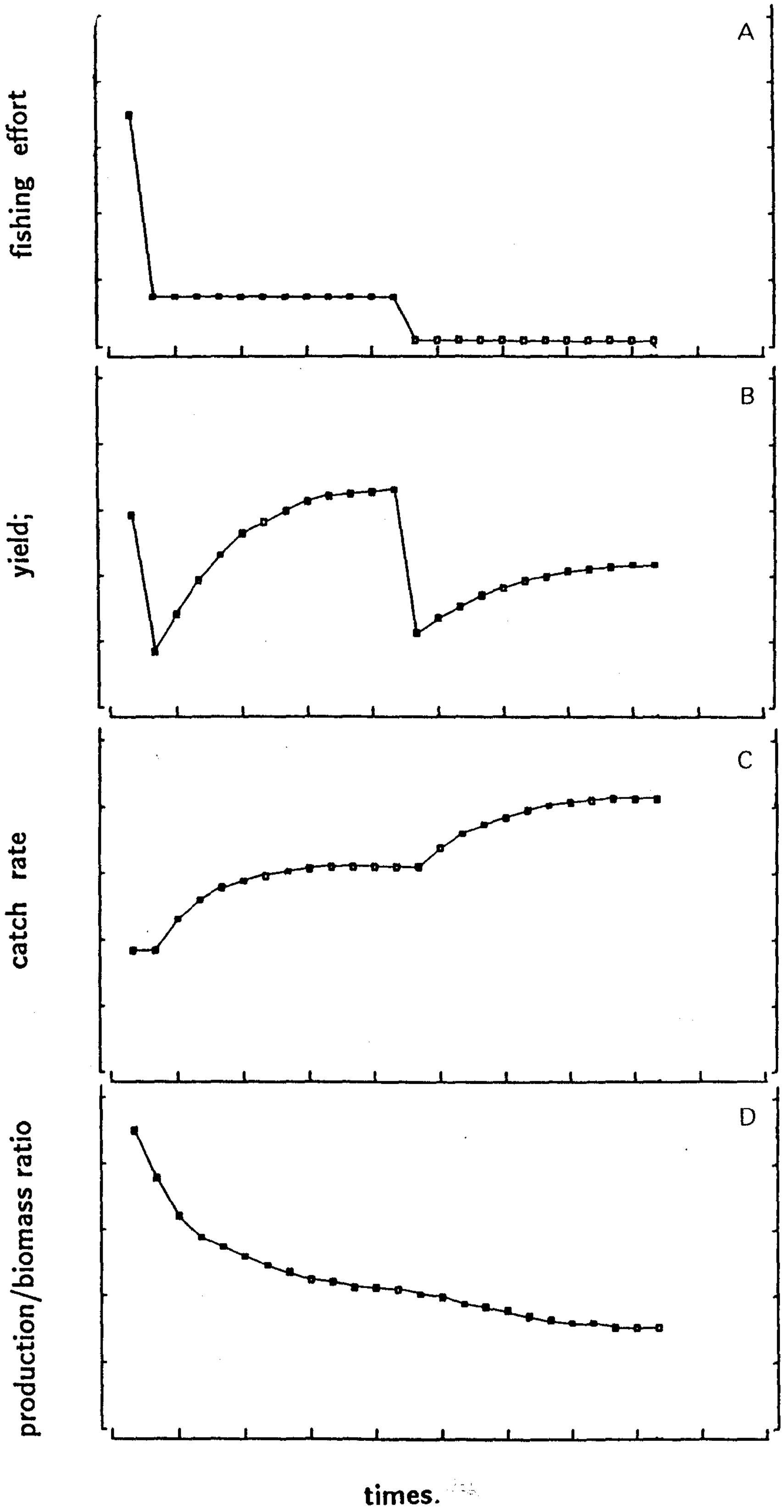


FIGURE 5. The transient state and transient paths. Response of the simulated stock when fishing effort decreases: A) fishing effort; B) yield; C) catch rate; and D) production/biomass ratio.

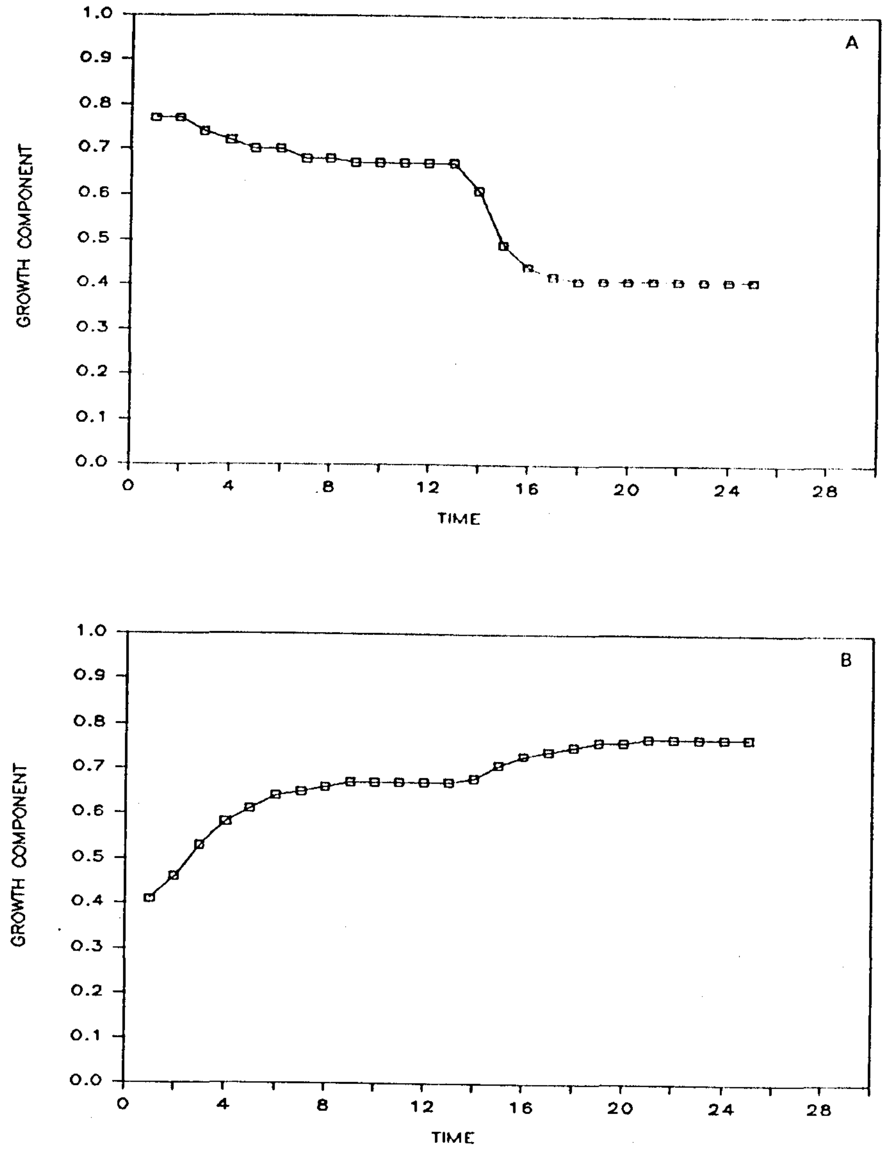


FIGURE 6. Response of the growth component of production when fishing effort varies: A) moving from low effort to higher effort levels; B) moving from high effort to lower effort levels.

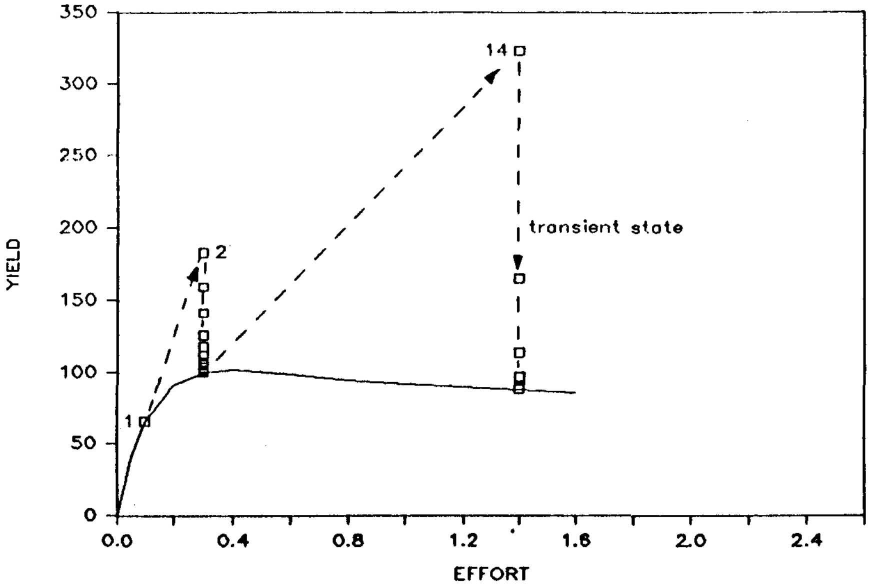


FIGURE 7. The transient path, as it appears in the yield/effort plane. The dashed line and arrows identify the temporal sequence. In this example, the simulated stock was at equilibrium in year 1. An increase of fishing effort in years 2 to 13 resulted in transient yields which are, at least momentarily, higher than the equilibrium yield. Fishing effort was further increased in year 14, for a substantial increase in transient yield. Finally, the transient yield levels in subsequent years decrease rapidly and stabilize around the equilibrium value.

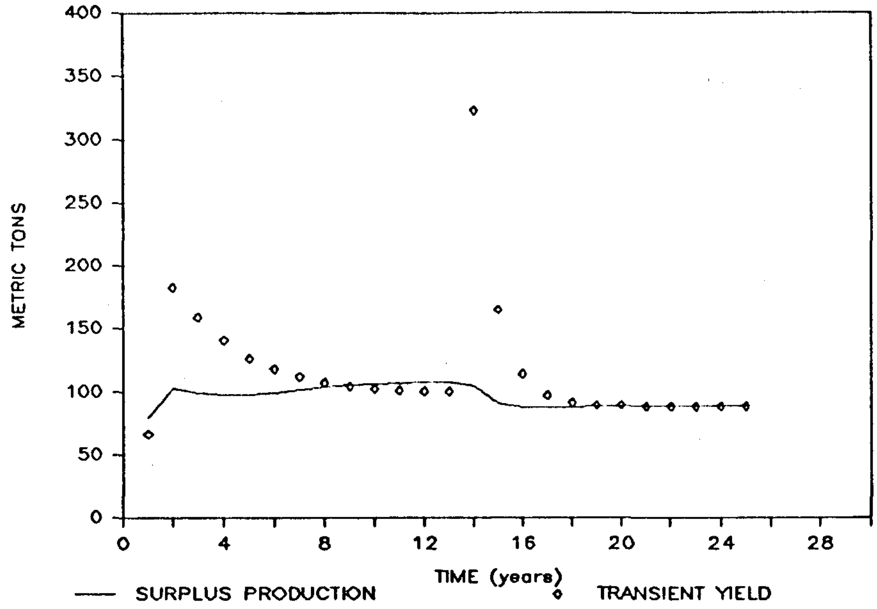


FIGURE 8. Surplus production (solid line) and yield (dots) for the simulated stock when fishing effort changes as in Figure 4A. Note that yield levels are equal to surplus production when the stock is in an equilibrium state but that yield is not equal to surplus production when the stock is in a transient state.

Surplus production and yield.

Surplus production is defined earlier in this chapter as a quantity which, if removed, will result in no change in stock level. Figure 8 illustrates the fluctuations of surplus production and yield for our simulated age-structured stock when fishing effort changes as in Figure 4A. The obvious conclusion is that yield levels are equal to surplus production levels when the stock is in an equilibrium state but that yield is not equal to surplus production when the stock is in a transient state. While this may seem obvious here, we will see below that equating surplus production and yield regardless of the stock status is a common procedure for estimating the parameters of production models.

Equilibrium state and environmental variability.

Even the most stable conditions in the marine ecosystem rarely lead to constant recruitment for fish stocks which are not exploited by man. For instance, recruitment is likely to vary in response to changes in temperature, current patterns and upwellings, and in response to changes in the biomass level of preys or predators. In a fish stock under exploitation, recruitment is also affected by changes in the age composition of the stock and by changes which occur as a result of changes in fishing pressure: e.g. high fishing mortalities reduce the relative importance of older fish, i.e. the fully mature fish. If recruitment varies around a given mean, say as $x \pm (0.2\ kx)$ where $k \sim \mathcal{N}(0,1)$, then the equilibrium yield and the equilibrium stock size appear as in Fig. 9. In other words, the equilibrium state is not described by a single curve but by a family of curves which describe an equilibrium zone. In fact, each equilibrium level leads to a different equilibrium curve. If a constant fishing mortality is applied to the stock, yield is expected on the long term to vary around a given mean. This behavior is illustrated in Fig. 10A. Thus, when recruitment is subject to environmentally determined variations, annual equilibrium yields are not expected to be constant. Attempting to remove a constant yield often leads to an unstable situation, as illustrated in Fig. 10 (C and D).

Effect of changes in recruitment and growth upon equilibrium yield.

We analysed above the response of equilibrium yield to effort and stock size when recruitment is constant or varies around a mean value. This situation may be representative, for example, of a snow crab fishery where size limits are such that only the males are exploited and the age at first capture is higher than the age at first maturity. Under these circumstances, recruitment is not likely to be affected by exploitation rates and variations in recruitment are mostly the result of environmental variability. The situation in which the spawning stock and the exploited stock overlap is described in Fig. 11 (cases C and D): the equilibrium yield increases rapidly as effort increases, reaches a maximum and decreases progressively thereafter as older age-groups are depleted. For most groundfish stocks in the Northwest Atlantic, the spawning stock and the exploited stock overlap, the degree of overlap being a function of mean age at first capture and of mean age of maturity (thus also a function of mesh size, fishing season, fleet behavior, etc.). Finally, the situation in which both growth and recruitment are affected by the effort level is represented in Fig. 11 (cases E and F): in this simulation, growth decreases progressively beyond a certain stock size. From these simulations, we conclude that the shape of the equilibrium yield/effort and equilibrium yield/stock size curves is affected by a wide variety of factors and, in particular, by the mechanisms controlling growth and recruitment. In general, the equilibrium yield/stock size curve is expected to be dome-shaped but the curve may be skewed.

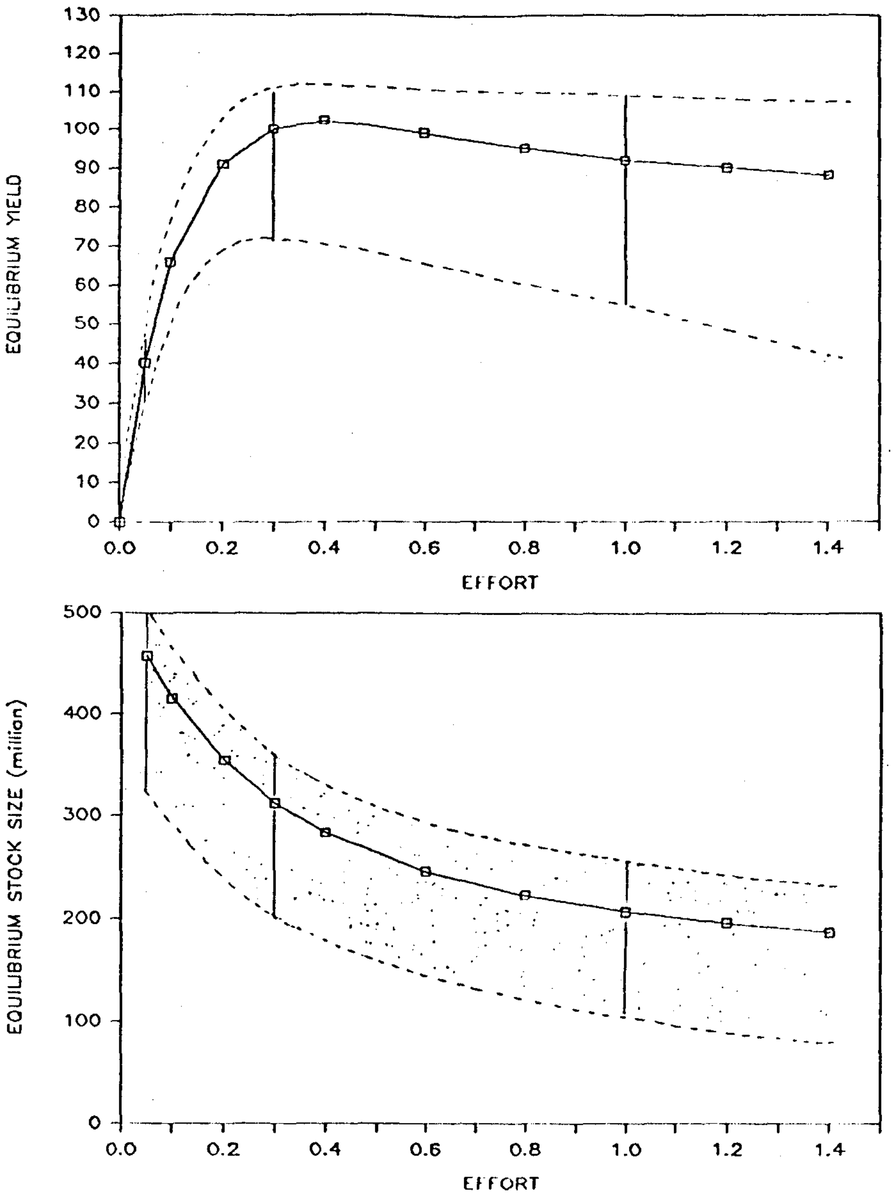


FIGURE 9. When recuitment varies around a given mean, the equilibrium state is not described by a single curve but by a family of curves which describe an "equilibrium zone". In other words, if a constant fishing mortality is applied to the stock, yield and stock size are expected to vary, on the long term, around a given mean.

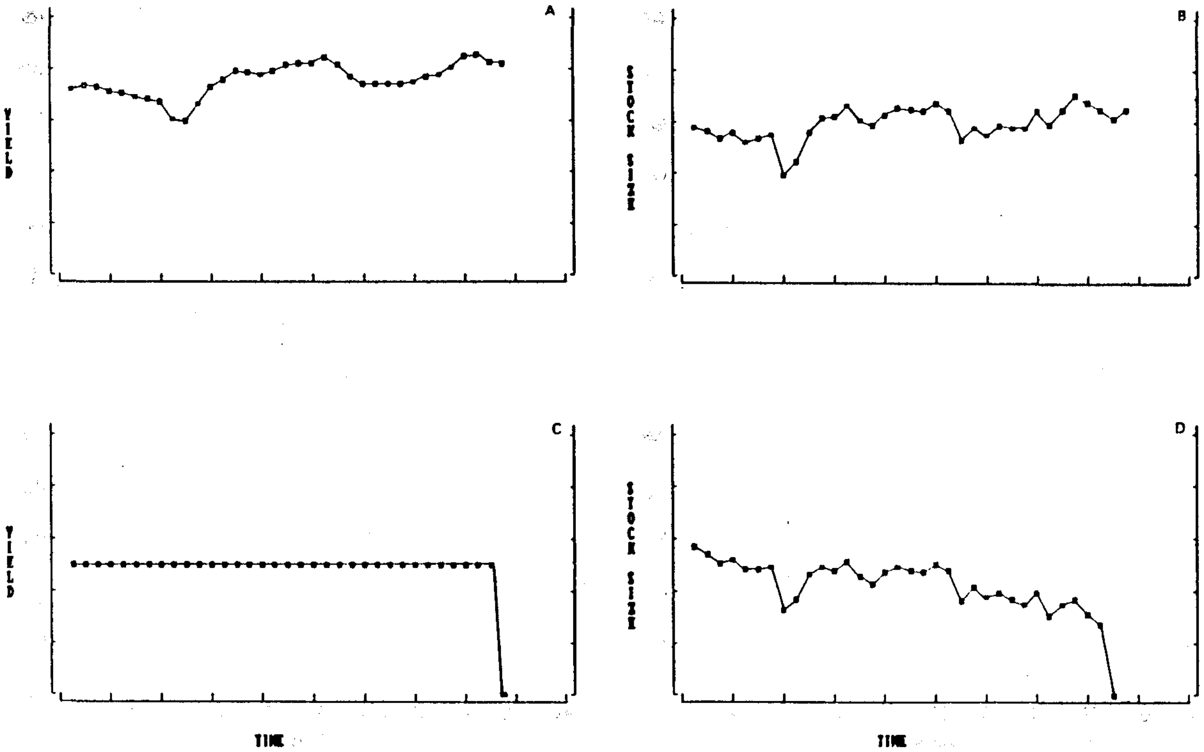


FIGURE 10. Variation of yield and stock size when recruitment is variable. For A) and B), a constant fishing mortality rate is applied. For C) and D), a constant yield is removed annually. Attempting to remove a constant yield leads to a declining stock size.

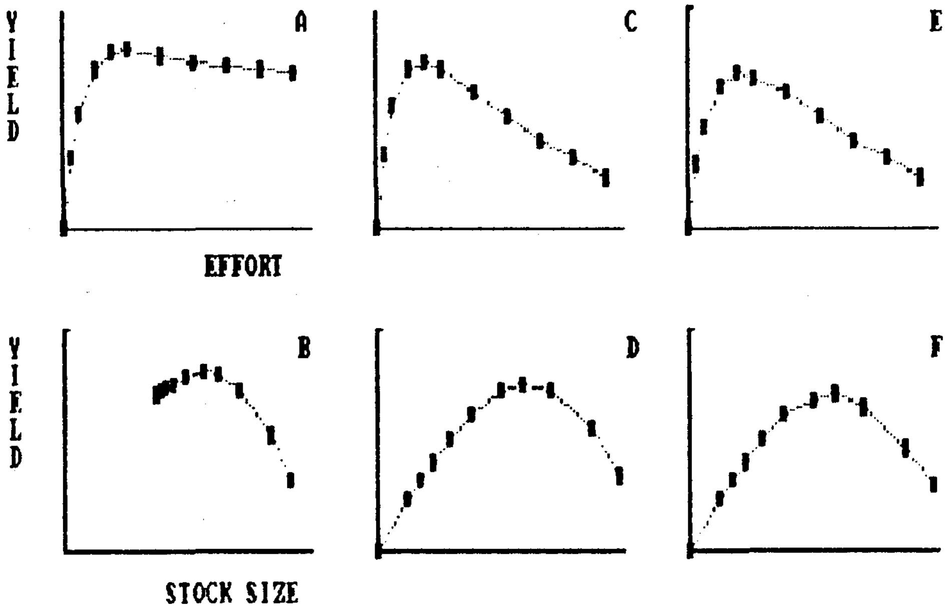


FIGURE 11. Yield from the simulated stock when:

- recruitment is constant or varies around a given mean (A and B);
- recruitment is affected by the effort level (C and D);
- both recruitment and growth are affected by the effort level: in this simulation, growth rate decreases progressively beyond a certain stock size (E and F).

The shape of the equilibrium yield/effort and equilibrium yield/stock size curves is affected by a wide variety of factors and, in particular, by the mechanisms controlling growth and recruitment.

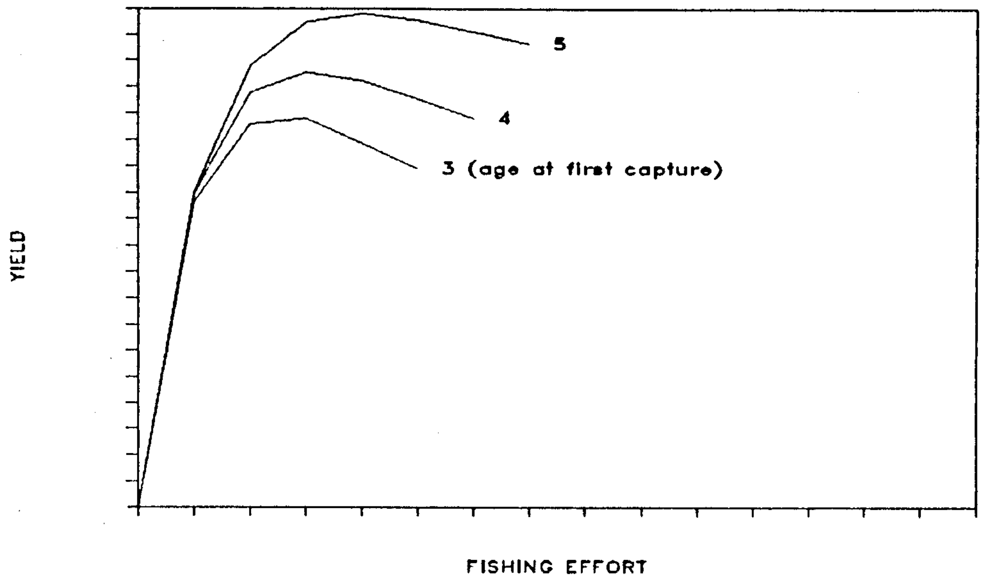


FIGURE 12. Response of the equilibrium yield curve to changes in mean age at first capture. From this, we conclude that changes in regulatory measures to control size limits, minimum mesh size or fishing seasons are all expected to influence equilibrium yield.

Effect of a change in age at first capture upon equilibrium yield.

While growth and recruitment can influence significantly the shape of the equilibrium yield curves, shape variations can also be introduced by changes in mean age at first capture (influenced by changes in size limit, mesh size and fishing season). For example, Fig. 12 illustrates the response of equilibrium yield changes in mean age at first capture. Consequently, changes in regulatory measures to control size limits, minimum mesh size or fishing seasons are all expected to influence equilibrium yield: in fact, any combination of effort levels and specific regulations on size limits, mesh size and fishing season will yield different equilibrium yield curves.

Summary.

The concepts of equilibrium state and transient state have been introduced above in the context of an age-structured fish population. If natural environment and recruitment are relatively stable, a given stock level will sustain a given yield; when this situation is realized, the stock is said to be in equilibrium and the yield obtained at that level is called "equilibrium yield". New production levels can be achieved by varying fishing pressure. A change in fishing pressure will perturb the equilibrium state of the stock and the stock will move towards a new equilibrium. The time necessary to move from one equilibrium level to a new one is referred to as the transition period. A stock in transition is often said to be in a transient state. From simulations, insight was sought on the shape of various equilibrium yield curves, on the behavior of transient yield under various conditions, as well as on the response of equilibrium yield to changes in mesh regulation, size limits, and fishing seasons. Our analysis of production revealed, for various fish stocks of the Northwest Atlantic, the relative contribution of recruitment and growth to production. The results of this section will serve as a basis for discussing the results of the next section on "classical production models".

Surplus production models: a general formulation.

For any small time interval (say Δt), we will assume that any change in stock size (say ΔB) will be the net result of the change in biomass resulting from growth, recruitment and natural mortality (say ΔB_P, or the production component) and the loss of biomass through fishing (say ΔB_Y, or the yield component). In other words,

$$\Delta B/\Delta t = \Delta B_P/\Delta t - \Delta B_Y/\Delta t \quad . \qquad (18)$$

We will also assume that the yield over Δt is a function of both the stock size and fishing mortality rate, say F:

$$\Delta B_Y/\Delta t = F_{t,t+1}\, \bar{B}_{t,t+1} \quad . \qquad (19)$$

If Δt tends to zero, we can rewrite (18) and (19) as

$$dB/dt = dB_P/dt - dB_Y/dt = B_P' - B_Y' \qquad (20)$$

and

$$dB_Y/dt = F(t)\, B(t) \quad , \qquad (21)$$

respectively. Since B_Y' represents the instantaneous yield at time t, $B_Y(t)$, the solution of B_Y', represents the cumulative yield from the stock up to time t. Thus we will simply write $Y(t)$ and Y' rather than $B_Y(t)$ and B_Y'. $B_P(t)$, the solution of B_P', represents a suitable production function and will be simply written $P(t)$. Finally, we write

$$B' = P' - Y' \quad , \qquad (22a)$$

$$Y' = F(t)\, B(t) \quad , \qquad (22b)$$

Equations (22a) and (22b) represent a system of differential equations. $B(t)$, the solution of B', represents the stock size at time t, while $Y(t)$, the solution of Y', represents the cumulative catch of the stock. $F(t)$ is a mortality function due to exploitation. As we shall see below, this basic form can be used to generate any of the classical production models.

The production function.

The production function describes the behavior of the unexploited stock. In other words, when $F(t) = 0$, $B' = P'$ governs population growth. A wide variety of expressions have been used to describe population growth. We can assume, for example, that the intrinsic rate of increase, say r, is constant. That is

$$B'/B = r \quad \text{or} \quad B' = rB = P' \quad . \qquad (23)$$

The solution of B' then becomes

$$B(t) = B_0\, e^{rt} \qquad (t \geq 0) \quad . \qquad (24)$$

where B_0 is the biomass of the stock at some arbitrary set time t=0. Alternatively, the logistic equation assumes that the rate of increase of the stock decreases at higher stock densities:

$$B'/B = r - bB \quad \text{or} \quad B' = B\,(r - \mathrm{b}B) = P' \quad . \qquad (25)$$

As B increases, the term $r-bB$ becomes smaller and the rate of increase of the stock decreases until it is zero. The stock size at which the rate of increase is zero is referred to as the carrying capacity of the environment: we refer to it as B_∞. Consequently, if B=0, $B_\infty = r/b$ and we can rewrite the logistic equation as

$$B' = rB - (r/\mathrm{B}_\infty)\, B^2 \quad . \qquad (26)$$

As B approaches B_∞, the rate of change of biomass (defined by the preceding equation) approaches zero and the biomass stops increasing. The solution of B' can be written as

$$B(t) = B_\infty / (1 + \mathrm{e}^{-r(t-t_0)}) \quad . \qquad (27)$$

The logistic equation yields an S-shaped curve. An important characteristic of this curve is that its inflexion point occurs at $B_\infty/2$. A more general form can be used in which the inflexion point can lie anywhere between zero and B_∞. That is

$$B' = rB - r\mathrm{B}_\infty (B/B_\infty)^{\mathrm{n}} \quad . \qquad (28)$$

A function of this form is often referred to as the Richards function. As indicated in Fig. 13, this population growth curve can take a wide variety of shapes depending upon the value of its parameters. The preceding models assume that the population is confined to a limited geographical distribution. If this distribution is unknown, it is appropriate to introduce the concept of "carrying capacity per unit area". In other words, if $B_\infty = KA$, the production function becomes

$$B' = rB - rKA\ (B/KA)^n \quad , \qquad (29)$$

where A refers to the area of interest. If the area fished is kept constant during a period t, say $A=A_t$, the carrying capacity is seen to be proportional to A_t:

$$B_{\infty,t} = K A_t \quad . \qquad (30)$$

The mortality function.

The mortality function $F(t)$ is often described by the relationship

$$F(t) = q(t)\ f(t) \quad . \qquad (31)$$

where $q(t)$ is the "catchability function" and $f(t)$, the fishing effort at time t. In practice, though, for each finite time interval over which yield statistics are integrated (usually a year), fishing effort is generally assumed to be constant. It follows that $f(t)$ appears as a step function that describes the effort as being constant over each time interval t, with abrupt changes at the end of each period. Consequently, the quantity f_t will represent the fishing effort for the time interval t and q_t, the catchability coefficient. In general, catchability is assumed constant over time and q_t is simply denoted as q.

If the fishing area is changing over time, then the instantaneous fishing mortality rate appears as a function of f_t and A_t. For a given stock and a given fishing effort, the instantaneous fishing mortality rate in the area fished will be assumed to be inversely proportional to the size of the fishing area:

$$F_t = a\ f_t\ /\ A_t \quad , \qquad (32)$$

where a is the proportionality factor.

A synoptic form.

As a means of introduction to historical forms of surplus production models, we will adopt the following synoptic form which is derived from equations (22), (29) and (32):

$$B' = r\ B(t) - r\ K\ A(t) \left[\frac{B(t)}{K\ A(t)} \right]^n - Y' \quad , \qquad (33a)$$

$$Y' = a\ f(t)\ B(t)\ /\ A(t) \qquad (A(t) > 0) \quad . \qquad (33b)$$

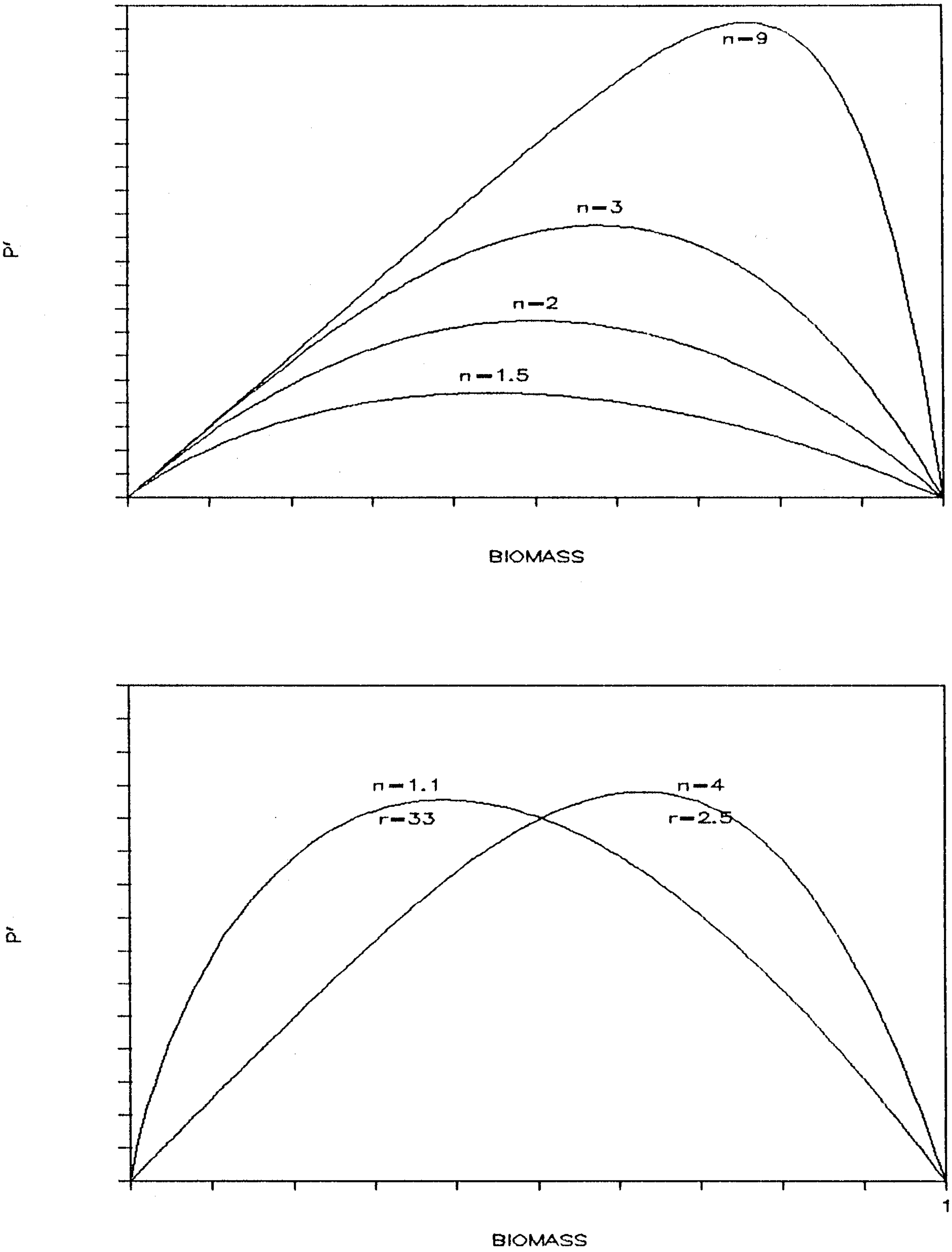


FIGURE 13. The production function.

