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

International Council for the Exploration of the Sea

C.M.1987/Assess:24

REPORT OF THE WORKING GROUP ON METHODS OF FISH STOCK ASSESSMENTS

Copenhagen, 9 - 16 June 1987

This document is a report of a Working Group of the International Council for the Exploration of the Sea and does not necessarily represent the views of the Council. Therefore, it should not be quoted without consultation with the General Secretary.

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

TABLE OF CONTENTS

---oo0oo---

1 INTRODUCTION

1.1 Participants

D.W. Armstrong	UK (Scotland)
E. Aro	Finland
V.K. Babayan	USSR
F. Borges	Portugal
R. Chevalier	France
R.J. Conser	USA
W. Dekker	Netherlands
R. Deriso	USA
J. Efimov	USSR
M. Fogarty	USA
A. Fonteneau	France
G. Gudmundsson	Iceland
T. Jakobsen	Norway
P. Kanneworff	Denmark
A. Kristiansen	Faroe Islands
F. Laloe	France
A. Laurec (Chairman)	France
J.J. Maguire	Canada
B. Mesnil	France
R.K. Mohn	Canada
O.K. Palsson	Iceland
J.G. Pope	UK (England)
D. Rivard	Canada
J.G. Shepherd	UK (England)
G. Stefansson	Iceland

Dr E.D. Anderson, ICES Statistician, attended part of the meeting.

1.2 Terms of Reference

It was decided at the 74th Statutory Meeting (C.Res.1986/2:5:17) that the Working Group on Methods of Fish Stock Assessments (Chairman: Mr A. Laurec) will meet at ICES Headquarters from 9-16 June to consider:

a) the development and applicability of stock-production models;

b) the utilization of research survey data;

c) the development and testing of statistical models for the joint analysis of catch-at-age and CPUE and/or survey data;

d) the effect of reduced reliability of fishery statistics on stock assessments, and the implications for management advice.

1.3 Agenda

A total of 11 working papers are summarized in Appendix A. They offered the basis for a discussion that took place during the first two days.

Practical work then started on case studies corresponding to the various terms of reference. This work required the adaptation of a large number of computer programs, the main ones being listed in Appendix D.

2 SURPLUS PRODUCTION MODELS

2.1 Background

Surplus production models have long been used in the assessment of exploited fish populations. These models are mathematically tractable and have minimal data requirements. In their most basic form, only a time series of catch and effort information is required to estimate the parameters of these non-age-structured models. In addition, surplus production models implicitly incorporate consideration of recruitment dynamics and, therefore, potentially can be used to evaluate the risk of recruitment overfishing. On the other hand, the models may be too simple and the underlying assumptions too restrictive to accurately represent the dynamics of fish populations.

Surplus production models have not been widely used within the ICES area. In part, this reflects the availability of relatively long time series of data on the age structure of many fish populations within this region that can be used in more complicated and presumably realistic models. The Working Group undertook an analysis to evaluate the performance of several surplus production models using simulated and real data sets. The ability of the models to recover the essential dynamics of the simulated population was used as the principal criterion for success. For the actual data sets, comparisons were made among the various models for a number of population parameters.

The net production of a population is defined as the difference between increases in biomass due to recruitment and growth and losses due to natural and fishing mortality. For an unexploited population at equilibrium, recruitment and growth are balanced by natural mortality. Surplus production models are predicated on the assumption that the population is regulated by density-dependent factors. In theory, harvesting the population reduces intraspecific competition and increases population production levels. This "surplus" production can be harvested without resulting in a change in population biomass levels. Additional assumptions underlying traditional surplus production models (Schaefer, 1954, 1957; Pella and Tomlinson, 1969; Fox, 1971) include:

1) Age-structure effects are assumed to be unimportant. It is implicitly assumed that the age structure of the population has a negligible effect on the factors affecting the production rate.

2) The population is assumed to respond instantaneously to changes in density. Time delays in production processes are not considered in the traditional forms of surplus production models, and the progeny are assumed to age instantaneously to the adult population.

3) The population is assumed to be closed or, alternatively, that immigration and emigration rates exactly balance. The population is assumed to be homogeneously distributed within the area. Extension of fishing areas to new or adjacent areas is not considered.

4) We assume that the catchability rate is constant and that fishing effort has been standardized to be proportional to instantaneous fishing mortality.

5) The fishing pattern has to remain constant. Changes in size limit regulations or gear regulations (e.g., mesh size) may violate this assumption.

Clearly, these assumptions are too simplistic to accurately reflect the dynamics of real populations. Surplus production models must be considered to be a crude representation of actual stock dynamics. Nevertheless, the models do embody the essential elements of the principal hypotheses regarding fish population regulation. Further, the traditional models can be modified to remove some restrictive and unrealistic features such as the assumption of no time delays, constant catchability, and spatially homogeneous populations (Fox, 1974; Freon, 1983). Laloe (WP 2) demonstrated a production model which considered environmental effects. Recent production models proposed by Deriso (1980) and generalized by Schnute (1985) embody a "collapsed" age structure comprising recruits and post-recruits. These models also treat the individual elements of production (growth, recruitment, and mortality) explicitly and more realistically than the traditional models.

The Working Group evaluated a sequence of increasingly detailed production models ranging from the simple traditional models of Schaefer and Pella and Tomlinson to the delay difference models of Deriso/Schnute and recent modifications due to Shepherd (WP 6). In addition, for the traditional models, the Working Group considered several approaches to parameter estimation ranging from simple methods which assume equilibrium conditions to more complicated methods which consider the non-equilibrium (transient) trajectory of the population (Rivard and Bledsoe 1978).

The principal distinction among the various models considered was the degree to which the individual components of production are treated in aggregated form. We refer to the traditional models of Schaefer, Pella and Tomlinson, and Fox as aggregated or "lumped" models. These models do not distinguish among recruitment, growth, and natural mortality. Further, the parameters of these models cannot be related to specific biological processes or mechanisms of population regulation. Accordingly, the parameters cannot generally be estimated using auxillary information based on biological studies. This point is important because it appears that the models are somewhat underdetermined when only catch and effort data are used for estimation. The delay-differential models proposed by Walter (1973) and expanded by Marchesseault <u>et al</u>. (1976) and Fogarty and Murawski (1986) attempt to treat recruitment separately from growth and natural mortality; however, the functional forms used to represent recruitment processes are simplistic. Finally, the delay-difference models of Deriso (1980) and Schnute (1985) treat each of the components of production individually. Further, these models are expressed in terms of parameters with specific biological interpretations which can, in principal, be estimated independently of catch and effort data. Auxillary information can, therefore, be used for estimation.

2.2 Theoretical Framework

The dynamics of an exploited species may be expressed as:

$$dB/dt = [R(B) + G(B) - M(B) - F(B) + n]B \qquad (2.2.1)$$

where R(B), G(B), M(B), and F(B) are per capita rate functions of recruitment, individual growth, natural mortality, and fishing mortality and n represents a random disturbance (Schaefer and Beverton, 1963). The traditional surplus production models of Schaefer (1954, 1957), Pella and Tomlinson (1969), and Fox (1971) treat recruitment, growth, and natural mortality in aggregate using a compensatory population function. The model then takes the simple form:

$$dB/dt = [\varphi(B) - F + n]B \qquad (2.2.2)$$

where $\varphi(B)$ is the compensatory function [e.g., logistic (Schaefer 1954, 1957), Richards (Pella-Tomlinson, 1969), or Gompertz (Fox, 1971) functions]. In practice, the stochastic differential equation model is often replaced by the corresponding deterministic form. The rate of change of yield is given by:

$$dY/dt = F\ B \qquad (2.2.3)$$

At equilibrium, for the deterministic model, we have:

$$\varphi(B) = F\ B \qquad (2.2.4)$$

which can readily be solved to find the maximum sustainable yield (sometimes referred to as the maximum equilibrium yield) and the level of fishing mortality or fishing effort at which yield is maximized.

The non-equilibrium or transient yield can also be studied directly. The short-run yield is given by:

$$Y(t) = \int F(t)B(t)dt \qquad (2.2.5)$$

Often, biomass estimates will not be directly available. In this case, catch per unit effort (CPUE) is assumed to be directly proportional to biomass. By definition, F = qE where q is the constant of proportionality between the instantaneous fishing mortality (the catchability coefficient) and standardized fishing effort (E). Therefore we have:

$$Y(t)/E(t) = qB(t) \qquad (2.2.6)$$

where Y(t)/E(t) is the catch per unit effort. The assumption of strict proportionality between F and E can be relaxed (e.g., Hilborn, 1979), although only at the expense of additional parameters and more complex fitting procedures.

It is implicitly assumed in the traditional surplus production models that there are no time delays between spawning and recruitment. Clearly, this cannot hold in general. Walter (1973) proposed a modification of the Schaefer and Fox models which explicitly considered time delays. This model may be expressed in general form as:

$$dB/dt = \{f[B(t)] + g[B(t-r)] - F\}B \qquad (2.2.7)$$

where $g[B(t-r)]$ is a function representing the effect of spawning biomass on recruitment. This assumes that there is no significant error in taking production to be defined by exploitable rather than spawning biomass. Closed-form solutions are not generally possible for the time-delay production model. Approximate solutions are possible, however. Marchesseault et al. (1976) and Fogarty and Murawski (1986) give applications of other time-delay models of this general form.

Deriso (1980) introduced an alternative approach in which each of the individual elements of production are treated separately. The general form of the model is:

$$B(t+1) = (1+g)s(t)B(t) + s(t)s(t-1)g[B(t-1)] + h[B(t+1-r)] \qquad (2.2.8)$$

where g is the Brody growth coefficient [exp(-K)], s is the survival fraction, and $h[B(t+1-r)]$ is the stock-recruitment function. The advantage of this formulation relative to traditional surplus production models is that the model is expressed in terms of parameters which can be estimated independently from CPUE or biomass data. For example, the Brody growth coefficient may be estimated independently from age and growth studies and included in the model as a fixed parameter. Alternatively, Bayesian methods can be used if prior estimates of some parameters and their variances are available. This general model formulation also allows specification of a more realistic recruitment function; traditional formulations implicitly include recruitment but in somewhat implausible functional form. One difficulty with this general approach is that it is somewhat difficult to obtain reasonable estimates for all of the parameters from catch and effort or biomass data alone. Fogarty and Murawski (1986) proposed a simplified model in which the growth and natural mortality terms were not separable without additional information. Shepherd (WP 6) provided results for a model in which natural mortality was specified in advance and growth and recruitment were treated in aggregate. The Shepherd model is based on the relationship:

$$B(t+1) = B(t) + P(t) - Y(t) \qquad (2.2.9)$$

where $P(t)$ is the net production to the exploited stock and all other terms are defined as before. The production-to-biomass ratio (P/B) is assumed to follow:

$$P/B = a/(1 + B/K) - M \qquad (2.2.10)$$

where a is the maximum rate of biomass increase, K is the biomass level at which density-dependent effects predominate, and M is the natural mortality rate. Natural mortality is assumed to be known. Further, Shepherd (WP 6) proposes that the parameter a, which is a measure of resilience, be estimated qualitatively based on known or inferred characteristics of the stock.

2.3 Case Studies

2.3.1 Generation of simulated data for production model comparison

An age-structured surplus production program was modified to produce data for the comparison of production models. The modifications were the inclusion of a stock-recruit relationship and the option for adding either measurement or process noise. The standard program requires the specification of weight at age, natural and fishing mortalities, and selectivity. The stock-recruit modification requires fecundity at age (FEC) to generate potential recruitment (PREC):

$$PREC = \Sigma N(a)\ FEC(a)$$

The potential recruitment is deduced by a Shepherd-style density-dependent expression. The fecundity coefficients above are analogous to Shepherd's parameter a. The critical density and shape parameters (k and g) are unchanged from his formulation:

$$REC = PREC/[1 + (B/k)^{g}]$$

Equilibrium values were obtained by finding stable age distributions over a range of fishing mortalities and then iteratively scaling the populations until recruitment was in equilibrium. The equilibrium yield versus fishing mortality and stock-recruitment curves are shown in Figures 2.3.1 and 2.3.2. The method of determining equilibrium yield is similar to Shepherd's (1982) method of combining yield-per-recruit and stock-recruitment relationships, except that the effective spawning biomass is not the same as the density-dependent biomass and both are functions of the age structure. A slightly domed stock-recruitment function was chosen which corresponds to an MSY of approximately 1,500 at a biomass of 5,500. The recruitment is 908 at MSY and the fishing mortality is just over 0.5.

After the parameters had been determined, a 20-year projection was run with the fishing effort increasing for ten years and then more slowly decreasing for ten years (see Table 2.3.1). Two more projections were carried out, the first with the addition of measurement noise and the second with process noise. In either case, the noise was log normal with a log standard deviation of 0.2. Measurement noise was added to numbers and catch at age, as well as effort, after the simulation. It was not added to weight at age. Process noise was added to fishing mortality, fecundity, and the density-dependency parameters. (It should have been added to natural mortality and weight at age, but was not.) The results of the simulations with measurement and process are in Table 2.3.2.

The simulated data sets had a larger dynamic range (F ranged from 0.3 to 1.25 in 20 years) and lower noise levels than are commonly seen in fisheries data. This means that the methods tested would have a relatively easy task compared to the real data situation and were not severely tested by the simulated data.

2.3.2 Estimation methods

The Working Group considered several methods of fitting traditional surplus production models using both equilibrium and non-equilibrium approaches. The Group employed a simple predictive regression of catch per unit effort on effort as the first method because this technique has been widely applied in fitting surplus production models. This method is problematical due to confounding of the dependent and independent variables and because the transient behavior of the system is not considered. The second method used the equilibrium approximation method suggested by Gulland (1961) based on averaging effort over k/2 years, where k is the number of significant year classes in the fishery. The third method employed the numerical integration method of Rivard and Bledsoe (1978) which directly takes into account the non-equilibrium (transient) stock dynamics. The Group also used the method of Schnute (1977) based on time-averaged regressors. This technique is also a non-equilibrium method. The final two methods were applied to models in which the individual components of production are treated in greater detail.

In the report, these four methods are referred to as: (1) equilibrium, (2) equilibrium approximation, (3) transitional, and (4) time average, respectively.

For methods in which an estimate of the catchability coefficient is produced, several additional population parameters were estimated in addition to the maximum sustainable yield (MSY) and effort level at MSY (E_{msy}). These were biomass at MSY (B_{msy}), the maximum production to biomass ratio (P/B), maximum biomass (B_{max}), current biomass (B_t), and current fishing mortality (F_t). It was possible to estimate these parameters only for the transitional method of Rivard and Bledsoe (1978) and the method of Shepherd (WP 6).

The Group considered the Deriso (1980) model as generalized by Schnute (1985). This method allows two estimation procedures: 1) a non-linear estimation procedure assuming process error only and 2) a simulation approach which assumes that the input data are subject to measurement error. The Working Group also applied the method of Shepherd (WP 6) as implemented in a computer algorithm provided for this meeting. This method fixes some parameters to reduce the estimation problem. A mapping of the sums of squares surface is used as a diagnostic tool in estimating the parameters. The Shepherd model was fitted for some stocks using two different functional forms for the recruitment-growth sub-model: 1) Beverton-Holt type and 2) Schaefer type.

2.3.3 Results for traditional production models

Results of the test runs on simulated data were particularly instructive. Comparisons among the various estimation methods for simulated data are given in Tables 2.3.3-2.3.8, and plots of the raw data and fitted equilibrium curves are provided in Figures 2.3.3-2.3.8. It should be noted that the transitional paths should also be considered and not simply the equilibrium curves as shown on these figures.

Several common themes emerge from a consideration of the model using the traditional model forms. First, the use of the equilibrium fit to the Schaefer model consistently resulted in overestimates of the maximum sustainable yield and the effort at MSY. An immediate consequence of this result is that the stock would be overexploited if the management strategy was based on results of the equilibrium fitting. The Schaefer model using the equilibrium approximation method also consistently overestimated MSY and E_{msy} for the simulated data. MSY estimates for the Pella-Tomlinson model were generally more consistent with the actual stock dynamics using both the equilibrium and equilibrium approximation methods. The methods, therefore, appear to be more robust to the estimation method per se than to the specification of the model structure. The simulated stock was generated using an underlying stock dynamic which differed considerably from the logistic form implicit in the Schaefer model. The greater flexibility afforded with the inclusion of a shape parameter in the Pella-Tomlinson model allows this model to mimic more complex stock dynamics. However, there are considerable estimation problems which result from the inclusion of the extra parameter due to the correlation among paramters, particularly m and q. One possible approach to reduce this problem would be to fix the shape parameter at a value consistent with known or assumed recruitment dynamics in much the same way that Shepherd (1982) suggested using ancillary information to fix the shape parameter of his 3-parameter stock-recruitment function.

The time-average method of Schnute performed somewhat better than the equilibrium and equilibrium approximation methods in estimating the actual MSY level, despite the fact that this method is based on the Schaefer model; however, this method consistently overestimated the E_{msy} level. A principal advantage of the Rivard-Bledsoe approach is that the transitional behavior of the stock is treated explictly and examination of the transitional path is very instructive.

All methods gave reasonably consistent estimates of MSY and E_{msy} for the actual data sets regardless of the model form and the estimation procedure. The single exception to this pattern was the estimates for North Sea cod using Schnute's (1977) time-average method which appeared to provide unreasonable results. It is, of course, not possible to evaluate the reliability of any of the methods for the actual data sets since the true stock dynamics are not known.

2.3.4 The Deriso/Schnute model

The Working Group was fortunate to have available a microcomputer implementation of the Deriso/Schnute delay-difference method (Schnute, 1985) written by Carl Walters. Since it was intended for didactic rather than operational use, it was difficult to carry out the necessary runs and extract the results. In addition, the software used was a preliminary version, not originally intended for the purpose for which it was here used, and the Working Group understands that important versions are under development.

The method utilizes a biomass-production representation, with the Deriso (1980) auto-regressive model for growth in weight, and explicit representation of the stock-recruitment relationship using the Deriso (1980) versatile-functional form, which includes the Schaefer, Beverton-Holt, and Ricker forms as special cases. It is, therefore, a delay-difference GMR-explicit model of very general form. Many other models considered are, in fact, special cases of this form. The model is fitted by automatic numerical optimization on any subset of its seven main parameters (in principle).

The results of these runs are, therefore, given in less detail than for the other methods, but are summarized in Tables 2.3.3-2.3.8. The Group's experience, which was confirmed by those members with previous experience with the method, was that, given good data and excellent starting values, the method could usually find a solution for any two of the three parameters q, A, and B. Attempts to solve for these three parameters simultaneously were usually unsuccessful.

Sequentially varying the parameters to be fitted did not necessarily lead to a converging solution and, on real data, was more likely to lead to divergence to extreme parameter values, even when the starting values were near to the correct solutions (insofar as these are known).

These results, therefore, confirm the general conclusion that it is not possible to determine more than one and a half parameters from stock-production data sets, and that there is a large class of possible alternative sets of parameter values which can fit the data, of which not all are reasonable or feasible. Automatic optimization of three parameters (or of two with user intervention) usually leads to solutions wandering in parameter space without noticeable benefit. It is, therefore, most important to explore the range of adequate solutions, which is time-consuming, using programs of this type. The difficulties encountered are common to most methods involving automatic fitting of multi-parameter models (Walters and Ludwig, 1981).

The results on specific data sets were:

a) Simulated data

On the exact data, if (and only if) given good starting values, the method easily found solutions close to the true ones. Cycling the parameters fitted or fitting three parameters, led to solutions departing from the starting values, failure to converge, or overflow failure. Where converged solutions were obtained, the estimates of MSY, etc. were generally reasonable, but the interpretation in terms of q (and, therefore, current biomass) was not.

