Fol. 41 Assess

This report not to be quoted without prior reference to the Council*

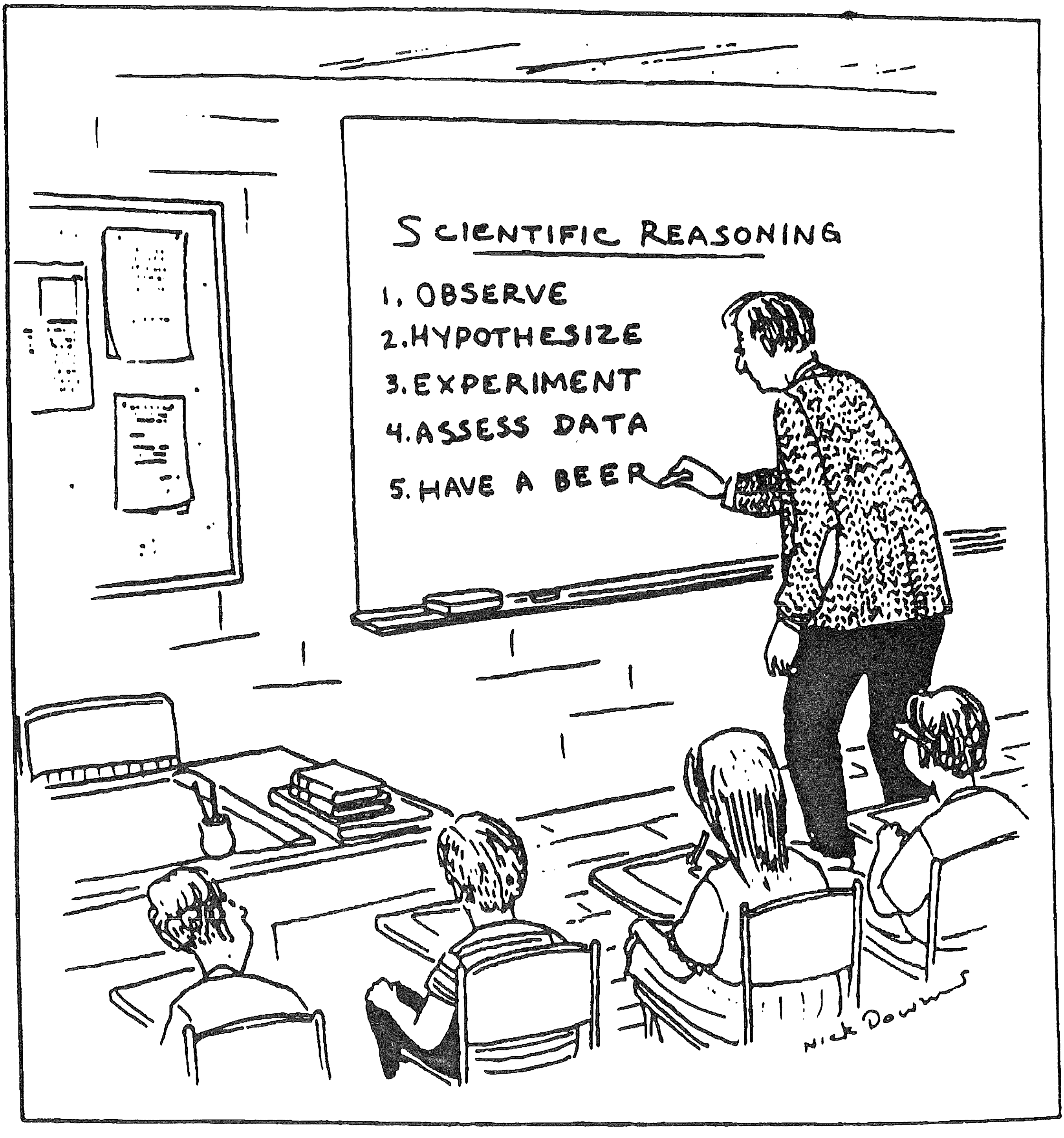
International Council for the
Exploration of the Sea

C.M.1993/Assess:12

REPORT OF THE WORKING GROUP ON METHODS OF FISH STOCK ASSESSMENT

Copenhagen, 3-10 February 1993

*General Secretary
ICES
Palægade 2-4
DK-1261 Copenhagen K
DENMARK


SCIENTIFIC REASONING
1. OBSERVE
2. HYPOTHESIZE
3. EXPERIMENT
4. ASSESS DATA
5. HAVE A BEER
Nick Downes

CONTENTS

1 INTRODUCTION

1.1 List of Participants

Vladimir Babayan	Russia
Tatiana Bulgakova	Russia
Doug Butterworth	Iceland
Robin Cook	U.K.(Scotland)
Chris Darby	U.K.(England)
Jean-Louis Durand	France
Guus Eltink	Netherlands
Rob Fryer	U.K.(Scotland)
Pierre Gagnon	Canada
Gudmundur Gudmundsson	Iceland
Svein Iversen	Norway
Tore Jakobsen	Norway
Eskild Kirkegaard (part-time)	Denmark
Hans Lassen (part-time)	Denmark
Manuel Meixide	Spain
Benoit Mesnil	France
Ransom Myers	Canada
Victor Restrepo	U.S.A.
Andy Rosenberg	U.S.A
John Shepherd	U.K. (England)
Gunnar Stefánsson (Chairman)	Iceland
Valentin Trujillo	Spain
Bernhard Vaske	Germany
Roger Bailey (part-time)	ICES
Henrik Sparholt (part-time)	ICES

1.2 Terms of Reference

It was decided at the 80th Statutory Meeting in 1992 (C.Res.1992/2:8:21) that the Working Group on Methods of Fish Stock Assessment (Chairman: Dr G. Stefánsson) would meet at ICES Headquarters from 3-10 February 1993 to:

a) investigate the use of risk analysis, especially how it might be useful in addressing the definition of safe biological limits;

b) investigate alternative assessment methods for short-lived species and advise on their usefulness;

c) investigate the appropriate use of shrinkage in tuning and advise how it should be implemented for assessment Working Groups.

d) investigate, using retrospective analyses, which regression methods are most appropriate for recruitment estimation, with particular reference to North Sea herring and Icelandic capelin, and advise on how these recruitment estimates should be brought forward into the predictions;

e) review the reports of the Workshop on the Analysis of Trawl Survey Data and the Planning Group for the Development of Multispecies, Multifleet Assessment Tools and indicate promising directions for future development.

1.3 Working Papers

Working papers were available on some of the topics. These are listed in Section 8.2

1.4 Notation

The Working Group adhered as far as possible to the standard notation used previously, expanded as necessary. An updated version is given in Appendix A.

1.5 Structure of the Report

The items in the Terms of Reference have been viewed in a rather general light. In most instances, a specific Term of Reference is really only a part of a larger issue. An attempt has been made to cover the corresponding topics in a reasonably comprehensive manner.

Section 2 of the report deals with possible solutions to problems associated with the management and assessment of short-lived species. For many short-lived species, the problems involved are not just connected with data and assessments, but are rather a part of the entire process: biology—assessment—management. Thus, examples are given of how short-lived species are handled in cases where formal management procedures have been used. It is also shown why it is essential to consider the entire process rather than just the assessment.

Section 3 on assessment methodology evaluates the use of shrinkage in assessments, recruitment prediction and integration of the two, giving advice on how to proceed on these issues.

Section 4 introduces several approaches to the analysis of stock-recruitment relationships, giving potential methods for determining SSB threshold levels. The Working Group agreed that risk analysis, as used by ICES working groups (i.e. in the form of computing probability profiles), cannot be used to obtain minimum biologically acceptable levels (MBALs). Other criteria must be used to obtain such levels, and stock-recruitment relationships play the single most important role. Therefore, several methods were considered for obtaining MBALs based on such relationships or indications thereof.

Section 5 introduces the methodology required for the evaluation of short-term and medium-term advice. This includes evaluation of management strategies and the uncertainty involved in predictions. The term risk analysis has been used in a broad sense within the ICES arena. In this report distinctions are made, particularly between short-term predictions and medium-term predictions.

2 SHORT-LIVED SPECIES

2.1 Introduction

Examples of problems raised in this context did not all appear to be related to life-span *per se*, as some of the species mentioned (e.g., sandeels, sardine) were harvested over age ranges extending beyond 5 years of age. Rather, the essential difficulty seems to relate to the provision of the short-term projections required for management purposes. This is frustrated, either because of the large contribution which the recruitment of the forthcoming year will make to the biomass of a genuinely short-lived species, or because of evaluation difficulties which arise for longer-lived species because of inadequacy or absence of certain data (such as recruit surveys).

The first of these situations occurs typically in anchovy fisheries, where the stock consists of a few year classes only. Even if regular surveys take place, much of the annual catch can have been taken before projections can be adjusted to take account of the results of the most recent survey. This is, therefore, as much a management as an assessment problem. The Working Group considered examples of two such stocks — South African anchovy and Icelandic capelin — for which "management procedures" (see Section 5.4) have been adopted to address such problems. These examples are summarised briefly below, followed by a description of how this approach might be applied to the anchovy stock in ICES Sub-area VIII. Examples given in Anon. (1992a) show that uncertainties about the level of F, M and recruitment lead to a 1/4 to 2-fold uncertainty about the current size of this resource in relation to the most recent spawning stock biomass estimate by egg survey.

For the second set of situations, there are usually sufficient data in principle to perform a full age-structured (VPA) (Gulland, 1965) assessment, but this is unsatisfactory for a number of reasons. A number of alternative approaches, which may prove helpful in such circumstances, are discussed below.

2.2 Examples of "Management Procedures" for Short-lived Species

2.2.1 Management of the South African anchovy resource

The South African anchovy is a short-lived species, with only three age classes contributing to the spawning stock. The TACs for the fishery are set on the basis of hydro-acoustic survey estimates of spawning biomass and recruitment (Butterworth and Bergh, in press). A particular complication is that the bulk of the catch is taken from the recruits of the year, and much of this component of the catch may already have been landed by the time the recruitment survey estimate for the year becomes available.

TACs are set directly from the survey estimates by means of very simple formulae, which correspond roughly to a constant proportional harvesting strategy. The initial TAC set at the start of the season is based on the results of the spawning biomass survey, and assumes that the recruitment for the forthcoming year will be equal to its historic average level. However, only a proportion (some 70%) of this TAC may be taken before it is revised later in the season in the light of the result of the subsequent recruitment survey. This is to guard against the possibility of below average recruitment, which might otherwise result in the revised TACs desired falling below catch levels already achieved.

The control parameters of the harvesting strategy (e.g. the proportion of biomass to be harvested) are chosen on the basis of the results of simulations projecting the stock forward for a period of 20 years under the proposed strategy for setting TACs, and the associated survey programme which provides the estimates from which these TACs are calculated. These simulations are conditioned on a recent assessment of the resource, and take the imprecision of the estimates of population dynamics parameters into account. The control parameter value choices are based on a consideration of factors such as the anticipated average annual catch, the extent to which catches will vary from one year to the next, and the trade-offs between these and other measures of performance.

Essentially, the problem of catches being taken from the recruits of the year before a survey estimate of the recruitment strength becomes available, is addressed by consideration of the assessed distribution of historic recruitments; this allows the probability that the initial TAC is set higher than turns out to be appropriate to be kept low. TACs are also subject to constraints intended to facilitate the smooth operation of the industry; e.g. the maximum decrease in TAC allowed from one year to the next is 25%, with account being taken of such constraints in the 20-year projection calculations.

Naturally, the whole process of choosing this "management procedure" (see Section 5.4) relies on the assumption that the TACs indicated will be adopted and enforced each year. In general, simulation tests of management procedures need to include tests of actual catches exceeding the TACs if this is a problem.

2.2.2 Management of the fishery on capelin in the Iceland-Greenland-Jan Mayen area

The capelin stock in the Iceland-Greenland-Jan Mayen area is a short-lived stock, maturing at ages 2-3 in the autumn and spawning in March (at ages 3-4). The spawning mortality is believed to be almost 100%.

The 2-3 group generally feeds in the northern part of the region, between Jan Mayen and Iceland, and starts on a return spawning migration in early autumn, appearing at the northern coast of Iceland in September-October. From there, the spawning migration to the south and southwest coast of Iceland begins in December-January. Most of the fishing takes place during the months October-March and is concentrated on the mature part of the population (Anon., 1993a). A part of the fishery takes place earlier, mainly in August, but this is also aimed at the 2-3 group.

Since the capelin is a migratory species, management is in accordance with international agreements which are binding to the parties involved. The management system is based on an aim to leave a minimum of 400,000 t spawning biomass at the end of the season.

When the maturing capelin migrates up to the northern coast, it sometimes mixes with juveniles. This leads to problems since the juveniles are recorded on the acoustic equipment of fishing vessels, but escape through the purse-seine. The effect of capture and escape in terms of mortality is completely unknown, but may potentially become high when repeated catches are made at the same location. Local management in Iceland, therefore, uses closed areas and time periods in order to reduce the catches of juveniles.

The stock estimate is obtained using acoustic surveys which usually take place in October and January. These surveys have proved to be internally consistent in most cases, with deviations of less than 5% (in numbers) between the January survey and the predicted January estimate based on the October survey. Any exceptions to this seem to correspond to years when the autumn survey was an underestimate and noted as such in survey reports. In general, the January survey thus seems to be the most reliable estimate available, but of course this is in the latter part of the season. The following management system has, therefore, been adopted. The system is based on the assumption that the acoustic estimates of maturing capelin are absolute stock estimates.

For a given season, August-March, a precautionary TAC needs to be set in order to enable an opening of the fishery in those years when capelin are abundant. This is done using a simple regression method connecting the acoustic estimates in one year to the estimates of the corresponding year classes from the year before, accounting for processes in the intervening period. The regression thus provides a way of obtaining estimates of the TAC which can be taken from the stock, leaving 400,000 t to spawn.

Since the prediction is quite variable, the precautionary TAC is reduced by roughly 30% from the predicted value. This corresponds to the maximum historical deviation between the predicted value and the final stock estimate.

Having obtained a precautionary TAC, the fishery can be opened in August. The October acoustic survey then yields a stock estimate which is usually satisfactory as a basis for the TAC for the entire season. In some circumstances weather or ice prevent completion of a satisfactory survey, in which case a repeat survey is needed. In any case, various pressures usually necessitate a second survey in January. This usually confirms the former estimate.

2.3 The Anchovy Fishery in Sub-area VIII

The proportions of the annual catch taken from this resource are roughly as follows:

Jan - Mar	10% juveniles and mature fish.
Apr - June	60% mature fish (spawning takes place during this period).
July - Sept	20% mature fish.
Oct - Dec	10% juveniles and mature fish.

An acoustic survey takes place each April just before spawning, followed by an egg survey in June.

The key problem is that any TAC set at the start of the year cannot take account of the size of recruitment the previous year, because the forthcoming April hydroacoustic survey provides the first estimate of that recruitment (Anon., 1992a).

A management scheme similar to that for the two fisheries described above seems possible for this case. An initial TAC would be set (conservatively) in January based on the previous year's survey results and the catch taken subsequent to these surveys. This would be updated as soon as the results of the April hydroacoustic survey become available. Clearly, the efficacy of such an approach depends critically upon associated administrative procedures. Unless TAC revisions can be adopted and announced quite soon after the survey results become available, the initial TACs have to be set rather conservatively. The values of the parameters of the equations linking the survey results to the TACs to be set would be evaluated by conducting simulations of the application of such management procedures to the resource over a certain time frame, and considering the anticipated results. These projections would need to be based on a recent assessment of the resource: a key aspect of this assessment exercise would be the estimation of summary statistics of the distribution of historic recruitment levels. They would also need to take account of the anticipated level of precision of future surveys.

2.4 Some Possible Alternatives to a Full Age-Structured (VPA) Assessment

Annual catch or even catch-at-age data are generally insufficient to allow a satisfactory assessment of the stock in the absence of additional information. Essentially, a time series of an index of relative abundance (at least) is a pre-requisite, although some inferences can be drawn given only a simple estimate of abundance provided that this is available in absolute terms. For example, Beddington and Cooke (1983) provide tables which relate an initial catch level to such a survey estimate, as a function of biological and technological parameter values (natural mortality, growth rate and age at first capture). Their calculations take account of recruitment variability, and their results are expressed in relation to the probability of (unintentionally) reducing the stock below a specified threshold within a certain period.

In circumstances where catch-at-age data are not available, or their level of precision is such that VPA methods are unable to perform adequately, some variant of a "dynamic" (or "non-equilibrium") production model may provide a superior alternative. Some discussion on such models may be found in the report of the 1987 meeting of this group (Anon., 1993c). The simplest versions of such models use a single variable only to categorise the state of the stock (usually taken to be the recruited biomass, B) and have the form:

$$B_{y+1} = B_y + g(B_y) - C_y + e_y \qquad (1)$$
$$U_y = qB_y + \mu_y \qquad (2)$$

where

- $g(B_y)$ is the surplus production in year y (typically a function with two parameters to be estimated or a recruitment index),
- C_y is the total catch (by weight) in year y,
- e_y is the "process" error,
- U_y is the relative index of abundance for year y (e.g. CPUE, or the result of a survey),
- q is the catchability (a parameter that can be estimated), and
- μ_y is the observation error.

Estimates of the model parameters (q and two parameters for the surplus production function) are usually obtained by means of an "observation error" estimator, e.g., minimize $\Sigma\mu_y^2$ assuming $e_y = 0$. This approach may prove unsatisfactory, however, if recruitment fluctuations (represented by the e_y) are of comparable magnitude to the biomass B_y. Typically a longish time series (and some data "contrast" - see, e.g., Walters, 1986) of the relative abundance index is required to allow adequate estimation of the three parameters. This process can be facilitated if the U_y are measures of absolute abundance, in which case the catchability parameter q must be near to 1. Packages are available which implement such assessment models, e.g., "PC-BA" (Punt, 1992).

Reservations about such simpler forms of these models are that they fail to make any allowance for the age-structured nature of the stock (and the implications thereof for its dynamics), or changes in the exploitation pattern over time. These can be addressed by extending the production model to an "age-structured production model" (e.g., de la Mare, 1989; Punt and Butterworth, 1992; Hilborn, 1990). Equation (1) above is then replaced by equations incorporating the full age structure of the stock, and allowance is made for the age structure of the catch by means of a selectivity function (which may change in time) to disaggregate annual total catches by age, or directly if age-breakdowns of annual catches are available. Parameters are estimated in a similar manner to that described above, though now certain further information (e.g., recruit surveys) can be incorporated more naturally into this process. The parameters of the surplus production function are replaced by those of the stock-recruitment relationship assumed. Thus, in comparison to VPA, this approach replaces estimation of the recruitment for every year by estimation of stock-recruitment function parameters. However, bias may be a concern if recruitment fluctuations are of comparable magnitude to the total biomass. Further extensions of this approach allow some account to be taken of process errors (the e_y) - e.g., Francis *et al.*(1992); Punt and Butterworth (1992); Punt and Japp (in press) - but have a level of complexity which probably renders them inappropriate as potential "off-the-shelf" assessment tools. Few examples exist where these methods have been shown to be better than the full age-based methods. Stock-recruitment functions and process error models have been included in age-based assessment methods such as Cagean (Deriso *et al.*, 1985).

Another set of methods which can take partial account of age structure are extensions of the de Lury approach (Rosenberg *et al.*, 1990; Conser, 1991). For short-lived species these rely on the availability of an index of abundance during the course of the season such as commercial CPUE, which enables the size of the resource to be assessed from an estimate of the rate of decline in the index induced by the fishery. Under a real-time management system, the fishery may then be closed when the stock size is estimated to have fallen below a threshold level.

2.5 Conclusions

Although conventional catch-at-age analysis such as VPA may not be the best assessment method for short-lived species, it may nevertheless be of some use in the absence of better techniques. A variety of problems may make such analyses poorly suited for an annual TAC management regime for short-lived species. This may include the lack of recruitment indices, the need for in-season management, etc. It would be desirable, therefore, to pursue actively some of the alternative methods discussed above to develop an assessment methodology which could be more readily applied in a revised management system.

In particular, the Working Group recommends that:

1. Catch-at-age analysis should continue until replaced by alternative assessment methods.

2. Working Groups such as the Norway Pout and Sandeel Working Group should examine all available data, especially monthly or quarterly CPUE data in order to determine, e.g., relationships between CPUE and abundance which would enable alternative management methods to be applied. Procedures and programs corresponding to various models for the analysis of CPUE and survey data exist and should be investigated.

Finally, many of the problems noted for short-lived species arise mainly under a system of control by TACs and are much less severe under alternative management regimes such as effort control, which may be more appropriate in such cases.

3 ASSESSMENT METHODOLOGY

3.1 Shrinkage

3.1.1 Theoretical concepts

In classical statistics, shrinkage pertains to prediction in multiple regression (Copas, 1983). Predictions made by regressing on some explanatory variables can often be 'improved' by shrinking them towards the mean of previous observations. Essentially, this can be thought of as obtaining a satisfactory compromise between an estimator with potentially high variance and low bias (that based on the multiple regression) and one with low variance and potentially high bias (the mean).

Many other estimation problems also present the choice between one estimator with high variance and low bias and another with low variance but potentially high bias. Again, taking a suitably weighted average of these estimators can provide a satisfactory compromise between bias and variance and has generally become known as 'shrinkage'.

Fryer *et al.* (WP 7) illustrate the compromise between bias and variance in the following simple example. Suppose a random variable Y is related to an explanatory variable X by

$$y = \alpha + \beta x + \epsilon,$$

where α, β are (unknown) parameters and ϵ is a normally distributed error term with zero mean and constant variance σ^2. Given n observations (x_i, y_i), $i = 1 \ldots n$, we wish to predict the expected value of Y at $X = x'$; namely

$$y' = \alpha + \beta x'.$$

Two possible estimators of y′ are

$$y_{LS} = \hat{\alpha} + \hat{\beta} x',$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the least squares estimates of α, β, and

$$y_{AM} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

the arithmetic mean of the y_i. The estimator y_{LS} is unbiased, whereas y_{AM} is generally biased. However, y_{LS} has a larger variance than y_{AM}.

One way of combining bias and variance is to consider the mean square error. Now

$$\begin{aligned} &\mathrm{MSE}[y_{AM}] \leq \mathrm{MSE}[y_{LS}] \text{ if } \tau^2 \leq 1 \\ &\mathrm{MSE}[y_{AM}] \geq \mathrm{MSE}[y_{LS}] \text{ otherwise} \end{aligned}$$

where

$$\tau^2 = \frac{\beta^2 S_{xx}}{\sigma^2}$$

and

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

It is convenient to think of τ^2 as the signal-to-noise ratio of the regression of y on x. Thus, to minimise mean square error, we should use y_{AM} if $\tau^2 \leq 1$ and y_{LS} otherwise. It is important to note that neither estimator is optimal for all values of τ^2. (Of course, a practical consideration is that the true value of τ^2 is unknown).

A third estimator - the 'shrinkage' estimator - is a weighted average of y_{AM} and y_{LS},

$$y_{SH} = (1 - \theta)y_{AM} + \theta y_{LS}$$

where $0 \leq \theta \leq 1$. This estimator includes y_{AM} and y_{LS} as the special cases $\theta = 0$ and $\theta = 1$ respectively.

Taking

$$\theta = \frac{\tau^2}{1 + \tau^2}$$

minimises the mean square error of y_{SH}, and is such that

$$\mathrm{MSE}[y_{SH}] \leq \min\{\mathrm{MSE}[y_{AM}], \mathrm{MSE}[y_{LS}]\}$$

for all values of τ^2. Note that

- when τ^2 is large - i.e. large signal-to-noise ratio - θ is close to 1 and y_{SH} is close to y_{LS},
- when τ^2 is small - i.e. small signal-to-noise ratio - θ is close to 0 and y_{SH} is close to y_{AM}.

The mean square errors of the three estimators can be written as

$$MSE[y_{LS}] = \sigma^2 \left(\frac{1}{n} + \frac{(x' - \bar{x})^2}{S_{xx}} \right)$$

$$MSE[y_{AM}] = \sigma^2 \left(\frac{1}{n} + \frac{(x' - \bar{x})^2}{S_{xx}} \tau^2 \right)$$

$$MSE[y_{SH}] = \sigma^2 \left(\frac{1}{n} + \frac{(x' - \bar{x})^2}{S_{xx}} \frac{\tau^2}{1 + \tau^2} \right)$$

and are shown in Figure 3.1.1 as a function of τ^2.

In practice, θ must be estimated from the data, and this causes some problems. In particular, the mean square error of y_{SH} is inflated, because it now includes some extra variability due to the estimation of θ. Consequently, y_{SH} is rarely the optimal estimator of y' for a particular value of τ^2. However, y_{SH} is generally 'close' to optimal for all values of τ^2, whereas y_{AM} and y_{LS} are sometimes 'far' from optimal.

In the example above, θ does not depend on the value of x' (ie the value of x used to predict Y). However, this is not always the case. For example, if there are errors in the explanatory variables, then the optimal value of θ decreases (ie more shrinkage than if there are no errors in the explanatory variables) and the appropriate value of θ depends on x'.

It is important to note that shrinkage estimators provide a compromise between bias and variance - ie bias not too big, variance not too big. However, the method of application will depend on the problem under consideration. Further, determining the level of shrinkage will depend on the quality and type of data available.

Of course, it is always possible to take a weighted average of two estimators, regardless of their biases and variances. Assuming the weights are not a function of the data, the resulting estimator has a bias that is a weighted average of the original biases and a variance that is less than or equal to the maximum of the two original variances. Whether this is a sensible thing to do depends on the problem in question.

During the meeting, the theory developed by Fryer *et al.* (WP 7) was extended to the case of Laurec-Shepherd tuning with one effort series. Although the results are extremely tentative, approximate CVs were estimated for age 3, 4, 5 Western Channel Sole for a retrospective analysis running from 1979 to 1989. These 'theoretical' CVs (Table 3.1.1) were generally between 0.3 and 0.4 for ages 3 and 4 and between 0.2 and 0.3 for age 5, and were similar to the 'optimal' CVs found by the retrospective analyses described in Section 3.5.

The results suggest that optimal CVs will vary with stock, age and effort in the most recent year. Further, the amount of shrinking is likely to increase with both errors in effort data and errors in catch at age data, particularly if the estimates of effort and catch in the most recent year are poor.

3.1.2 Shrinkage in VPA tuning

The CPUE data from a fleet can be related to the VPA results by

$$u_{ay} = q_a + c_{ay} - f_{ay} + \epsilon_{ay}$$

where the small letters denote log-values of the respective variable.

The values of f_{ay} are obtained from VPA and treated as exact. The catchabilities are estimated from the observed values for y=1,2,...,t-1. By inserting the estimated value of q_a the fishing mortality rates in the last year are obtained from the observed values as

$$f_{at} = q_a + c_{at} - u_{at}.$$

A simple model of the fishing mortality rates is

$$f_{ay} = f_{a0} + \delta_{ay}$$

where f_{a0} is a constant value and δ_{ay} residual. The values of f_{a0} can be estimated and a weighted average of f_{a0} and the estimate obtained from the observed values in the last year has a lower mean square error than the estimate from the observed values alone. The optimal weights depend upon the ratio between the variance of the residuals ϵ_{ay} and δ_{ay}, respectively (Gudmundsson, WP 8).

The variance associated with CPUE data is often high so shrinkage could potentially improve the estimation of terminal Fs considerably. Shrinkage can also improve the estimates if more than one set of fleet data is included or if more sophisticated methods are applied such as the extended survivors analysis. However, as the influence of measurement errors in the final year is reduced, the optimal weight of the estimated average would be lower.

Misspecification of models leads to systematic errors in estimates derived from them. With the CPUE data the main risk is usually that catchabilities change systematically over time. The introduction of shrinkage reduces the effect of such misspecifications.

On the other hand, shrinkage produces systematic errors if the assumption of a constant mean of the fishing mortality rates is wrong. In the time series method this assumption is tested against more general models and is in fact rejected for a large proportion of actual stocks (Gudmundsson, 1987; in press). This does not imply that shrinkage should be abandoned, but that the mean values of the fishing mortality rates should be estimated only from values in the most recent past. As a result of this it is difficult to estimate the variance from the data, but the problem is examined empirically in Section 3.5.

Notice that in the statistical literature the word shrinkage is applied to prediction of future values of the dependent variable (corresponding to u_{ay} above), but it also has the effect of moving the value predicted by straightforward application of the regression towards an estimated mean value. In this context shrinkage is rarely useful with less than three independent variables whereas the optimal weight attached to an estimated mean of the fishing mortality rate would generally be reduced by adding a new set of CPUE data.

3.1.3 Shrinkage and time series analysis

In the time series method, the catch-at-age values are treated as dependent variables and the equation

$$C_{ay} = [F_{ay}/Z_{ay}][1 - \exp(-Z_{ay})]N_{ay}\exp(-Z_{ay}) + \epsilon_{ay}$$

is combined with a time series model of $\log F_{ay}$ which does not assume a constant mean value.

The catch at age observations contain information about changes in fishing mortality rates, even in the last year. There is a "shrinkage" effect similar to that described for the VPA tuning methods. In the time series method it weighs the indication of changes in the last year from the catch at age data against predictions of the Fs from the time series model. The time series predictions are conservative, but usually they are not as simple as the "same as last year" or the "same as the average in the past". CPUE data can be included in the analysis, but they are not indispensable for detecting changes in the last year. The parameters of the time series model are estimated from the data. (Gudmundsson, 1987; in press).

As an example of the ability of the time series method to estimate sharp changes of fishing mortality in the last year the Working Group used a simulated data-set with very large changes with time, jointly for all ages, copied from a simulated data-set from Fournier and Archibald (1982). Random variations with standard deviation 0.1 were added to the log Fs and the standard deviations of the catches were also 0.1 for the best observed ages and higher for the oldest fish. The results are presented in Figure 3.1.2. For comparison, untuned Xsa4 was run with shrinkage and the retrospective results are given in Figure 3.1.3. (This is an exceptionally unfavourable data-set for that method.) Note that only the last two years for each retrospective assessment are given for the time series method.

3.2 Recruitment Estimation

To complement the theoretical investigations on methods for combining several recruitment indices (Rosenberg *et al.*, 1992, Gudmundsson, WP9), and on whether shrinkage should be used (Fryer *et al.*, WP 7), an empirical evaluation of the various regression methods and options available to working groups for recruitment estimation has been carried out on actual data. This was made in the form of a retrospective analysis of how year-class strengths predicted by the RCT3 program compare with the VPA estimates obtained in the most recent assessments, i.e. using a similar approach to that used for retrospective evaluation of tuning methods.

3.2.1 Retrospective analysis

The RCT3 program has been run on a selection of data-sets known to have been somewhat problematic, using each of the three regression methods available in that software, viz., calibration, predictive and functional regression (the latter is implemented but not explicitly proposed to the user), with or without shrinkage. The other options proposed by the program were adopted consistently across runs for each stock and generally were the proposed defaults, such as a CV of 0.2 of the VPA mean for shrinkage, minimum of 3 points in regressions, no exclusion of surveys with small variance, etc. However, most of the time series used in this comparison were rather short and, except for the North-east Arctic cod data, no time taper was considered.

Whichever regression method is used, the RCT3 software performs a log-transformation on both the VPA and the recruitment indices, i.e. it fits a power rather than a linear relationship, in which the power (called "slope" in the outputs) is expected to be close to one. An important point to notice is that the recruitment indices taken from the relevant working group reports sometimes had to be rescaled in order to become significantly larger than the constant 1.0 added to them by the program prior to log-transformation. One should not be surprised that the data and results here may differ from those in the reports.

The closeness/discrepancy between RCT3 and VPA estimates for each method can be examined graphically (albeit with difficulty) and has also been measured by the root mean square logarithm of the ratios (RCT / VPA) over the years in which both VPA and RCT estimates are available; these "scores" are given in the bottom row of the tables of results. The smaller the figure, the better the RCT estimates by the method considered match, on average, the recruitments eventually obtained by VPA.

Due to a limitation of the spreadsheet software used for plotting the results, only 5 options could be graphed in addition to VPA. Thus, results from the predictive regression with shrinkage had to be omitted from the figures and are only given in the tables.

Western English Channel Sole (Division VIIe)

Six series of indices from 4 surveys are available for the period 1978-1991, and VPA estimates of 1-year-olds are thought to be sufficiently converged prior to 1989 (Table 3.2.1). Since survey indices are scarce in the earlier years, valid comparisons can only be made for the 1984-1988 year classes, but it was felt of interest to include the VPA estimate of the apparently strong 1989 year class in the comparison. Note, however, that RCT3 did not use that estimate in fitting the regressions. The results are given in Table 3.2.2 and Figure 3.2.1.

Over this short time series, all methods track the changes in recruitment rather well and, although they all have a slight tendency to overestimate recruitment, one cannot conclude that there is a systematic effect. The most extreme results are given by the calibration method without shrinkage, but the lowest score is obtained by the predictive regression with shrinkage which cannot deal with the abrupt changes observed in that stock, for reasons discussed below. The best score is for the functional regression without shrinkage. The differences between the shrinkage and no-shrinkage options are largely attributable to the fact that the surveys are poorly correlated with recruitment (one actually takes place in the northeast part of the Channel), so the mean is often given a predominant weight in the final estimation.

Irish Sea Plaice (Division VIIa)

Eight series of indices from 4 surveys are available for the 1974-1991 year classes, and VPA estimates of 1-year-olds up to the 1988 year class are used in the regressions (Table 3.2.3). Several of the surveys used are rather poor indicators of recruitment as indicated by their very small r-squares and, when the shrinkage option is turned on, the mean usually receives the largest weight.

The 1980-1988 year classes are considered in the comparison tests (Table 3.2.4 and Figure 3.2.2). Here again, the calibration method without shrinkage gives the most extreme variations, although it performs slightly better in terms of root mean square log-ratio than both methods using predictive regressions. The best scores are obtained by the calibration with shrinkage and the functional regression without shrinkage. It is noteworthy, however, that most methods overestimated recruitment of the 1984-1988 year classes, although they all detected the drop in 1986-1988. Surprisingly, the calibration with shrinkage does better in that respect than the predictive regression without shrinkage.

Icelandic Cod

Recruitment data are available from commercial CPUEs of age-3 fish from 1983-1991, and for ages 1-4 from surveys carried out since 1984 (Table 3.2.5). Sufficiently converged VPA estimates of 3-year-olds are available for the 1980-1987 year classes and, to allow for a minimum number of points in the regressions, comparisons can only be made for the 1984-1987 year classes. Caution is warranted in interpreting such a small set.

The results (Table 3.2.6 and Figure 3.2.3) conform with expectation, namely: the methods involving shrinkage respond to variations in recruitment but with some delay (particularly for 1984-1985), and the calibration without shrinkage exaggerates the fluctuations. For the 1987-1990 year classes, all methods are fairly consistent but, for the 1991 year class, the "shrunk" methods predict recruitment to be about average whereas the other methods indicate a sharp decrease. The scores probably do not make much sense here. They indicate, however, that the functional regression without shrinkage performs best, followed by the predictive regression without shrinkage. If there has been a problem with the recruitment estimation for this stock, it may have arisen because all indices used by the Working Group were small compared to the constant "1" added by the program, resulting in a very weak signal on the log scale.

North Sea Herring

During its 1991 and 1992 meetings, the Herring Assessment Working Group for the Area South of 62°N experienced some problems with recruitment forecasts, particularly due to differences in regressions using raw or log-transformed survey indices. The problem was further addressed by the Workshop on the Analysis of Trawl Survey Data (Anon., 1992d), but the emphasis at that meeting was on the comparison of the standard IYFS index with various elaborations of this index, and all the evaluations were made on log-transformed data. As stated above, the RCT3 program systematically performs a log-transformation of data on both axes, so the current exercise is of little relevance to the issue as it emerged initially.