This model includes two control variables: the fishing effort, $f(t)$, and the total fishing area, $A(t)$. Because the size of the fishing area appears explicitly in this form, the model can be used to simulate the exploitation of a stock under expansion of the total fishing area. This situation is typical of an expanding or developing fishery. Adopting a synoptic form will facilitate the discussion of classical stock production models.

The equilibrium model.

If fishing removes the surplus production from the stock at the same rate as it is produced, the stock size remains constant and the stock is held in equilibrium. In order to find the levels of equilibrium yield, we set $B'=0$ in equation (33a):

$$0 = r B^* - rKA (B^*/KA)^n - (afB^*/A) \quad , \qquad (34)$$

where B^* is used to indicate that the stock is at equilibrium. Because fishery statistics are generally produced over finite time intervals, we will use the subscript t to identify the equilibrium state during each time interval t. Consequently, we rewrite equation (34) as

$$r B^* - rK A_t (B^*/KA_t)^n = a f_t B^* / A_t = Y_t^* \quad , \qquad (35)$$

where Y_t^* is the equilibrium yield over the interval t. The maximum value of Y_t^* is given by

$$Y_{max}^* = rK A_t (n-1) / n^{n/(n-1)} \quad . \qquad (36)$$

The maximum equilibrium yield, often referred to as the maximum sustainable yield, will be written hereafter as Y_{MEY}.

Equation (35) defines surplus production under equilibrium conditions as a function of stock size. Surplus production can also be related to the instantaneous rate of fishing mortality, say $F(t)$:

$$Y_t^* = K A_t (F_t^{n-1} - (F_t^n / r))^{1/(n-1)} \quad . \qquad (37)$$

The rate of fishing mortality at maximum equilibrium yield can be calculated as (see Fletcher, 1978)

$$F_{MEY} = Y_{MEY} \, n^{n/(n-1)} / K A_t n \quad . \qquad (38)$$

The case where $n=2$ is illustrated in Fig. 14. In this case,

$$Y_t^* = K A_t F_t (1 - (F_t / r)) \quad , \qquad (39)$$

and the equilibrium levels become negative when $F_t>r$, since F_t, A_t and K are positive quantities. In other words, the actual stock is doomed to extinction if the instantaneous fishing mortality is greater than the instantaneous rate of increase of the stock at densities approaching zero. The complete response of the equilibrium yield to the control variables F_t and A_t is represented in Fig. 14A. Given a fixed mortality F_t, the equilibrium yield appears to be proportional to the actual size of the fishing area.

From equation (39), the equilibrium yield can be expressed as a function of f_t:

$$Y_t^* = a K f_t (1 - (a f_t / r A_t)) \quad , \qquad (40)$$

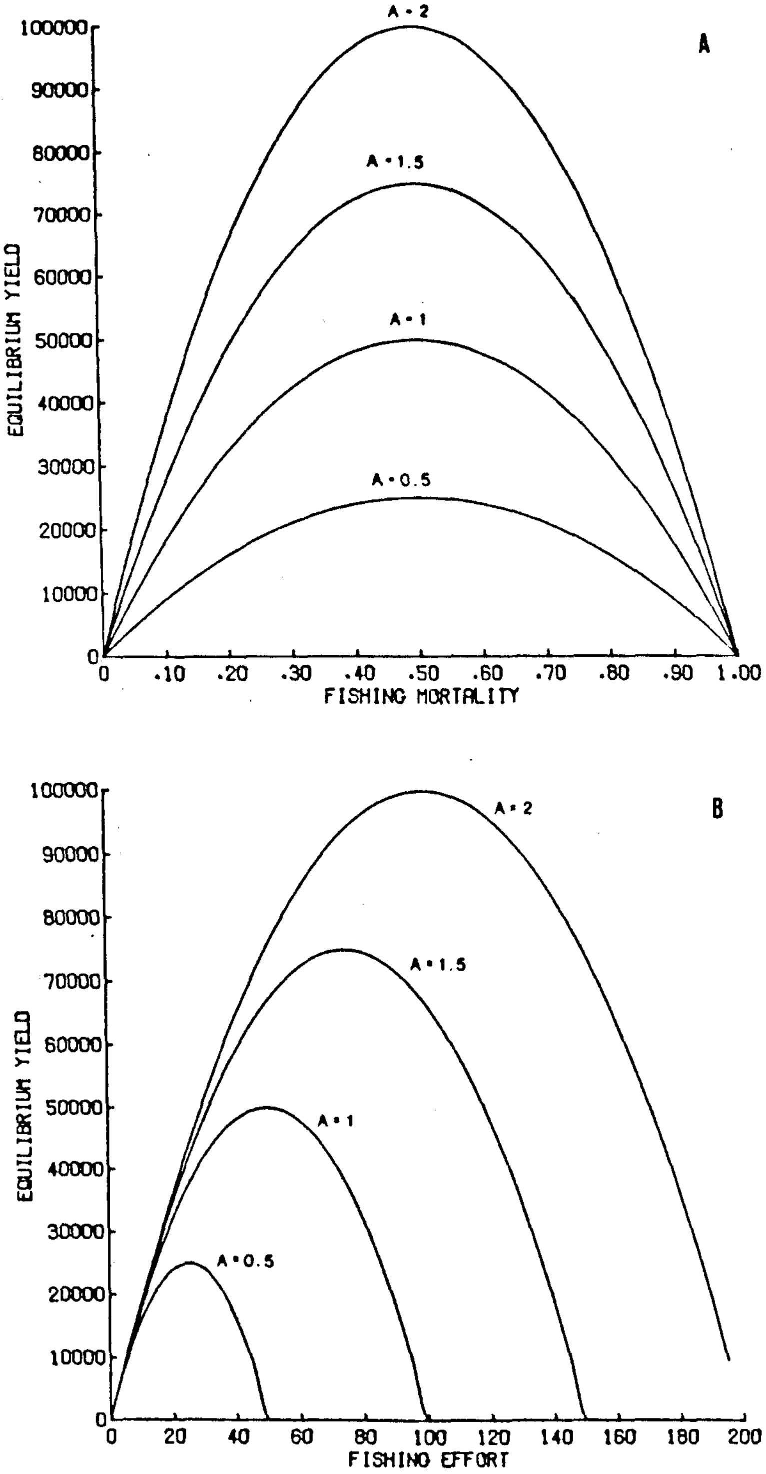


FIGURE 14. The synoptic form: response of the equilibrium yield to the instantaneous fishing mortality (A) and to the fishing effort (B) when the total fishing area is expanding.

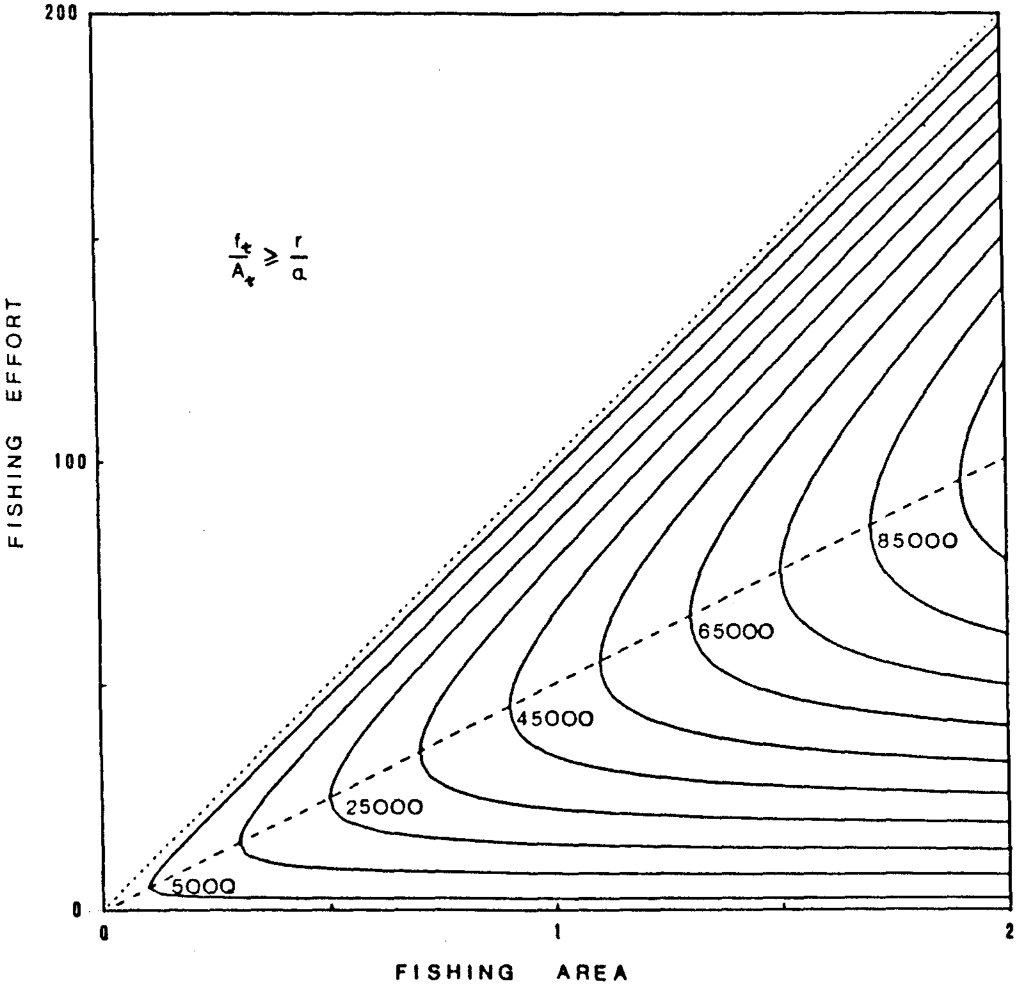


FIGURE 15. Isopleths for the equilibrium yield. The dashed line represents the locus of all points which provide the "maximum equilibrium yield" for a given fishing area. The region above the dotted line represents the locus of all points which lead to extinction of the exploited stock within the fishing area A_t.

The response of the equilibrium yield to the control variables f_t and A_t is illustrated in Fig. 14B. As the total fishing area increases, the maximum productivity Y_{MEY} increases also. Of equal interest are the isopleths for the equilibrium yield, as given in Fig. 15 (solid lines). Each isopleth represents the locus of all points (A_t, f_t) which generate equal values for the equilibrium yield. In turn, the dashed line represents the values of f_t which provide the maximum sustainable yield for a fixed area A_t. That is,

$$f_t = r A_t / 2 a \quad . \qquad (41)$$

As in Fig. 14B, we also observe that the maximum productivity of an area is increasing as the size of the fishing area increases. However, we can now chose a series of points (A_t, f_t) which provide equal values for the equilibrium yields. For instance, by moving along the isopleths, we see that an expansion of the total fishing area allows a reduction of the total effort without loss in the long term yield (below the dashed line in Fig. 15). At low fishing levels, we also observe that the equilibrium yield is less affected by the size of the total fishing area; in fact, a wide range of values for A_t will generate equal equilibrium yields. Finally, the upper section of Fig. 15 (above the dotted line) represents the values of the fishing effort and the sizes of the total fishing area which lead, when employed simultaneously, to extinction of the exploited stock within the given fishing area A_t. This region is the locus of all points where $(f_t/A_t) \geq (r/a)$.

Classical models

The synoptic model developed above will be used to introduce certain production models commonly used for fisheries management. The "classical" models can be seen as a particular case of the synoptic form.

The Graham-Schaefer model.

The earliest form of surplus production systems as applied to fisheries data is due to Graham (1935). Graham suggested a parabolic form for the production function. Schaefer (1954) suggested to use indices of mean stock biomass for estimating surplus production on an annual basis. Schaefer's work extended the applicability of the model. The Graham-Schaefer model uses the parabolic production function and assumes a constant fishing area. Consequently, if n=2 in the synoptic model and if $B_\infty = KA$, we obtain the parabolic form of Graham-Schaefer:

$$Y^* = rB^* - (r/B_\infty) (B^*)^2 \quad . \qquad (42)$$

Fletcher (1978) reformulated the model so as to introduce the maximum equilibrium yield, say Y_{MEY}, as a parameter. That is

$$Y^* = 4 \, Y_{MEY} \, (B^*/B_\infty) - 4 \, Y_{MEY} \, (B^*/B_\infty)^2 \quad . \qquad (43)$$

The maximum productivity, i.e. Y_{MEY}, which is in fact the intrinsic rate of increase of the stock at the inflexion point of the production function, now appears explicitly in the equation and replaces r, the intrinsic growth rate at low stock densities.

The following catch per unit effort against effort relationship is underlying the Graham-Schaefer model:

$$Y^*/F^* = B_\infty - (B_\infty / r) F^* \quad . \tag{44}$$

Substituting qf^* for F^*, we find

$$Y^*/f^* = a - b f^* \quad , \tag{45}$$

where $a = q B_\infty$ and $b = q B_\infty / r$. In other words, the equilibrium catch per unit effort decreases linearly as a function of the equilibrium effort (Fig. 16).

The Gulland-Fox model.

Gulland (1961) suggested to relate the yield per unit of effort on a stock in a given year and the average fishing effort for earlier years. Gulland suggests that the average effort for earlier years is closer to the equilibrium effort (f^*). In other words, $\bar{f}_t$ over a certain number of years is assumed to approximate the equilibrium effort. Fox (1970) developed an algorithm for fitting a negative exponential relationship of the form $y = a\, e^{-bx}$ to data treated by Gulland's method. The Gulland-Fox model is thus based on a curvilinear relationship between catch per unit effort in a given year and the average effort for earlier years (Fig. 16):

$$\ln (Y/f) = a - b \bar{f} \quad . \tag{46}$$

If U_∞ denotes the equilibrium catch per unit effort when the population is at its maximum equilibrium size, the Gulland-Fox model is written in our notation as

$$Y^*/f^* = U_\infty\, e^{-f^*/f_{\text{MEY}}} \quad , \tag{47}$$

where f_{MEY} is the equilibrium fishing effort which provides maximum equilibrium yield. We can also express equilibrium yield as a function of stock biomass; that is,

$$Y^* = B^* k (\ln B_\infty - \ln B^*) \quad . \tag{48}$$

One of the characteristics of the exponential production model is that stock size which produces maximum equilibrium yield is always 37% (i.e. 1/e) of the maximum stock size. It can be shown that the synoptic model approximates the Gulland-Fox model when n tends towards 1 and $B_\infty = KA$.

The Pella-Tomlinson model.

The Pella-Tomlinson production model uses the Richards form of the production function and assumes a constant fishing area. Consequently, if the fishing area is constant and $B_\infty = KA$ in the synoptic model, we obtain

$$Y^* = r B^* - r B_\infty (B^*/B_\infty)^n \qquad (n \neq 1) \quad . \tag{49}$$

As in the Graham-Schaefer model, we can reformulate the model so that the maximum equilibrium yield appears explicitly in the equation. That is

$$Y^* = \gamma Y_{\text{MEY}} (B^*/B_\infty) - \gamma Y_{\text{MEY}} (B^*/B_\infty)^n \tag{50}$$

where

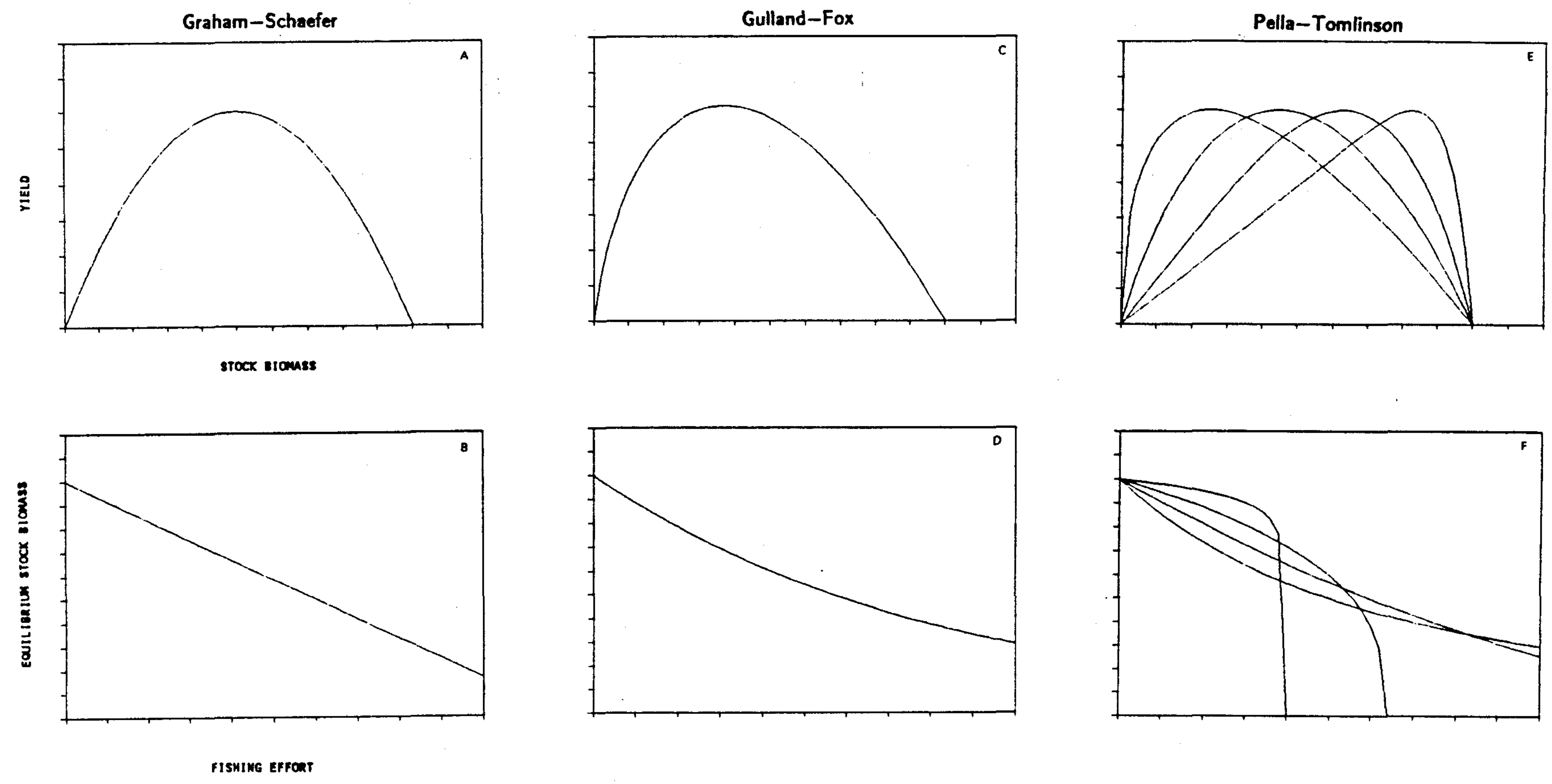


FIGURE 16. Classical models: the Graham–Schaefer model (A and B); the Gulland–Fox model (C and D); the Pella–Tomlinson model (E and F).

$$\gamma = n^{n/(n-1)} / (n-1) \qquad (n \neq 1) \quad .$$

The exponent *n* controls the location of the inflexion point in the production function. The Pella-Tomlinson model is thus a general model capable of generating a wide variety of curves (Fig. 16). For example, when $n=2$, equation (49) yields a Graham-Schaefer model. Similarly, the Pella-Tomlinson model approximates the Gulland-Fox model when *n* tends towards 1.

The catch per unit effort against effort relationship underlying the Pella-Tomlinson system appears as

$$Y^*/f^* = [\, U_\infty^{\,n-1} - (\, U_\infty^{\,n} / \gamma\, Y_{\mathrm{MEY}} \,)\, f^* \,]^{1/(n-1)} \quad . \qquad (51)$$

Underlying assumptions and departures from them.

The above models assume a production function in which the unexploited population reaches a maximum value and remains at this constant density if not perturbed. In general, unexploited fish populations fluctuate simply because of environmental variability and trends: climatic changes and increased environmental stress due to pollution are examples of perturbing factors. Environmental homeostasis thus appears as an underlying assumption for classical models. This assumption becomes increasingly important as the period covered by the analysis increases. Classical models assume that fishing effort is the only variable factor affecting the fish stock. The variation in recruitment in response to environmental changes and variations due to interactions with other species represent obvious departures from this assumption.

No age structure. Classical models are global models which do not take into account differences between age-groups and interactions between age-groups. Population is treated as a whole, with no specific reference to its age structure. As indicated by Gulland (1969), population growth depends as much, and if not more, on the age structure of the stock than on its total biomass. Nevertheless, the production functions used in classical models treat all individuals of the stock as equal. Because of this inability to consider relative changes in the age composition of a stock, these models are expected to mimic populations or stocks showing a stable age structure; deviations from the stable age structure are thus assumed to have negligible effects on the production rate (Poole, 1974). In practice, changes in fishing effort generate changes in the age structure of the population and the number of age-groups being fished. In addition, the response of the rate of increase of the stock to an increase in density is, in these models, instantaneous. In natural populations, because of the presence of many age-groups, the contribution of recruitment to annual production is lagged.

Closed population. No provision is made for migrations. Consequently, these models must be applied to a closed population. Classical models also assume no modification of the geographical distribution of the stock or no extension of the area being fished to new or adjacent areas.

Equilibrium state. In most applications of classical models, the concept of equilibrium state must apply to the population being considered. By definition, equilibrium is achieved when the rate of increase of a stock submitted to a constant fishing effort reaches or approaches zero. At equilibrium, a population will not change in size or characteristics (Beverton and Holt, 1951). Equilibrium conditions are not realized when fishing effort changes rapidly.

Constant catchability. In applying surplus production models to generate management advice, problems also arise with the definition of fishing effort. Due to changes in gear types or fleet composition, standardization of fishing effort to arrive at a unique effort index is often necessary. Over long time periods, an absolute measure of fishing effort may itself constitute a poor index of fishing mortality because of the presence of changes or trends in the efficiency of a fleet. The effect of learning factors and technological innovations on the effective fishing mortality is difficult, and often impossible, to quantify. It is also difficult to obtain a measure of fishing effort which is revealing of the fishing mortality on the stock for species which exhibit a schooling behavior, like herring, mackerel and capelin. Since direct measures of catch per unit effort -for example, catch per set - are only indicative of the density and may be of the size of a particular school (which factors may be independent of total stock size), a measure of search for fish must also be considered when applying surplus production models to these stocks. For such species, the problem is often compounded by the phasing in of technological developments (like sonar equipment) during the period of interest. If technology make it so easy to locate schools that locating one school when stock size is low is only slightly more difficult than locating schools when stock size is high, the measure of effort will be insensible to variations in fishing mortality. This limits the applicability of classical production models for such species.

The stock is defined as a static entity. In general, fish are only partly recruited to the fishery, only certain age-groups being subject to fishing pressure. However, changes in fishery regulations (e.g. mesh size regulation, season closures, area closures) may induce important changes in the definition and characteristics of the exploited portion of the population and thus change the characteristics of the production function and the mortality function (through catchability). The application of classical models further assumes that fishing effort is relatively well distributed over the stock or that catch rates are indicative of the stock size. This assumption is unlikely to be met in an expanding fishery or when stocks overlap at certain times of the year.

In view of the above, classical models should not be considered more than a convenient, but crude, representation of a very complicated process. They are, however, the only tool available for fish species which cannot be aged or in cases when only catch and effort data are available. An important point to remember is that only a few parameters are used to define the growth characteristics of the stock and the characteristics of the exploitation function, and that these parameters are time independent (i.e. invariant in relation to time). The production functions inherent to classical models make no distinction between the effect of density on the growth rates and on death rates. In natural populations, density can be expected to affect growth and larval survival, or both. Instead, the production functions assume a global "net" rate of increase.

The transient system

In order to relate directly annual catch and annual fishing effort, classical models were formulated under the assumption of equilibrium conditions. Equilibrium states can be achieved, as we saw, when fishing effort is kept constant or changes only slowly over the period covered by the analysis. However, it is also of interest to describe population changes during a transition period, for example when the stock moves from one equilibrium state to a new equilibrium (e.g., as a result of moving from one fishing level to a new one). From the manager's viewpoint, the transition period can be characterized in different manners. How long will it take to move from one equilibrium state to a new one? What will the intermediate levels of the stock be during the transition period? These questions are particularly important since fish stocks rarely are in equilibrium, being constantly perturbed by varying fishing effort levels. It is thus important to study the exploitation of a fishery under non-equilibrium conditions and to develop strategies for attaining the desired stock levels which consider the transient states. Walter (1976) introduced a framework, based on the Graham-Schaefer model, to calculate the transient yield.

Transient yield levels.

Transient yield levels can be obtained by integrating equation (22b). That is,

$$Y_t = \int_t F(t)\, B(t)\, dt \qquad . \qquad (52)$$

If the fishing effort, say f_t, is constant over the period t and directly proportional to the instantaneous rate of fishing mortality, then

$$Y_t = q\, f_t \int_t B(t)\, dt \qquad . \qquad (53)$$

The stock biomass at time t, $B(t)$, is itself given by the solution of equation (22a). Following Walter (1976), we assume that B' has the Schaefer form; that is,

$$B'/B = r - (r/B_\infty)\, B - q\, f_t \qquad . \qquad (54)$$

Substituting the solution of (22a) in (51), we obtain an expression which relates the transient yield to fishing effort:

$$Y_t = (qfB_\infty/r)\,(r-qf) - (qfB_\infty/r)\,\ln(B_0/B_1) \qquad , \qquad (55)$$

where B_0 identifies the biomass at the beginning of the period t and B_1, the biomass at the end of the period. If the value of the equilibrium yield is substituted by Y_t^* in (55), we write

$$Y_t = Y_t^* - (qfB_\infty/r)\,\ln(B_0/B_1) \qquad , \qquad (56)$$

The transient yield, Y_t, will be either greater or smaller than the equilibrium yield, Y_t^*, depending on whether B_0 is greater or smaller than B_1. For example, if the optimum effort level (f_{MEY}) is applied to a virgin stock (B_0=B_∞), a yield greater than the equilibrium yield is expected for the first period t. The trajectory of the yield in subsequent years appears in Fig. 17. There is a striking similarity between this trajectory and the transient path resulting from the simulation of an age-structured stock under exploitation (Fig. 7 and 8). The reader notes that B_1, the biomass at the end of the period, depends entirely on B_0, the initial state, and effort. In fact, B_1 can be calculated from the solution of (51), with initial conditions $B(0)$=B_0. That is,

$$B_1 = (1-x)B_\infty \,/\, [\, 1 - (\, 1 - (1-x)(B_\infty/B_0)\,)\, e^{-(1-x)r}\,] \qquad , \qquad (57)$$

where $x = qf/r$. The solution for B_1 may be substituted into (55) to cast the transient yield as a function of f, B_0, q, r and B_∞.

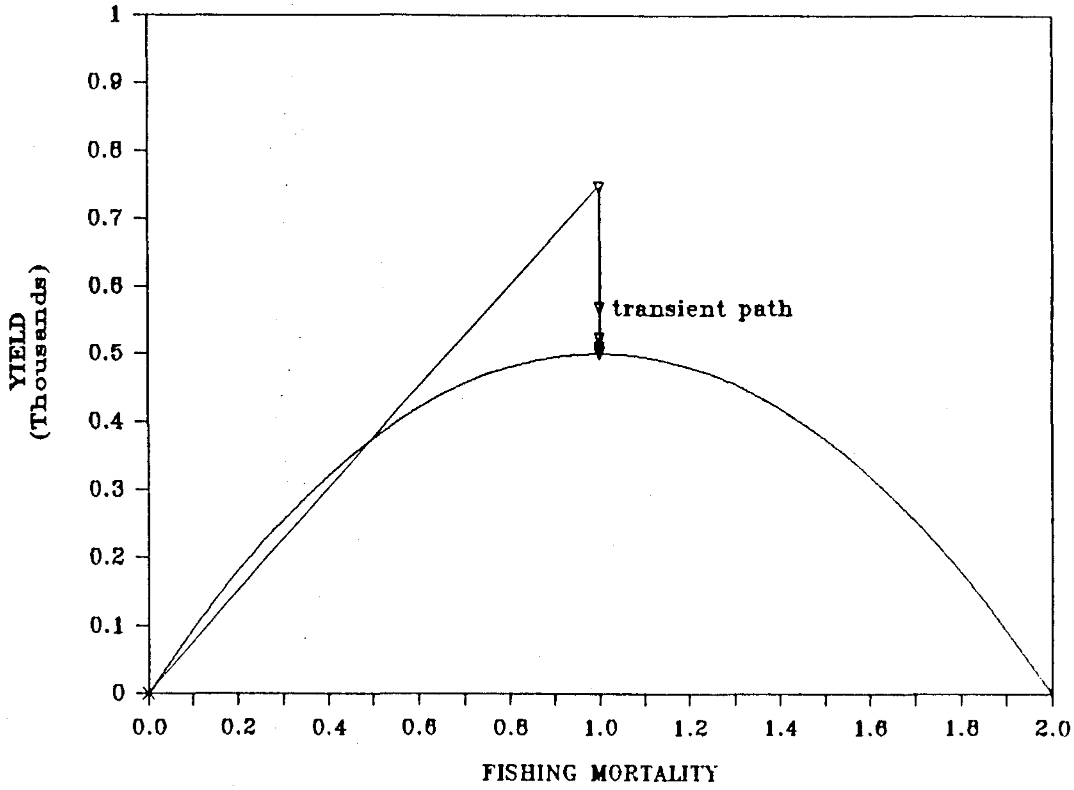


FIGURE 17. The transient path: trajectory of the yield when F_{MSY} is applied to a virgin stock. This trajectory is similar to the transient path resulting from the simulation of an age–structured stock under exploitation (Figures 7 and 8).

Recovery time for overexploited stocks.

It is also instructive to calculate the time it takes to move from a given stock level to a new one. The length of the transition period, say t, may be obtained by solving equation (53) for B_t, given the initial level B_0. During the transition period, the fishing effort will be assumed constant. The solution of equation (53), under this set of conditions, is given by

$$B_t \;=\; (1-x)B_\infty \;/\; [\; 1 - (\; 1- (1-x)(B_\infty/B_0)\;)\; e^{-(1-x)rt}\;] \qquad , \qquad (58)$$

Equation (58) can be used to calculate the time required to reach the equilibrium stock level which gives the maximum equilibrium yield. Consequently, setting $B_t=B_\infty/2$ in (58) and solving for t, we find

$$t \;=\; [\; 1/(1-x)r\;]\; \ln\; [\; (\; (1-x)(B_\infty/B_0) - 1\;)\; /\; 1-2x\;] \qquad , \qquad (59)$$

Since equation (59) is minimized when x=0, the shortest length for the transition period is given by

$$t \;=\; [\; 1/r\;]\; \ln\; [\; (B_\infty/B_0) - 1\;)\;] \qquad . \qquad (60)$$

The monitoring of catch and effort

In the above models, effort appears as an important variable. Fishing effort is, in fact, our only control: in other words, changes in the stock and in yield can only be induced by changes in fishing effort. The monitoring of fishing effort is thus an important and essential element of the application of surplus production models. In general, two sources can be used for the monitoring of fishing effort: commercial records (e.g. logbooks) and surveys independent of commercial data. Sometimes, both sources can be used.

In general, it is possible to find a measure of fishing effort which is easy to monitor: "hours fished", "number of hook-hours", or "number of crab pots used", are examples of such measures. The problem of defining the unit effort to be measured takes another dimension when it is necessary to ensure that a given unit effort keeps the same meaning over the course of many years. When applying surplus production models, the statistics of 15, 20 or more years are often utilized. Both the measure of fishing effort and its meaning in terms of effective fishing mortality must remain constant over the time period considered. Consequently, we have to find a measure of fishing effort which is representative of fishing mortality during the period of interest. Ideally, if estimates of fishing mortality were available for the whole period covered by the analysis, these could be used to establish the desired relationship between production and fishing mortality. Since direct estimates of fishing mortality are difficult to obtain and generally unavailable for long time periods, we must use an index of fishing mortality, namely a measure of fishing effort. In most situations, the problem of finding an index of fishing mortality will be complicated by the presence of more than one gear or fleet component in the fishery. In these cases, standardization techniques have to be applied in order to arrive at a single index of fishing effort. Standardization techniques are discussed elsewhere in this manuel.

Parameter estimation

Surplus production models relate catch and fishing effort. From observations on both catch and effort, we can estimate the parameters of the production function and derive from these estimates the characteristics of the *equilibrium* curve. It should be emphasized here that exploited stocks generally depart from equilibrium conditions; consequently, the observations available on catch and effort refer to transition periods. A wide variety of methods have thus been developed in order to correct the observed transient points by moving these closer to the equilibrium curve. Such methods are said to be based on equilibrium approximations. The classical forms, as written above, assume that the population is in equilibrium and therefore need strictly to be applied to observations coming from a period during which the fishing effort and other factors are constant, a condition which is unlikely to be met (Gulland, 1978). The failure to meet it can be tackled in different ways, as we discuss hereafter.

Direct estimation by assuming periodic equilibrium.

Schnute (1977) distinguishes between "continuous equilibrium", i.e. the population biomass is constant for all time, and "periodic equilibrium", i.e. the biomass returns to the same level at the beginning of each year. The assumption of continuous equilibrium leads to production models of the forms given by equations (35) and (40). "Periodic equilibrium" constitutes a much weaker assumption than "continuous equilibrium": all that is needed is that the population returns to the same biomass level at the beginning of each year (Schnute , 1977). For example, Schaefer (1954) estimated the coefficients of his model by relating statistics which were readily available, i.e. the average catch per unit effort over year i, say $\bar{U}_i$, and the total annual effort f_i, in the following manner:

$$\bar{U}_i = a + b\, f_i \quad . \qquad (61)$$

An equation of this form arises from the assumption of periodic equilibrium.

Gulland's method of equilibrium approximations.

As a means to estimate the parameters of the Graham-Schaefer model, Gulland (1961, 1969) suggests a method in which the catch per unit effort in a given year is related to the fishing effort averaged over a given number of years (say k) that a year class contributes significantly to the catch. That is,

$$U_i = g(\bar{f}_i) + e \quad , \qquad (62)$$

where

$$\bar{f}_i = \sum_{j=i-k+1}^{i} f_j \,/\, \mathrm{k} \quad . \qquad (63)$$

Gulland suggests that the number of years included in the summation be equal to half the duration of the exploited phase (see Cadima, 1976) and argues that the line thereby obtained will be very close to the equilibrium relationship. Gulland's method of equilibrium approximations can be applied to anyone of the classical models. For the Graham-Schaefer model, we obtain the following linear model,

$$U_i = a + b\,\bar{f}_i + e_i \quad . \qquad (64)$$

in which *a* and *b* can be estimated by simple regression techniques. In theory, for the production curve to be valid, a negative slope (*b*) is expected. Then the estimates of *a* and *b* can be used to calculate the maximum equilibrium yield, say MEY, the catch per unit effort at MEY and the fishing effort which gives MEY:

$$Y_{\text{MEY}} = -\,a^2 / 4b \quad , \qquad (65)$$

$$U_{\text{MEY}} = a / 2 \quad , \qquad (66)$$

$$f_{\text{MEY}} = -\,a / 2b \quad . \qquad (67)$$

We can also apply Gulland's method of equilibrium approximation to estimate the parameters of the Pella-Tomlinson model. In this case, a model of the form

$$U_i^{n-1} = a + b\,\bar{f}_i + e_i \qquad (68)$$

is obtained. But here, three parameters must be estimated and the model is nonlinear in *n*. Parameter *n*, which controls the shape of the production curve, can be estimated along with parameters *a* and *b* by using a search method. For example, in a method proposed by Fox (1971), a set of values that covers the range of possible values for *n* is first provided. For each value *n*, a simple linear regression of U_i^{n-1} against f_i is performed and the value of *n* which produces the minimum sum of squares of residuals is chosen as the new *n*. In order to increase precision, a new set of values that surround the best value of *n* obtained in the first step is defined; for each value of the new set, the same procedure is repeated and the minimum sum of squares criterion is used to obtain a "best" estimate from these. Since the model is not defined for n=1, the method, if implemented on a computer, should only allow values of *n* different than one. The final estimates for parameters *a*, *b* and *n* can be used to calculate the maximum equilibrium yield, the catch per unit effort at MEY and the fishing effort which gives MEY. That is,

$$U_{\text{MEY}} = (a/n)^{1/(n-1)} \quad , \qquad (69)$$

$$f_{\text{MEY}} = a\,(1-n) / bn \quad , \qquad (70)$$

$$Y_{\text{MEY}} = f_{\text{MEY}}\,U_{\text{MEY}} \quad . \qquad (71)$$

Alternatively, since the model to estimate is nonlinear, it is also possible to use nonlinear estimation procedures to estimate *a*, *b* and *n* from equation (68). Nonlinear regression methods, which involve iterative processes, are available in most computer installations for solving a wide variety of estimation problems. The steepest descent, the Gauss method and the Marquardt method are examples of such iterative methods which can be applied to this estimation problem.