Very similar results were obtained with the noisy data sets, except that failure was more common. It seemed that the options for allowing for measurement or process error worked better on data sets where the errors were of the opposite type, which is a bit strange.

b) Pacific halibut

Good starting assumptions were available for this data set, and the method had no difficulty returning to these if perturbed slightly. Other starting assumptions led to different results, depending on which parameters were optimized. The method generally failed to converge unless the starting assumptions were very well considered. Significantly different results were obtained using the measurement- and process-error options.

c) North Sea cod

Given reasonable starting assumptions, the method converged to a solution which gives an unreasonable estimate of MSY and biomass (by at least a factor of 10).

d) Southern horse mackerel

No converged solutions were obtained for this stock (the program usually stopped due to execution errors in the first few iterations). The true solution (and, therefore, good starting assumptions) are not known for this stock, and other methods (including eyeball analysis) indicate that the data are not consistent with a stock-production model because of secular changes.

2.3.5 Surplus production models - Shepherd's method

Shepherd's working paper "Towards improved stock-production models" (WP 6) presents a non-equilibrium production model which is described by the three essential parameters: catchability and two production terms. The production parameters are resilience α', and pristine biomass B_{max}. The product of resilience and natural mortality is the maximal P/B ratio at zero biomass. Natural mortality is not estimated in the procedure but rather supplied by the user. Ranges of two other parameters are selected to ensure that only "reasonable" values are used. The final parameter (only) is then determined by fitting to the data. In the simplest case, the fit is obtained simply by constraining the model to pass through the mean estimated production and biomass. A goodness-of-fit map is produced to aid the user in parameter estimations.

The method is constructed in terms of net production, yield, and biomass rather than a catch and CPUE. The formulae and their derivations are not presented here except for the equation for MSY. It was noticed that the equation (Equation 6 in WP 6) did not produce the same values as the author's computer program, which in fact uses a different equation. The MSY in the computer program is calculated from:

$$MSY = \alpha' M \; B_{msy} \; (\frac{1-B_{msy}}{B_{max}}) / \sqrt{1 + \alpha'}$$

where α' is the resilience and B_{max} the virgin or pristine biomass.

The same mapping and fitting procedure can also be used with other production models (including that of Schaefer). This is done either explicitly or by setting the natural mortality to a large number, say 1,000, and the resilience to a small number such that their product is the desired maximum estimated P/B ratio. An example of a Schaefer fit is shown in Figure 2.3.13.

The procedure was reprogrammed into APL and run on a micro-computer. The standard six data sets were run by a user who was unaquainted with the stocks from which they came and had not previously used the model. Because the parameter estimation is interactive, better results would be expected from a user who is familiar with the stocks. Also, ancillary information would aid in the choice of appropriate parameter values. Natural mortality was taken as 0.2 for all runs and the terminal biomass was picked such that MSY would be in the vicinity of the largest catch in the catch history (through this is not a recommended procedure). Results are summarized in Table 2.3.2. Figures showing the fit production curve and scatter of data points are given in Figure 2.3.9-2.3.15. In the simulated data runs, both MSY and B_{msy} were underestimated, the former by about 20% and the latter by about 40%. The results were poorest for the measurement error scenario. When the measurement-error data were rerun using biomass in place of CPUE, the program performed much better. As these observations are based on a single stochastic run, it is impossible to make general conclusions from this observation. The underestimation is an expected bias, given the very crude fitting procedure used in the present implementation, and probably not a fundamental feature.

It was observed that the residual surface was a most useful output. The minimum of the surface was banana-shaped. The sides of the minimum were steeper when the solution was constrained to a Schaefer fit.

2.3.6 <u>Attempts to fit halibut (1932-1986) catch/effort data with a model with uncatchable quantities of biomass</u>

The model (Working Paper 2) used is a Schaefer model where the "qfB" term is replaced by $qf(B-\alpha B_{max})$ and H is a function of α, the latter being the proportion of pristine biomass which is not accessible:

$$H(\alpha) = H_0(1-\alpha)$$

$$dB/dt = H[\alpha(i)]B_t(B_t - B_{max}) - qf(i)[B_t - \alpha(i)B_{max}]$$

with i being the year from 1932 (i = 1) to 1986 (i = 55).

$\alpha(i)$ is fitted by:

$$\alpha(i) = A_0 + i(A_1/55) + i(iA_2)/(55 \times 55)$$

The other parameters are: MSY (α = 0), F_{msy} (α = 0), B_{max}, and B_o (initial biomass).

The criterium to be minimized is $SC = \sum_{i=1}^{55} [(P_i - P_i)/P_i]^2$.

(The program makes adjustment in non-equilibrium conditions, using the subroutine E04FDF of NAG Library.)

Results are:

MSY (α = 0)	= 88.8	
F_{msy} (α = 0)	= 509	
B_{max}	= 503	
B_o	= 236	(with SC = 0.22)
A_0	= 0.34	
A_1	= 0.36	
A_2	= -1.11	

The square root of SC/55 is 0.06, giving the relative mean difference between observed and fitted catches. The value of 100(SCT-SC)/SCT, where

$SCT = \sum_{i=1}^{55} [(P_i - P)/P]^2$ is 94, which indicates a good fit.

Table 2.3.9 gives the observed and fitted catches, biomass at the end of the years, catchabilities, values of the α coefficient, and the differences between observed and fitted catches.

This good fit may be related to the high numbers of parameters incorporated in the model. External information about the existence and importance of an unaccessible biomass may be necessary in practice to reduce linearity problems. In such a case, however, the suggested model may prove useful, to account for catch and effort relationships that would be difficult to explain.

The principal feature is the existence of two "stable" periods separated by a transition period (see Figure 2.3.16).

The first period was characterized with α values between 0.2 and 0.3, high MSY effort, an MSY of about 60, and relative independence between catch and effort. In the 1960s, increasing effort could lead to increasing catches by accessing to new resources, that is, quick decreases in α values. At the end of this transition period (1972), the fishery was in a large overexploitation situation in a Schaefer-type model. The decrease in effort led to the present MSY effort level. The fishery would be now on the way to reach MSY equilibrium, which could be of about 90.

2.4 Conclusions

Several general conclusions can be made regarding the use of surplus production models. First, it is clear that the number of parameters that can be estimated from catch and effort data alone is limited. John Shepherd has in fact suggested that only one and a half parameters can actually be estimated. The Working Group suggests that the "one-and-a-half rule" be kept firmly in mind when attempting to fit surplus production models. More complicated production models with more parameters are particularly difficult to reliably estimate without ancillary information. The models proposed by Deriso (1980) and Schnute (1985) are framed in terms of biologically meaningful parameters which can be estimated independently of catch and effort data. It is clear that use of this auxillary information is essential in estimating the parameters of these methods. This philosophy can be extended for any of the more traditional methods. For example, independent estimates of q can be made and used directly in fitting these production models. Similarly, fixing the shape parameter in the Pella-Tomlinson model to be consistent with known or inferred recruitment dynamics appears to be desirable.

Despite their apparent simplicity, the traditional surplus production models performed reasonably well on simulated data, although E_{msy} tended to be overestimated. The Pella-Tomlinson model appears to be sufficiently flexible to mimic complex stock dynamics. However, auxiliary information should be used in fitting this model. In principle, the delay-difference models which treat recruitment, growth, and mortality individually are preferable. However, they almost invariably will require the use of auxiliary information.

The Working Group recommends that special care be given to consideration of the sums of squares or maximum likelihood surface when using any of the "automatic" fitting techniques. Correlations among parameter estimates can lead to nonsensical results; again, the use of auxillary information can be used to resolve some ambiguities indicated by an examination of the surface.

A careful consideration of the underlying assumptions of the models should be made. For example, Laloe (WP 2) has clearly demonstrated the problems which result when an expansion of the fishing grounds has occured. Similarly, changes in fishery regulations during the time span under consideration will result in a violation of the assumption of constancy of exploitation patterns. Changes in catchability with changes in gear type or population density must also be considered. If least squares or other objective fitting criteria are employed for estimation, the assumptions of the method must also be considered. For example, are the residuals independent? Autocorrelation in the residuals will affect inferences on the reliability of the parameters.

Rivard (1987a) suggests a general strategy for fitting surplus production models: choose a robust estimation procedure for initial estimation. The Gulland equilibrium approximation method appears more robust than other methods when the number of observations is small. If this method produces estimates of MSY and E_{msy} which are within the range of the historical series, more complicated procedures can be tried which directly account for the transient population dynamics. In fitting these non-linear models, several sets of starting values should be tried to guard against local minimum problems. Use independent estimates of the parameters where possible. Examine the parameter estimates and their standard errors. Are the coefficients conceptually acceptable with regard to sign and statistically significant? If not, the model should be discarded. Plot the results and analyse the transient path in relation to the equilibrium curve. Remember that the equilibrium curve and the actual (non-equilibrium) data may be quite different. Deviations from the equilibrium curve may be attributed to the occurence of dominant year classes or changes in fishing patterns.

Consideration of these issues should go a long way towards removing difficulties associated with the application of surplus production models in the past. Despite the potential limitations of these models, they can be used to provide insight into the basic stock dynamics which are not considered in some analytical methods (e.g., yield per recruit). The ideal approach would appear to be the use of models with full age structure and explicit consideration of recruitment dynamics. The models of Deriso and Schnute provide an intermediate approach when comprehensive data on the age structure of the population are not available; these methods may be particularly useful when used in conjunction with ancillary information.

3 ESTIMATION OF RECRUITMENT THROUGH ABUNDANCE INDICES

3.1 Background

Research survey sampling schemes have usually been based upon spatial strata. The sampling variances have been calculated (when they have been calculated) using the corresponding formulae. When the strata considered show a high within-stratum heterogeneity, high variances result for the abundance indices. Reducing the geographical extension of each stratum would reduce the variance, but it becomes increasingly difficult to obtain enough observations in every stratum. It appears that stratification methods tend to consider any spatial variation within a stratum as a perturbing noise, whilst it may really correspond to biological characteristics, which can be partially reproduced from year to year.

After the construction of an abundance index from a survey, procedures must be derived for estimating the recruitment on the basis of past relationships between recruitment (generally estimated through VPA) and corresponding abundance indices.

The calibration of a single series of research survey indices against VPA year-class strengths was dealt with at a previous meeting of this Working Group (Anon., 1984). This has not eliminated all of the problems, and assessment working groups have had to face several difficulties when trying to estimate recruitment.

The questions concern five main topics:

1) Is it helpful to search for consistency between the past observed values for recruitment and the present estimates?

2) How should the estimates coming from different sources be combined?

3) Should the slopes of the regression lines be forced to be 1?

4) Is it legitimate to consider the results from VPA as error free?

5) Should possible trends in catchability be considered?

3.2 Theoretical Considerations

Although the following discussion will refer to the estimation of recruitment, most of the remarks would be relevant for any estimation of abundance, i.e., for any individual age group, exploited or not.

3.2.1 Definition of an abundance index from a research survey

Such an abundance index is usually defined by using the estimation formulae corresponding to stratified sampling schemes. Other possibilities could be considered. The most promising ones are related to various mapping procedures. A simple trend-surface-analysis technique was discussed during the meeting (Houghton, pers. comm.). It makes it possible to take into account the geographical macroscale distribution of the fish. In addition to global abundance indices, it provides indications on the apparent distribution, which will help future interpretations. This will be especially interesting when several years are considered. It is possible to consider a response surface relating the apparent abundance to space and time. The existence of terms corresponding to space x year interactions will show changes in the spatial distribution which will have to be taken into account when estimating year-class strengths.

Another related technique involves the fitting of a multiplicative model when, year after year, the hauls are set at the same locations. This creates a large number of parameters, since the space effects will be described by as many parameters as set locations. It would probably be preferable to reduce the dimensionality by assuming that the space effect can be described by some simple functions of latitude, longitude, and possibly depth. This is done by trend or response-surface techniques.

Another possibility is afforded by kriging and related methods (Matheron, 1965). A connection can be established with response-surface techniques by using so-called universal kriging. This technique considers that the existing estimated spatial distribution results from the combination of a trend, described by some simple function, and a random process, the structure of which can be characterized by a variogram (essentially the mean square difference as a function of the distance between points), which is closely related to a spatial autocorrelation function. Response-surface fitting by least squares is directly related to universal kriging (when the variogram is limited to the so-called nugget effect, i.e., the random component is white noise).

Whatever the technique used, it appears to be very important to map the results of research surveys in order to characterize the main features of the spatial distribution, the differences from species to species and possibly from year to year.

3.2.2 Estimation of a year-class strength from abundance indices

Whatever the technique used, a logarithmic transformation will be considered. On the logarithmic plots, VPA estimates will be put on the x axis and research survey indices on the y axis. In this case, the calibration line corresponds to the regression line where y is predicted from x. Whenever considering the other regression line that will predict x from y, the method will be called a predictive one. This may not be the best convention (it differs from that used previously by the Working Group), but is used for consistency with background papers.

Point 1

Points 1 and 2 can be related. The past observed values bring by themselves, regardless of their use to calibrate the other abundance indices, information about the recruitment one is trying to estimate. When a single series of surveys is considered, two basic estimations can be considered: the historical average (or more precisely the geometrical mean of past values, since logarithmic transformations should be performed) and the estimation suggested by the simple calibration (inverting the regression equation to predict survey indices from VPA). Working Paper 9 shows that this leads, when the series of recruitment estimates is considered as normal white noise, to the traditional predictive regression line. This in fact is equivalent to "shrinking" values that would be obtained through calibration towards the historical geometrical mean considered as a pole. Such a shrinking can also be considered when several abundance indices are simultaneously considered for calibration. Using the Kalman filter, as previously discussed by the Working Group (Anon., 1985a; Pope, 1986), corresponds to another possibility to take advantage of the past series of recruitment estimates. The two points of view can be easily related. The key question is, in fact, to know whether or not it is useful to consider the past series, and especially its average value, as valuable first information.

Point 2

It appears that the simplest combination can be offered by weighted averages. Any weighting should take into account the variance of the different estimators and the length of the corresponding series to avoid attraction by indices corresponding to short time series that will create good fittings which are likely to be unreliable. Working Paper 4 gives a very simple way for combining different indices. It considers, for each index, the empirical calibration line and, for each past observed value of recruitment, the error that would have been commited using this line to estimate recruitment from the abundance index. These errors are squared and then averaged. After correction by a multiplicative factor equal to (n-2)/n, if n is the number of points available for the calibration, this will give an estimate of the mean square error. Weights given to the different indices will be proportional to those estimates of mean square error. The length of each time series does not appear directly in the weighting, but the n-2/n correction factor should avoid biases in the estimation of variances.

Working Paper 10 fits a multiplicative model to various abundance indices, separating year effects from fleet effects (each index being associated by convention to a "fleet"). It also tries simultaneously to estimate the unknown variances associated with the various fleets by using an iterative least-squares procedure. In this technique, the abundance index given for past years by VPA is considered as just another fleet index, the variance of which is also estimated (see following discussion of Point 4).

The maximum likelihood approach can be generalized (see Appendix E) and provide estimates of the last year's recruitment through a "multicalibration" procedure that can also consider the historical geometrical mean, if required. The Kalman filter approach can also automatically take into account the existence of several abundance indices and the historical geometrical mean.

The different variances associated with the various indices are not only useful for a possible weighting. If several recruitment or abundance indices are to be used directly in VPA tuning (see Section 4), estimates of the respective variances may be required. On the other hand, it must also be kept in mind that estimating variances throught short time series is statistically very difficult, if not dangerous. Extreme weightings, giving a very high influence to an individual index, should be avoided. The danger of getting, "by chance", a very low estimate for an individual variance becomes progressively higher when the number of indices increases, as will happen if highly disaggregated data are used. Another reason for avoiding the multiplication of disaggregated abundance indices is the fact that weighting by the reciprocal of variances is optimal only when covariances in the errors from one index series to another one are negligible. This will not necessarily be true when several indices are obtained in a similar way (e.g., several vessels operating at the same time of the year in neighbouring areas can be affected in a similar way by hydrographic events). Finally, it should be recalled that, due to the statistical difficulties of estimating variances, especially when other parameters such as regression coefficient are simultaneously estimated, any direct information will be highly valuable.

Point 3

The problems related to Point 3 (slope of the regression lines) can be viewed from various ways. Several reasons argue for slopes equal to 1. First of all, for the sake of simplicity, it appears reasonable to assume that CPUE is proportional to abundance, at least for research survey vessels. Assuming a slope equal to 1 will reduce the number of unknowns in the fitting procedures and consequently reduce the variability of the estimations. A number of simple statistical tools (e.g., basic linear models) can be more easily used, and the integration of abundance indices within VPA tuning procedures will become much easier (see, for instance, GLIM, ANOVA, or CAGEAN in Section 4). On the other hand, on a number of experimental diagrams, plotting abundance indices against VPA results, "convincing" departures from a slope of 1 can be observed for the slopes of the regression lines. One must, however, avoid being convinced too easily. Testing the statistical significance of an apparent departure from the simplest hypothesis will be helpful. It cannot also be excluded that a real departure could be due to errors in the VPA as an estimate of the true abundance. Misreading of the ages or density-dependent natural mortality could, for instance, create such phenomena. In such cases, the relationship with the true abundance could well show a slope equal to 1 on the logarithmic diagram, even when that with VPA results does not.

The problems will be especially severe if calibration lines have a slope less than 1. In such a context, extreme estimated values far from the historical average can be obtained for recruitment. This would make it dangerous to accept values different from 1 for the slope without shrinking the estimators towards the historical geometrical mean. However, up to now in most examples, this has not been the case. This experience is confirmed by the case studies discussed in the following subsection and would suggest that the risks introduced by freely estimated slopes are not very severe.

Point 4

Point 4 has been touched upon several times in the previous paragraphs. VPA outputs obviously do not really give error-free estimates of abundance. Trying to estimate an extra unknown variance will, however, complicate a problem which is not especially simple. In fact, the only attempt to deal with this problem corresponds to Working Paper 10. An intermediate way could correspond to techniques admitting an assumed level of variance on VPA estimates and then checking the sensitivity of the results to the considered variance. In general, it appears that the variance of VPA estimates of abundance, at least on the first ages, for past years will be small compared to the errors affecting the other indices of abundance.

Point 5

Trends in catchability have been dealt with in a more general context during a previous meeting. From a statistical point of view, it brings one back to the classical choice between reductions in biases and increases in variances. Denying a possible trend in catchability can introduce biases, since such changes can and must occur. On the other hand, including terms describing changes in catchability with time will increase the number of parameters, and so the variance problems, in a way which may be dangerous, especially when flexible functions, allowing for rapid changes, are considered. Working Paper 10 introduced a weighting procedure which, by reducing the influence of "old" data in the model fitting procedures, could reduce the problems created by trends in catchability. For very short time series, trends in catchability should not have much impact, and down-weighting should not be necessary. In other situations (e.g., beyond 10 years), it appears worthwile to use such a weighting. This eliminates, in part, the worst consequences of changing catchability without destabilizing the estimation procedure.

3.3 Case Studies

The methods available have been tested and compared using three data sets: North Sea cod, North Sea haddock, and Irish Sea cod. The performances were compared in two different utilizations: prediction of the 1985 year class and step-through-time validations.

The maximum likelihood calibration method implemented during the meeting was explored more extensively with consideration of different options and combinations thereof in each run: multi-calibration without additional constraint, concentration on the historical mean, Cleveland-type weighting to emphasize recent vs earlier observations $W(y) = \{1-[d(y)/max(d)]^3\}^3$ where d(y) is the number of years of the yth data point from the most recent year, see Cook, WP 10), and forcing the surveys-to-VPA relationships to be linear (slope of the log-log fit forced to 1). Code numbers for these options are listed in Table 3.3.1.

Shepherd's weighted calibration method (WP 5) has been used as well as a variant based on predictive regression lines instead of calibration lines. This in fact induces a shrinkage effect towards the historical geometric mean.