Nevertheless, a data-set (Table 3.2.7) was compiled for 1-ring (read: 2-group) herring using the IYFS indices (means of all rectangle means) and VPA estimates for the 1980-1990 year classes given in the report of the 1992 Herring Working Group (Anon., 1992e) (n.b.: the index for the 1990 year class in that report differs from the figure used by the Trawl Survey Workshop, but this has no importance for the present purpose). The year classes prior to 1980 were not included since the survey procedures were not completely standardized at that time.

The "herring problem" is clearly reflected in the results (Table 3.2.8 and Figure 3.2.4) which show a rather large discrepancy among methods and with VPA. All methods fail adequately to match the drop in abundance of the 1985-1987 year classes and the upsurge of the 1988 year class shown by VPA, although the estimate of the latter is still uncertain due to poor convergence of the VPA (*cum* F < 0.6). Moreover, the methods involving shrinkage missed the large 1985 year class, and it is no surprise that their scores in terms of root mean square log ratios are the poorest overall. The methods without shrinkage have similar scores.

Northeast Arctic Cod

As documented in the report of the relevant working group, recruitment indices for this stock are available from a number of surveys carried out over a variable range of years during the period 1955-1992 (Table 3.2.9; note the rescaling). These were regressed against VPA estimates of the 3-year-olds from the 1957-1986 year classes, and comparisons were made with RCT3 predictions for the 1972-1991 year classes (Table 3.2.10 and Figure 3.2.5).

Although they sometimes depart from the VPA estimates, all methods give fairly consistent estimates of recruitment over the period. This is reflected in their scores which are all similar and can be taken to be sensible with such a long time series, in contrast with the previous examples. It can be noted, however, that the methods involving predictive regression perform slightly worse than the others.

3.2.2 Summary and conclusions

It appears quite difficult to draw any firm conclusion from these comparisons since the way in which the various methods perform depends not only on their intrinsic properties, but also on specific features of the data to which they are applied. Thus, no method seems to come out as universally better than the others. A tentative way of summarising the results is to tabulate the ranks, in terms of increasing root mean square log ratios, that each method achieved in each of the cases examined, as presented in the text-table below:

Stock/ Method	CAL+ SH	CAL- SH	FUN+ SH	FUN- SH	PRE+ SH	PRE- SH
VIIe Sole	4	3	5	1	6	2
VIIa Plaice	1	4	3	2	5	6
Icelandic Cod	4	3	5	1	6	2
North Sea Herring	5	2	4	1	6	3
NE Arctic Cod	1	4	2	3	6	5
Overall	15	16	19	8	29	18

Great caution should be exercised in interpreting this table, since differences in rank may be disproportionate in comparison with differences in absolute values of the scores. In addition, all applications of shrinkage used a single common value (0.2) of the CV of the VPA mean, and the methods might rank differently if an appropriate value was used in each case. It is thus advisable to refer to the specific assumptions each regression method makes about the error structures in the variables.

- The calibration mode of regression assumes that the errors in the VPA estimates are negligible compared to the errors in the survey indices. This is often the case in recruitment estimation, in view of the generally large variance of survey results, unless VPA is badly affected by errors in the catch-at-age data due to poor sampling, aging errors or occasional misreporting, for example. Despite its reasonably good score, the calibration without shrinkage often produces rather extreme variations, and the results confirm that shrinkage is necessary when using calibration regression. It is noteworthy that calibration with shrinkage performed best in both cases where the data series were long enough to make the comparisons of some significance.

- The functional regression is relevant when both variables are subject to error, but one has to provide an estimate of the ratio of the respective error variances, which is not an easy problem. As currently implemented, RCT3 assumes that variances in log-VPA and log-indices are of similar magnitude. This is a very specific option in the large family of functional regressions, and its relevance for general application is questionable.

- The predictive mode of regression is the one which is most commonly used in other contexts. It assumes that most of the errors apply to the 'predicted' variable compared to errors on the explanatory variable. As mentioned earlier, this is probably not valid for the present problem in most circumstances. The OLS model does not deal explicitly with errors on the explanatory variable and, if these exist, they result in bias on the estimates of the slope, the effect of which is similar to that of shrinkage. Moreover, when several indices are used in that way, some shrinkage applies to the predictions inferred from each index and, when these are eventually combined, the mean contributes several times to the final estimate. It is, therefore, no surprise that in these examples the predictive regression with shrinkage performed worst, as the sort of "two-stage shrinkage" makes it unable to match sudden changes in recruitment.

Obviously, this exercise has been based on a very restricted set of cases and some caution is called for. However, the results are consistent with previous conclusions from this Working Group in 1987 and by Rosenberg et al. (1992), that calibration with shrinkage is the preferred method among the class of regression estimators. It is recommended that working groups routinely evaluate the performance of their recruitment estimations using the retrospective analyses facility which has always been available in RCRTINX2 and RCT3, just as they do for retrospective evaluations of VPA tuning.

3.3 Integration of Recruitment Estimation and VPA

At present most ICES working groups use VPA for the estimation of current population size for all age groups except a few of the youngest, and a separate regression method for estimating the abundances of the youngest, recruiting age groups. The regression estimates are generally used in preference to the VPA estimates for the recruits because the latter are usually based on poorly sampled catch-at-age data which are known to be unreliable. Nevertheless, in some cases the VPA-based estimates may have some utility, and it would be preferable to include them in the estimation procedure with a weight appropriate to their precision, rather than to ignore them completely. This is especially true for intermediate ages where both estimates may have comparable precision. Whilst it is possible to deduce usable variance estimates from the standard program outputs, and carry out the combination calculation manually, this is rarely done, and a less labour-intensive method would be desirable.

There is no difference in principle between the methods used for tuning the VPA and recruit index analysis, as both are based on a calibration regression model: the procedure for the adult ages simply assumes a log slope of one (constant catchability) for the older ages, whilst the slope is estimated for the recruits, allowing catchability to vary with stock size in a density dependent manner. The problem of combination is thus handled gracefully by more recent methods such as XSA and ADAPT, which allow the incorporation of recruit index and survey data, and permit the appropriate model to be used for the different age groups. This ensures that all data are used once only (avoiding possible duplication which may otherwise occur), and eliminates the combination problem since all the estimates are made and used together in a consistent manner (Shepherd, 1991a).

The Working Group considers that the use of such a combined estimation is far preferable to maintaining the tradition of separate VPA and recruit index estimation, and to developing more elaborate methods of *post hoc* combinations. The Working Group, therefore, recommends that assessment working groups investigate the XSA option in the VPA tuning package for this purpose.

It is, however, very important that working groups continue to scrutinise the analysis of the data, especially for the recruiting ages, very carefully. Further enhancements to the XSA output are in hand to assist this process. It may also be instructive to continue to use RCT3 for diagnostics. It should also be noted that in order to allow all available data for pre-recruit ages to be included, working groups may have to extend the age range of the data file to include all the youngest ages, inserting zero catches in the catch number file as necessary. Natural mortality estimates for these youngest ages will also be required. Since these only provide an appropriate re-scaling of the estimates, however, this need not be the cause of too much grief.

3.4 Updating VPA with Recent Survey Data

At present the estimation of the current state of the stock is normally done with a tuned VPA procedure of some sort (i.e. including XSA and ADAPT). The VPA algorithm depends on the availability of catch-at-age data, so this procedure of calibration and estimation can only produce estimates of stock size up to and including the end of the last year for which catch data are available. Survey data which became available after that can only be used in an *ad hoc* way to update the assessment. In some cases this can mean that survey data cannot be utilised properly for up to a year after they become available, which is clearly undesirable, especially as the recent evolution of the stock is often a matter of considerable interest and debate. Some procedure for making proper and efficient use of recent survey data is, therefore, highly desirable.

Some methods based on direct maximum likelihood or least squares estimation such as CAGEAN (Deriso *et al.*, 1985), ADAPT (Gavaris, 1988), and ITCOTIO (Pope and Stokes, 1989) can, if necessary, be augmented to include the missing catch-at-age data as additional parameters, and thus be used in this way without much practical difficulty. It is not obvious that this is the best way to proceed, however, because it may simply lead to catches being computed which are consistent with the log-catch ratios from the most recent surveys. The recent fishing mortality estimates are thus wholly determined by the recent surveys, which might, therefore, just as well have been used directly. Any previous information on population size and fishing mortality has effectively been ignored. This is not quite what is wanted and, in the spirit of Bayesian priors and Kalman filters, one may well wish for something a little more refined.

In effect this means

a) using information on recent F-at-age values as a basis for prior estimates along with the new estimates,

b) using forward projections of the previous survivors at currently estimated rates of F along with the new estimates.

This may be done by the inclusion of extra terms in the maximum likelihood methods to represent these prior expectations. The procedure required, however, is very similar to that involved in applying shrinkage to estimates of mean F in VPA tuning. There should in fact be no difficulty in extending the tuning procedures to allow for the incorporation of more recent survey data in this way. The user would need to supply expected F-multipliers (varying with age if a mesh change had taken place) for the most recent years. These would be used to bring forward all available estimates in time, and thus to generate estimates of survivors in the usual way. Any discrepancies between the assumed Fs and those implied by the surveys would be apparent in the residual tables.

3.5 Retrospective Testing of Tuning Methods

At the 1991 meeting of the Working Group on Methods of Fish Stock Assessment (Anon., 1991) it was discovered that shrinking the predicted terminal fishing mortality towards the mean was quite effective in reducing the retrospective bias problem in the stocks on which it was tried, and that it also seemed to be useful in reducing random variation in the predicted F values. A theoretical explanation of the latter property is now available (Section 3.1) but the application of shrinkage to reduce retrospective bias is still an *ad hoc* procedure.

The Working Group was asked to investigate the appropriate use of shrinkage in tuning and advise how it should be implemented for the use of working groups. To address this question, retrospective analysis was carried out on two stocks which had proved troublesome at the 1991 meeting (Western Channel, Division VIIe sole, and NAFO Division 4VsW Cod) (Anon., 1991), and for the simulated data-set No. 5 from the Reykjavik meeting (Anon., 1993c). This was done using Laurec-Shepherd tuning and several variants of XSA, including that now implemented as part of the standard VPA tuning package as well as the time series-based method (TSER) (for the simulated data-set only). The VPA analyses were carried out with no shrinkage, and with the "cv" parameter (the log standard error specified to be attached to the mean F) ranging from 0.1 to 0.5 in steps of 0.1. It should be noted that low values of this parameter imply strong, and high values weak, shrinkage.

In all, more than fifty retrospective analyses were carried out for each stock, and the results for each run were summarized by a page of figures and one of tabulated results, as at the 1991 meeting (Anon., 1991). These results are too voluminous to include in the report, but examples are given in Tables 3.5.1 to 3.5.10 and Figures 3.5.1-3.5.10 of the unshrunk, overshrunk, and optimally shrunk results for each stock. The full set of results are summarized in Tables 3.5.11 to 3.5.16. These show the percentage of estimates at each age which appear to be in error by a log ratio of more than 0.5 ("outliers"), the mean log ratio ("bias") and the root mean square log ratio (i.e., r.m.s. prediction error, "s.e."). The latter estimates are as usual multiplied by 100 and may be regarded as percentage errors. The comparison for the real data-set is with the final run of the series, whilst for the simulated data-set it is with the "truth".

In the "basic" version of shrinkage implemented for the Laurec-Shepherd x XSA2 procedures, a CV for mean F is specified by the user. This may be referred to as "hard" shrinkage. Clearly, if the variability of F (at any age) is larger than that assumed, it would be appropriate to use the higher observed CV (and, therefore, to shrink less). This is here referred to as "soft" shrinkage.

Further, shrinkage has hitherto been applied only to the terminal F estimates (i.e., those for the last year and the oldest age). However, it is known that separable VPA, which determines terminal F and year class strength from estimates of catch at age and smoothed (separable) fishing mortalities over the whole cohort, is a relatively robust method of analysis, especially when the survey/CPUE data are of poor quality. This is effectively the logical conclusion of the shrinkage process (ignoring the tuning data and using just catch at age and some smooth model for fishing mortality). It is, therefore, possible to explore another "universal" flavour of shrinkage, in which population at age is estimated from catch at age and a smoothed (running mean) estimate of F for all ages and years, and these are included in the analysis in the usual way (i.e., as though they were estimates from surveys). This is easily done in the XSA method (but is not possible for the standard tuning methods). It could be regarded as taking a solution similar to that from separable VPA as a prior, and modifying it in the light of information from survey and/or CPUE data.

A summary of shrinkage terminology used in this report is given below:

strong: giving much weight to the mean, by specifying a small CV value.

weak: giving little weight to the mean, by specifying a large CV value.

hard: using the specified CV value only to determine the strength of shrinkage.

soft: using the observed CV (when higher than that specified) to determine the strength of shrinkage.

marginal: using shrinkage only on the terminal F values, at the margins of the catch-at-age array.

universal: using shrinkage to the mean F throughout the catch-at-age array, thus biassing the solution towards one with a slowly changing exploitation pattern.

The methods used were:

L/S: standard Laurec-Shepherd tuning.

XSA2: extended survivors, as previously tested at the Reykjavik (Anon., 1993c) and St. John's (Anon., 1991) meetings.

XSA4: the XSA4 variant of extended survivors analysis (not generally available) which shrinks to an exponentially weighted running mean F, and allows for hard or soft, and marginal or universal shrinkage options (HM, HM and SU options were tested, as hard universal shrinkage was already known to give extremely variable results on poor quality data).

The results indicate that weak shrinkage (CV ≈ 0.5) is not only adequate but preferable. Strong shrinkage (CV ≈ 0.1) can easily create a biased retrospective pattern in which the direction of bias is reversed. This is not at all surprising for stocks in which quite rapid changes of fishing mortality have taken place.

The conclusion and message for working groups is that weak shrinkage should be preferred, that a CV of 0.5 may be a sensible starting value, and that the version implemented in the standard tuning package is adequate. Although shrinkage is sometimes very effective in reducing both bias and variance, as perceived by a retrospective analysis, this cannot be guaranteed. Routine retrospective testing with several values of CV is desirable and has now been provided as a "push button" option in the VPA tuning package, for both L/S and XSA2 procedures. Files which can be read directly by the SAS tabulation and plotting routines now provided at ICES may be produced automatically. Working groups are urged to make use of these facilities, and explore the effect of shrinkage on their assessments, before making a choice. It is expected that weak shrinkage will be sufficient and preferred in most cases.

It should be stressed that shrinkage is a recognised procedure for reducing the variance of predictions. Where a retrospective pattern shows bias, shrinkage may still help, for example if the bias is in the terminal estimates. However, if it is the converged VPA estimates which are biased, shrinkage can make the situation worse. Thus, if shrinkage is used to "cure" retrospective patterns (as opposed to variability), the sources of bias should be investigated zealously to avoid aggravating an already serious problem.

These results confirm the conclusion of the Reykjavik Workshop (Anon., 1993c) that the simulated data-set No. 5 is not so "badly behaved" as many real data-sets. With methods now available this data-set can be analyzed with little difficulty to a high precision. The Working Group reiterates that new or improved methods should, as a minimum, be tested on this data-set, and should be discarded or amended if they do not perform well. More difficult standard simulated data-sets are required for future use.

3.6 Conclusions

The Working Group concluded that future work should emphasize development and testing of integrated methods for the entire assessment process, which include the years and age groups traditionally used in assessments along with younger ages (pre-recruits) and also the year following the last catch data year, if surveys exist for that year.

A low level of shrinkage in VPA tuning is found to be beneficial in most cases and rarely (if ever) detrimental. The effect of various options can now easily be tested retrospectively using programs (VPA and SAS) available at ICES.

4 STOCK-RECRUITMENT RELATIONSHIPS AND MBALs

4.1 Stock-recruitment, general

Several questions on the relationship between stock and recruitment were analysed using the databases compiled by Myers and co-workers (Myers *et al.*, WP 10) and P. Mace (unpubl.). These databases consist of estimates of spawning stock biomass, recruitment, catch, fully recruited fishing mortality, taxonomic information, life-history parameters, and some basic parameters relevant to fisheries management. Also included are units and the source and reference in an ASCII format that is readable by any programming language or statistical package. The data have been read into an Splus object for easy access. For the purposes of this meeting, the 72 stocks with at least 20 years of concurrent stock and recruitment data were used in the analysis. When the Myers database is finished in a few months, it will be freely available over Internet to all interested users. A summary and analysis of the database will soon be published as a technical report.

The Working Group agreed that it is very useful to have estimates of spawning stock biomass, recruitment and associated variables together in one database. The Working Group suggested that the data-sets should be provided to the various working groups for quality control. It was also suggested that the working groups provide comments on the reliability of each series. It is recommended that the results from assessments be kept in a standard format, and that a stock and recruitment database be kept up to date. It is suggested that an updated database be kept at ICES (see Section 4.7).

Several preliminary conclusions from the analysis were presented from work soon to be published by Myers and co-workers. These conclusions are:

A relationship between stock and recruitment commonly occurs

In general, there is a surprising amount of evidence for a relationship between stock and recruitment. A simple chi-square test (Table 4.1.1) indicates a significant relationship for a number of these stocks (20% of the stocks at the 5% significance level). This is a very weak test, and an examination of the parametric fits indicates much stronger evidence (Myers, in prep.).

Depensation is undetectable in most stock and recruitment data

The hypothesis was examined that fish populations can exhibit multiple stable states, and that they may collapse suddenly because of depensatory recruitment, i.e. increased mortality or lower *per capita* reproductive output at low stock sizes. Of the 105 populations examined, only one, Icelandic spring-spawning herring, showed statistically significant evidence of depensation; however, it was the only population studied that became commercially extinct. Previous empirical claims of the existence of depensation lacked a firm statistical basis, and this analysis indicates that such depensation is undetectable in the data on a large number of stocks. This result calls into question the theoretical claims that the collapse of fish stocks can be attributed to depensatory recruitment.

Overcompensation and stability

The hypothesis was examined that fish populations exhibit overcompensation, i.e. that recruitment is a decreasing function of spawning stock biomass at larger stock sizes and that the equilibrium population would be stable. To test overcompensation, the fit of the Shepherd (1982) stock recruitment model was estimated:

$$R = \frac{\alpha S}{1 + \left(\frac{S}{K}\right)^{\gamma}}$$

The model was re-estimated with the overcompensation parameter, gamma, constrained to be 1. A likelihood ratio test was used in a significance test. 105 stocks from the Myers *et al.* (WP 10) compilation of stock and recruitment data were used in the analysis. Thirteen of these showed statistically significant overcompensation.

Clear evidence of overcompensation of recruitment is apparent in the data; however, it does not appear to be generally a very important phenomenon for most marine fish within the stock levels that were observed. Overcompensation appears to be most common among species in which cannibalism of young can be an important source of mortality. However, it can also occur in species in which cannibalism appears to be rare and not important, e.g. herring.

The most serious difficulty in the estimation of the stock-recruitment function is due to the time series bias in the parameters caused by the dependence of the stock on earlier recruitment, i.e. large recruitment usually leads to large stock sizes (Hilborn and Walters, 1992). This type of bias is most important when there is little contrast in stock size; in these cases the estimates of stock and recruitment parameters must be treated with caution. This bias would tend to make overcompensation appear to occur more often than it actually does and would make populations appear more unstable than they actually are.

There is another bias in the estimation of the replacement line, i.e. the spawning stock biomass that would result from any level of recruitment at zero fishing mortality. It has been assumed that there is no density-dependent growth or mortality after recruitment to the fishery has occurred. This is clearly false for many populations (Millar and Myers, 1990), and would have the effect of making the replacement line curve upwards, i.e. one in which the first and second derivatives are positive. This would tend to make the population dynamics around the equilibrium point less stable.

4.2 Stock and Recruitment: Biological Reference Points for Fishing Mortality

Biological reference points such as F_{high} and F_{med} which are based on percentiles of the observations are liable to depend on the range of stock sizes which have been observed. Indeed, as pointed out by T. Jakobsen (pers. comm.), if all the observations relate to low stock sizes (among those stock sizes attainable by the stock), F_{med} may well be a better estimate of the fishing mortality which will lead to collapse, than of one which is sustainable. Similarly, if recruitment is very variable at low stock size, the estimate of F_{high} may be inflated above that which corresponds to eventual collapse (Sissenwine and Shepherd, 1987).

Thus, whilst these percentile-based estimates are easily derived and better than no estimate at all, it would be very desirable to deduce more rigorously-based estimates of dangerous and safe levels of fishing mortality. Ideally, the estimate of F at which collapse is to be expected, hereby denoted by F_{ult}, following the suggestion of Pope (WP. 6), should be deduced from the slope at the origin of a fitted stock-recruitment relationship. A comprehensive set of such estimates has been reported by Myers *et al.* (WP. 10). However, they find that dubiously high values of this slope are regularly produced when Beverton and Holt and Shepherd-type relationships are fitted, although the problem is less acute for fitted Ricker relationships.

Pope (WP.6) has suggested that there is merit in working with plots of the reciprocals of recruitment and SSB, especially if the values are normalised by estimates of the maximum (average) recruitment attainable (by the unfished stock) and the SSB corresponding to this recruitment with no fishing (the virgin biomass). On such plots the Beverton and Holt relationship is a straight line, passing through a predetermined intercept at (1,1). Only the slope of this line remains to be determined, and it yields directly an estimate of the maximum sustainable ratio by which biomass-per-recruit may be reduced (for which 20% has been suggested as a reasonably safe lower limit). The plot also emphasises and gives enhanced influence to points arising near the origin of the stock-recruitment plot, which is entirely appropriate. It may, however, do strange things to the error distributions associated with the observations so that robust procedures for fitting the stock-recruitment relationship are really needed. Finally, this presentation has the merit that data for different stocks are reduced to a common scale, and may be superimposed legitimately. Thus, if there is any common behaviour among species or families, it may be detected by such a presentation of the data. An example of this figure from Pope (WP. 6) is given as Figure 4.2.1. Pope further suggests that one half of F_{ult} would be an appropriate target for sustainable management, as he shows that it is guaranteed to yield at least half the maximum number of recruits (under either the Ricker or Beverton and Holt model). Thus, F_{ult} may provide a superior replacement for F_{high}, and $F_{ult}/2$ a replacement for F_{med}. These estimates should be relatively insensitive to the range of stock sizes observed in the historic data.

Finally, Pope suggests that in seeking threshold values like F_{ult}, it would be appropriate to determine values which apply to the worst run of years on record, and cites a biblical precedent for examining seven-year running means. This considerably reduces the scatter on the reciprocal plot, and his suggestion implies an appropriate non-parametric method for fitting the S/R line, since one may set it to pass through the least favourable point. In practice, the worst 90 percentile would probably be preferable.

4.3 Stock and Recruitment Analysis without VPA, and the Stability of F_{med} and Related Reference Points

In WP5 Pope investigates the observation (T. Jakobsen, pers. comm.) that the F_{med} biological reference point is more stable than might have been expected under changes of natural mortality (cf. previous discussion by the Working Group on the stability of reference points in Anon. (1985). To do so he derives a method for deducing indices of SSB and recruitment from CPUE or survey data only (i.e. without total catch or catch-at-age data). A similar procedure was used by Sparholt (1992) in a Working Paper to ACFM in 1992. These estimates are correctly dimensioned, but the scaling depends on an undetermined ratio of the catchabilities of recruits to that of mature fish. Furthermore, Pope shows that similarly scaled estimates of steady-state SSB-per-recruit may be derived from the same data, for any level of fishing effort represented in the data-set, including the current level. The techniques involve log-catch ratios and are closely related to those used by Shepherd and Nicholson (1991) for multiplicative modelling of such data.

In this way it is possible to determine whether current effort is higher or lower than that required to reach F_{med} (or any related measure deducible directly from the stock-recruitment plot in terms of SSB-per-recruit, including F_{ult} as defined above). This provides the essential signal required by managers, i.e. should effort be decreased, or may it be increased, together with a very rough estimate of the magnitude of the change, if it is legitimate to assume that SSB/R is roughly inversely proportional to effort over a small range.

The estimate of the reduction factor may be considerably refined if catch-at-age and effort data are available, still without performing VPA, and thus without requiring an estimate of M, and it is also possible to compute yield-per-recruit in the same way (and subject to the same scaling) over the range of effort data observed. Thus, potentially, the effort corresponding to F_{max} may also be determined if it lies within the range of effort observed.

Pope then shows that, although these estimates are constructed without adopting any value for M, changes in the perceived trend of F and recruitment at different levels of fishing mortality would cause errors in these estimates. These errors are analogous to those induced by a wrong choice of M. The F_{med}-type estimates are, however, relatively robust under a change of M because a cancellation occurs, whilst those for F_{max} do not benefit from such a cancellation and are much more sensitive to M.

This explains the observations of T. Jakobsen (pers. comm.), but does not of course remove the difficulty mentioned above of sensitivity to the ranges of biomass observed. Nevertheless, this work does show that it is possible to go much further than has hitherto been realized in providing conventional management advice even where much of the data normally required is lacking. These methods may be particularly useful where it is possible to carry out research surveys, but difficult or impossible to obtain accurate statistics from the commercial fishery.

4.4 MBALs (Minimum Biologically Acceptable Levels)

Myers *et al.* (WP 10) examine stock and recruitment data from 72 stocks (Myers, in prep.) and spawning stock biomass per recruit data (Mace, in prep.) to develop procedures for estimating critical spawning biomass thresholds. The selected data-sets each contained at least 20 data points. Eight methods for estimating the critical spawning biomass level were tried. Six of the methods relied on fitted stock-recruitment relationships, where the parameters were estimated using maximum likelihood assuming lognormal errors. Of these, two estimated the point where recruitment was 50% of its maximum value on the fitted Ricker and Beverton and Holt relationships, respectively. The other four estimated the critical point as 20% of the estimated virgin biomass obtained from the intersection of fitted stock-recruitment relationships or mean recruitment with the replacement line at F = 0. Two further methods based on the work of Serebryakov (1991) as detailed by Shepherd (1991b), were examined. One estimate was from the spawning stock biomass at the point where the 90th percentile of the survival ratio (R/S) intersected the 90th percentile of recruitment in each data-set and the other was the spawning biomass at the intersection of the 90th percentile of the survival ratio with mean recruitment.

In general, there will be no best method for estimating critical spawning levels for all stocks. However, a number of simple criteria can be used to determine if the estimated critical point is sensible. An important indicator is the linear slope of the log-transformed points (or preferably the linear slope assuming log-normal errors) above and below the estimated critical point. A simple decision diagram (Figure 4.4.1) has been developed to interpret these slopes. If the slope above the estimated critical biomass is positive, and the slope below is also positive, the critical point estimate is sensible if the slope above is less than the slope below. If the slope above is greater, the critical point is probably at too low a spawning stock biomass. If both slopes are negative, the estimated critical point is too conservative. Finally, if the slope above the estimate is positive and the slope below is negative, the data are pathological and the estimate is not sensible. Using these measures, the Serebryakov method using 90th percentiles often gives a sensible result, but can be very variable. Several of the other methods perform nearly as well and on a case by case basis may be better at times.

Another criterion is that recruitment below the estimated critical point should be on average lower than above the critical point. In other words, there should be an expected impact of allowing spawning stock biomass to drop below the critical level. The two methods based on the fitted stock-recruitment relationships and the associated point where recruitment is 50% of the expected maximum both do well with respect to this criterion when sensible estimates have been made based on the slope criteria. The results on this criterion for the other methods are more variable.

The ability of any of the methods to estimate a reasonable critical level is closely related to the range of observed spawning biomass. If the only observations are all near the origin, with little information on the level of maximum recruitment, then the critical level should probably be near or above the highest observed spawning biomass. On the other hand, if the slope of all the points is negative, and the range of data only covers high biomasses, then a reasonable critical level may be near or below the lowest observed spawning biomass. It is important to check the range of the data and, to put it on a common scale, it can be plotted as a proportion of the estimated virgin biomass.

Many of the methods are extremely sensitive to the range of the data observed. For example, the Serebryakov methods will give very different answers for the same stock if only the data corresponding to the lower half of the observed biomass range is used in the estimation. This is an undesirable property since it implies that the threshold level decreases as the average level of exploitation increases. On the other hand, the methods using the 50% of maximum recruitment point on fitted curves do not have this problem, but are often unable to obtain reasonable estimates.

Suggested Procedure for the Analysis

The comparative study provides some general guidance on procedures for estimating MBALs, but the analysis for any particular stock may result in one estimation method being preferred over another. Based on the comparative study there are a series of recommended steps which should be followed for the analysis of a given stock. Two example diagnostic sheets were prepared by the Working Group to illustrate the methods (Figures 4.4.2 and 4.4.3).

1) Shepherd, Ricker and Beverton and Holt stock-recruitment relationships should be fitted (using maximum likelihood or other statistical approach).

2) A replacement line at F = 0 should be calculated from the spawning biomass per recruit curve using growth, maturation and selectivity data. The replacement line is a line through the origin whose slope is given by the inverse of the spawning biomass per recruit at F = 0. Note that this replacement line assumes that growth and maturation schedules are constant over the entire range of spawning stock biomass including the "virgin" biomass level and usually requires extrapolation beyond the observed spawning stock biomass levels. These estimates of virgin biomass should be regarded as tentative and may only be appropriate for scaling the data.

3) The following MBAL estimates should be plotted on the stock-recruitment graph.

    a) the SSB at 50% of maximum recruitment on the fitted stock-recruitment relationship,

    b) the Serebryakov level using the 90th percentile of survival and the 90th percentile of recruitment,

    c) 20% of the virgin biomass estimated from the intersection of the replacement line and the fitted stock-recruitment relationship,

    d) 20% of the virgin biomass estimated from the intersection of the replacement line and mean recruitment.

4) Tabulate for each of these estimates the ratio of mean recruitment above and below MBAL. Estimation methods which show a clear reduction in recruitment below the MBAL are preferable. Graph the probability of recruitment in the upper and lower quartiles of the data over the range of observed spawning biomass.

5) Calculate the range of the data as a proportion of virgin biomass (using estimate d). A wider range of data provides a better basis for estimating MBAL. Note, however, that the estimate of virgin biomass should be viewed with caution because of the assumed stationarity in life history parameters extrapolated to higher stock biomass (see point 2 above).

6) Calculate and tabulate the linear slope of the data above and below each MBAL estimate assuming lognormal errors. Discard any slope estimates which include less than five data points. These slopes help one to judge whether the MBAL estimates are sensible as in Myers *et al.* (WP10). In general, if the slope below the MBAL estimate is positive and if the slope above is less than the slope below, the estimate is sensible. If the slope above is greater than the slope below, it is risky and if both slopes are negative it is over-conservative (Figure 4.4.1). The calculation of these slopes can be done over the range of observed biomass to indicate the most appropriate MBAL (Figures 4.4.2 and 4.4.3).

7) In general, if the slope over the entire range of data is negative, MBAL should be estimated at or slightly below the lowest observed spawning biomass. This should be interpreted in the light of the range of the data observed. In this circumstance it is expected that only relatively high biomasses with respect to the estimated virgin level have been observed.

8) If the slope over the entire range of data is positive and only relatively low spawning biomass levels have been observed then MBAL should be above the highest observed spawning biomass.

9) MBAL estimates should be chosen only if they are sensible with respect to point (6) above using the slopes of the data either side of the estimates. The estimate should be at a point below which recruitment declines or the probability of good recruitment declines using the calculations in point (4) above. The MBAL estimate should be considered tentative if only a small range of spawning biomass has been observed.

10) If the fitted stock-recruitment relationship(s) are reasonable, i.e. do not result in extremely high slopes at the origin or predict maximum recruitment well outside the range observed, the estimate as in (a) above is to be preferred because it will not be so dependent on changing data and because it has the best theoretical underpinning. Note that the (a) estimate should still be compared to tests 7-9.

11) If the (a) estimate is judged to be unacceptable, the choice between b - d should be made using the criteria listed above.

4.5 Caveats

Stock-recruitment relationships are almost always very noisy and, therefore, long-time series are required in order to detect such relationships. For this reason assessment working groups should regularly run a final VPA as far backwards in time as possible (although such long-time series should not be used for tuning purposes).

One consequence of the use of such long-time series is that important environmental changes may have taken place during the period. For example, the Icelandic cod data include data from before, during and after the period 1965-1970. That particular period was one of severe ice conditions which are likely to have had major effects on the ecosystem from plankton upwards.

Such effects must be kept in mind when stock-recruitment data are analysed.

4.6 Future Work

The estimation of stock-recruitment relationships is an important area for future work. The Working Group noted that inclusion of prior information to facilitate estimation of the slope at the origin is a particularly promising line of attack and should be considered in the near future. Some work along these lines done during the meeting gave promising results.

Some of the methods suggested in this section can easily be implemented using a ruler and a hand calculator. In order to implement maximum likelihood estimation of parameters within assessment working groups, programs need to be made available. A SAS program to estimate the parameters assuming lognormal errors in recruitment (i.e. using non-linear least squares on log R) would be adequate for this purpose.

4.7 Conclusions (Stock and Recruitment)

The analysis of stock and recruitment data carried out at the meeting showed that several of the novel methods tried out seem to be quite promising. In particular, the profiles of the probability of good and poor recruitment seem to capture many of the essential features, and may often assist in determining MBALs. The fitting of several stock-recruitment relationships paying proper attention to the error structure is also often informative.

The Working Group recommends that assessment working groups should regularly re-analyse stock and recruitment data and an example of a set of analyses which are likely to be useful are given in Figures 4.4.1 and 4.4.2. These analyses are commended to stock analysts as worth trying.

The Working Group wishes to record that the rather extensive analysis of these data would not have been possible without access to the database assembled and made available by R.A. Myers. The Working Group recommends that such data from ICES stocks be routinely collected and maintained in a standard format accessible on the Internet. The Working Group further recommends that the format be standardised with that used by other collection agencies and that data from as many other sources as possible (notably North America) also be kept in the ICES system.

5 MANAGEMENT ADVICE

5.1 Risk Analysis: Generalities and Indistinct Terminology

In the last few years there has been considerable interest in extending scientific advice for fisheries management to take proper account of the uncertainty inherent in assessments of the state of the stocks and uncertainty about their future course (e.g., because of unknown future recruitment). These uncertainties mean that the effects of various management options and procedures can only be determined in terms of probabilities. The presentation to managers of the probability of various outcomes in terms of the state of the fishery and the resource for different management scenarios has been generally referred to as risk analysis, and several recent scientific meetings have been devoted to this subject (NAFO, 1991; Canadian Department of Fisheries and Oceans, 1992; Alaska Sea Grant, 1992). In the U.S., the National Marine Fisheries Service has set up a Risk Assessment Working Group to pursue the topic, and a

Theme Session will be devoted to risk analysis at the 1993 ICES Statutory Meeting.

Several contributors to the Methods Working Group and to its e-mail conference prior to the meeting had pointed out that there is a substantial literature on decision theory and risk analysis which defines "risk" as the expected loss for a specified loss function. A loss (or conversely a utility) function quantitatively expresses the value a manager gives to different attributes. In a fishery context, a loss function may include the yield foregone compared to some reference level, spawning biomass or recruitment foregone compared to a reference level or economic measures of fishery performance over specified time scales. In decision theory, "risk" can only be evaluated when a loss or utility function has been chosen or, in other words, after the relative importance of attributes of the fishery and resource have been quantified.