The variance of the regression coefficients and the confidence intervals of the predicted line can be calculated from the standard equations arising from the linear theory (see, for example, Draper and Smith, 1966). From these, the statistical significance of the regression coefficients can be assessed. This step is, too often, omitted when fitting production models. Non-significant estimates suggest that data do not contain enough information to support the model. The quantities Y_{MEY} and f_{MEY} are simple functions of the regressions coefficients and, consequently, their variances can be derived from the variance-covariance matrix of the regression coefficients. We de-emphasize their use, however, as they may lead to erroneous conclusions if used independently. Instead, the confidence intervals around the regression curve should be calculated and plotted against the effort index. The confidence regions thereby obtained provide a better appreciation of the location of the equilibrium yield over the full range of effort values.

Parameter estimation under non-equilibrium conditions.

Time averaged regressors. Schnute (1977) developed a dynamic form for the Schaefer model and proposed linear and nonlinear regression methods for estimating its parameters. The dynamic equations have a relatively complex form and relate the average catch per unit effort in year *i* (U_i) to the average catch per unit effort in year i-1, the total effort in year *i*-1 (f_{i-1}) and the total effort in year *i* (f_i). Schnute's linear formulation, which assumes that conditions change slowly in the fishery, yield the following estimates for the parameters of the Schaefer model (assuming annual statistics are used):

$$\hat{r} = \frac{\bar{x}\,(S_{xy}S_{yx}-S_yS_{xz}) + \bar{y}\,(S_{xy}S_{xz}-S_xS_{yz}) + \bar{z}\,(S_xS_y-S_{xy}{}^2)}{(\,S_xS_y - S_{xy}{}^2\,)} \quad , \qquad (72)$$

$$\hat{B}_\infty = \hat{r}\,(S_x\,S_y - S_{xy}{}^2\,)^2 \,/\, (S_{xy}S_{yz}-S_yS_{xz})\,(S_{xy}S_{xz}-S_xS_{yz}) \quad , \qquad (73)$$

$$\hat{q} = (S_{xy}S_{yz}-S_yS_{xz}) \,/\, (S_xS_y-S_{xy}{}^2) \quad , \qquad (74)$$

where

$$S_x = \sum_{i=1}^{N}(\,x_i - \bar{x}\,)^2 \qquad \text{(same form for } S_y \text{ and } S_z\text{)}$$

$$S_{xy} = \sum_{i=1}^{N}(\,x_i - \bar{x}\,)(\,y_i - \bar{y}\,) \qquad \text{(same form for } S_{xz} \text{ and } S_{yz}\text{)} \quad ,$$

$$\bar{x} = (\,\sum_{i=1}^{N} x_i\,)\,/\,N \qquad \text{(same form for } \bar{y} \text{ and } \bar{z}\text{)} \quad ,$$

$$x_i = 0.5\,(f_{i-1} + f_i) \quad ,$$

$$y_i = 0.5\,(U_{i-1} + U_i) \quad ,$$

$$z_i = \ln\,(U_i \,/\, U_{i-1}) \quad .$$

Because the Schnute's algorithm attempts to model the dynamic process of approach to equilibrium, there is no need to apply the equilibrium approximation method of Gulland to the catch and effort data. Because this method uses time averages of catch rate and effort, it is said to be based on time-averaged regressors. Uhler (1979) also uses time averaged regressors for estimating the parameters of the Schaefer model.

By approximating the solution of the underlying system of differential equations. Equations (33a) and (33b) represent a system of differential equations which governs the dynamic model. In other words, the solution of this system provides a complete description of the transient states. As the solution of (33) does not have a close form, one must rely on discrete approximations to the solution: the Runge-Kutta algorithm for solving differential equations is one example of such approximation. An example of the transient solution for the Pella-Tomlinson model is illustrated in Fig. 17. The solution of (33) has the general form:

$$\hat{Y}_i = g(f_1, f_2, \ldots, f_i;\ \theta) \quad , \qquad (75)$$

where θ is the column vector formed of the parameters of the production model used. The function g is thus a linear or nonlinear function of the parameters θ. Now let

$$e_i = y_i - g(f_1, f_2, \ldots, f_i;\ \theta^*) \quad , \qquad (76)$$

where θ^* is the true value of θ. For any value $\hat{\theta}$ of θ, the differences will be referenced as the residuals. If the probability distribution function of the errors e_i is known, estimates of the unknown parameters can be sought by the maximum likelihood method. For convenience, the errors are often assumed to follow a normal distribution $\mathcal{N}(0,\mathbf{C})$, with known covariance matrix $\mathbf{C}$. In this case, a maximum likelihood criterion reduces to a weighted least squares criterion with the weights, say w_i, given by the elements of the inverse of the covariance matrix. Thence, the parameter estimation problem can be stated as follows: find the value $\hat{\theta}$ of θ that minimizes the weighted least squares criterion

$$S(\hat{\theta}) = \sum_i w_i \left(y_i - g(f_1, f_2, \ldots, f_i;\ \hat{\theta}) \right)^2 \quad , \qquad (77)$$

where the w_i are the statistical weights. In most cases, however, the covariance matrix is unknown. In order to obtain least variance estimates, the use of weights which are as close as possible to the elements of $\mathbf{C}^{-1}$ is desirable. Following Fox (1971a), we will consider an error term e_i proportional to the estimated population size; in terms of yield, this is equivalent to the form

$$y_i = \hat{y}_i - \hat{y}_i\, e_i \quad . \qquad (78)$$

Then the residuals are described by the relationship

$$e_i = (y_i - \hat{y}_i) \ / \ \hat{y}_i \quad . \qquad (79)$$

In other words, $w_i = \hat{y}_i^{-2}$. This particular casting of the error term enables us to estimate simultaneously the weights w_i and θ. We can find $\hat{\theta}$ by writing the normal equations

$$[\ \partial\, S(\hat{\theta}) \ / \ \partial\, \theta_k\ ] = [\ 0 \ldots 0\ ] \quad , \qquad (80)$$

where the subscript k identifies the k^{th} parameter. When g is linear, θ can be calculated via linear regression methods. In our dynamic production system (see equation 33), however, g appears as a nonlinear function of θ and the normal equations must be solved numerically. For example, Rivard and Bledsoe (1978) examined the use of a derivative-free version of the Levenberg-Marquardt numerical optimization algorithm, together with a fourth order Runge-Kutta differential equation solver, to estimate the parameters of the Pella-Tomlinson model. The general structure of the algorithm used is shown by the flow diagram of Fig. 18. They applied this estimation process to a five-parameter version of the Pella-Tomlinson model, in which MEY, B_∞, n, q and B_0 (which arises from the solution of equation (33a)) were estimated, and to a three-parameter version in which only MEY, n and q required estimation.

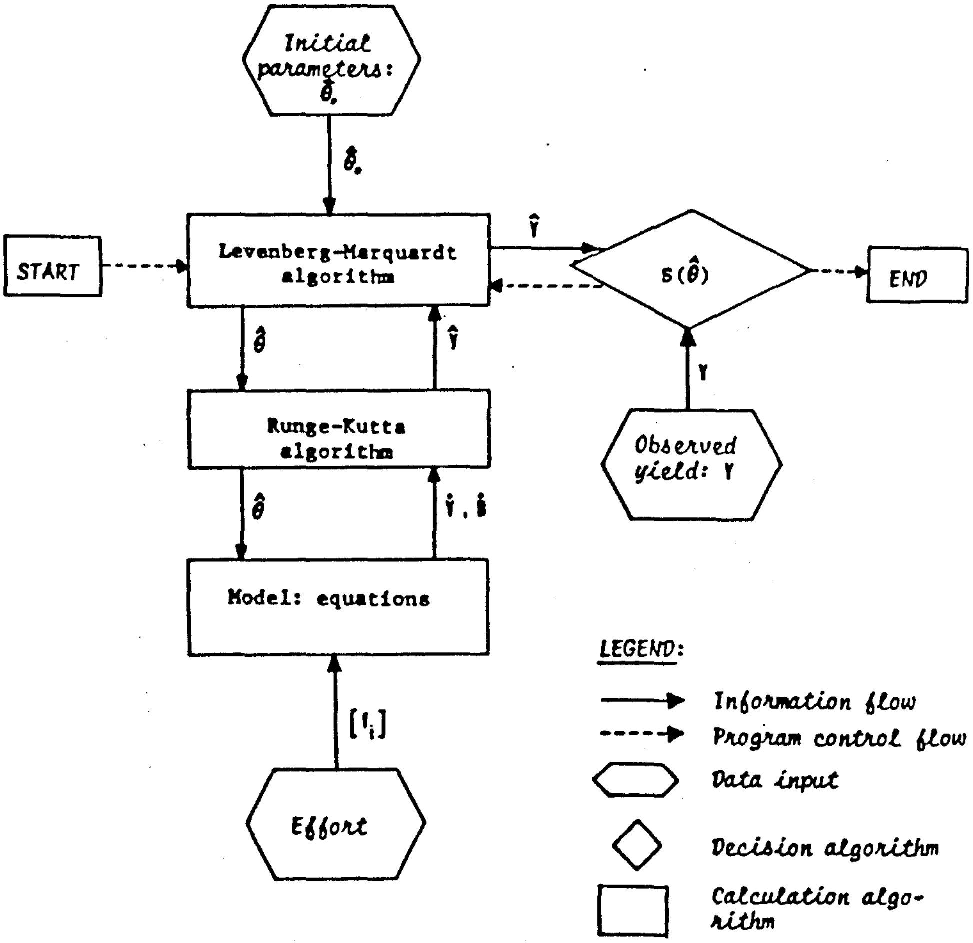


FIGURE 18. Information flow diagram for the computer program written to estimate the coefficients of the generalized production model.

Projections.

Production models have been described as having two basic forms: the equilibrium form and the transient form. We can project yield or catches from either form. However, the projections established from the equilibrium form should be seen as the long term average, i.e. the yield level which the stock will achieve at equilibrium given a reference effort level. Due to environmental changes affecting recruitment and growth and to changes in effort level, the equilibrium level may never be achieved or may be achieved only temporarily. In addition, this level may not be sustainable, as illustrated in Fig. 10. The projected transient yield will be somewhat different from the equilibrium projection and will describe the yield levels during the transition periods. However, the underlying assumption is that transient levels are affected only by changes in fishing effort. In practice, transient levels may depart significantly from transient projections if, for example, recruitment in these years is much weaker or much stronger than average recruitment levels.

A reference point for the projections: 2/3 f_{MEY}.

A reference point is needed for projecting catches from the transient form of production models. For a number of years, the maximum equilibrium yield was seen as the obvious target level which management should pursue on a yearly basis. The application of a constant quota leads, however, to loss of opportunities in years of good recruitment and accentuates stock decline in years of bad recruitment. When Total Allowable Catches (TAC) were established on many stocks of the Northwest Atlantic in the early seventies, F_{MAX} and F_{MEY} were identified as target levels or reference points for the projections. However, due to the critical conditions of a number of groundfish stocks, more conservative reference points were soon adopted: $F_{0.1}$ was suggested for projections from age-structured models and 2/3 F_{MEY}, for production models. While projections at $F_{0.1}$ were using age-structured models and were thus clearly transient projections, the projections at 2/3 F_{MEY} were made from production models assuming equilibrium conditions. It is only recently that the transient form of production models have been used for projecting transient yield levels.

Examples

The purpose of this section is 1) to gain insight into the adequacy of various production models for describing stock dynamics, 2) to evaluate the performance of different estimation procedures and 3) to apply various models to stocks for which the data base allows a cross-checking of the estimates obtained. It would be unfair to use data generated from a given production model to evaluate the performance of different estimation procedures: such an approach could benefit the model which was chosen for generating the data and, thus, could bias the results of the analysis. A different approach, which is the approach taken here, is to build a model which mimics, as closely as possible, the dynamics of known stocks and to use the model for generating catch and effort data. This model would provide the necessary data for estimating the parameters of various production models.

Three age-structured populations were thus simulated over 24 years by using the projection algorithm of Rivard (1982). Growth patterns and recruitment were chosen to reflect the characteristics of haddock, cod and redfish. These, as we can see from Fig. 1, cover a wide range of turnover rates and include very dynamic as well as less dynamic stocks. We applied to each population a pattern of fishing mortality rates simulating a developing fishery for the first 13 years, followed by a progressive decline of fishing mortality to the vicinity of the $F_{0.1}$ level. Then, sensitivity analyses and Monte Carlo simulations were used to gain insight into both the ability of a given model to describe the stock dynamics and the performance of various estimation procedures.

Sensitivity analysis.

For this analysis, the time series of simulated data for each stock was split as follows:

-case 1: complete time series (years 1-24);
-case 2: the first 12 years (years 1-12);
-case 3: the last 12 years (years 13-24).

Within each period, the effect of the following sources of error on the estimation of Y_{MEY} and f_{MEY} was assessed:

-A: no error;
-B: 20% underestimation of effort in the last year of the series considered;
-C: 20% underestimation of catch in the last year of the series considered.

This arrangement will provide insight both into the influence of data series obtained from different phases of a fishery and the influence of selected data points on the parameter estimates. For each of these cases, the parameters of the Schaefer and of the Pella-Tomlinson models were estimated by using various methods. Table 1 summarizes the results of this sensitivity analysis. While it is recognized that this analysis is not exhaustive, it does provide insight into the importance of various errors on parameters estimates. Since the simulated data are "error-free", one would expect to find close agreement with the "parametric" values. Note that because an age-structured model was used here for generating the data, the "parametric" values are themselves unknown and have been approximated from the average recruitment over the time period considered.

Effect of stock dynamics. None of the estimation methods worked satisfactorily for the haddock data set. The simulated stock had few age groups contributing to the fishery and recruitment was variable. Under these conditions, the dynamics of any stock is such that catch rates are not indicative of changes in fishing pressure but are more likely to respond to changes in recruitment level. For the cod and redfish data series, variations in recruitment are dampen by the presence of a large number of age groups in the stock (see Fig. 1A). Consequently, changes in stock size are relatively more dependent upon changes in fishing effort rather than upon changes in recruitment level. As the estimation of production models is based on the relationship between catch rate (i.e. an index of stock size) and fishing effort, production models are more likely to be applicable to the cod and redfish types than to the haddock type.

Number of data points and time period considered. Both the number of observations and the time period considered have a significant impact on the estimates of MEY and f_{MEY}. Reducing the number of data points increases the sensitivity of parameter estimates to errors in input data. Using only part of the time series of observations, namely data coming from either the "developing phase" or the "stabilizing phase" of a fishery, leads to substantial biases in the estimates of MEY and f_{MEY}: the redfish type was particularly sensitive to the reduction in the number of observations and to the time period considered. For our simulated redfish stock, more than 20 age groups had a significant contribution to the catch in any given year. For such a stock, the length of the "transition phase" is expected to be rather long: in other words, extreme recruitment levels are expected to influence stock size for a relatively long period. Consequently, when using a production model for describing the dynamics of an age-structured stock, residuals show series of positive or negative values (Fig.19). For the redfish type, our analysis revealed series of 9-10 residuals having the same sign; for the cod type, the series included 4-6 consecutive residuals of the same sign. Because of the presence of systematic patterns in the residuals, parameter estimates obtained by fitting production models by standard regression techniques are likely to be biased when a low number of observations are available for the fit. This suggests that the minimum acceptable number of observations which must be included in the analysis is a function of the number of age-groups which are present in the stock in any given year.

Estimation methods. The estimation methods which are based upon equilibrium approximations or which use a transient form of production models outperformed the regression of catch rate against effort, which assumes that the observations are from an equilibrium state. Gulland's method of equilibrium approximations worked best, overall, for estimating the parameters of the Schaefer form and of the Pella-Tomlinson form (Fig. 20). The estimates of MEY obtained from the regression of catch rate against effort were sensitive to errors in input data, to the number of observations and to the time period chosen for the fit. For the redfish type (all data points), the Schnute's method and the nonlinear estimation lead to significant biases in the estimates of MEY when the Schaefer model was applied; these estimates were also very sensitive to errors in input data. For cod (all data points), the Gulland, Schnute and nonlinear methods underestimated f_{MEY} systematically.

Even though the number of cases analysed is limited, the results of this sensitivity analysis are important as they point at potential sources of error under known conditions. The conditions simulated are close to conditions which someone would expect to encounter in the analysis of data from a real fishery. In certain cases (e.g. the haddock example), the shape of the equilibrium Schaefer or Pella-Tomlinson curve will be inappropriate for describing the expected concavity of the underlying equilibrium curve over the range of available fishing effort values. In fact, because of the weak concavity underlying the models which were simulated, f_{MEY} is expected to be more difficult to estimate than MEY: the sensitivity analysis confirms this point. Finally, because production curves correspond to simple models which do not take into consideration the age composition of the stock, all models estimated suffered from lack of fit, as revealed by the presence of systematic patterns in the residuals.

Comparing results from different estimation procedures: Monte Carlo simulations.

When undertaking the assessment of a given fishery, it is common to find divergent estimates for the parameters of production models . As pointed out by Mohn (1980), "this may be due to discrepancies within the data, i.e. different catch and effort statistics for various sectors or gear types within a fishery" or due to the use of different estimation methods. Mohn (1980) also notes that the results depend upon the dynamics of the stock being studied. Similarly, Stocker and Hilborn (1981) observed that the relative performance of the methods depends upon the exploitation history of the stock. Rivard and Bledsoe (1978) suggest that nonlinear estimation methods, when applied to the dynamic form of the Pella-Tomlinson model, will tend to be inefficient in the absence of external constraints. While the model exhibits a convenient flexibility with a minimum number of parameters, the system often provides more flexibility than we care to have.

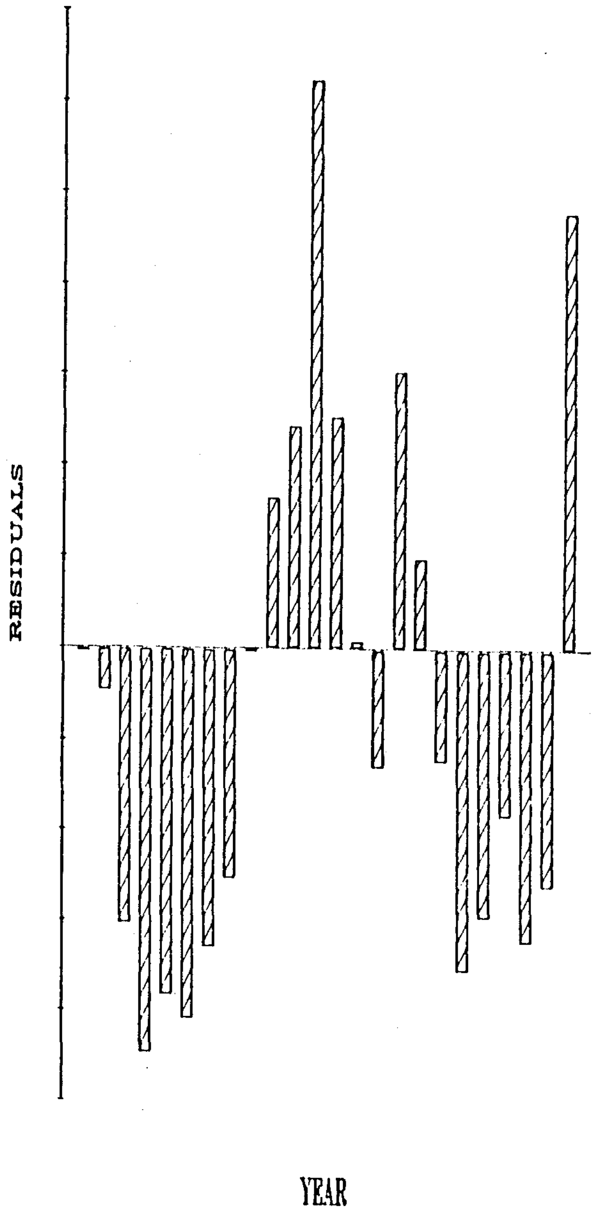


FIGURE 19. Plot showing the residuals obtained when estimating the parameters of the Graham–Schaefer model with a non–linear estimation procedure. The presence of residuals having the same sign is the result of the age–structure of the simulated stock: the transient model cannot mimic the complex behavior of an age–structured fish population.

TABLE 1. Sensitivity analysis for three stocks having the characteristics of haddock, cod and redfish stocks.

				haddock type		cod type		redfish type	
Period (years)	Error type	Model	Estimation method	Yield at MEY	Effort at MEY	Yield at MEY	Effort at MEY	Yield at MEY	Effort at MEY
1-24	TRUE VALUES	age-structured		24	944	102	40	42	99
1-24	No error	Schaefer	U vs E	-11	-62	129	57	43	29
			Gulland	41	154	101	33	45	36
			Schnute	-2	8	101	29	78	86
			Nonlinear	N.A.	N.A.	101	29	23	49
		Pella-Tomlinson	U vs E	-2	-10	135	51	53	49
			Gulland	293	912	100	33	43	31
	Yield in last year under-estimated by 20%	Schaefer	U vs E	-11	-61	129	57	42	29
			Gulland	42	161	100	33	45	36
			Schnute	-.6	5	98	29	-31	-59
			Nonlinear	N.A.	N.A.	100	29	10*	35*
		Pella-Tomlinson	U vs E	-2	-9	134	51	51	47
			Gulland	344	1074	98	34	43	31
	Effort in last year under-estimated by 20%	Schaefer	U vs E	-11	-62	128	55	43	29
			Gulland	40	151	102	33	45	35
			Schnute	-5	11	104	28	48	42
			Nonlinear	N.A.	N.A.	102	29	28	43
		Pella-Tomlinson	U vs E	-2	-9	134	49	53	49
			Gulland	305	949	101	33	43	29
1-12	No error	Schaefer	U vs E	-4	-23	216	96	52	32
			Gulland	-5	-26	97	29	1191	1001
			Schnute	646	69	98	22	-9	-13
			Nonlinear	N.A.	N.A.	99	21	N.A.	N.A.
		Pella-Tomlinson	U vs E	-.5	-3	208	77	124	420
			Gulland	-8	-39	99	29	846	589
	Yield in last year under-estimated by 20%	Schaefer	U vs E	-6	-31	177	76	45	27
			Gulland	-10	-43	93	26	58	43
			Schnute	60	30	185	13	42	34
			Nonlinear	N.A.	N.A.	154	14	N.A.	N.A.
		Pella-Tomlinson	U vs E	-.9	-5	183	70	106	1370
			Gulland	-6	-25	98	28	51	33
	Effort in last year under-estimated by 20%	Schaefer	U vs E	-3	-16	335	156	54	33
			Gulland	-3	-16	102	32	-31	-30
			Schnute	1	-16	106	30	45	9
			Nonlinear	N.A.	N.A.	102	28	N.A.	N.A.
		Pella-Tomlinson	U vs E	-1	-7	269	102	101	147
			Gulland	-7	-39	102	32	-66	-90
13-24	No error	Schaefer	U vs E	-10	-69	98	43	-4	-10
			Gulland	51	279	111	43	-53	-77
			Schnute	10	32	98	33	-.1	-1
			Nonlinear	N.A.	N.A.	104	36	N.A.	N.A.
		Pella-Tomlinson	U vs E	-6	-40	119	660	<<<<	<<<<
			Gulland	31	134	180	1147	-55	-64
	Yield in last year under-estimated by 20%	Schaefer	U vs E	-8	-59	98	44	-3	-7
			Gulland	282	1664	135	61	-6	-14
			Schnute	-.5	4	94	35	-17	-34
			Nonlinear	N.A.	N.A.	99	4	N.A.	N.A.
		Pella-Tomlinson	U vs E	-3	-25	123	759	<<<<	<<<<
			Gulland	40	181	121	47	-63	-2600
	Effort in last year under-estimated by 20%	Schaefer	U vs E	-12	-86	97	40	-13	-21
			Gulland	31	159	103	33	38	37
			Schnute	13	48	120	29	-1	-4
			Nonlinear	N.A.	N.A.	111	35	N.A.	N.A.
		Pella-Tomlinson	U vs E	-7	-49	114	587	-5	-7
			Gulland	25	105	134	530	33	28

Footnote: <<<< : beyond reasonable range
N.A. : not attempted
* : none of the convergence criteria were met after 10 iterations
(The nonlinear estimation was constrained within the domain of positive parameter values)

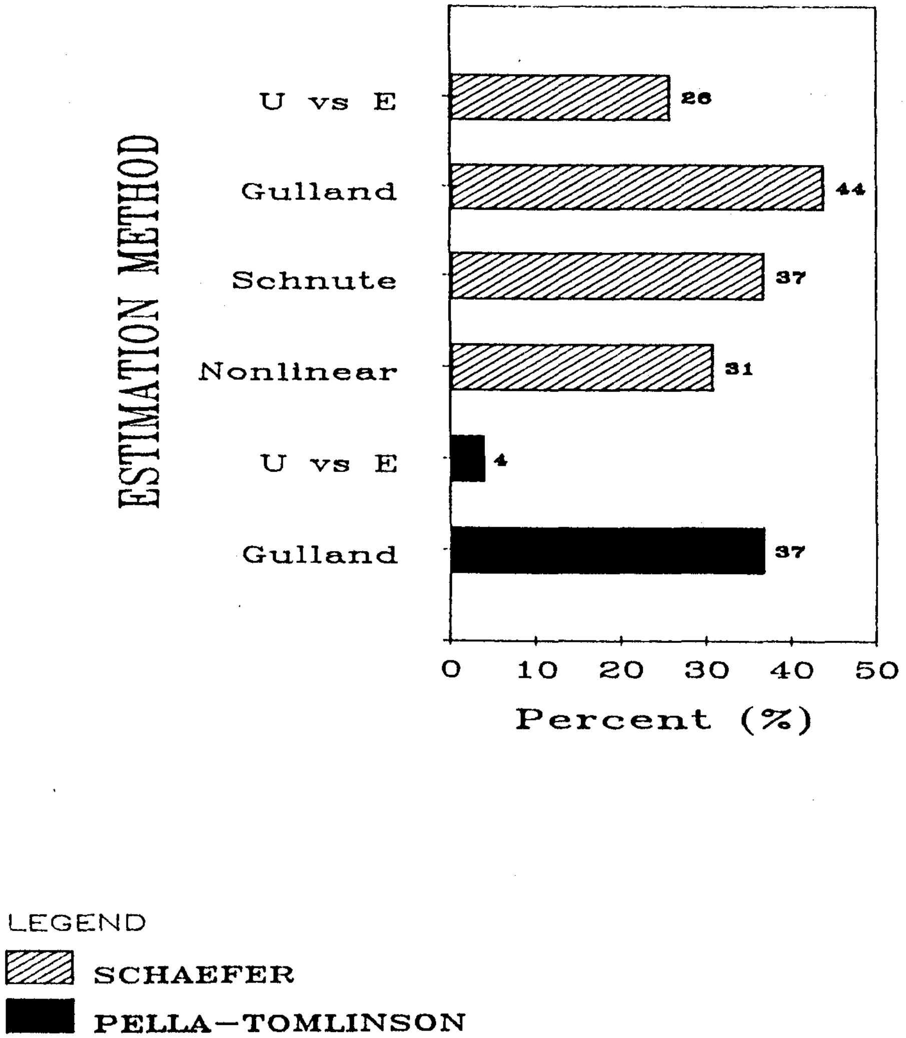


FIGURE 20. Relative performance of various procedures for estimating MEY, as established from the results of the sensitivity analysis appearing in Table 1. The percentages reported identify the cases for which estimates of MEY were within 15% of the parametric value.

Rivard and Bledsoe (1978), Uhler (1979) and Mohn (1980) present the results of some Monte Carlo simulations made to evaluate the performance of various estimation procedures. While these authors provide insight into potential sources of concern for the estimation of production models, none provide insight on the biases or uncertainties expected in an age-structured system. The results of a Monte-Carlo simulation in which an age-structured stock is fished at various F-levels appears in Table 2: the means and standard deviations of MEY and f_{MEY} corresponding to 100 trials are reported, along with the range for the correlation coefficient, and the number of times a negative estimate of MEY or f_{MEY} resulted within the 100 trials. Catch and effort data obtained from the simulation of the redfish and cod fisheries were used in this analysis. Error (noise) was added independently to effort and to catch in the following manner:

$$f_t^1 = f_t + k\, f_t\, e_t$$

$$Y_t^1 = Y_t + k\, Y_t\, e_t'$$

where e_t and e_t' are drawn from a normal distribution $\mathcal{N}(0, 1)$. The coefficient k, which serves to introduce noise levels, was assigned the values of 5%, 10% and 20%.

For the cod data, all methods tended to have low bias for estimates of MEY but showed considerable bias for those of f_{MEY}. This is so because the age-structured population which was used in the simulation implies a low concavity for the surplus production curve over the range of fishing effort used. For the Schnute's method, the correlation coefficient varied as a function of the noise level in input data and high correlation translated into good fits. Schnute's method lead to unsatisfactory estimates at higher noise levels, as indicated by the large standard deviation of the estimates. For the Gulland method, the estimates of MEY obtained at all noise levels were relatively close to the true value and correlation coefficients were relatively high (higher than 0.5 in all cases; greater than 0.8 in most cases). Because the Gulland method averages data over a number of years, this method produces estimates which are more stable and the coefficients of correlation are less affected by the degree of noise present in the data; averaging tends to smooth out errors in "annual" data and, thus, leads to more "stable" estimates.

For the redfish data, the estimates of MEY and f_{MEY} obtained by the Schnute's method were excessively "unstable" and their distribution lead to very large standard deviations; negative estimates were frequent. Estimates obtained by the Gulland method were stable over a wide range of noise levels and showed relatively small standard deviations (Fig. 21). These results are consistent with the results of Mohn (1980). For high noise levels, the frequency distribution of MEY estimates showed relatively long tails (Fig. 21), an indication that unrealistic estimates were quite frequent. While such estimates were generally associated with low coefficients of correlation (smaller than 0.2) and, thus, would not be retained as being statistically acceptable, there were a few cases where "unrealistic" estimates of F_{MEY} generated relatively high correlations.

This Monte-Carlo simulation points at a major difficulty in the estimation of parameters for production models from catch and effort data: namely, biases in the estimated effort at MEY. When data come from a stock which has experienced relatively constant exploitation rates, catch and effort may contain very little information on the effort level which would give the maximum equilibrium yield. When a management strategy dictates relatively constant effort levels, such as the application of $F_{0.1}$, catch and effort contain information on a relatively small range of the yield/effort curve and it becomes impossible to infer the shape of that curve from the limited data series.

Table 2A. Results of various estimation procedures for estimating the parameters of the Schaefer model. For each noise level, 100 replicates of the effort and catch rates series obtained by simulating an age-structured population were constructed and used for estimating the parameters of the models.

SCHAEFER MODEL

			Gulland						Schnute					
	Noise added to	Noise level	MEY		f(MEY)		R	NNEG	MEY		f(MEY)		R	NNEG
			MEAN	SD	MEAN	SD	(range)		MEAN	SD	MEAN	SD	(range)	
	Parametric	N.A.	102	-	40	-	1.00	0	102	-	40	-	1.00	0
COD	Effort	0%	101	-	33	-	.97	0	101	-	29	-	.96	0
		5%	103	1.5	33	.9	(.94-.96)	0	101	1	29	.8	(.92-.97)	0
		10%	103	2.5	33	1.6	(.92-.98)	0	101	2.3	29	1.6	(.89-.99)	0
		20%	103	5.3	33	3.2	(.86-.98)	0	101	4.1	29	3.2	(.63-.98)	0
	Catch rate	0%	101	-	33	-	.97	0	101	-	29	-	.96	0
		5%	101	1.1	33	.7	(.93-.98)	0	101	1.3	29	.8	(.73-.94)	0
		10%	101	2.3	33	1.7	(.80-.95)	0	101	3.1	29	1.4	(.44-.80)	0
		20%	102	4.8	34	2.9	(.52-.93)	0	110	35.5	25	3.1	(.11-.59)	0
	Catch rate and effort	0%	101	-	33	-	.97	0	101	-	29	-	.96	0
		5%	103	1.6	33	1.1	(.90-.97)	0	101	1.7	29	1.1	(.73-.90)	0
		10%	102	4.0	33	2.5	(.80-.97)	0	102	3.8	29	2.5	(.42-.78)	0
		20%	105	7.1	33	4.0	(.58-.91)	0	107	37.3	26	4.9	(.07-.68)	1
	Parametric	N.A.	42	-	99	-	1.00	0	42	-	99	-	1.00	0
REDFISH	Effort	0%	45	-	36	-	.48	0	78	-	86	-	0.36	0
		5%	46	1.7	36	1.8	(.43-.54)	0	100	114	113	140	(.36-.37)	2
		10%	46	4.1	37	4.5	(.36-.59)	0	80	327	89	403	(.36-.39)	17
		20%	48	7.8	38	8.4	(.27-.75)	0	35	216	32	266	(.36-.40)	25
	Catch rate	0%	45	-	36	-	.48	0	78	-	86	-	0.36	0
		5%	46	3.1	36	3.5	(.36-.62)	0	76	432	81	525	(.19-.43)	16
		10%	45	5.1	35	6.0	(.29-.66)	0	41	294	42	366	(.08-.47)	32
		20%	53	24.3	43	25.7	(.07-.68)	1	-54	423	-50	274	(.01-.38)	52
	Catch rate and effort	0%	45	-	36	-	.48	0	78	-	86	-	0.36	0
		5%	47	5.1	37	5.7	(.27-.58)	0	27	220	23	269	(.17-.44)	26
		10%	48	11.2	39	12.7	(.12-.67)	0	-40	305	-59	376	(.06-.46)	37
		20%	3	571.8	-12	634.5	(.00-.78)	6	-53	638	-83	636	(.03-.30)	39

NOTE: NNEG refers to the number of negative estimates obtained in 100 simulations.

Table 2B. Results of various estimation procedures for estimating the parameters of the Pella-Tomlinson model. For each noise level, 100 replicates of the effort and catch rates series obtained by simulating an age-structured population were constructed and used for estimating the parameters of the models.

PELLA-TOMLINSON MODEL

Gulland

	Noise added to	Noise level	MEY MEAN	MEY SD	f(MEY) MEAN	f(MEY) SD	R (range)	NNEG
	Parametric	N.A.	102	-	40	-	1.00	0
	Effort	0%	100	-	33	-	.97	0
		5%	100	1.4	33	1.0	(.95-.98)	0
C		10%	100	2.8	34	2.3	(.92-.98)	0
		20%	100	6.4	37	24.3	(.84-.98)	0
O								
	Catch	0%	100	-	33	-	.97	0
D	rate	5%	100	2.3	33	1.0	(.91-.97)	0
		10%	100	3.8	34	8.0	(.80-.96)	0
		20%	105	6.9	38	28.3	(.48-.90)	0
	Catch	0%	100	-	33	-	.97	0
	rate and	5%	100	2.5	33	1.5	(.92-.98)	0
	effort	10%	102	5.1	35	9.0	(.78-.97)	0
		20%	104	9.8	34	4.3	(.41-.90)	0
	Parametric	N.A.	42	-	99	-	1.00	0
R	Effort	0%	43	-	31	-	.51	0
		5%	43	1.3	31	1.3	(.46-.58)	0
E		10%	43	3.0	32	2.8	(.39-.63)	0
		20%	44	5.1	32	4.8	(.33-.77)	0
D								
	Catch	0%	43	-	31	-	.51	0
F	rate	5%	43	1.9	31	2.2	(.36-.58)	0
		10%	44	5.8	31	5.0	(.17-.69)	0
I		20%	44	34.1	29	29.0	(.05-.63)	4
S	Catch	0%	43	-	31	-	.51	0
	rate and	5%	43	2.3	31	2.6	(.38-.60)	0
H	effort	10%	47	15.1	41	77.0	(.07-.74)	0
		20%	41	204.0	27	172.5	(.00-.77)	5

NOTE: NNEG refers to the number of negative estimates obtained in 100 simulations.