GLIM and Kalman filter results could not be compared since they were based on VPA estimates using constant natural mortality at age.

Cook's method (WP 10) could only be compared in 1985 year-class predictions.

3.3.1 Retrospective analysis

This consisted of using the methods on stepwise increasing time series and predicting successively the strength of the incoming year class, with comparison against the estimate eventually obtained by VPA, as if they had been used by working groups over the years.

Using North Sea cod data from the 1987 North Sea Roundfish Working Group report (Anon., 1987a), the various options of likelihood techniques were compared for the 1973-1984 year classes, and with Shepherd's estimates for the 1981-1984 year classes (Tables 3.3.2 and 3.3.3), due to lack of time.

For the maximum likelihood estimates, the lowest log residual is obtained when the historical mean is taken as a pole. Down-weighting the earliest survey points does not significantly change the residuals. It can be seen on Table 3.3.3 that all options systematically underestimate the strength of the 1977-1982 year classes.

Both of Shepherd's estimates give a better fit of predicted year-class strength to VPA estimates, but their relative advantage is inverted when errors on logarithms or on straight estimates are considered. Each corresponds to a different loss function (see Working Paper 9).

For North Sea haddock (results not shown), the best fit is obtained when the log index/log VPA relationship is forced to be linear; apparently, down-weighting the older observations gives higher residuals. For this stock, the likelihood methods seem to overestimate the recruitment.

A possible explanation of the problems encountered with the maximum likelihood calibrations on the North Sea stocks is the strong influence afforded by the IYFS, which is the longest series, but in which the catchability has significantly changed over the years. Shepherd's *ad hoc* technique seems more efficient in correcting the effects of such a trend.

The Irish Sea cod data, taken from the 1987 Irish Sea and Bristol Channel Working Group report (Anon., 1987b), were treated in two different ways with regard to the indices provided by the pre-recruit gadoid surveys: indices given for the eastern and western areas separately and also combined for the total stock.

In Table 3.3.4, only the totals are considered for the survey series. Shepherd's estimates again give the lowest residuals and among the maximum likelihood estimators, those in which the historical mean is taken as a pole perform comparatively better, while those in which a linear relationship is forced give the largest residuals.

When the separate indices for the eastern and western Irish Sea are considered instead of the totals, the relative performance of the estimators is not changed, but they all give larger residuals than when only the totals are considered. In cases when indices are split spatially, it seems preferable to aggregate them for the total stock area.

3.3.2 Comparison of 1985 estimates for North Sea cod recruitment

The results obtained by simple calibration over the various individual survey indices, as well as those obtained by the different combined techniques, appear in Table 3.3.5.

The differences in the results suggested that the various combined methods may perform quite differently. The variability between the estimates given by individual fleet calibration does suggest in fact that the choice of the weighting factors will have important consequences. A comparison of the weighting factors is made possible by Table 3.3.6. In fact, these coefficients are not similar since Cook's technique operates in a different way. However, they do show that Shepherd's technique gives a much higher weight to Scottish groundfish surveys.

Likelihood techniques give results in a range coherent with those of Shepherd's method at least when slopes are not forced to 1. The high estimates obtained with slopes forced to 1 can be related to the fact that other calibration lines have slopes less than 1 (VPA being on the x axis, survey indices on the y axis).

Cook's method provides a lower estimate than all other techniques. Taking into account the standard deviation provided by Cook's technique would lead to a 95% confidence interval ranging from 470 to 679. This interval includes the other estimates, except for those corresponding to a slope forced to 1.

Finally, it must be pointed out that the retrospective analysis suggests that, at least for North Sea cod, useful recruitment estimates can be built from the survey indices (see Tables 3.3.2 and 3.3.3). Since one can expect a progressive increase in the standardization of operating procedures and improvement of the preprocessing techniques, it seems that research surveys will in the future contribute efficiently in providing necessary auxiliary information to catch-at-age analyses.

3.4 Discussion

3.4.1 Shepherd's and other techniques

The discrepancy between the results obtained by the various methods in the case studies suggests that choosing between them is not a minor problem. Cook's method seems to be in a development stage and should be pursued. Maximum likelihood techniques appear to be developed on a firmer theoretical ground than Shepherd's *ad hoc* technique. However, they are based on a number of assumptions that could well be violated in practice. On the other hand, Shepherd's method, if not optimal in a precise meaning, does not appear to contain any major risk.

It appears that this method should be used until more work has been conducted on the others. It could, however, be useful to implement within Shepherd's techniques the possibility of forcing slopes to 1, as well as introducing weightings.

3.4.2 Retrospective analysis

Whatever method is used, retrospective analysis should be systematically conducted. If users agreed to consider several techniques, such a procedure would offer a basis for a choice. Simulation or resampling techniques could also be useful, but it will be difficult to reproduce the real complexity of the departures from the basic assumptions.

3.4.3 Preprocessing the survey stocks

The fitting of response surfaces and the use of mapping techniques should be developed.

Calculating sampling variances from research surveys could be useful, but great care must be taken in interpreting them. When year after year the hauls occur at the same locations, a sampling variance calculated on the basis of a stratification scheme can well be an overestimate of the variance of survey indices considered as estimates of annual relative abundance. On the other hand, this variance error will also contain other components than those related to sampling (e.g., changes in catchability). A comparison of retrospective errors and sampling variances could be useful.

When very high retrospective errors appear for a survey, it will be legitimate to reanalyze the basic data and the preprocessing techniques. Great care must, however, be taken to avoid reprocessing that would lead to dangerous practices, resulting in meaningless excellent correlations with VPA results, mainly due to the fact that the data had been reprocessed precisely to maximize this correlation.

When several survey indices are available, a balance must be found between the drawbacks of aggregation, which can destroy information, and the statistical risks related to high numbers of survey indices. Going, for instance, beyond ten indices should be avoided before more studies have been conducted. Spatially split indices should be combined.

3.4.4 Weightings

It may be wise, when estimated variances appear to be very high for some indices, to eliminate the corresponding ones, while refining the weightings for the remaining ones. Refining could consist in just taking equal weights, or at least rebalancing the coefficients. Simulations would be useful to check this procedure.

3.4.5 Admitting errors in VPA

Fitting a multiplicative model, as suggested by Working Paper 10, appears to be the best way for allowing for variance in VPA results. The iterative procedure used is not, however, guaranteed to converge to an optimal solution and may "focus" inappropriately on one series or another. The attempt developed by Cook should be further developed, and may be linked to maximum likelihood studies. It could be validated through retrospective and simulation procedures.

The robustness of techniques which do not take into account errors in VPA to the existence of such errors should be checked. All techniques should also be tested in a context of errors in VPA corresponding to white noise but also to more complicated time series, including trends and autocorrelations. This is especially necessary when taking into account the most recent years for the calibration.

3.4.6 Slopes/shrinking

When time series are very short (e.g., less than 6 points), their slope should be forced to 1. But in such a case, shrinking towards the geometrical mean should be simultaneously useful. Departures from slopes equal to 1 must be considered. They do seem to reduce retrospective errors. However, statistical significance tests should not be neglected.

3.4.7 Trends in catchability

It should be avoided, unless statistically demonstrated as being highly necessary, to allow for changes in catchability. It appears preferable to use a weighting, as suggested by Cook (WP 10), or maybe to break long series into shorter ones, considering that a new fleet, with a new catchability, is replacing the old one. Retrospective analysis of catchability by survey, as performed in the North Sea Roundfish Working Group, would help for such splitting.

4 INTEGRATED STATISTICAL ANALYSIS OF CATCH-AT-AGE AND AUXILIARY DATA

4.1 Introduction

The need to carry out combined analyses of catch-at-age and auxiliary data has been recognized for many years. The auxiliary data in question are usually CPUE data from either commercial fisheries or research surveys (or both).

The matter has been discussed in the previous reports of the ad hoc Working Group on the Use of Effort Data in Assessments (Anon, 1984) and in all previous reports of this Working Group (Anon., 1984, 1985a, 1986a). The Working Group recognized at the outset that it would be most desirable to use well-founded statistical models for this purpose and to ensure that proper fitting procedures were used (see Anon., 1984, particularly Appendix F).

Unfortunately, although several workers have attempted to construct and fit such models, no practical procedure has yet emerged for routine use. The methods of Pope and Shepherd (1982), Gudmundsson (1986), and similar ones have all either had difficulty in locating satisfactory solutions or required inordinate amounts of computer time. The most practicable procedure to date is probably that of Doubleday (1981), but the statistical optimality is questionable.

For this reason, the usual procedure within ICES working groups has been to use so-called ad hoc methods for tuning VPAs (see Anon., 1986 and references cited therein), which are capable of coping with the rather extensive data sets (more than 10 years, ages, and fleets) common in the North Sea and elsewhere in the ICES area. This is in spite of the known problems of such methods, notably:

a) the absence of a firm statistical basis;

b) doubts as to whether all parameters estimated are indeed estimable (i.e., whether the solutions are unique);

c) their sensitivity to noise in the most recent data, particularly if CPUE for only one fleet or survey is available.

These deficiencies have been reduced to some extent by the development of methods which take account of the historic precision of the various data sets (i.e., from a weighted mean using variances) and which permit the down-weighting of old (and possibly no longer appropriate) data.

These modifications, however, do not strike at the essence of the problem, which is:

a) to select a plausible family of prior models for the processes involved;

b) to allow for the existence, size, and nature of errors in all the data sets available;

c) to clarify the estimability of the parameters of the models and ensure uniqueness of the solutions;

d) to find reasonably efficient fitting algorithms, so that the methods are capable of being used in a working group environment where many stocks must be examined in a few days. In practice, this means that a 10-age, 10-year, 10-fleet problem should be solvable in less than 1 hour on a microcomputer equipped with a floating point co-processor.

More recently, there have been further developments in integrated statistical models which may provide a basis for progress. The CAGEAN method developed by Deriso is based on a model similar to that used by Gudmundsson (WP 7), and is also available as a reasonably well-tested portable computer program. Pope and Stokes (WP 3) have used a standard statistical package (GLIM) for linearized (multiplicative) approximations of the process equations and have been particularly successful in identifying aliasing (non-estimability in the parameters). Finally, Gudmundsson has proposed a random walk model which [unlike those of Deriso et al. (1985) and Stokes (WP 3)] does not require the assumption that fishing mortality is separable.

The principle questions which need to be addressed are, therefore:

a) Is the assumption of separability necessary or desirable?

b) Is it permissible or desirable to allow catchability to vary for some or all fleets/surveys?

c) Can appropriate weightings be used to take account of the varying precision of the data?

d) Can the estimation of recruitment from surveys be incorporated within the same statistical analysis as is applied to older age groups?

e) Should one allow for non-linearity of the index/abundance relationships?

f) Are there any data relevant to determining selection on the oldest ages? If not, what effect do more-or-less arbitrary assumptions about these parameters have on the results?

The Working Group was not able to deal with all these points in the time available, but considerable effort was devoted to item (b) in particular and to investigating the applicability of CAGEAN to a typical ICES data set.

4.2 Theoretical Considerations

Earlier work on least squares fits to catch-at-age data (Doubleday, 1981; Pope and Shepherd, 1982) indicated that there was insufficient information in catch-at-age data alone to estimate all the mortality terms F(y) and S(α) of a separable VPA model. The problem was most succinctly posed by Shepherd and Nicholson (1986). They observe that $\ln C(a,y) \simeq YC(y-a) + Y(y) + A(a)$, where YC, Y, and A are year-class, year, and age factors and that there is a degeneracy in the design matrix for this problem such that any solution YC(y-a), Y(y), A(a) may be replaced equally well by an alternative solution:

$$YC(y-a) + L(y-a)$$

$$Y(y) - Ly$$

$$A(a) + La$$

where L is an arbitrary factor.

The problem of estimating assessment parameters from catch-at-age data is thus to constrain the value of L (i.e., the trend in the year effect) by using suitable auxiliary data (CPUE, effort, survey) or by making additional assumptions about the parameters. This section discusses some developing approaches.

General linear models

Working Paper 3 contains details of four methods for the statistical fitting of catch-at-age data and auxiliary data. The use of the statistical package GLIM for this purpose was a common theme.

Method 1 was an extension of the simple year-age-year-class ANOVA of catch-at-age data made by Shepherd and Nicholson (1986). The method simultaneously fitted ln catch-at-age data by year-age-year-class factors and ln English groundfish survey catch at age by $\widehat{age}$-year-class factors, where the age factors and year-class factors were common to both data sets and where $\widehat{age}$ specified the difference between catch and survey selection.

Results include relative year class, relative year effect (fishing mortality), and two age factors. This method produces quite sensible interpretations of North Sea cod, but, of course, does not produce the normal assessment parameters.

Methods 2 and 3 need not concern us here.

Method 4 was a multifleet separable effort tuning approach where the catch equation was rendered linear and interpretable by using evolved values of cum Z (cumulative mortality) as an offset in the fit to a linear model.

It is known as the "if thy cum Z offendeth thee, cast it out" method (ITCOTCIO). Its error structure is essentially similar to that of the CAGEAN model (Deriso et al., 1985), and it may prove a useful approach to thinking about multifleet tuning models. In its original form, it was slow to converge and convergence was rather brittle, but both problems are largely solved in Working Paper 4.

Working Paper 4

This paper was an update of some progress made with Method 1 and Method 4 of the previous paper. Method 1 is extended to a multi-fleet separable form with effort tuning and the possibility of catchability change with time and perhaps also with age. The structure indicated that allowing catchability to change on all fleets resulted in a degeneracy in the structural matrix (cf. Shepherd and Nicholson, 1986). It thus gives working groups the very clear advice: WHEN TUNING VPAS WITH CPUE OR EFFORT DATA, DO NOT ALLOW CATCHABILITY TO VARY ON THE EFFORT DATA OR CPUE DATA OF ALL FLEETS. YOU MUST! MUST! MUST! SPECIFY AT LEAST ONE AGE OF ONE FLEET FOR WHICH THE CATCHABILITY DOES NOT CHANGE!!! The paper also shows updates of Method 3 which result in the same lesson. The results from an improved form of the ITCOTCIO model are shown which indicate the need for sensible restrictions on catchability change as noted above. Both models indicate the near linearity of catch-at-age data and hence the usefulness of the ANOVA analogy for giving insight into more complex tuning methods. Both papers seek to help provide insight into the problem rather than to provide practical algorithms.

In particular, the ITCOTCIO model may provide a useful analogy to the CAGEAN model to which it is conceptually similar.

Non-linear models

Non-linearity occurs in log catch-at-age models usually through terms describing cumulative mortality. We can write logarithms of catch as:

$$\ln C(a,y) \simeq YC(y-a) + F(y) + S(a) - [Z(1) + \ldots + Z(a-1)]$$

for a separable fishing mortality model

$$F(a,y) = \exp\,[F(y) + S(a)]$$

Those exponential terms occur in Z and they induce the non-linearity in most catch-at-age models.

Deriso et al. (1985) describe a model and accompanying software package CAGEAN which estimates parameters of non-linear catch-at-age models. Auxiliary information, such as fishing effort data, is used in the procedure to constrain the time trend of log F. A weighting factor λ controls the magnitude of the constraint. The principal assumptions of CAGEAN are that (1) fishing mortality is separable and (2) fishing effort is proportional to true fishing mortality up to a log-normal random variation, as in the model of Fournier and Archibald (1982).

Extensions of CAGEAN to multi-gear data are trivial in theory, but experience is only one realization. Two gear types seem to pose no practical difficulty, but more research is needed for higher numbers of gear types. The two-gear model can be used for an integrated stock assessment where one gear is chosen to be commercial catch-at-age data aggregated over commercial gear types and where the second gear is chosen to be a survey catch-at-age data set. The objective function to be minimized for this problem can be described by the following sum:

minimize RSQ(log commercial catch at age)

+ $\lambda 1$ x RSQ (log survey catch at age)

+ $\lambda 2$ x RSQ (log commercial fishing effort)

+ $\lambda 3$ x RSQ (log survey fishing effort)

where RSQ denotes a residual sum of squares between predicted and observed quantities. Roughly speaking, $\lambda 3$ controls the extent to which survey CPUE is made proportional to survey catch per unit predicted survey fishing mortality rate, while $\lambda 1$ controls the extent to which predicted abundance is forced to agree with predicted survey catch per unit survey fishing mortality rate. We set $\lambda 2$ to a value of zero in our applications described later.

Coefficients for λ must be supplied by the analyst. Deriso et al. (1985) describe the indeterminacy of λ for maximum likelihood functions of the sort considered above. As a consequence, CAGEAN provides a set of hypotheses about abundance time trends where each hypothesis corresponds to a vector of assumed λ coefficients.

Times-series models

Statistical methods for fish stock assessment from catch-at-age data have defined fishing mortality rates uniquely by a number of parameters. Separability of age and year effect is usually assumed. In time-series models of fishing mortality rates (Gudmundsson, WP 7), all Fs are regarded as time series. Their statistical properties are determined by three parameters and assumptions about the correlation structure.

There is no need to assume separability. But the model has both the option of strict separability and random variation of the Fs around a separable pattern.

Given some initial values, the time-series models provide a prediction of the next values of F. These are used to predict the stocks and catches. The actual catches are compared to the predicted ones, and the stocks and fishing mortality rates are adjusted in accordance with the observed catch prediction error before the next values are predicted. The appropriate correction for a given set of catch prediction errors depends both on the properties of the time-series model and the magnitude of the measurement errors of the catches. The estimation procedure (maximum likelihood) seeks the model which produces the best retrospective catch predictions and will tend to find interpretations in which fishing mortality changes are as little as possible from year to year.

This estimation can be carried out without any further information except the rate of natural mortality, but the method will underestimate recent fishing mortality changes unless auxiliary information is also used. The estimated standard deviations appear to give a fair assessment of the accuracy. For actual stocks that have been examined so far, the range of standard deviations for the terminal Fs have been in the range of 10% to over 30%.

The accuracy can be increased by introducing further measurements related to the stocks and fishing mortality rates. Gudmundsson (WP 11) describes joint analysis of catch-at-age data and CPUE from separate fleets or research vessel surveys. The selection is estimated and supposed to be constant during the estimation period. Catchability may be defined as constant or modelled as a time series, but if variations are allowed, the uniqueness of the solutions needs to be examined. Recruitment can be included in a similar way.

The estimation of time-series models takes much longer time than for models of similar size where the pattern of Fs is fixed by the estimated parameters.

Residual analysis

Least squares or maximum likelihood analysis of catch-at-age data assumes certain statistical properties of the residuals, usually that they are independent, normally distributed, and, possibly after appropriate weighting, with equal variances. We do not expect these assumptions to be strictly true, but it is important to detect major discrepancies. Application of these methods should, therefore, be accompanied by analysis of the observed residuals.

In least squares analysis, abnormally large residuals for a particular age or fleet spoil the accuracy. This can often be remedied by weighting.

Gross departure from normality expressed by large kurtosis may be the result of outliers which should be left out or modified.

Correlation between residuals at different ages within the same year are taken into account in some methods. They may often be relatively harmless even if they are left unattended.

Highly significant positive correlations with time or within cohorts strongly indicate that the estimated model is seriously misspecified. For further discussion of residuals, see Gudmundsson (1986).

4.3 Case Studies

In the limited time available at the meeting, it was only possible to make limited studies of the performance of the various methods, and more detailed investigations will need to be conducted between meetings by interested members. Of the methods available, the multifleet ANOVA, the ITCOTCIO, GLIM, and CAGEAN models were run during the Working Group meeting, and only limited comparisons of the results obtained from these were possible with the time-series models (Working Paper 7, Working Paper 11) and with ad hoc tuning methods applied by the North Sea Roundfish Working Group.