In contrast to the decision theory approach, the term "risk" has been used more loosely in fisheries to mean "the probability of something bad happening" (Francis, 1991). Risk analysis under this usage would consist of the calculation of probabilities associated with stock abundance falling below a specified level within a given time period, for example. In other words, risk is the probability of an adverse event or a mishap resulting from management action (or inaction). The calculation of the probability of a mishap is, of course, conditional upon the type of variability included in the calculation and the risk is conditional upon the unspecified importance a decision maker (manager or scientist) attaches to the mishap.

The Working Group supports the need to avoid confusion in terminology, and in general agrees that it would be desirable for scientific advice for managers to go beyond simple calculation of the probability of a mishap in developing risk analyses in future. In some circumstances, it may be possible to conduct a formal analysis of expected loss with a specified loss (utility) function or functions or for a number of likely loss functions. This is to be encouraged. However, in other cases, it may only be feasible to present the probabilities of mishaps with respect to a number of measures of fishery and resource status without specifying tradeoff between them. It is quite clear that no rigid terminology for "risk" and "risk assessment" is likely to be widely accepted and fishery scientists in ICES will continue to use these terms to mean the probability of a mishap. As long as the advice given continues to evolve in such a way that methods are developed for incorporating uncertainty in the advice in a useful way that is interpretable by decision makers, the strict terminology is of secondary importance.

The Working Group found it convenient to refer to diagrams of probabilities of mishaps and benefits as "probability profiles" and recommends this usage. When presenting such diagrams, the conditional nature of the probabilities should be clearly stated; also, what time scale they apply over, what sources of variability have been included and what assumptions have been made. There are usually five types of uncertainty to be considered:

1) uncertainty in the parameters defining the current state of the stock.

2) uncertainty in the future values of relevant quantities such as recruitment.

3) uncertainty in specification of the models and their parameters defining the population dynamics (natural mortality rate, growth rate, etc.).

4) uncertainty in future assessments of the state of the stock.

5) uncertainty arising from imprecise management controls (TACs being exceeded for example).

In principle, it is desirable that all of these should be taken into account in a risk assessment, but it is still useful to perform the analysis on a subset in many cases. At present, few analyses include all sources of uncertainty, but many are available which describe only measurement error (1) or recruitment (2) uncertainty. These should be regarded as an important step forward and are to be encouraged rather than criticized for leaving some factors out.

Another important point is that assessments of the probabilities of mishaps, and the losses associated with them, are important in their own right. While a logical next step may be to minimize expected losses to determine a "risk averse" strategy in a decision theoretic approach, it may be more useful in giving scientific advice to managers to evaluate the contributions to overall loss (or utility) separately or, in other words, to produce probability profiles of important attributes. The Working Group considers that a major effort is required to implement methods for probability profiling as a regular component of stock assessments.

Probability profile studies can take two forms because there is a concern both for the short-term possible outcomes of management measures and the long-term robustness of management strategies.

5.2 The Short-Term Possible Outcomes of Management Measures

It is well recognized that the results from fisheries assessment methods suffer from three sources of uncertainty. There is imprecision in the data, uncertainty in the model assumptions and variability in the estimates (the model predictions do not match the data precisely). Probability profiles can show the effects of this uncertainty on the distribution of parameter estimates. For example, the distribution of projected catch values reflects the probability that the fishing mortality resulting from a given total catch will be greater than the target fishing mortality.

The calculation of probability profiles can be done in many ways. Analytical methods calculate the distribution of the parameter estimates by assuming a distribution shape for the errors in the input data and calculating an approximation of the resulting non-linear estimation variance and covariance. Parametric resampling methods (Monte-Carlo) also assume a distribution for the variability in the input data but get the distribution of the estimated parameters by generating random realizations from the assumed distribution of the input data. Non-parametric re-sampling methods produce new random input data by adding to the corresponding predicted value an error selected randomly from the original empirical distribution of the residuals. Many combinations of these methods are possible (Restrepo *et al.*, 1992).

5.2.1 Sensitivity Analysis

Sensitivity analysis is not directly related to the question of estimating probability profiles but is relevant in so far as it considers the effect of uncertainty in model input (the parameters) on model output (or state variables). Calculating the probability profile of projected catch, for example, can be thought of as estimating how much the effect of uncertainty in the current population affects the predicted catch (the state variable). A particular method, the Fourier Amplitude Sensitivity Test (FAST) (Cukier *et al*, 1978) seems appropriate in this context. The analysis requires that for each parameter (eg fishing mortality or population size) a range of uncertainty be defined. The method then chooses sets of parameter values from the uncertainty ranges according to specific criteria so that a full range of combinations of parameter values is sampled. The choice of parameter sets is such that the amount of variability in the state variable can be partitioned analytically into the contribution from each parameter. Thus it is possible to determine which parameters contribute most to the variability in the state variable. By choosing uncertainty ranges which correspond to the variances of the parameters the method will also give estimates of the variance of the state variable. This variance could be used to plot a probability profile provided the frequency distribution of the input parameters is realistic.

FAST is not strictly an appropriate technique for calculating probability profiles but it may be useful in determining a suitable formulation for doing so. By performing FAST first, it may be possible to simplify a simulation study by eliminating those parameters which contribute very little to the state variable. As can be seen in the example in Section 5.2.5 below, if it can be shown that only a few parameters contribute to the variability of projected catch then the estimation of variance of unimportant parameters may be unnecessary.

5.2.2 Covariance matrix from statistical analysis of catch-at-age data

The covariance approach is a very simple three step procedure that requires the fitting of a statistical model to the data. The steps are:

STEP 1: Carry out an assessment using the chosen statistical model.

STEP 2: Calculate the parameter covariance matrix. This can usually be done directly from the Hessian matrix which forms part of the minimization procedure (e.g. see Seber and Wild, 1989).

STEP 3: Translate the parameter covariance matrix into the variance of the desired quantity. This can often be done using a linear approximation method such as a delta method (a finite difference approximation).

Many analyses of catch-at-age data involve the fitting of a statistical model with an explicit objective function. The assumptions about the error structure of the data and the fitting procedure generally permit the estimation of the parameter covariance matrix. Examples of existing methods in which this is in principle possible are ADAPT (Gavaris, 1988), CAGEAN (Deriso *et al.*, 1985), Time series analysis (Gudmundson, WP8) and XSA (Shepherd, 1991a). Given this matrix it is possible to estimate the variance of any quantity derived from the parameters using, for example, a delta method. Thus, in the case of ADAPT, for example, the variances of the survivors can be used directly to compute the variance of the predicted yield given an estimate of recruitment and its variance.

One of the main advantages of this approach is that of speed, since alternative methods using bootstrapping or simulation can take considerable computational time. However, considering only the variability in the estimated parameters can result in underestimation of all the

5.2.2.1uncertainty contributing to the variability of the calculated quantity. Natural mortality, for example, is seldom estimated from catch-at-age analysis but must certainly contribute to the variability of projected populations and yield. Covariance matrices calculated from non-linear minimization may also be poor estimators of the true distributions.

5.2.3 Uncertainty estimates from Monte Carlo simulation of the assessment process

Monte Carlo simulation is a generalized method for obtaining uncertainty estimates from any model. Restrepo et al. (1992) describe the use of such simulations in quantitative stock assessments and their application in obtaining probability profiles for management recommendations. The basic Monte Carlo procedure is as follows:

STEP 1: Quantify the uncertainty in the inputs used for the assessment. This can often be achieved from statistical analyses of the raw data. For instance, variances can be estimated for the catch estimates and for the abundance survey estimates that go into the VPAs. In some cases, however, uncertainty in the inputs cannot be estimated statistically. For example, natural mortality is often input as an assumed quantity, not an estimated one, and thus the uncertainty associated with it must also be assumed. In addition to obtaining the variance estimates for the inputs, Monte Carlo simulation requires an idea about the shape of the input error distributions. Often these follow from statistical theory (e.g. the catch compositions at age are sometimes assumed to follow a multinominal distribution, and the abundance indices may be lognormally distributed). In other cases, the shapes of these probability density functions (pdf's) must be assumed (e.g. the uncertainty in M may be uniformly distributed, implying that all values within a range are equally likely).

STEP 2: Repeat the assessment process numerous times drawing random sets of inputs. The objective of this is to generate a large number (e.g. 1000) of plausible data-sets by drawing values at random from all the input pdf's. Then, for each data-set, the entire assessment is carried out with the model being used (e.g. Separable VPA, XSA, ADAPT, etc.). The end result of this step is the production of a large number of plausible assessment results.

STEP 3: Repeat all additional analyses and projections for each set of assessment results. Additional results may include the computation of the current status of the fishery with respect to common reference points (e.g. what is the ratio of current F to $F_{0.1}$?). The current F is available from each set of results in step 2; the reference point can also be estimated for each set by using the set-specific values of the variables required. For instance, to compute $F_{0.1}$ one would need the set-specific values of M, the current selectivity pattern and the weight-at-age vector. For each set one would then compute the ratio $F_{status\ quo}/F_{0.1}$ and the distribution of these values would describe the uncertainty in the ratio. The same process is followed for the projections upon which scientific advice will be based. The projections are carried out in the same manner as for the point estimates, once for each data-set obtained in step 2. Here it is again important to keep track of set-specific values that affect the outcomes (e.g. M). Recruitment forecasts used for the projections should also be associated with a variance. The simplest way to do this with few assumptions is to pick randomly one of the recruitment values estimated in the assessment process. Alternatively, one could fit a stochastic stock-recruitment relationship to the estimated time series in each data-set, and generate recruitment forecasts from that model.

A main advantage of Monte Carlo simulation in this context is that it is a very flexible way to account for possible departures from model assumptions. Another advantage is that it is a method to obtain uncertainty estimates for models that do not have an explicit objective function in a statistical minimization framework (e.g. *ad hoc* tuning). The main disadvantages of Monte Carlo simulation for use in working groups are that a) many decisions and analyses must be made in preparing the information about input uncertainty distributions and b) they are very intensive computationally, with a single run requiring several hours on a fast PC.

5.2.4 Uncertainty estimates from bootstraps

Bootstrapping (Efron, 1982) is related to Monte Carlo simulation but requires many fewer assumptions. Estimates of the uncertainty in the assessment are obtained from the model fit to the data rather than being dependent on the specification of uncertainties for all inputs. The basic bootstrap procedure is as follows:

STEP 1: Carry out the base assessment. Obtain the initial model fit and compute the residuals (e_{ij}).

STEP 2: Generate numerous data-sets by resampling the residuals. The non-parametric way of doing this is to generate new observations from

$$y_{obs_{ij}} = y_{pred_{ij}} + e_{ik} \qquad (1)$$

In this case y_{ij} may refer to the jth value of the ith available index of abundance. The new y observations are thus obtained by adding the predicted y and a residual chosen at random from the set estimated in step 1. The parametric way of doing this is to assume a distribution type for the residuals in step 1 (e.g. lognormal) and estimate the mean and variance for their distribution. The resampling would then be carried out by randomly choosing residuals drawn from this distribution rather than from the limited set of observations in step 1.

STEP 3: Repeat the assessment process for each data-set.

STEP 4: Repeat all additional analyses and projections for each set of assessment results. This and the preceding step are identical to the last two in the Monte Carlo procedure. Note, however, that many of the variables (e.g. M and all other data considered to be fixed) will not change in value from data-set to data-set.

The bootstrap procedure, like Monte Carlo simulations, can be used for methods without an explicit statistical objective function. The bootstrap is less flexible in the sense that it is still conditional on most inputs being known precisely; in this sense it is very similar to the method in Section 5.2.2. However, the bootstrap can be used with very few assumptions, especially if non-parametric resampling is carried out as in step 2 above. In terms of speed, bootstraps are computationally intensive like Monte Carlo simulations.

5.2.5 Examples of probability profiles for North Sea cod

In order to exemplify the methods described in the previous sections, probability profiles were calculated for a simulated projected catch of North Sea cod in 1992 given catch at age data and research vessel survey data for 1982-1991 from Anon. (1993b). The software available at the meeting was such that it was necessary to limit the analysis to ages 3-10 so as to achieve some comparability between methods. Thus, this simulation does not correspond to the accepted forecast for this stock but simply illustrates the methodology.

For the covariance method recruitment in 1992 at age 3 was taken as the geometric mean recruitment estimated for the previous ten years and the sample variance used for an estimate of its uncertainty. For the Monte Carlo and bootstrap methods recruitment was selected at random from the ten previous recruitment values. In a second set of runs recruitment was selected as the geometric mean of the last ten estimated recruitment values in the Monte Carlo and bootstrap methods, and a variance equivalent to this procedure used on the covariance method. This simulated recruitment was estimated with high precision. The probability profile plotted was the cumulative probability distribution of the projected catch at *status quo* fishing mortality. It is equivalent to the probability that the fishing mortality in 1992 will be greater than the fishing mortality in 1991. This is because the probability of a given catch from the distribution is equal to the probability that F=F*status quo*.

The analysis of the cod data is for illustration only and is not intended for any other purpose.

Fast Analysis: Uncertainties in the input parameters are given in Table 5.2.1. These give an uncertainty range of plus or minus two standard deviations. Results of the sensitivity analyses are shown in Figure 5.2.5.1. The results show that for high recruitment variance, the variability in the catch is due almost entirely to recruitment. When recruitment variance is low, the uncertainty in the population estimate at age 3 becomes important. This means, in the case of high recruitment variance, that all the methods are likely to perform similarly since they all made a similar assumption about recruitment.

Covariance method: As an example, a particular statistical model has been used here to estimate the covariance matrix using the methodology described in Cook *et al.* (1991). It assumes that all the errors are in the catches. Briefly, the model is based on the assumption that fishing mortality F is a product of a year effect, f, and an age effect, s, so that;

$$F_{ay}=s_a f_y \tag{2}$$

where a and y are subscripts for age and year respectively. It is then possible to write down the catch, C_{ay} in terms of F_{ay} and recruitment, R, in year y-$a+1$, ie

$$C_{ay}=g(s_a,f_y,R_{y-a+1}) \tag{3}$$

Subject to certain constraints the parameters can be estimated by minimizing the sum of squares;

$$\sum_a \sum_y [\log(\overline{C_{ay}})-\log(C_{ay})]^2 \tag{4}$$

This model has many parameters so the covariance matrix is very large. For reasons of space, only the CVs of the Fs and populations in 1991 are given in Table 5.2.2. It is important to understand that the estimated CVs are conditional on the model assumptions and constraints. These are elaborated in Cook *et al.* (1991).

Bootstrap and Monte Carlo methods: The bootstrap and Monte Carlo methods repeated an ADAPT assessment with varying inputs. The parameters estimated were the 1992 populations at ages 4 to 7 and the age-specific catchability for all indices. The fishing mortality for ages 7 to 10 in 1991 were constrained to be the same as for age 6. Within each iteration the recruitment for catch projections at *status quo* fishing mortality was chosen at random from the estimated recruitment of previous years.

For the Monte Carlo simulations the natural mortality was chosen uniformly at random between 0.14 and 0.28, the catches were assumed to be distributed lognormally with CVs between 3 and 10 percent and the abundance indices were also assumed to be lognormal with CVs of 40 percent.

The resulting CVs in the terminal year are given in Table 5.2.2.

The estimated probability profiles from the methods are shown in Figures 5.2.2 and 5.2.3. When recruitment has high variance, the profiles are all very similar both in their location and slope. For the bootstrap and Monte Carlo methods, the profiles show inflexions. This is due to the fact that, as applied here, these methods draw on an empirical frequency distribution for recruitment taken from the estimated recruitment in the VPA (see Figure 5.2.4). Only ten values were used in the data-set which is unlikely to give a smooth frequency distribution. In the covariance method, recruitment was assumed to be lognormal. The principal reason for the similarity of the profiles is the fact that the recruiting age group dominates the catch in this example. Hence its variability dominates the calculations as adumbrated in the sensitivity analysis.

When recruitment variability is low (Figure 5.2.3) the bootstrap and Monte Carlo methods still agree very closely and, not surprisingly, give a much steeper profile. The covariance method is still similar but has a noticeably lower slope. This is due to the fact that the particular model used in the covariance method only uses the catch data and so there is very little information in the data on the youngest age in the most recent year. This age group, therefore, has a high estimated variance. Since it makes an important contribution to the projection (Figure 5.2.1) it makes a large contribution in the probability profile.

Despite important differences in the methods applied here, the estimated profiles are very similar which is encouraging. It does not mean, however, that this would be generally true. The covariance approach would be a desirable tool to use in working group environments because of the speed of computation. However, the use of such a method should be supported by more thorough studies involving Monte Carlo and bootstrap studies where a more comprehensive investigation of uncertainty can be made.

5.3 Medium-term Projection and Advice

5.3.1 Methods for medium-term simulations

Management strategies can be very complex, and can involve many regimes, exceptions and complex interactions. The only way to assess the long-term performance of management strategies on such systems is via simulation. Detailed simulations will require the modelling of fish stock dynamics, management measures and the resulting exploitation behaviour. This exercise may help to identify more precisely the consequences of a particular strategy.

Medium-term simulations should be stochastic and explicitly include variability in (a) the current status of the stock, and (b) the likely trajectory of future recruitment. Uncertainty in current stock status can be estimated with any of the methods in Section 5.2. Owing to the dangers involved when assuming that recruitment is independent of stock size, the simulated future recruitment should be drawn from a stochastic stock-recruitment relationship. Such a relationship can be obtained by fitting a relationship to the data and estimating the variance around the fitted relationship (see Section 4) or using kernel methods (Evans and Rice, 1988; Skagen, 1991; Cook and Forbes, WP2). The projections should be made so that each stock size trajectory will differ from the others due to random outcomes.

In their simplest form, these simulations can be used to examine the medium-term performance of simple management strategies. For instance, one can estimate the probability that the stock will go below a threshold level at least once during a 20-year period, given that it continues to be fished at F*status quo*. Such projections would be carried out in a manner analogous to Step 1 of Section 5.2.3.

More complicated simulations are required to evaluate the medium- and long-term performance of management strategies under more realistic situations. It is likely that the management measures taken from year to year over a medium-term horizon will vary, depending on the perceived status of the stock every few years. Thus, it is important that the future assessments of the stock be simulated as well. There are many ways to simulate this interaction between future stock status and future management measures and only two are provided in the paragraph below. Note that other components of the simulation (uncertainty in current status, stock-recruitment) should be taken into consideration as explained at the beginning of this section.

(A) **Without actual assessment updates**. At any point in time in the simulations, the status of the stock will be known (i.e., the software should keep track of all state variables of interest such as abundance, catches, age structure, etc.). The perceived stock status at that point in time can be generated by drawing from a distribution of possible estimates. For instance, if the fishing mortality for a given trajectory in year 2000 is F = 1.2, one could randomly draw an estimated F from, say, a lognormal distribution with mean 1.2 and CV = 0.2. This estimate of F would be used in place of an assessment. Then, given the known catch, one could estimate the corresponding stock size. Depending on that perceived outcome, the management regulation for the following simulation year would be decided.

(B) **With actual assessment updates**. This would be similar to (A) above, except that the simulated data would be used as input to an assessment model. Since uncertainty should play an important role, these inputs should be subjected to measurement error (e.g. the catch or relative abundance would be sampled from distributions centred around the simulated value). The assessment results would then provide a perceived estimate of stock status which would in turn affect the following year's management advice.

A particular advantage of method (B) over method (A) is that the former can more easily track the benefits of an increasingly longer time series of data in terms of reducing assessment uncertainty. In this sense, method (B) can more realistically simulate flexible management control laws (Section 5.4).

5.3.2 Assessment working group advice

In the light of the current state of most fish stocks assessed by ICES, it is clear that simple short-term advice does not cover all aspects of the problem. Short-term advice captures the fact that many stocks are overexploited according to any reasonable definition of the term. However, this form of advice does not capture the fact that a fairly large number of stocks show a stock-recruitment relationship. Thus, higher yields would be expected at higher stock sizes. The effect of such relationships would be expected to appear within a few years of a build-up of the stock. Medium-term advice should, therefore, be considered a regular part of the work done by assessment working groups.

In order to implement medium-term advice, it will be necessary to make software available to do the relevant computations. The Methods Working Group noted that software already exists for this purpose (see Section 5.3.1), but this software needs to be modified and adapted to the specific output recommended in this report (see Section 5.3.3).

The Methods Working Group, therefore, recommends that an *ad hoc* study group led by R. Cook (Aberdeen) be set up to develop the medium-term prediction and simulation software as indicated in this report in order that the assessment working groups may be able to make use of it as a regular part of the assessments.

Output from the software should include (but not be restricted to) expected annual yields and spawning stock sizes for the time period in question along with fractiles of the distribution of these estimates.

5.3.3 Presentation of medium-term management advice

The Working Group considers that the medium-term consequences of the various management options are not very clear in the present form of the advice of ACFM, and suggests that some improvements could be made. The present section commencing "Continued fishing at current levels of fishing mortality" often simply repeats in words information already presented in the option table, and could be replaced by a new section on the medium-term "consequences". In the immediate future this could simply contain text setting out the expected consequences along the following lines.

> **"Medium-term Projections**
>
> Over the next five/ten/fifteen years the consequences of these management options are likely to be:
>
> A: Gradual recovery of SSB to levels above the MBAL after three or four years provided recruitment improves from the current low level.
>
> B: SSB remains close to the MBAL for the foreseeable future.
>
> C: Continued decline of SSB to levels well below the MBAL with increased risk of recruitment failure."

These statements have been framed making reference to MBAL, but this concept needs to be critically reviewed (see Section 4).

In some cases sufficient information already exists in the Working Group reports for this to be done in the near future. In other cases ACFM would need to request the assessment working groups to prepare and present the necessary catch forecasts. The use of deterministic forecasts for this purpose is, however, really rather unsatisfactory, as they often (but not always) depend crucially on unknown future recruitment. As described above and in Section 5.2, methods for preparing appropriate stochastic forecasts are now available, and the Working Group recommends that ACFM should encourage assessment working groups to adopt and use these methods as soon as possible (see Section 5.3.2)

It should be stressed that these forecasts need to take account of the uncertainty of recruitment, by Monte Carlo methods, and should incorporate the information on the probability distribution of recruitment at the appropriate stock levels as described in Section 4, thus taking due account of any stock-recruitment relationship indicated by the data.

When the results of such forecasts become available, it would be possible to expand the section in the ACFM report on medium-term projections considerably.

A suggested presentation of the results of such calculations is given in Figures 5.3.1-5.3.5. These give the trajectories of medium-term yield and SSB under various selected management options, with the uncertainty indicated by the appropriate upper and lower percentiles, e.g. quartiles. In addition to this it is suggested that estimates of the inter-annual variability of yield (as percentage change), and the cumulative yield over the selected time horizon should be given for each option. This would provide a much firmer basis for the textual advice proposed above, which should probably be retained.

The Working Group recognises that this could expand the ACFM report by around half a page or a full page for each stock for which it is done, but considers that this would not be inappropriate if it assists in conveying an important message which is currently not reaching managers at all clearly.

5.4 Management Procedures

The development of "management procedures" for the management of fish stocks recognizes that this process involves more than assessment exercises. Essentially four steps are involved in the actual overall process:

i) the stock "generates" data each year (e.g. catch-at-age data, CPUE, survey results);

ii) these data are input to an assessment process which estimates *(inter alia)* the size and productivity of the resource;

iii) the results of the assessment process are used to formulate a control measure, such as a TAC (through a "catch control law") or a fishing effort level;

iv) the control measure impacts the dynamics of the stock, and hence the results of i - iii when the whole procedure is repeated each following year.

The management procedure approach argues that all these steps have to be considered, and in combination rather than separately. This is both because of the interaction between the steps, and because the anticipated consequences of certain management measures can sensibly be considered only in terms of their application and updating over a period of time rather than for a single year only.

Alternative candidate management procedures are evaluated by means of computer simulation (e.g. Anon., 1990), which mimics steps i) - iv) above. Thus:

a) a computer model of the stock and fishery is developed, which each "year" generates data of the form available for the actual fishery - thus, for example, survey results are output, which incorporate typical measurement errors;

b) an updated assessment of the stock is calculated each simulated "year" from the data generated and those data only (the "assessor" does not know the true state of the stock) - by means of VPA, for example;

c) the control measure is calculated each "year" from the results of b) - an $F_{0.1}$ TAC for example;

d) the TAC, say, is fed back to the computer model of the stock, where this information is used in updating the numbers at age for the following year.

Steps a) - d) are repeated for a number of years chosen to be appropriate for the time-scale of the stock's dynamics: 10-20 years would be appropriate for most fish species. At the end of this simulation period, the computer model of the stock provides information on its true status to allow statistics of the anticipated performance of the management procedure to be developed. Typical such statistics might include the spawning stock biomass at the end of the period, the total catch taken during the period, and the extent to which the catch taken (or fishing effort applied) has varied from one year to the next.

Eventually, managers have to select between different candidate procedures on the basis of the trade-offs which they exhibit between such attributes. The simulations need to be repeated a number of times because of stochastic effects - recruitment variation and measurement errors, for example - so that performance statistics are expressed as parameters (means or percentiles) of the resultant distributions of values for chosen attributes.

This evaluation process provides a framework to take explicit account of the inevitable uncertainties in the state of knowledge of a resource. Although primary calculations to choose a procedure make use of a computer model of the stock which is based on the current "best" assessment of the resource, it is essential that they be repeated for plausible variations of this assessment. The purpose of such "robustness tests" is to determine whether the procedure under consideration provides performance statistics which are reasonably insensitive to such variations (which should reflect the degree of uncertainty - both structural and as regards the imprecision of parameter estimates - in the current "best" model of the resource).

The Working Group noted that there has recently been increased interest in the possibility of using defined management procedures within the ICES area (Horwood and Griffith, 1992). It was noted that techniques now exist for the evaluation of candidate management strategies according to multiple criteria as discussed above. The Working Group considers that the development of longer-term management strategies is possible and much needed, and recommends that within ICES the subject should be carried forward by the Working Group on Long-Term Management Measures.

6 REPORT REVIEW

6.1 Earlier Reports of the Working Group on Methods of Fish Stock Assessment

The results in this report demonstrate that the statement in the report of the 1989 meeting (Anon., 1993c) that methods such as TSA which do not use CPUE/survey data "have no chance of detecting sharp changes of fishing mortality in the last year" is incorrect: TSA1 in particular detects such changes with remarkable accuracy. Clearly such methods have considerable practical potential, and the Working Group would strongly support work to produce a version of this method which could be used operationally.

6.2 Report of the Planning Group for the Development of Multispecies, Multifleet, Assessment Tools

The Planning Group had basically two questions to address:

a) The dissemination of multi-species tools to area-based working groups and,

b) The definition of fleets, data formats and analytical software for area-based working groups.

It became clear at an early stage of the Planning Group meeting that the terms of reference could not be dealt with as they stood because of resource implications in national institutes and in the ICES Secretariat. In the case of multi-species software, for example, only one institute had the expertise and resources for development and much of this was being funded externally. It was not felt that other institutes could commit resources to the project. Furthermore, the most commonly used multi-species tools are very data demanding and presently only exist for the North Sea and Baltic areas. This means that only two assessment working groups would be able to benefit from new software. Thus, it did not appear to be a priority to devote international resources to this type of development.

So far as analytical software was concerned the Planning Group pointed out that the main needs are in short-term prediction and long-term analysis. Forecasting programs need to be better interfaced with other analytical software. Long-term (multi-species) analysis is required but should be developed under the umbrella of the Working Group on Long-Term Management Measures.

In the case of fleet definitions and data formats for area-based groups, it became clear that the existing data storage format within IFAP was not suitable for the development of a fleet-based data structure. The design of a new data structure has substantial implications for IFAP. The Planning Group did not feel able to pursue this issue since the IFAP steering group would need to be involved. In view of the development of an SAS-based data management system by the Danish Institute for Fisheries and Marine Research (DIFMAR), which is to be fleet-based, it appeared prudent to await the completion of that package before subjecting IFAP to substantial re-development. This package should be completed by the end of 1993 as part of an EC-AIR contract. In the meantime it would be better to concentrate resources on optimising the performance of the present IFAP system. The Planning Group suggested that area-based working groups should also consider a simplified STCF exchange format as a basis for the exchange of their own data.

The Methods Working Group generally endorsed the conclusions of the report.

6.3 Report of the Workshop on the Analysis of Trawl Survey Data

Aspects of the work carried out by the Trawl Survey Workshop have been discussed under Section 3.2. The Working Group recommends that the report, edited by G. Stefánsson, should be published in the Cooperative Research Report series.

7 RECOMMENDATIONS

The Working Group recommends that :

(1) Catch-at-age analysis of short-lived species should continue until replaced by alternative methods (see Section 2);

(2) Working Groups dealing with short-lived species (such as the Norway Pout and Sandeel Working Group) should examine all available data, especially monthly or quarterly CPUE data, in order to determine e.g. relationships between CPUE and abundance which would enable alternative management methods to be applied (see Section 2).

(3) When recruitment indices are included in the tuning data, the XSA estimates of terminal population (including the recruiting year classes) be retained for predictions. These estimates should be cross-checked using RCT3 and, if there are discrepancies, these should be investigated using retrospective analysis.

(4) Assessment working groups should regularly reanalyse stock and recruitment data (see Section 4) and follow the steps described in Section 4.4 to estimate MBALs.

(5) Stock and recruitment data for ICES stocks be routinely collected and maintained in a standard format, accessible on the Internet. The format should be standardized with that used by other collection agencies and the data from as many other sources as possible (notably North America) should also be kept in the ICES system (see Sections 4.1 and 4.7).

(6) A small *ad hoc* study group led by Dr R Cook (Aberdeen, UK) be set up to develop the medium-term prediction and simulation software as indicated in Section 5 in order that the assessment working groups may be able to make use of it as a regular part of the assessments.

(7) ACFM encourage assessment working groups to adopt and use methods available for preparing appropriate short- and medium-term stochastic forecasts as described in Sections 5.2 and 5.3.3 as soon as possible.

(8) The evaluation, by simulation studies, of longer-term management procedures and strategies should be assigned to the Working Group on Long-term Management Measures (LTMMWG) as a major task for the future. The LTMMWG will wish to take note of the proposed EC Study Group on this subject.

(9) The suggestions made in Section 5.4 and relevant other sections of this report, and that of the LTMMWG, should be discussed at the Theme Session on Risk Analysis at the 1993 Statutory Meeting.

(10) The report of the Workshop on the Analysis of Trawl Survey Data, edited by G Stefánsson, should be published in the Cooperative Research Report series.

(11) A low level of shrinkage (CV = 0.5) should be used as a starting point in VPA tuning and other values should be explored using retrospective analysis. This level of shrinkage is found to be beneficial in most cases and rarely detrimental (see Section 3).

(12) The Working Group on Methods of Fish Stock Assessment, at a meeting in 1995, should consider alternative assessment methods based on limited data.

8 REFERENCES AND WORKING PAPERS

8.1 References

Anon. 1985. Report of the Working Group on Methods of Fish Stock Assessments. ICES, Coop. Res. Rep., No. 133.

Anon. 1990. Report of the Scientific Committee. Annex E. Report of the Sub-Committee Management Procedures. Rep. int. Whal. Commn. 40:94-118.

Anon. 1991. Report of the Working Group on Methods of Fish Stock Assessments. ICES, Doc. C.M.1991/Assess:25.

Anon. 1992a. Working Group on the Assessment of Mackerel, Horse Mackerel, Sardine and Anchovy. ICES, Doc. C.M.1992/Assess:17.

Anon. 1992d. Report of the Workshop on Analysis of Trawl Survey Data. ICES, Doc. C.M.1992/D:6.

Anon. 1992e. Report of the Herring Assessment Working Group for the Area South of 62°N. ICES, Doc. C.M.1992/Assess:11.

Anon. 1993a. Report of the Working Group on Atlanto-Scandian Herring and Capelin. ICES, Doc. C.M.1993/Assess:6.

Anon. 1993b. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak. ICES, Doc. C.M.1993/Assess:5.

Anon. 1993c. Reports of the Working Group on Methods of Fish Stock Assessments. ICES Coop. Res. Rep., 191.

Beddington, J.R., and J.G. Cooke. 1983. The potential yield of fish stocks. FAO. Fish. Biol. Tech. Pap. 242:1-47.

Butterworth, D.S., and M.O. Bergh (in press). The development of a management procedure for the South African anchovy resource. Can. J.Fish. Aquat. Sci. Spec. Publ. (in press).

Conser. R.J. 1991. A DeLury model for scallops incorporating length-based selectivity of the recruiting year class to the survey gear and partial recruitment to the commercial fishery. 12th NEFC Stock Assessment Workshop, Woods Hole (MA), USA, June 1991 Working Paper No. 9 18 p.

Cook, R.M., P.A. P.A. Kunzlik and R.J. Fryer. 1991. On the quality of North Sea cod stock forecasts. ICES Journ. of Marine Science, 48:1-13.

Copas, J. B. 1983. Regression, prediction and shrinkage (with discussion). J. R. Statist. Sec. B 45 (3) 311-354.

Cukier, R.I., H.B. Levine, and K.E. Schuler. 1978. Non-linear sensitivity analysis of multiparameter model systems. Journ. Computational Physics, 26:1-42.

de la Mare, W.K. 1989. The model used in the Hitter and Fitter program. Rep. Int. Whal. Commn. 39: 150-151.

Deriso, R.B., T.J. Quinn II and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. and Aquatic Science, 42:815-824.

Efron, B. 1982. The jackknife, bootstrap and other resampling plans. Society for industrial and applied mathematics, Philadelphia.

Evans, G.T. and J.G. Rice. 1988. Predicting recruitment from stock size with the mediation of a functional relation. Journ. du Cons. Int. Explor. Mer, 44:111-122.

Fournier, D. and C.P. Archibald. 1982. A general theory for analysing catch-at-age data. Can. J. Fish. Aquat. Sci., 39:1195-1207.

Francis, R.I.C.C. 1991. Risk analysis in fishery management. NAFO Scientific Council Studies, No. 16, 143-148.

Francis, R.I.C.C., Robertson, D.A., Clark, M.R., and Coburn, R.P. 1992. Assessment of the QMA 3B orange roughy fishery for the 1992/93 season. New Zealand Assess. Res. Document 1992/4.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. Canadian Atl. Fish. Sci. Adv. Comm. (CAFSAC) Res. Doc. 88/29. 12p.

Gudmundsson, G. 1987. Time series models of fishing mortality rates. ICES, Doc. C.M. 1987/D:6, mimeo.

Gudmundsson, G. In press.Time series analysis of catch-at-age observations. Applied statistics.

Gulland, J. A. 1965. Estimation of mortality rates. Annex to Arctic Fisheries Working Group Report. ICES, C.M. Gadoid Fish Comm., (3), 9pp.

Hilborn, R. 1990. Estimating the parameters of full age-structured models from catch and abundance data. Bull. Int. North Pac. Fish. Commn. 50: 207-213.

Hilborn, R.and C.J. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York.

Horwood, J. and D. de G. Griffith. 1992. Management strategies and objectives for fisheries. Privately published, pp. 38.

Millar, R.B. and R.A. Myers. 1990. Modelling environmentally induced change in growth for Atlantic Canada cod stocks. ICES, Doc. C.M.1990/ G:24.

Pope, J. G. and T.K. Stokes. 1989. The use of multiplicative models of separable VPA, integrated analysis and the general VPA tuning problem. Am.Fish. Soc. Symp., 6:92-101.

Punt, A.E. 1992. PC-BA User's Guide (Version 1.1). Benguela Ecology Programme (South Africa). No.24 (37 pp).