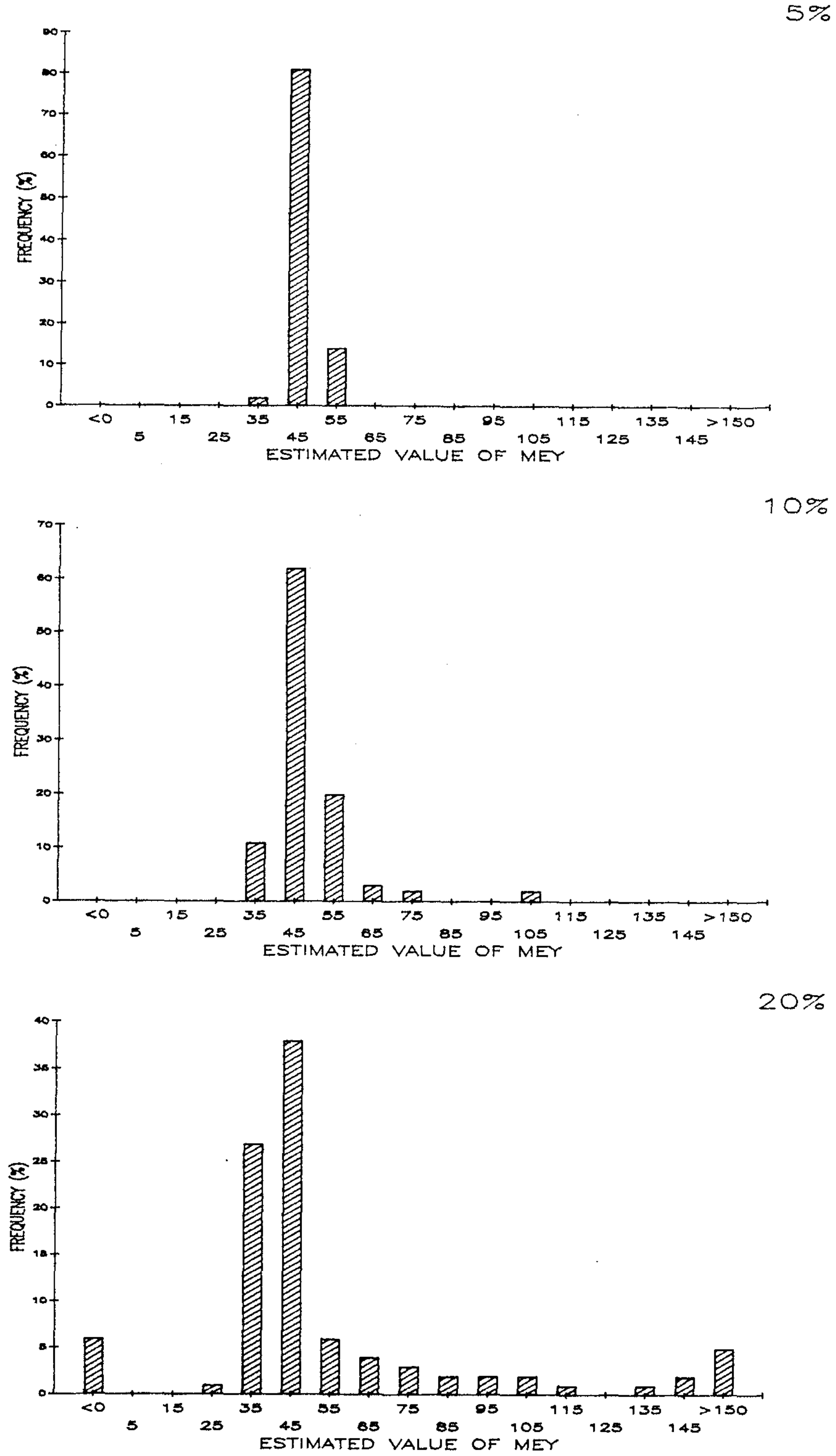


FIGURE 21. Frequency distribution of MEY estimates when normally distributed error is added to catch rate and effort data. Frequency distribution is for estimates of MEY from the Schaefer model by the Gulland method, as applied to the redfish type.

Examples from the Northwest Atlantic.

In practice, the application of production models to real data leads to many options regarding the form of the production model, the catch effort (or catch rate) data series, the estimation method and the criteria for goodness of fit. When several data series are available, the biologist must select either the relevant data series or derive a standard catch rate series. In this Section, various production models are applied to selected stocks for which the data base allows a cross-checking of the estimates obtained. We must remember, however, that the models estimated with catch and effort data coming from a given time period are describing at best the average productivity over the time period considered. The average productivity is itself a function of the average conditions (environmental or climatic conditions, recruitment, gear selectivity, etc.) which prevailed during the time period considered.

Cod (4T-Vn).

Five different catch and effort series from the 4T-Vn cod stock (Table 3) have been used to estimate the parameters of various production models by different methods: Table 4 summarizes the results. The first catch rate series (32 observations) is derived from the total catch and fishing mortality estimated by cohort analysis (Maguire and Waiwood, 1982). For that series, estimates of MEY are within 11% of the average production calculated from yield per recruit and the 1950-81 average recruitment. The nonlinear estimation did not converge for the Pella-Tomlinson system when applied to series A. Series B has 23 observations representing catch per trip from the 1953-1975 period, as defined by Lett (1978). Even though the Gulland method applied to series B for the Pella-Tomlinson model led to a correlation coefficient of 0.8, someone would reject the resulting model because the estimated (MEY, f_{MEY}) point is completely outside the range of observed catches and effort. Similarly, the estimates of the Schaefer model by the nonlinear procedure must also be rejected. Finally, the Schnute method did not provide a significant regression with series B. For the other three catch and effort series (series C, D, and E), there were less than 14 observations and only the U vs E regressions and, in one case, the Gulland method provided estimates of MEY which were within the range of observed catch levels. The estimates of MEY and f_{MEY} obtained from the *U* vs *E* regression are expected to be biased, however, because the method does not correct the transient points to approximate equilibrium conditions. Due to the low number of observations, the confidence intervals around the equilibrium curve would suggest a low confidence in the models obtained with data series C, D and E. In summary, data series A and B provided estimates of MEY varying from 61,400 to 71,200 (average 67,500 t): this mean value is very close to the average maximum productivity calculated from yield per recruit and the 1950-1981 average recruitment, i.e. 64,000 t. Nonlinear estimation was not attempt on the shorter series.

We estimated also the parameters of the Schaefer and the Pella-Tomlinson models by applying various methods to the first and last 14 years of Series A, respectively (Table 4B). There is a tendency for the estimates of the first 14 years to be somewhat higher than the estimates obtained with the last 14 years . In addition, the estimates obtained for the last 14 years are comparable to those obtained from series C, D, and E. This suggests that the average productivity of the seventies was lower than the average productivity of the 1950-1963 period. This is also apparent in cohort analysis.

Table 3. Catch and effort, or catch and catch rate series used for fitting production models to cod in NAFO Division 4T-Vn.

Year	Catch (t)	A	B	C	D	E
1950	44,023	0.179	-	-	-	-
1951	34,827	0.118	-	-	-	-
1952	41,956	0.147	-	-	-	-
1953	58,911	0.185	8.96	-	-	-
1954	63,901	0.257	9.29	-	-	-
1955	65,227	0.200	8.59	-	-	-
1956	104,469	0.259	10.44	-	-	-
1957	89,131	0.272	10.00	-	-	-
1958	86,582	0.498	8.01	-	-	-
1959	70,720	0.765	9.18	-	-	-
1960	66,013	0.394	5.14	-	-	-
1961	65,583	0.314	7.17	-	-	-
1962	66,664	0.175	7.77	-	-	-
1963	70,202	0.386	7.13	-	-	-
1964	60,547	0.446	6.28	-	-	-
1965	63,027	0.591	5.48	-	-	-
1966	54,851	0.405	3.65	-	-	-
1967	41,314	0.340	4.32	-	-	0.29
1968	46,551	0.349	6.51	-	-	0.25
1969	47,819	0.299	7.38	-	-	0.34
1970	64,459	0.433	7.26	6.7	20.4	0.28
1971	56,375	0.385	4.99	6.5	13.4	0.26
1972	65,291	0.555	4.98	7.4	10.1	0.28
1973	49,608	0.567	3.77	5.7	11.6	0.26
1974	47,255	0.778	2.52	4.8	6.7	0.22
1975	41,231	0.600	2.84	6.0	7.9	0.24
1976	32,729	0.457	-	2.5	7.8	0.21
1977	22,219	0.208	-	3.5	11.7	0.24
1978	37,880	0.255	-	5.6	12.7	0.30
1979	51,202	0.247	-	7.6	15.2	0.48
1980	57,270	0.246	-	8.1	22.6	0.52
1981	59,325	0.272	-	12.3	35.6	-

A: fishing mortality from cohort analysis (Maguire and Waiwood, 1982).
B: catch/trip from Lett, 1978.
C: TC2 catch/trip from Maguire and Waiwood (1982).
D: TC3 catch/trip from Maguire and Waiwood (1982).
E: TC3 t/hr from Sinclair and Maguire (1981), OT+DS combined.

Table 4A. Estimation of MEY and f(MEY) for various catch rate series, production models and estimation methods. The MEY, as calculated from yield per recruit and the 1950-1981 average recruitment, is 64,000 t. A six-year running average was used for approximating equilibrium in the Gulland method. For the nonlinear estimation, Boo, MEY and q were estimated in the Schaefer form (no constraints were applied); in addition, n was estimated in the Pella-Tomlinson form.

		A			B			C		
Model	Estimation method	MEY	f(MEY)	R	MEY	f(MEY)	R	MEY	f(MEY)	R
Schaefer	U vs E	69,705	0.414	0.74	68,342	10,793	0.79	51,943	8,444	0.64
	Gulland	71,221	0.318	0.73	69,730	7,266	0.76	(115,759)	(4,906)	0.96
	Schnute	70,300	0.312	0.42	68,566	7,130	0.36	(96,640)	5,009	0.69
	Nonlinear	65,167	0.322	0.52	84,792	(6,002)	0.74	--	--	--
Pella-Tomlinson	U vs E	61,393	0.414	0.81	66,137	10,212	0.82	52,935	8,446	0.63
	Gulland	70,698	0.315	0.74	(110,099)	(3,824)	0.80	(125,159)	(4,645)	0.96
	Nonlinear	64,504**	0.130**	0.65	**	**	**	--	--	--

		D			E		
Model	Estimation method	MEY	f(MEY)	R	MEY	f(MEY)	R
Schaefer	U vs E	57,566	4,155	0.69	53,808	(232,791)	0.51
	Gulland	(87,663)	2,896	0.96	62,225	132,563	0.90
	Schnute	(6,612,010)	2,230	0.54	(-62,910)	(79,078)	0.57
	Nonlinear	--	--	--	--	--	--
Pella-Tomlinson	U vs E	52,002	4,534	0.75	53,630	228,250	0.51
	Gulland	(136,897)	(1,191)	0.98	(-oo)	(-oo)	N.A.
	Nonlinear	--	--	--	--	--	--

A: fishing mortality estimates from cohort analysis (Maguire and Waiwood, 1982): 32 observations.
B: catch/trip from Lett, 1978: 23 obs.
C: TC2 catch/trip from Maguire and Waiwood (1982): 12 obs.
D: TC3 catch/trip from Maguire and Waiwood (1982): 12 obs.
E: TC3 t/hr from Sinclair and Maguire (1981), OT+DS combined: 14 obs.

N.B.: ** convergence not achieved
( ) identify estimates which are outside the range of observed values for catch and effort.

Table 4B. Estimation of MEY for the first and last 14 years of the catch and effort series A: cod in Div. 4T-Vn.

		First 14 years			Last 14 years		
Model	Estimation method	MEY	f(MEY)	R	MEY	f(MEY)	R
Schaefer	U vs E	89,903	473	0.78	56,489	481	0.75
	Gulland	82,127	370	0.56	(69,296)	336	0.85
	Schnute	77,977	357	0.29	60,504	351	0.80
Pella-Tomlinson	U vs E	85,427	475	0.83	53,821	483	0.77
	Gulland	93,092	377	0.54	(70,576)	365	0.86

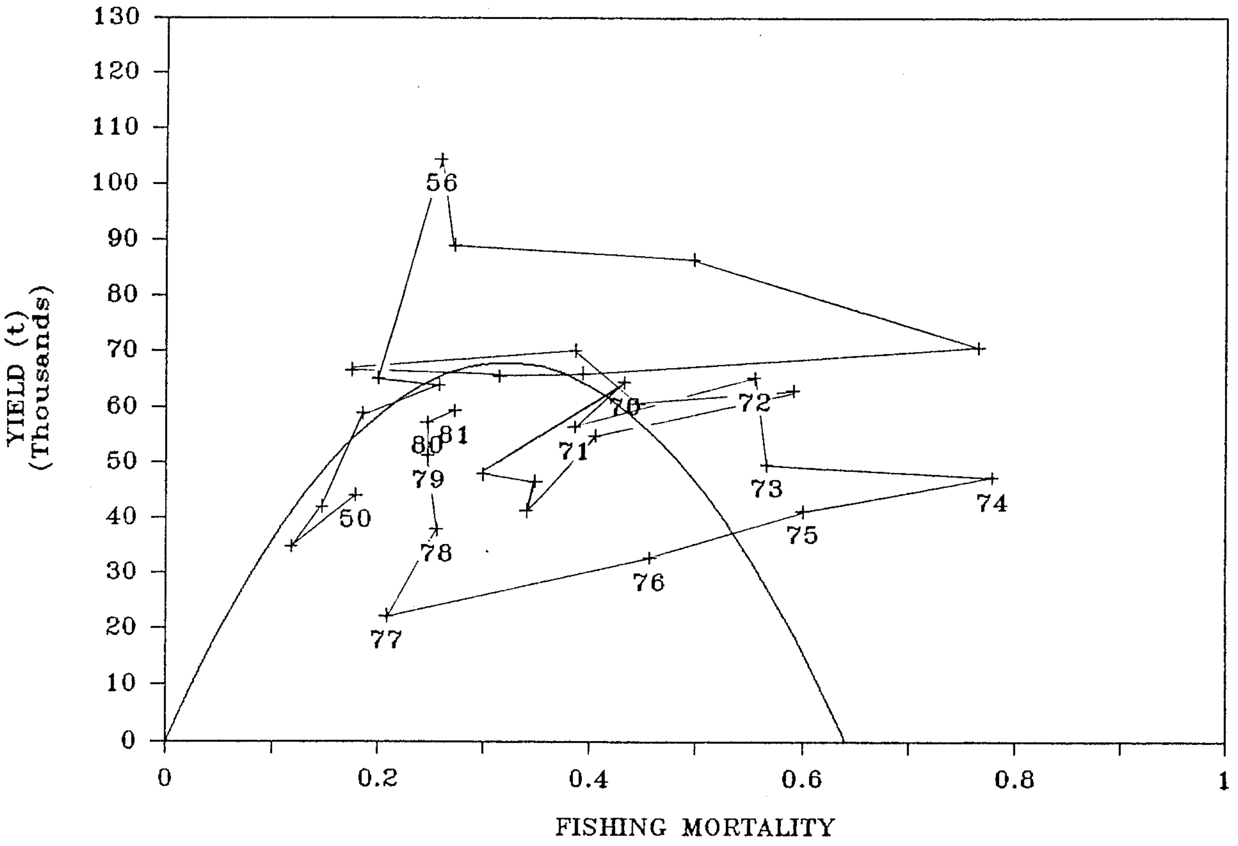


FIGURE 22. Cod 4T–Vn. Equilibrium curve resulting from fitting the Schaefer model with data series A: the estimates obtained from the Gulland, the Schnute and the nonlinear methods were averaged. The transient observed points or the observed transient path are also represented.

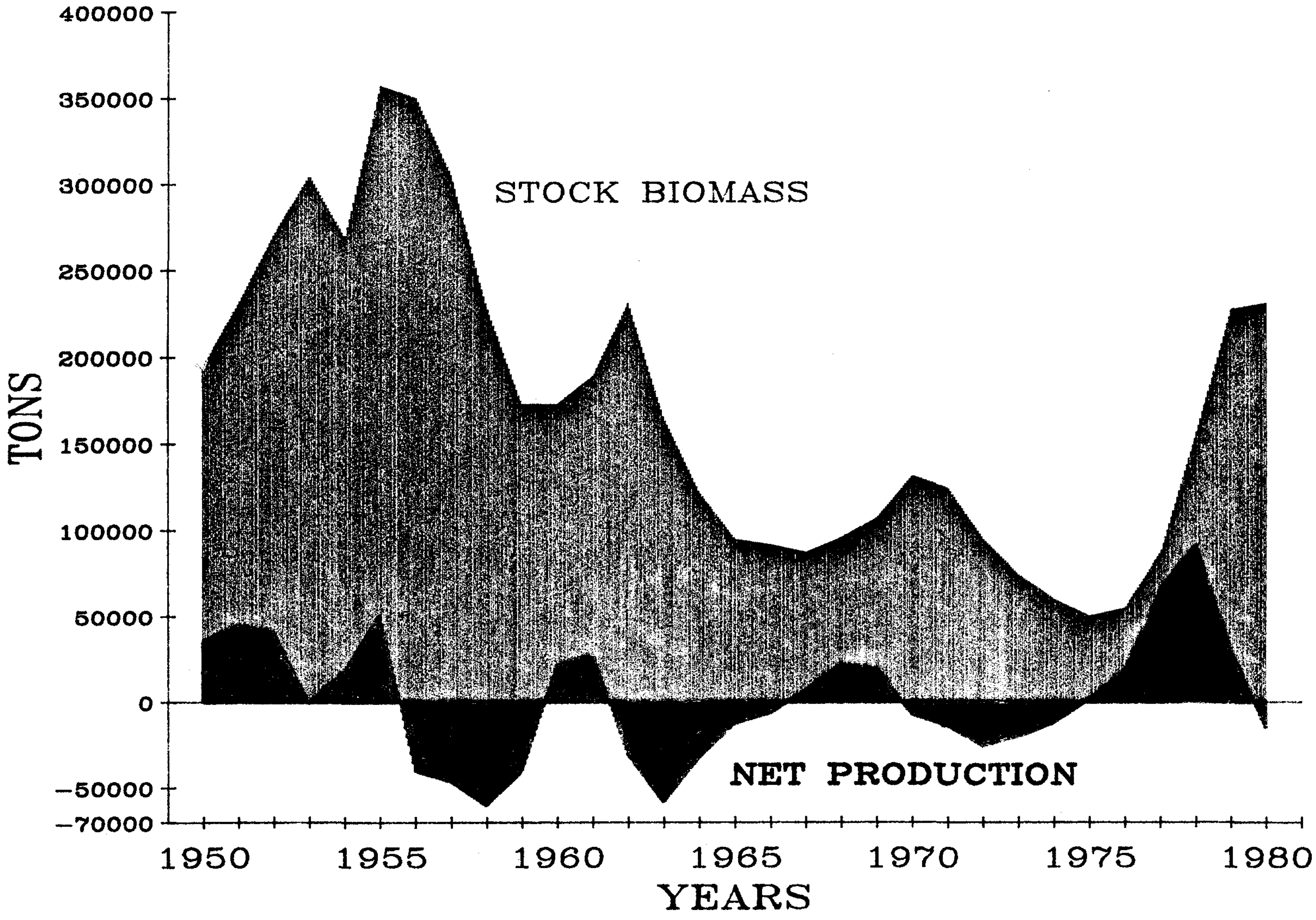


FIGURE 23. Temporal variations of stock biomass and net production for cod in 4T–Vn. Net production has been back–calculated for 1950 through 1980 from cohort analysis. Removals consistently exceeded the surplus production in the late fifties, the mid–sixties and the early seventies: this led to an important decline in stock size.

The equilibrium curve resulting from fitting the Schaefer model with data series A is represented in Fig. 22: the estimates obtained from the Gulland, the Schnute and the nonlinear methods were averaged. The observed transient path, which is defined by the observed transient points, follows the expected behavior of a stock under non-equilibrium conditions: for example, the 1977-1981 path follows the expected recovery of the stock after a period of overfishing (compare this recovery with Fig. 5B). The 2/3 F_{MEY} point is 0.21; this value is close to the $F_{0.1}$ level of 0.20 per year. The average fishing mortality for the 1978-1981 period was 0.23, which level would generate an equilibrium yield level of 67,000 t. The equilibrium yield at $F_{0.1}$ is 59,000 t. The transient projections (obtained from the fitting of the transient form) for 1982 at 2/3 F_{MEY} are 57,900 t. The projections made in Maguire and Waiwood (1982) from an age-structured algorithm were 54,900 t at $F_{0.1}$, assuming an average recruitment for 1982.

The temporal variations of stock biomass and net production for the cohort analysis presented in Maguire and Waiwood (1982) appear in Fig. 23. The late fifties, the mid-sixties and the early seventies are periods of negative net production. During these periods, removals by the fishery exceeded the surplus production. The first two periods were followed by only brief periods of positive net production and stock rebuilding. The periods of negative net production correspond to a yield spiralling downwards in Fig. 22, an indication of stock decline. This pattern is the result of uncontrolled increases in fishing pressure. The move to a new management regime in the mid-seventies, and particularly the pursuit of $F_{0.1}$ since 1977, led to a phase of positive net production and rapid stock recovery. A steady increase in stock size since 1977 is evidenced in both graphs (Fig. 22 and 23). The response of the transient path around the equilibrium curve (Fig. 22) is consistent with the phases of negative and positive net production (Fig. 23).

Haddock (4X).

The catch and catch rate series presented in Table 5 have been used to estimate the parameters of various production models by different methods. Results appear in Table 6. The first data series is the mean population biomass calculated from cohort analysis (O'Boyle et al., 1982). Catch per tow from the U.S. fall bottom trawl surveys and the Canadian summer bottom trawl surveys were also used. The first two series had 20 and 19 observations, respectively, and provided estimates of MEY ranging from 25,161 to 35,790 mt. The highest estimate, which was obtained with series A from the Schnute method, must be rejected as it is outside the range of the observations for most of the time series (in fact, since 1963). The mean of the remaining nine estimates is 27,980 t. For comparison, we calculated the average maximum productivity of 27,272 mt from yield per recruit and the 1962-81 average recruitment. The canadian summer trawl survey (12 observations) gave estimates of MEY which were outside the range of observed values of catch and effort for the Gulland and the Schnute methods. These estimates were thus rejected.

From our knowledge of the fishery, we can explain the transient path in the following manner. In response to the strong 1963 year-class, effort increased significantly in the late sixties. The 1966-69 yields relied mostly on that year-class and were above equilibrium levels both because of the change in fishing effort and higher than average recruitment (Fig. 24). Subsequently, a reduction of effort occurred in the early seventies. This reduction in effort led to a slow recovery of the stock and yield started to increase and to move towards a new equilibrium level, as expected. The 1980 and 1981 yields are above equilibrium, as a result of higher than average recruitment in the

Table 5. Catch and catch rate series used for fitting production models to haddock in NAFO Division 4X.

		A	B	C
YEAR	Catch (t)	Mean population biomass	U.S. fall bottom trawl survey (kg/tow)	Canadian summer trawl survey (kg/tow)
1962	18,550	129,022	-	-
1963	22,665	141,355	67.91	-
1964	32,042	176,163	31.43	-
1965	25,536	152,499	31.82	-
1966	38,863	150,878	58.65	-
1967	33,267	122,929	34.91	-
1968	30,624	96,946	28.53	-
1969	29,612	74,817	14.59	-
1970	17,583	62,999	17.66	22.95
1971	17,168	53,013	24.10	27.50
1972	13,283	55,802	24.45	13.24
1973	12,958	59,601	17.27	12.77
1974	13,037	62,161	20.74	39.46
1975	18,148	75,246	13.86	21.29
1976	17,366	86,466	21.94	22.21
1977	21,224	93,711	75.29	90.45
1978	26,578	115,346	53.59	26.46
1979	24,642	123,912	55.55	41.98
1980	28,580	131,616	33.47	57.10
1981	30,627	160,366	56.57	44.68

Table 6. Estimates of MEY and effort at MEY for various catch rate series, production models and estimation models. Exemple is for haddock in Div. 4X. The MEY, as calculated from yield per recruit and the 1962-1980 average recruitment (48.7 million), is 27,272 t. The Gulland method uses a four-year running average for approximating equilibrium.

Model	Estimation method	A			B			C		
		MEY	f(MEY)	R	MEY	f(MEY)	R	MEY	f(MEY)	R
Schaefer	U vs E	25,429	0.28	0.52	29,519	953	0.68	25,259	593	0.84
	Gulland	25,161	0.20	0.82	28,705	805	0.55	(43,053)	442	0.91
	Schnute	35,790	0.16	0.77	29,495	635	0.21	(33,484)	485	0.26
Pella-Tomlinson	U vs E	27,422	0.27	0.52	29,622	952	0.67	22,249	519	0.87
	Gulland	25,879	0.15	0.83	30,577	831	0.52	(52,864)	344	0.94

A: mean population biomass from cohort analysis (O'Boyle et al., 1982); 20 observations.
B: U.S. fall bottom trawl survey (kg/tow); 19 observations.
C: Canadian summer bottom trawl survey (kg/tow); 12 observations.

N.B.: ( ) identify estimates which are outside the range of observed values for catch and effort.

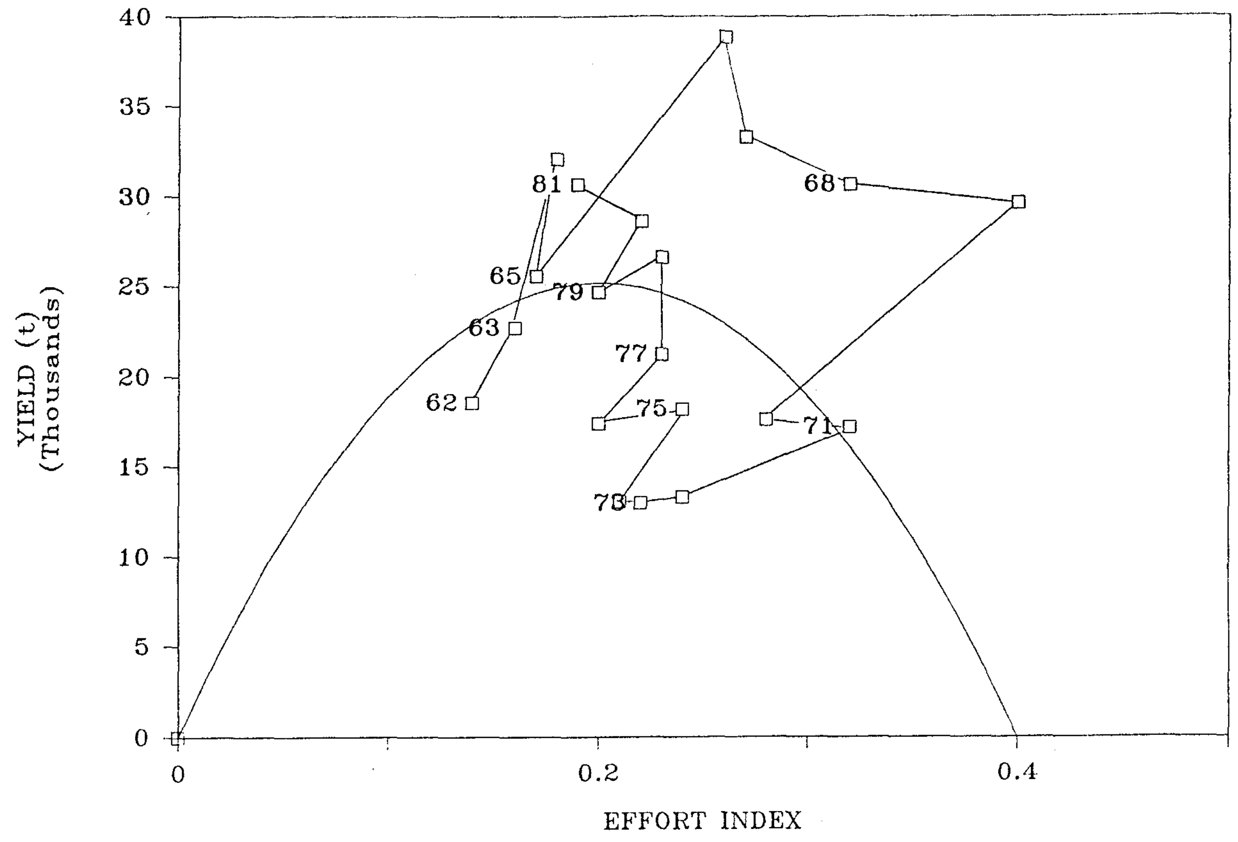


FIGURE 24. Haddock 4X. Equilibrium curve resulting from fitting the Schaefer model with data series A: the estimates were obtained from the Gulland method. The transient observed points or the observed transient path are also represented.

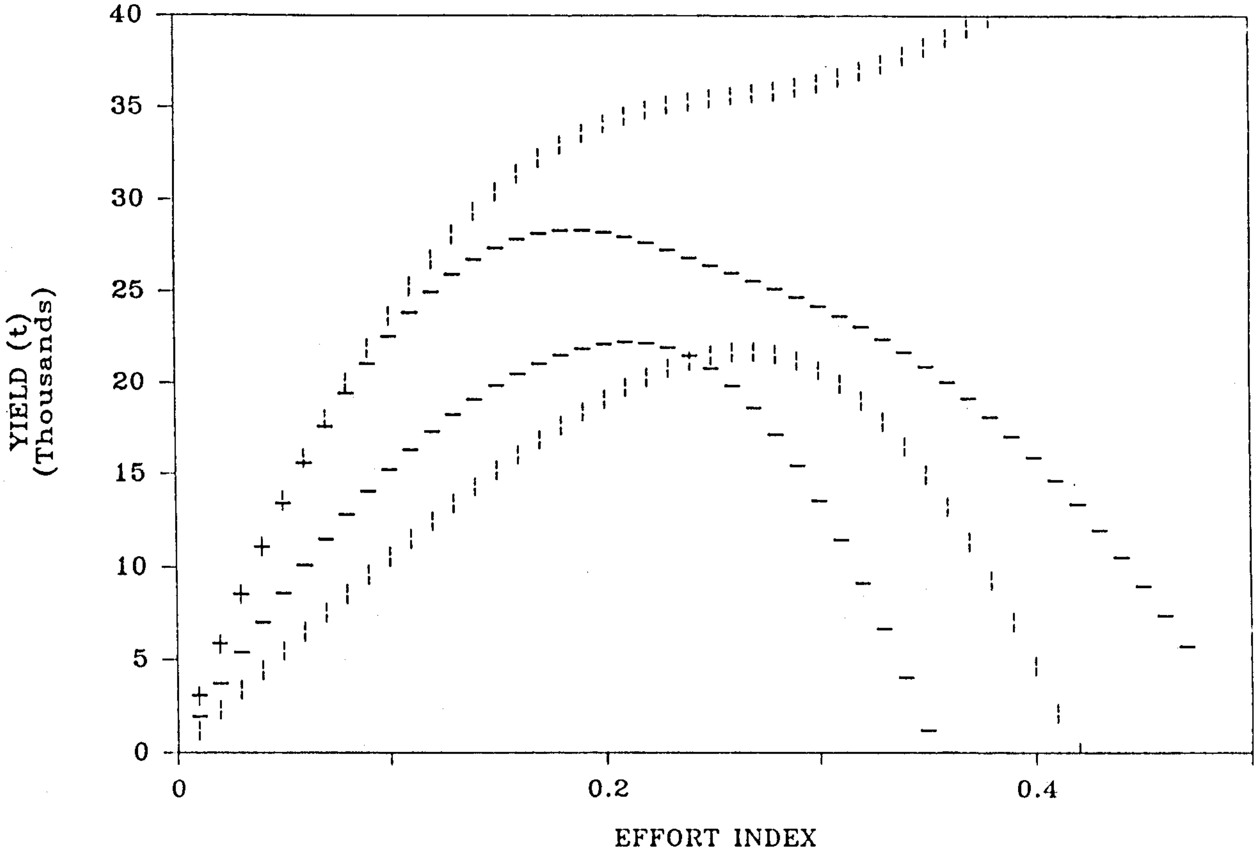


FIGURE 25. Haddock 4X. Confidence intervals (at the 95% level) for the production models resulting from the fitting of the Schaefer model by the Gulland method with data series A (identified by "–") and B (identified by "|"). These confidence intervals correspond to those arising from the linear theory.

mid to late seventies. If effort is maintained at the 1981 level, yield is likely to decrease and to fluctuate around the equilibrium level, in response to the variations of recruitment.

The confidence intervals for the production models resulting from the fitting of the Schaefer model by the Gulland method with data series A and B are plotted in Fig. 25 . In order to represent the two intervals on a single set of axes, the effort values calculated from series B were standardized to series A. The confidence intervals suggest that the model obtained from series B is not well estimated: the confidence intervals diverge quickly beyond the maximum, an indication that effort at MEY is not well estimated. Definitely, more data are required beyond the maximum for estimating the location of the descending limb of the curve with some confidence. However, both models provide similar estimations for the location of the equilibrium level for low effort values (below 0.2). For the model estimated from series A, the location of effort at MEY is also not well estimated: a wide range of values provides similar yield levels. In other words, it is not possible from these results to indicate whether the current effort level is below, at or slightly above the F_{MEY} level.

Practical guidelines for stock assessment.

With the expansion of the monitoring programs for most groundfish stocks of the Northwest Atlantic, stock assessment relies more and more on age-structured models. This search for models which mimic more closely a fish population is accompanied by a move away from production models and analysis of production. Production models are used only when the availability of data precludes the use of other models. Understanding production mechanisms is, however, the key to the understanding of stock dynamics. Production, including net and surplus production, is implicit to all models. For example, calculating production from cohort analysis (Rivard, 1980) is an attempt to look at the production underlying an age-structured model.

The simulations performed at the beginning of this chapter served to illustrate that production is a function of a number of factors and can take, at equilibrium, a wide variety of shapes when plotted against effort. Simple production models also offer a variety of shapes for the production function. The transient formulations expand their applicability and remove the need of applying them to equilibrium conditions. There are, however, a number of problems inherent to the estimation of simple production models from catch and effort data. Because these global models (even in their transient form) are not capable of simulating the complex behavior of age-structured populations, the residuals often show trends or systematic patterns, an indication that some of the assumptions regarding error structure are violated.

It is suggested that the following points be addressed when applying production models to catch and effort data:

A) Are the basic assumptions met? For example: presence of migrations; expansion of the fishing grounds (e.g. in a new and developing fishery); existence of changes in regulation (e.g. mesh size) which may affect the definition of the stock; ability of catch rates to reflect changes in stock size.

B) Which data series can be used and do they contain enough observations? Because residuals are often not independent, we must use more observations than dictated by a normal regression case. As a practical reference point, a minimum of fifteen years of observations is recommended but this number must be increased considerably for long lived species such as redfish. Finally, each data series must be discussed in terms of its potential biases.

C) Which estimation procedure will be used? The Gulland method is more robust than the Schnute method or the catch-effort regression when a small number of observations is available. Apply it first. If the estimates of MEY and f_{MEY} are within the range of observed catches and effort, respectively, then the nonlinear methods or the Schnute method can be used. When using nonlinear estimation procedures for estimating the parameters of the transient system, use different sets of starting values to see if the iterative process converges to the same results. It is also recommended to minimize the number of unknowns by using the approximations given in Rivard and Bledsoe (1978) for B_{∞} and B_0. Alternatively, if some parameters (e.g. q) can be estimated independently, use this independent estimate and estimate only the remaining parameters of the model.