ANOVA model

The ANOVA model was implemented on data for North Sea cod and Pacific halibut. For North Sea cod, the model was run on catch-at-age data at two different levels of aggregation. Run 1 used commercial data aggregated to total international level and research vessel data from the English groundfish survey. Run 2 used catch-at-age data for Scottish seiners, Scottish trawlers, Scottish light trawlers, and all other commercial fleets with research vessel data from the IYFS, English and Dutch groundfish surveys, and the Federal Republic of Germany shrimp trawl fishery (a total of eight "fleets").

The method can directly treat catch-at-age and catch-per-unit-effort data as separate entities. The ANOVA model is limited in the size of implementation to about 175 parameters. This means, for example, that results from only eight fleets, seven ages, and nine years could be comfortably integrated. Moreover, the GLIM package is somewhat slow. This might be overcome by using a different STATs pack (SAS or SPSS).

ITCOTCIO model

The ITCOTCIO model, which also runs on GLIM, suffers from similar limitations and is extremely slow for a large implementation due to the need to iterate 10-20 times, which makes it 10-20 times as slow as the ANOVA. Moreover, the ITCOTCIO was unable, in its present implementation, to consider changes in selection.

The ITCOTCIO model was run on data for Pacific halibut.

SURVIVORS model

The SURVIVORS model (Doubleday, 1981) was run using total international commercial catch-at-age data and research vessel data from the English groundfish survey.

The results are shown in Figures 4.3.7 and 4.3.8. They indicate good agreement between the North Sea Roundfish Working Group parameter estimates. This was expected because of the convergence of the VPA and the high fishing mortalities on this stock. Nevertheless, the agreement was still good for recent years with the SURVIVORS estimates being slightly higher than the North Sea Roundfish Working Group estimates.

CAGEAN model

Considerable difficulty was encountered in implementing the CAGEAN model on the NORD computer. (It is thought that the program currently on the NORD is correct, but further testing is required.) Because of the loss of time caused by these difficulties, only a restricted series of implementations was carried out. In its current form, CAGEAN assumes a constant value of natural mortality rate for all ages and years, whereas it is becoming increasingly common in ICES assessment working groups to use age-specific natural mortality rates.

Tests were carried out to assess the effect of varying the parameters $\lambda 1$ and $\lambda 3$. In addition, the age range was varied and, within any defined age range, the ages for which selectivities were fixed were also varied.

4.3.1 Test runs on Pacific halibut

Pacific halibut catch-at-age data were available for the years 1967-1982 with appropriate weight-at-age and fishing effort data. Parts of these data were analyzed by the ANOVA, ITCOTCIO, CAGEAN, and TSA methods.

The ANOVA and ITCOTCIO methods were run on data from 1974-1982 because there was a change in selectivity at that time. CAGEAN was run for the full data set, while Working Paper 7 gives results from 1967-1977.

Figure 4.3.1 compares the trends in fishing mortality estimated by the four methods from 1974-1982.

The CAGEAN and ITCOTCIO models give very similar results, while the ANOVA and TSA have a less variable trend.

Figure 4.3.2 compares the relative year-class strength estimates for the ANOVA, ITCOTCIO, and CAGEAN models. All three models show similar trends in year-class strength.

Figure 4.3.3 compares the exploitation pattern estimates for the ITCOTCIO and CAGEAN models. These show some divergence probably due to an inappropriate choice of terminal value in the ITCOTCIO model. The results of the three figures indicate a close correspondence between the results of the CAGEAN and ITCOTCIO models which might reasonably also be inferred from their similar structure and treatment of errors. The parallel nature of these two models should be explored since they could well prove complementary. The CAGEAN model is used to make practical estimates and the ITCOTCIO model to examine the near-linear structure of estimates. In the time available, it was not possible to consider *status quo* TAC estimates or other final outputs from the models.

4.3.2 Test runs on North Sea cod

Eight runs of the CAGEAN model were performed on data for North Sea cod. Each of these runs used total international catch-at-age data and English groundfish survey research vessel data.

Run	$\lambda 1$	$\lambda 3$	Age(s) for which selectivities fixed	Highest age
1	2.0	0.5	9-10	10
2	2.0	1000	9-10	10
3	1.0	1000	9-10	10
4	1.0	0.5	9-10	10
5	0.25	0.5	9-10	10
6	0.25	0.5	7	7
7	0.25	1000	7	7
8	2.0	1000	7	7

The trials with large $\lambda 3$ just correspond to usual tuning with survey indices (no error assumed in the survey effort), while other runs consider the possibility of poor standardization of the survey effort.

Estimates of total biomass, mean fishing mortality, and recruits at age 1 obtained from these runs are shown in Tables 4.3.1-4.3.3 and Figures 4.3.4-4.3.6. Corresponding estimates obtained by the North Sea Roundfish Working Group are also shown.

Within this set of realizations, widely differing results were obtained in both the time trend and the magnitude of estimates of fishing mortality and biomass. Estimates of recruitment were less sensitive to variation in the input parameters. Comparison with the North Sea Roundfish Working Group results were complicated by the fact that these incorporate the assumption of age-specific natural mortality rates of 0.2 or higher values.

In addition, one run (Run 9) was made using commercial fishery catch-at-age data for Scottish trawlers, Scottish seiners, Scottish light trawlers, Scottish Nephrops trawlers, English trawlers, English seiners, and all other commercial gears; English groundfish survey data were also included. This implementation thus used disaggregated data for eight fleets.

The main value in carrying out this run is that it demonstrates that CAGEAN can be successfully implemented on highly disaggregated data.

It should be stressed that the runs described above were carried out with the intention of gaining experience in running CAGEAN and obtaining some insight into the sensitivity of the model to changes in important parameters. Much more experimentation will be required before any decision can be made on whether CAGEAN can be adopted as a working tool within ICES.

It is apparent, however, from the limited experience gained at this meeting that some modifications of the program are desirable. Preliminary suggestions for modifications are:

i) Include a facility to allow input of age-specific values of natural mortality.

ii) Compute spawning biomass in harmony with the ICES standard SSBs. This will require input of age-specific maturity data.

iii) Input and output of the program should be made compatible with the ICES standard formats and procedures.

iv) Bivariate frequency table of observed catches vs estimated catches as well as analysis of corresponding residuals.

4.3.3 Discussion

The activities of the Working Group were influenced rather more than had been anticipated by the introduction of new statistical models: multiplicative models for separable VPA (Working Papers 3 and 4), CAGEAN, and time-series models of fishing mortality rates (Working Papers 7 and 11).

The multiplicative models and CAGEAN are based on the assumption of separability. This is a very restrictive assumption which may, however, be well founded for individual fleets. With these methods, it may, therefore, often be advisable to work with catches disaggregated by fleets.

CAGEAN has facilities to split the time intervals into blocks if changes in selectivity are supposed to have occurred. The fishing mortality rates are supposed to be constant within each year above a certain age. Effort data are needed for at least one fleet.

The estimation procedure in CAGEAN is least squares, and the weights of the data sets for catch at age and effort data are determined a priori.

The logarithmic transformation has been widely applied in statistical fish stock assessment. Obviously the logarithmic values will not be normally distributed at all levels of aggregation. Problems of non-normality and unequal variances probably increase with the disaggregation of catches between many fleets.

The connection of the multiplicative methods to GLIM can have valuable advantages, e.g., for examining the effects of different transformations. Unlike CAGEAN and the time-series methods, it imposes no constraints on the variation of fishing mortality rates with age. This could presumably easily be changed.

The accuracy of the time-series method depends mainly on the accuracy of the catch-at-age data and, unless good CPUE data are available, the variability of the actual Fs from year to year. Separability is not required. Portable programs for the time-series analysis have not been produced, and the method is based on statistical concepts which are unfamiliar to many biologists.

The value of statistical methods is greatly reduced if the statistical properties of the data differ drastically from the distributions that are assumed implicitly or explicitly in the estimation procedure. This applies also to simplifying assumptions like constant catchability or separability; they increase the precision if they are a reasonable approximation of the actual situation, but if not, they lead to serious errors. Analysis of residuals along the lines discussed in Section 4.2 should become a routine part of the statistical analysis of catch-at-age data and be reported together with the other results. The residuals represent a mixture of measurement errors and random variations in fishing mortality rates. In the time-series method, the variance of the measurement errors is estimated separately from other random elements.

CAGEAN is now available to working groups and others engaged in fish stock assessments. We recommend its use alongside with traditional methods. It is important to collect experience on how far its premises apply to various stocks. For this purpose, its use should be accompanied by analysis of residuals (see Section 4.2).

We have nothing new to contribute on the subject of *ad hoc* VPA tuning except that, in future years, it would be interesting to have an ICES implementation of the SURVIVORS method. It is essential that some constraints be put on estimated changes in catchability. These could be of the form of fixing them for at least one fleet.

It is felt that it would be valuable if this Working Group carried out and presented fish stock assessments and compared the results of various methods. In fact, an attempt at this was made at this meeting, but setting up the programs on the available computers took too much time, so fewer results were obtained than had been expected and less time was available to examine them. We should try to organize this better before the next meeting so that we are able to analyze several data sets using several methods. Some effort is needed to ensure that these sets together represent the main problems encountered in practical work. The following aspects should be included:

1) measurement errors in observed catches and effort;

2) random variations of Fs around a separable pattern;

3) changes in selectivity;

4) changes in catchability.

(Some aspects of simulation are considered in Examples 1 and 2 in Working Paper 7.)

5 CONSEQUENCES OF REDUCED RELIABILITY IN FISHERIES STATISTICS

5.1 Background

In recent years, several stock assessments have been seriously hampered by the lack of reliable, official statistics (ICES Statistician, 1986). However, many working groups have used confidential data supplied by their national representatives. In most cases, the impact of using data of unknown reliability could not be evaluated by the assessment working groups themselves.

Thus, this Working Group studied the effect of reduced reliability of fisheries statistics on stock assessments in general.

5.2 Theoretical Considerations

5.2.1 Approach taken by the Working Group

Although it is possible to predict the effect of changes in input data on the outcome of an assessment analytically, the Working Group preferred to assess the effect of reduced reliability by considering a case study. The basic approach taken was two-fold:

i) The sensitivity of assessment results to reduced reliability of the input data was directly estimated by a sensitivity analysis.

ii) A simulation of different scenarios of how data could have been corrupted by misreportings, and of how this would have misled the regular assessment procedures.

These simulations were restricted to misreportings in landings data and did not address the problems associated with under-sampling, which may also reduce the reliability of data used in assessments.

5.2.2 Data set used for sensitivity analysis and simulations

Sensitivities were calculated using data presented in the North Sea Flatfish Working Group report for 1985 and 1986 (Anon., 1985b, 1986b). The method used is described in Rivard (1982).

Simulated data were generated from the 1972 population numbers and recruitment from the most recent VPA (Anon, 1985b). A constant natural mortality over age and time was used and the exploitation pattern over age was taken from the 1986 North Sea Flatfish Working Group report. The 1984 weights at age (from the 1986 North Sea Flatfish Working Group report) were used for all years. Yearly fishing mortalities were set close to the highest of those for ages 3 and 4 in the most recent VPA. Effort data were generated from fishing mortalities using $q = 0.0001$.

The assessment procedures used were not identical to the procedures taken by the North Sea Flatfish Working Group; their technique contains some subjective expert decisions (fine tuning of the VPA on several CPUE series, with no a priori weight attached to the different series) which the present Working Group did not feel capable of reproducing effectively. Thus, the procedure given in Rivard (1983) was taken. Basically, this procedure consists of a cohort analysis, with fine tuning of the estimated biomass on CPUE data (linear regression through the origin).

5.2.3 Types of misreportings and scenarios tested in simulations

In this section, possible reasons for corruption of official statistics are briefly summarized, and major outlines for simulation runs are extracted from them.

Within the ICES area, the most commonly used management strategy to regulate a fishery is to confine the total catch volume to some level considered to lead towards a gradual improvement in the state of the stocks; no restrictions on effort or fishing capacity are advised. Thus, a structural overcapacity exists, leading to prolonged friction between allowed catch and realizable catch.

In practice, some evidence might exist that the following types of misreporting do occur:

a) Catch and/or effort of certain trips are (partly) not reported.

b) Catch and/or effort of certain trips are reported to stem from a different area.

c) Catch and/or effort of certain trips are reported to belong to a different species.

d) Catch of the higher-valued market categories is selectively underreported.

e) Incidental high catches due to strong year classes may be underreported to circumvent taxes.

Based on these types of misreportings, the Working Group devised a set of 11 scenarios thought to reveal the effect of misreportings as clearly as possible. It should be stressed that the simulated scenarios are not thought to be realistic, but instructive.

The following scenarios were used (summarized in Table 5.2.1):

0) The basic data set as described in the previous paragraph. This data set was taken to represent the truth.

1) A constant underreporting of catch and effort in all years of 20%, irrespective of the age composition (market category). Since this type of misreporting is very consistent, it was expected to have only minor effect on the assessment; it was included only for completeness.

2) Correct reporting of catch and effort in all years, except for the last year, in which both catch and effort are underreported by 20% irrespective of age.

3) Deteriorating reporting of catch and effort: starting 6 years before the last data year, underreporting increased every year by 5%. Again, catch and effort are assumed to be misreported proportionally and irrespective of age.

4) The ratio of reported to unreported catch and effort is assumed to be proportional to the ratio of officially reported to unreported catches as given in the 1985 North Sea Flatfish Working Group report (Anon, 1985b), i.e., this scenario explores what would have happened if the Working Group had used the official statistics.

5) Age-specific underreporting of catches: it was noted that underreporting of higher valued market categories might be worthwhile to circumvent the catch restrictions as well as the income tax. The assumed percentages of underreporting (in numbers) are listed in Table 5.2.1. This age-specific underreporting is assumed to have taken place in all years. It should be kept in mind that the high percentage of underreporting in the older age groups affects only a small catch volume and thus would have been only a minor part of the total catch weight. Efforts are assumed to be correctly reported.

6) Same as 5, but the underreporting is assumed to have occurred only in the last year.

7) Same as 3, combined with 5, i.e., in the last 6 years, there has been an increasing trend to misreport preferentially the older ages, up to 30% of the oldest age in the last year. Efforts are assumed to be correctly reported.

   To study the effect of differential misreporting of catch and effort, three scenarios were included in which catches were assumed to be correctly reported, but efforts to be underreported. Although this may not be a very likely case, it might show the impact of differential misreporting straightforwardly. Furthermore, this scenario also covers possible changes in effort quality without problems in the reporting as such.

8) The first case with differential misreporting of efforts took the correct catches, and 20% underreporting of efforts in all years.

9) Alternatively to 8, efforts were assumed to be correctly reported in all years except for the last year, in which they were underreported by 20%. Again, reported catches were assumed to be correct.

10) Finally, catches were assumed to be correctly reported, while there had been an increasing trend in underreporting of effort from 0 to 30% over the last 6 years.

5.3 <u>Results of Case Studies</u>

5.3.1 <u>Sensitivity analysis</u>

The application of sensitivity analysis to North Sea sole provided insight into the convergence properties of cohort analysis under various conditions and on the potential sources of bias for the estimation of recruitment, stock size, and fishing mortalities. It was found that recruitment estimates are very sensitive to the initial values of fishing mortalities in the last year (Figure 5.3.1). This sensitivity decreases quickly as one goes back in time and as recruitment estimates become more sensitive to the initial estimate of natural mortality. The sensitivity of recruitment to M remained relatively small throughout the time period covered by the analysis.

The sensitivities of calculated recruitment to individual catches are low, except for the current year of catch data. Thus, casual misreporting of catches prior to the current year is not an important source of error in the estimation of recruitment by cohort analysis. If it persists from year to year, misreporting could influence considerably recruitment estimates. However, recruitment figures calculated by cohort analysis would still provide, in that case, a good relative index of recruitment.

Finally, a change in the reporting practice for the current year may also generate spurious trends in recruitment. Thus, the accuracy of sampling estimates of catch in the current year, particularly for younger fish, as well as an analysis of possible changes regarding the reporting (and/or discarding) practice for the current year, should be given prime consideration in the interpretation of trends in the calculated recruitment.

In any assessment, the usefulness of cohort analysis must be evaluated in terms of its ability to produce estimates of stock size and year-class size having desirable statistical properties. Sensitivity analysis provides indications of the importance of a given error in input data for the calculation of recruitment, stock size, and fishing mortalities.

It should be noted that the sensitivity coefficients calculated here for recruitment correspond to the sensitivity of absolute recruitment estimates. The sensitivity of relative changes in recruitment was not analyzed by the Working Group.

5.3.2 Simulation studies

The results of the simulation runs with data sets corrupted by simulated misreportings are summarized in Tables 5.3.1-5.3.5 and Figures 5.3.1-5.3.6. In interpreting these tables, it should be kept in mind that, in an actual misreporting case, unlike the present simulations, one has no outside information on the truth, e.g., in Scenario 5, spawning stock biomass appears to be low compared to the "truth", but this has always been the case, so an assessment working group would have no way of knowing this.

Scenarios 1 and 8 (constant underreporting of catch and effort and effort, respectively) appear to have almost no effect on the assessment at all: exploitation rates are estimated correctly and TACs do predict catches as far as they will be reported.

Increasing underreportings [either sudden (Scenario 2) or as a smooth trend (Scenario 3)] have a very small effect on estimates of $F_{0.1}$ and F_{max}, but current exploitation rate and _status quo_ catch are underestimated. Note, however, that the errors in the estimates are smaller than the error in the catch and effort reportings.

Age-dependent underreporting apparently transforms the long-lived species into a short-lived, heavily-exploited species (Scenario 5).

If the misreporting starts abruptly (Scenarios 6 and 7), the working group may detect that from the changes in exploitation pattern in the converged part of the VPA, but generally not for the most recent years.

Surprisingly, $F_{0.1}$ and F_{max} are correctly estimated in all cases considered, indicating much greater effort reductions than are actually needed. However, this kind of misreporting would lead someone to believe that the stock is in worse shape or condition than it really is (especially in the case corresponding to Scenario 5). One may doubt the disadvantage for most of the stocks assessed.

Finally, disproportional effort misreporting (or equivalently increasing effort quality) does not affect estimates of $F_{0.1}$ or F_{max}, but does seriously affect accompanying TACs. Prolonged misreporting, however, converges to Scenario 8, in which all estimates are correct.

5.4 Conclusions

The sensitivity analysis indicates that cohort analysis (without calibration through the use of an independent index of abundance) provides reliable indices of recruitment and fishing mortality for the "far past". However, these indices may show spurious trends in the recent years. Sensitivity analysis may be helpful in determining which period of a chosen time series is particularly sensitive to a given parameter. A routine examination of sensitivities is desirable and should be considered as an important source of information for the interpretation of trends in calculated quantities.

From the simulation studies, it appears that the assessment method used is rather robust to misreportings (errors in estimates are smaller than errors in misreportings, and convergence in time tends to correct estimates) unless the effort series used are not consistent with the catch reportings.

There are, therefore, three responses which assessment working groups may need to take when data quality deteriorates:

1) When there is substantial misreporting, it should be made clear that any forecasts based on assumed unallocated catches <u>include</u> a proportion of unallocated catches. Managers should be told the size of this proportion and advised to make an appropriate downward adjustment before setting TAC regulations, if the situation is likely to persist.

2) Where it becomes difficult or impossible to determine a best estimate of an intermediate quantity (e.g., of current F or stock size), it may be necessary to explore a feasible range of values and base the advice on whichever value leads to the lowest forecast catches.

3) Where the confidence intervals of estimated quantities, such as catch forecasts, become very wide (because of deteriorating sampling), it would be desirable to give upper and lower estimates (maybe corresponding approximately to upper and lower quartiles) as well as the central estimate, and to advise managers to select an option in the lower part of this range.