Punt, A.E., and D.S. Butterworth. 1992. Examination of the implementation of "enhanced" stock reduction analysis for the Chatham Rise orange roughy fishery. Doc. submitted to the New Zealand Fishery Industry Board (35 pp).

Punt, A.E., and D.W. Japp. (In press.) Stock assessment of the kingklip (*Genypterus capensis*) resource off South Africa. S. Afr. J. Mar Sci. (in press).

Restrepo, V.R., J.M. Hoenig, J.E.Powers, J.W. Baird and S.C. Turner. 1992. A simulation approach to risk and cost analysis, with applications to swordfish and cod fisheries. Fish. Bull. 90:736-748.

Rosenberg, A.A., G.P. Kirkwood, J.A. Crombie and J.R. Beddington. 1990. The assessment of stocks of annual squid species. Fisheries Research, 8:335-350.

Rosenberg, A.A., G.P. Kirkwood, R.M. Cook and R.A. Myers. 1992. Combining information from commercial catches and research surveys to estimate recruitment: a comparison of methods. ICES, Journ. Mar. Sci. 49:379-387.

Seber, G.A.F. and C.J. Wild. 1989. Non-linear regression. John Wiley, New York.

Serebryakov, V.P. 1991. Predicting year class strength under certainties related to survival in the early life history of some North Atlantic commercial fish. NAFO Sci. Counc. Studies 16: 49-56.

Shepherd, J.G. 1982. A versatile new stock-recruitment relationship for fisheries, and the construction of sustainable yield curves. J.Cons. int. Explor. Mer, 40:67-75.

Shepherd, J.G. 1991a. Extended Survivors' Analysis: an improved method for analysis of catch-at-age data and catch-per-unit effort data. Working Paper presented at meeting of Methods Working Group, 20-27 June 1991, see ICES Doc. C.M. 1991/Assess:25.

Shepherd, J.G. 1991b. Report of special session. NAFO Sci. Counc. Studies, 16:7-12.

Shepherd, J.G., and Nicholson, M.D. 1991. Multiplicative modelling of catch-at-age data and its application to catch forecasts. ICES, J. Cons. int. Explor. Mer. 47: 284-294.

Sissenwine, M.P., and Shepherd, J.G. 1987. An alternative perspective on recruitment overfishing and biological reference points. Can. J. Fish. Aquat. Sci. 44: 913-918.

Skagen, D. 1991. Stock prediction using stochastic recruitment numbers with empirical stock-dependent distribution. ICES, Doc. C.M.1991/H:28.

Sparholt, H. 1992. Note on haddock and whiting assessment in the North Sea. Working Doc. ACFM Nov. 1992.

Walters, C. [J.] 1986. Adaptive mnagement of renewable resources. New York, MacMillan (374 pp).

8.2 Working Papers

WP1. Cook, R.M. and S.A. Reeves. An assessment of North Sea industrial fish stocks with incomplete catch-at-age data.

WP2. Cook, R.M. and S.T. Forbes. Unsafe biological limits for North Sea cod.

WP3. Beek, van, F. SPLIR, a risk analysis model to assist in finding appropriate long term levels of fishing mortality.

WP4. Butterworth, D. Current initiatives in the management of the S.A. anchovy and Antarctic krill resources.

WP5. Pope, J.G. Thoughts on the stability of estimates of F_{med}/F_{low} and F_{max}/F_{now} relative to changes in natural mortality.

WP6. Pope, J.G. A consideration of ways of introducing stock recruitment constraints into pragmatic fisheries management.

WP7. Fryer, R., R. Cook and L. Hastie. The B Team talk about shrinkage.

WP8. Gudmundsson, G. Time series analysis of catch at age data without effort measurement.

WP9. Gudmundsson, G. Applications of recruitment indices.

WP10. Myers, R.A., A.A. Rosenberg, P. Mace, N. Barrowman and V. Restrepo. In search of recruitment overfishing thresholds.

Related document

Anon. 1985. Problems in the management of short-lived species. From ACFM, Discussion Paper, November 1985.

Table 3.1.1 'Theoretical' CVs for Western Channel Sole

	Age 3	Age 4	Age 5
1978	0.05	0.14	0.01
1979	0.14	0.12	0.08
1980	0.21	0.19	0.19
1981	0.25	0.22	0.26
1982	0.38	0.33	0.29
1983	0.41	0.39	0.24
1984	0.33	0.33	0.18
1985	0.39	0.38	0.22
1986	0.34	0.34	0.19
1987	0.05	0.05	0.02
1988	0.16	0.15	0.08
1989	0.37	0.36	0.20

Table 3.2.1 Western Channel Sole RCT data.

```
"Western Channel sole recruits"
6,14,2  ,"(no. surveys, no. years, vpa col no.)"
1978, 4924, -11, -11, -11 ,   11, -11  , -11,
1979, 8441, -11, -11,  272,  135, -11  , -11,
1980, 4773, -11, -11,  107,   77, -11  ,  408,
1981, 3873, -11,  50,  200,    3,  260,  127,
1982, 6034,  41,  69,   46,    2,  331,  204,
1983, 6580,  45, 122,   38, -11 , 1386,  376,
1984, 3368,  21,  49, -11 , -11 ,  220,   90,
1985, 5251,  30,  57, -11 , -11 ,  497,  141,
1986, 3080,  17,  44, -11 ,    4,  420,   96,
1987, 2968,  20,  25,   36,    8,  823,  180,
1988, 2168,  17,  27,    2,    8,  290,   82,
1989,-11,    79, -11,  777,   25,  530,  229,
1990,-11,   -11, -11,   25,   21,  447,  450,
1991,-11,   -11, -11,   46, -11 ,  170, -11,
UK7e2
UK7e3
Fr7d0
Fr7d1
UK7d0
UK7d1
```

Table 3.2.2 Western Channel Sole (VIIe). RCT3 Retrospective Analysis.

Year-Cl	VPA	CAL +SH	CAL NSH	FUN +SH	FUN NSH	PRED +SH	PRED NSH
1984	3368	5222	2368	5109	3334	5220	4045
1985	5251	4566	4257	4575	4348	4615	4420
1986	3080	3598	3284	3770	3540	3914	3718
1987	2968	3744	3533	3877	3727	4006	3886
1988	2168	2933	2675	3073	2881	3275	3116
1989	7631	8728	10494	7513	8325	6829	7358
1990		5289	6193	4955	5208	4692	4802
1991		3913	3071	3790	3303	4007	3755
RMSLR 84-89		0.258	0.241	0.267	0.180	0.300	0.225

Table 3.2.3 Irish Sea Plaice RCT data.

```
"Irish sea plaice recruits"
6,18,2  ,"(no. surveys, no. years, vpa col no.)"
1974,11180,-11,-11,-11,352,473,-11,-11,-11,
1975,17254,-11,308,726,1775,1711,8.18,-11,-11,
1976,19167,78,877,190,1648,650,14.56,-11,-11,
1977,23226,32,641,1110,1744,3018,6.06,-11,-11,
1978,20768,237,348,4046,5588,1161,19.09,-11,-11,
1979,15585,757,3003,2330,1925,1897,3.37,-11,-11,
1980,8497,17,98,323,940,844,3.4,-11,-11,
1981,21525,18,585,3125,1371,1538,12.9,-11,-11,
1982,21330,1250,1195,4061,1796,2358,22.18,-11,-11,
1983,22422,262,1983,2995,2208,1683,-11,-11,-11,
1984,16235,508,2635,2649,2281,970,17.9,-11,-11,
1985,18995,430,2520,2246,1959,2145,19.71,-11,-11,
1986,20025,1033,2074,4886,4264,2945,29.71,-11,29776,
1987,10945,173,2624,4053,2961,914,38.78,12727,11168,
1988,5797,397,506,553,610,134,14.01,5998,6985,
1989,-11,31,438,271,480,-11,9.65,24855,14079,
1990,-11,216,873,-11,-11,-11,8.31,11052,-11,
1991,-11,-11,-11,-11,-11,-11,40.37,-11,-11,
"ssoct0"
"ssjun1"
"ssoct1"
"ssjun2"
"ssoct2"
"irmay1"
```

Table 3.2.4 Irish Sea Plaice. RCT3 Retrospective Analysis.

Year-Cl	VPA	CAL +SH	CAL NSH	FUN +SH	FUN NSH	PRED +SH	PRED NSH
1980	8497	16403	15694	15712	15160	17566	17604
1981	21525	17343	18797	17172	17700	16746	17032
1982	21330	22490	29080	21434	23430	19234	20047
1983	22422	21001	25207	20230	21477	18875	19424
1984	16235	21207	23609	20420	21217	19264	19643
1985	18995	22060	25680	20853	21826	19331	19743
1986	20025	26646	35072	23842	25770	20944	21692
1987	10945	21173	23577	20228	20867	19170	19455
1988	5797	10753	5287	13125	12058	15017	14613
1989		9738	4166	10886	9382	13584	13046
1990		15360	11361	15198	14740	15798	15798
1991		16723	1186595	18626	30174	16378	17402
RMSLR 80-88		0.405	0.428	0.418	0.407	0.457	0.452

Table 3.2.5 Icelandic Cod RCT data.

```
Icelandic COD. Predicting 3-group.
5 12 2
'Ycl' 'VPA' 'CPUE'  'SUR4'  'SUR3'  'SUR2'  'SUR1'
1980   229     30    -11     -11     -11     -11
1981   141    170     50     -11     -11     -11
1982   145    210     29      38     -11     -11
1983   336   1620     79      96      46     -11
1984   299    760     95     111      53     180
1985   175     60     55      82      31     160
1986   86      30     16      27      11      50
1987   159     20     34      26      27      40
1988   -11     70     28      30      17      70
1989   -11   -11     -11      46      23      90
1990   -11   -11     -11     -11      25      60
1991   -11   -11     -11     -11     -11      20
```

Table 3.2.6 Icelandic Cod. RCT3 Retrospective Analysis.

Year-Cl	VPA	CAL +SH	CAL NSH	FUN + SH	FUN NSH	PRED +SH	PRED NSH
1984	299	222	426	225	314	217	263
1985	175	228	234	225	228	227	231
1986	86	126	66	132	90	143	109
1987	159	138	132	139	134	143	138
1988		123	117	128	123	131	127
1989		156	152	158	154	159	155
1990		163	159	163	160	164	161
1991		157	44	153	63	157	82
RMSLR 84-87		0.285	0.280	0.294	0.161	0.332	0.206

Table 3.2.7 North Sea Herring RCT data.

```
Herring in North Sea +IIIa - 1 ringers
1 11  2
1980   85610  1293
1981  169767  1797
1982  153421  2663
1983  158214  3416
1984  275885  3667
1985  335217  5717
1986  278343  4192
1987  152919  3468
1988  176851  2146
1989  119454  2433
1990    -11   2339
IYFS-1rg
```

Table 3.2.8 North Sea Herring - 1 ring / IYFS RCT3 Retrospective Analysis.

Year-Cl	VPA	CAL +SH	CAL NSH	FUN +SH	FUN NSH	PRED +SH	PRED NSH
1984	275885	150186	238354	151264	203725	148333	182057
1985	335217	208865	437840	209553	358699	201346	305523
1986	278343	231465	273946	228463	261098	223223	250201
1987	152919	214288	226356	212860	222430	210781	218959
1988	176851	149344	129066	151434	135955	155202	142145
1989	119454	165441	153458	166725	158045	169026	162000
1990		155304	141456	156801	146984	159555	151717
RMSLR 84-89		0.382	0.261	0.380	0.254	0.395	0.278

Table 3.2.9 Northeast Arctic Cod RCT data.

```
NORTHEAST ARCTIC COD : recruits as 3 year-olds (inc. data for ages 0,1,2 & 3)
16,35,2             (No. of surveys, No. of years, VPA Column No.)
1957,  791, -11, -11, -11, -11, 120,  160, -11 ,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1958,  919, -11, -11, -11, -11, 160,  240, -11 ,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1959,  730, -11, -11, -11, -11, 180,  140, -11 ,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1960,  473, -11, -11, -11, -11,  90,  190, -11 ,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1961,  339, -11, -11, -11, -11,  20,   20, -11 ,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1962,  778, -11, -11, -11, -11,  70,   40, -11 ,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1963, 1583, -11, -11, -11, -11, 210, 1200, -11 ,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1964, 1293, -11, -11, -11, -11, 490,  450, -11 ,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1965,  170, -11, -11, -11, -11,  10,   10, -11 ,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1966,  112, -11, -11, -11, -11,  20,   10,   20,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1967,  197, -11, -11, -11, -11,  10,   10,   40,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1968,  405, -11, -11, -11, -11,  70,   10,   20,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1969, 1016, -11, -11, -11, -11, 110,   60,  250,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1970, 1818, 230, 640, 600, 420, 700,  850, 2510,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1971,  525,  70,  90,  60,  30, 370,  240,  770,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1972,  622,  50,  40, 340, 150, 540,  170,  520,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1973,  614, 160,  50, 150,  20, 700,   50, 1480,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  -11
1974,  348,  10,  10,  40,  10,  60,   10,  290,   -11,  -11,  -11,  -11,  -11,  -11,  -11 ,  -11 ,  104
1975,  639, 600,  10, 440,  10, 930,   40,  900,   -11,  -11,  -11,  -11,  -11,  -11,  -11 , 8820,  797
1976,  199,  10,  10,  10,  10,  40,   10,  130,   -11,  -11,  -11,  -11,  -11,  -11,  450, 2350,  109
1977,  140,  10,  10,  20,  10,  20,   10,  490,   -11,  -11,  -11,  -11,  -11,  -11,  280,  140,  -11
1978,  158,  10,  20,  10,  10,  10,   30,  220,   -11,  -11,  -11,  -11,  -11,  -11,  160,  -11 ,   58
1979,  158,  10,  10,  10,  10,  10,   80,  400,   -11,  -11,  -11,  -11,  -11,  -11, -11  ,  730,   71
1980,  169,  10,  10,  10,  10,  10,   80,  130,   -11,  -11,  232,  -11,  -11,  107,   30,   40,   17
1981,  386,  10,  10,  10,  10,  40,   40,  100,   -11,  177, 1220,  -11,  268,   73,   10,  150,  174
1982,  508,  10,  80,  30, 130,  80,  100,  590,  2590, 3660, 1620, 1450, 1130,  991, -11  , 5060,  550
1983,  894,  40,  90, 110,  70, 450,  410, 1690, 21700, 6470, 6790, 4990, 4520, 2970, 23820, 8780, 1246
1984,  282,  10,  10,  20,  80,  70,  150, 1550,   390, 4030, 2330, 2390, 1810, 1410,  690, 5780,  126
1985,  230,  30, 100,  20,  30,  40,   60, 2460,  5620, 3870, 1800,  409, 1080,  332, 6250,  470,   79
1986,  216,  10,  20,  10,  10,  20,   50, 1370,   253,  635,  379,  415,  166,  154,   10,  230,   31
1987,  -11,  10,  10,  10,  10,  10,   10,  170,    38,  127,  258,   31,   27,   82,   10,   90,   32
1988,  -11,  10,  10,  10,  10,  70,   10,  330,    71,  489,  370,   36,   83,  288, -11  ,  580,  145
1989,  -11,  10,  10,  40,  10,  70,  100,  380,  1220, 2127, 1704,  615, 1009,  -11, 1450, 4340,  490
1990,  -11,  60,  10,  40,  40,-11 , -11 , 1230,  3567, 4822,  -11, 1316,  -11,  -11, 2770, 10040,  -11
1991,  -11,  30,  60, -11, -11,-11 , -11 , 2300,   997,  -11,  -11,  -11,  -11,  -11, 2500,  -11 ,  -11
R-1-1      USSR Bottom trawl index, area I,   age 1
R-2B-1     USSR   "      "     "     "  IIb, age 1
R-1-2      USSR   "      "     "     "  I,   age 2
R-2B-2     USSR   "      "     "     "  IIb, age 2
R-1-3      USSR   "      "     "     "  I,   age 3
R-2B-3     USSR   "      "     "     "  IIb, age 3
INTOGP     International 0-group survey
N-BST1     Norwegian Barents Sea,  Bottom trawl survey, age 1
N-BST2     Norwegian   "     "       "      "     "     age 2
N-BST3     Norwegian   "     "       "      "     "     age 3
N-SVT1     Norwegian Svalbard area   "      "     "     age 1
N-SVT2     Norwegian   "      "      "      "     "     age 2
N-SVT3     Norwegian   "      "      "      "     "     age 3
N-BSA1     Norwegian Barents Sea   Acoustic survey      age 1
N-BSA2     Norwegian   "     "       "       "          age 2
N-BSA3     Norwegian   "     "       "       "          age 3
```

Table 3.2.10 Northeast Arctic Cod. RCT3 Retrospective Analysis.

Year-Cl	VPA	CAL +SH	CAL NSH	FUN +SH	FUN NSH	PRED +SH	PRED NSH
1972	622	1076	1256	993	1107	913	996
1973	614	818	867	793	828	764	792
1974	348	288	262	322	304	357	341
1975	639	567	571	580	584	571	574
1976	199	279	257	320	306	365	354
1977	140	263	245	294	282	325	315
1978	158	215	198	244	232	270	260
1979	158	202	190	224	216	242	235
1980	169	142	131	162	154	180	173
1981	386	198	191	202	197	214	210
1982	508	490	511	455	465	414	420
1983	894	819	865	765	798	635	654
1984	282	481	485	467	470	457	460
1985	230	423	427	413	415	404	406
1986	216	167	158	193	188	217	213
1987		136	128	158	153	186	182
1988		230	225	237	235	257	255
1989		377	385	355	359	348	351
1990		428	473	395	412	359	368
1991		389	538	381	426	334	349
RMSLR 72-86		0.392	0.409	0.404	0.407	0.433	0.429

Table 3.5.1 Western Channel Sole. XSA2 unshrunk.

	Age 2	3	4	5	6	7	8	All	Recruits	Partial recruits	Fully recruited
F ratio											
70 < p	5	2	1	1	2	3	5	19	5	2	12
50 < p <= 70	1	2	4	6	4	5	4	26	1	2	23
30 < p <= 50	1	4	4	2	3	1	1	16	1	4	11
10 < p <= 30	0	1	0	1	1	2	2	7	0	1	6
-10 < p <= 10	3	3	4	3	3	2	1	19	3	3	13
-30 < p <= -10	2	1	0	0	0	0	0	3	2	1	0
-50 < p <= -30	0	0	0	0	0	0	0	0	0	0	0
-70 < p <= -50	1	0	0	0	0	0	0	1	1	0	0
p <= -70	0	0	0	0	0	0	0	0	0	0	0
Outliers											
|p| > 50	7	4	5	7	6	8	9	46	7	4	35
|p| < 50	6	9	8	6	7	5	4	45	6	9	30
Total	13	13	13	13	13	13	13	91	13	13	65
Mean	44	39	37	37	43	51	58	44	44	39	45
Std.	67	35	28	26	30	31	35	38	67	35	30

Table 3.5.2 Western Channel Sole. XAS2 shrinkage CV=0.1.

	Age 2	3	4	5	6	7	8	All	Recruits	Partial recruits	Fully recruited
F ratio											
70 < p	0	0	0	0	0	0	1	1	0	0	1
50 < p <= 70	0	1	0	1	0	1	0	3	0	1	2
30 < p <= 50	0	0	1	1	3	1	1	7	0	0	7
10 < p <= 30	2	1	2	1	0	1	3	10	2	1	7
-10 < p <= 10	2	3	3	4	3	2	2	19	2	3	14
-30 < p <= -10	1	2	1	0	0	2	0	6	1	2	3
-50 < p <= -30	2	0	0	0	1	0	0	3	2	0	1
-70 < p <= -50	2	3	2	0	1	1	1	10	2	3	5
p <= -70	4	3	4	6	5	5	5	32	4	3	25
Outliers											
|p| > 50	6	7	6	7	6	7	7	46	6	7	33
|p| < 50	7	6	7	6	7	6	6	45	7	6	32
Total	13	13	13	13	13	13	13	91	13	13	65
Mean	-69	-61	-68	-71	-76	-80	-81	-72	-69	-61	-75
Std.	98	103	118	139	134	149	157	126	98	103	136

Table 3.5.3 Western Channel Sole. XAS2 shrinkage CV=0.5.

	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	All	Age groups: Recruits	Age groups: Partial recruits	Age groups: Fully recruited
F ratio											
70 < p	2	0	0	1	0	0	1	4	2	0	2
50 < p <= 70	1	1	0	0	1	1	0	4	1	1	2
30 < p <= 50	2	2	2	1	2	1	2	12	2	2	8
10 < p <= 30	1	3	5	3	2	5	2	21	1	3	17
-10 < p <= 10	1	4	5	6	6	3	6	31	1	4	26
-30 < p <= -10	2	3	0	1	1	2	1	10	2	3	5
-50 < p <= -30	4	0	1	1	1	1	0	8	4	0	4
-70 < p <= -50	0	0	0	0	0	0	1	1	0	0	1
p <= -70	0	0	0	0	0	0	0	0	0	0	0
Outliers											
|p| > 50	3	1	0	1	1	1	2	9	3	1	5
|p| < 50	10	12	13	12	12	12	11	82	10	12	60
Total	13	13	13	13	13	13	13	91	13	13	65
Mean	10	11	9	9	10	10	12	10	10	11	10
Std.	48	26	22	28	28	27	36	31	48	26	28

Table 3.5.4 4VSW Cod. XSA2 unshrunk.

	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	All	Age groups: Recruits	Age groups: Partial recruits	Age groups: Fully recruited
F ratio											
70 < p	0	1	0	0	0	0	0	1	0	1	0
50 < p <= 70	1	0	1	0	0	0	0	2	1	1	0
30 < p <= 50	0	0	0	0	0	0	0	0	0	0	0
10 < p <= 30	1	0	1	1	1	1	0	5	1	2	2
-10 < p <= 10	4	3	1	3	1	2	2	16	4	7	5
-30 < p <= -10	0	2	1	1	1	0	2	7	0	4	3
-50 < p <= -30	0	0	2	0	3	2	2	9	0	2	7
-70 < p <= -50	1	1	1	2	0	1	1	7	1	4	2
p <= -70	2	2	2	2	3	3	2	16	2	6	8
Outliers											
|p| > 50	4	4	4	4	3	4	3	26	4	12	10
|p| < 50	5	5	5	5	6	5	6	37	5	15	17
Total	9	9	9	9	9	9	9	63	9	27	27
Mean	-29	-34	-35	-38	-45	-45	-39	-38	-29	-36	-43
Std.	81	76	62	46	43	44	37	55	81	60	40

Table 3.5.5 4VSW Cod. XSA2 shrinkage CV=0.1.

	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	All	Age groups: Recruits	Age groups: Partial recruits	Age groups: Fully recruited
F ratio											
70 < p	5	2	0	0	0	0	0	7	5	2	0
50 < p <= 70	0	1	0	0	0	0	0	1	0	1	0
30 < p <= 50	0	1	2	0	1	0	0	4	0	3	1
10 < p <= 30	0	1	2	2	1	2	2	10	0	5	5
-10 < p <= 10	2	2	4	6	4	3	1	22	2	12	8
-30 < p <= -10	1	1	1	1	1	2	3	10	1	3	6
-50 < p <= -30	1	1	0	0	2	2	1	7	1	1	5
-70 < p <= -50	0	0	0	0	0	0	2	2	0	0	2
p <= -70	0	0	0	0	0	0	0	0	0	0	0
Outliers											
|p| > 50	5	3	0	0	0	0	2	10	5	3	2
|p| < 50	4	6	9	9	9	9	7	53	4	24	25
Total	9	9	9	9	9	9	9	63	9	27	27
Mean	101	31	12	3	-6	-9	-20	16	101	15	-11
Std.	128	48	20	12	26	26	30	65	128	32	27

Table 3.5.6 4VSW Cod. XSA2 shrinkage CV=0.4.

	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	All	Age groups: Recruits	Age groups: Partial recruits	Age groups: Fully recruited
F ratio											
70 < p	4	0	0	0	0	0	0	4	4	0	0
50 < p <= 70	1	1	0	0	0	0	0	2	1	1	0
30 < p <= 50	0	0	0	0	1	0	0	1	0	0	1
10 < p <= 30	0	2	1	1	0	2	2	8	0	4	4
-10 < p <= 10	2	3	5	3	3	3	1	20	2	11	7
-30 < p <= -10	1	2	2	3	2	1	3	14	1	7	6
-50 < p <= -30	0	1	1	2	3	3	3	13	0	4	9
-70 < p <= -50	0	0	0	0	0	0	0	0	0	0	0
p <= -70	1	0	0	0	0	0	0	1	1	0	0
Outliers											
|p| > 50	6	1	0	0	0	0	0	7	6	1	0
|p| < 50	3	8	9	9	9	9	9	56	3	26	27
Total	9	9	9	9	9	9	9	63	9	27	27
Mean	58	5	-8	-11	-13	-12	-16	0	58	-4	-14
Std.	89	26	16	21	25	24	22	45	89	22	23

Table 3.5.7 Reykjavik Simulation 5. XSA2 unshrunk.

	Age									Age groups		
	3	4	5	6	7	8	9	10	All	Recruits	Partial recruits	Fully recruited
F ratio												
70 < p	0	0	0	0	0	0	0	0	0	0	0	0
50 < p <= 70	0	0	0	0	0	0	0	0	0	0	0	0
30 < p <= 50	1	0	0	0	0	0	0	0	1	1	0	0
10 < p <= 30	1	1	0	0	0	0	0	0	2	1	1	0
-10 < p <= 10	8	7	7	4	2	6	4	4	42	8	14	20
-30 < p <= -10	3	6	6	9	11	5	5	3	48	3	12	33
-50 < p <= -30	1	1	2	2	2	3	4	4	19	1	3	15
-70 < p <= -50	0	0	0	0	0	1	2	4	7	0	0	7
p <= -70	1	0	0	0	0	0	0	0	1	1	0	0
Outliers												
|p| > 50	1	0	0	0	0	1	2	4	8	1	0	7
|p| < 50	14	15	15	15	15	14	13	11	112	14	30	68
Total	15	15	15	15	15	15	15	15	120	15	30	75
Mean	-9	-9	-15	-15	-18	-20	-25	-31	-18	-9	-12	-22
Std.	27	14	13	11	11	17	19	20	18	27	14	17

Table 3.5.8 Reykjavik Simulation 5. XSA2 shrinkage CV=0.1.

	Age									Age groups		
	3	4	5	6	7	8	9	10	All	Recruits	Partial recruits	Fully recruited
F ratio												
70 < p	0	0	0	0	0	0	0	0	0	0	0	0
50 < p <= 70	1	0	0	0	0	0	0	2	3	1	0	2
30 < p <= 50	3	0	0	0	0	2	2	2	9	3	0	6
10 < p <= 30	3	2	2	3	3	2	2	2	19	3	4	12
-10 < p <= 10	4	8	6	7	7	3	4	2	41	4	14	23
-30 < p <= -10	2	5	6	5	4	7	3	2	34	2	11	21
-50 < p <= -30	0	0	1	0	1	1	4	4	11	0	1	10
-70 < p <= -50	2	0	0	0	0	0	0	1	3	2	0	1
p <= -70	0	0	0	0	0	0	0	0	0	0	0	0
Outliers												
|p| > 50	3	0	0	0	0	0	0	3	6	3	0	3
|p| < 50	12	15	15	15	15	15	15	12	114	12	30	72
Total	15	15	15	15	15	15	15	15	120	15	30	75
Mean	5	-2	-7	-3	-5	-1	-5	-2	-2	5	-5	-3
Std.	34	13	15	13	14	26	26	40	24	34	14	25

Table 3.5.9 Reykjavik Simulation 5. XSA2 shrinkage CV=0.5.

	Age								All	Age groups		
	3	4	5	6	7	8	9	10		Recruits	Partial recruits	Fully recruited
F ratio												
70 < p	0	0	0	0	0	0	0	0	0	0	0	0
50 < p <= 70	0	0	0	0	0	0	0	0	0	0	0	0
30 < p <= 50	1	0	0	0	0	0	0	0	1	1	0	0
10 < p <= 30	4	2	1	1	1	1	1	1	12	4	3	5
-10 < p <= 10	3	6	6	4	7	7	3	4	40	3	12	25
-30 < p <= -10	4	7	7	9	5	4	7	4	47	4	14	29
-50 < p <= -30	3	0	1	1	2	2	4	5	18	3	1	14
-70 < p <= -50	0	0	0	0	0	1	0	1	2	0	0	2
p <= -70	0	0	0	0	0	0	0	0	0	0	0	0
Outliers												
|p| > 50	0	0	0	0	0	1	0	1	2	0	0	2
|p| < 50	15	15	15	15	15	14	15	14	118	15	30	73
Total	15	15	15	15	15	15	15	15	120	15	30	75
Mean	-5	-6	-12	-11	-13	-14	-19	-21	-13	-5	-9	-16
Std.	24	12	13	12	14	19	16	20	17	24	13	16

Table 3.5.10 Reykjavik Simulation 5. TSA.

	Age								All	Age groups		
	3	4	5	6	7	8	9	10		Recruits	Partial recruits	Fully recruited
F ratio												
70 < p	0	0	0	0	0	0	0	0	0	0	0	0
50 < p <= 70	0	0	0	0	0	0	0	0	0	0	0	0
30 < p <= 50	0	0	0	0	0	0	0	0	0	0	0	0
10 < p <= 30	1	4	2	3	2	7	4	5	28	1	6	21
-10 < p <= 10	8	6	8	7	8	3	6	5	51	8	14	29
-30 < p <= -10	1	0	0	0	0	0	0	0	1	1	0	0
-50 < p <= -30	0	0	0	0	0	0	0	0	0	0	0	0
-70 < p <= -50	0	0	0	0	0	0	0	0	0	0	0	0
p <= -70	0	0	0	0	0	0	0	0	0	0	0	0
Outliers												
|p| > 50	0	0	0	0	0	0	0	0	0	0	0	0
|p| < 50	10	10	10	10	10	10	10	10	80	10	20	50
Total	10	10	10	10	10	10	10	10	80	10	20	50
Mean	3	6	4	6	7	11	11	11	7	3	5	9
Std.	9	6	8	6	4	5	7	7	7	9	7	6

Table 3.5.11

The percentage of terminal F estimates (all ages) which are in error by a log ratio of more than 0.5, the mean log ratio and r.m.s log ratio (s.e)

Western Channel (VIIe) Sole

OUTLIERS	SHRINKAGE CV 0.1	0.2	0.3	0.4	0.5	unshrunk
L/S	45	21	19	16	14	15
XSA2	51	38	23	15	10	51
XSA4 HM	44	23	14	12	19	51
XSA4 SM	18	15	14	15	19	51
XSA4 SU	7	7	10	8		51

BIAS	SHRINKAGE CV 0.1	0.2	0.3	0.4	0.5	unshrunk
L/S	-50	-14	-5	3	7	18
XSA2	-72	-42	-12	4	10	44
XSA4 HM	-41	-12	9	19	23	44
XSA4 SM	22	18	20	22	23	44
XSA4 SU	4	-1	-2	0		44

SE	SHRINKAGE CV 0.1	0.2	0.3	0.4	0.5	unshrunk
L/S	106	42	36	32	32	36
XSA2	126	85	49	33	31	38
XSA4 HM	60	43	31	29	30	38
XSA4 SM	30	31	30	30	30	38
XSA4 SU	25	27	28	29		38

Table 3.5.12

The percentage of terminal population estimates (all ages) which are in error by a log ratio of more than 0.5, the mean log ratio and r.m.s log ratio (s.e) .

Western Channel (VIIe) Sole

		SHRINKAGE CV					
OUTLIERS		0.1	0.2	0.3	0.4	0.5	unshrunk
	L/S	37	21	20	15	11	12
	XSA2	39	41	26	15	10	39
	XSA4 HM	43	23	14	10	12	38
	XSA4 SM	14	12	13	14	12	38
	XSA4 SU	4	7	8	9		38

		SHRINKAGE CV					
BIAS		0.1	0.2	0.3	0.4	0.5	unshrunk
	L/S	46	13	5	-1	-5	-14
	XSA2	70	42	13	-3	-9	-41
	XSA4 HM	40	14	-6	-16	-21	-41
	XSA4 SM	-20	-16	-18	-20	-21	-41
	XSA4 SU	-2	3	4	2		-41

		SHRINKAGE CV					
SE		0.1	0.2	0.3	0.4	0.5	unshrunk
	L/S	108	44	38	33	32	33
	XSA2	126	90	53	36	32	38
	XSA4 HM	65	46	32	28	29	34
	XSA4 SM	29	30	28	28	29	34
	XSA4 SU	24	27	28	28		34

Table 3.5.13

The percentage of terminal F estimates (all ages) which are in error by a log ratio of more than 0.5, the mean log ratio and r.m.s log ratio (s.e).

4Vs W Cod

		SHRINKAGE CV					
OUTLIERS		0.1	0.2	0.3	0.4	0.5	unshrunk
	L/S	25	29	29	32	36	55
	XSA2	16	14	13	11	14	41
	XSA4 HM	25	21	16	16	24	40
	XSA4 SM	24	24	24	24	24	40
	XSA4 SU	22	22	22	19	19	40

		SHRINKAGE CV					
BIAS		0.1	0.2	0.3	0.4	0.5	unshrunk
	L/S	-4	-11	-18	-26	-32	-54
	XSA2	16	13	7	0	-5	-38
	XSA4 HM	43	34	23	11	3	-38
	XSA4 SM	-19	-18	-17	-17	-19	-38
	XSA4 SU	3	3	4	2	1	-38

		SHRINKAGE CV					
SE		0.1	0.2	0.3	0.4	0.5	unshrunk
	L/S	68	65	61	59	57	51
	XSA2	65	59	51	45	39	55
	XSA4 HM	76	71	66	60	55	55
	XSA4 SM	50	47	46	46	46	55
	XSA4 SU	50	48	46	44	44	55

Table 3.5.14

The percentage of terminal population estimates (all ages) which are in error by a log ratio of more than 0.5, the mean log ratio and r.m.s log ratio (s.e) for all ages.

4Vs W Cod

OUTLIERS	SHRINKAGE CV 0.1	0.2	0.3	0.4	0.5	unshrunk
L/S	25	22	21	24	29	48
XSA2	21	16	14	14	12	32
XSA4 HM	20	17	14	12	15	32
XSA4 SM	15	17	17	16	16	32
XSA4 SU	12	12	12	12	10	32

BIAS	SHRINKAGE CV 0.1	0.2	0.3	0.4	0.5	unshrunk
L/S	2	7	12	28	22	38
XSA2	-14	-11	-8	-4	0	33
XSA4 HM	-31	-25	-17	-9	-1	33
XSA4 SM	13	13	13	13	14	33
XSA4 SU	0	-1	-1	-1	1	33

SE	SHRINKAGE CV 0.1	0.2	0.3	0.4	0.5	unshrunk
L/S	66	62	58	50	54	60
XSA2	56	51	44	38	34	67
XSA4 HM	70	66	61	56	51	65
XSA4 SM	57	56	56	57	59	65
XSA4 SU	58	58	56	57	59	65

Table 3.5.15

The percentage of terminal F estimates (all ages) which are in error by a log ratio of more than 0.5, the mean log ratio and r.m.s log ratio (s.e)

Reykjavik simulated data set 5.