D) Validity of regression estimates? Are the regression coefficients significant? If not, reject the resulting model. Also, compare the estimates obtained from different data series or from different methods. If the estimates of MEY and f_{MEY} are considerably outside the range of observed values, reject them as they do not contain information on the location of the optimum. Finally, draw confidence intervals around the equilibrium curve.

E) Trends or patterns in the residuals (often referred to as "runs" or as "serially correlated residuals"), may indicate that the resulting parameter estimates are biased. The bias is generally small with a large number of observations (say greater than thirty observations) but may be significant for small data series.

F) Plot the results and analyse the transient path in relation to the equilibrium curve. This is better achieved with a graphical representation of the resulting model. Can deviations from equilibrium be explained in terms of strong/weak year-classes, changes in fishing intensity, etc.?

G) Proceed with short term projections at the desired effort level by using the transient form of the model. If the parameters used come from the fitting of a transient form, discuss the short term projections in light of the patterns observed in the residuals. These patterns can usually be linked to extreme conditions, such as good or poor recruitment years and will often provide an indication of the likely importance and direction of bias in the projections. Alternatively, Stocker and Hilborn (1981) show that the use of simple time series models works very well for short term forecasts. They conclude that the stock production models are no better in forecasting next year's catch rates than a time series model. Time series models provide no usable information, however, for formulating long-term management objectives.

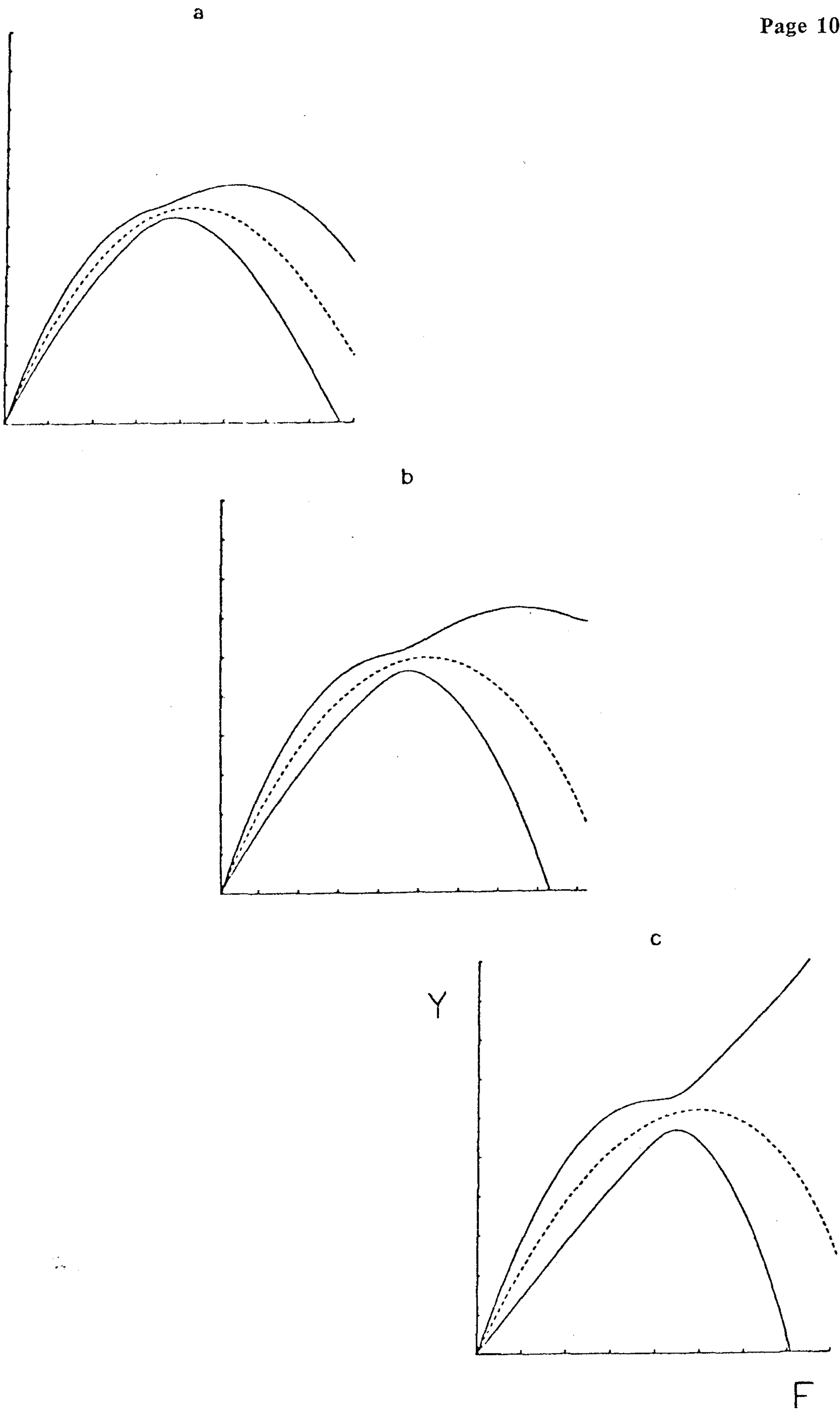


FIGURE 26. Confidence regions for the equilibrium yield when only a limited number of observations is available beyond the maximum.

The fitting of production models assumes that changes in catch rates are due to changes in effort. When there is little change in effort, the model will not provide information on the likely changes in catch rates in response to changes in effort. The current management objectives, for instance the pursuit of $F_{0.1}$, is unlikely to improve our ability to use production models for assessment purposes.

Finally, it is important to understand that the resulting estimates contain information on average productivity only for the time frame covered by the observations. Short time frames may be strongly influenced by the peculiarities of recruitment and growth during that period. Production models describe the response of yield to fishing effort given this average historical productivity. Changes in environmental conditions may affect drastically future productivity and, thus future yield from the stock.

Always keep in mind that the "information content" of your data series may not be sufficient to make strong inferences regarding the actual location of the equilibrium curve for a certain range of effort values. This is illustrated in Fig. 26. When only a limited number of observations are available beyond the maximum, confidence intervals may diverge quickly beyond the maximum. However, the location of equilibrium levels at low effort values may still be well estimated: for example, see how tight the confidence intervals are below the maximum in Fig. 26. Consequently, the ascending portion of the curve may be used for assessing the response of yield at low effort levels but the model would provide no useful information for the formulation of long term strategies at higher effort levels. In the absence of a precise estimation of equilibrium levels beyond the maximum, production models may still provide information on the response of the stock within the range of historical effort levels.

References

Allen, K.R. 1971. Relation between production and biomass. J. Fish. Res. Bd. Can. 28: 1573-1581.

Cadima, E.L. 1976. Synthetic models. FAO Fisheries Circular No 701.

Bagenal, T.(Ed.) 1978. Methods for assessment of fish production in fresh waters. IBP Handbook No. 3. Blackwell Scientific Publications, Oxford. 365 p.

Chapman, D.G. 1967. Statistical problems in the optimum utilization of fishery resources. Int. Statist. Inst., Bull. 42(1): 268-290.

Doubleday, W.G. 1976. Environmental fluctuations and fisheries management. Int. Comm. Northwest Atlantic Fisheries, Selected Papers 1975: 141-149.

Draper N.R. and H. Smith. 1966. Applied regression analysis. John Wiley & Sons Inc., New York. 407 p.

Fletcher, R.I. 1978. On the restructuring of the Pella-Tomlinson system. Fish. Bull., 76(3): 515-521.

Fox, W.W., Jr. 1970. An exponential surplus-yield model for optimizing exploited fish populations. Trans. Amer. Fish. Soc., 99(1): 80-88.

Fox, W.W., Jr. 1971. Random variability and parameter estimation for the generalized stock-production model. Fishery Bulletin 69: 569-580.

Gavaris C.A. and D.B. Atkinson. 1982. Assessment of redfish in divisions 4RST. CAFSAC Res. Doc. 82/30. 24 p.

Graham, M. 1935. Modern theory of exploiting a fishery and application to North Sea trawling. J. Cons. Int. Expl. Mer 10: 264-274.

Gulland, J.A. 1961. Fishing and the stocks of fish at Iceland. Min. Agr. Fish and Food (G.B.), Fish. Investig., Ser. II, 23(4): 1-32.

Gulland, J.A. 1969. Manual of methods for fish stock assessment. Part 1. Fish Population Analysis. FAO Man. Fish. Sci. 4, 154 p.

Gulland, J.A. 1978. Assessment of a fishery. **In** Bagenal T. (Ed), Fish Production in Fresh Waters. IBP Handbook No. 3. Blackwell Scientific Publications. Oxford. 365 p

Jensen, A.L. 1973. Relation between simple dynamic pool and surplus production models for yield from a fishery. J. Fish. Res. Bd. Can. 30(7): 998-1002.

Larkin, P.A. and W.E. Ricker. 1964. Further information on sustained yields from fluctuating environments. J. Fish. Res. Bd. Can. 21(1): 1-7.

Lett, P.F. 1978. A comparative study of the recruitment mechanisms of cod and mackerel, their interaction, and its implication for dual stock management. Ph. D. Thesis, Dalhousie University. 125 p.

Maguire, J.J. and K.G. Waiwood. 1982. The 1982 assessment of cod in NAFO Division 4T and Subdivision 4Vn (Jan.-Apr.). CAFSAC Res. Doc. 82/22.

Mathew, C.P. 1970. Estimates of production with reference to general surveys. Oikos 21: 129-133.

Mohn, R.K. 1980. Bias and error propagation in logistic production models. Can. J. Fish. Aquat. Sci. 37: 1276-1283.

O'Boyle, R.N. and G.N. White III. 1982. An evaluation of the 4X haddock population characteristics during 1962-81 with yield projected to 1983. CAFSAC Res. Doc. 82/53.

Pella, J.J. and P.K. Tomlinson. 1969. A generalized stock production model. Inter-Amer. Trop. Tuna Comm., Bull. 13(3): 419-496.

Poole R.W. 1974. An introduction to quantitative ecology. McGraw-Hill Inc., U.S.A. 532 p.

Ricker, W.E. 1958. Maximum sustained yields from fluctuating environments and mixed stocks. J. Fish. Res. Bd. Can. 15(5): 991-1006.

Rivard, D. and L.J. Bledsoe. 1978. Parameter estimation for the Pella-Tomlinson stock production model under non-equilibrium conditions. Fish. Bull., 76(3): 523-534.

Rivard, D. 1980. Back-calculating production from cohort analysis, with discussion on surplus production for two redfish stocks. CAFSAC Res. Doc. 80/23.

Rivard, D. 1982. APL programs for stock assessment (revised). Can. Tech. Rep. Fish. Aquat. Sci. 1091: 146 p.

Schaefer, M.B. 1954. Some aspects of the dynamics of populations important to the management of the commercial fisheries. Inter-Amer. Trop. Tuna Comm., Bull. 1(2): 26-56.

Schaefer, M.B. 1957. A study of the dynamics of the fishery for yellowfin tuna in the eastern tropical Pacific Ocean. Inter-Amer. Trop. Tuna Comm., Bull. 2(6):245-285.

Schnute, J. 1977. Improved estimates from the Schaefer production model: theoretical considerations. J. Fish. Res. Board Can. 34: 583-603.

Sinclair M. and J.J. Maguire. 1981. 4TVn (Jan.-Apr.) cod assessment. CAFSAC Res. Doc. 81/41.

Stocker, M. and R. Hilborn. 1981. Short term forecasting in marine fish stocks. Can. J. Fish. Aquat. Sci. 38: 1247-1254.

Tautz, A., P.A. Larkin and W.E. Ricker. 1969. Some effects of simulated long-term environmental fluctuations on maximum sustained yield. J. Fish. Res. Bd. Can. 26(10): 2715-2726.

Uhler, R.S. 1979. Least squares regression estimates of the Schaefer production model: some Monte Carlo simulation results. Can. J. Fish. Aquat. Sci. 37: 1284-1294.

Walter, C.G. 1973. Delay-differential equation models for fisheries. J. Fish. Res. Bd. Can. 30(7): 939-945.

Walter, C.G. 1976. Non-equilibrium regulation of fisheries. International Commission for the Northwest Atlantic Fisheries, Selected Papers No 1, 193 p.

Waters, T.F. 1969. The turnover ratio in production ecology of freshwater invertebrates. Am. Nat. 103: 173-185.

The Development and Utility of Sequential Population Analysis in Stock Assessments

by **G.H. Winters**

Science Branch,
Department of Fisheries and Oceans,
P.O. Box 5667,
St. John's, Newfoundland
A1C 5X1

Introduction

Current fisheries management policies depend largely on real time estimates of stock size, the proportion of the annual removal from the stock by fishing, and the relationship of this removal rate to the long-term sustained productivity of the stock. The number of fish removed by fishing activity can be obtained from landing statistics; stock size estimates may then be made from catches, provided the exploitation ratio, F/Z is known, since formally

$$C_t = (1 - e^{-Z}) N_t F / Z \quad .$$

This ratio has traditionally been determined by assuming (or being given) some value for the natural mortality rate, M, and then calculating a value for Z from the change in average density (catches per unit of effort) of fish of a particular year-class in consecutive years. If M does not vary significantly, this approach should provide reasonable estimates of stock size and fishing mortality rates providing that the rate of change in stock abundance can be adequately reflected in the sample data on catch and effort from which density estimates are derived. This in turn assumes that the average annual performance characteristics of the fleet have remained the same in relation to the distribution of stock density or that appropriate conversion factors can be formulated to provide comparable effort measurement units. However, in many recent fisheries (particularly since the late 1960's) it has become extremely difficult to compare directly fleet performance characteristics from year to year because of the introduction of various types of new gear, equipment etc., which in combination with subtle changes in fishing strategy, have changed the efficiency of fishing. Any such unaccounted changes in the behaviour of the fleet towards the resource make effort comparisons (and hence density changes) unreliable. Even in fisheries where the catch capacity has remained the same from year to year, effort statistics are difficult to collect and compare because of international involvement and the use of mixed gears.

As a result of these difficulties in obtaining reliable effort measurements, in recent years fisheries biologists have depended mainly on sequential population analyses for stock assessment purposes. This method allows estimates of stock size and fishing mortality rate to be made from catch data independently of effort measurement. The more common expression of sequential population analyses, the term "virtual population analyses" was coined by Fry (1949, 1957) in referring to the sum of a cohort of fish present in the water at any given time that are destined for subsequent capture in the fishery. This sum is obtained by summing backwards from the terminal year all the contributions that a particular year-class has made to the fishery. Because the virtual population method or its modifications have become of paramount importance in fisheries management problems and since no comprehensive review of the method is available in the literature, this chapter outlines the development of the method and appraises its utility in fisheries research.

Historical Development of the Sequential Population Concept

The concept of virtual populations was first introduced by Derzhavin (1922) who constructed a table of the probable absolute age composition of sturgeon (*Acipenser stellatus*) catches in the Kura River from 1881 to 1915 from age composition data averaged over a period of years (Ricker 1958). Vertical summation of the columns of his tables represented the virtual populations or complete contribution of various year-classes to the fishery. Under assumptions of stable age distribution, constant recruitment and negligible natural mortality, Derzhavin estimated the total commercial stock of sturgeon present in the sea in a particular year as consisting of that year's catch plus the next year's catch less the number of Age I fish in it, plus the next year's catch diminished by the number of Age I and Age II fish, and so on. Thus if x_t represents the proportion of age t fish in the catch and r the oldest age involved, the Derzhavin expression for the virtual population at age then becomes (Ricker 1948)

$$N = C_1 + C_2(1-x_1) + C_3(1-x_1-x_2) + \dots + C_r(1-x_1-x_2- \dots -x_{r-1}) \quad . \tag{1}$$

This expression can be interpreted as the complete contribution of a particular year-class only under the conditions stated above. If only the exploitable stock size were required and the fish did not recruit, say until Age III, then the first two terms of expression (1) would be omitted. It may also be noted that since natural mortality was assumed to be negligible for sturgeon, Derzhavin's calculation supplied an estimate of the actual rather than the virtual population.

Unlike Derzhavin, Fry (1949) did not require the assumptions of stable age distribution, negligible natural mortality or constant recruitment as continued sampling provided him with yearly estimates of age compositions and this, together with catch data, enabled him to compute the total contribution of a particular year-class of fish as it passed through the fishery. Fry termed this total contribution of a year-class its virtual population because of the analogy with the virtual image of physics.

From age structure and catch data of Lake Opeongo trout, Fry summed backwards the successive catches of each year-class (V_t) thus providing complete virtual population sizes for each year of the fishery. Two statistics were readily derived from such data; (1) a minimum estimate of population size, i.e. V_t, and (2) by calculating the percentage of the virtual population removed in each fishery year, a maximum estimate of the rate of exploitation can be determined, i.e. $u(\max) = C_t/V_t$. Fry however was more interested in obtaining estimates of fishing mortality rates in each year. This was accomplished by first estimating the total mortality rate Z from survival rates as calculated from virtual population sizes. Survival rates during each year of the fishery were calculated as follows:

$$S_n = V_{n+1} / V_n \quad , \tag{2}$$

where V_n = number of fish of year-class which are caught in year n or later. That is

$$V_n = E_n N_n$$

and

$$V_{n+1} = E_{n+1} N_{n+1}$$

where E_n is the exploitation ratio, the proportion of fish of year-class which was alive at the beginning of year n and subsequently caught, i.e. $E = F / Z$. But if

$$N_{n+1} = N_n e^{-Z} \quad ,$$

with constant mortality for consecutive years,

$$S_n = E N_n e^{-Z} / E N_n = e^{-Z} \tag{3}$$

and $Z = -\ln S$. Fry estimated natural mortality by plotting values of Z against estimates of the annual fishing effort (which were available from his data series) and taking the intercept value (i.e. $F = 0$) as being an estimate of M. Fishing mortality rates were then simply obtained from Z by subtraction of M.

Various Methods of Sequential Population Analyses

Method of Ricker (1948)

The mechanics of sequential computation were first described by Ricker (1948) in an application to the partially-recruited ages of a halibut stock. Using estimates of M and fishing mortality in the oldest age of a cohort, Ricker computed the terminal population size (N_{i+1}) from the Baranov (1918) catch equation as follows:

$$N_{i+1} = C_{i+1} / u_{i+1} \quad , \qquad (4)$$

where C_{i+1} = catch at age i+1 and

$$u_{i+1} = (1 - e^{-Z_{i+1}}) F_{i+1} / Z_{i+1} \quad .$$

N_i is then calculated from the following relationships. By definition,

$$N_i = C_i / u_i \quad \text{and} \quad N_i = N_{i+1} / S_i \quad ;$$

therefore

$$C_i / u_i = N_{i+1} / S_i$$

whence

$$N_{i+1} / C_i = S_i / u_i \quad . \qquad (5)$$

For a given value of M, u has a unique solution which may be derived by trial and error or from a set of tables relating F to u/S.

The calculation procedure for this method is shown in Appendix I based on the data in Table 1 and the calculations displayed in Table 2.

Method of Gulland

The fish catch in any year can be expressed as a function of fishing and natural mortality rates during the year and the population at the end of the year. If natural mortality is assumed to be constant and an initial estimate of F in the old age-groups is also assumed, Gulland (1965) demonstrated that it was possible to determine the varying fishing mortality during the life span of each year-class by calculating back from year to year from the older to the younger fish.

Thus if a year-class is i years old in year t

$$r_i = N_{i+1} / C_i \quad ,$$

Table 1. Population model of a cohort of fish subjected to arbitrary fishing mortality rates. M has been assumed constant at .20.

Age	Actual population	Total deaths	Catch (C_i)	Virtual population (V_i)	F	Z	Z^1 (- log S)
3	70000	15484	3097	33014	.05	.25	.10
4	54516	14131	4710	29917	.10	.30	.17
5	40385	14631	8129	25207	.25	.45	.39
6	25754	11620	7747	17078	.40	.60	.60
7	14134	5562	3337	9331	.30	.50	.44
8	8572	4315	3082	5994	.50	.70	.72
9	4257	1675	1005	2912	.30	.50	.42
10	2582	1422	1066	1907	.60	.80	.82
11	1160	523	349	841	.40	.60	.54
12	637	378	294	492	.70	.90	.91
13	259	154	120	198	.70	.90	.93
14	105	58	43	78	.60	.80	.80
15	47	24	17	35	.50	.70	.67
16	23	15	12	18	.80	1.00	1.10
17	8	8	6	6			

Table 2. Estimation procedures for F_i and N_i by the methods of Ricker (1948) and Gulland (1965).

	Method of Ricker				Method of Gulland							
Age	C_i (1)	$\frac{N_{i+1}}{C_i}$ (2)	F_i (3)	N_i (4)	$S^1{}_i = \frac{V_{i+1}}{V_i}$ (5)	$\frac{S_i}{a}$ (6)	$\frac{F}{Z}$ (E_{i+1}) (7)	$\frac{SZ}{Fa}$ (r_i) (8)	S_i (9)	$\frac{Fa}{Z}$ (Ea) (10)	F_i (11)	N_i (12)
3	3097	17.60	0.05	70000	.906	9.64	.55	17.56	0.78	.04	.05	70000
4	4710	8.63	0.10	54515	.842	5.33	.63	8.53	0.74	.09	.10	54516
5	8129	3.20	0.25	40385	.678	2.11	.66	3.17	0.64	.20	.25	40385
6	7747	1.87	0.39	25754	.546	1.20	.66	1.82	0.55	.30	.40	25754
7	3337	2.59	0.30	14135	.642	1.79	.70	2.57	0.61	.24	.30	14134
8	3082	1.39	.50	8572	.487	0.95	.69	1.38	0.50	.36	.50	8572
9	1005	2.57	0.30	4257	.655	1.90	.74	2.57	0.61	.24	.30	4257
10	1066	1.09	0.60	2584	.441	.79	.73	1.08	0.45	.41	.60	2581
11	349	1.80	0.40	1163	.585	1.41	.78	1.82	0.55	.30	.40	1159
12	294	0.87	0.71	628	.402	0.67	.77	0.87	0.40	.47	.71	632
13	120	0.87	0.71	257	.394	0.65	.76	0.86	0.40	.47	.71	256
14	43	1.07	0.60	104	.449	0.82	.76	1.07	0.44	.42	.61	103
15	17	1.31	0.52	46	.514	1.06	.81	1.31	0.49	.37	.52	46
16	12	0.63	0.88	22	.333	0.50	.80	0.63	0.34	.54	.89	22
17	6	-	0.80	7.5	-	-	-	-	-	-	-	8

i.e. r is the ratio of the population at the end of the year to the catch during the year. However, as shown previously, $N_{i+1} = N_i S_i$ and $C_i = a\ N_i\ F/Z$ (Baranov's catch equation) where a is annual total mortality rate and r_i becomes

$$r_i = N_i S_i \ / \ N_i \ a \ (F/Z) \quad .$$

Consequently,

$$r_i = S_i Z_i \ / \ F_i a_i = S_i / E_i \ a = S_i / u_i \quad . \qquad (6)$$

A first estimate of the survival of a year-class at age i is given by the ratio of the virtual population of that year-class at the beginning and end of the year, i.e.

$$S_i^1 = V_{i+1} \ / \ V_i \quad .$$

Under conditions of constant F expression, (6) becomes

$$r_i = S \ / \ E_{i+1} \ a \quad .$$

Thus r_i is a simple function of the survival rate during the year (as determined from virtual populations) and the exploitation ratio E_{i+1} to which fish are subjected in the following year.

Under conditions of constant F, the exploitation ratio E_i applicable to fish at the beginning of year i will be the sums of the proportions of fish alive at the beginning of the year caught during that year, and subsequently, i.e.

$$E_i = E\ a + S_i \ E_{i+1} \quad . \qquad (7)$$

Thus, if values of M and (initially) E_{i+1} are assumed, r_i can be determined and values for F_i can be calculated or read directly from work tables of $SZ \ / \ FA$ against F. Estimates can then be observed in succession of r_{i-1} F_{i-1} and E_{i-1} ... etc. The calculation procedure and a worked example for Gulland's method are shown in Table 2.

Method of Jones

The virtual population model of Gulland may be used for estimating rates of fishing mortality only on year-classes that have passed completely through the fishery. This obviously precludes using data from recent year-classes. The dynamic capacity of modern fishing fleets, however, makes it extremely desirable to have estimates of F as up to date as possible. Jones (1961, 1968) presented a method that allows the calculation of F when a year-class is still present in the fishery. This is accomplished by using the ratio of the catch of one year-class in two consecutive years to determine a starting value for r_i, eg. C_i and C_{i+1}.

Let C_i = catch of a year-class at age i and C_{i+1} = catch of same year-class at age i + 1.

Then

$C_i = E_i \, a_i \, N_i$ (from Baranov catch equation)

and

$$C_{i+1} = E_{i+1}\, a_{i+1}\, N_{i+1} \quad .$$

Since $N_{i+1} = N_i\, S_i$, the ratio

$$C_{i+1} \,/\, C_i \;=\; E_{i+1}\, a_{i+1}\, N_i\, S_i \,/\, E_i\, a_i\, N_i$$

and by eliminating N_i from the numerator and denominator, we get

$$C_{i+1} \,/\, C_i \;=\; S_i\, (\, E_{i+1}\, a_{i+1}\,) \,/\, (\, E_i\, a_i\,)$$

Dividing both sides by $E_{i+1}\, a_{i+1}$ we get

$$C_{i+1} \,/\, C_i\, E_{i+1}\, a_{i+1} \;=\; S_i \,/\, E_i\, a_i \quad . \qquad (8)$$

This is the same as expression (5) for r_i of the Gulland method.

Thus at a given value for M and an estimated starting value for $E_{i+1}\, a_{i+1}$, a value can be calculated for F or read from a work table (Schumacher 1970). The calculation procedure and a worked example for this method are shown in Table 3.

Method of cohort analyses

Cohort analysis is a modified form of population analyses developed by Pope (1972) who considered the method an approximation to Gulland's virtual population analysis which is usable at least up to values of $M = 0.3$ and $F = 1.2$. The method is based on estimates of population sizes of a cohort of fish at successive age intervals from which estimates of total mortality can be derived. By simple subtraction of M (assumed or determined), values of F can be obtained for each age-group.

Pope's cohort analysis is based on the approximate formula

$$N_i \;=\; C_i\, e^{M/2} \;+\; N_{i+1}\, e^{M} \quad , \qquad (9)$$

where N_i = population of a year-class at the i^{th} birthday

C_i = catch of that year-class at age i.

Thus using (9) as a recurrence relationship

$$N_i \;=\; C_i\, e^{M/2} + C_i\, e^{3M/2} + C_i\, e^{5M/2} + N_t\, e^{(t-i)M}$$

If the year-class has not passed completely through the fishery, C_i refers to the catch in year i only. In this case

$$N_i \;=\; C_i\, Z_i \,/\, F_i\, a_i \quad .$$

When the year-class has passed completely through the fishery, C_i refers to the catch in year i and all subsequent years, i.e. $C_i = V_i$. In this case

$$N_i \;=\; V_i\, Z_i \,/\, F_i \quad .$$

Table 3. Estimation procedure for F and N by the methods of Jones (1968) and Pope (1972). Starting values for M and F are .20 and .80 respectively. Adjustments have been made for rounding errors.

Age	Method of Jones (1968)					Method of Pope (1970)				
	$\frac{C_{i+1}}{C_n}$ (1)	$\frac{Fa}{Z}$ $(E_{i+1}\ a_{i+1})$ (2)	$\frac{\left(\frac{C_{n+1}}{C_n}\right)^Z}{Fa}$ (3)	F (4)	N_i (5)	$C_i e^{M/2}$ (6)	$N_{i+1} e^M$ (7)	N_i (8)	lN $\frac{N_{i+1}}{N_i}$ (9)	F (10)
3	1.521	.086	17.686	.050	70000	3422.8	67088.1	70510	.250	.050
4	1.726	.201	8.587	.100	54516	5205.5	49721.7	54927	.300	.100
5	.953	.301	3.166	.250	40385	8984.2	31724.6	40709	.449	.249
6	.431	.236	1.826	.400	25754	8562.0	17412.0	25974	.600	.400
7	.924	.360	2.567	.300	14134	3688.1	10567.7	14256	.499	.299
8	.326	.236	1.381	.500	8572	3406.2	5245.9	8652	.700	.500
9	1.061	.412	2.575	.300	4257	1110.7	3184.3	4295	.499	.299
10	.327	.299	1.094	.598	2587	1178.1	1529.0	2607	.801	.601
11	.842	.457	1.842	.397	1167	385.7	784.3	1170	.600	.400
12	.408	.452	.903	.690	643	324.9	318.4	643	.903	.703
13	.361	.394	.916	.680	265	132.6	128.1	261	.910	.710
14	.389	.328	1.186	.562	110	47.5	57.4	104	.803	.603
15	.698	.435	1.604	.445	52	18.8	28.2	47	.710	.510
16	.500	.506	.988	.644	27	13.3	9.8	23		.600*
17					12			8		.600

*assumed

In either case

$$F_i = \ln [\, N_{i+1}/N_i \,] - M \qquad . \qquad (10)$$

The calculation procedure and a worked example for cohort analyses are shown in Table 3.

Doubleday's (1981) Survivors Method

Doubleday (1981) developed a general formula of sequential population analysis which represented abundance at age as the sum of integrated catches and integrated survivors, adjusted for natural mortality. This general formula is as follows

$$N_i(t) \;=\; \int_t^T \exp \int_t^T M(y)\, \mathrm{d}y \;\; C_i(\tau)\, \mathrm{d}\tau \;+\; S_i \exp \int_t^T M(y)\, \mathrm{d}y \qquad , \qquad (11)$$

where S_i = survivors of year-class i are date T

T = terminal date (year).

Thus the population is divided into one component corresponding to future catches and another component corresponding to survivors. Equation (11) is equivalent to Pope's cohort analyses when catches occur at mid-year and includes also Gulland's formula if catches are distributed throughout the year according to an exponential decline. Since population estimates are mainly required for catch projections, Doubleday's main interest was on the survivors and only on historical abundance as a means of calibrating an independent abundance index such as research vessel survey data. Doubleday's method therefore involves an iterative procedure by which a maximum likelihood estimate of survivors as derived from a least squares analysis of the relationship between an abundance index (usually research vessel surveys) and sequential population estimates. The reader is referred to Doubleday (1981) for further details on the implementation procedures. Output from the survivors model includes historical population numbers, estimated survivors and their variances as well as the age specific coefficients of variation for the terminal year population numbers. Advantages of the survivors method are that repeated surveys provide multiple estimates of the same year-class and therefore there is not the same dependence on terminal fishing mortality rates as in the Gulland or Pope methods. In addition, the model can accomodate date-dependent natural mortality rates and, since survivor estimates have variances associated with them, several independent estimates of survivors can be efficiently combined. The major disadvantage of the Survivors Method is the requirement for a time series of population abundance index such as research vessel data which often have high variability associated with them.

Method of Collie and Sissenwine (1983)

Collie and Sissenwine (1983) developed a modified DeLury model which was capable of estimating population size from research vessel survey information on relative abundance. The model can be applied to either age-structured or non-age-structured catches and essentially calculates a catchability coefficient using a method which simultaneously smooths the relative abundance indices by accounting for measurement variability in research vessel data. Considering only age-structured data, the DeLury Method (1947) as modified by Allen (1966) becomes:

$$N_{t+1,i+1} = [\, N_{t,i} - C_{t,i} \,] \exp(-M) \quad , \qquad (12)$$

where $N_{t,i}$ and $C_{t,i}$ refer to the age-specific populations size (in numbers) and catch respectively in year t. The survey abundance $P_{t,i}$ is related to the population size by a catchability coefficient q, which is constant for all age-groups such that:

$$P_{t,i} = q\, N_{t,i} \quad . \qquad (13)$$

Substituting for $N_{t,i}$ in equation (12) and multiplying by q, we obtain the following:

$$P_{t+1,i+1} = [\, P_{t,i} - q\, C_{t,i} \,] \exp(-M) + e_{t,i} \quad , \qquad (14)$$

where $e_{t,i}$ is the equation error.

The measured relative abundance $P_{t,i}$ (logarithmically transformed) is related to the "true" relative abundance by:

$$P_{t,i} = P_{t,i} \exp(n_{t,i}) \quad , \qquad (15)$$

where $n_{t,i}$ are normally distributed random variables with mean one. Equations (14) and (15) are then solved simultaneously for q by a non-linear least-squares estimation procedure. Estimates of population size (with confidence intervals) can then be calculated by dividing the survey catch rates by q.

The major advantage of Collie's and Sissenwine's model is that it provides an objective basis for estimating population size and fishing mortality rates in the terminal year. In addition, it can also be used for those stocks for which catches and/or survey data are not aged. Its major disadvantage is that q cannot be estimated separately for each age-group and therefore estimates of population size by age will be biased but the total population estimates should be unbiased.

Methods of Separable VPA

Doubleday's (1976) Least Squares Approach

The previous models of sequential population analysis evaluate the catches of a particular cohort as a unit. Their major disadvantage is that n+1 independent estimates have to be derived from n equations which utilize the catch-at-age data from the n ages of a year-class. Solution to this problem requires making an assumption about at least one of the parameters which usually takes the form of guessing the value for terminal fishing mortality rate. Doubleday (1976) introduced a method of analysing catch-at-age data by linearized least squares which takes into account stochastic variability in the observed catch data and attempts to extract the additional information which can be gained by analysing various cohorts simultaneously. A key element of Doubleday's approach was the separation of the instantaneous fishing mortality rate (F) into its two components, i.e. the year effect (fully recruited F or effective fishing effort) and the age effect (partial recruitment rate, say S) as suggested by Pope (1974).

Given that

s_i = logarithm to base e of availability (i.e. partial recruitment) in age i;

f_t = logarithm of effective effort multiplier (i.e. fully recruited F) in year t;

${}_iF_t = \exp(s_i + f_t)$;

r = age at first recruitment;

Doubleday transformed the Baranov catch equation by taking logarithms and expanding $\ln {}_iF_t$ as $s_i + f_t$:

$$\begin{aligned}\ln {}_iC_t \;=\; & \ln({}_rN_{t-i+r}) - (i-r)M + (s_i + f_t) \\ & - \sum_{n-r}^{i-1} \exp(s_i + f_{t-i+1}) - \ln(\exp(s_i + f_t) + M) \\ & + \ln(1 - \exp(-\exp(s_i + f_t) - M)) + {}_i\varepsilon_t \quad . \qquad (16)\end{aligned}$$

In equation (16), ${}_i\varepsilon_t$ represents the sampling error in the observed $\ln({}_iC_t)$ and is assumed to have a zero mean and constant variance for all ages and years.

An initial set of estimates of s_i and f_t for all ages and years are used to expand the non-linear terms in (16) by Taylor series. A second algorithm, based on logarithms of catch ratios, is used to provide starting values for an iterative procedure which applies standard least squares to the resulting linear approximations in an iterative manner until two successive estimates of all parameters differ by less than 10^{-2}.

The advantages of this least squares approach over the traditional methods of sequential population analyses are as follows (Doubleday 1976):

-1) The arbitrary choice of terminal F is removed since there are now fewer parameters than observations.

-2) The residual mean square provides an indication of how well the catches are explained.

-3) The variance estimates provide an indication of the reliability of the parameter estimates.

-4) The goodness of fit of the model to the data is provided as output.

-5) Variance estimates associated with stock size estimates quantify the amount of information in the catch data about stock sizes.

Doubleday's method has, however, some serious disadvantages (Doubleday 1976):

-1) There is no guarantee of a unique solution and therefore independent data are required to evaluate the output.

-2) The method requires a long time series of well sampled catches during which partial recruitment has not changed greatly. The analysis of variance table may mean very little for data sets having less than 10-15 years.