All these responses would have the effect of implying lower allowable catches as the data quality deteriorates, without preempting the right of managers to decide on the acceptable level of risk. This would have the incidental advantage of concentrating the minds of managers and fishermen on the need to maintain and improve the quality of the data.

5.5 Recommendations

1) A routine examination of sensitivities of cohort analysis is desirable and should be extended to include more elaborate outputs (e.g., standardized marginal yield and status quo TAC).

2) Software to calculate sensitivities should be made available within ICES.

3) Sensitivity studies should be enlarged to cover the full set of error analyses.

4) As the effect of underreporting fishing effort is to increase forecast catches, reliable effort data are vital for a correct assessment. Consequently, effort and catch series used in an assessment should be consistent.

5) Working groups should clearly state whether TACs do or do not include a proportion of unallocated catch.

6 CONCLUSIONS

6.1 Immediate Recommendations

* Stock-production models are capable of giving useful preliminary analyses for stocks for which detailed data are not available. They usually give reasonable estimates of MSY, but the interpretation of the state of the stock is usually highly uncertain.

* Non-equilibrium models (especially those of the delay-difference type) are preferable in principle, but do not necessarily yield more reliable results in practice. Equilibrium models may give valid results on favourable data sets (i.e., those

with low recruitment variability and high contrast in effort and stock size), but may give unreliable or infeasible results on less adequate data. The data usually employed for stock-production analysis are generally sufficient to determine only about one and a half parameters out of the minimum of three normally required for a non-equilibrium analysis (catchability and two for the production function). It is, therefore, important to acquire as much additional information as possible to constrain the solutions within the multiplicity of possible ones. Reparameterization of the models in terms which are easily understood or may be estimated by analogy is helpful.

The extent of the range of plausible solutions should be explored, and the mapping of goodness-of-fit criteria over feasible parameter ranges is strongly recommended in preference to automatic fitting procedures, which may yield highly variable, confusing, and infeasible results. Fitting more than one or two parameters automatically is very dangerous, and results for ranges of other specified major parameters should be computed.

* Stock-production methods are not valid if exploitation patterns change for the data sets used and should not be employed where this is believed to have occurred.

* Residuals should be examined! They may lead to important insights about the effects of secular (e.g., climatic) changes.

* More elaborate models do not necessarily perform better than simple ones, and the simplest non-equilibrium delay-difference models are to be preferred.

* Plot the data, but be aware of catch/effort plots, since the data follow transient trajectories. Catch/CPUE plots are more closely related to what is fitted by non-equilibrium models.

* Response-surface techniques including both spatial and year effects should be applied to the construction of abundance estimates from research survey data and compared to automatic mapping methods.

* In the immediate future, Shepherd's ad hoc technique (Working Paper 5) should be recommended for use by assessment working groups for combining several abundance indices.

* Retrospective analyses should be systematically conducted for recruitment estimates and VPA tuning as well.

* Future development of statistically based methods for estimating recruitment is highly desirable. Special attention should be paid to the influence of possible errors on VPA estimates and the variance and biasses of the final estimate.

* The most available "constant catchability" data are almost certainly those from research surveys, and such data will become of increasing importance. Existing surveys should, therefore, be maintained as a high priority, and great care should be taken to ensure that their standardization is preserved.

Survey indices for older ages should be routinely provided for all standard age groups.

* Working groups are warned that allowing catchability for all fleets to vary in VPA tuning methods or integrated analysis is likely to lead to incorrect or unstable results. Catchability should always be held constant for at least (one age group for) one fleet or survey. This requires a modification to the present ICES tuning module, which presently only permits either all or none of the catchabilities to vary, and this should be implemented as soon as possible.

* In addition, the F values on the oldest ages should not be set arbitrarily, as they may influence the results when the auxiliary data are not highly informative (e.g., if catchability is allowed to vary). They should be set with care (e.g., to the average of those for several younger age groups in each year). This is an option in the ICES standard VPA suite.

* Integrated statistical models (of catch-at-age and auxiliary data) are free of some difficulties associated with *ad hoc* tuning methods and are in principle preferable to them. It is recommended that assessments should be based on such techniques as soon as operational methods can be implemented and tested.

* The CAGEAN model is the most practicable procedure available at present, and it is recommended that (with the permission of the author) this should be integrated as an additional subroutine within the ICES VPA suite, in order to facilitate its use on standard data sets and permit the production of standard outputs and files.

* The methods based on general linear models are conceptually acceptable, and it should be possible to improve the efficiency of the calculations by fiting the same models directly using NAG subroutines rather than the GLIM package. This should be investigated and, if successful, the procedures should also be implemented as subroutines in the ICES VPA suite.

* ANSI FORTRAN 77 programs for the time-series models of fishing mortality rates would be appreciated, as well as directions for users who are unfamiliar with GLIM on how to apply the multiplicative methods.

* In the meantime, assessment working groups are advised to continue to use these *ad hoc* tuning methods which combine according to variances (and thus also permit the inclusion of survey data). Attention should be concentrated on using data for fleet/surveys for which catchability is believed not to have changed. The utility of data sets for which catchability must be allowed to vary is believed to be low (see above survey indices) and more attention should be paid to standardization of effort and CPUE data before they are analyzed.

* Sensitivities should be calculated and examined on a routine basis. These should concentrate on a sensitivity analysis of

the final product (i.e., the advice) to the various inputs. Software should be adapted to make it as easy as possible.

* When misreporting is suspected, the data sets should be adjusted and the assessment completed with the adjusted values. The robustness of the advice to the adjustment should be evaluated.

* Effort and catch data series used for assessment should be consistent with one another.

* Working groups should consider the effects of misreporting and reduced precision of sampled data and make clear any necessary adjustments to their catch forecasts. The proportion of their estimates due to unallocated catches should be made clear, and an indication of the range of the estimates should be provided wherever possible.

* The work of the Group was greatly facilitated by the availability of the ICES microcomputer and its connection to the NORD machine. The help of the ICES staff was highly appreciated in this connection. The work would, nevertheless, have been impossible without the additional IBM-compatible microcomputers brought to the meeting by Working Group members, and ICES is strongly recommended to acquire several more IBM-compatible machines as soon as funds can be made available. These could be of a lower specification than the existing machine.

6.2 Future Work

6.2.1 Dissemination of the results

Among the Working Group's objectives is the development of more efficient techniques, evaluation of the various methods, and dissemination of its conclusions within assessment working groups. The Group strongly feels that priority must now be given to the last task. The Group noted the process of assimilation of its advice by assessment working groups and ACFM and recommends that national institutes should be encouraged to:

a) send members of the Methods Working Group to regular assessment working group meetings;

b) send appropriate members of assessment working groups to the Methods Working Group.

Publication of the Working Group reports in the Cooperative Research Report should also be continued.

The Working Group notes that methods cannot be adopted in practice unless appropriate software is provided on appropriate machines, and encourages its members (and others) to write portable software, contribute this to ICES, and collaborate with the ICES staff with its integration on the ICES system.

The Secretariat will be requested to make available the services of its staff to assist in the implementation of new methods into the ICES VPA suite.

6.2.2 Special workshop

The Working Group foresees the need to return to the utilization of integrated statistical methods for the analysis of catch-at-age and auxiliary data and to review the experience with their experimental use in the intervening time. In particular, the Working Group should address the question of the integration of the recruitment estimation process.

In order to avoid the problems due to undue time spent in adapting software and constructing and implementing data sets, the Working Group strongly recommends that a special Workshop be held before its next meeting.

The details concerning the suggested organization are found in Appendix F.

ACFM should consider this recommendation, and a decision should be taken as soon as possible.

6.2.3 Next Working Group meeting

The Working Group noted the need for improved methods for the construction of survey indices from raw station data, and also for the further development of CPUE estimates based on detailed analysis of disaggregated data (as opposed to simple aggregation). These problems are closely related, and the Working Group, therefore, proposes that the principal topic for consideration at its next meeting should be: "Construction of CPUE and survey indices by detailed analysis of spatially disaggregated data".

As suggested in Appendix F for the special Workshop, it will be necessary to concentrate on previously-chosen methods, associated to an operational software, and to select data sets prior to the meeting. Such choices would take place by correspondance, under the responsibility of the Chairman.

7 REFERENCES

Almeida, F.P., Fogarty, M.J., Clark, S.H., and Idoine, J.S. 1986. An evaluation of precision of abundance estimates derived from bottom trawl surveys off the Northeastern United States. ICES, Doc. C.M.1986/G:91.

Anon. 1984. Reports of the _ad hoc_ Working Group on the Use of Effort Data in Assessments and of the Working Group on Methods of Fish Stock Assessments. ICES, Coop. Res. Rep. No. 129.

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

Anon. 1985b. Report of the North Sea Flatfish Working Group. ICES, Doc. C.M.1985/Assess:13.

Anon. 1986a. Report of the Working Group on Methods of Fish Stock Assessments. ICES, Doc. C.M.1986/Assess:10.

Anon. 1986b. Report of the North Sea Flatfish Working Group. ICES, Doc. C.M.1986/Assess:13.

Anon. 1987a. Report of the North Sea Roundfish Working Group. ICES, Doc. C.M.1987/Assess:15.

Anon. 1987b. Report of the Irish Sea and Bristol Channel Working Group. ICES, Doc. C.M.1987/Assess:13.

Bard, Y. 1974. Nonlinear parameter estimation. Academic Press. N.Y. 341 p.

Clark, S.H. 1979. Application of bottom trawl survey data to fish stock assessment. Fisheries 4: 9-15.

Deriso, R.B. 1980. Harvesting strategies and parameter estimation for an age-structured model. Can. J. Fish. Aquat. Sci. 37: 268-282.

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

Doubleday, W.G. 1981. A method of estimating the abundance of survivors of an exploited fish population using commercial fishing catch-at-age and research vessel abundance indices. Can. Spec. Pub. Fish. Aquat. Sci. 38:164-178.

Dupont, W.G. 1983. A stochastic catch-effort method for estimating animal abundance. Biometrics 39:1021-1033.

Fogarty, M.J. and Murawski, S.A. 1986. Population dynamics and assessment of exploited invertebrate stocks. p. 228-244. <u>In</u> Jamieson, G.S. and Bourne, N. (ed.) North Pacific Workshop on stock assessment and management of invertebrates. Can. Spec. Publ. Fish. Aquat. Sci. 92.

Fogarty, M.J., Idoine, J.S., Almeida, F.P., and Pennington, M. 1986. Modelling trends in abundance based on research vessel surveys. ICES, Doc. C.M.1986/G:92.

Fournier, D. and Archibald, C.P. 1982. A general theory for analyzing catch-at-age data. Can. J. Fish. Aquat. Sci. 39: 1191-1207.

Fox, W.W. 1970. An exponential surplus-yield model for optimizing exploited fish populations. Trans. Am. Fish Soc. 99:80-88.

Fox, W.W. 1971. Random variability and parameter estimation for the generalized stock-production model. Fish. Bull., U.S. 69:569-580.

Fox, W.W. 1974. An overview of production modelling ICCAT workshop on tuna population dynamics, NANTES, FRANCE. Rec. Doc. Scient. Cicta., 3:142-156.

Fox, W.W. 1975. Fitting the generalized stock-production model by least square and equilibrium approximation. Fish. Bull., U.S. 73:23-37.

Freon, P. 1983. Production models as applied to sub-stocks depending on upwelling fluctuations. In Proceedings of the expert consultative to examine changes in abundance and species composition of veritic fish resources. FAO Fish Rep. No. 291, Vol. 3.

Gray, J.F. 1978. Response of cahort analysis to change in proportion of catch reported. CAFSAC Res. Doc. 78/15, 7 p.

Gudmundsson, G. 1986. Statistical considerations in the analysis of catch-at-age observations. J. Cons. int. Explor. Mer. 43: 83-90.

Gulland, J.A. 1961. Fishing and the stocks of fish at Iceland. Fishery Invest., Lond., Series 2, 23(4), 52 p.

Halldorsson, O., Reynisson, P., and Stefansson, G. 1986. A method for estimating terminal F's from a series of acoustic surveys. ICES, Doc. C.M.1986/H:62.

Hilborn, R. 1979. Comparison of fisheries control systems that utilize catch and effort data. J. Fish. Res. Board Can. 36: 1477-1489.

ICES Statistician. 1986. Progress report 1985/1986. ICES, Doc. C.M.1986/D:19.

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

Jensen, A.L. 1984. Logistic surplus production model with explicit terms for growth mortality and recruitment. Trans. Am. Fish. Soc. 113:617-626.

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

Laurec, A. 1986. Les méthodes Delta en halieutique: évaluation des sensibilités, approximation des biais et variances à l'aide de développements limités. Institut Francais de Recherche pour l'Exploration de la Mer, Rapport DRV-86.002 RH/NTES, 64 p.

Marchessault, G.D., Saila, S.B., and Palm, W.J. 1976. Delayed recruitment models and their application to the American lobster (Homarus americanus) fishery. J. Fish. Res. Board Can. 33:1779-1787.

Matheron, G. 1965. Les variables aléatoires regionalisées et leur estimation. Masson et Cie. Paris, 305 p.

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

Myers, R.A. and Rosenberg, A.A. 1981. Estimation using research survey data and commercial catch data. NAFO, SCR Doc. 85/91.

Pope, J.G. 1986. A possible use of the Kalman filter for making routine updates of recruitment indices. ICES, Doc. C.M.1986/ D:5.

Pope, J.G. and Shepherd, J.G. 1982. A simple method for the consistent interpretation of catch-at-age data. J. Cons. int. Explor. Mer. 40:176-184.

Pope,J.G. and Shepherd, J.G. 1984. On the integrated analyses of catch-at-age and groundfish survey on CPUE data. ICES, Doc. C.M.1984/G:16.

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

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

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

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

Rivard, D. 1983. Effects of systematic, analytical, and sampling errors on catch estimates: a sensitivity analysis. In Doubleday, W.G. and Rivard, D. (ed.) Sampling commercial catches of marine fish and invertebrates. Can. Spec. Publ. Fish. Aquat. Sci. 66:114-129.

Rivard, D. 1987a. Biological production and surplus production models. Unpublished manuscript, Fisheries Research Branch, 200 Kent Street, Ottawa, Ontario, Canada K1A OE6.

Rivard, D. 1987b. Effects of systematic, analytical and sampling errors on the results of cohort analysis: a brief overview. Department of Fisheries and Oceans, Fisheries Research Directorate, 200 Kent Street, Ottawa, Ontario, Canada K1A OE6.

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

Roff, D. 1983. Analysis of catch effort data: a comparison of three methods. Can. J. Fish. Aquat. Sci. 40:1496-1106.

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

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

Schaefer, M.B. and Beverton, R.J.H. 1963. Fishery dynamics - their analysis and interpretation. In Hill, M.N. (ed.) The sea: Ideas and observations on progress in the study of the seas. Vol. 2, p. 464-483. Interscience Publ., N.Y.

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

Schnute, J. 1985. A general theory for analysis of catch and effort data. Can. J. Fish. Aquat. Sci. 42:414-429.

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. and Nicholson, M.D. 1986. Use and abuse of multiplicative models in the analysis of fish catch-at-age data. The Statistician. 35:221-327.

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

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

Tsoa, E., Schrank, W.E., and Roy, N. 1985. Generalizing fisheries models: an extension of the Schaefer analysis. Can. J. Fish. Aquat. Sci. 42:44-50.

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

Walter, G.G. 1973. Delay-differential equation models for fisheries. J. Fish. Res. Board Can. 30:939-945.

Walter, G.G. 1976. Non-equilibrium regulation of fisheries. ICNAF, Sel. Pap. 1:129-140.

Walters, C.J., and Ludwig, D. 1981. Effects of measurement errors on the assessment of stock-recruitment relationships. Can. J. Fish. Aquat. Sci. 38:704-710.

Table 2.3.1 Parameter values used in data simulation.

M = 0.2, Shepherd's K = 6000, g = 2.

Age	Selectivity	Fecundity	Weight
1	0.01	0.00	0.6
2	0.10	0.18	0.9
3	0.50	0.54	2.0
4	1.00	1.35	4.3
5	1.00	2.70	6.7
6	1.00	4.05	8.6

MSGY = 155 t, B_{msy} = 5,670, F_{msy} = 0.53, Rec_{msy} = 900.

Table 2.3.2 Results of simulations.

Year	No noise				Measurement error				Process error-20%			
	Y	E	B	R	Y	E	B	R	Y	E	B	R
1980	1,010	30	5,943	908	1,153	17	6,522	1,021	1,266	39	5,943	908
1981	1,479	40	6,574	908	1,378	43	6,465	837	1,288	37	6,285	923
1982	1,791	50	6,629	930	1,892	35	5,046	692	1,917	55	6,558	885
1983	1,906	60	6,258	935	1,892	64	6,980	1,011	1,725	55	6,041	864
1984	1,920	70	5,787	923	1,703	58	5,380	848	2,150	81	5,828	1,017
1985	1,911	80	5,362	897	2,008	64	5,257	947	1,891	89	5,127	966
1986	1,878	90	4,963	869	2,078	86	5,207	1,010	1,637	84	4,686	776
1987	1,813	100	4,568	840	2,060	99	4,490	1,144	2,287	139	4,749	1,003
1988	1,721	110	4,178	807	2,098	96	4,595	900	1,445	110	3,777	813
1989	1,617	120	3,812	768	1,582	110	3,408	769	1,494	115	3,810	866
1990	1,447	125	3,472	725	1,317	117	3,473	573	1,561	121	3,660	634
1991	1,314	120	3,204	681	1,227	125	2,918	639	1,106	81	3,300	554
1992	1,228	115	3,057	645	1,116	211	3,112	475	1,244	81	3,488	639
1993	1,165	110	2,959	629	1,146	109	3,099	699	1,689	145	3,394	684
1994	1,108	105	2,893	620	977	119	2,823	468	1,059	128	2,786	893
1995	1,066	100	2,865	614	891	124	2,718	595	835	89	2,751	564
1996	1,040	95	2,877	613	898	76	2,806	516	1,244	104	3,177	559
1997	1,027	90	2,921	617	1,140	100	2,775	547	1,158	89	3,147	653
1998	1,020	85	2,995	628	1,043	97	2,658	580	1,090	88	3,099	678
1999	1,020	80	3,096	642	1,196	80	3,679	555	936	73	3,187	779

Table 2.3.3. Population parameters derived from various estimation methods for production models using simulated data with no measurement or process error. See text for description of estimation methods. MSY is the maximum sustainable yield, E_{msy} is the effort level at MSY, F_{msy} is the fishing mortality rate at MSY, B_{msy} is the biomass at MSY, (P/B) is the maximum production to biomass ratio, B_{max} is the maximum biomass, B_t is the current biomass, F_t is the current fishing mortality, q is the catchability coefficient, and m is a shape parameter.

Estimation method		MSY	E_{msy}	F_{msy}	B_{msy}	P/B	B_{max}	B_t	F_t	q $\times 10^{-2}$	m
Equilibrium		1704	74.5	-	-	-	-	-	-	-	2
		1535	68.8	-	-	-	-	-	-	-	1.1
Equil. approx.		1644	73.2	-	-	-	-	-	-	-	2
		1489	53.9	-	-	-	-	-	-	-	0.61
Transitional		1629	73.0	-	1501.4	-	3003	1101	1.19	1.49	2
		1415	62.0	-	3778.0	-	11305	-	-	0.60	0.83
Time average		1575.8	66.0	-	-		-	-		-	-
Deriso/Schnute	P	1250	-	0.15	8342	-	-	-	-	0.29	-
	M	1931	-	0.20	9663	-	-	-	-	0.14	-
	M	1083	-	0.15	7227	-	-	-	-	0.68	-
Shepherd (B-H)		1282	-	-	3896	1.2	14202	2000	0.51	0.60	-
(SCH)		1778	-	0.30	5929	0.6	11857	2000	-	-	-
Actual		1551	53.0	0.53	5670	0.58	10312	3000	0.8	0.01	-

P indicates process error model for Deriso/Schnute method.
M indicates measurement error model for Deriso/Schnute method.
B-H indicates that a Beverton-Holt model was used for the Shepherd model.
SCH indicates that a Schaefer-type model was used for the Shepherd model.