OUTLIERS		SHRINKAGE CV 0.1	0.2	0.3	0.4	0.5	unshrunk
	L/S	12	6	6	8	8	9
	XSA2	5	2	1	1	2	7
	XSA4 HM	3	1	1	1	2	7
	XSA4 SM	3	3	2	1	2	7
	XSA4 SU	2	2	1	0	0	7

BIAS		SHRINKAGE CV 0.1	0.2	0.3	0.4	0.5	unshrunk
	L/S	-7	-8	-9	-10	-10	-11
	XSA2	-2	-9	-12	-13	-13	-18
	XSA4 HM	-5	-12	-14	-14	-13	-17
	XSA4 SM	-13	-14	-14	-14	-13	-17
	XSA4 SU	-9	-11	-12	-12	-12	-17

SE		SHRINKAGE CV 0.1	0.2	0.3	0.4	0.5	unshrunk
	L/S	35	28	26	26	26	27
	XSA2	24	19	17	17	17	18
	XSA4 HM	19	17	16	17	17	18
	XSA4 SM	17	17	17	17	17	18
	XSA4 SU	15	15	15	16	17	18

Table 3.5.16

The percentage of terminal population estimates (all ages) which are in error by a log ratio of more than 0.5, the mean log ratio and r.m.s log ratio (s.e)

Reykjavik simulated data set 5.

OUTLIERS	SHRINKAGE CV 0.1	0.2	0.3	0.4	0.5	unshrunk
L/S	16	12	11	9	9	10
XSA2	11	6	6	6	5	8
XSA4 HM	6	5	6	4	4	8
XSA4 SM	6	6	6	4	4	8
XSA4 SU	6	5	5	4	4	8

BIAS	SHRINKAGE CV 0.1	0.2	0.3	0.4	0.5	unshrunk
L/S	5	5	6	5	6	7
XSA2	1	7	10	11	11	16
XSA4 HM	3	10	12	12	11	16
XSA4 SM	11	12	13	12	11	16
XSA4 SU	8	9	10	10	10	16

SE	SHRINKAGE CV 0.1	0.2	0.3	0.4	0.5	unshrunk
L/S	49	43	40	38	38	30
XSA2	39	36	35	35	35	35
XSA4 HM	35	34	35	35	35	36
XSA4 SM	35	35	35	35	35	36
XSA4 SU	33	33	34	34	34	36

Table 4.1.1 Chi-square evaluation of 72 stocks from Myers *et al.* (WP10) stock and recruitment database. Note that the p-values for the chi-square p-value uses Yates' continuity correction, so the p, if all cells have the same entry, will not be 0, e.g. hake from South Africa.

	below median SSB	above median SSB
above median Recruitment	a	b
below median Recruitment	c	d

If the relationship between Recruitment and SSB is positive then the p-value of the Chi square test is signed positive.

ID	Stock	a	b	c	d	p-value
ANCHOCAL	Northern anchovy California	7	5	6	7	- 0.835
BWHITNA	Blue whiting Northern ICES	6	4	4	6	- 0.655
COD1	Cod NAFO 1	5	10	11	5	+ 0.107
COD2J3KL	Cod NAFO 2J3KL	5	8	9	5	+ 0.339
COD3NO	Cod NAFO 3NO	5	9	9	5	+ 0.257
COD3Ps	Cod NAFO 3Ps	6	7	7	6	+ 1
COD4TVn	Cod NAFO 4TVn	8	10	11	8	+ 0.625
COD4VsW	Cod NAFO 4VsW	7	8	9	7	+ 0.862
COD4X	Cod NAFO 4X	11	9	10	11	+ 0.873
CODBA2224	Cod Baltic Areas 22 and 24	3	7	7	3	+ 0.18
CODCS	Cod Celtic Sea	3	7	7	3	+ 0.18
CODFAPL	Cod Faroe Plateau	9	5	5	9	- 0.257
CODICE	Cod Iceland	7	10	10	7	- 0.493
CODIS	Cod Irish Sea	8	3	3	8	- 0.0881
CODNEAR	Cod North East Arctic	6	13	13	6	+ 0.0516
CODNS	Cod North Sea	7	6	7	7	- 0.853

continued ...

Table 4.1.1 (cont'd)

CODVIa	Cod ICES VIa	7	4	5	7	- 0.525
HAD4TVW	Haddock NAFO 4TVW	4	15	15	4	+ 0.00118
HAD4X	Haddock NAFO 4X	7	5	5	7	- 0.683
HAD5Z	Haddock NAFO 5Z	9	20	20	9	+ 0.00864
HADFAPL	Haddock Faroe Plateau	10	3	4	10	- 0.0334
HADICE	Haddock Iceland	8	6	6	8	- 0.705
HADNEAR	Haddock North East Arctic	7	12	13	7	+ 0.15
HADNS	Haddock North Sea	7	8	8	7	1
HADVIa	Haddock VIa	5	7	7	5	+ 0.683
PHAKEVANCON	Pacific hake W. US. + Canada	7	8	8	7	1
HERRCC	Herring Central Coast B.C.	10	9	9	10	1
HERREBER	Herring Eastern Bering Sea	9	4	4	9	- 0.117
HERRGM	Herring Gulf of Maine	6	5	6	6	0.842
HERRIspr	Herring Iceland (Spring spawners)	3	8	9	3	+ 0.0613
HERRIsum	Herring Iceland (Summer spawners)	5	16	17	5	+ 0.00137
HERRNOR	Herring Norway (Spring spawners)	7	12	13	7	+ 0.15
HERRNS	Herring North Sea	8	12	13	8	+ 0.276
HERRNSG	Herring North Strait of Georgia	7	12	12	7	+ 0.194
HERRNWCVI	Herring North West Coast Vancouver Island	8	11	11	8	+ 0.516
HERRPRD	Herring Prince Rupert District	9	10	10	9	+ 1
HERRQCI	Herring Queen Charlotte Islands	9	10	10	9	+ 1
HERRSEALA	Herring S.E. Alaska	9	6	6	9	+ 0.465
HERRSSG	Herring Southern Strait of Georgia	9	10	10	9	- 1
HERRSWCVI	Herring South West Coast Vancouver Island	11	8	8	11	+ 0.516
MACK2-6	Mackerel NAFO 2 to 6	6	8	8	6	+ 0.705
MENATLAN	Atlantic Menhaden U.S. Atlantic	8	9	10	8	+ 0.869
PHALCANUS	Pacific halibut Pacific	11	12	13	11	+ 0.886
PLAICIS	Plaice Irish Sea	7	6	6	7	+ 1
PLAICKAT	Plaice Kattegat	3	8	8	3	+ 0.0881
PLAICNS	Plaice North Sea	12	4	5	12	- 0.0232
PMACKCAL	Pacific mackerel Southern California	7	11	11	7	+ 0.317
POLLFA	Pollock or saithe Faroe	9	5	5	9	- 0.257
POLLICE	Pollock or saithe Iceland	6	7	7	6	+ 1
POLLNEAR	Pollock or saithe North East Arctic	3	7	8	3	+ 0.128
POLLNS	Pollock or saithe North Sea	4	6	7	4	+ 0.518

continued...

Table 4.1.1 (cont'd)

POLLVI	Pollock or saithe ICES VI	7	3	4	6	- 0.369
SAHAKE	Hake South Africa 1.6	5	5	5	5	0.655
SAPILCH	Southern african pilchard South Africa	5	10	11	5	+ 0.107
SARDCAL	Pacific sardine California	1	14	15	1	+ 7.16e-06
SHAKE5Ze	Silver hake NAFO 5Ze	4	12	13	4	+ 0.0091
SHAKEMAB	Silver hake Mid Atlantic Bight	4	12	13	4	+ 0.0091
SOCKADAM	Sockeye salmon Adams Complex,B.C., Canada	1	18	18	1	+ 2.09e-07
SOCKBIRK	Sockeye salmon Birkenhead River, B.C., Canada	8	10	11	8	+ 0.625
SOCKCHIK	Sockeye salmon Chilko River, B.C., Canada	3	16	16	3	+ 9.89e-05
SOCKHFLY	Sockeye salmon Horsefly River, B.C., Canada	1	18	18	1	+ 2.09e-07
SOCKRINL	Sockeye salmon Rivers Inlet, B.C., Canada	8	8	10	10	0.737
SOCKSKEE	Sockeye salmon Skeena River, B.C., Canada	8	14	15	8	+ 0.102
SOCKSTEL	Sockeye salmon Stellako River, B.C., Canada	2	17	17	2	5.57e-06
SOCKSTUA	Sockeye salmon Early Stuart Complex, B.C., Canada	5	14	14	5	+ 0.00944
SOLEIS	Sole Irish Sea	7	3	3	7	- 0.18
SOLENS	Sole North Sea	11	6	6	11	- 0.17
SOLEVIIe	Sole ICES VIIe	2	9	9	2	+ 0.0105
WHITNS	Whiting North Sea	6	7	7	6	+ 1
WHITVIa	Whiting ICES VIa	8	4	5	8	- 0.313
WPOLLEBS	Walleye pollock E. Bering Sea	6	5	6	7	1
WPOLLGA	Walleye pollock Gulf of Alaska	7	3	4	7	- 0.27

Table 5.2.1 Uncertainties used in FAST analysis.

Age	Terminal Fs	Terminal Ns
3	1.18	1.42
4	1.17	1.61
5	1.17	1.80
6	1.18	1.93
7	1.21	2.08
8	1.14	2.23
9	1.45	2.32
10	1.76	2.55

Table 5.2.2 Coefficients of variation (CVs) used in the various methods for calculating probability profiles.

	Terminal Fs							
Method/Age	3	4	5	6	7	8	9	10
cov/SepVPA	.09	.09	.09	.09	.10	.07	.23	.38
Boot/ADAPT	.26	.18	.14	.24	.24	.24	.24	.24
Monte/ADAPT	.20	.20	.18	.25	.25	.25	.25	.25
	Terminal Populations							
cov/SepVPA	.21	.30	.40	.47	.54	.62	.66	.78
Boot/ADAPT	.18	.12	.06	.11	.11	.11	.11	.11
Monte/ADAPT	.13	.13	.09	.13	.14	.16	.16	.17

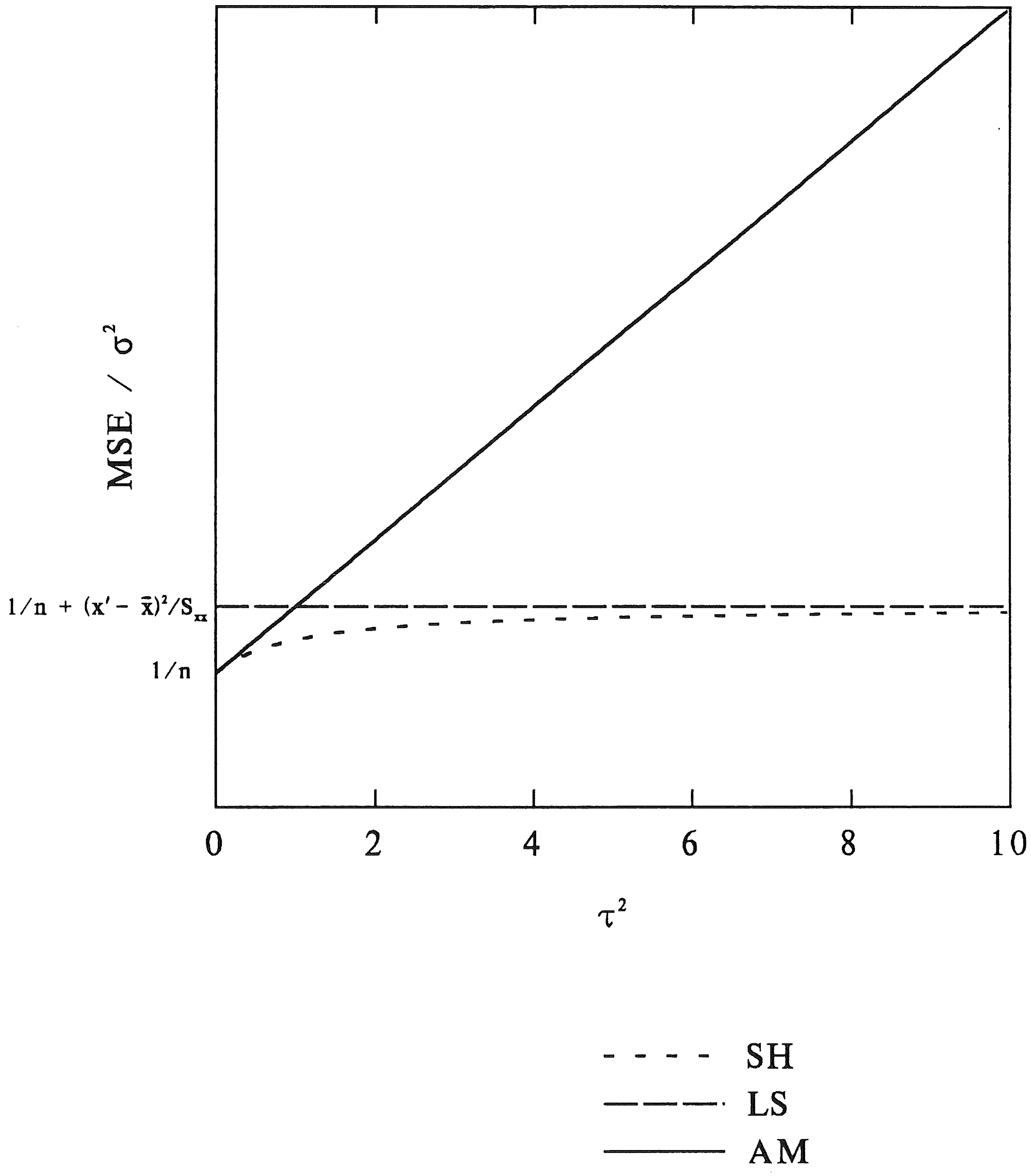


Figure 3.1.1 Mean square errors of three estimators of random variable y' as a function of τ^2. (SH - shrinkage; LS - Laurec-Shepherd; AM - arithmetic mean).

Figure 3.1.2

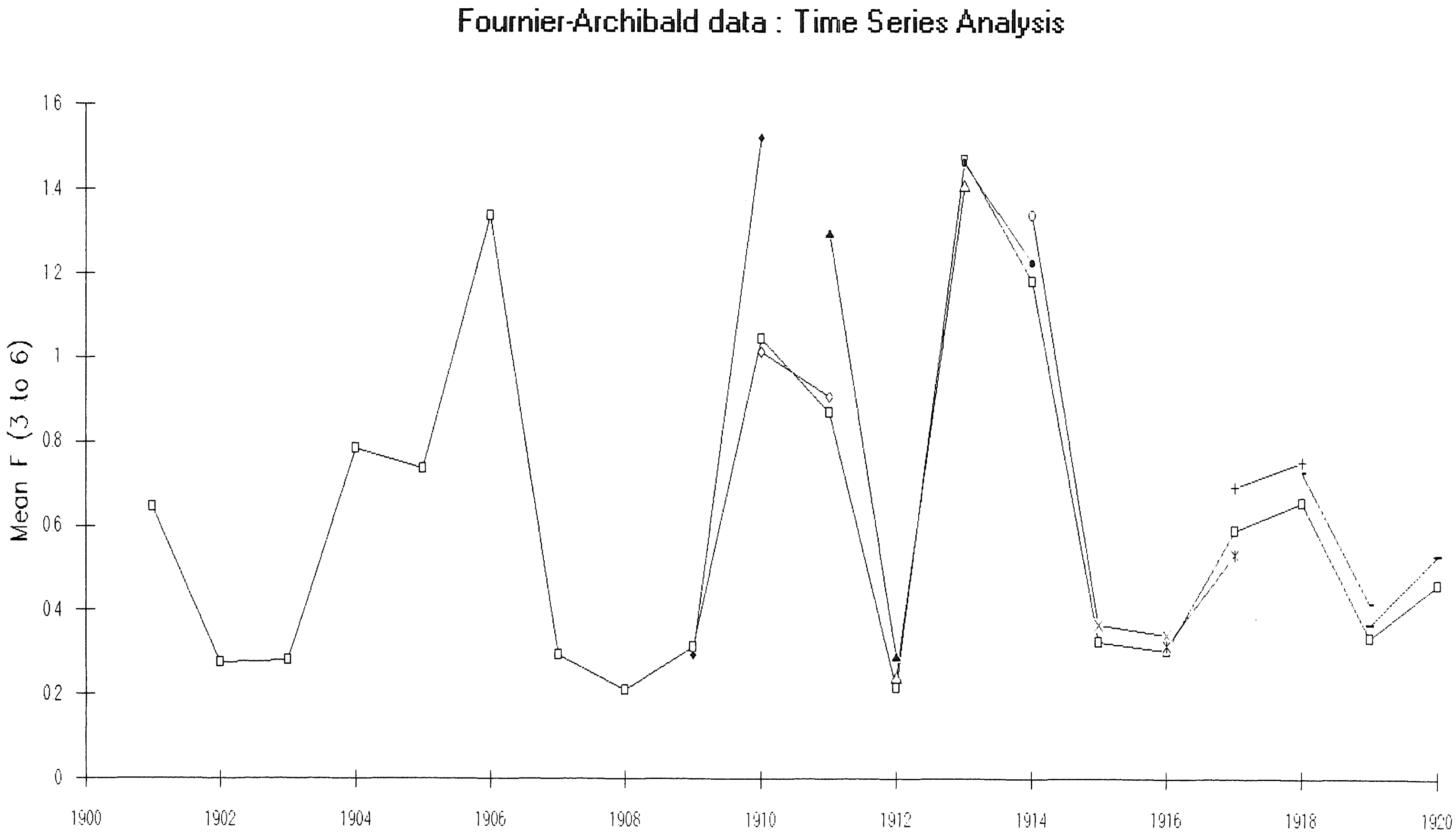

Fournier-Archibald data : Time Series Analysis
Mean F (3 to 6)
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1900
1902
1904
1906
1908
1910
1912
1914
1916
1918
1920

Figure 3.1.3

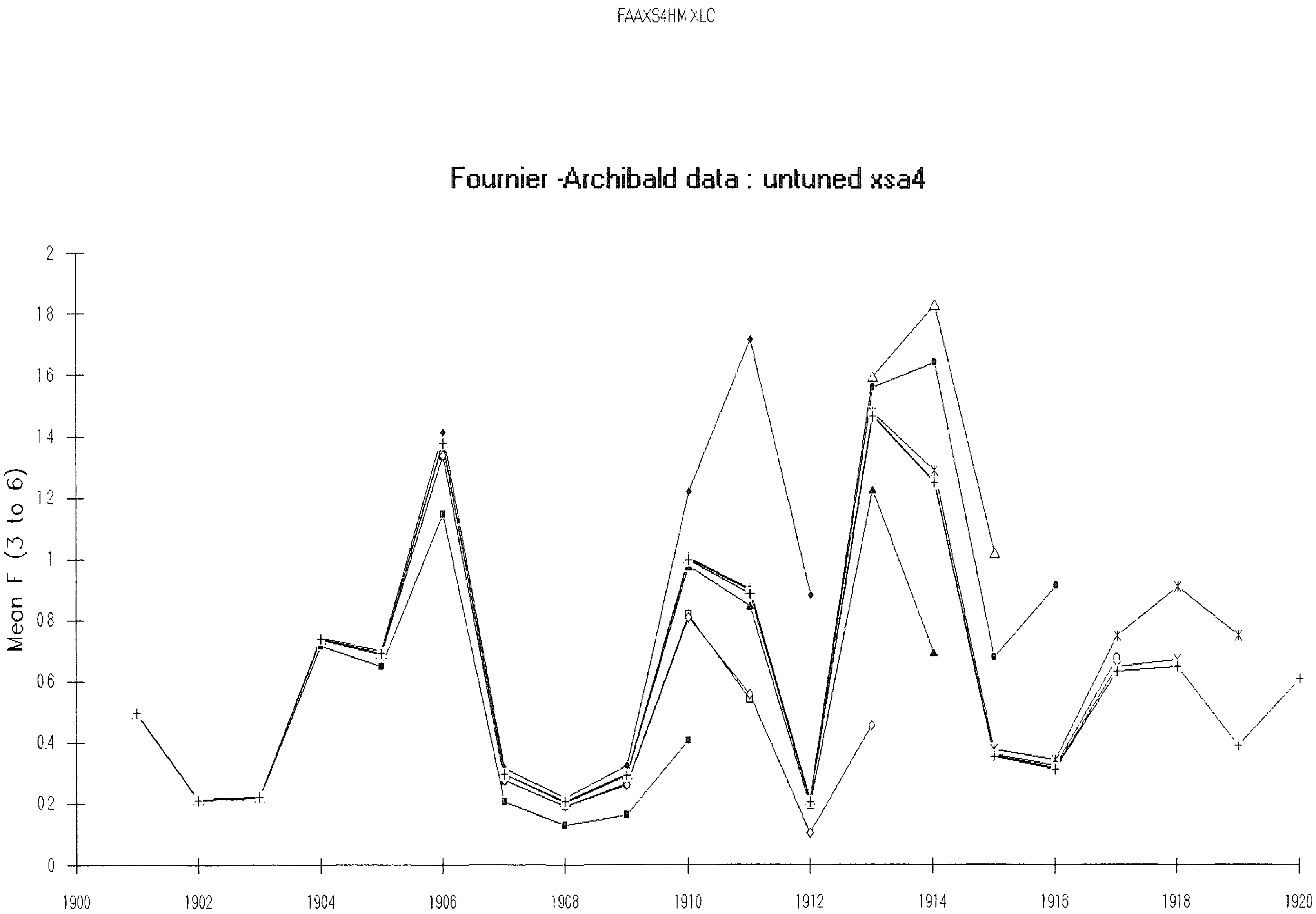

FAAXS4HM.XLC
Fournier -Archibald data : untuned xsa4
Mean F (3 to 6)
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1900
1902
1904
1906
1908
1910
1912
1914
1916
1918
1920

Figure 3.2.1

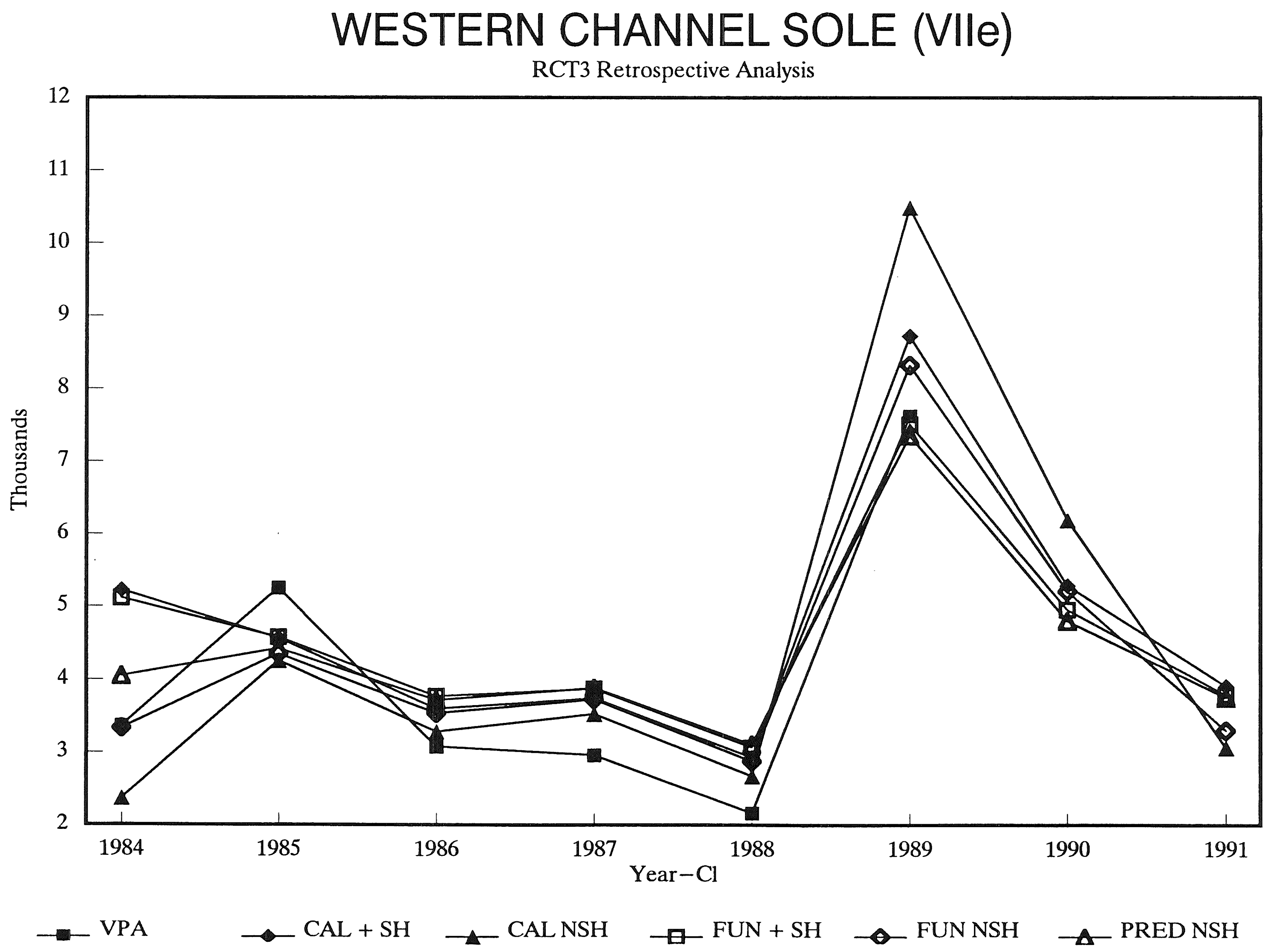

WESTERN CHANNEL SOLE (VIIe)
RCT3 Retrospective Analysis
Thousands
12
11
10
9
8
7
6
5
4
3
2
1984
1985
1986
1987
1988
1989
1990
1991
Year–Cl
VPA
CAL + SH
CAL NSH
FUN + SH
FUN NSH
PRED NSH

Figure 3.2.2

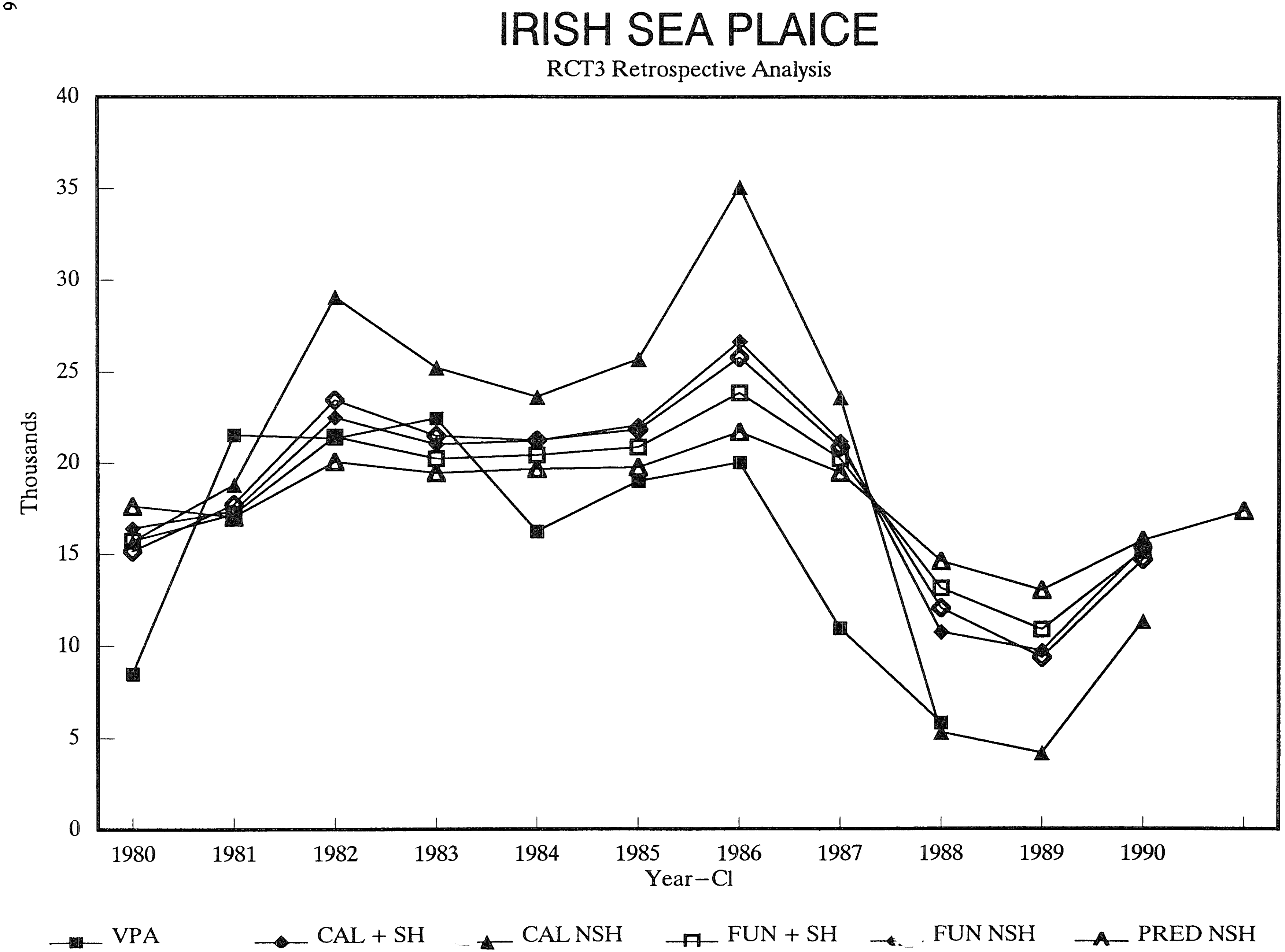

IRISH SEA PLAICE
RCT3 Retrospective Analysis
40
35
30
25
20
15
10
5
0
Thousands
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
Year–Cl
VPA
CAL + SH
CAL NSH
FUN + SH
FUN NSH
PRED NSH

Figure 3.2.3

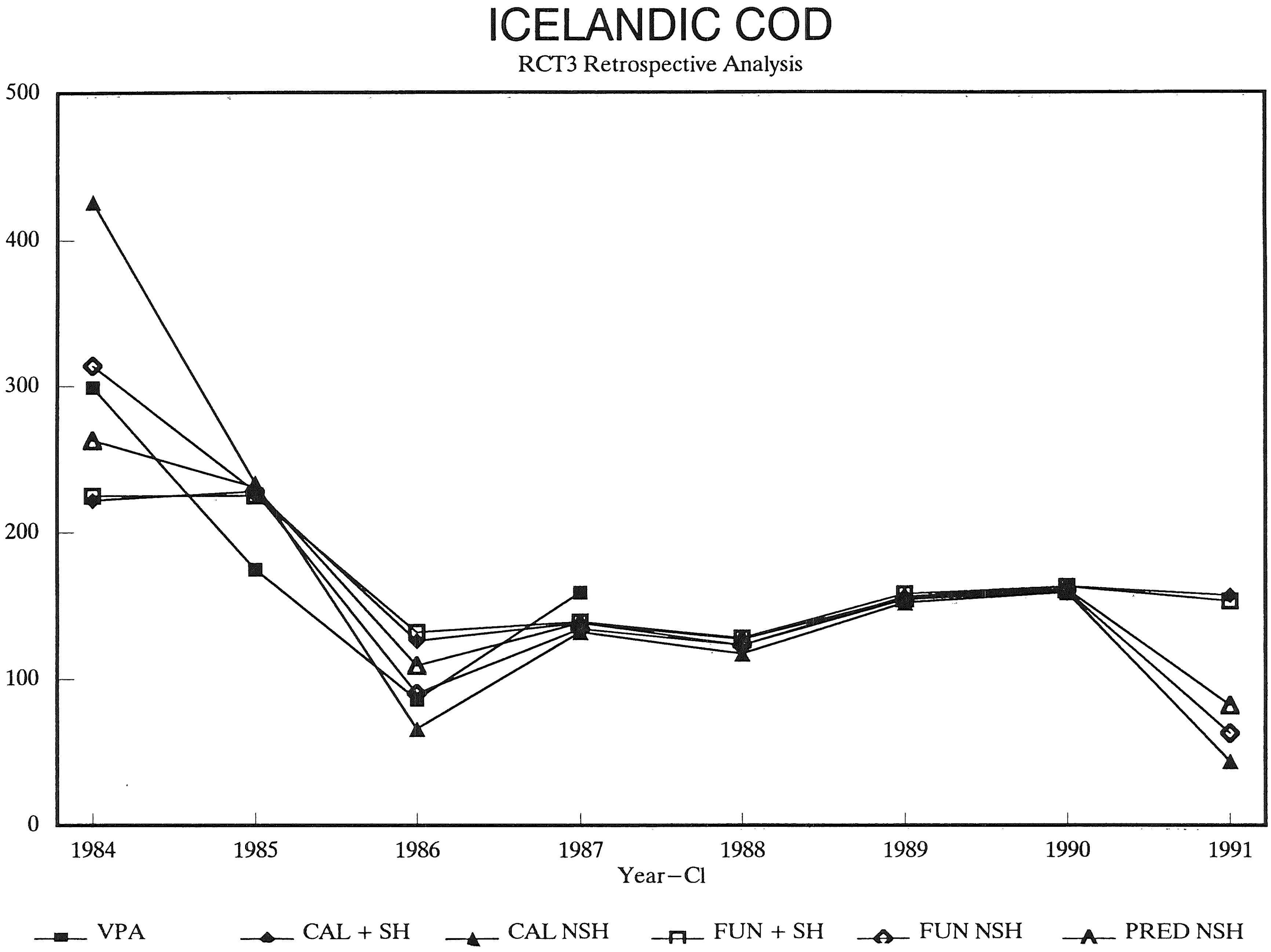
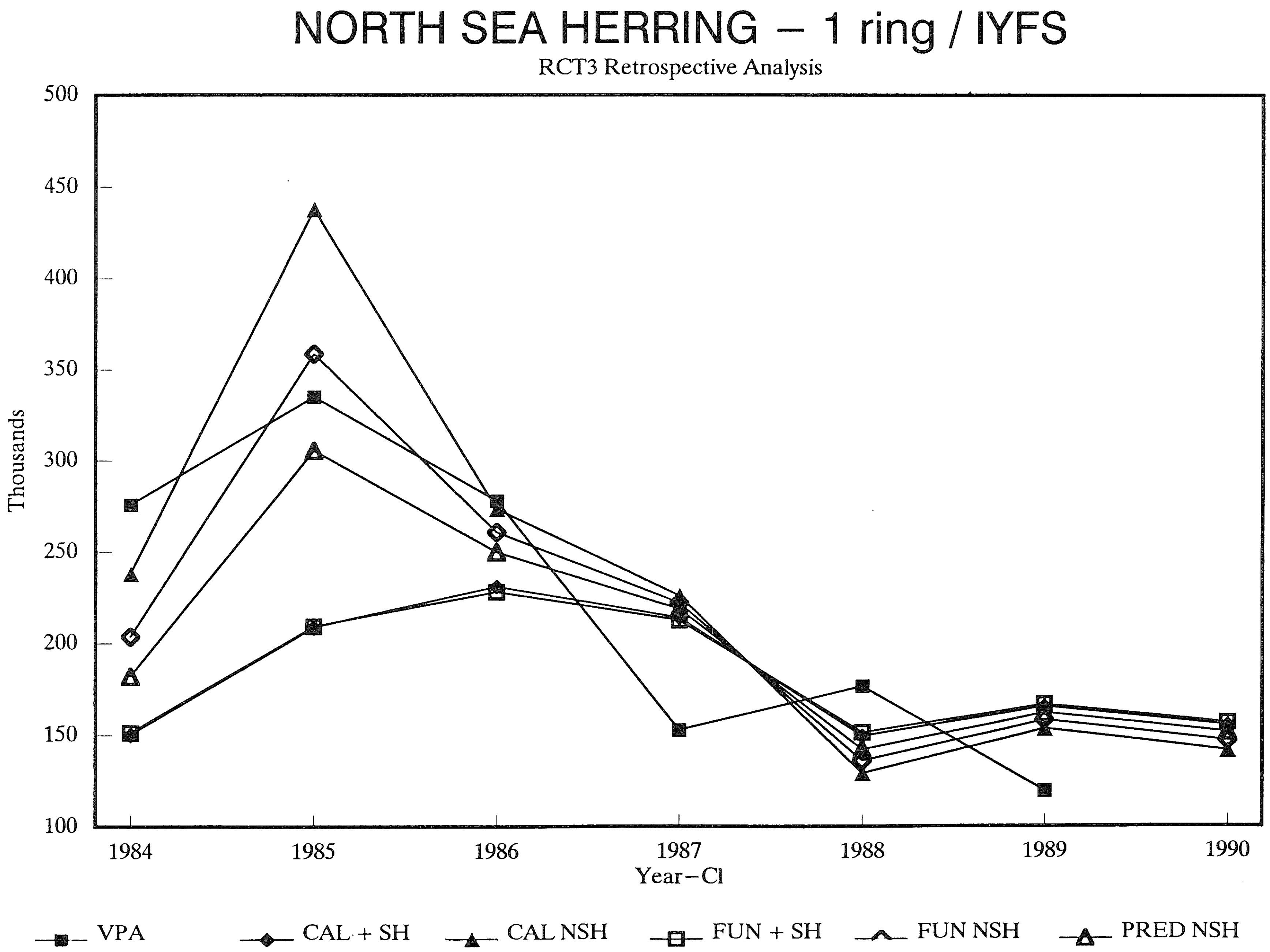
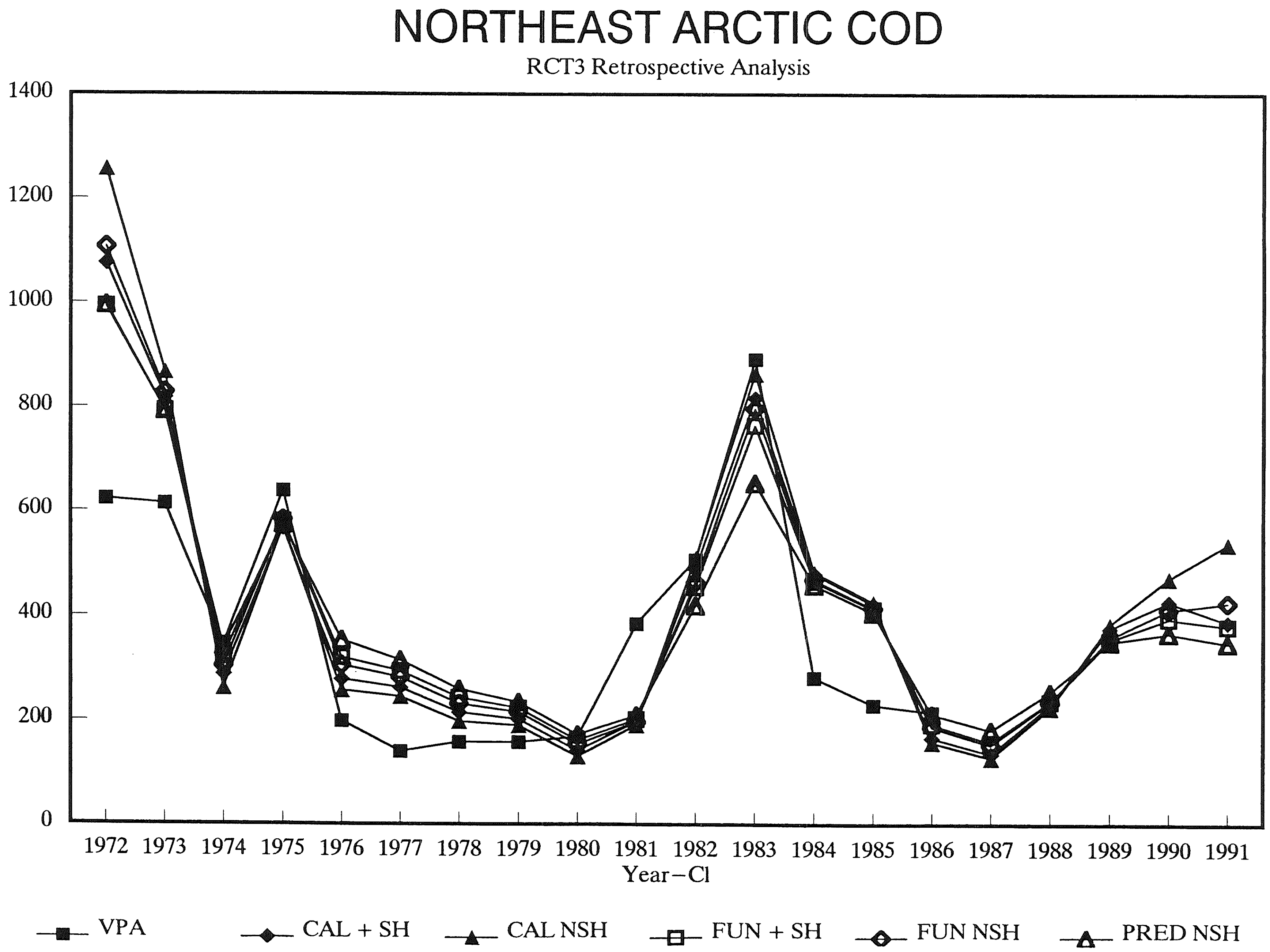