-3) Confidence limits for fishing mortality estimates can be quite large even though the catches are well explained.

-4) Natural mortality is assumed to be constant in time and independent of age.

Although these disadvantages restrict the use of the model, its output may provide an insight into changes in the fishery, particularly with regard to variation in selection at age and fishing mortality rates.

Method of Fournier and Archibald (1982)

Fournier and Archibald (1982) considered Doubleday's (1976) model to be somewhat unsatisfactory due to the assumption by Doubleday (1976) that the variance in $\ln({}_iC_t)$ was constant for all ages and years. They proposed an alternative, more flexible, model based on estimates of the latest catch (in numbers) and on estimates of the percentage of the catch at each age rather than the number of fish caught at each age, as is commonly used in most other sequential population models. The authors considered this approach to have several advantages in that it was possible to take into account the errors in estimating both the total catch and the percentage at age and, also, errors in age estimation can be easily incorporated into the model. Considering

O_t = estimated number of fish caught in year t

C_t = actual number of fish caught in year t

P_{it} = actual percentage of the fish of age i caught in year t

S_{it} = number of fish observed to be of age i in year t

and making the following assumptions

-1) The random variables S_{it} and O_t are independent;

-2) $O_t = C_t \exp(\varepsilon_t)$ where the sampling errors ε_t are independent, normally distributed random variables with mean 0 and variance σ^2 ;

-3) S_{it} is derived from a random sample of the catch which is aged without error;

then the basic model of Fournier and Archibald (1982) is the log-likelihood function

$$\Sigma_{it}\ S_{it} \ln P_{it} - (\Sigma_t\ 0.5\ [\ (\ln O_t - \ln C_t) / \sigma\ ]^2\ ) - n \ln \sigma \quad , \qquad (17)$$

where n = number of years of fishery information. The link between the observations S_{it} and O_t and the exploitation rate is provided by the Baranov catch equation. The reader is referred to Fournier and Archibald (1982) for the details of this coupling process. Two methods were used to estimate fishing mortality rates. One approach used the two factor model of Doubleday (1976), i.e. $\ln F = S_i + F_t$ and a second approach, considered more reliable by the authors, involved estimating fishing mortality from a curve defined by $\ln F_{it} = b_1 j(s) + b_2 j(s)^2 + F_t$, where j is a rescaling factor.

Pope and Shepherd's Separable VPA

Pope and Shepherd (1982) extended the approach of Pope (1974) and Doubleday (1976) in which fishing mortality on particular age-group (i) in a particular year (t) is considered to be a product of the year effect (i.e. fully recruited F) and the age effect (i.e. partial recruitment):

$$f_{it} = F_i S_t \quad .$$

Pope (1974) and Doubleday (1976) attempted to solve for both F_i and S_t using a least squares approach. Their approach was not entirely satisfactory however, since there was no guarantee of a unique solution. The solution was to some extent elucidated by Pope (1977) who noted the close correpondence of the matrix of log catch ratios to a two-way analysis of variances. This observation led Pope (1979) to state that there is insufficient information in the catch-at-age matrix to uniquely define F_i and S_t unless values of these parameters are assumed for the terminal year. Accepting this constraint, Pope and Shepherd (1982) devised a least squares method for estimating F_i and S_t separately and which is commonly referred to as a separable VPA. The reader is referred to Pope and Shepherd (1982) for details of the algorithm development and methods of calculation.

Pope and Shepherd (1982) tested their method on simulated data generated from a separable $F_i S_t$ pattern. Their results indicated that while the separable VPA could find the true solution for "clean" (i.e. without sampling error) data, it was unable to distinguish the true solution from other consistent solutions when random noise with a 10% coefficient of variance was introduced into the test data. From such data, Pope and Shepherd (1982) have concluded that major advantages of the separable VPA is that it eliminates the need for trial and error estimation of appropriate values for input terminal fishing mortality rates. The number of *a priori* parameters is reduced from (t+i-1) to three, i.e. M, terminal fishing mortality rates and terminal selection values. However, external information on stock abundance is required to discriminate the best choice of terminal F. Finally, as with the methods of Doubleday (1976) and Fournier and Archibald (1982), use of the separable VPA model requires that the overall exploitation pattern has remained constant over the range of years included.

Multispecies Sequential Population Models (Legion Analysis)

Method of Pope (1979)

In this model the food of the predator is derived from the potential prey species proportional to each prey biomass weighted by a coefficient which represents the preference of the prey as a food source for the predator. The model is initiated by specifying fishing mortalities for the oldest age of each cohort in the terminal year (T). These mortalities are then used to calculate the average populations in the terminal year since

$$\bar{N}_T = C_T / F_T \quad . \qquad (18)$$

If X = annual ration size, and

Y = % of X which is derived from the specified prey species included in the analysis,

then, the total ration of the j^{th} cohort is

$$\bar{N}_T(j)\ X(j)\ Y(j) / 100 \quad . \qquad (19)$$

Thus the amount of the ration derived from the i^{th} species ages is proportional to

$$\bar{N}_T(i)\ A(i,j)\ X\ W(i) \quad , \qquad (20)$$

where $A(i,j)$ is a coefficient which expresses the preference of predator cohort j for prey cohort i and $W(i)$ is the average weight of the prey cohort. The numbers of each cohort in the terminal year eaten by each other cohort is then calculated using equations 18-20 from which N_T, the population at the beginning of the year, is calculated. For previous years (i.e. non-terminal), population sizes are estimated from a modification of Pope's (1972) cohort analyses, i.e.

$$N_t = (C_t + D_t) \exp(M/2) + N_{t+1} \exp(M) \quad , \qquad (21)$$

where C_t = number of fish caught in year t

D_t = number of fish consumed in year t

M = natural mortality from sources other than predation.

Initial values of $\bar{N}_t(i)$ are estimated as

$$\bar{N}_t(i) = N_{t+1}(i) + 0.5\ C_t(i) \exp(M/2) \quad . \qquad (22)$$

Estimates of $D(i)$ are then calculated thereby allowing $N_t(i)$ to be estimated using equation (21). A new estimate of $\bar{N}_t(i)$ is then derived, as well as a second estimate of $D_t(i)$. This iterative process is continued until

$$\Sigma\ (Z_t^C(i) - Z_t^P(i))^2 \leqslant 0.1 \times 10^{-8}$$

where

$Z_t^C(i)$ = total mortality of the i^{th} cohort in the t^{th} year as estimated in the current iteration

$Z_t^P(i)$ = total mortality of the i^{th} cohort in the t^{th} year as estimated in the previous iteration.

According to Pope (1979), the iterative procedure results in a very rapid convergence.

It is obvious from the above description that the data requirements for such a multispecies approach is rather large. In addition to the requirements for simple cohort analysis (age specific catch and weight matrices for each predator-prey species for each year as well as estimates of terminal F for the most recent year and the oldest age-group), information is required on the annual food consumption, by age-group, for each predator prey fish species as well as non-predation induced natural mortality rates. It would also be useful to have information on the stomach contents for each predator fish according to prey species by age-group including other food for each calendar year. Nevertheless, by making some simple assumptions about diet preference and prey selection, Pope's Legion Analysis can produce a simultaneous assessment of several interacting stocks.

Method of Helgason and Gislason (1979)

One of the constraints of Pope's (1979) Legion Analysis relates to the biomass of food species not considered in the model (i.e. other food). Pope (1979) assumed that a fixed percentage of the food of each predator was obtained from species not considered in the model. This could lead to a serious depletion of prey stocks since predators will continue to take the same fraction of prey species even though these may be depleted by outside predation such as by man. Helgason and Gislason (1979) overcame this difficulty by defining "other food" as one constant biomass specific to each predator. In addition to the data requirements of Pope's Legion Analysis, Helgason and Gislason's model requires information on the annual biomass consumption rate for each predator-prey species as well as a power function relating the species specific food consumption rate to fish body weight. Simultaneous sequential population analyses are then carried out using the version of cohort analysis developed by Gulland (1965).

The reader is referred to Helgason and Gislason (1979) for details on the development of the algorithms and the computational steps necessary to run the model. However, Dekker (1982) has demonstrated that the approximation procedures as proposed by Helgason and Gislason (1979) does not guarantee a unique solution and therefore the results of their model would be difficult to interpret.

Method of Per Sparre (1980) (Legion Analysis)

Per Sparre's Legion Analysis is based on a combination of the features of Pope's (1979) and Helgason and Gislason (1979) models. It includes the partitioning of Z into three parts, i.e.

$$Z = F + M_1 + M_2$$

where F = fishing mortality;

M_1 = natural mortality due to predation;

M_2 = residual natural mortality due to other sources than predation.

For a particular age-group (i) of a species (s) in a given year (t), Legion Analysis may be carried out using the three basic equations

$$N_{t+1}(i+1) = N_t(i) \exp(-Z_i) \quad , \qquad (23)$$

$$C_t(i) = F(i) \bar{N}_t(i) \quad , \qquad (24)$$

$$D_t(i) = M_1(i) \bar{N}(i) \quad . \qquad (25)$$

M_1 is calculated from an algorithm expressing the contribution of the prey species to the total food consumption of the predator in a manner similar to that used by Helgason and Gislason (1979). The suitability index of a particular prey species may be calculated from a purely empirical basis (as in Pope 1979) or from the functional relationship described by Helgason and Gislason (1979). Legion Analysis assumes that total ecosystem biomass is known and remains constant from year to year. The "other food" component is then calculated for each year by subtracting the estimated fish biomass of the interacting species from the ecosystem biomass. This approach is somewhat different than that of Helgason and Gislason (1979) who assumed that the "other food" component remained constant from year to year whereas the ecosystem biomass may vary annually.

Method of Majkowski (1981)

The multispecies model of Majkowski has features which are common to both Pope's (1979) model and that of Helgason and Gislason (1979). Annual food biomass consumption for each predator prey fish species is calculated from species-specific proven functions relating annual food consumption to fish body weight in similar fashion to Helgason and Gislason (1979). Pope's cohort analysis is used for the simultaneous sequential population analyses and the food preference for each predator species is determined from empirical data as in Pope (1979). The "other food" component, which necessitated external estimation in the model of Helgason and Gislason (1979), is estimated internally by the Majkowski model. In addition, the fraction of the annual food consumption by the predator, which is derived from a specific prey species, is estimated internally for each year by Majkowski whereas Pope treats this fraction as being constant from year to year. The reader is referred to Majkowski (1981) for further details on algorithm development.

Comparison of the Various Methods of Sequential Computation Analysis

The deterministic methods of sequential computation presented above (i.e. Fry's, Jones', Gulland's and Pope's) are sometimes referred to in the literature as virtual population analyses (VPA). Only Gulland's method uses virtual populations and his iterative computations are very different from Fry's use of the method. Furthermore, all of these models rely on two algorithms which have a long history in fisheries assessments: the catch equation of Baranov (1918) and the exponential survivor model of Beverton and Holt (1957). Simultaneous solutions of these two equations was first described in the sequential computation model of Ricker (1948). It is interesting to note that neither the Jones nor the Gulland methods represent a significant theoretical innovation over the Ricker method; expression (6) of the Gulland method and expression (8) of the Jones method are identical to expression (5) of the Ricker method. The major difference between these methods is the manner in which the survival rate is estimated. Gulland has used the ratio of the virtual populations in two consecutive years as an initial estimate of S, i.e. Fry's method, whereas Jones utilized the ratio of the catches over the same period. Both the Jones and the Gulland estimates of initial survival are based on the assumption that $V = (F/Z) N$ which is true only if the exploitation ratio (F/Z) does not change as the year-class passes through the fishery. This is usually not the case. The major differences in Pope's model is that instead of iteratively solving the two simultaneous equations, it makes use of a simple approximation to the survival model. The Ricker method and Pope's method have additional advantages of convenience and simplicity and are therefore recommended for universal use. Nevertheless, all of the above methods of traditional VPA suffer from the major disadvantage that independent measures of abundance are necessary if such models are to be used as a basis for catch prognoses. Insofar as retrospective analyses are concerned, all of the traditional models of VPA will produce a fairly accurate description of population trends provided that catches are a substantial proportion of total removals. In addition, these simple models require no assumptions regarding the catchability coefficient and the effects of a violation of the assumption of constant M are likely to be small if the fluctuations in M are random (Ulltang 1977). Systematic variation in M (eg. with time, age, year-class size) will, however, produce significant biases in parameter estimates (Ulltang 1977).

The various methods of separable VPA described above were developed not only to provide a stochastic image of stock dynamics but also to eliminate the arbitrary manner in which input parameters (terminal F, partial recruitment) were defined. However, all of these methods are severely constrained by the assumption that partial recruitment must have remained relatively constant during the period of review and there are few fisheries in which this requirement can be shown to be met. In addition, both Doubleday's and Pope and Shepherd's models have failed to eliminate the need for external information as a basis for properly evaluating output parameters. Furthermore, the conclusion has been made that Pope and Shepherd's method should not be used at all because the parameter estimation procedure is not simultaneous (Megrey 1983). Given the constraint of a constant exploitation pattern, the Fournier and Archiblad (1982) method is perhaps the most useful of the three methods of separable VPA. It has the flexibility of being able to include in its parameter estimation information on abundance and/or fishing effort indices as well as a stock recruitment relationship. Furthermore, it is capable of estimating natural mortality rates and has the option of including as extra information the impact of errors in the ageing procedure.

A review of all the methods of single-species VPA leads to the general conclusion that an appropriate model should (1) possess the flexibility of including information on independent estimates of abundance as well as date-dependent natural mortality, (2) should not be constrained by changes in PR, (3) should include, as output, estimates of current population size and confidence limits on these estimates and (4) should be easy to use such that the underlying assumptions can be readily evaluated. The Doubleday Survivors Method fits all of these requirements and is particularly appropriate when prospection is the major objective of population estimation. Although the Survivors Method is usually used in conjunction with age-structured survey data, which is often highly variable, other indices of abundance such as commercial catch rates can be used either separately or combined with survey data. Thus, the only arbitrary decision which has to be made is the choice of years (i.e. calibration block) to be included in the least squares estimation of the age-specific catchability coefficients. Ideally, the calibration block should be insensitive to the input parameters in the terminal year but this is not always the case. The Collie and Sissenwine Method overcomes the problem by calculating q directly from the survey and catch data. Thus, when used together, the Survivors Method and that of Collie and Sissenwine should provide a powerful objective tool for estimating terminal population sizes from sequential computation analysis.

The various models of multispecies sequential population analysis (Legion Analysis) are all based on similar algorithms for the solution of simultaneous VPAs of more than one fish stock and the attractive feature of these models is that natural mortality is not a fixed input parameter but is determined by inter and intraspecific predation. The additional information required refers to a suitability index of each prey age-group for each predator age-group, and the rate of total food consumption. As shown by Sparre (1980), empirical stomach content data can be converted into an estimate of the suitability index by means of an iterative procedure. Given such prey composition data, Legion Analysis can provide useful insights into the importance of predator-prey interactions in controlling populations (Sparre 1980). Their major disadvantage is that empirical estimates of average annual food composition data for a population by age-groups makes very large demands on the sampling scheme involved. Given the almost complete lack of such data for most predators, it is not surprising that there is a large discrepancy between the theoretical advances in this field in relation to the practical application of the various models. Furthermore, the use of such models for catch projection would probably involve complicated fine-tunning procedures since the models output should adequately reflect not only trends in the predator's abundance but also those of the various prey species. Thus, in the near future, the major significance of such multispecies models will be as a means of identifying the major departures of single-species assessments from the assumptions of the independent non-interactive stocks in an equilibrium environment.

Natural Mortality Estimation

Estimates of historical stock size from sequential population analyses essentially represent the accrued deaths in the population due to fishing effort and natural causes. Consequently, the level of natural mortality rate (M) used in such analyses will have a significant impact on the absolute estimate of population size as well as stock trajectory. It is therefore important that M be well-estimated and a review of the various methods of estimating M is appropriate.

Catch Curve Analysis

The statistical analyses of relative age distribution (i.e. catch curves) are useful to derive estimates of natural mortality in several ways. For unexploited populations, the slope of a catch curve is a direct estimate of M. In the case of exploited populations, it is necessary to have catch curves for several years in addition to fishing effort data over the period of years of contributing fishing mortality. Then, an estimate of M is derived by regressing values of Z against the average fishing effort appropriate to each catch curve. The intercept of such a regression line is then an estimate of M.

It should be pointed out that the slope of a catch curve will provide an unbiased estimate of mortality only under the following assumptions: a) there is no temporal change in survival rates; b) recruitment is uniform for each age; c) the sample is taken randomly and accurately aged; d) vulnerability (catchability) and M are independent of age. In the simplest case, it is assumed that the age distribution is geometric, i.e. constant recruitment, uniform age-specific survival rates. In this form, the catch curve is derived from a regression of $\log_e N_t$ against age. The more usual assertion, however, is that a population at any point in time is the result of random recruitment and mortality processes. In this case, the estimate of mortality is more appropriately obtained from the slope of the regression of $\log_e N_t - (1 / (N_t+1))$ on age (Chapman and Robson 1969).

In most exploited populations, it is frequently observed that both fishing mortality and recruitment vary systematically over time. This will often introduce a bias into the catch curve which can only be removed when independent data in the form of recruitment estimates (R_t) and fishing effort (f_t) are available. If the temporal change in fishing effort and/or recruitment (when cross-sectional age-compositions are used) is systematic, a regression of $\log_e f_t$ (or $\log_e R_t$) on time will provide an estimate of the instantaneous rate of change in these parameters which can then be subtracted from Z^1, the original slope estimate. Generally, however, recruitment and fishing effort changes over time are not systematic but highly variable. If recruitment is variable over time, a regression of $\log_e N_t - \log_e R_t - (1/(N_t+1))$ on age will provide an accurate estimate of Z; similarly, in the case of variable effort, a regression of $\log_e N_t - \log_e f_t - (1/(N_t+1))$ on age will adjust the slope to its true value.

Age Composition Analyses

Mortality rates may be extracted from the age frequency of a catch sample by the methods of Heincke (1913), Chapman and Robson (1960) and Robson and Chapman (1961). Application of these techniques requires that the age distribution be geometric, i.e. constant recruitment and uniform age-specific recruitment rates, or that independent data are available which would allow adjustment for these biases. Heincke's formula weights successive ages as their abundance and is particularly useful when age determinations of the older fish are unreliable (Ricker 1975). Heincke's estimate of survival (S_H) is defined as:

$$S_H = (N - N_0) / N$$

where N = number in the sample
N_0 = number of animals in the youngest (fully recruited) age-group.

According to Robson and Chapman (1961), an unbiased estimate of survival from a geometric age distribution may be obtained from

$$S_c = T / (N + t - 1)$$

where $T = N_1 + 2N_2 + 3N_3 \ldots$, and N = number in the sample. The Robson and Chapman estimation is very robust and is not as sensitive to changes in recruitment as Heincke's method. It does, however, place substantial weight on the older age-groups and may, therefore, be significantly affected by ageing errors.

Catch and effort Analyses

Paloheimo (1961) simplified an interactive procedure of Beverton and Holt (1957) by referring the calculation of Z from catch/effort data to time intervals that include half of one year and half of the next, under the assumption that both natural and fishing mortality occur simultaneously throughout the year. In this case, for each cohort (or combination of cohorts),

$$\tilde{Z} = \log_e [(C/f_1) / (C/f_2)] \quad .$$

$\tilde{Z}$ thus approximates the mean instantaneous mortality rates of two adjacent years. Paloheimo (1961) also defined

$$\bar{f} = (f_1 + f_2) / 2$$

Then a regression of $\bar{Z}$ on $\bar{f}$ is computed thereby providing an intercept estimate of *M*.

Silliman (1943) proposed a method of estimating *M* for those exploited populations of fish whose history includes two periods of fairly stable (although different) levels of fishing effort, each persisting long enough to give a reliable estimate of *Z*. Assuming that *M* is the same in both periods, we have

$$Z_1 = F_1 + M$$

and

$$Z_2 = F_2 + M$$

but

$$(F_1 / F_2) = (f_1 / f_2)$$

Thus, since both Z_1 and Z_2 are known, the above equations can be solved directly for F_1, F_2 and *M*.

A variety of authors have used a regression of fishing mortality (*F*) from cohort analyses against fishing effort as an intercept estimate of *M*. This method, however, is not valid since the effect of *M* on population size (and therefore *F*) as one proceeds from the oldest to the youngest age in cohort analyses is not constant but increases in magnitude. Therefore, with a given set of historical catch-at-age data, changes in *M* will not only affect the intercept but also the slope of a plot of *F* (as derived from cohort analyses) on effort.

Estimating *M* from Life History Parameters

The association between natural mortality and life history parameters of fish was first studied by Beverton and Holt (1959) who related the Von Bertlanffy growth parameters (growth rate, asymptotic length and longevity) to estimates of *M* given in the literature. Pauly (1980) extended this approach by developing an empirical multiple regression model which expressed *M* as a function of asymptotic size, growth rate and mean environmental temperatures for 175 fish stocks. Myers and Doyle (1983) developed a life history model to predict *M* based on the hypothesis that there is a trade off between reproduction, growth and survival and that fitness maximizes lifetime fecundity. The basic data required are (1) growth rate over the natural life span, (2) life-span fecundity data, (3) age and size at maturity, and (4) energy content of the eggs. Estimates of *M* by the model of Myers and Doyle (1983) correspond quite well with estimates of *M* by independent means. Sensitivity analyses indicated that the model was very sensitive to the surplus energy parameters but was not very sensitive to the other life history parameters. The model, however, cannot be used for short-lived species (less than 4 years) which spawn only during a short time period of the year.

Roff (1984) expanded the approach of Beverton and Holt (1959) by developing an algorithm which predicted *M* from the Von Bertlanffy parameters L_∞ (asymptotic length), *k* (rate of growth) and L_T (length at maturity). Estimates of *M* by the method were good in agreement with values published in the literature for a variety of species. Because of the sensitivity of *M* to small variations in *k*, Roff (1984) cautioned against the general use of his model.

References

Allen, K.R. 1966. Some methods for estimating exploited populations. J. Fish. Res. Bd. Can. 23: 1553-1574.

Baranov, F.I. 1918. On the question of the biological basis of fisheries. Nauchnyi issledovatelskii ikhtiologicheskii Institut, Izvestiya 1(1): 81-128

Beverton, R.J.H., and S.J. Holt. 1957. On the dynamics of exploited fish populations. U.K. Mm. Agric. Fish. Invest. (Ser. 2) 19: 533 p.

Beverton, R.J.H., and S.J. Holt. 1959. A review of the lifespan and mortality rates in nature and their relations to growth and physiological characteristics. CIBA Found. Colloq. Ageing 5: 142-147.

Chapman, D.G., and D.S. Robson. 1960. The analysis of a catch curve. Biometrics 16: 354-369.

Collie, J.S. and M.P. Sissenwine. 1983. Estimating population size from relative abundance data measured with error. Can. J. Aquat. Sci. Vol. 40(11): 1871-1879.

Dekker, W. 1982. Theoretical aspects of the multispecies vertical population analysis. Int. Comm. Explor. Mer. (ICES) CM 1982/G:32

DeLury, D.B. 1947. On the estimation of biological populations. Biometrics. 3:145-167.

Derzhavin, A.N. 1922. The stellate sturgeon (*Acipenser stellatus*, Pallas) a biological sketch. Biull. Bakinskoi Ikhttiologicheskoi Stantsii 1: 1-393.

Doubleday, W.G. 1976. A least squares approach to analyzing catch-at-age data. Int. Comm. Northw. Atl. Fish. Res. Bull. 12: 69-81.

Doubleday, W.G. 1981. A method of estimating the abundance of survivors of an exploited fish population using commercial catch-at-age and research vessel abundance indices. Can. J. Fish. Aquat. Sci. Spec. Pub. 58: 164-178

Fournier, D., and C.P. Archibald. 1982. A general theory for analyzing catch-at-age data. Can. J. Fish. Aq. Sci. Vol. 39(8): 1195-1207.

Fry, F.E.J. 1949. Statistics of a lake trout fishery. Biometrics 5: 27-67.

Fry, F.E.J. 1957. Assessments of mortality by use of virtual populations. Proc. of the Joint Sci. Meeting of ICNAF, ICES, and FAO on fishing effort, the effects of fishing on resources, and the selectivity of fishing gear. Contribution No. 15 (mimeo).

Gulland, J.A. 1965. Estimation of mortality rates. Annex to Arctic Fisheries Working Group Report. ICES, C. M. 1965, Doc. No. 3 (mimeo).

Heincke, F. 1913. Investigations on the plaice. General Report I. The plaice fishery and protective measures. Preliminary brief summary of the most important points of the report. ICES Rapp. Proc.-Verb., 16: 1-67.

Helgason, T., and G. Gislason. 1979. VPA-analysis with species interaction due to predation. Int. Comm. Explor. Mer. (ICES) CM 1979/G:52.

Jones, R. 1961. The assessment of the long-term effects of changes in gear selectivity and fishing effort. Mar. Res. No. 2, p. 1-19.

Jones, R. 1968. Appendix to the report of the Northwestern Working Group, ICES C.M. 1968, Doc. No. 20 (mimeo).

Majkowski, J. 1981. Application of multispecies approach for assessing the population abundance and age structure of fish stocks. Can. J. Fish. Aquat. Sci. Vol. 38(4): 424-431.

Megrey, B.A. 1983. Review and comparison of three methods of cohort analysis. Nat. Mar. Fish. Ser., Northw. and Alaska Fish. Centre, Prog. Rept. 83-12. 26 p.

Murphy, G.I. 1965. A solution of the catch equation. J. Fish. Res. Bd. Canada 22(1): 191-202.

Myers, R.A., and R.W. Doyle. 1983. Predicting natural mortality rates and reproduction-mortality trade-offs from fish life history data. Can. J. Fish. Aquat. Sci. 40: 612-620.

Paloheimo, J.E. 1961. Studies on estimation of Mortalities. I. Comparison of a method described by Beverton and Holt and a new linear formula. J. Fish. Res. Bd. Can. 18: 645-662.

Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters and mean environmental temperatures in 175 fish stocks. J. Cons. Int. Explor. Mer. 39: 175-192.

Pope, J.C. 1972. An investigation of the accuracy of virtual population analyses. Intern. Comm. Northw. Atl. Fish. Res. Bull. No. 9: 65-74.

Pope, J.C. 1974. A possible alternative method to vertical population analysis for the calculation of fishing mortality from catch-at-age data. Int. Comm. Northw. Atl. Fish., Res. Doc. No. 20, Ser. No. 3166 (mimeo).

Pope, J.C. 1979. A modified cohort analysis in which constant natural mortality is replaced by estimates of predation levels. Int. Comm. Explor. Mer. (ICES) CM1979/H: 16.

Pope, J.C. 1980. Estimating mortality rates in fish populations. Trans. Amer. Fish. Soc. 109: 378-386.

Pope, J. G., and J.G. Shepherd. 1982. A simple method for the consistent interpretation of catch-at-age data. J. Cons. Explor. Mer. 40: 176-184.

Ricker, W.E. 1948. Methods of estimating vital statistics of fish populations. Indiana Univ. Publ. Sci. Ser. No. 15, 101 p.

Ricker, W.E. 1958. Handbook of computations for biological statistics of fish populations. Fish. Res. Bd. Can. Bull. No. 119, 300 p.

Robson, D.S., and D.G. Chapman. 1961. Catch curves and mortality rates. Trans. Amer. Fish. Soc. 93: 215-226.

Roff, D.A. 1984. The evolution of life history parameters in teleosts. Can. J. Aquat. Sci. Vol. 41: 989-1000.

Silliman, R.P. 1943. Studies on the pacific pulchard or sardine (*Sardinops caerulae*). A method of computing mortalities and replacements. U.S. Fish. Wildl. Serv. Spec. Sci.: Rep 24: 10 p.

Sparre, P. 1980. A good function in fisheries (legion analyses). ICES CM 1980/G:40.

Ulltang, O. 1977. Sources of errors in and limitations of Virtual Population Analysis (Cohort Analysis). J. Cons. Int. Explor. Mer. 37: 249-260.

Catch Projections

by **D. Gascon**

Institut Maurice Lamontagne,
Mont-Joli, Québec
G5H 3Z4

Introduction.

Catch projections are normally the last step in the stock assessment process. They represent the core of the information provided to the fisheries managers by the assessment biologists and they serve to establish the management policies - usually under the form of TAC's - given some pre-agreed management objectives (e.g. $F_{0.1}$, $F_{\max}$, etc.).

As far as assessment methods go, catch projections are by far the least controversial and are commonly perceived as straightforward extensions of conventional analytical techniques. However, the projections are subject to errors resulting from uncertainties in the input parameters: population numbers - or terminal fishing mortality -, catch at age, mean weight at age, recruitment, etc. The errors associated with the catch projections are usually unknown although Rivard (1981) and Pope and Gray (1983) have made a strong case for identifying the source and size of these errors to maintain the uncertainties about projection estimates within acceptable bounds.

Catch projections are ultimately used in the establishment of TAC's. In this paper, I shall discuss only those approaches based on age-structured demographic models, while the general production methods are discussed in Rivard (1988). I will describe first the computational methods (with the data requirements and underlying assumption) to be followed by a description of the sources of error and their methods of estimation.

Computational Methods.

Two approaches for formulating TAC's based on catch projections exist: targeted TAC's and status-quo TAC's (Pope 1983). A targeted TAC is the catch corresponding to a specific level of removal, either in terms of quota (e.g. a pre-agreed TAC) or in terms of fishing mortalities (e.g. $F_{0.1}$, $F_{\max}$, etc.). Status-quo TAC's are formulated in terms of relative fishing mortalities such that

$$F_{t+\alpha} = k\, F_{t-\beta} \quad , \qquad (1)$$

where k is a proportionality factor and α and β are time intervals (usually 1 or 2 years).

Targeted TAC is the standard procedure used by CAFSAC with the target set at the $F_{0.1}$ level while status-quo TAC's are commonly used by the International Council for the Exploration of the Sea (ICES). The computational methods are exactly the same for both, although short-cut techniques exist to determine status-quo TAC's. These will be dealt with briefly in a separate section as they are not yet widely accepted.

Data requirements.

Current analytical methods merely project population numbers in time according to a simple exponential decline model, and estimate future catches using a catch equation and an age-specific allocation of fishing mortalities (partial recruitment). Hence two types of input parameters are required. The parameters describing the population structure for which only initial estimates at time t are required, while the remaining parameters require estimates for each of the projected years (usually, year t+1 and t+2). Note that τ and ι refer respectively to the intervals t, t+1 and i, i+1 and that the dot (•) represents the summation over all the index.

The input parameters required by the current projection techniques are as follows:

$N_{i,t}$	Age specific population numbers in the initial year t; *and/or*
$F_{\iota,\tau}$	Age specific instantaneous rates of fishing mortalities in initial year;
$w_{i+0.5;\, t,\, t+1,\, t+2}$	Yearly estimates of age-specific mean weights (taken as mid-year values);
$r_{i;\, t,\, t+1,\, t+2}$	Yearly estimates of partial recruitment;
$M_{\iota;\, \tau,\, \tau+1,\, \tau+2}$	Yearly estimates of age specific instantaneous rate of natural mortality;
$N_{b;\, t+1,\, t+2}$	Yearly estimates of recruitment in numbers at the first age (b) recruited to the fishery;

Target parameters.

$F_{\bullet;\, \tau+1,\, \tau+2}$	Target fishing mortalities in each year of projection;
	or
$CB_{\bullet;\, t+1,\, t+2}$	Target catch biomass in each year of projection;
	or
	A combination of both, such that all years in the projection are represented by one target each.

Methods to estimate the demographic parameters ($N_{i,t}$, $C_{i,t}$, $F_{\iota,\tau}$) are given in Winters (1988). These are normally from the last year for which sampling data are available. Because some of the input parameters cannot be estimated, some assumptions must be made about their nature. For instance, natural mortality ($M_{\iota,\tau}$) is assumed to be constant for all ages and in all years, and equal to some pre-agreed value. The values chosen for recruitment are usually the mean of historical values, although from time to time, other *ad hoc* estimates are used (e.g. mean from a range of years) in response to misgivings about the data in the most recent year. The choice of method seems to stem from an attempt to compromise between errors due to time trends and errors due to random fluctuations. For those parameters (e.g. mean weights and partial recruitment) where random year to year fluctuations are small (or at least thought to be so) but where time trends are likely to occur (density dependent growth, technological improvements, etc.), the most recent estimates are used in order to minimize differences due to directional changes. On the other hand, where random fluctuations are presumed to be more important than time trend (e.g. for $N_{b,t}$), historical averaging is used. These are somewhat arbitrarily accepted conventions which have proved useful with past experience.

Recruitment is normally projected as the average of historically observed recruitments. A geometric mean is used here, owing to the log-normal distribution of the observations (usually). It is also possible to use other types of statistical estimators based on the time series (see Anom. 1984). However, recruitment indices from research surveys are sometimes available for short-term (1 or 2 years) projections. These indices can be used directly. However, because of uncertainties about relative catchabilities, partial recruitment, etc., they are best used as relative indices in a predictive relationship appropriately time-lagged with population number estimates at the first age (b) recruited to the fishery. The index is then used to predict numbers at age b in future years of cohorts for which the recruitment index is available. As most stocks recruit to the fisheries at an age ranging from 2 to 4, these indices cannot usually predict recruitment for more than 1 or 2 years.

Targeted TAC's.

Assumptions.

The model for catch projection is based on the following 2 equations. The first one assumes an exponential population decline:

$$N_{i,t+1} = N_{i,t}\, e^{-Z_{\iota,\tau}} \quad , \qquad (2)$$

and the second one is based on the Baranov's catch equation

$$C_{i,t} = F_{\iota,\tau}\, N_{i,t} \,(\, 1 - e^{-Z_{\iota,\tau}} \,)\, /\, Z_{\iota,\tau} \quad , \qquad (3)$$

where

$$Z_{\iota,\tau} = F_{\iota,\tau} + M_{\iota,\tau}$$

and

$$F_{\iota,\tau} = r_{i,t}\, F_{\bullet,\tau} \qquad \text{(Ricker, 1975)}$$

The assumptions of this model are identical to those of sequential population analysis, which is based on the same equations. Fishing and natural mortalities should occur concurrently and be distributed uniformly over the year.

Computations.

Catch projections normally follow some sort of sequential population analysis. Therefore $C_{i,t}$, $N_{i,t}$ and $F_{\iota,\tau}$ are known, although projections could conceivably be produced following other types of population estimations (e.g. groundfish surveys, etc.), in which case $F_{\iota,\tau}$ will not be known. Consequently, software packages usually compute fishing mortalities in year *t* using the Newton Raphson solution of (3). The computational steps for the method are given in Appendix I.

Population numbers in subsequent years are sequentially computed by solving (2) for the target values of *F*. The corresponding catch numbers are obtained from (3). Catch biomass is then simply obtained as the sum of the product of age specific mid-year weights and catch numbers at age.

TAC's are normally set for the following year, that is they are set in the year t+1 for year t+2 using data from year t. Thus, the target for the first year of projections (and possibly other subsequent years) is often a TAC rather than a fishing mortality level. In such cases, an arbitrary but realistic value of F is used as a seed in an iterative process. Equations (2) and (3) are then solved iteratively for F, adjusting the values of fishing mortalities in the successive iterations by the factor:

target catch biomass / calculated catch biomass

until the differences between the two values is less than a predetermined level.

Status—quo TAC's.

Assumptions.