Table 2.3.4 Population parameters derived from various methods for production models for simulated data with process error noise added. See text for description of fitting methods. Definitions of population parameters identical to those in Table 2.3.3.

Estimation method	MSY	E_{msy}	F_{msy}	B_{msy}	P/B	B_{max}	B_t	F_t	q $x10^{-2}$	m
Equilibrium	1647	86.4	-	-	-	-	-	-	-	2
	1765	86.7	-	-	-	-	-	-	-	2.7
Equil. approx.	1650.9	75.0	-	-	-	-	-	-	-	2
	1523	49.0	-	-	-	-	-	-	-	0.5
Transitional	1483.1	60.0	-	3709.8	-	10240	2021	0.49	0.67	0.97
Time average	1650.6	68.8	-	-	-	-	-	-	-	2
Deriso/Schnute P	1200	-	0.2	6000	-	-	-	-	1.00[1]	-
Shepherd (B-H)	1303	-	-	3257	1.6	13027	2000	0.47	0.60	-
(SCH)	2155	-	0.3	7184	0.6	14369	2000	-	-	-
Actual	1551	53.0	0.53	5670	0.58	10312	3000	0.80	0.01	-

Deriso/Schnute measurement error model failed to converge.

[1] Fixed.

Table 2.3.5 Population parameters derived from various estimation methods for production models for simulated data with measurement error. See text for description of fitting methods. Definitions of population parameters identical to those in Table 2.3.3.

Estimation method	MSY	E_{msy}	F_{msy}	B_{msy}	P/B	B_{max}	B_t	F_t	q $x10^{-2}$	m
Equilibrium	1993.5	79.9	-	-	-	-	-	-	-	2
	1531.1	56.8	-	-	-	-	-	-	-	0.60
Equil. approx.	1837	77.1	-	-	-	-	-	-	-	2
	1630	44.7	-	-	-	-	-	-	-	0.61
Transitional	1657	71.0	-	2060	-	4406	1179	0.91	1.14	1.69
Time average	1588	78.4	-	-	-	-	-	-	-	2
Deriso/Schnute P	2764	-	0.30	9224	-	-	-	-	0.26	-
P	1496	-	0.20	7492	-	-	-	-	1.00[1]	-
Shepherd (B-H)	1203	-	-	2597	2.0	11212	4000	0.30	0.40	-
(SCH)	1703	-	0.30	5678	0.6	11356	2000	-	-	-
Actual	1551	53.0	0.53	5760	0.58	10312	3000	0.8	0.01	-

Deriso/Schnute method failed to converge for measurement error method.
Two different process error runs were made.
[1] Fixed.

Table 2.3.6 Population parameters derived from various estimation methods for production models for North Sea cod. See text for description of estimation methods. Definitions for population parameters are identical to in Table 2.3.3.

Estimation method	MSY	E_{msy}	F_{msy}	B_{msy}	P/B	B_{max}	B_t	F_t	q $\times 10^{-3}$	m
Equilibrium	242.5	742.6	-	-	-	-	-	-	0.125	2
	253.4	764.2	-	-	-	-	-	-	-	-
Equil. approx.	237.5	647.2	-	-	-	-	-	-	-	2
	244.2	692.6	-	-	-	-	-	-	-	-
Transitional	250.4	571.0	-	351.2	-	707.7	184	1.14	0.125	2
	247.9	616.0	-	322.0	-	567.2	184	1.14	-	-
Time average	1090.0	340.7	-	-	-	-	-	-	-	2
Deriso/Schnute	1560	-	0.1	15609	-	-	-	-	5.000	-
Shepherd	252	382.0	-	766.0	1.2	2793.0	200	1.06	-	-

Table 2.3.7 Population parameters derived from various estimation methods for production models for horse mackerel. See text for description of estimation methods. Definitions of population parameters identical to those in Table 2.3.3.

Estimation method	MSY	E_{msy}	F_{msy}	B_{msy}	P/B	B_{max}	B_t	F_t	q	m
Equilibrium	213	3.3	-	-	-	-	-	-	-	2
Equil. approx.	176	3.2	-	-	-	-	-	-	-	-
Time average	192	2.4	-	-	-	-	-	-	-	2
Shepherd (B-H1)	138	-	0.46	298	2.0	1288	130	0.55	0.54	-
(B-H2)	210	-	0.25	849	0.8	2749	200	0.36	0.35	-
(SCH1)	146	-	0.60	243	1.2	485	130	0.55	0.54	-
(SCH2)	145	-	0.40	361	0.8	723	200	0.36	0.35	-

Transitional method did not run.
Deriso/Schnute method failed to converge for process and measurement error methods.

Table 2.3.8 Population parameters derived from various estimation methods for production models for Pacific halibut. Definitions of population parameters identical to those in Table 2.3.3.

Estimation method	MSY	E_{msy}	F_{msy}	B_{msy}	P/B	B_{max}	B_t	F_t	q $\times 10^{-4}$	m
Equilibrium	70.0	954	-	-	-	-	-	-	-	2
	70.4	921	-	-	-	-	-	-	-	2.4
Equil. approx.	68.7	81.2	-	-	-	-	-	-	-	2
	68.6	815	-	-	-	-	-	-	-	1.9
Transitional	70.3	652	-	415.4	-	830	540	0.12	2.6	2
	75.0	370	-	845.8	-	3381	585	0.11	2.4	0.50
Time average	74.3	605	-	-	-	-	-	-	-	2
Deriso/Schnute	72.0	-	0.25	288.0	-	-	-	-	0.36	-
Shepherd (B-H)	73.0	-		182.0	1.6	727	150	0.43	0.94	-
(SCH)	74.0	-	0.37	198.0	0.8	395	200	-	-	-

Table 2.3.9 Description of the Pacific halibut fishery data set.

Year	Effort	Observed Catch	Fitted Catch	Biomass	Catchability	Alpha	Obs.-Fit. Catches
1932	1031.	49.	48.	246.	.000709	.346	1.
1933	888.	44.	46.	257.	.000660	.351	-2.
1934	746.	44.	44.	270.	.000700	.356	0.
1935	748.	47.	48.	278.	.000669	.360	-2.
1936	724.	47.	50.	284.	.000663	.363	-2.
1937	624.	47.	46.	294.	.000718	.365	2.
1938	692.	49.	53.	296.	.000642	.367	-4.
1939	624.	49.	49.	301.	.000700	.368	1.
1940	568.	50.	47.	308.	.000733	.368	3.
1941	640.	51.	54.	307.	.000651	.368	-3.
1942	653.	53.	55.	306.	.000672	.366	-2.
1943	613.	52.	52.	307.	.000693	.365	0.
1944	559.	50.	49.	312.	.000708	.362	1.
1945	558.	54.	51.	314.	.000727	.358	3.
1946	479.	53.	46.	321.	.000800	.354	7.
1947	523.	53.	53.	322.	.000701	.349	1.
1948	596.	60.	60.	316.	.000695	.344	0.
1949	562.	56.	57.	315.	.000682	.338	-1.
1950	556.	56.	57.	314.	.000677	.330	-1.
1951	576.	55.	60.	311.	.000639	.323	-5.
1952	598.	57.	62.	306.	.000637	.314	-5.
1953	586.	56.	61.	304.	.000631	.305	-5.
1954	578.	62.	62.	302.	.000698	.295	0.
1955	453.	60.	51.	311.	.000806	.284	8.
1956	542.	72.	65.	308.	.000775	.273	8.
1957	496.	59.	61.	309.	.000674	.261	-2.
1958	527.	69.	67.	306.	.000711	.248	2.
1959	572.	63.	73.	298.	.000596	.234	-10.
1960	561.	67.	72.	294.	.000641	.220	-5.
1961	599.	75.	77.	286.	.000674	.205	-2.
1962	631.	81.	81.	276.	.000695	.189	0.
1963	673.	83.	85.	264.	.000679	.172	-2.
1964	838.	91.	100.	259.	.000632	.155	-9.
1965	969.	98.	104.	211.	.000654	.137	-6.
1966	887.	90.	89.	198.	.000702	.118	1.
1967	1024.	101.	97.	176.	.000725	.099	5.
1968	922.	92.	84.	165.	.000760	.078	8.
1969	843.	83.	78.	161.	.000733	.057	5.
1970	739.	76.	73.	162.	.000713	.036	2.
1971	892.	85.	91.	146.	.000650	.013	-6.
1972	896.	83.	86.	131.	.000670	.000	-3.
1973	918.	81.	79.	118.	.000714	.000	2.
1974	1005.	81.	76.	102.	.000732	.000	4.
1975	994.	64.	66.	91.	.000674	.000	-2.
1976	888.	55.	55.	88.	.000690	.000	0.
1977	765.	48.	48.	92.	.000701	.000	1.
1978	955.	51.	59.	85.	.000604	.000	-7.
1979	731.	44.	45.	92.	.000673	.000	-1.
1980	665.	44.	45.	102.	.000688	.000	0.
1981	691.	49.	51.	111.	.000667	.000	-2.
1982	539.	50.	45.	130.	.000765	.000	5.
1983	425.	47.	43.	160.	.000754	.000	4.
1984	411.	50.	50.	191.	.000686	.000	0.
1985	439.	59.	62.	215.	.000664	.000	-3.
1986	460.	65.	71.	231.	.000629	.000	-7.

Table 3.3.1 Characteristics of the various likelihood methods.

Methods	Shrinking to geom. mean	Cleveland weighting	Slopes forced to 1
1			
2	+		
3		+	
4			+
5	+	+	
6		+	+
7			+
8	+	+	+

Table 3.3.2 Compared performances of maximum likelihood and Shepherd's estimates of year-class strength in retrospective validation.

Option		Year classes			
		1973-1984		1981-1984	
Number	Name	A	B	A	B
1	Basic max. likelihood calibration	0.391	0.298	0.285	0.320
2	Concentration on GM	0.197	0.266	0.155	0.263
3	Cleveland weighting	0.399	0.314	0.300	0.308
4	Slopes forced to 1	0.338	0.346	0.483	0.339
5	GM + weights	0.260	0.288	-	-
6	Weight + slopes 1	0.352	0.333	-	-
7	GM + slopes 1	0.305	0.311	0.443	0.327
8	GM + weight + slopes	0.315	0.323	-	-
	Shepherd-calibration	-	-	0.075	0.119
	Shepherd-prediction	-	-	0.152	0.093

A = Square root of mean square log error.
B = Square root of mean square error divided by mean recruitment (straight values).

Table 3.3.3 North Sea cod. Comparison of year-class strengths obtained by different calibration methods (see Table 3.3.1 for option codes).

Year class	Option							
	VPA	1	2	3	4	7	Shepherd's-C	Shepherd's-P
1973	234	253	263	243	211	262	-	-
1974	426	423	413	426	546	452	-	-
1975	208	206	223	195	187	256	-	-
1976	710	475	455	470	819	558	-	-
1977	427	353	353	311	378	365	-	-
1978	454	375	368	306	313	320	-	-
1979	800	628	627	628	505	505	-	-
1980	271	90	240	97	239	256	-	-
1981	556	532	526	529	457	455	539	527
1982	276	175	260	167	134	141	284	303
1983	552	764	743	750	729	721	638	600
1984	93	84	90	84	54	58	93	122

Table 3.3.4 Irish Sea cod. Compared performances of recruitment estimates for 1981-1984 year classes (see Tables 3.3.1-3.3.2 for option codes).

Option	A	B	1981	1982	1983	1984
1	0.345	0.397	2754	8006	5047	7835
2	0.330	0.375	2902	7810	5060	7594
3	0.317	0.355	2665	7511	5127	7903
4	0.534	0.586	1921	9444	3810	8276
5	0.294	0.325	2861	7230	5136	7655
6	0.484	0.545	2066	9006	4194	8733
7	0.261	0.247	2350	3611	4414	6868
8	0.241	0.227	2515	3347	4706	6891
Shep-C	0.182	0.191	3347	4478	4870	7033
Shep-P	0.238	0.214	4175	4461	4995	6512
VPA	-	-	2922	4375	6819	6849

Table 3.3.5 Comparison of the various estimates for North Sea cod recruitment at age 1 (1985 year class).

Method		Log estimate	Linear estimate
Individual fleet calibration[1]	1	5.7680	320
	2	6.2580	522
	3	6.0180	411
	4	6.2287	507
	5	6.3621	579
	7	6.6352	761
Shepherd calibration		6.4663	643
Shepherd prediction		6.4452	630
Cook method		6.3345	564
Maximum likelihood	1	6.4345	623
	2	6.4003	602
	3	6.4394	626
	4	6.6720	790
	5	6.4036	604
	6	6.6771	794
	7	6.5944	731
	8	6.5985	734

[1] 1 = English groundfish survey, age 0;
2 = Demersal groundfish survey, age 0;
3 = IYFS, age 1;
4 = English groundfish survey, age 1;
5 = Demersal groundfish survey, age 1;
7 = Scottish groundfish survey, age 1.

Table 3.3.6 Comparison of Shepherd's and Cook's weights (1985 year class).

Survey[1]	Shepherd	Cook
1	0.0384	0.0885
2	0.2217	0.1767
3	0.0204	0.0556
4	0.0948	0.1683
5	0.1046	0.2205
7	0.5276	0.2905

[1] 1 = English groundfish survey, age 0;
2 = Demersal groundfish survey, age 0;
3 = IYFS, age 1;
4 = English groundfish survey, age 1;
5 = Demersal groundfish survey, age 1;
7 = Scottish groundfish survey, age 1.

Table 4.3.1 Estimates of recruitment at age 1. North Sea cod from CAGEAN runs (based on seven commercial gears and one survey) and from the 1987 report of the North Sea Roundfish Working Group.

Parameter	Run 1	2	3	4	5	6	7	8	9	North Sea Roundfish WG 1987
λ_1	2.0	2.0	1.0	1.0	0.25	0.25	0.25	2.0	0.25	
λ_3	0.5	1000	1000	0.5	0.5	0.5	1000	1000	1000	
SR	9-10	9-10	9-10	9-10	9-10	7	7	7	7	
Year										
1977	344	241	274	321	313	379	386	525	310	710
1978	225	185	199	237	220	231	228	265	212	427
1979	196	190	197	230	229	227	215	246	197	454
1980	332	459	414	352	372	365	351	398	285	800
1981	137	208	193	133	145	145	146	156	129	271
1982	311	320	309	291	298	290	302	320	335	556
1983	198	159	151	166	153	145	151	167	152	276
1984	355	371	351	336	301	320	331	448	353	552
1985	23	40	41	43	43	49	51	48	233	93
1986	66	318	336	352	365	529	539	433	542	730

Table 4.3.2 Estimates of mean fishing mortality for North Sea cod from CAGEAN runs (based on seven commercial gears and one survey) and from the 1987 report of the North Sea Roundfish Working Group.

Parameter	Run 1	2	3	4	5	6	7	8	North Sea Roundfish WG 1987
λ_1	2.0	2.0	1.0	1.0	0.25	0.25	0.25	2.0	
λ_3	0.5	1000	1000	0.5	0.5	0.5	1000	1000	
SR	9-10	9-10	9-10	9-10	9-10	7	7	7	
Year									
1977	0.71	0.74	0.73	0.71	0.71	0.90	1.12	0.35	0.72
1978	1.05	0.80	0.87	0.85	0.78	1.07	1.36	0.42	0.81
1979	0.72	0.53	0.63	0.68	0.66	0.96	1.21	0.32	0.70
1980	0.60	0.41	0.52	0.70	0.71	0.44	1.12	0.28	0.78
1981	0.70	0.65	0.63	0.74	0.75	0.47	1.15	0.31	0.77
1982	0.89	1.31	1.22	0.81	0.87	1.16	1.38	0.34	0.90
1983	0.82	0.80	0.88	0.81	0.84	1.09	1.34	0.33	0.89
1984	1.17	1.02	0.90	0.86	0.88	1.04	1.28	0.29	0.88
1985	1.05	0.86	0.94	0.81	0.87	0.91	1.07	0.25	0.85
1986	2.50	1.00	1.01	0.85	0.98	0.81	0.93	0.23	0.91

Table 4.3.3 Estimates of biomass ('000 t) for North Sea cod from CAGEAN runs (based on seven commercial gears and one survey gear) and from the 1987 report of the North Sea Roundfish Working Group.

Parameter	Run									North Sea Roundfish WG 1987
	1	2	3	4	5	6	7	8	9	
λ_1	2.0	2.0	1.0	1.0	0.25	0.25	0.25	2.0	0.25	
λ_3	0.5	1000	1000	0.5	0.5	0.5	1000	1000	1000	
SR	9-10	9-10	9-10	9-10	9-10	7	7	7	7	
Year										
1977	479	430	419	482	459	439	432	900	459	704
1978	512	421	438	505	486	507	502	842	576	705
1979	401	409	413	469	481	485	466	828	407	702
1980	459	577	533	532	554	528	498	960	489	884
1981	520	769	665	533	559	533	517	1042	392	739
1982	521	764	691	498	519	497	502	1018	509	734
1983	466	446	435	451	443	416	431	972	427	558
1984	499	498	455	467	437	422	436	975	520	633
1985	329	381	383	384	352	370	381	943	342	406
1986	180	365	351	374	351	472	490	1000	667	632

Table 5.2.1 Scenarios used in testing the robustness of VPA to misreportings.

Scenario		Years	Ages	Catches	Effort	CPUE
0	No misreporting	-	-	-	-	-
1	Constant misreporting	All	All	20%	As catch	-
2	Pulse misreporting	Last	All	20%	As catch	-
3	Year trend in misreporting	Trend	All	0-30%	As catch	-
4	Observed misreporting	All	All	Observed	As catch	-
5	Age trend in misreporting	All	Trend	0-99%	-	Calculated
6	Pulse, age trend	Last	Trend	0-99%	-	Calculated
7	Age and year trend in misreporting	Trend	Trend	0-99%	-	Calculated
8	Constant effort misreporting	All	All	-	20%	-20%
9	Pulse effort misreporting	Last	All	-	20%	-20%
10	Trend effort misreporting	Trend	All	-	0-30%	0-30%

In cases where misreportings varied with year, the following percentages of misreporting were used:

Year	1978	1979	1980	1981	1982	1983	1984
Misreporting	0	5	10	15	20	25	30

In cases where misreportings varied with age, the following percentages of misreporting were used:

Age	1	2	3	4	5	6+
Misreporting	10	20	30	60	80	99

Table 5.3.1 Exploitation patterns as estimated from the data of the various scenarios.

Scenario		Age													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	No misreporting	0.01	0.34	1.00	1.00	0.89	0.80	0.77	0.77	0.76	0.66	0.58	0.51	0.51	0.51
1	Constant misreporting	0.01	0.34	1.00	1.00	0.89	0.79	0.77	0.75	0.75	0.70	0.56	0.51	0.51	0.51
2	Pulse misreporting	0.01	0.34	1.00	1.00	0.90	0.81	0.79	0.78	0.78	0.74	0.59	0.54	0.54	0.54
3	Year trend in misreporting	0.01	0.34	1.00	1.00	0.90	0.83	0.80	0.79	0.75	0.57	0.55	0.54	0.54	0.55
4	Observed misreporting	0.01	0.34	1.00	1.00	0.93	0.83	0.82	0.82	0.82	0.72	0.67	0.56	0.56	0.56
5	Age trend in misreporting	0.01	0.43	1.00	1.00	0.91	0.70	0.68	0.67	0.62	0.57	0.45	0.45	0.45	0.45
6	Pulse, age trend	0.01	0.31	0.93	0.98	0.99	1.00	0.97	0.96	0.96	0.83	0.82	0.80	0.80	0.80
7	Age and year trend in misreporting	0.01	0.34	1.00	1.00	0.92	0.80	0.78	0.78	0.73	0.67	0.53	0.53	0.53	0.53
8	Constant effort misreporting	0.01	0.34	1.00	1.00	0.89	0.80	0.77	0.77	0.76	0.66	0.58	0.51	0.51	0.51
9	Pulse effort misreporting	0.01	0.34	1.00	1.00	0.89	0.80	0.77	0.76	0.71	0.57	0.52	0.51	0.51	0.51
10	Trend effort misreporting	0.01	0.34	1.00	1.00	0.89	0.79	0.77	0.75	0.75	0.70	0.56	0.51	0.51	0.51

Table 5.3.2 Recruitment at age 1 (thousands) estimated from the data of the various scenarios.