ICELANDIC COD
RCT3 Retrospective Analysis
500
400
300
200
100
0
1984
1985
1986
1987
1988
1989
1990
1991
Year−Cl
VPA
CAL + SH
CAL NSH
FUN + SH
FUN NSH
PRED NSH

Figure 3.2.4



NORTH SEA HERRING – 1 ring / IYFS
RCT3 Retrospective Analysis
500
450
400
350
300
250
200
150
100
Thousands
1984
1985
1986
1987
1988
1989
1990
Year–Cl
VPA
CAL + SH
CAL NSH
FUN + SH
FUN NSH
PRED NSH

Figure 3.2.5



NORTHEAST ARCTIC COD
RCT3 Retrospective Analysis
1400
1200
1000
800
600
400
200
0
1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991
Year – Cl
VPA
CAL + SH
CAL NSH
FUN + SH
FUN NSH
PRED NSH

Figure 3.5.1

Stock: Western Channel Sole

Procedure: XSA2 Unshrunk

Retrospective analysis
s7exsa2.s0

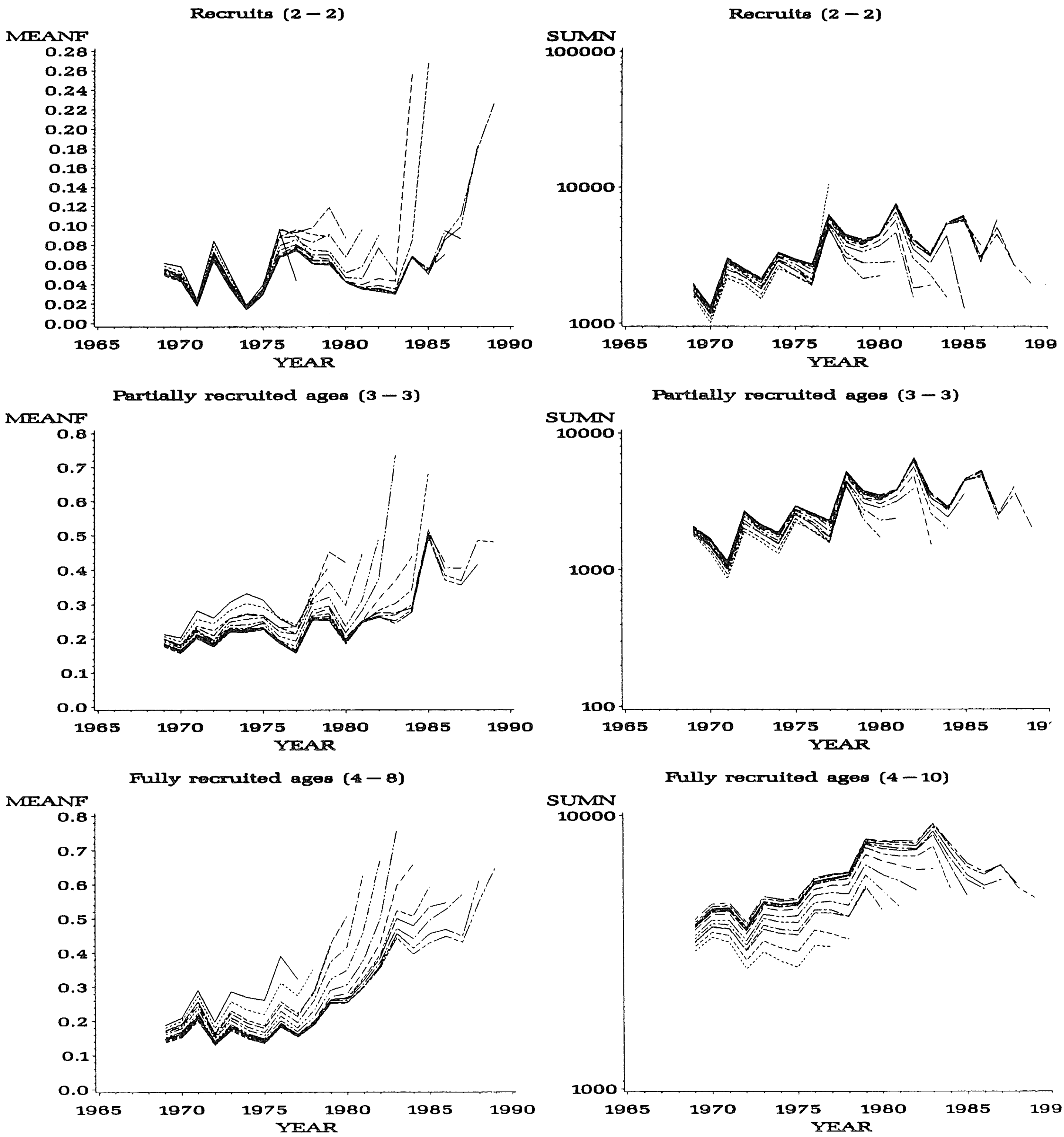


Figure 3.5.2

Stock: Western Channel Sole

Procedure: XSA2 Shrinkage CV = 0.1

Retrospective analysis

s7exsa2.s1

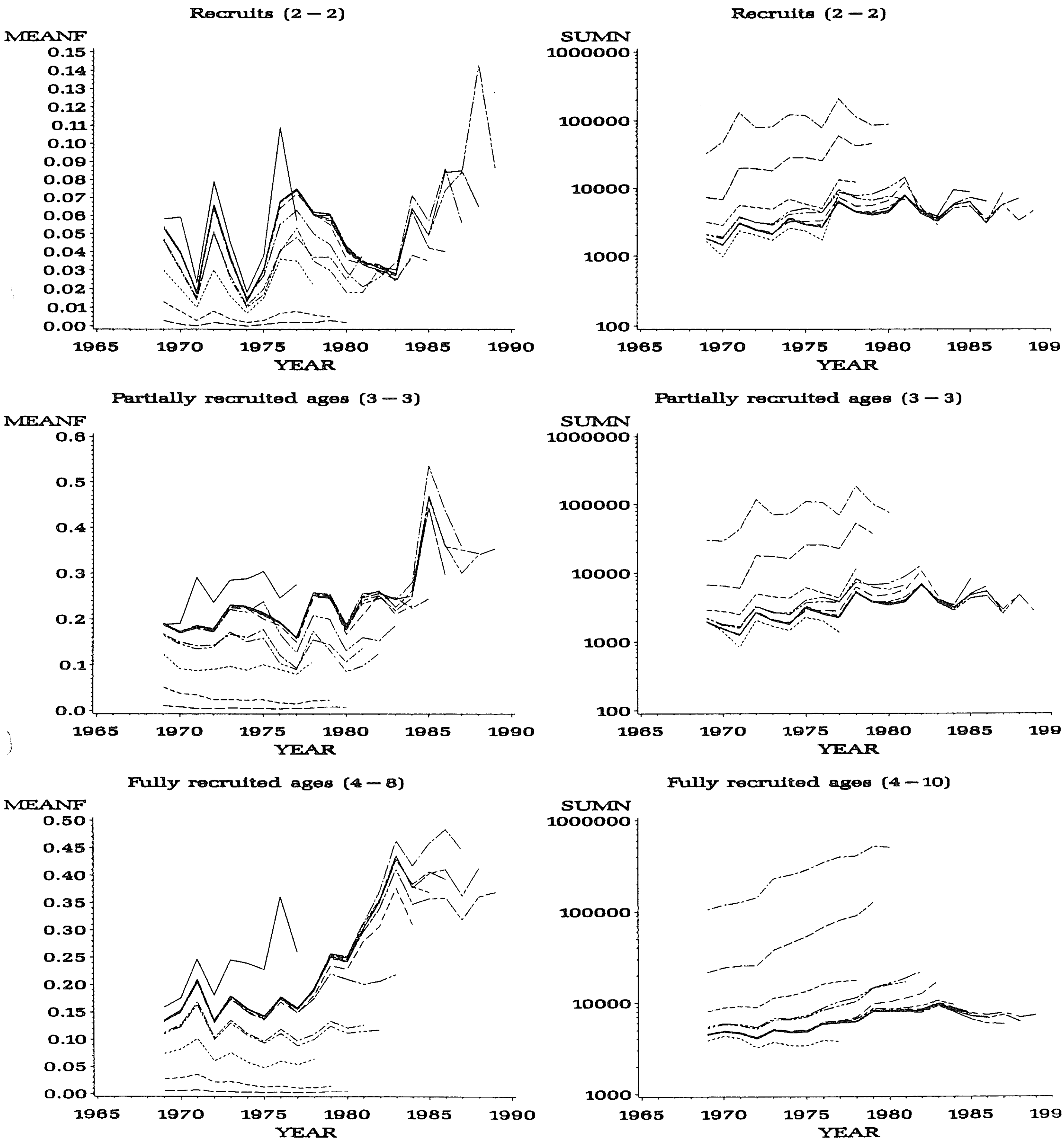


Figure 3.5.3

Stock: Western Channel Sole

Procedure: XSA2 Shrinkage CV = 0.5

Retrospective analysis

s7exsa2.s5

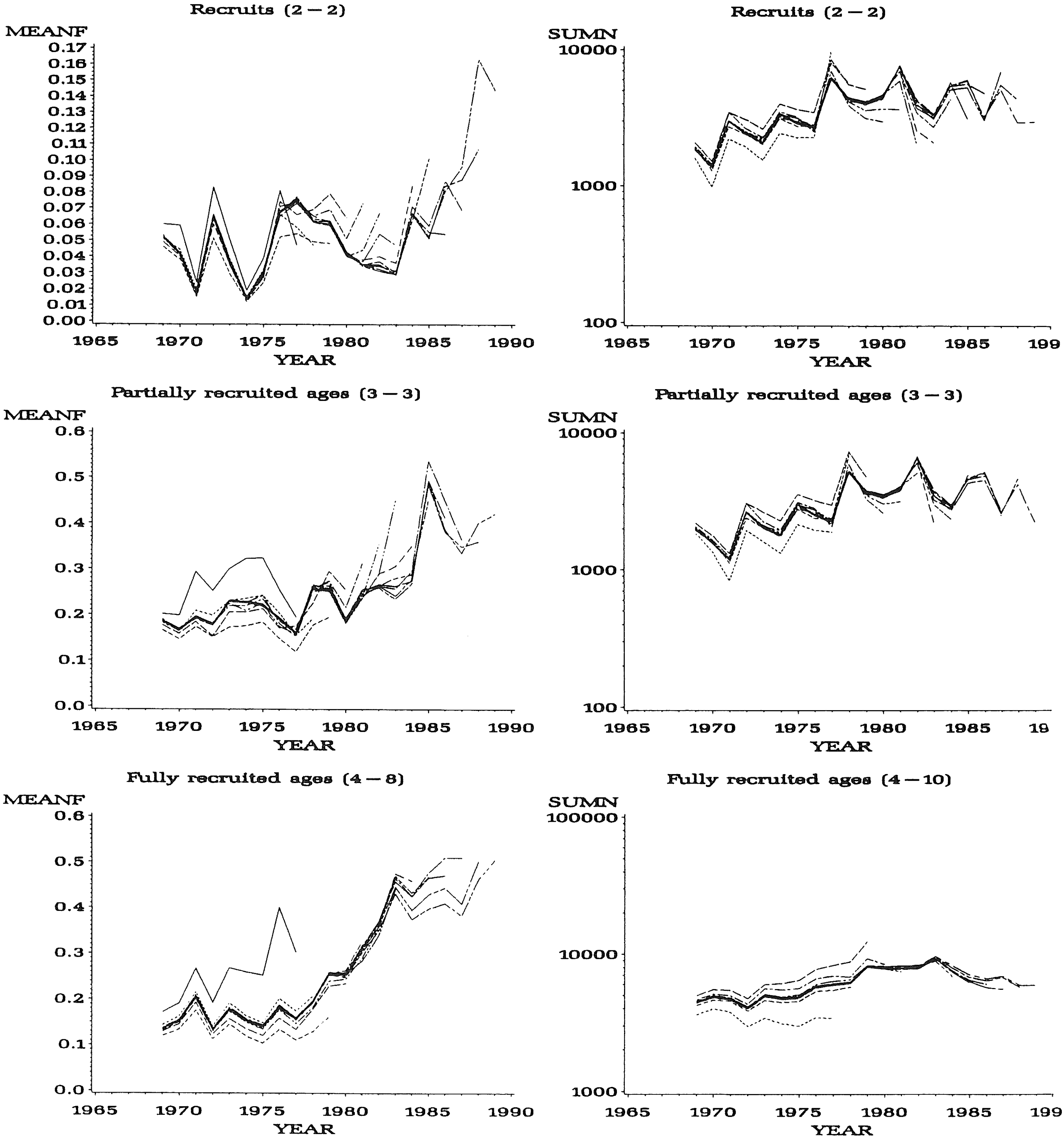


Figure 3.5.4

Stock: 4Vs W Cod

Procedure: XSA2 Unshrunk

Retrospective analysis

c4vxsa2.s0

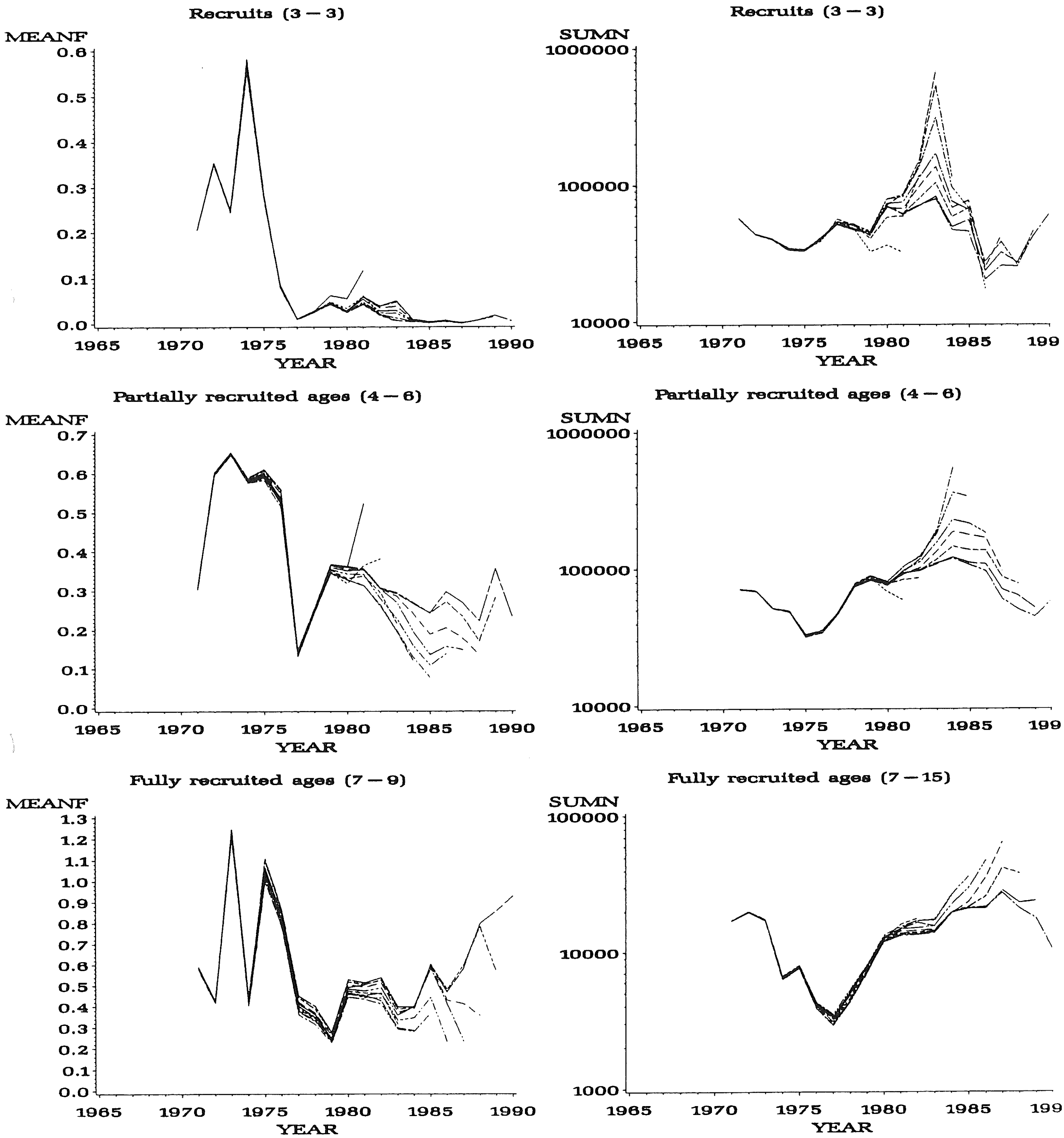


Figure 3.5.5

Stock: 4Vs W Cod

Procedure: XSA2 Shrinkage CV=0.1

Retrospective analysis
c4vxsa2.s1

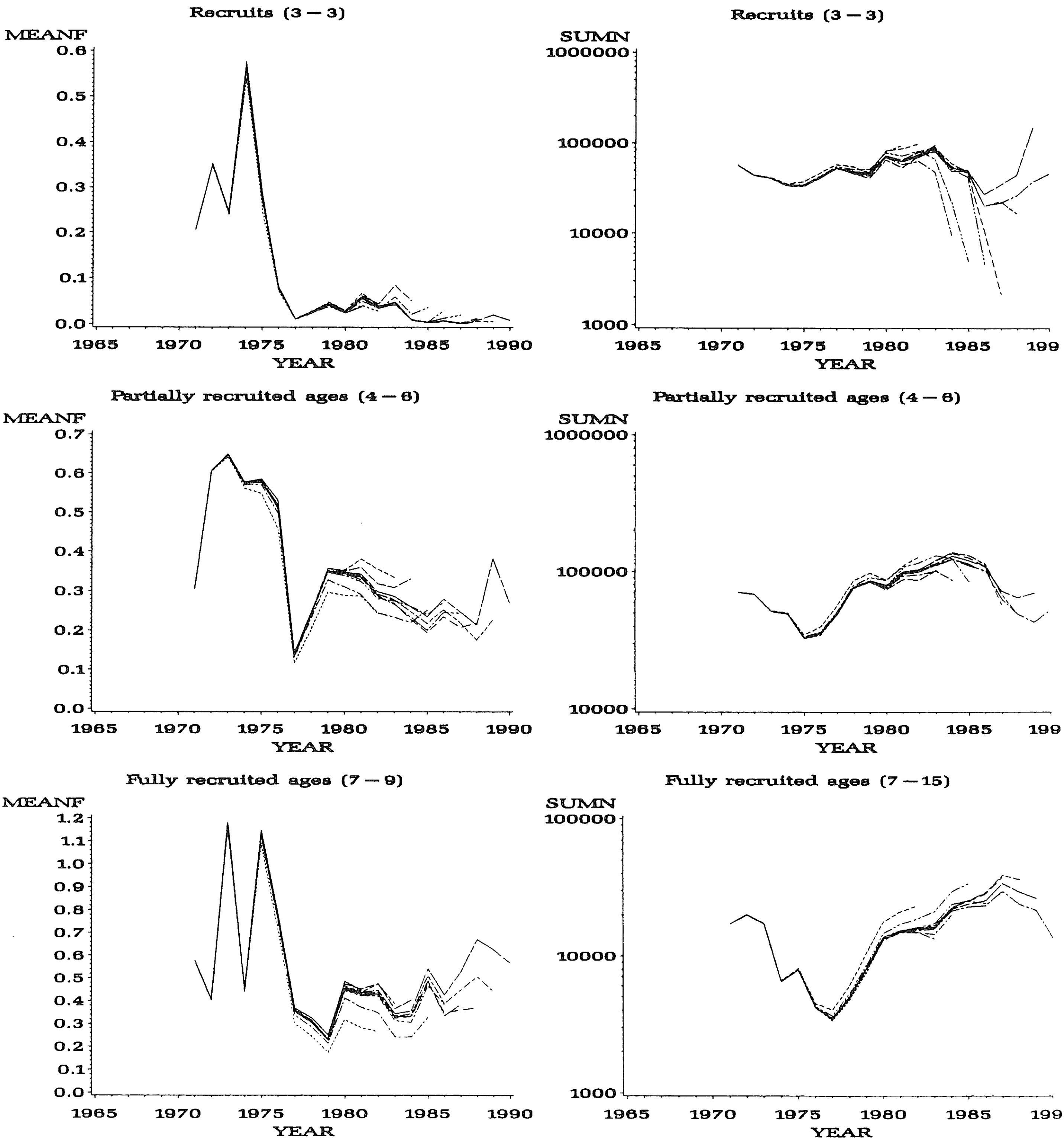


Figure 3.5.6

Stock: 4Vs W Cod

Procedure: XSA2 Shrinkage CV=0.4

Retrospective analysis

c4vxsa2.s4

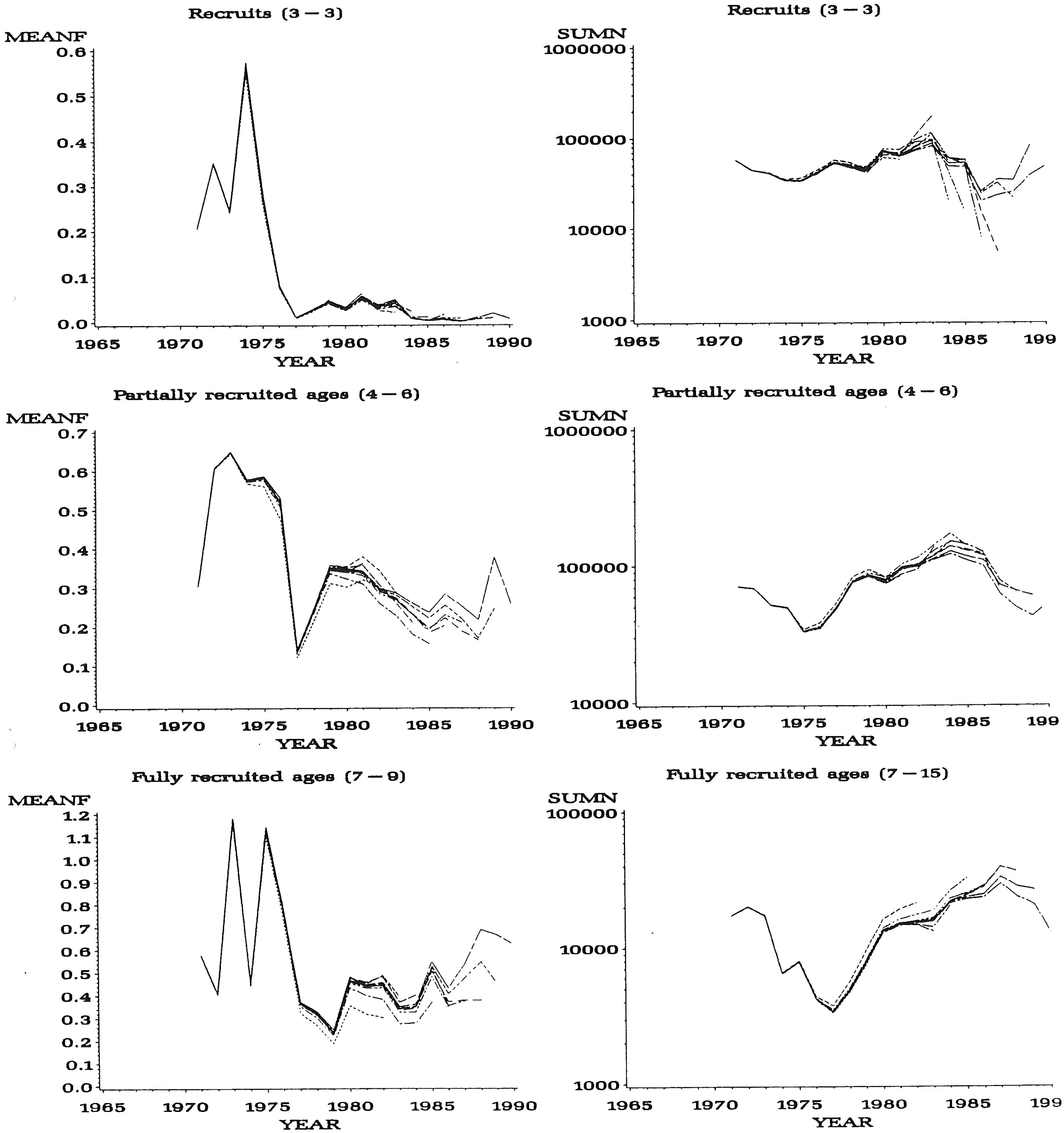


Figure 3.5.7

Stock: Reykjavik Simulation 5

Procedure: XSA2 Unshrunk

Retrospective analysis
sim5xsa2.s0

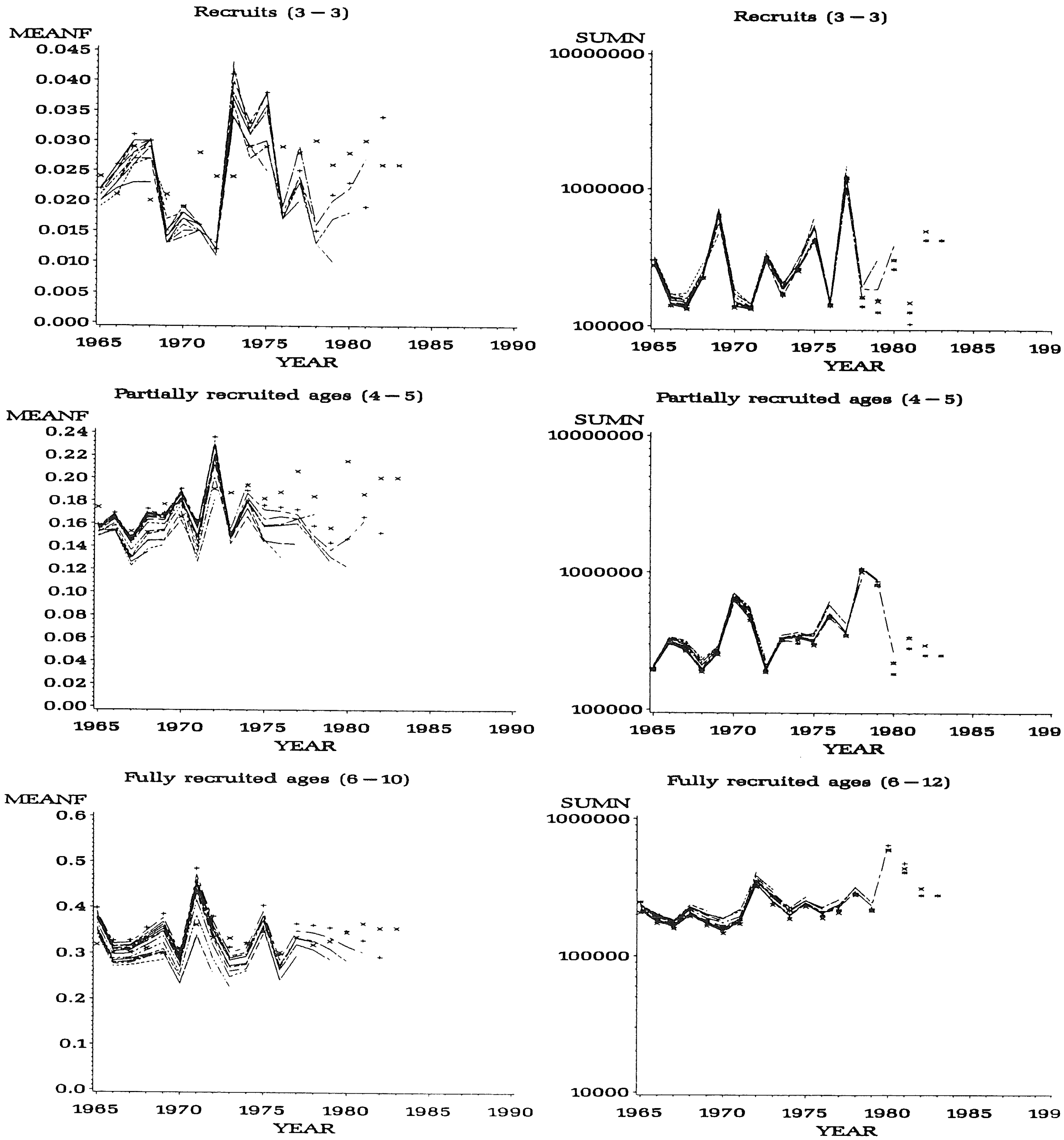