Pope (1983) has proposed two short cut techniques to estimate status-quo TAC's in the special case where the proportionality factor between F values is 1. Somewhat more complicated formulations could be derived in the case where this factor differs from 1. Since these methods do not require an *a priori* knowledge of fishing mortalities, or population numbers, they bypass a great deal of the usual assessment process, and inasmuch are more like a comprehensive assessment technique than simple projection methods. I shall review them only very briefly here as they have not yet received acceptance and because status-quo TAC's are not customarily used by CAFSAC. Readers are referred to the original paper by Pope (1983) for a detailed description of these methods. Both methods allow one to estimate TAC's and their variance such that $F_{t+2} = F_t$. They are derived from the ANOVA generalization of separable virtual population analysis (SVPA, see Pope and Shepherd 1982):

$$\ln (C_{i+1,t+1} / C_{i,t}) = \ln (F_{T+1} / F_T) + \ln (S_{i+1} / S_i) + 0.444 \, F_{T+1} \, S_{i+1} + 0.556 \, F_T \, S_i - M \quad , \qquad (4)$$

where S is the partial recruitment vector of SVPA. The assumptions of this technique are those of SVPA, which are identical to those of standard sequential population analysis and projection methods. SVPA adds a further assumption, that the fishing mortality interpretations of catch at age is the result of an age effect (partial recruitment) and a year effect, the fully recruited fishing mortalities (Pope and Shepherd 1982) such that:

$$F_{\iota,T} = S_{i,\bullet} \, F_{\bullet,T} \quad . \qquad (5)$$

Since this is generally assumed in standard catch projections, equation (5) does not impose additional constraints on the methods.

Leapfrog TAC's.

By subtracting $\ln(C_{i+1,t} / C_{i,t-1})$ from $\ln(C_{i+1,t+1} / C_{i,t})$, one can express catch at time t+1 as a function of catches and ratios of F values at time t-1 and t .

$$C_{i+1,t+1} = \frac{C_{i+1,t}\, C_{i,t}}{C_{i,t-1}} \left[\frac{F_{T-1}}{F_T} \right]^2 \quad . \qquad (6)$$

Replacing the ratio of fishing mortalities by the ratio of effort (*f*) values (assuming $F = qf$), and with recruitment estimates for year *t* and *t*+1, one can then estimate the TAC's at time *t*+2 knowing only the catch and mean weights at age, the effort values, and recruitment for 2 years.

ANOVA TAC's.

The derivation of ANOVA TAC's is somewhat more involved. It is derived from the analogy of SVPA with ANOVA. Separable virtual population analysis produces an interpretation of the catch at age matrix decomposed into essentially pure year and age effects (equation 4). By algebraic manipulations and cancellation of factors, it is then possible to estimate the year effects (α_t) of SVPA as a function of fishing mortalities:

$$\alpha_t = 0.5 \ln \left(F_{\bullet,T+1} \,/\, F_{\bullet,T}^{\,3} \right) + \text{constant} \quad . \qquad (7)$$

The ratio of fishing mortalities can be replaced, as in leapfrog TAC, by the ratios of effort values. It is then possible to relate the α_t's of SVPA to the function of effort and predict α_t's of future years for given effort ratios. Catch at age for future years is then obtained from:

$$C_{i+1,t+1} = C_{i,t}\, e^{\left(\alpha_t + \beta_i + \mu + \sigma^2/2 \right)} \quad , \qquad (8)$$

where β_i, μ and σ^2 are respectively the age effects, mean and variance of SVPA. A recruitment index is also needed to estimate $C_{b,t+1}$.

This formulation requires, as does the leapfrog TAC, a linear relationship between fishing mortality and fishing effort (for a discussion of this relationship, see Beddington 1979). Further, the 3rd and 4th terms of (4) can approximately cancel only for F=1.0 and smaller or larger values of fishing mortality result in departures from this approximation. Departures become important for values of $F \leq 0.3$ and $F \geq 1.7$. A complete separable virtual population analysis is required in order to derive the relationship between the α_t's and effort function, although the choice of terminal *F* and *S* and of natural mortality has little influence on the relationship and the TAC's. ANOVA TAC produces projection estimates with smaller confidence intervals than those produced by Leapfrog TAC (Pope 1983).

Precision of Projection Estimates.

The precision of catch projections is dependent on the uncertainties about the input parameters, although in practice, since deterministic models are used, little attention is paid to this problem. Consequently, the projections are often provided in very definite terms with an appearance of precision. Several sources of errors can be identified:

Model errors. This is an unavoidable consequence stemming from the need for simplification and generalization in model building. I shall not consider this source of error further.

Sampling errors. Data on catch at age, mean weights at age, and in some cases, numbers at age and partial recruitment are obtained through sampling and the variances associated with these parameters can be computed (e.g. Gavaris and Gavaris 1983; Smith and Maguire 1983).

Errors due to uncertainties about projected parameters. As stated in the section on data requirements, an assumption of constancy is normally formulated for mean weights at age, partial recruitment, and recruitment. If systematic departures from this assumption occur, the accuracy of catch projections will decline with time. This source of error is impossible to assess quantitatively *a priori.*

Errors due to uncertainties about level of terminal fishing mortality. The determination of an appropriate level of terminal fishing mortality, and hence of stock size in the first year of projections, is a highly recursive and subjective process which does not lend itself to the estimation of variances (Pope 1983). A possible mean of assessing the relative importance of this source of error is to make a retrospective examination of how well - or bad - past assessments were at estimating terminal F. Because of the converging properties of sequential population analysis (Pope 1972), the estimates for the population parameters in past years become less dependent upon the input terminal fishing mortality as the time elapsed increases. Thus, estimates of fishing mortality by year from SPA will be relatively freed (nearly entirely so after 3 or 4 years) of the errors associated with the determination of F_t. It is then possible to obtain a rough estimate of the size of these errors by comparing the values of fishery mortalities selected in successive years to those generated in a recent SPA, using the terminal fishing mortality deemed the most appropriate at the time. Rivard (1981) estimated the relative error in stock size for the southern Gulf of St. Lawrence cod stock at 24% based on three different assessments (i.e. differences due to the interpretation of the data by various authors in successive years). Further, by making the comparison between the estimates made in three successive years for the stock size of 1977, he came to a relative error ranging from 1% to 50%, with a mean of 24% for various stocks of groundfish in the Northwest Atlantic. I repeated the same exercise over a longer time interval, using the data for all stocks of Gadidae assessed by CAFSAC, NAFO and the North Sea groundfish working group of the CIEM. The results are presented in Figure 1. In the assessments used to generate Figure 1, the estimates of fishing mortalities in the assessment year diverged from 12% to 112% from the most recent (and presumably more accurate) estimate for that year from SPA. The average relative errors associated with the determination of F_t (using the technique described above) ranged from 24% to 56% on the 6 years for which data are available. Because it appears that the fishing mortalities are systematically underestimated (i.e. stock size is overestimated), the relative errors in stock size would be in the order of 50% to 100%, a somewhat greater range than reported by Rivard (1981).

Estimation of Variances

Rivard (1981) provides formulas to compute conditional variances (assuming M is constant) for stock size and catch biomass given the variances in initial stock size, recruitment and in mean weights at age. These formulations assume also that there are no time trends, and that only random-sampling errors occur on projected parameters. The variance of stock size at time t is then :

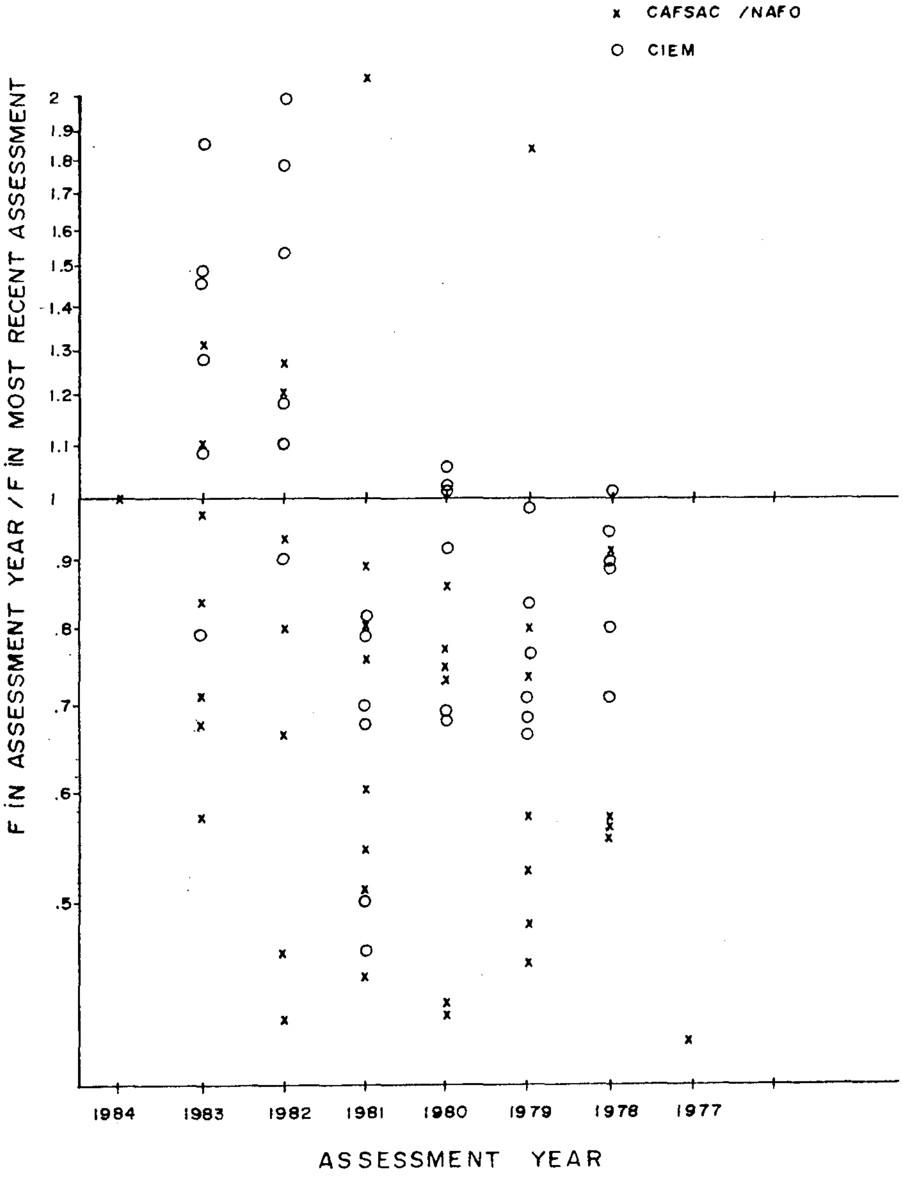


FIGURE 1. Examination of the accuracy of past determinations of fishing mortality at year by comparing with the most recent SPA estimates of fishing mortality in those years for various stocks of Gadidae assessed by CAFSAC, NAFO and ICES.

$$\mathrm{VAR}[\, N_{\bullet,t} \,] = \sum_{i=b}^{\max} \mathrm{VAR}[\, N_{i,t} \,] \, (\, \partial N_{\bullet,t} \, / \, \partial N_{i,t} \,) + \sum_{T=t+1,t+2} \mathrm{VAR}[\, R_T \,] \, (\, \partial N_{\bullet,t} \, / \, \partial R_T \,) \quad , \qquad (9)$$

and the variance in yield (catch biomass or TAC):

$$\mathrm{VAR}[\, Y_{\bullet,t} \,] = \sum_{i=b}^{\max} \mathrm{VAR}[\, N_{i,t} \,] \, (\, \partial Y_{\bullet,t} \, / \, \partial N_{i,t} \,) + \sum_{T=t+1,t+2} \mathrm{VAR}[\, R_T \,] \, (\, \partial Y_{\bullet,t} \, / \, \partial R_T \,) + \sum_{i=b}^{\max} \mathrm{VAR}[\, W_i \,] \, (\, \partial Y_{\bullet,t} \, / \, \partial W_i \,) \quad , \qquad (10)$$

Pope and Garrod (1975) provide similar formulations relating the variance of the TAC to the variances of initial stock size, recruitment, weights at age and fishing mortality (the variance of F being derived from the variance of effort assuming $F = qf$). While it is relatively simple to obtain the variance of most parameters (from sampling and/or observed historical data), variances for $F_{\bullet,t}$ and $N_{\bullet,t}$ are difficult to estimate and the use of these formulas is generally not possible in the context of annual assessments.

Sensitivity Analysis

Although it may not yet be possible to estimate accurately the magnitude of the relative errors associated with catch projections, the relative influence of each source of error can be assessed by sensitivity analysis (Rivard 1982). Sensitivity is the relative response of a model to small perturbations of input parameters; sensitivity coefficients smaller than 1 indicate a relative insensitivity of the model to perturbations in that parameter, while sensitivity greater than 1 indicate the opposite. Rivard (1981) performed a sensitivity analysis on 3-year catch projections using the cod stock in the southern Gulf of St. Lawrence (4TVn) as an example. The projections for this stock were rather insensitive to errors in natural mortality, partial recruitment, and in recruitment while they were quite sensitive to mean weights at age and initial stock size. Pope and Gray (1983) used Monte Carlo simulations in an attempt to partition the relative errors in catch projections among potential sources (fishing effort, recruitment, and catch at age) for three North Sea groundfish stocks and thus identify the influential variables.

The relative importance of these sources appear to be stock-dependent and it is difficult to draw any general conclusions from the above-mentioned simulations. The precision of projections for stocks in which recruits contribute significantly to the catch will be more affected by the precision of recruitment estimates (e.g. North Sea stocks, see Pope and Gray 1983), while for those in which recruits contribute little, other variables will be more influential (e.g. 4TVn cod, see Rivard 1981). Furthermore, because projected values of cohorts size are entirely determined by initial size estimates for that cohort, short term projections will be mostly affected by the uncertainties in initial population size, while long term projections are mostly affected by uncertainties in recruitment.

Common wisdom usually states that short term projections are more accurate than long term ones, and it is customary to add a cautionary note to long term projections, stating the possible large uncertainties associated with them. This may be more of a myth than a reality. It depends on the relative precision of initial population size estimates and recruitment at year estimates. Monte Carlo simulations of 10-year catch projections that I made using data from the northern Gulf of St. Lawrence cod stock (assuming variable recruitment and uncertain population size) produced a decline in the coefficient of variation of TAC's from 26% in the first year of projection to 20% in the tenth year. This result obviously depends on the choice of the degree of variability for the input parameters. The variability in recruitment was generated from the observed historical variations. The coefficient of variation for fishing mortality in year t was 28.5%, a value in line with historical data (see Errors due to uncertainties about the level of terminal fishing mortality). The values of F were generated from a log-normal distribution such that $0.5\ F_a \leq F \leq 2\ F_a$ (where $F_a = 0.275$ is the value of fishing mortality selected at assessment time). Clearly then, current short term projections are not necessarily more precise than long term ones. Random variations on other parameters will not affect the relative precision of short and long term projections (although they produce additional error). However, if temporal changes occur, the accuracy of projections will decline as they become more remote in time.

Conclusions

Because catch projections depend so much on precision of the previous steps in assessment, and that little latitude is left at this stage, it is difficult to draw general conclusions without involving most of the assessment process.

Clearly, a great many uncertainties exist around the projection estimates as they are now formulated. Most of these are inherited from uncertainties in the previous steps of the assessment; improvements in these steps would be immediately reflected in the projections. More attention should be given to the definition of the degree of precision of estimates and the relative importance of the sources of error. This could serve a dual purpose. First, the influential variables for each stock could be identified and the application of corrective measures could be aimed at those variables which would most improve the estimates in order to bring them within acceptable bounds. Second, it would also allow the development of guidelines as to when changes in regulation are required (i.e. implemented TAC's) following changes in projection estimates. Current procedures in use by CAFSAC are based on the rule of thumb that a change of less than 10% in projection does not warrant a change in TAC.

Although it is conceivable that future improvements in catch projections may take place in the form of more sophisticated models including growth functions, recruitment functions, etc., the present degree of imprecision in the data seems to preclude this kind of development in the immediate future. Rather, current trends seem to indicate a move toward simpler models with a greater degree of reliability (Pope 1983; Anom. 1984).

References

Anom. 1984. Report on the working group on methods of fish stock assessments. International Council for the Exploration of the Sea. C.M. 1984/ Assess: 19.

Beddington, J.R. 1979. Harvesting and population dynamics, in R.M. Anderson, B.D. Turner and L.R. Taylor (ed.). Population dynamics. The 20th Symposium of the British Ecological Society. London, 1978. Blackwell Scientific Publications.

Gavaris, S. and C. Gavaris. 1983. Estimation of catch at age and its variance for groundfish stocks in the Newfoundland Region, p. 178-182. *In* W.G. Doubleday and D. Rivard (ed.). Sampling commercial catches of marine fish and invertebrates. Can. Spec. Publ. Fish. Aquat. Sci. 66.

Pope, J.G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. ICNAF Res. Bull. 9: 65-74.

Pope, J.G. 1983. Analogies to the status-quo TAC's: their nature and variance. p.99-113. *In* W.G. Doubleday and D. Rivard (ed.). Sampling commercial catches of marine fish and invertebrates. Can. Spec. Publ. Fish. Aquat. Sci. 66.

Pope, J.G. and D.J. Garrod. 1975. Sources of error in catch and effort quota regulations with particular reference to variations in the catchability coefficient. ICNAF Res. Bull. 11: 17-30.

Pope, J.G. and D. Gray. 1983: An investigation of the relationship between the precision of assessment data and the precision of total allowable catches. p. 151-157. *In* W.G. Doubleday and D. Rivard (ed.). Sampling commercial catches of marine fish and invertebrates. Can. Spec. Publ. Fish. Aquat. Sci. 66.

Pope, J.G. and J.G. Shepherd. 1982. A simple method for the consistent interpretation of catch at age data. J. Cons. Int. Explor. Mer. 40: 176-184.

Ricker, W.E. 1975. Computations and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Canada No 191, 382 p.

Rivard, D. 1981. Catch projections and their relation to sampling error of research surveys. p. 93-109. *In* W.G. Doubleday and D. Rivard (ed.). Bottom trawl surveys. Can. Spec. Publ. Fish. Aquat. Sci. 58.

Rivard, D. 1982. APL programs for stock assessment (revised). Can. Tech. Rep. Fish. Aquat. Sci. 1041, 146 p.

Rivard, D. 1988. Biological production and production models. *In* Rivard D. [Ed], Collected papers on assessment methods. CAFSAC Res. Doc. 88/61: 44-110.

Smith, S.J. and J.J. Maguire. 1983. Estimating the variance of length composition samples. p. 165-170. *In* W.G. Doubleday and D. Rivard (ed.). Sampling commercial catches of marine fish and invertebrates. Can. Spec. Publ. Fish. Aquat. Sci. 66.

Winters, G.H. 1988. The development and utility of sequential population analyses in stock assessments. *In* Rivard D. [Ed.], Collected papers on assessment methods. CAFSAC Res. Doc. 88/61: 111-136.

Appendix I

The Newton-Raphson's solution for F in Baranov's catch equation (equation 3) can be found as follows (see Rivard 1982):

1: Select an initial approximation for the $F_{\iota,T}$ ($F_{\iota,T}^{(0)} = r_{i,t}$ for a possible approximation) and solve (3) for $F_{\iota,T}^{(0)}$.

2: If $C_{i,t}^{(k)} - C_{i,t} \leq$ predetermined constant (e.g. 0.01) for all i's, then stop and set $F_{\iota,T} = F_{\iota,T}^{(k)}$, otherwise use the Newton-Raphson's method to compute a new approximation of $F_{\iota,T}$:

$$F_{\iota,T}^{(k+1)} = F_{\iota,T}^{(k)} - ((C_{i,t} - C_{i,t}^{(k)}) / (C_1 + C_2))$$

where

$$C_1 = N_{i,t} (F_{\iota,T}^{(k)} / (F_{\iota,T}^{(k)} + M_{\iota,T})) e^{-(F_{\iota,T}^{(k)} + M_{\iota,T})}$$

$$C_2 = N_{i,t} (M_{\iota,T} / (F_{\iota,T}^{(k)} + M_{\iota,T})^2)$$

where k is the number of successive approximations and $C_{i,t}^{(k)}$ and $F_{\iota,T}^{(k)}$ are respectively the k^{th} approximations of $C_{i,t}$ and $F_{\iota,T}$.

3: Return to step 2.

Yield per Recruit Analysis

by

R.K. Mohn

Biological Sciences Branch
Fisheries and Oceans
Government of Canada
P.O. Box 550
Halifax, Nova Scotia, Canada
B3J 2S7

Introduction.

Yield-per-recruit (Y/R) is the name given to a family of management models which are generally applied to age structured data. The underlying idea of this type of analysis is to follow the change in biomass of a cohort from recruitment until it is no longer available to the fishery. Growth causes the biomass to increase while natural and fishing mortalities cause it to decrease. Growth, and resultant yield, need not be constrained to weight. We shall present examples where yield is in terms of only the portion of the animal or is expressed in economic terms. As yield per recruit analyses are age-structured models (we will ignore length based models for the present.), growth, natural mortality (*M*) and fishing mortality (*F*) are dependent upon age. The usual form of analysis is to estimate or assume the growth and natural mortality. Various magnitudes and (age) patterns of the fishing mortality are then tried until an optimum solution is found.

Traditionally, two types of overfishing have been identified (Cushing 1981), growth overfishing and recruitment overfishing. Growth overfishing is the situation when the animals are fished before they have reached their potential yield. Recruitment overfishing is when the ability of the stock to reproduce itself has been jeopardized by excessive fishing pressure. Yield per recruit models are the principal tools in assessing growth overfishing. The usual practice is to define the fishing pattern (a result of selectivity and availability for each age-group) from virtual population analysis (VPA) and then determine what level of fishing mortality (*F*) results in the highest yield (F_{max}) or what magnitude of *F* results in a more conservative level such as $F_{0.1}$. These targets F levels will be defined below.

Yield per recruit analysis is often taken one step further; in order to determine a total allowable catch (TAC), the target *F* from the above analysis is multiplied by the selectivity and an estimate of the standing stock. This practice usually violates the stable age distribution assumption and will be given some attention below.

Description of method.

Principal Assumptions

The biomass of an unexploited cohort of fish will tend to increase as the individuals grow until a certain age when the mortality rate exceeds the growth rate. Y/R determines what pattern and intensity of fishing,when applied to this cohort, will result in the optimal yield. We will not define what is meant by optimal at this point. This analysis is dependent upon a knowledge of growth (or weight) at age as well as natural and fishing mortalities as a function of age. The mortality rates are incorporated into an exponential model as seen Winters (1988) and Rivard (1988):

$$N_{a+1} = N_a \exp [-Z_a] \quad , \qquad (1)$$

where Z_a , the total mortality rate for age *a* is defined as $F_a + M_a$. The other fundamental equation is Baranov's catch equation which asserts that removals over a period (or age) are proportional to average abundance over that period.

$$\overline{C}_a = F_a \overline{N}_a \qquad . \qquad (2)$$

F_a is often expressed as (F S_a), which expression presents explicitly the concepts of pattern (S_a) and magnitude (F) for the fishing mortality.

The exponential model of (1) implies that the average numbers over age *a* are

$$\overline{N}_a = (1 - \exp [Z_a]) N_a / Z_a \qquad . \qquad (3)$$

The above formulae are in terms of numbers of animals and catch in numbers. They may be converted to biomass, B_a, and yield (in weight), Y_a, by multiplying by the average weight for each age.

$$Y_a = \overline{W}_a C_a = \overline{F}_a B_a \qquad , \qquad (4)$$

where

$$\overline{B}_a = \overline{W}_a \overline{N}_a \qquad . \qquad (5)$$

The average weight for each age may be determined empirically or from a growth model. The principal distinction among the various Y/R analyses is the manner in which growth is incorporated.

Growth and natural mortality, the natural parameters of our model, must be measured or inferred. They describe the specific system (fishery) to be analyzed. The fishing mortality is generally not known beyond the pattern as a function of age. The age-dependent pattern is often called selectivity (of the gear) or partial recruitment (to the gear). Often, a simplified fishing pattern, expressed in terms of age of first capture, will be examined in the course of an analysis. The magnitude of the fishing mortality is unknown and, consequently, various levels of fishing mortality must be assumed for the Y/R analysis.

Data Requirements

The data requirements for Y/R are relatively undemanding. Weight and mortality as a function of age are required to establish the natural parameters. Weight at age is one of the easier field measurements to make for commercial species which can be aged. Natural mortality as a function of age is generally not available and a constant value over all ages is used. The pattern over age of the fishing mortality may be estimated from VPA results or from the behaviour of the gear used.

A sample set of data is presented in Table 1. It is from the 1986 Georges Bank scallop assessment (Mohn *et al.*, 1987). Natural mortality is assumed to be 0.1 for all ages. These data will be used for sample calculations with the three basic Y/R methods presented below. This table contains height (the scallop equivalent to length), weight and selectivity as functions of age. This data will be used below as a basis for comparison of various Y/R models, such as the *Baranov*, the *Thompson and Bell* and the *Beverton and Holt* models. The extensions to the basic techniques require more detailed data which are given in Table 2. These were derived from the same scallop stock but are compiled on a quarterly basis.

Table 1. Scallop data used in sample calculations of basic analysis.

Age	Height	Weight	Selectivity
3	70.5	5.1	.113
4	93.8	11.8	.750
5	110.0	19.2	1.000
6	121.1	25.8	.705
7	128.8	31.2	.500
8	134.1	35.3	.375
9	137.7	38.3	.375
10	140.2	40.5	.350
11	141.8	42.0	.320

Table 2. . Quaterly scallop data with seasonal and economic factors.

Age	Length	Weight	Sel.	Economic	Seasonal
3	61.23	3.11	0.05	0.95	0.011
3.25	63.22	3.44	0.1	0.95	0.062
3.5	74.57	5.73	0.1	0.95	0.028
3.75	83.13	8.03	0.2	0.95	0.011
4	87.3	9.34	0.4	0.95	0.075
4.25	89.23	10	0.6	0.95	0.413
4.5	96.26	12.64	1	1	0.188
4.75	102.35	15.29	1	1	0.075
5	105.51	16.8	1	1	0.1
5.25	107.02	17.55	1	1	0.55
5.5	111.6	19.99	1	1	0.25
5.75	115.81	22.42	1	1	0.1
6	118.08	23.81	0.75	1	0.071
6.25	119.18	24.5	0.75	0.8	0.388
6.5	122.23	26.49	0.66	0.8	0.176
6.75	125.13	28.49	0.66	0.8	0.071
7	126.72	29.63	0.5	0.8	0.05
7.25	127.5	30.2	0.5	0.8	0.275
7.5	129.55	31.73	0.5	0.8	0.125
7.75	131.54	33.26	0.5	0.8	0.05
8	132.65	34.13	0.375	0.8	0.038
8.25	133.19	34.57	0.375	0.8	0.206
8.5	134.58	35.69	0.375	0.8	0.094
8.75	135.94	36.82	0.375	0.8	0.038
9	136.7	37.47	0.375	0.8	0.038
9.25	137.08	37.79	0.375	0.8	0.206
9.5	138.03	38.6	0.375	0.8	0.094
9.75	138.96	39.41	0.375	0.8	0.038
10	139.48	39.88	0.35	0.8	0.035
10.25	139.74	40.11	0.35	0.8	0.193
10.5	140.39	40.68	0.35	0.8	0.088
10.75	141.02	41.26	0.35	0.8	0.035
11	141.38	41.58	0.32	0.8	0.032
11.25	141.56	41.75	0.32	0.8	0.176
11.5	142	42.15	0.32	0.8	0.08
11.75	142.44	42.55	0.32	0.8	0.032

Target Fishing Levels (*F*).

The two most frequently used target *F*s are F_{max} and $F_{0.1}$. F_{max} is the fishing mortality that results in the maximum yield per recruit for given natural parameters, recruitment age and selectivity pattern. $F_{0.1}$ is a more conservative fishing level which was presented by Gulland and Boerema (1973). Figure 1 shows how it is defined. The slope at the origin, which is the virgin catch rate or more specifically the yield per unit fishing mortality for a virgin stock, is estimated from the Y/R curve. The point where 10% of this slope is tangent to the Y/R curve is found and the *F* at that point is $F_{0.1}$. This is analogous to the concept of marginal yield which is found in economic literature.

F_{max} and $F_{0.1}$ may be found in a number of ways. We first defined a standard *F* series, say .0001, .0002, .1, .2, .3, ... , 1.1, 1.2, etc., To find F_{max}, the highest Y/R from the *F* series and its two neighbors are fit to a parabola. The maximum of the parabola is then found algebraically. If the parabola is expressed as $(aX^2 + bX + c)$, then the maximum is $(-b/2a)$. Similarly, for $F_{0.1}$, a parabola is fit through the origin and the first two points of our *F* series, which is why they are so small. The slope at the origin is the parameter b. The three points whose slope is closest to one-tenth of the slope at the origin are then found. And, similarly to the maximum, the $F_{0.1}$ point is found by equating the slope of the parabola $(2aX + b)$ to the 0.1 slope and solving for X. There are no standard procedures to locate these target F's and the method used can affect the target *F*s obtained. A listing of the APL functions which were used in the preparation of sample results, as well as a listing of APL functions to find F_{max} and $F_{0.1}$, is presented in the Appendix.

Classical Models

Three methods of Y/R are presented as being representative of this type of analysis, those of Baranov, Beverton-Holt and Thompson-Bell. Baranov's (1918) method is not in general use today because of its restrictive assumptions about growth. Contributions from a cohort are summed over length instead of age which was the subscript in equations 1 to 5. The most recent of the three methods, that of Beverton and Holt (1957) has a more realistic growth model (von Bertalanffy) but is sufficiently complicated that analyses generally require the use of a computer. The Thompson and Bell (1934) analysis is the least constrained of the models, both in terms of growth or selectivity. It also is generally performed on a computer.

For the scallop example (see Table 3), the output for these three methods is in terms of meat yield per recruit. The weights in Table 1 are meat weights in grams. For each of the F levels in the F series provided, the yield per recruit is calculated. It is a common practice to construct a tabular representation of a such yields over a range of recruitment ages. The rows of such a table are *various ages of entry to the fishery* and the columns are *the series of F levels* . If the same output values are plotted as surface, or as a contour plot, the result is known as a yield isopleth (Figure 2).

Table 3. Target F's and their yield-per-recruit values for Baranov's, Thompson-Bell's and Beverton-Holt's methods.

Method	Recruit Age	Sel*	F_{max}	$Y(F_{max})$	$F_{0.1}$	$Y(F_{0.1})$
Baranov	3	K	.131	16.2	.052	32.2
	4		.137	19.4	.054	32.7
	5		.141	21.4	.055	32.8
Thompson-Bell	3	K	.284	11.9	.169	11.5
	4		.396	14.6	.191	13.0
	5		.574	17.0	.281	15.6
Thompson-Bell (with plus group)	3	K	.209	13.6	.074	12.9
	4		.271	15.6	.078	13.4
	5		.408	17.5	.150	16.4
Beverton-Holt	3	K	.319	14.1	.176	13.2
	4		.439	16.8	.254	15.9
	5		.640	18.9	.288	17.1
Beverton-Holt (with plus group)	3	K	.230	15.5	.076	14.1
	4		.320	17.5	.080	14.5
	5		.477	19.2	.152	17.6
Thompson-Bell Quarterly	3	D	.617	12.8	.382	12.1
	4		.664	13.3	.395	12.5
	5		.989	15.6	.553	14.6
Thompson-Bell Seasonal Growth	3	D	.600	12.2	.377	11.6
	4		.645	12.8	.391	12.0
	5		.957	15.1	.516	14.0
Thompson Bell Economic	3	D	.684	11.8	.400	11.1
	4		.749	12.4	.451	11.7
	5		1.240	14.7	.596	13.4
Thompson Bell Seasonal Harvest	3	D	.591	12.2	.375	11.6
	4		.638	12.8	.389	12.0
	5		.985	15.4	.551	14.4

* K denotes knife-edge selectivity and D denotes dome shaped as contained in Table 2.

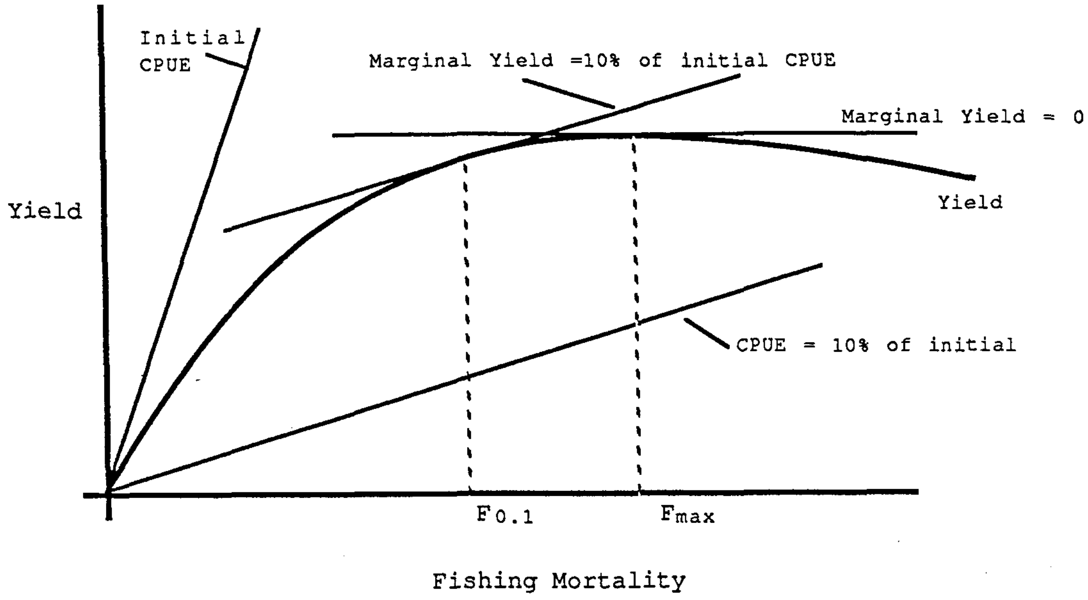


Figure 1. Graphical representation of F_{max} and $F_{0.1}$. The steeper straight line from the origin is the slope of the yield curve for a virgin stock. One tenth of the slope of this line is shown as '10% of initial'. The intersection of the higher line parallel to it and the yield curve defines the $F_{0.1}$ point.

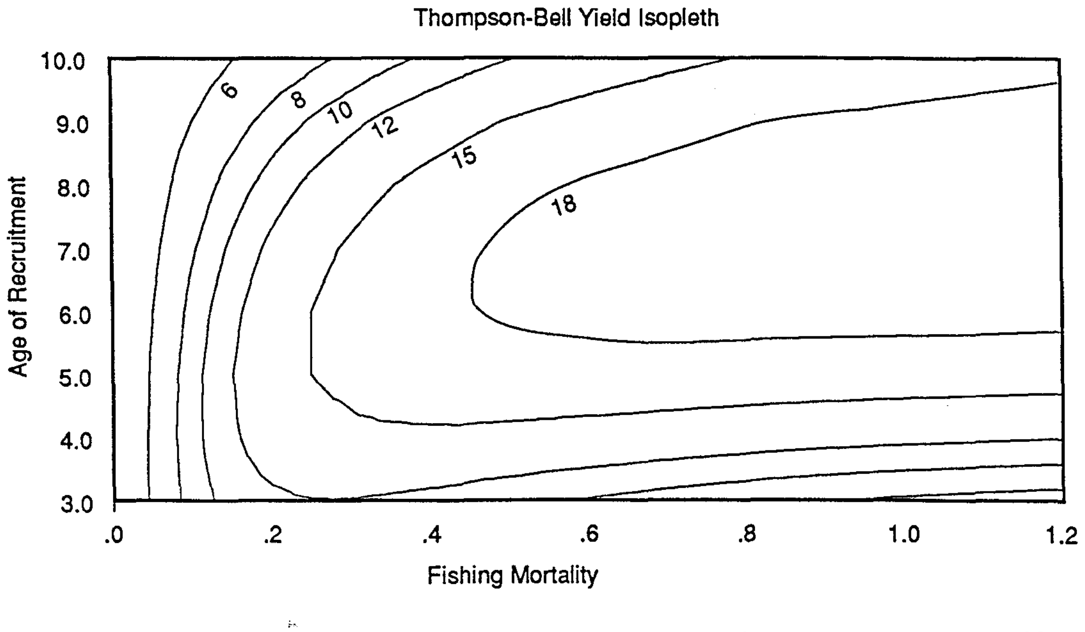


Figure 2. Yield isopleth using the scallop growth data in Table 1 and a knife-edge selectivity. The lines of constant yield are labeled in terms of yield per recruit in grams of meat.