Scenario		Year												
		1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984
0	No misreporting	77868	107136	111255	42157	114383	140553	47536	12342	159268	155947	169222	199106	56503
1	Constant misreporting	62295	85710	89044	33697	91541	112515	38052	9887	127402	124708	135279	159164	45167
2	Pulse misreporting	77818	107036	111175	42063	114122	140094	47313	12266	157266	152326	161911	193160	53916
3	Year trend in misreporting	77387	106049	109538	41137	109126	129093	41672	10352	129531	126390	142483	176425	49453
4	Observed misreporting	72873	94239	85348	27845	66221	89372	40440	11704	144503	128871	133735	160062	44165
5	Age trend	41085	55784	56023	21241	57302	71619	24208	6360	86370	82673	88448	75104	19870
6	Pulse, age trend	77213	105830	109565	41380	110411	133008	43753	10784	130592	112818	123917	155164	48341
7	Age and year trend	77259	105794	109191	40918	109131	131073	43293	11004	140105	135577	149235	177001	51515
8	Constant effort misreporting	77868	107136	111255	42157	114383	140553	47536	12342	159268	155947	169222	199106	56503
9	Pulse effort misreporting	78036	107470	111917	42480	115640	142402	48461	12714	167012	170972	199561	251962	73080
10	Trend effort misreporting	78092	107581	111889	42361	115703	142072	48770	12852	169658	175928	209568	269359	78527

Table 5.3.3 Spawning stock biomass (t) estimated from the data of the various scenarios.

	Scenario	Year 1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984
0	No misreporting	41568	26792	24605	25361	27724	22953	25237	30080	24299	17300	25101	30280	36989
1	Constant misreporting	33263	21440	19691	20295	22192	18372	20200	24087	19469	13876	20114	24251	29603
2	Pulse misreporting	41414	26659	24460	25204	27548	22755	25048	29812	23973	16944	24408	28938	36697
3	Year trend misreporting	41016	26301	24018	24599	26631	21595	23140	26450	20496	14204	20365	25282	33799
4	Observed misreporting	38929	24253	21109	19616	18901	14314	15855	24113	22647	15403	20670	23441	29942
5	Age trend	5411	2547	3222	4233	4746	2741	4479	5908	3159	1238	5674	7060	8907
6	Pulse, age trend	40889	26182	23865	24408	26431	21427	23236	26684	19901	12246	14941	11252	15633
7	Age and year trend	41034	26310	24005	24538	26502	21409	22929	26286	20081	13470	19620	23189	28858
8	Constant effort misreporting	41568	26792	24605	25361	27724	22953	25237	30080	24299	17300	25101	30280	36989
9	Pulse effort misreporting	41573	26806	24648	25465	27951	23266	25784	30993	25470	18640	27763	35651	48019
10	Trend effort misreporting	41578	26814	24666	25504	27990	23289	25812	31131	25682	18899	28465	37247	51491

Table 5.3.4 Fishing mortality averaged over the ages, estimated from the data of the various scenarios.

Scenario		Year												
		1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984
0	No misreporting	0.607	0.587	0.585	0.610	0.527	0.565	0.524	0.619	0.534	0.537	0.696	0.564	0.568
1	Constant misreporting	0.607	0.587	0.585	0.610	0.527	0.565	0.524	0.618	0.534	0.536	0.695	0.563	0.568
2	Pulse misreporting	0.609	0.589	0.587	0.612	0.529	0.568	0.527	0.623	0.540	0.547	0.713	0.589	0.456
3	Year trend in misreporting	0.613	0.596	0.596	0.625	0.545	0.596	0.569	0.668	0.570	0.556	0.688	0.506	0.434
4	Observed misreporting	0.640	0.641	0.673	0.784	0.614	0.665	0.300	0.308	0.571	0.600	0.842	0.544	0.506
5	Age trend	1.385	1.486	1.429	1.425	1.268	1.495	1.242	1.414	1.394	1.630	1.423	1.151	1.099
6	Pulse, age trend	0.614	0.598	0.599	0.629	0.547	0.597	0.563	0.690	0.643	0.744	1.114	1.499	0.622
7	Age and year trend	0.613	0.596	0.597	0.627	0.548	0.602	0.573	0.687	0.604	0.611	0.799	0.642	0.611
8	Constant effort misreporting	0.607	0.587	0.585	0.610	0.527	0.565	0.524	0.619	0.534	0.537	0.696	0.564	0.568
9	Pulse effort misreporting	0.607	0.587	0.583	0.607	0.522	0.557	0.513	0.602	0.511	0.501	0.634	0.481	0.438
10	Trend effort misreporting	0.607	0.586	0.583	0.606	0.521	0.557	0.513	0.599	0.507	0.493	0.617	0.459	0.408

Table 5.3.5 Projections estimated from the data of the various scenarios (TACs for 1986, SSB of 1984).

	$F_{0.1}$		F_{max}		F_{last}		$\frac{dY}{dE} \times \frac{E}{Y}$
Scenario	F	TAC	F	TAC	F	TAC	
0 No misreporting	0.211	11577	0.348	16278	0.58	20627	-17%
1 Constant misreporting	0.211	9294	0.349	12984	0.58	16494	-17%
2 Pulse misreporting	0.206	11672	0.340	16360	0.47	19301	-12%
3 Year trend misreporting	0.209	10923	0.346	15251	0.45	17424	-10%
4 Observed misreporting	0.200	9318	0.332	13117	0.52	16341	-15%
5 Age trend	0.225	3448	0.368	4794	1.10	6913	-22%
6 Pulse, age trend	0.176	6883	0.295	10006	0.66	14760	-20%
7 Age and year trend	0.210	9855	0.347	13798	0.62	17945	-16%
8 Constant effort misreporting	0.211	11577	0.348	16278	0.58	20627	-17%
9 Pulse effort misreporting	0.215	15629	0.355	21703	0.45	24566	-10%
10 Trend effort misreporting	0.211	16699	0.349	23258	0.42	25646	-3%