Figure 3.5.8

Stock: Reykjavik Simulation 5

Procedure: XSA2 Shrinkage CV=0.1

Retrospective analysis

sim5xsa2.s1

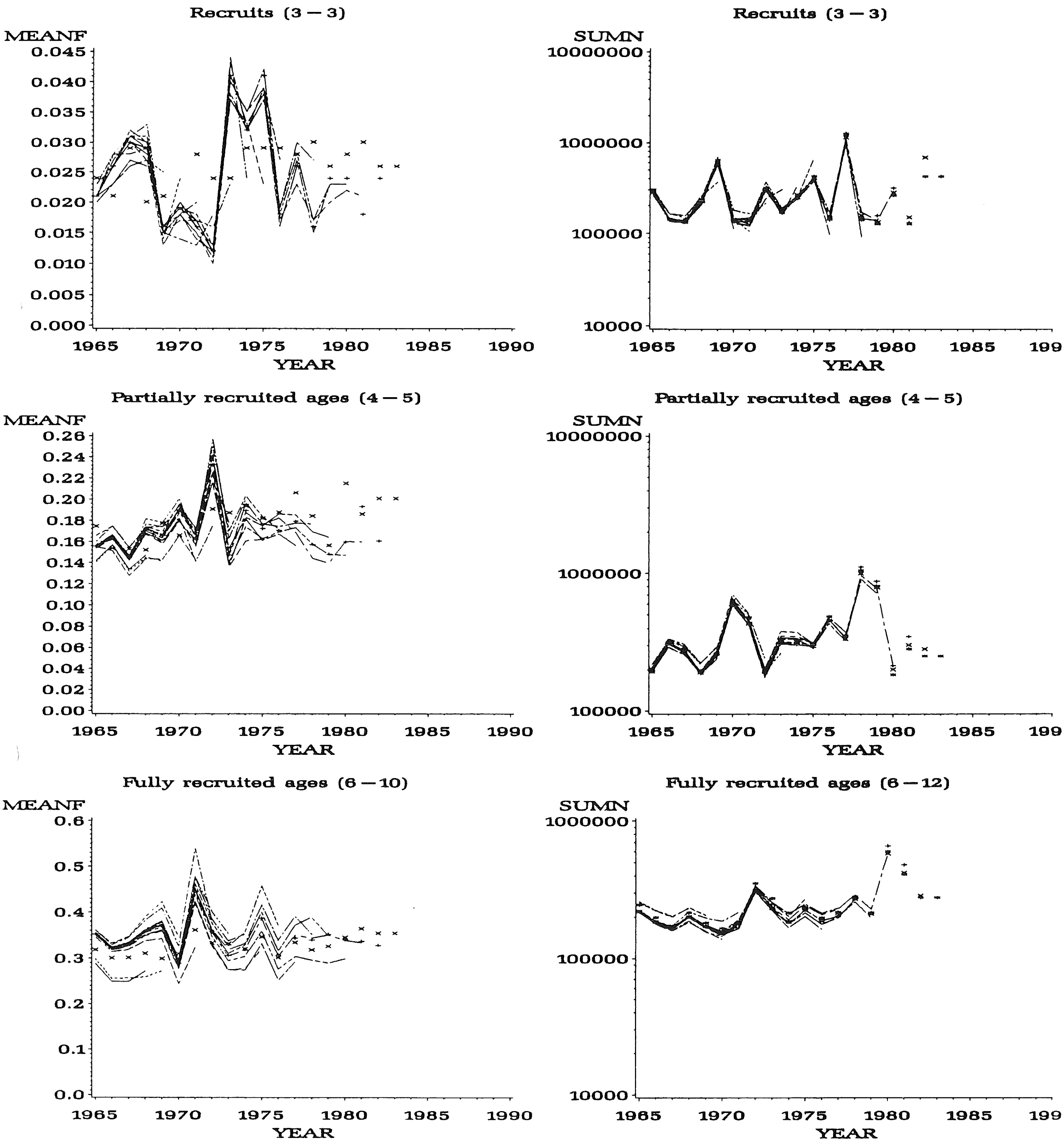


Figure 3.5.9

Stock: Reykjavik Simulation 5

Procedure: XSA2 Shrinkage CV=0.5

Retrospective analysis
sim5xsa2.s5

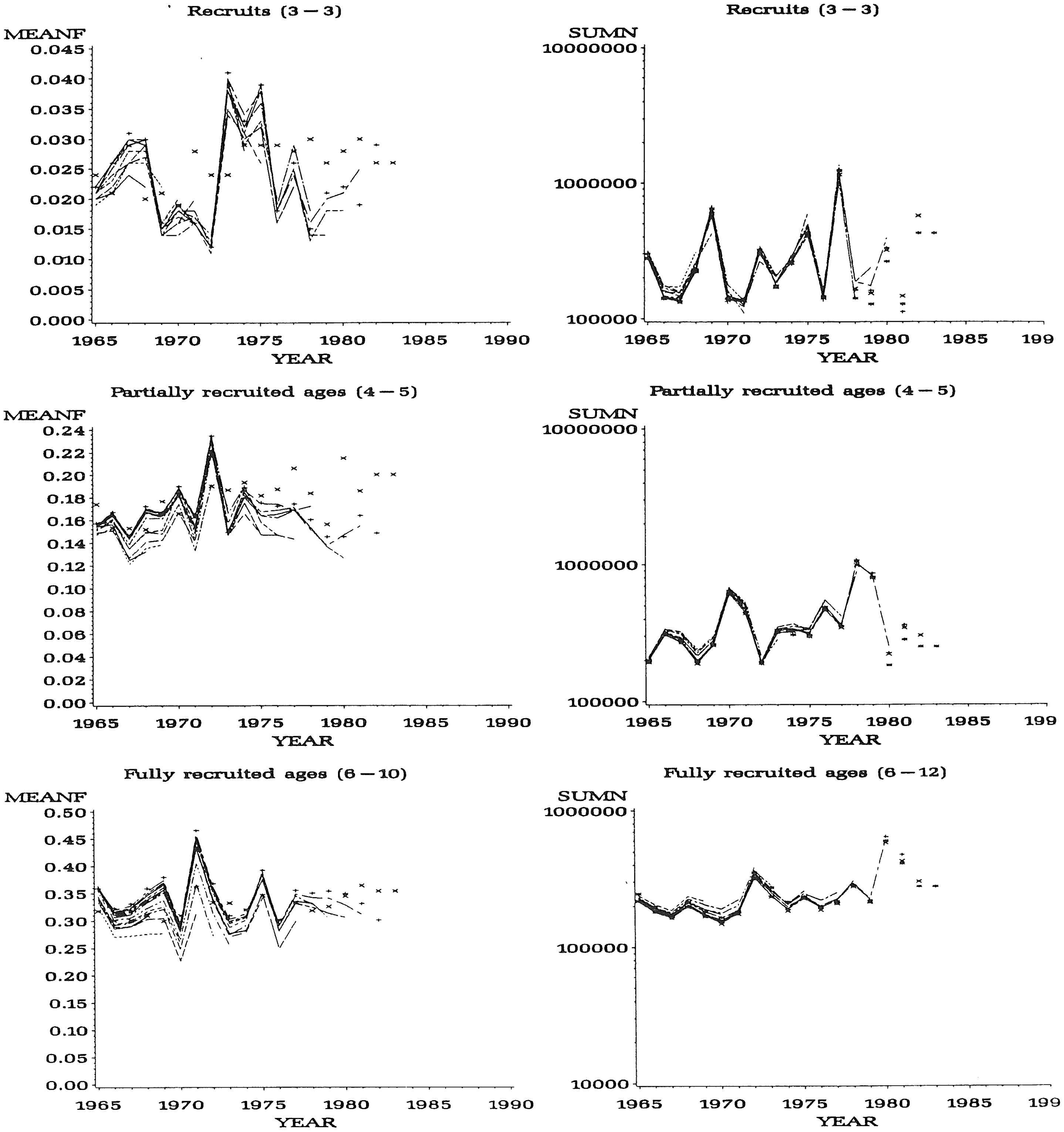
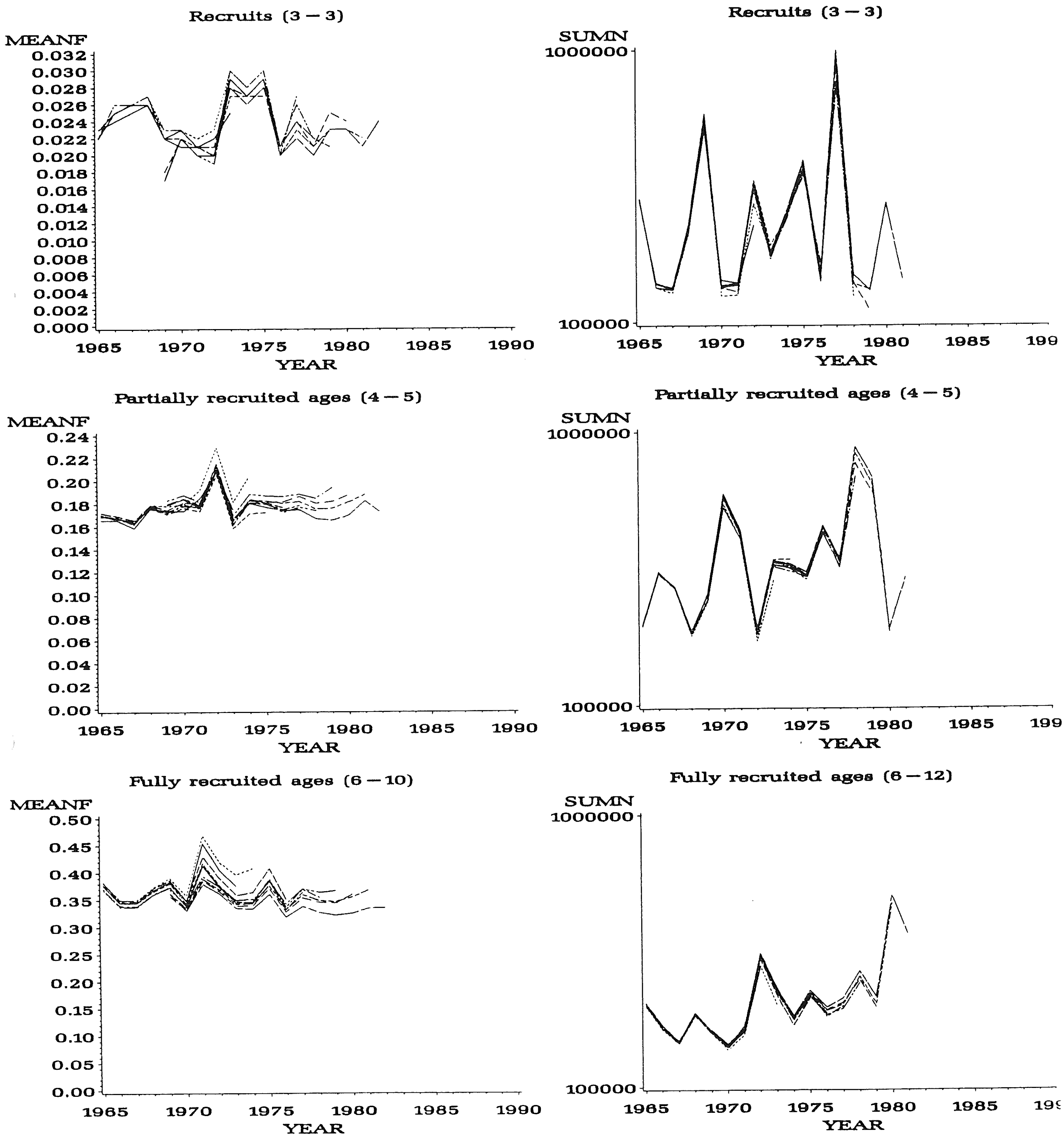


Figure 3.5.10

Retrospective analysis
simul.txt

Stock: Reykjavik Simulation 5

Procedure: TSA



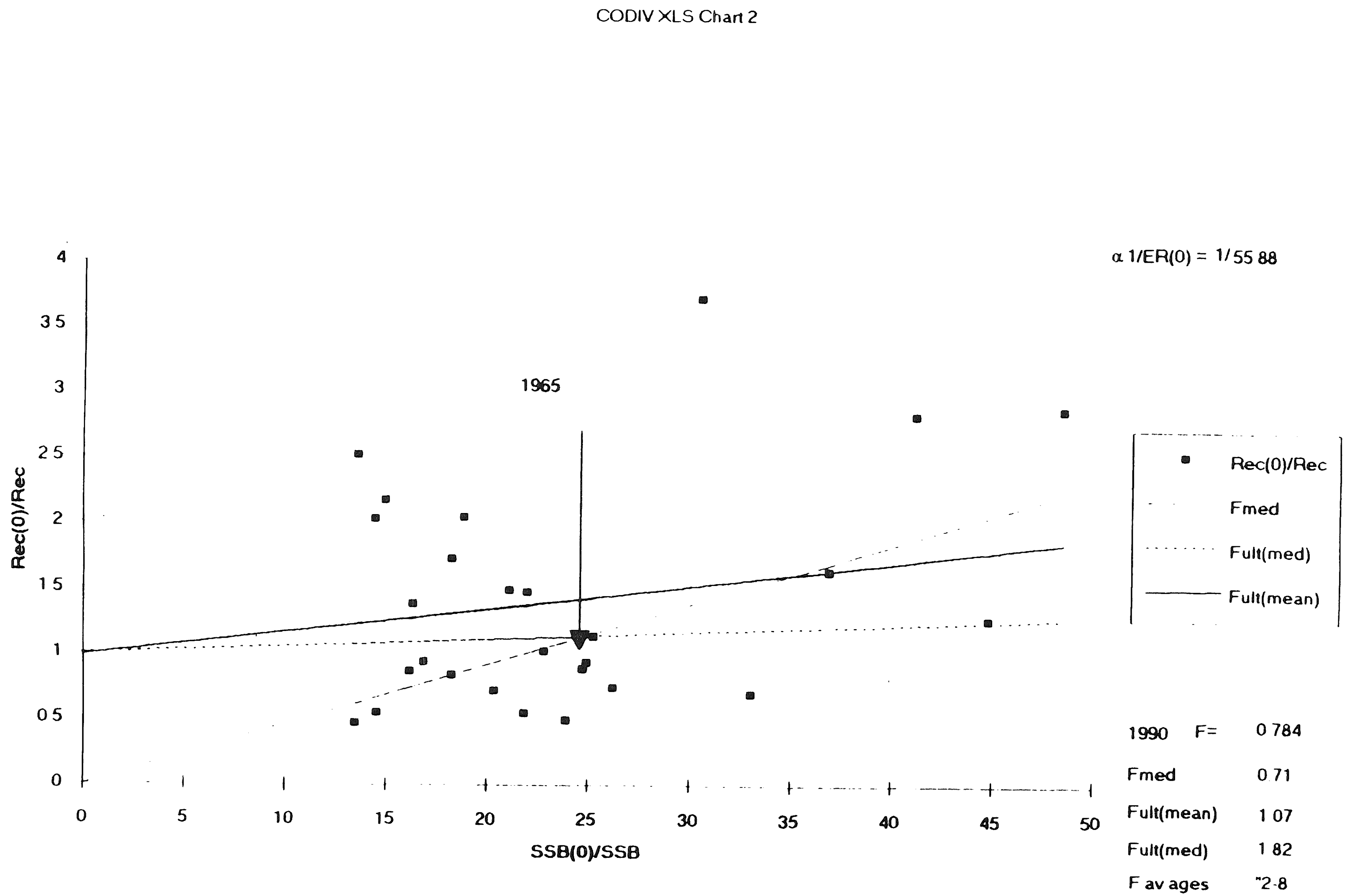


Figure 4.2.1 Example of a plot of the reciprocal of recruitment against the reciprocal of spawning stock biomass showing derivation of biological reference points (from Pope WP 6).

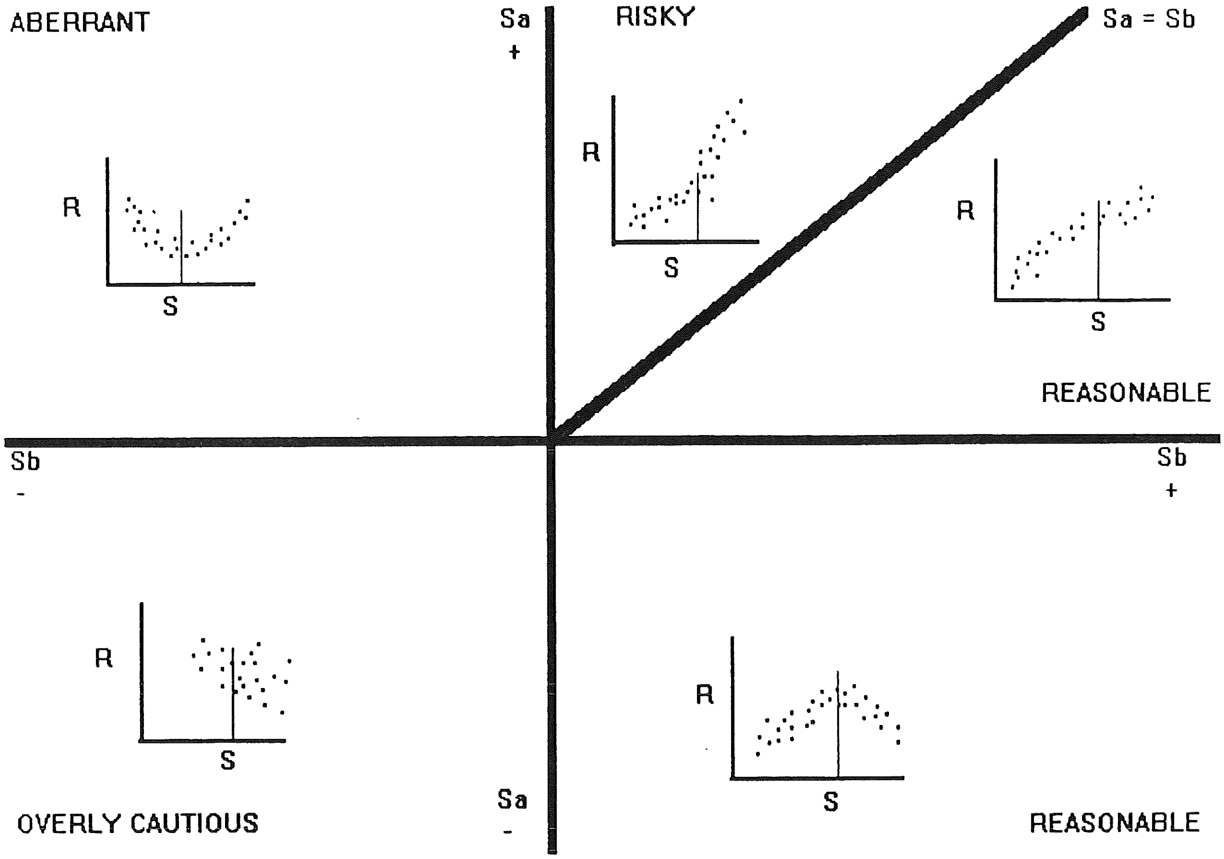
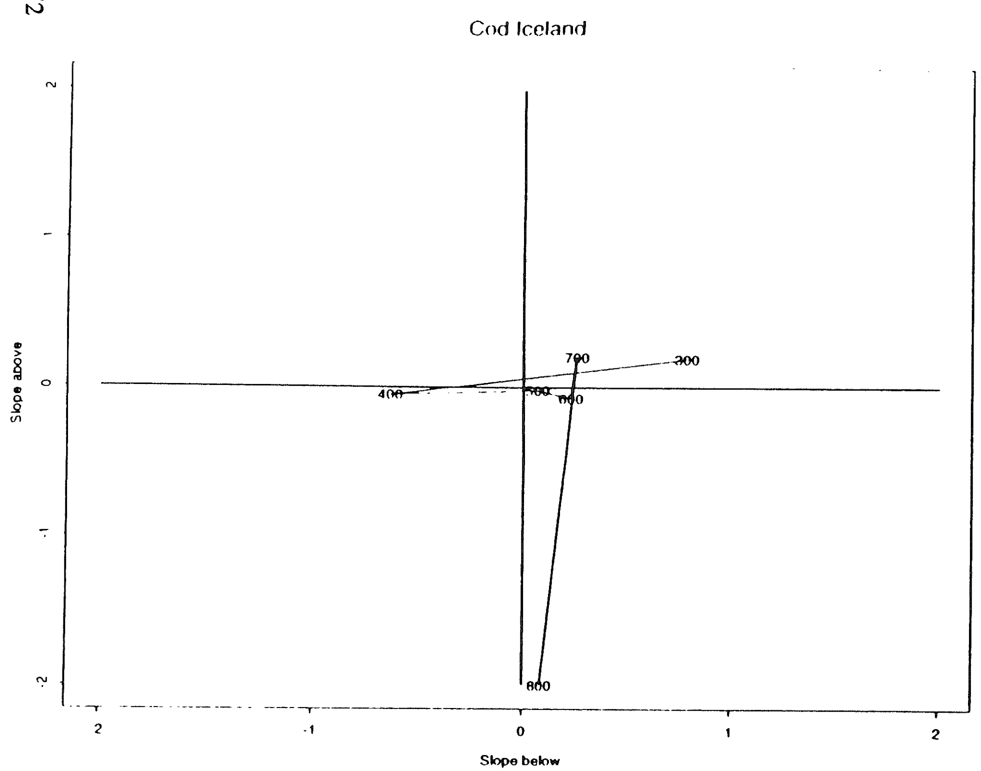
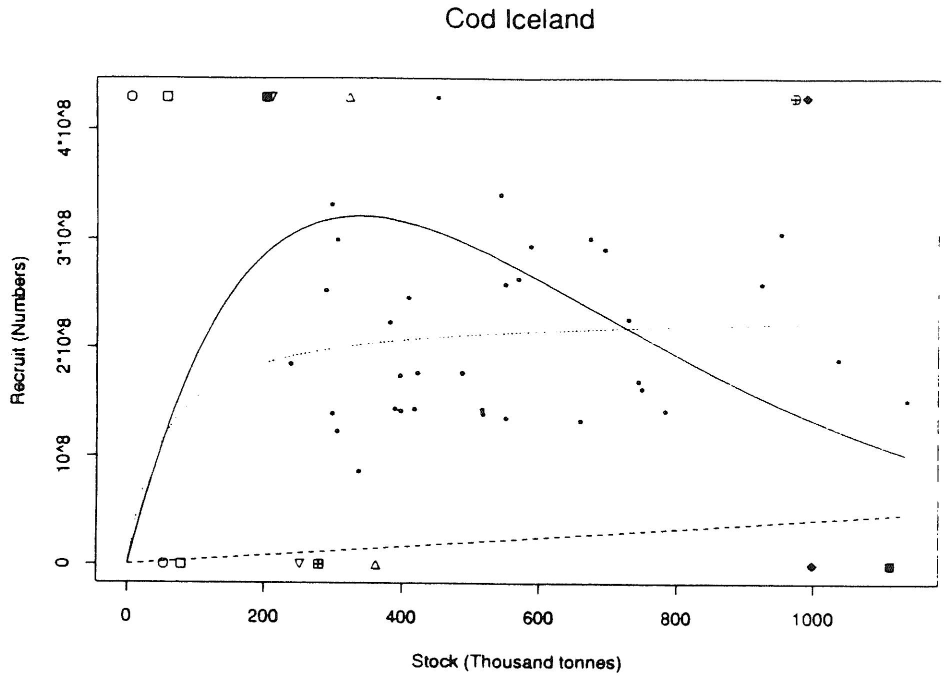


Figure 4.4.1 Decision diagram for slope above and below threshold biomass.





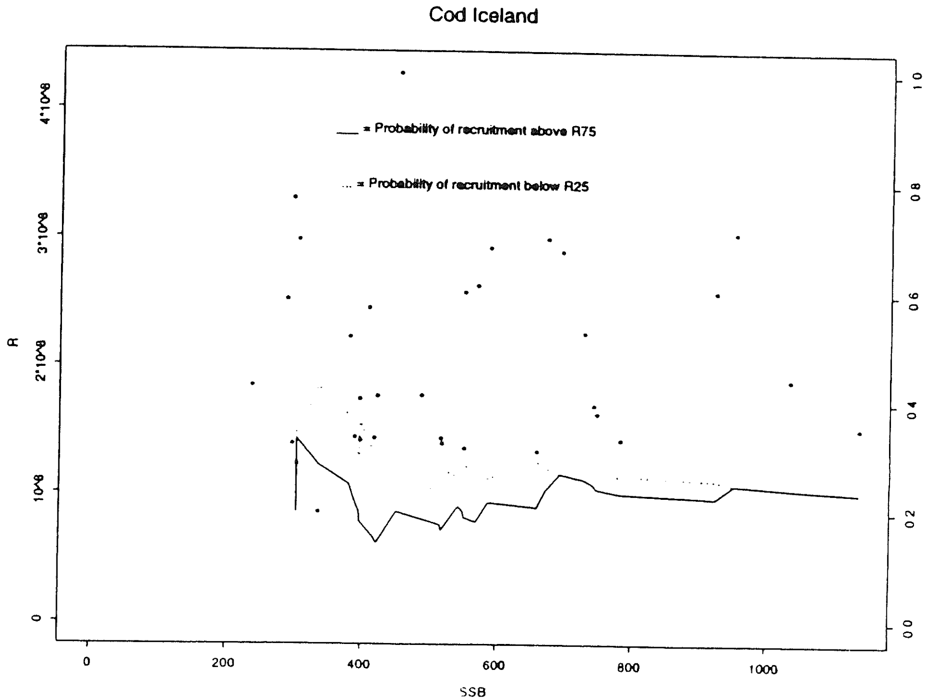
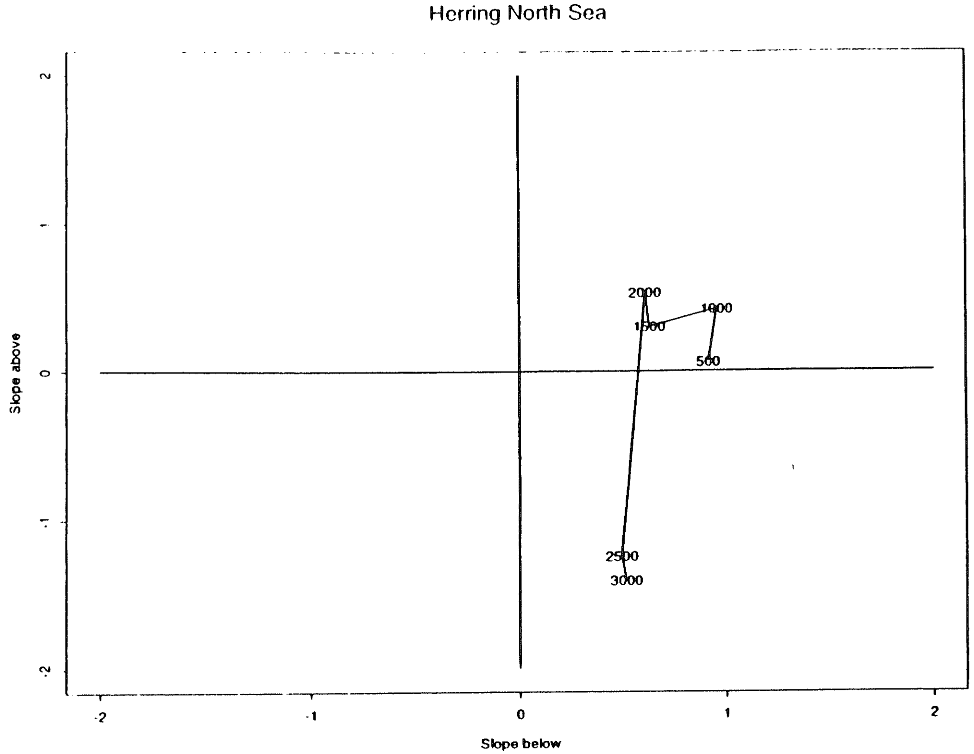


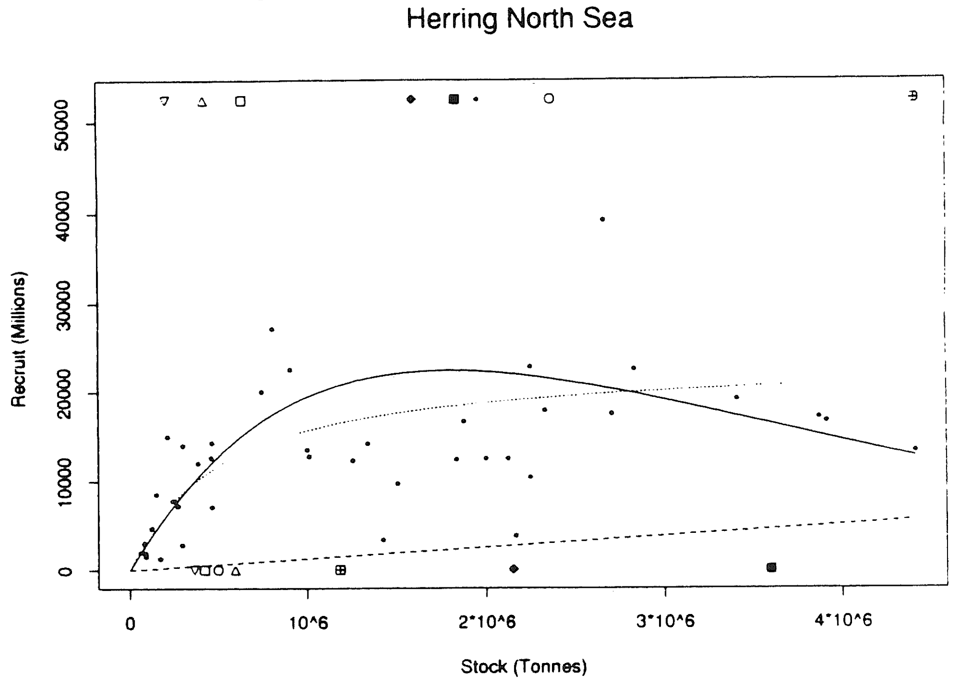
Figure 4.4.2 Icelandic cod - diagnostic sheet for estimating MBAL.

Virgin biomass from the Rmnv method	4988
Minimum observed SSB as proportion of virgin	0.0475
Maximum observed SSB as proportion of virgin	0.2281

Sbv:	361.77	ratio of mean above & below	1.054
BH50:	52.40	ratio of mean above & below	Cannot be estimated
RK50:	78.37	ratio of mean above & below	Cannot be estimated
Rmnv:	997.759	ratio of mean above & below	0.7993

Slope of replacement line: 41684.03





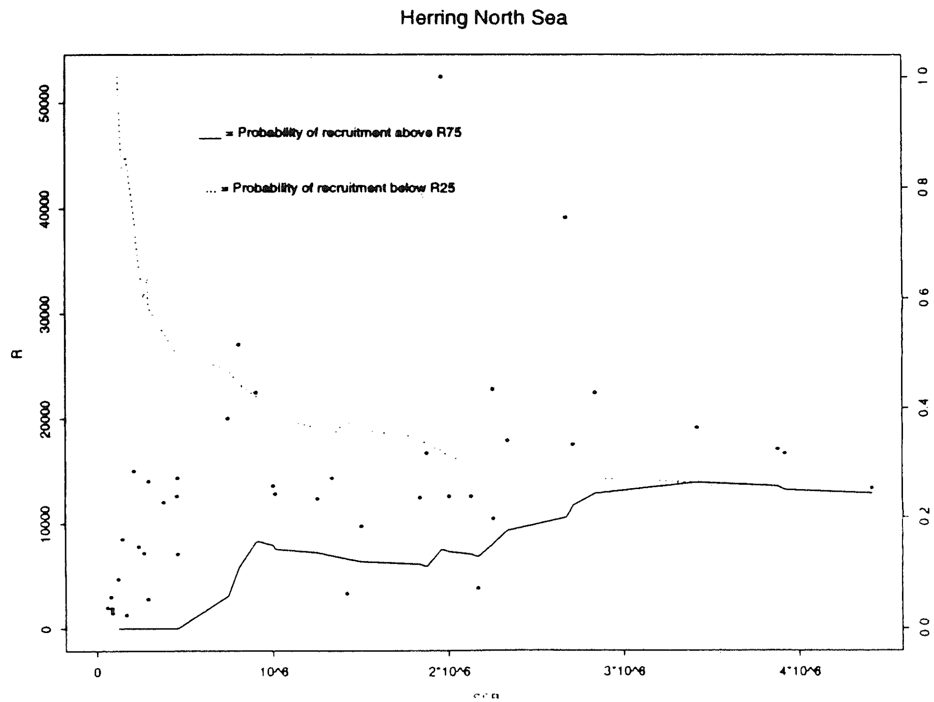


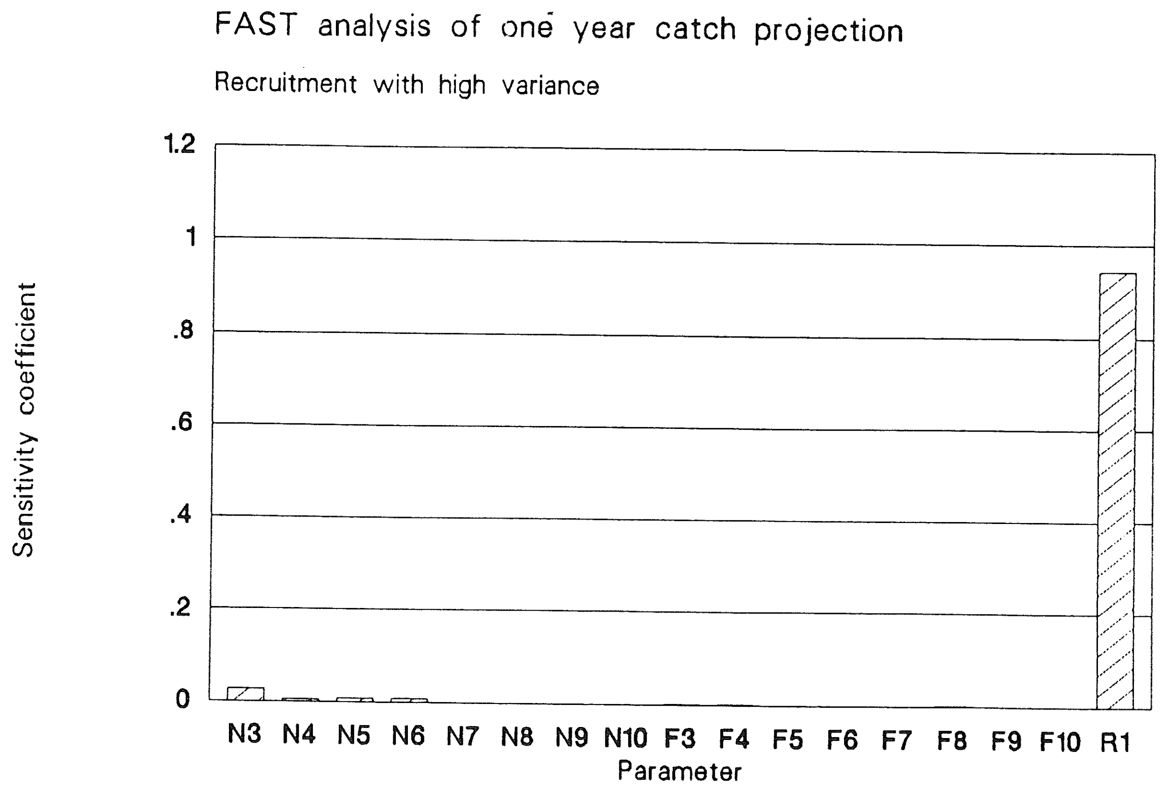
Figure 4.4.3 North Sea herring - diagnostic sheet for estimating MBAL.