Baranov's method

Baranov's (1918) method assumes that the growth, in terms of, is a linear function of age after recruitment. The linear relationship between age and length allows the summation over the life of the cohort to be made in terms of lengths. This method also assumes that the weight is proportional to the cube of the length. The program provided in the listing of the APL function BAR2 requires length (actually the height for scallops) and weight at age data. It first performs a linear regression between length and age. The linear coefficient, *d*, will be used to convert length after recruitment length in time after recruitment age for the purpose of calculating survivorship. The mean weight at age is divided by the mean length cubed to estimate the parameter which is used to convert the length at recruitment to a weight. An integration factor, which is mainly determined by survivorship (Baranov's *q*), completes the calculation. Equation 4 is the basis for this method but is slightly altered in that the averaging is over the time the animal is fished (recruitment age or length until infinity) instead of a year as was implicit above. The total removals after recruitment is the average standing stock times the rate of total mortality (which is analogous to equation 2 except that all removals are being considered and the average is over the recruited life span). Furthermore, in steady state, these removals must be equal to the recruitment rate, say *R*:

$$\text{Removals} = R = Z\,\overline{N} \quad , \qquad (6)$$

which may be rearranged to express $\overline{N}$ in terms of R and Z :

$$\overline{N} = R/Z. \qquad (7)$$

$$Y = F\,\overline{W}\,\overline{N} \quad . \qquad (8)$$

Although $\overline{N}$ is known, $\overline{W}$ requires some calculus and its derivation is not included. It may be found in Ricker (1975 p.248). However,the resultant formula is

$$Y = F\,R\,a\,L^3\,Q\,S\,/\,Z \quad , \qquad (9)$$

where Q, the factor from the integration, appears as

$$Q = 1 + (3\,d\,/\,L\,Z) + (6\,d^2\,/\,(LZ)^2) + (6\,d^3/\,(LZ)^3) \quad , \qquad (10)$$

and S is the fraction surviving from the minimum recruitment size L_0 to any of the trial L's

$$S = \exp[\ -M\,(L - L_0)\,/\,d\ ] \quad . \qquad (11)$$

The sample data is run at three recruitment lengths, 70.5, 93.8 and 110 mm and at using the standard range of fishing mortalities presented with the methods for determining F_{max} and $F_{0.1}$. The recruitment lengths correspond to 3, 4 and 5 year old animals (mid-season size). It should also be noted that selectivity is not included in this procedure and recruitment is so called knife-edge.

The Thompson and Bell Method

The method of Thompson and Bell (1934) is, conceptually and algebraically, more simple than Baranov's or Beverton and Holt's which follows. Unlike both of them, which incorporate an analytical expression for the integral over the fished life of the harvested species, the Thompson-Bell method simply sums the contribution from each age class individually. It is quite similar to stock projections seen in Gascon (1988) or to VPA calculations going ahead in time instead of backwards. Consequently, growth does not have to be parameterized by a linear growth function and isometric length-weight, as in Baranov's method, or as a von Bertalanffy growth function, as in Beverton-Holt Y/R. One proceeds by using equations 1, 3 and 4. An initial number of recruits, say 1000, are assumed to start the process. The average number for the first recruited age class over the first year is found from Equation 3. The yield from that year is found with equation 4 and, finally, the number surviving from Equation 1. The process is then repeated for the second year in the fishery, etc.

A problem exists for the oldest age class in the Thompson-Bell calculation and that is whether or not to treat it as a 'plus' group. If it is a plus age group, that means that all animals of that age and older are contained in that group. For example, if the oldest category in Table 1 were 11+, it would signify that these values were for all animals of age 11 and older. This would mean that Equation 3 was no longer appropriate, as the averages are for from one age to the next, i.e. not for an entire class of ages. This is corrected by changing (3) to

$$\overline{N} = N_a / Z_a \quad . \qquad (12)$$

It is worth noting that fishing mortality, natural mortality and weight may all be functions of age. This allows the Thompson and Bell method to incorporate selectivity (or age-dependent natural mortality) which Baranov's method does not handle.

The Beverton and Holt Model

Beverton and Holt (1957) presented a method which uses a realistic growth model, i.e. the von Bertalanffy model for length-at-age and an isometric relationship to link length and weight. Also, the Beverton and Holt model differs from Baranov's method in that the integral is over time instead of length. The yield expression is quite similar to Baranov's and, again, the reader is referred to Ricker (1975 p. 251) for details of derivation.

$$Y = F\, R\, W_{\infty}\; Q'\, S \,/\, Z \quad . \qquad (13)$$

where the integration term Q' is

$$Q' = (\,1 - \exp[\,-Z(t-t_0)\,]\,)\,/\,Z\; .....) \quad . \qquad (14)$$

The three differences between equations (13)-(14) and (9)-(10) are: that the integration terms differ (Q and Q'); that $a\,L^3$ of (9) which is the weight at the size of entry to the fishery has been replace by the asymptotic weight W_{∞} from von Bertalanfy-isometric growth model; and that the $L\text{-}L_0$ as been expressed in terms of time $t\text{-}t_0$, the interval from recruitment age to age of first capture. Similarly, the survivorship term of Equation (13) replaces $(L\text{-}L_0)/d$ with $t\text{-}t_0$. Beverton-Holt allows one to change the onset of knife-edge recruitment but not the selectivity pattern.

The results of the three methods of Y/R are shown in Table 3. In that example, delaying the entry into the fishery leads to higher yield for all methods. For comparison, the F_{max} and $F_{0.1}$ values are also calculated and are compiled into Table 3. The Thompson-Bell results are generally higher than those from the other two methods, especially for the *F* at $F_{0.1}$, while the others have their maxima (for *F*) at considerably higher levels. The most striking entries are for the Thompson-Bell method when age of entry is 5. Because the Y/R does not reach a maximum, the quadratic fit does an extrapolation to what are obviously unreliably high values.

Extensions.

Stock Projections

Once the target F has been determined, it may be used to estimate a catch level. The catch equation in terms of yield (Equation 4) is used with an estimate of the standing stock biomass and the target F as input and the resultant yield provides the recommended catch level (see Gascon 1988). The target F was derived with a stable age distribution. It is very unlikely that an exploited stock will be in the stable age distribution for the target F and, therefore, the recommended catch level may not correspond to the target F. Deviations from a stable age-distribution can result from a number of causes; the most significant would be recruitment variability or changing fishing practices. Sinclair et al. (1983) investigate the consequences of applying a stable-age derived F to a non-stable age standing stock. No general solution to the problem has appeared but, in qualitative terms, if the stock has more young animals than the stable distribution, than the target F will be too high and conversely if the population is older than the stable age distribution. The reader is warned of the situation but, unfortunately, a solution cannot be presented. If recommended catches based on Y/R target F's are applied over a number of years, they will converge to proper values and, in the long term, the advice will be appropriate (consistent with the model). Because the analysis runs over the recruited life of the cohort, it does not provide answers as to the optimal harvest for a given stock for the short term. Mohn (1984) presents a method for estimating the contribution from each age-class in the standing stock which may be of value when generating advice for a stock which is far from the stable age distribution.

Economic Inputs.

The examples presented above are for a scallop stock and the yield is in terms of the meat of the scallop. It is relatively simple, especially with the Thompson-Bell method, to expand the analysis with economic considerations. For example, if the price of scallop meats depended on there size, which is indeed the case, one could multiply the cost per unit meat times the meat weight at age and thereby convert yield into revenue instead of meat. The economic price factors shown in Table 2 show the relative price for various sizes of meats. When doing the Thompson-Bell calculations, the product of this factor and the meat weight is used in place of weights alone. All other aspects of the calculations are unchanged. The conversion to economic parameters in the Beverton-Holt model requires a re-fitting of the von-Bertalannfy growth model with the cube root of the price factors multiplying the length-at-age. This is deemed to be a less desirable model to use because of this complication and because the resultant relationship may not fit the von Bertalanffy model if the prices vary significantly with age. The results are shown in Table 3 (Thompson-Bell Economic) and the assumed discounting of larger animals (Table 2) results in higher target F's.

Another economic factor is the cost of fishing. In the simplest terms, one may assume that fishing cost is proportional to fishing effort. Also, as is commmonly done, we can assume that the fishing mortality is porportional to fishing effort. Then, a final assumption is that the revenue from the catch is porportional to the catch magnitude (i.e. constant prices). Now, if one divides the yield-per-recruit at a given *F* by the *F*, the result will be comparable to a profit rate - how much you get per dollar spent in fishing. As an example, consider Table 3 (Thompson-Bell Quaterly method) and an age of entry of 4. If $Y(F_{max})/F_{max}$ (= 20.0) is compared to $Y(F_{0.1})/F_{0.1}$ (=31.6), we see that the $F_{0.1}$ strategy is almost 70% more efficient. In other words, although F*max* results in a slightly higher yield, it requires a lot more effort to realize it. More complex models incorporating the operations of the fishing fleet and the costs of the processing operation could be developed, as illustrated by Huson *et al.* (1984).

Seasonal Factors

Scallops exhibit a seasonal growth pattern, as do all fish in temperate waters. The data in Table 1 ignore seasonal growth but these effects are included in Table 2. The Thompson-Bell method is the only one of the three presented above which can include seasonal effects. When this method is applied to scallop data, it is seen that the effects on the target F's is very small (see, in Table 3, the rows for Quaterly versus Seasonal Growth Thompson and Bell). This analysis may be further 'seasonalized' by including seasonal fishing patterns. Very little scallop fishing takes place during the first 3 months of the year. How does the distribution of effort affect the target F's derived from Y/R. Again with the Thompson-bell method, one needs only to correct the selectivity by a seasonal coefficient. Column 4 of Table 2 has been multiplied by the season distribution 10, 55, 25 and 10%. These coeffcients have an average value of 1 so as not to distort the total F. Again, the affect is not large when one considers a seasonal distribution of effort on our sample data set.

The use of quarterly defined growth and mortality data in yield per recruit analysis is not as straight forward as it may appear. This is mainly the result of the way in which selectivity (or partial recruitment) is traditionally normalized. The accepted practice is to normalize the selectivity such that the largest element has a magnitude of one (this is also called fully recruited age). When a stock projection is performed, the selectivity pattern is multiplied by a fishing intensity to form the mortality at age. In transforming *annual figures* to or from *seasonal figures*, the rescaling of the selectivity means that one must rescale the F in a compensatory fashion. For example, if one were taking annual values to quarterly ones and it was known that the distribution of fishing mortalities F by quarters was 10, 45, 30 and 15%, then the fully recruited age would become seasonal values (i.e. 0.1, 0.45, 0.3 and 0.15) which would be renormalized to 0.22, 1, 0.66 and 0.33 . To compensate for the renormalization, the fishing intensities would have to be multiplied by 2.22 (=1/0.45). Similarly, if the seasonal selectivities for the fully recruited annual age-class were 0.4, 0.6, 1 and 0.8, then the fishing intensity would have to be divided by 2.8 (=0.4+0.6+1+0.8) to maintain the same rate of removals.

Considering changes in gear selectivity or fish availability.

Selectivity and partial recruitment are used almost interchangably to describe the age distribution of the fishing mortality. Selectivity is generally associated with the characteristics of fishing gear while partial recruitment is considered to be a characteristic of the fish. The gear 'selects' the fish and the fish are 'recruited' to the gear. In the case of a trawl, one can think of it as the small fish passing through the net (pre-recruits), the large fish all being caught (fully recruited) and those in between having an intermediate probability of being caught (partially recruited). But the formal role of these two inter-related concepts in the catch equation is that of a coefficient to the fishing mortality and no distinction is made as to whether it is a fish or a gear based characteristic. Fishermen may concentrate their effort on a specific year-class and, in this case, the selectivity of the gear will not solely determine the age distribution of the catch. To deal with this situation, Mohn (1986) introduced a behavioral parameter into the catch equation which was a coefficient of F but could be function of abundance. This was done because scallop regulations required at that time that the average size of the catch was above a minimum but the individual animals could be of any size. This regulation resulted in a practice known as blending in which a few large animals could be blended with a number of smaller ones to bring the average over the legal limit. The behaviour of the fleet then became a function of the relative abundance of the various age classes. If large animals were rare, proportionally more effort would have to be expended to blend up the catch. The fishing mortality defined in Equation 2 now becomes

$$F_a = S_a \beta_a F \quad , \qquad (15)$$

where β_a is a function of the relative abundance of the stock structure. In the case of schooling fish, particularly pelagics, β_a would be used to model the fishermen's practice of targetting strong year classes. The behavioral parameter cannot be used with either Baranov's or Beverton-Holt's methods. It also cannot be used directly with Thompson-Bell's. What is required is that Thompson-Bell be used as a stock projection with constant recruitment. A trial catch is found and, if it does not have the appropriate characteristics, β_a is iteratively changed and new trial catches are estimated until the desired catch results. In the case of blending, this is done by treating the mortality on animals above the average limit as a block and those beneath as a second block. If the average catch is too small, the block above the limit receives a larger proportion of the mortality (effort) and that below is adjusted such that the total magnitude of the mortality is unaffected. The details of such calculations are not be included here. Similarly, for the case when strong year classes are targeted, the behavioral parameter is adjusted so that the total catch per unit effort is maximized, subject to size limitations. This again is done on a iterative-trial projection basis.

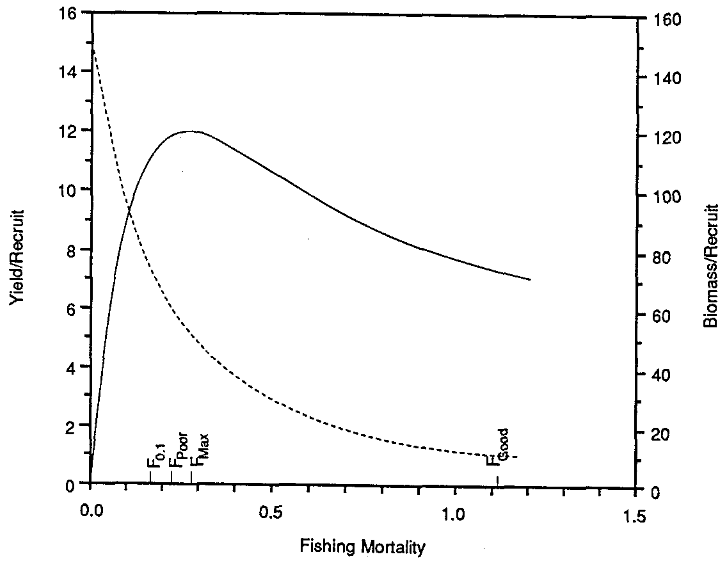
Linkages to recruitment models.

Y/R analysis gives only a relative (i.e. per recruit) index. Above we saw that the results must be multiplied by the standing stock to convert it into real units, such as tons. Another method to convert it to absolute units would be to multiply it by an estimate of the recruitment which is essentially what is done in stock projections (see Gascon 1988). If recruitment values - or a recruitment model - were available, one could link it to Y/R techniques to give insights in terms of absolute values.

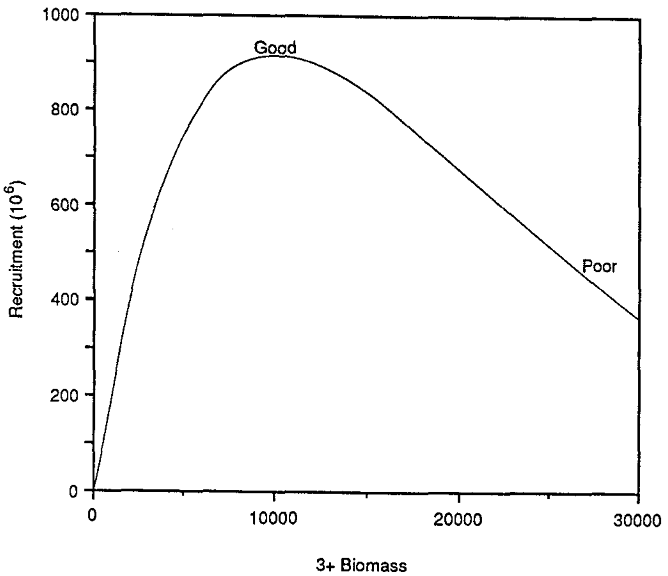
Anon. (1984) presents a method of linking Y/R analysis to a stock-recruitment model. The purpose of this analysis was to develop alternative target F's which would take into account the recruitment resilience of the stock. This is done by recalling that, for a given *F*, there is a resultant Y/R and a resultant stable age distribution. The biomass of the stable age distribution may be thought of as the biomass per recruit. At the same time, if one considers that the slope from the origin to any point on a stock-recruitment curve gives the recruitment per biomass, the reciprocal of the latter gives biomass per recruit (Figure 3). This approach assumes that the stock is in a steady state. The results of Anon. (1984) was the definition of two new target *F*'s. If one assumes recruitment is going to be good, this can be converted into a steady state recruit per biomass from the stock recruit relationship in a region of the curve where the production of the recruits is high (i.e. the region marked GOOD in the lower portion of Figure 3). This slope is inverted to give a biomass per recruit and the F position of the inverted slope gives a target *F* which has been called F_{GOOD}. Similary, the expectation of poor recruitment will yield a lower target *F*, F_{POOR} in Figure 3. F_{GOOD} is seen to be considerably higher that F_{max}. These concepts have been recently examined in Sissenwine and Shepherd (1987).

Steady state predictions incorporating recruitment may also be done more directly by solving iteratively the stock-recruit and yield per recruit relationships. This is done by first finding the steady state biomass for each F level in a series, say F_{TEST}. For each standing stock, a trial recruitment is calculated using the stock-recruitment function. The trial recruitment is compared to the number of recruits in the standing stock. The standing stock is scaled by a correction factor (constant over ages) until the trial recruitment matches the recruits in the steady state standing stock. The trial recruitment in then compared to the recruitment used to generate the standing stock and this factor is used to scale the yields originally calculated. The resultant yields are the steady state solutions of production or it may be thought of as the steady state yield per recruit times the steady state recruitment.

Variable recruitment, deterministic or stochastic, can also be used in conjunction with Y/R analysis. The calculations are the same as for a stock projection; the difference is the emphasis of the exercise. In a yield-per-recruit context, such projections would be the basis of comparison of various target F's derived from Y/R. These projections may be run with historical recruitment indices. In this case, the results address the question "what would have been the yield if say $F_{0.1}$ had been followed instead of what was actually done" (assuming that the recruitment would not have been affected by the changed fishing pattern). The other use is more predictivive in nature. In either case, Y/R matched with dynamic recruitment functions can be used to estimate the stability of the yield series under various harvesting strategies. As the fishing mortality is reduced, the standing stock increases and, unless this destabilizes recruitment, the year to year variation in the catch should also stabilize. Stability of catch would have economic benefits, particularly for the processing sector. Figure 4 and Table 4 show 16 year (restrospective) projections when the recruitment is the series found by VPA for the previous 16 years. The three curves are for 1) the yield using the fishing mortality of 1981 (that year was chosen as it was a year of high fishing pressure on relatively young animals), 2) a projected F_{max} yield and 3) a projected $F_{0.1}$ yield. Two indices of variability are presented in the table, the standard deviation of the yield and the year-to-year deviation. In a stock displaying highly variable recruitment, such as scallops, the stability increase for the relatively conservative $F_{0.1}$ is not large.



3.a Thompson-Bell with knife-edge selectivity



3.b Ricker recruitment function

Figure 3. Linking yield-per-recruit analysis with an arbitrary stock recruitment function. The slope from the origin to the area marked 'Good' or 'Poor' in the lower Figure is inverted and located on the upper to define the new target F's, F_{GOOD} and F_{POOR}. For example, the slope to 'Good' is approximately 900 million animals per 10,000 t meat biomass. This corresponds to a biomass per recruit of 11 grams and a F_{GOOD} of 1.12 .

Discussion

Yield-per-recruit analysis has two major components. The first is a descriptive component in which the natural and fishing influences are analyzed to produce a target *F*. Given a fish that grows a certain way, has a certain mortality pattern and is harvested with a gear having certain characteristics, a target *F* may be derived which describes how theses forces may be balanced to achieve the objective of the target *F*. The second component is predictive: the yield per recruit serves to define a target *F* which is applied to a stock estimate to predict the allowable catch or to predict the yield variability given a certain recruitment function.

The Thompson-Bell method is endorsed as the most versatile and the simplest (conceptually) of the methods reviewed.

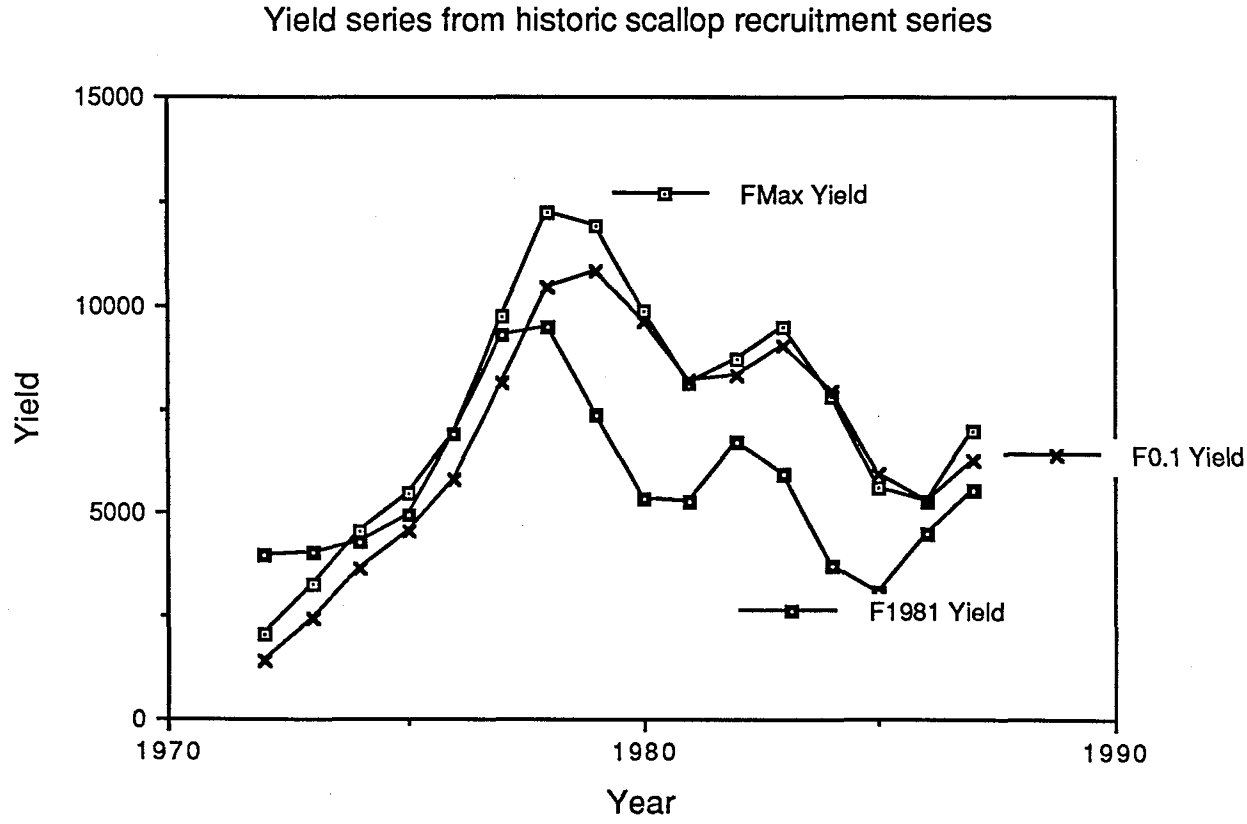


Figure 4. Yield series from historic recruiment series. The input data and average yields are given in Table 4.

Table 4. Input values for and results of yield projections with historical Georges Bank scallop recruitment series (1972-1987).

Recruitment series:

475 529 732 1197 1220 780 497 416 888 761 270 211 471 836 409 483

Selectivity for F_{max} and $F_{0.1}$:

0.11 0.75 1.00 0.70 0.50 0.375 0.375 0.35 0.32

Selectivity for F_{1981}:

0.45 1.00 0.92 0.53 0.31 0.13 0.22 0.48 0.55

	Target	**Mean yield (t)**	**Standard Deviation**	**Serial Deviation**
F_{max}	0.625	7,400	2,800	1,400
$F_{0.1}$	0.383	6,700	2,700	1,200
F_{1981}	1.200	5,700	1,800	1,100

References

Anon. 1984. Int. Coun. Explor. Sea Coop. Res. Rep. 129, 134 p.

Baranov, F.I. 1918. On the question of the biological basis of fisheries. Nauchn. Issled. Ikhtiologicheskii Inst. Izv. 1: 81-128.

Beverton, R.J.H. and S.J. Holt. 1957. On the dynamics of exploited fish populations. U.K. Min. Agric. Fish., Fish. Invest. (Ser. 2) 19:533 p.

Cushing, D.H. 1981. Fisheries biology: a study in population dynamics. Univ. Wisc. 295 p.

Gascon, D. 1988. Catch Projections. *In* Rivard D. [Ed.], Collected papers on assessment methods. CAFSAC Res. Doc. 88/61: 135-146.

Gulland, J.A. and L.K. Boerema 1973. Scientific advice on catch levels. Fish. Bull. 71: 325-335.

Huson R.M., D. Rivard, W.G. Doubleday and W.D. McKone 1984. Impact of varying mesh size and depth of fishing on the financial performance of an integrated harvesting/processing operation for redfish in the Northwest Atlantic. N. Am. J. Fish. Man. 4:32-47.

Mohn, R.K. 1986. Generalization and recent usages of yield per recruit analysis. *In* Jamieson, G.S. and N.F. Bourne [Ed.], Can. Spec. Publ. Fish. Aquat. Sci. 92: 318-325.

Mohn, R.K., G. Robert and D.L. Roddick. 1987. Georges Bank scallop stock assessment - 1986. CAFSAC Res. Doc. 87/9, 23 p.

Ricker, W.E. 1975. Computations and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Canada No 191, 382 p.

Rivard, D. 1988. Biological production and production models. *In* Rivard D. [Ed], Collected papers on assessment methods. CAFSAC Res. Doc. 88/61: 44-110.

Sinclair, M., R. O'Boyle and T.D. Iles. 1983. Consideration of the stable age distribution assumption in "analytical" yield models. Can. J. Fish. Aquat. Sci. 40: 95-101.

Sissenwine, M.P. and J.G. Shepherd. 1987. An alternative perspective on recruitment overfishing and biological reference points. Can. J. Fish. Aquat. Sci. 44: 913-918.

Thompson, W.F. and F.H. Bell. 1934. Biological statistics of the Pacific halibut fishery. 2. Effect of changes in intensity upon total yield and yield per unit gear. Rep. Int. Fish. (Pacific Halibut) Comm. 8: 49 p.

Winters, G.H. 1988. The development and utility of sequential population analyses in contemporary assessment science. *In* Rivard D. [Ed.], Collected papers on assessment methods. CAFSAC Res. Doc. 88/61: 111-136.

Appendix A. Computer listings of programs used in the yield per recruit chapter. The listings are not well annotated and some expertise would be required to implement them.

BARANOV YIELD PER RECRUIT

```
Y←BAR2
⍝
LW←⍎LWN←' Length-weight N×2 'DEFAULT LWN
M←' M 'DEFAULT M
RL←'RECRUITMENT LENGTH 'DEFAULT RL
NA←⍴LR←'Length range of 1st capture 'DEFAULT LR
NF←⍴FR←'F Range 'DEFAULT FR
C1←(⍳1↑⍴LW)RLIN LW[;1] ⍝ FIND COEFFICIENTS FOR LINEAR GROWTH
d←' LINEAR GROWTH COEF. 'DEFAULT C1[2]
a←(+/LW[;2])÷+/LW[;1]*3 ⍝ FIND a S.T. W = aL*3
Q←1+(3÷LZ)+(6÷LZ*2)+6÷(LZ←(LR÷d)∘.×FR+M)*3
SV←*-M×⍉(NF,NA)⍴((LR-RL)÷d)
Y←1×SV×((NA,NF)⍴FR)×a×Q×((LR*3)∘.×(⍴FR)⍴1)÷(NA,NF)⍴FR+M
```

BEVERTON-HOLT YIELD PER RECRUIT

```
Y←BEVHOLT
⍝ BEVERTON-HOLT YIELD PER RECRUIT  RICKER SECTION 10.2
LW←⍎LWN←' Length-weight N×2 'DEFAULT LWN
M←' M 'DEFAULT M
RA←' RECRUITMENT AGE 'DEFAULT RA
NA←⍴AR←'Age range of 1st capture 'DEFAULT AR
NF←⍴FR←'F Range 'DEFAULT FR
K←' K 'DEFAULT K
LINF←145.7 ⍝ FROM VONB FIT OF SCALLOPS MASTA
T0←'T NOUGHT 'DEFAULT T0
LA←(TL←' OLDEST AGE 'DEFAULT TL)-AR
C1←(⍟LW[;1])RLIN⍟LW[;2] ⍝ ALLLOMETRIC REGRESSION
WINF←'W INFINITY 'DEFAULT WINF←(*C1[1])×LINF*C1[2]
EK←*-K×⍉(NF,NA)⍴AR-T0 ⍝ EXP OF K×R USED TO FIND Q
Z←M+F←(NA,NF)⍴FR
CO←1-*-LA∘.×Z[1;]∘.+0 1 2 3×K
Q←(((÷Z),[2.5](¯3×EK÷Z+K)),((3×EK*2)÷Z+2×K)),((-EK*3)÷Z+3×K) ⍝RICKER 10.19
Q←+/CO×Q
Y←1×F×(⍉(NF,NA)⍴*-M×AR-RA)×WINF×Q
```

BASIC THOMPSON-BELL YIELD PER RECRUIT

```
[ 1] Y←QDD
[ 2] ⍝THOMPSON BELL
[ 3] N←1↑⍴WS←⍎WSN←' Weight-selectivity Nx2 'DEFAULT WSN
[ 4] M←' M 'DEFAULT M
[ 5] RA←' RECRUITMETN AGE 'DEFAULT RA
[ 6] NA←⍴AR←'Age range of 1st capture 'DEFAULT AR
[ 7] NF←⍴FR←'F Range 'DEFAULT FR
[ 8] PLUSON←'PLUSON 'DEFAULT PLUSON
[ 9] FM←(NA,NF,N)⍴FR∘.×WS[;2]
[10] EZM←1-*-ZM←(FM←(2 1 3⍉(NF,NA,N)⍴AR∘.≤RA+¯1+⍳N)×FM)+(NA,NF,N)⍴M
[11] CZM←+\0,0 0 ¯1↓ZM
[12] EZM[;;N]←PLUSON⌈EZM[;;N] ⍝ Oldest age plus group?
[13] N←1×*-CZM
[14] C←1×FM×EZM×(*-CZM)÷ZM
[15] Y←C+.×WS[;1]
```

THOMPSON-BELL WITH STOCK-RECRUIT FUNCTION

```
[ 1] Z←QDP
[ 2] ⍝THOMPSON BELL Y/R WITH S-R FUNCTION
[ 3] NAG←1↑⍴WS←⍎WSN←' Weight-selectivity Nx2 'DEFAULT WSN
[ 4] M←' M 'DEFAULT M
[ 5] KG←'SHEPHERD K + G 'DEFAULT KG ⍝ Shepherd S-R function
[ 6] FOG←'FECUNDITY OGIVE'DEFAULT FOG
[ 7] ZM←(FR∘.×F)+((⍴FR)⍴1)∘.×(⍴F)⍴M
[ 8] NF←⍴FR←'F Range 'DEFAULT FR
[ 9] RA←'RECRUITMENT AGE 'DEFAULT RA
[10] NA←⍴AR←,'AGE RANGE OF 1ST CATPURE 'DEFAULT AR
[11] PLUSON←'PLUSON 'DEFAULT PLUSON
[12] Z←(NA,NF)⍴ARI←0
[13] ARLOOP:→(NA<ARI←ARI+1)/AREND
[14] FM←FR∘.×WS[;2]×AR[ARI]≤RA+¯1+⍳NAG
[15] EZM←1-*-ZM←FM+(NF,NAG)⍴M
[16] CZM←+\0,0 ¯1↓ZM
[17] EZM[;NAG]←PLUSON⌈EZM[;NAG] ⍝ OLDEST AGE PLUS GROUP OR NOT
[18] N←1000×*-CZM
[19] C←1000×FM×EZM×(*-CZM)÷ZM
[20] Y←C+.×WS[;1]
[21] B←N+.×W
[22] I←0
[23] NT←N
[24] TOP:→(10<I←I+1)/END
[25] RECT←(NT+.×FOG)÷1+(NT+.×W÷KG[1])*KG[2]
[26] NT←NT×⍉(⌽⍴N)⍴(NT[;1]÷RECT)*¯1
[27] →TOP
[28] END:Y←Y×RECT÷1000000
[29] Z[ARI;]←Y
[30] →ARLOOP
[31] AREND:
```

THOMPSON-BELL WITH HISTORIC RECRUITMENT SERIES

```
QDRUN
⍝DYNAMIC VERSION OF Y/R (REALLY STOCK PROJ)
N←1↑⍴WS←⍎WSN←' Weight-selectivity Nx2 'DEFAULT WSN
⍝M←' M ' DEFAULT M
RV←' RECRUITMENT VECTOR 'DEFAULT RV
P←PIN ⍝←'INITIAL POP ' DEFAULT PIN
NY←⍴FV←'F VECTOR 'DEFAULT FV
⍝ PLUSON←'PLUSON ' DEFAULT PLUSON
STAT←(2,⍴RV)⍴Y←0
TOP:→((⍴RV)<Y←Y+1)/END
Z←M+WS[;2]×FV[Y]
C←FV[Y]×WS[;2]×P×(1-*-Z)÷Z
STAT[;Y]←(+/C),WS[;1]+.×C
P←RV[Y],¯1↓P×*-Z
→TOP
END:0 ROUND+/2 ¯10↑STAT
```

FUNCTION TO FIND FMAX WITH QUADRATIC FIT

```
Z←QMAX X;I;C
I←¯1 0 1+2⌈(1↑⍋X[;2])⌊¯1+1↑⍴X ⍝ SELECT MAXIMUM POINT AND NEIGHBORS
C←LAGR X[I;]
→(0>1↑C)/OK
Z←X[1↑⍋X[;2];]
→0
OK:Z←¯0.5×C[2]÷C[1]
Z←Z,+/(Z*2 1 0)×C
```

FUNCTION TO FIND 0.1 WITH QUADRATIC FIT

```
Z←OP2 X;SL;I;C
⍝FIND OP1 POINT BY MATCHING 0.1 SLOPE AND QUADRATIC SLOPE
SL←0.1×1↑1↓LAGR X[⍳3;]
DYDX←÷÷/(1 0↓X)-¯1 0↓X
I←¯1 0 1+2⌈+/SL<DYDX
C←LAGR X[I;]
Z←(SL-C[2])÷2×C[1]
Z←Z,+/(Z*2 1 0)×C
```

LAGRANGIAN INTERPOLATION FUNCTION

```
Y←LAGR X;TMP
⍝LAGRANGIAN INTERPOLATION - RETURNS POLYNOMIAL COEFFICIENTS
⍝ X IS 3×2 MATRIX OF F AND YIELDS
TMP←X[;1]∘.*(1↑⍴X)-⍳1↑⍴X
Y←(⌹TMP)+.×X[;2]
```

DATA INPUTTING FUNCTION

```
R←P DEFAULT X
⍝PROMPTS FOR INPUT AND RETURNS DEFAULT VALUE IF NONE
⍞←P
' CURRENT VALUE '
→(0=1↑0⍴X)/NUMERIC
X
→(0<⍴R←,⍞)/END
→SUBS
NUMERIC:X
→(0=⍴R←⍞)/SUBS
R←⍎R
→END
SUBS:R←X
END:
```