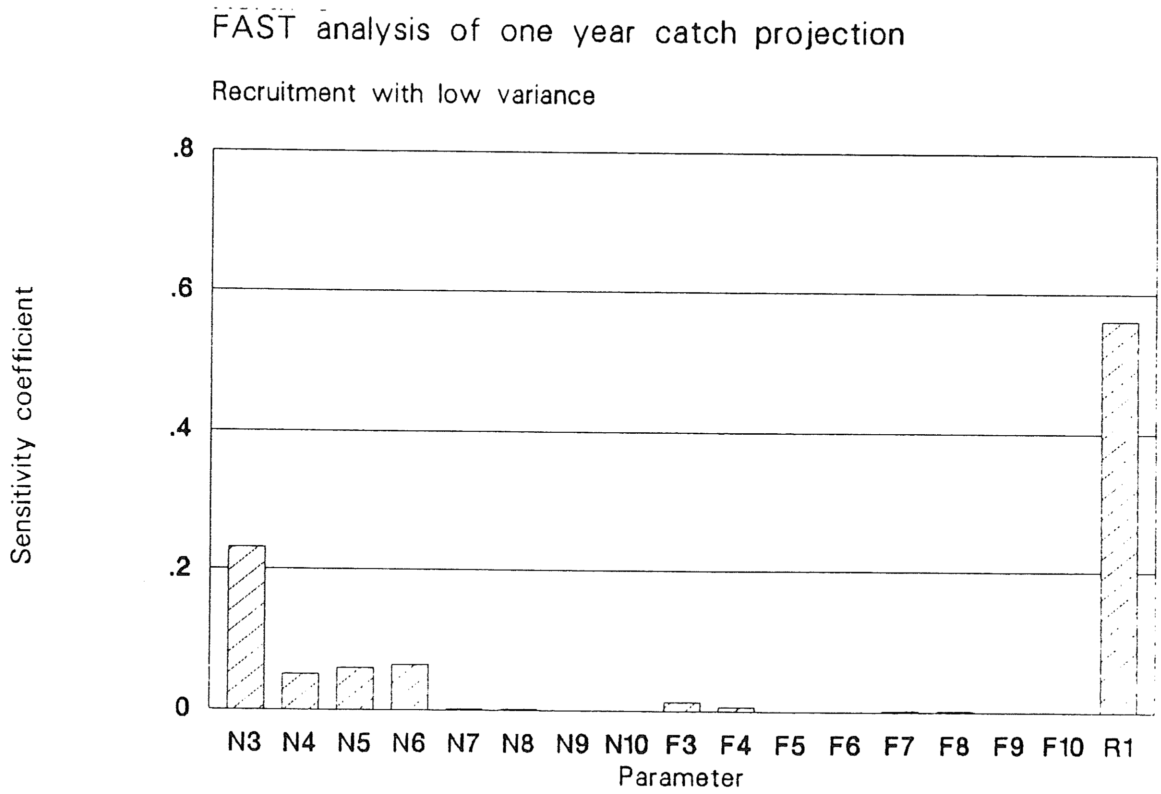
Virgin biomass from the Rmnv method	10751649
Minimum observed SSB as proportion of virgin	0.0053
Maximum observed SSB as proportion of virgin	0.4107

Sbv:	591787	ratio of mean above & below	2.44
BH50:	493716	ratio of mean above & below	2.44
RK50:	415862	ratio of mean above & below	2.70
Rmnv:	2150329	ratio of mean above & below	1.52

Slope of replacement line: 0.0012658

Figure 5.2.1 Output from sensitivity analysis. Ns refer to terminal populations indexed by age and Fs to terminal Fs at the same ages. R1 is recruitment at age 3 in 1992.





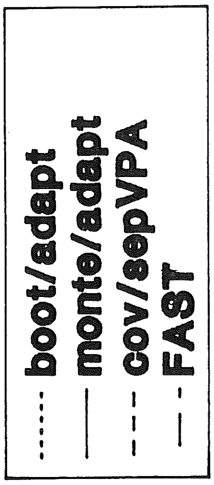


Figure 5.2.2 Probability profile for the simulated cod forecast example when recruitment variability is high in the recruiting year class. The profile is illustrative only.

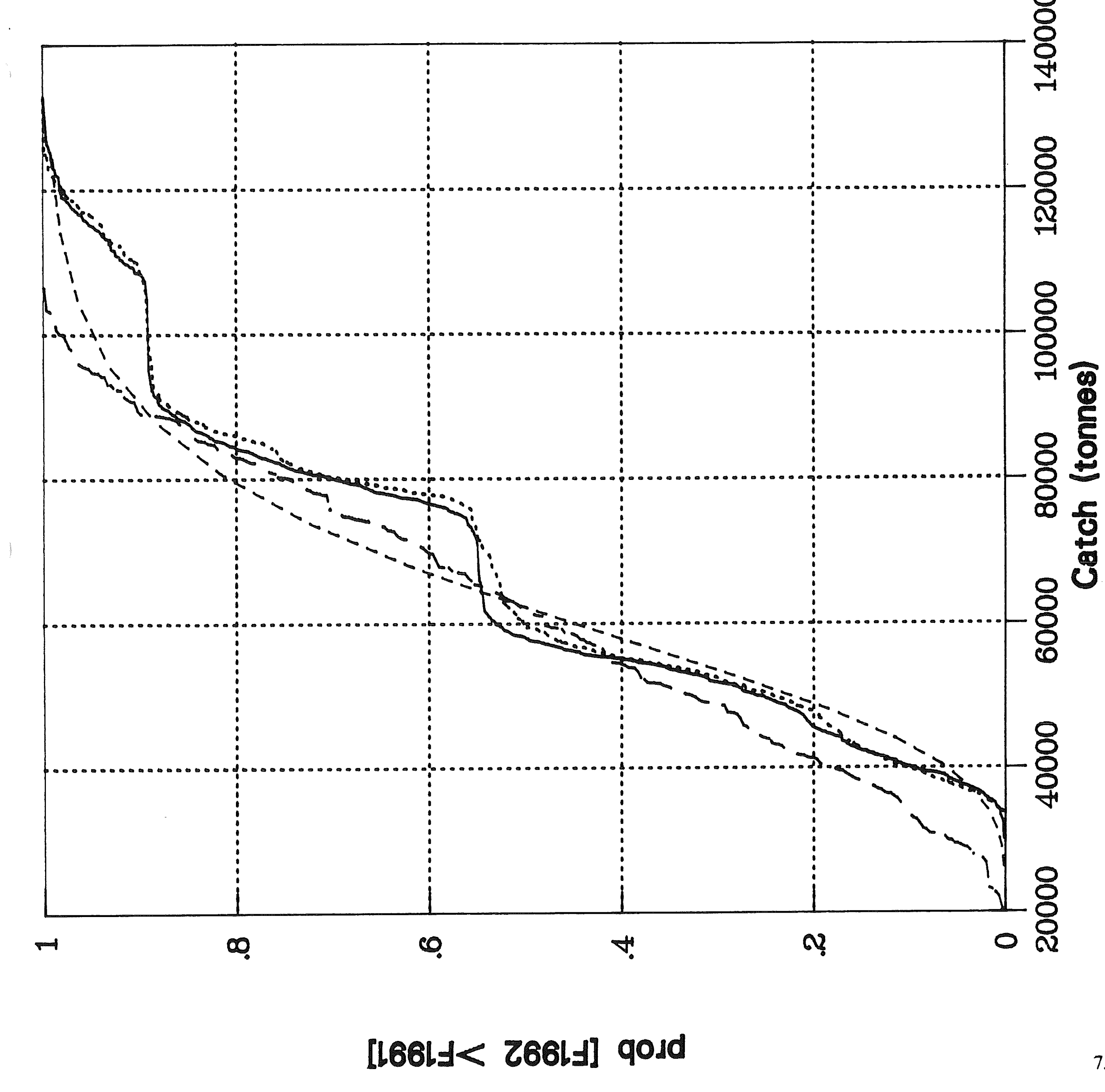
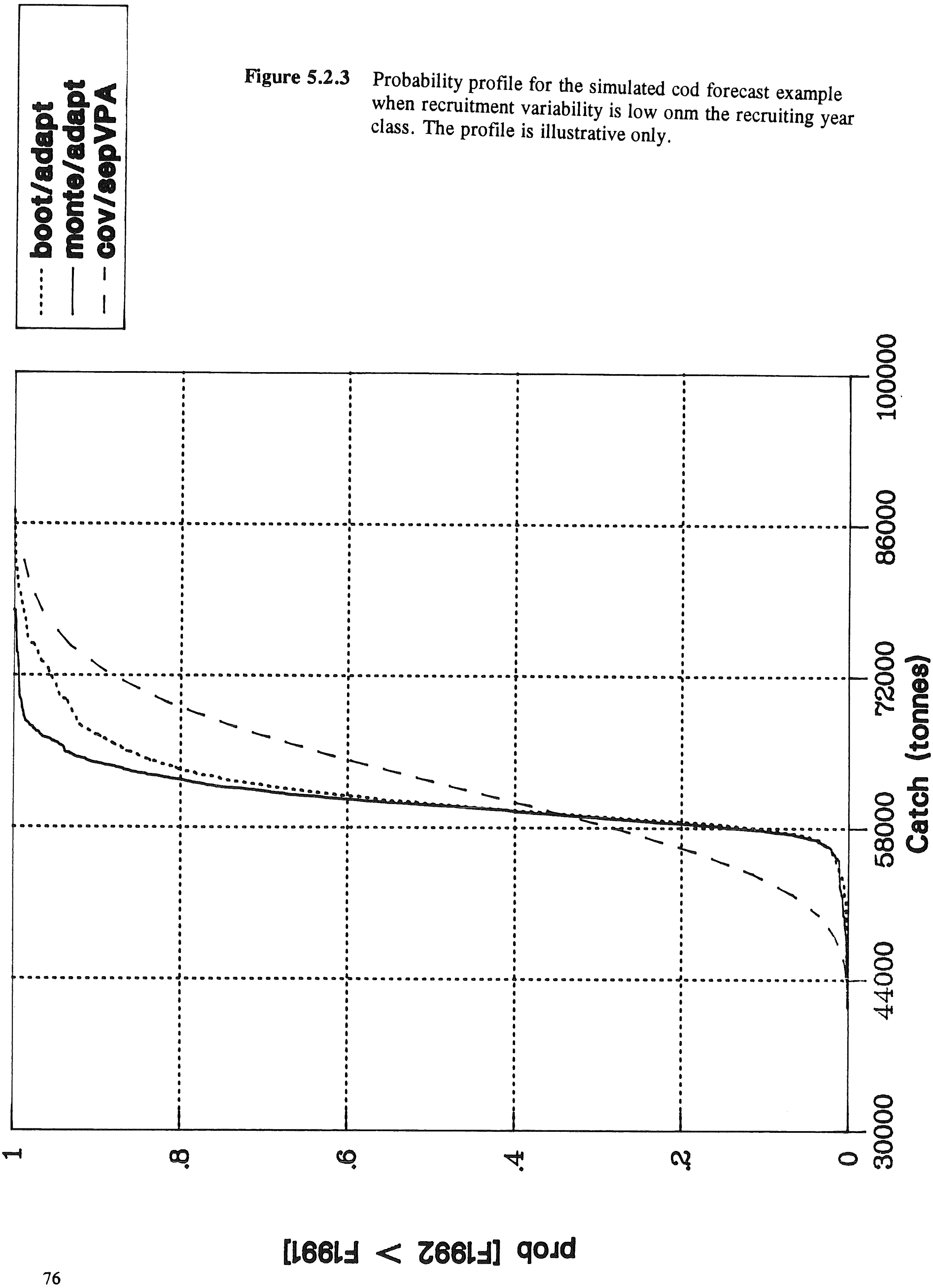


Figure 5.2.3 Probability profile for the simulated cod forecast example when recruitment variability is low onm the recruiting year class. The profile is illustrative only.



Bootstrap

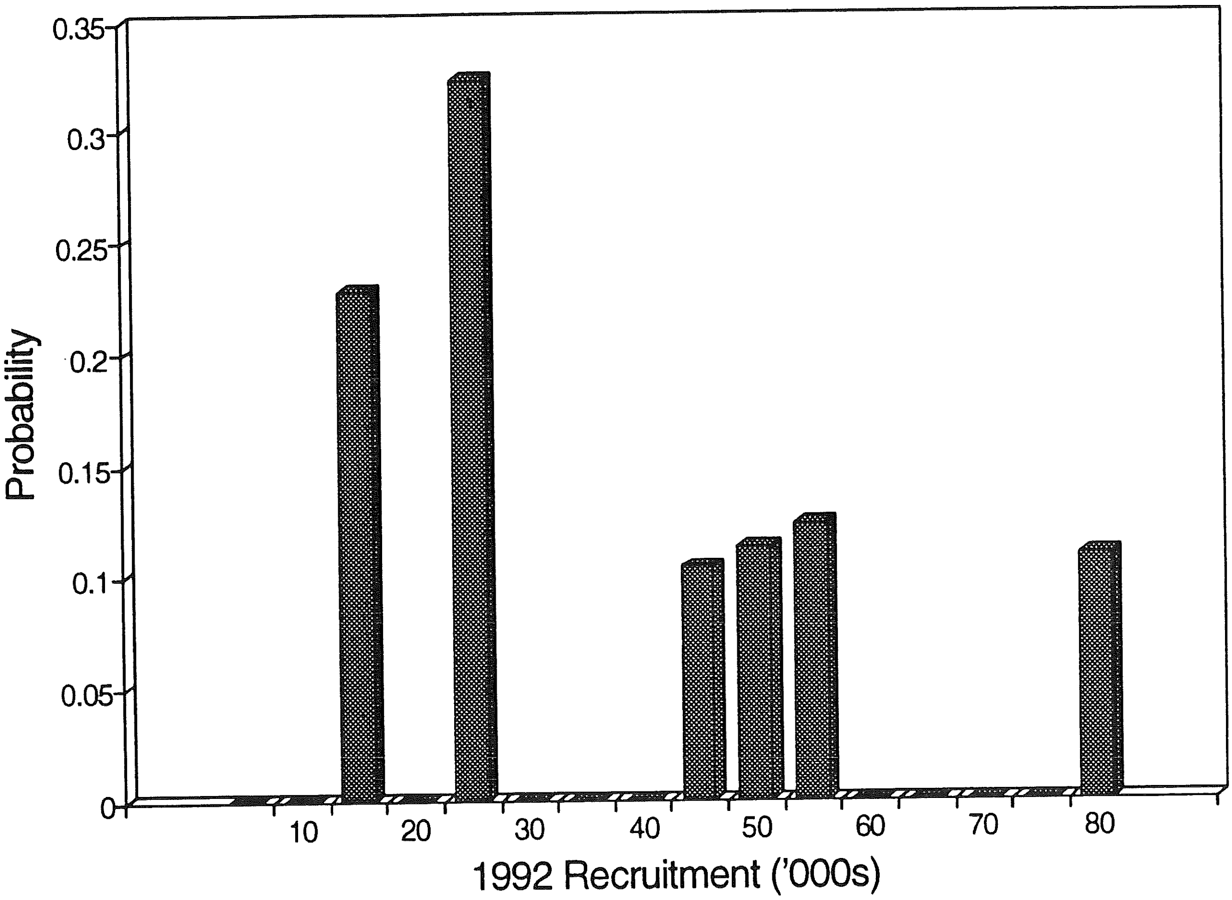


Figure 5.2.4 Observed distribution of recruitment estimates used in bootstrap technique.

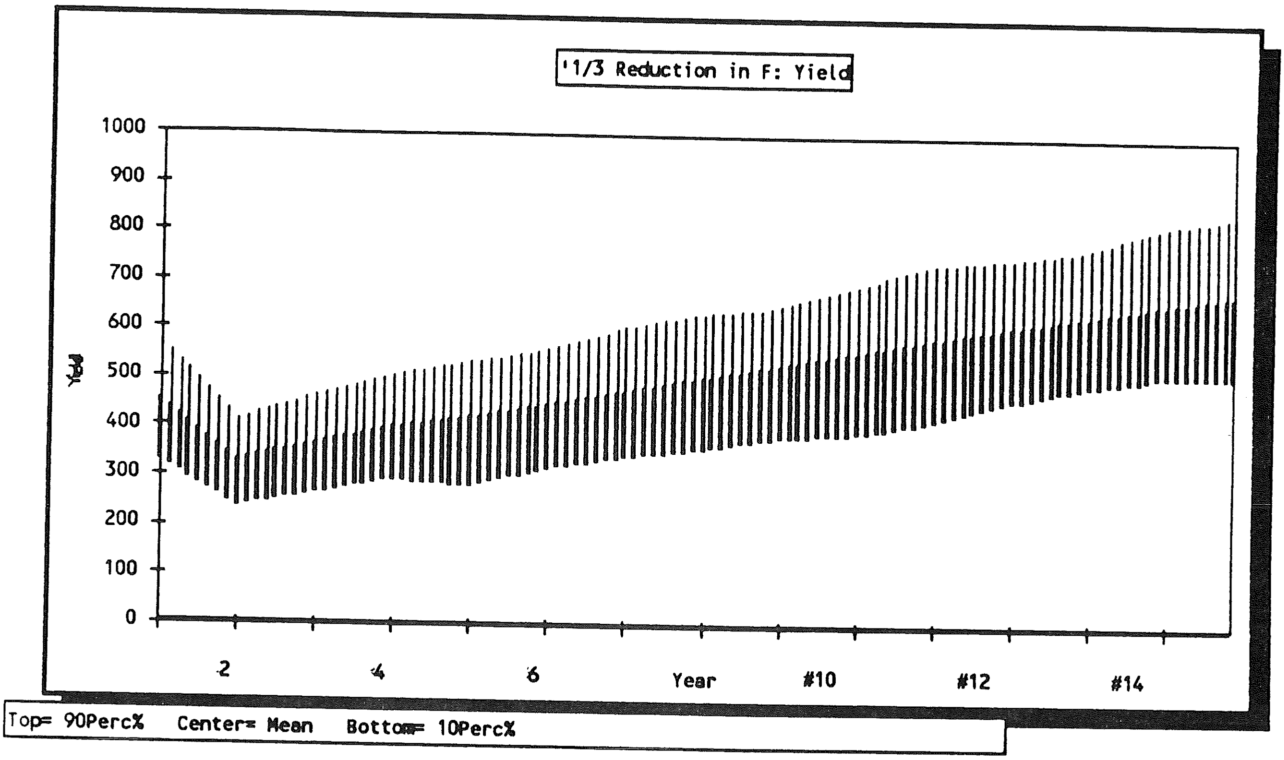


Figure 5.3.1 Medium-term (15 year) prediction of yield with upper and lower 90% confidence bars.

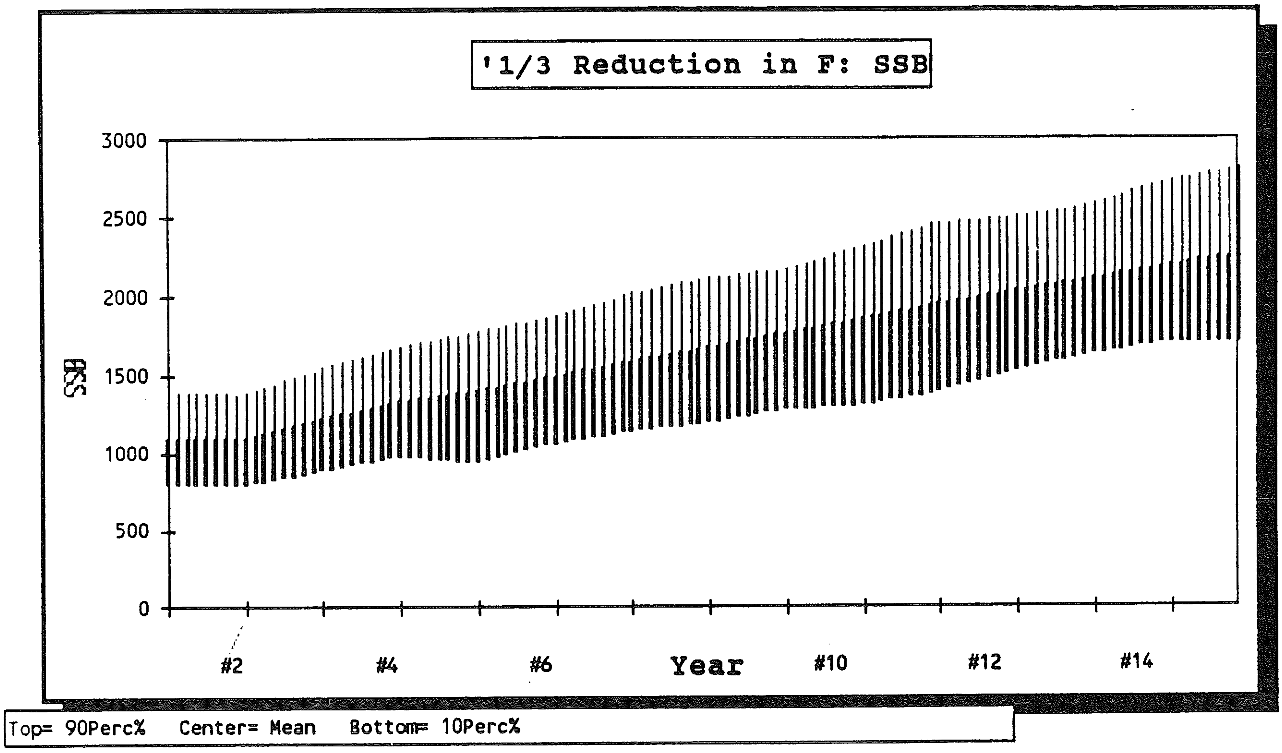


Figure 5.3.2 Medium-term (15 year) prediction of SSB with upper and lower 90% confidence bars.

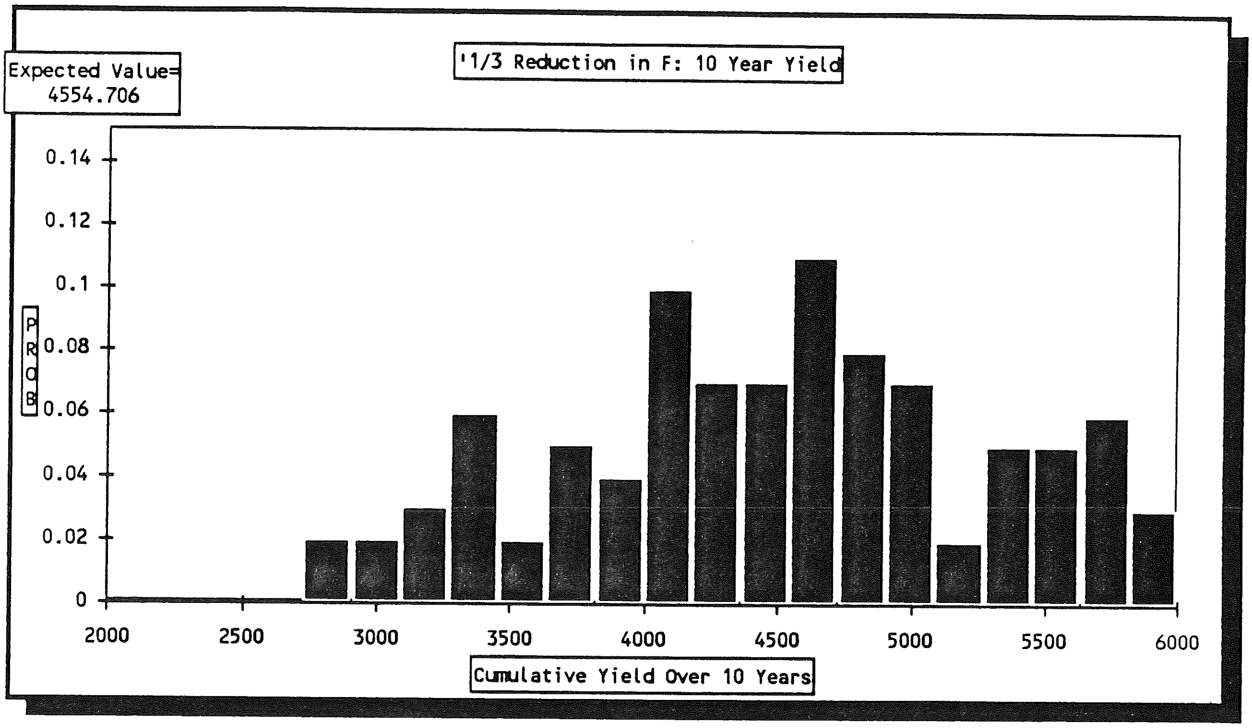


Figure 5.3.3 Histogram of simulated total yields over a 10 year period.

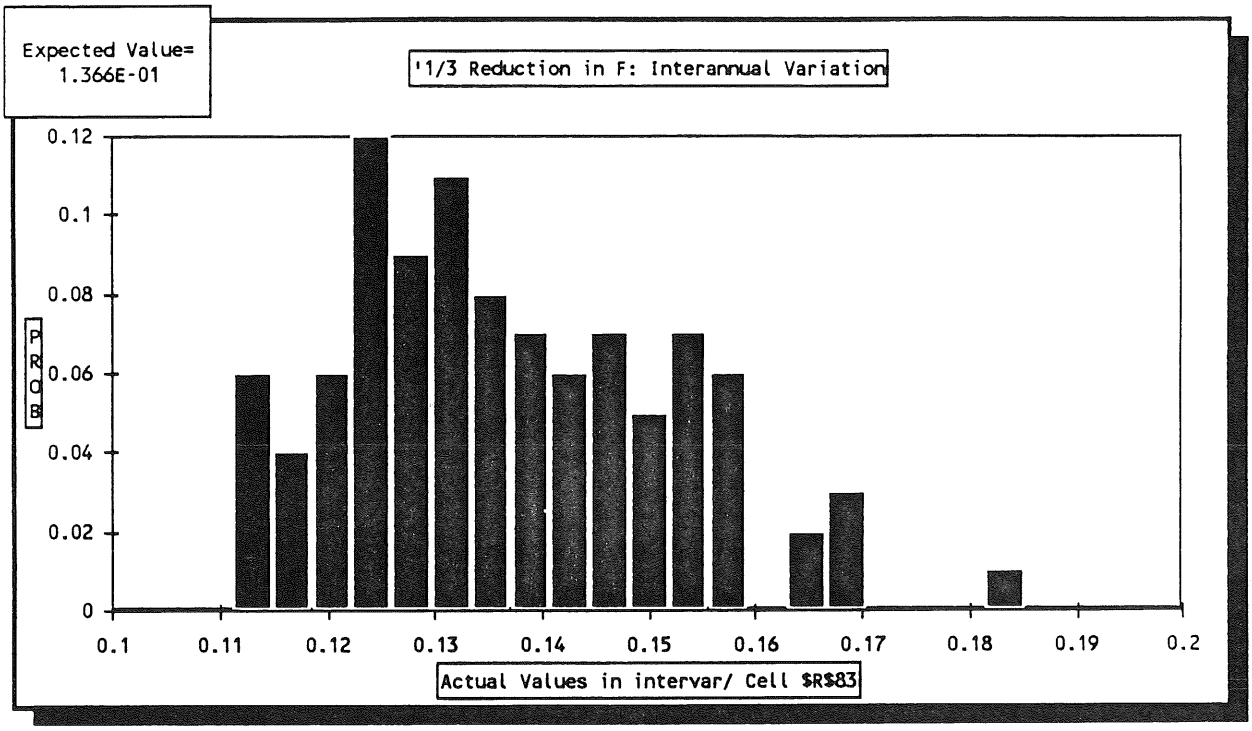


Figure 5.3.4 Histogram of simulated interannual variations.

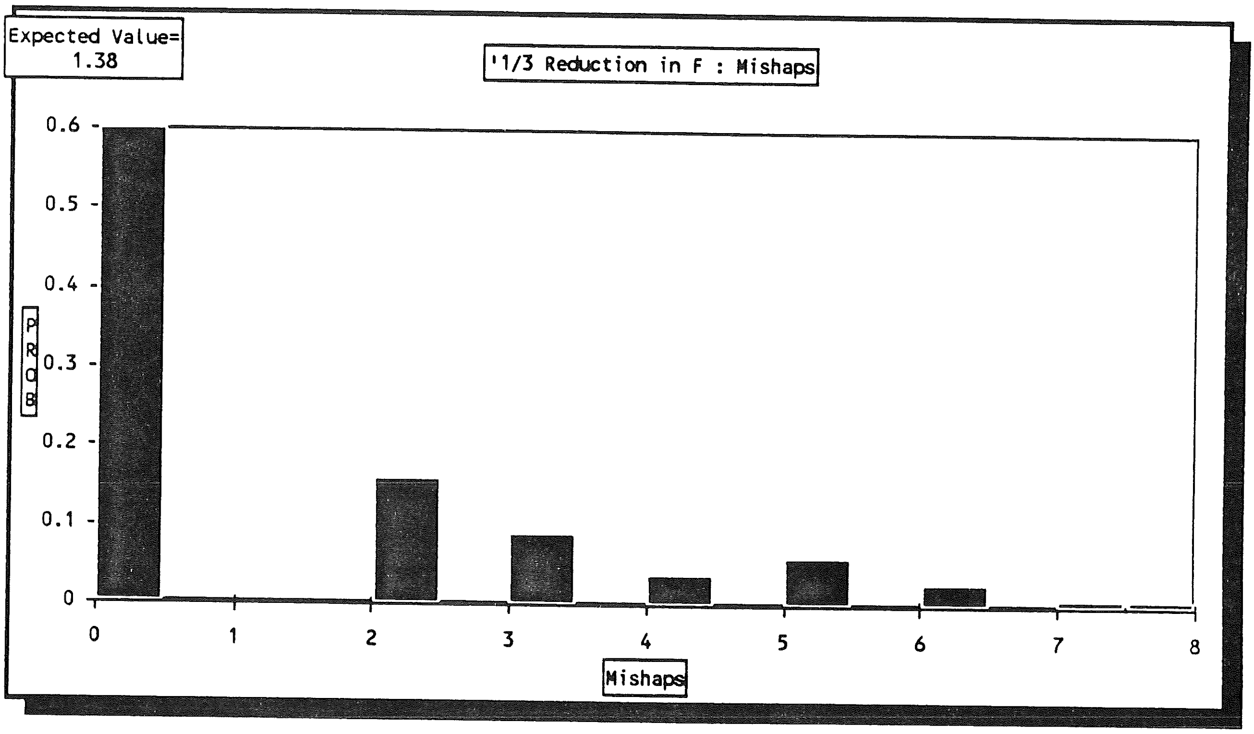


Figure 5.3.5 Probability distribution of the number of mishaps.

Appendix A: Notation

NOTE: This standard (and largely mnemonic) notation is followed sofar as possible, but not slavishly. Other usages and variations may be defined in the text. Array elements are denoted by means of either indices or suffices, whichever is more convenient. The same character may be used as both an index or a variable, if no confusion is likely.

Suffices and Indices

y	indicates	year
f	"	fleet
a	"	age
t	"	last (terminal) year
g	"	oldest (greatest) age group
l	"	length
k	"	year class
$	"	summation over all possible value of index (usually fleets)
#	"	summation over fleets having effort data
@	"	an average (usually over years)
*	"	a reference value

Quantities (all may have as many, and whatever, suffices are appropriate).

C(y,f,a)	Catch in numbers (including discards)
E(y,f)	Fishing effort
F(y,f,a)	Fishing mortality
F (y,f)	Separable estimate of overall fishing mortality
s	
q	Catchability coefficient (as in F=qE)
Y	Yield in weight
W	Weight of an individual fish in the catch
W	Weight of an individual fish in the (spawning) stock
S	
B	Biomass
P	Population number (also fishing power)
E	Fishing effort
U	Yield or landings per unit of effort
C	Catch in weight of fish (including discards)
W	
N	Stock in numbers of fish
F	Instantaneuos fishing mortality rate
M	Instantaneuos natural mortality rate
Z	Instantaneuos total mortality rate
S	Selection coefficient defined as the relative fishing mortality (over age)
R	Recruitment
f	Relative F (e.g., F/F*)
y	Relative yield (e.g., Y/Y*)
d	Fraction discarded
b	Fraction retained (b=1-d)
h	Hang-over factor

G	Instantaneous growth rate (in weight)
L	Landings in numbers (excludes discards)
l	Length
$l\infty$	Von Bertalanffy asymptotic length
K	Von Bertalanffy "growth rate"
r	Recruit index
MSY	Maximum sustainable yield
Fmsy	Fishing mortality associated with MSY
Emsy	Fishing effort associated with MSY
Bmax	Pristine stock biomass
m	Shape parameter for various surplus production models

APPENDIX B

SUMMARY OF REPORTS OF ICES WORKING GROUP ON THE METHODS OF FISH STOCK ASSESSMENT (AND ASSOCIATED MEETINGS)[1]

Summary of topics

Topic	1981	1983	1984	1985	1987	1988	1989	1991	1992	1993
1 Application of separable VPA	-	M	r	-	-	-	-	m	-	-
2 Simpler methods of assessment	-	-	M	M	i	-	-	i	-	m
3 Measures of overall fishing mortality	-	-	-	-	-	-	-	i	-	-
4 Use of CPUE effort and survey data in assessments	M	M	r	r	M	M	m	i	-	i
5 Need for two-sex assessment	-	-	-	-	-	-	-	-	-	-
6 Computation and use of yield per recruit	-	M	m	i	-	-	-	-	-	-
7 Inclusion of discards in assessments	-	-	-	-	-	-	-	-	-	-
8 Methods for estimation of recruitment	-	-	M	r	M	-	-	-	-	M
9 Density dependence growth, mortality, etc.)	-	-	-	-	-	-	-	-	-	-
10 Linear regression in assessments	-	-	M	-	m	-	-	-	-	-
11 Effect of age-dependent natural mortality	-	-	-	M	-	-	-	-	-	-
12 Stock-production models	-	-	-	-	M	-	-	-	-	-
13 Utilization of research survey data	-	-	-	-	M	M	m	i	M	m
14 Use of less reliable fishery statistics	-	-	-	-	m	-	i	i	-	i
15 Construction of survey and CPUE indices from disaggregated data	-	-	-	-	-	-	M	i	-	-
16 Implications of timing of WG meetings	-	-	-	-	-	-	m	-	-	-
17 Testing of age-balanced methods of analysis	-	-	-	-	-	M	m	M	-	-
18 Effects of management measures on CPUE	-	-	-	-	-	-	-	m	-	-
19 Evaluation and development of diagnostics	-	-	-	-	-	-	-	M	-	-
20 Application of length-based methods	-	-	-	-	-	-	-	m	-	-
21 Extension of time series of stock and recruitment	-	-	-	-	-	-	-	m	-	-
22 Problems with weight-at-age	-	-	-	-	-	-	-	-	-	-
23 Evaluation of uncertainty and risk	-	-	-	-	-	-	-	-	-	M
24 Shrinkage	-	-	-	-	-	-	-	-	-	M
25 Stock-recruitment relationships	-	-	-	-	-	-	-	-	-	M
26 Retrospective analysis	-	-	-	-	-	-	-	-	-	m
27 Minium Biologically Acceptable Levels (MBALs)	-	-	-	-	-	-	-	-	-	M

[1]See List of Meetings on page 86.

M: Major topic; m = minor topic; r = reprise;
i: incidentally considered

DATES, LOCATIONS AND REPORTS OF PREVIOUS MEETINGS OF THE ICES WORKING GROUP ON THE METHODS OF FISH STOCK ASSESSMENT (AND ASSOCIATED MEETINGS)

Date	Place	Report Title	Citation	
			C.M. Paper	Cooperative Research Report
1981	Copenhagen	ICES Ad Hoc WG on the Use of Effort Data in Assessments	1981/G:5	129 (1984), 1-66
1983	Copenhagen	ICES WG on Methods of Fish Stock Assessment	1983/Assess:17	129 (1984), 67-134
1984	Copenhagen	ICES WG on Methods of Fish Stock Assessments	1984/Assess:19	133 (1985), 1-56
1985	Copenhagen	ICES WG on Methods of Fish Stock Assessments	1986/Assess:10	157 (1988), 1-92
1987	Copenhagen	ICES WG on Methods of Fish Stock Assessments	1987/Assess:24	191 (1993), 1-77
1988	Reykjavik	ICES Workshop on Methods of Fish Stock Assessment	1988/Assess:26	191 (1993), 78-172
1989	Nantes	ICES WG on Methods of Fish Stock Assessments	1990/Assess:15	191 (1993), 173-249
1991	St. John's	ICES WG on Methods of Fish Stock Assessments	1991/Assess:25	
1992	Woods Hole	ICES Workshop on the Analysis of Trawl Survey Data	1992/D:6	
1993	Copenhagen	ICES WG on Methods of Fish Stock Assessment	1993/Assess:12	