Figure 2.3.1 Equilibrium Yield vs. F-Stimulated data

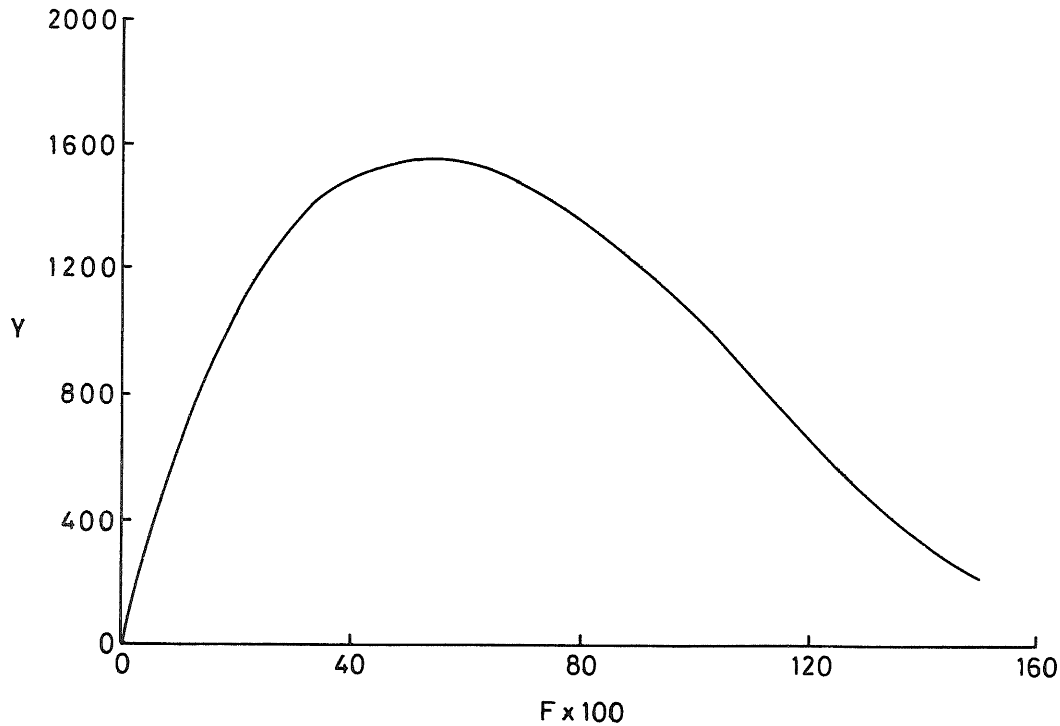


Figure 2.3.2 Equilibrium Stock-Recruit-Simulated data

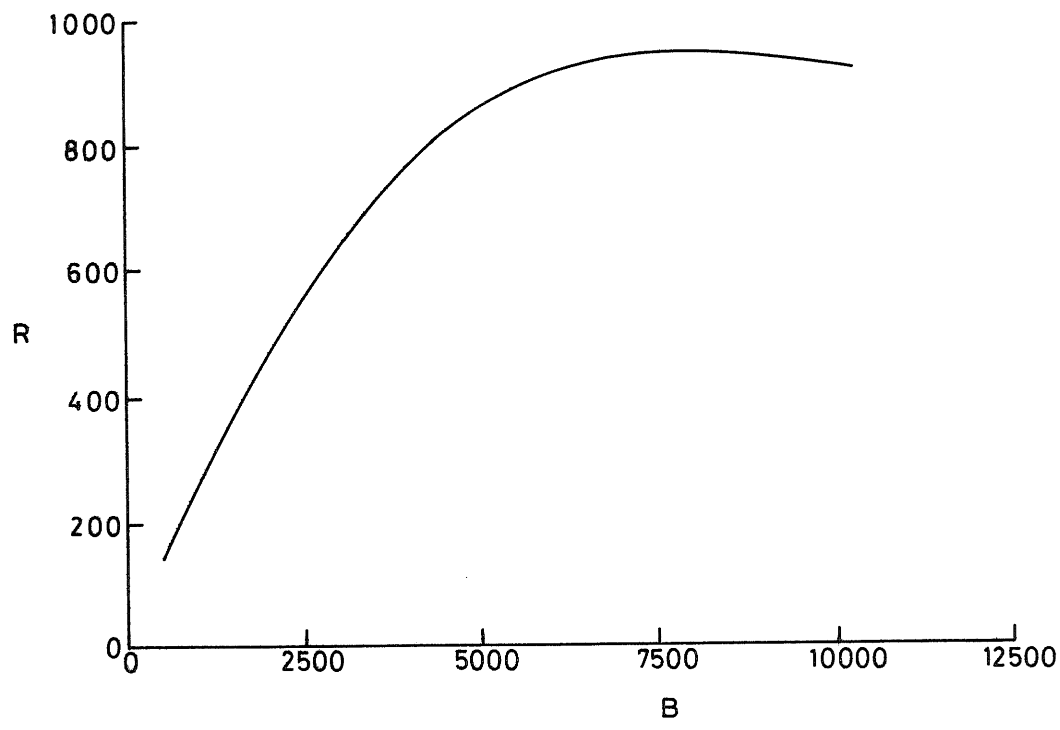


Figure 2.3.3 SIMULATED DATA — NO NOISE

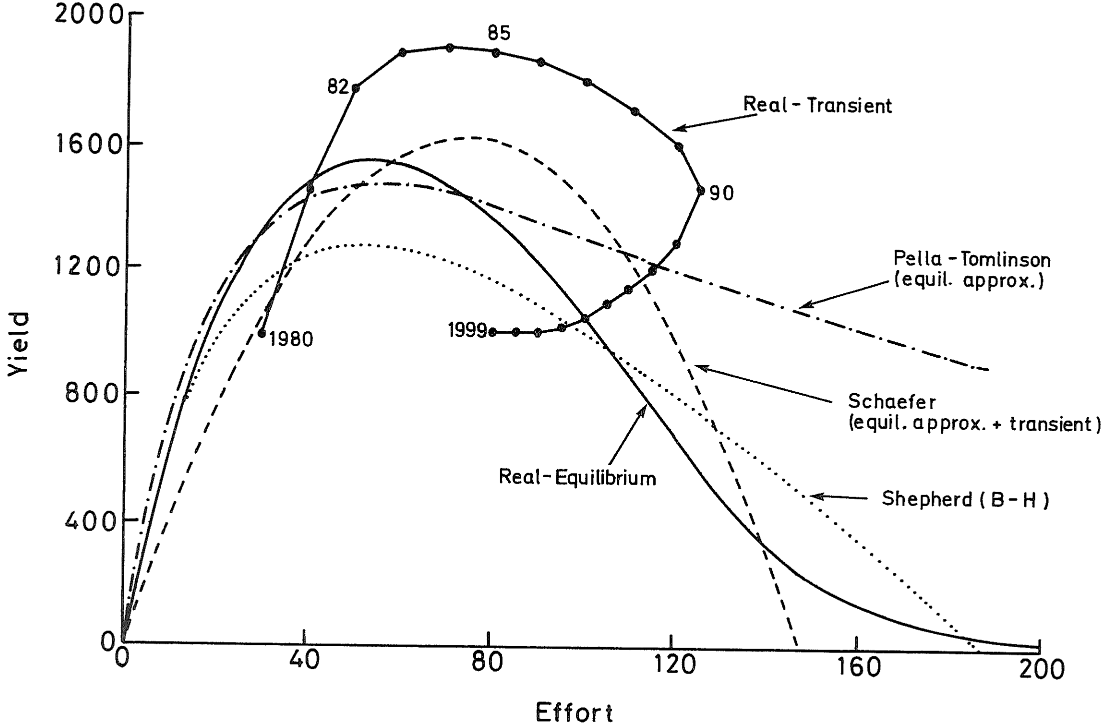


Figure 2.3.4 SIMULATED DATA — MEASUREMENT NOISE

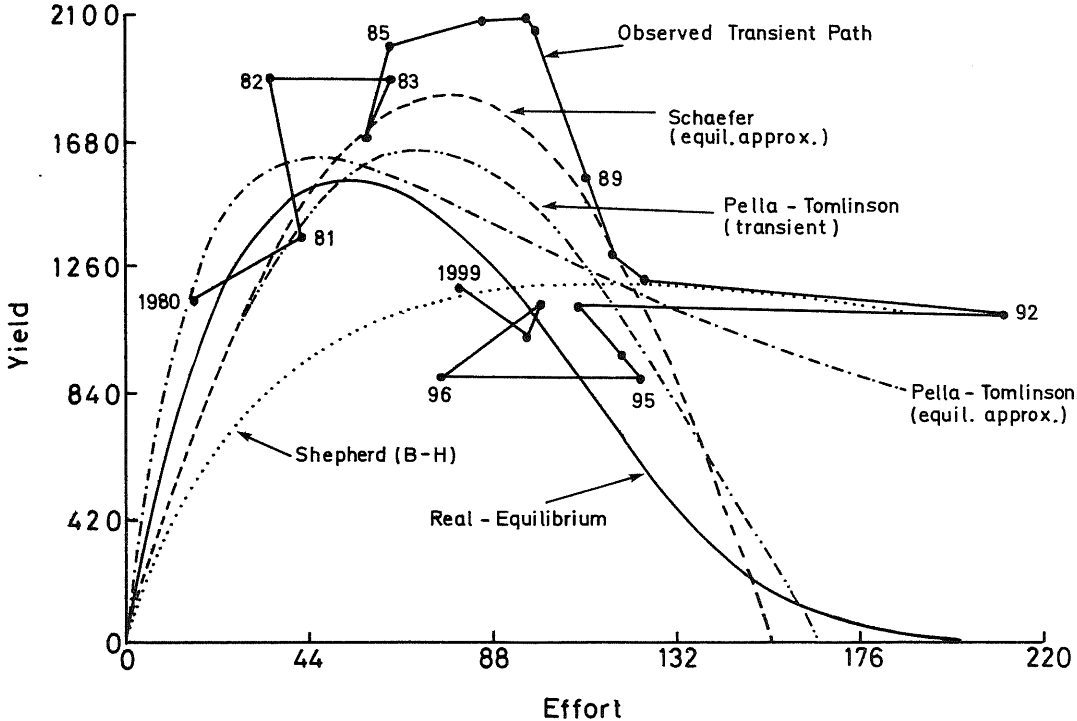


Figure 2.3.5 SIMULATED DATA — PROCESS ERROR

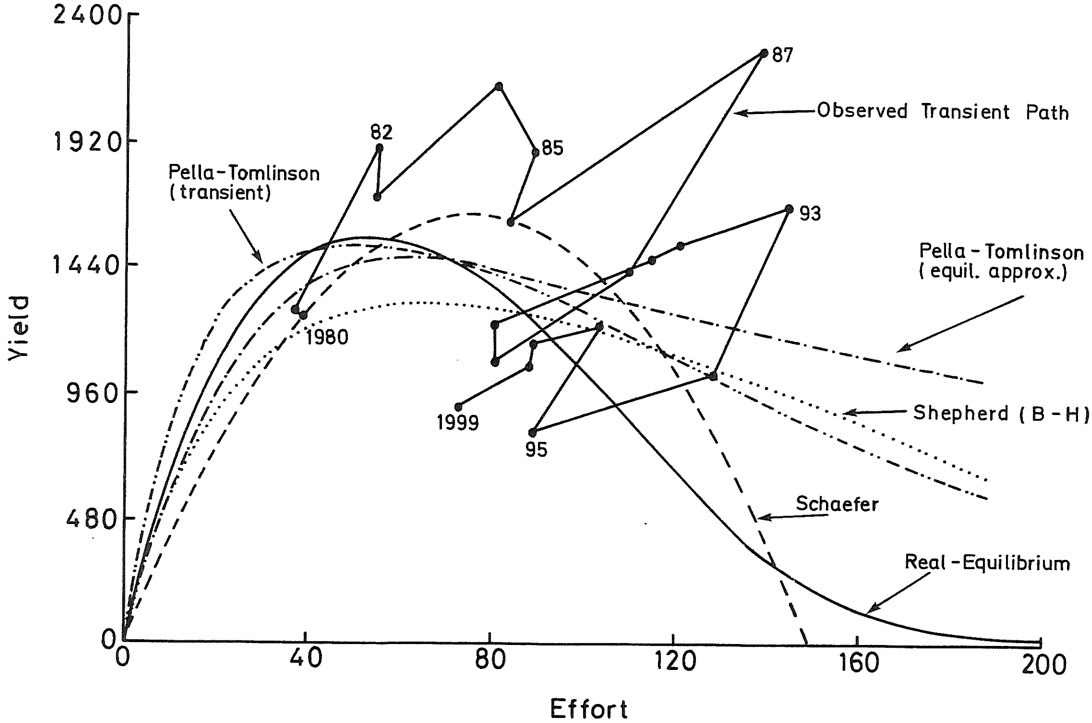


Figure 2.3.6 NORTH SEA COD

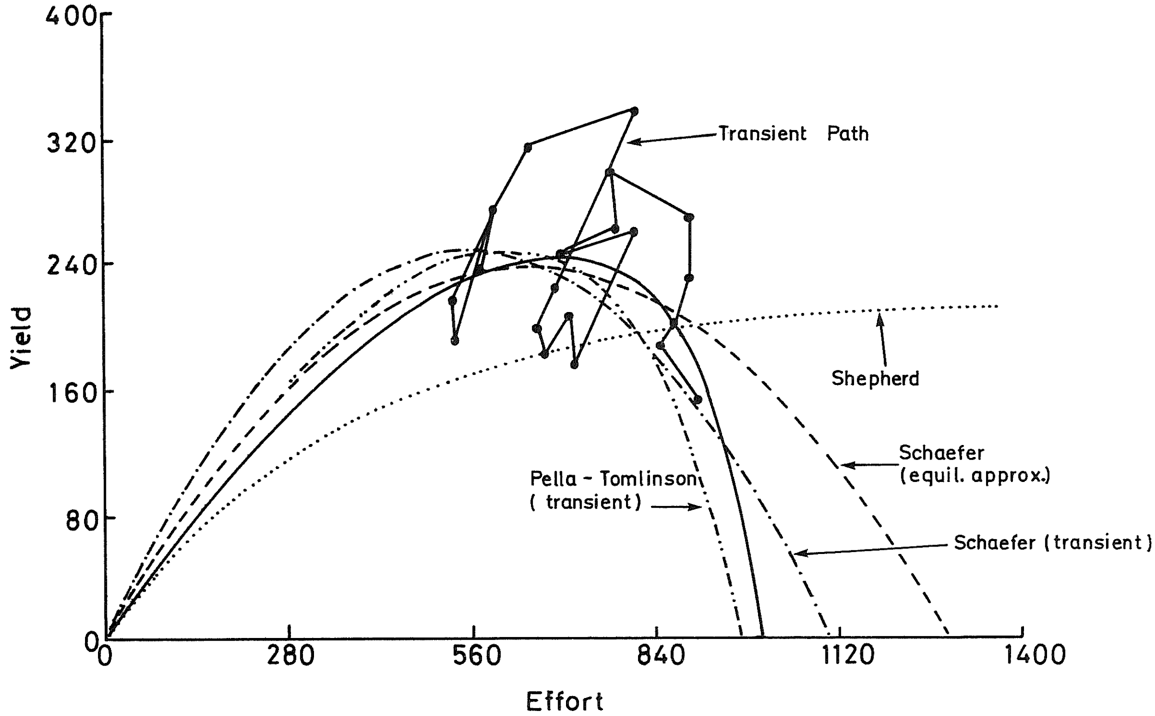


Figure 2.3.7 S. HORSE MACKEREL

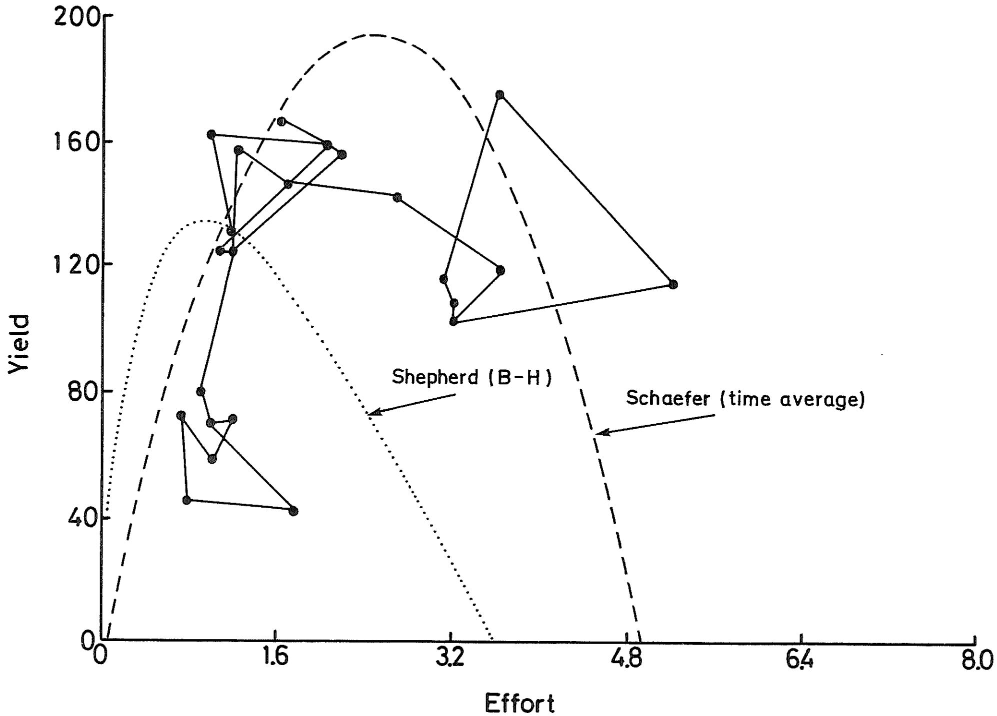


Figure 2.3.8 PACIFIC HALIBUT

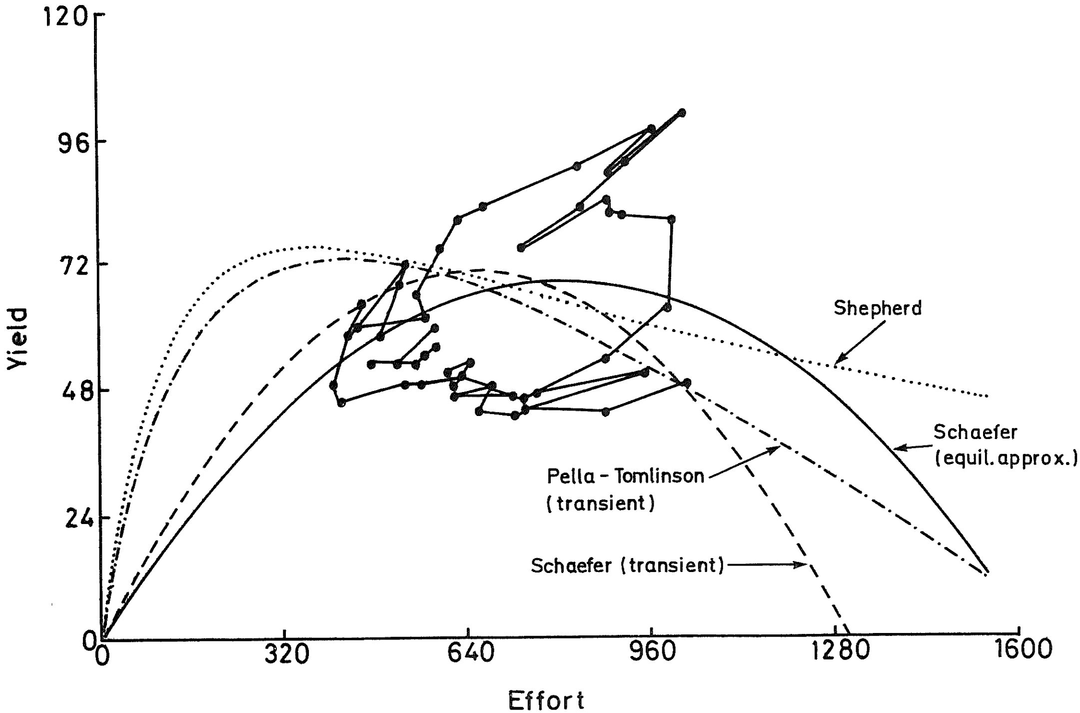


Figure 2.3.9 SIMULATED DATA NO NOISE

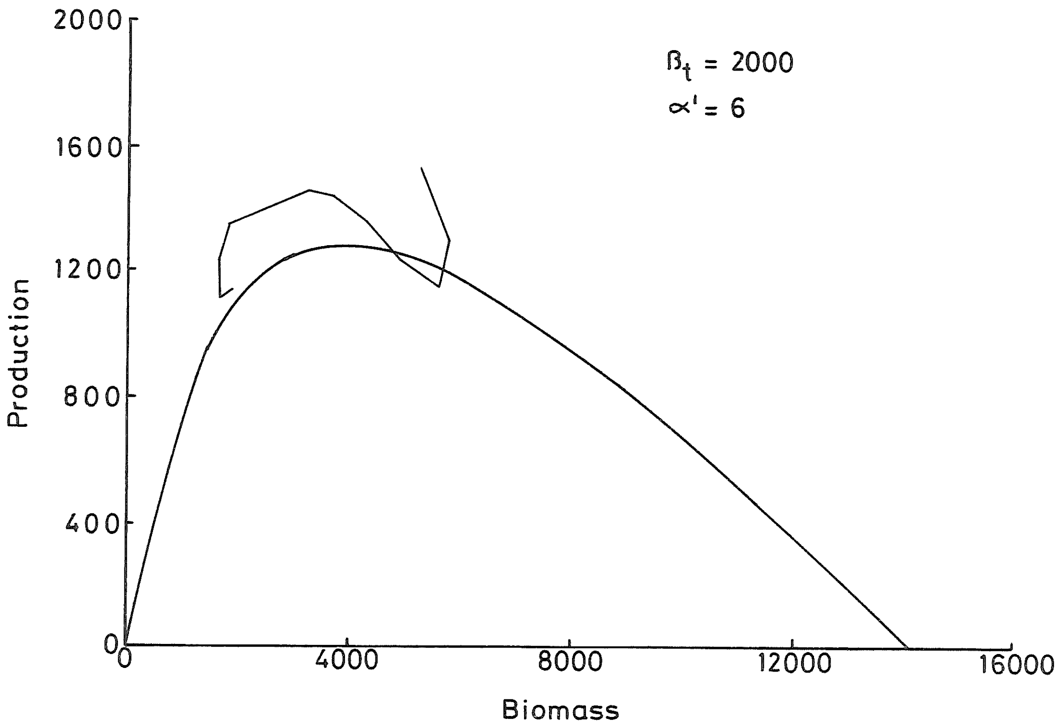


Figure 2.3.10 SIMULATED DATA MEASUREMENT NOISE

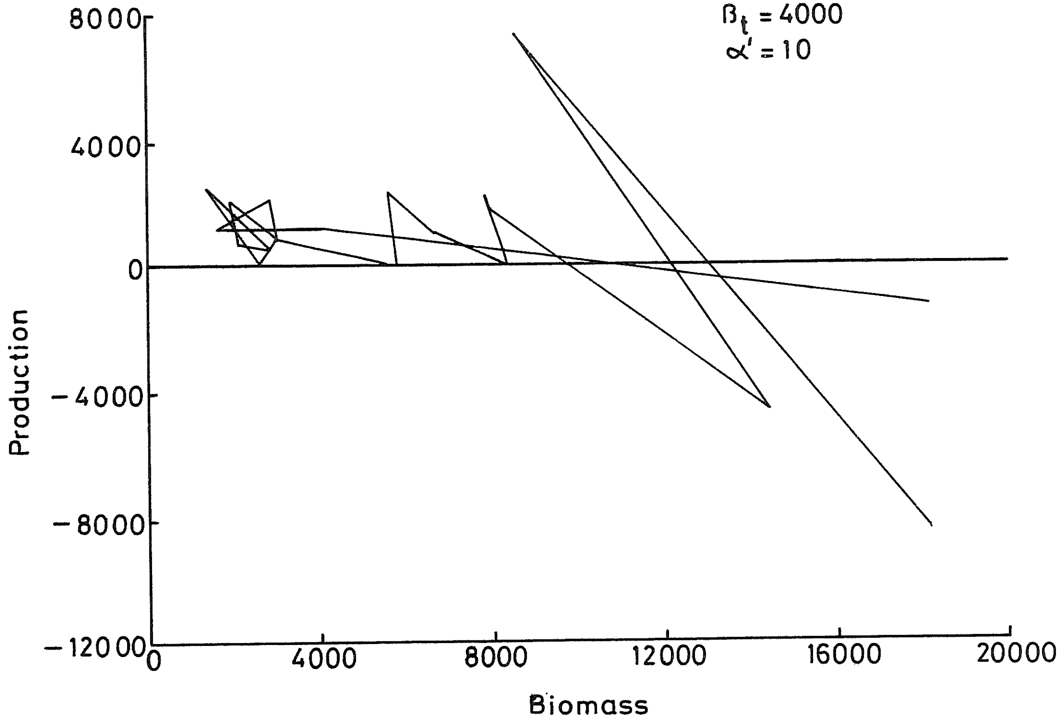


Figure 2.3.11 SIMULATED DATA PROCESS NOISE

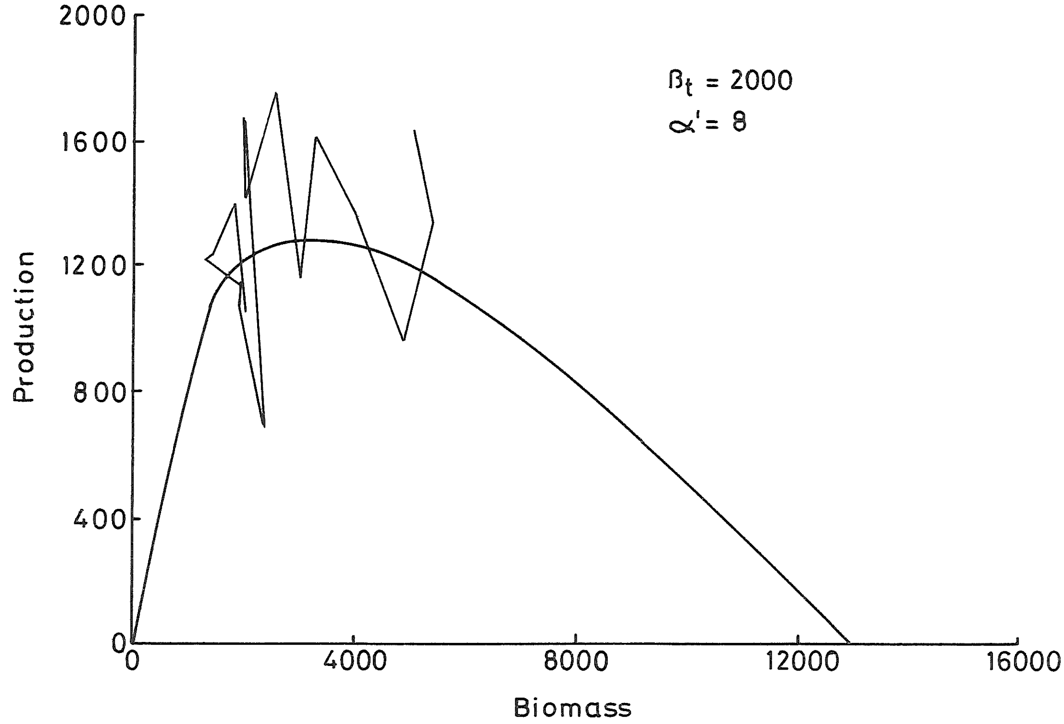


Figure 2.3.12 COD

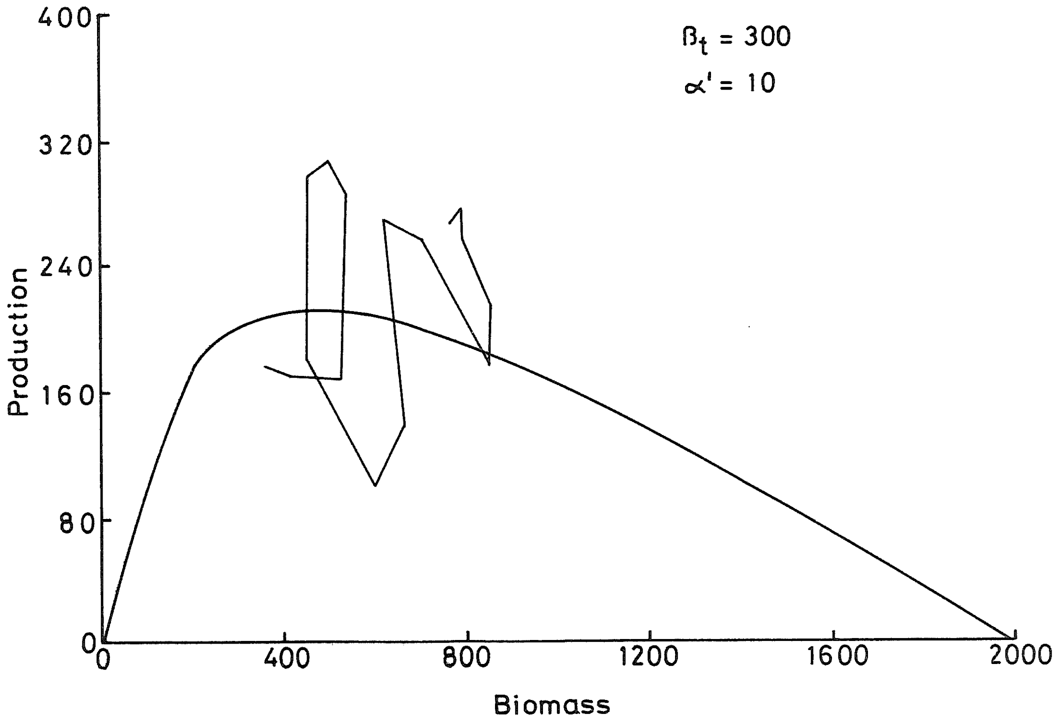


Figure 2.3.13 HALIBUT

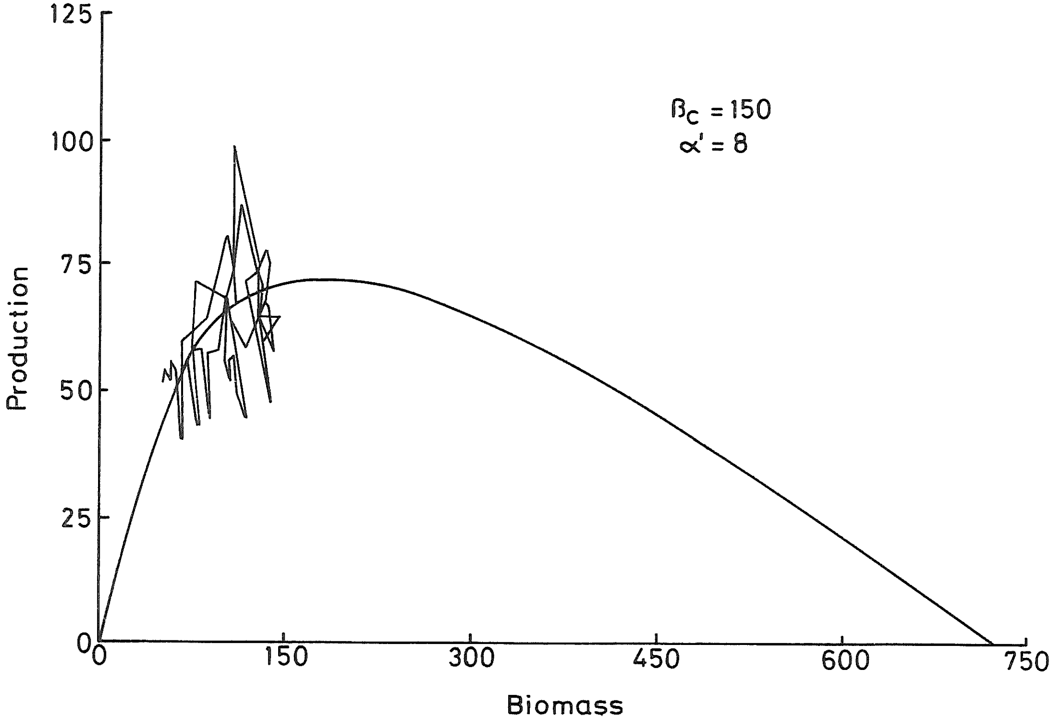


Figure 2.3.14 S. HORSE MACKEREL

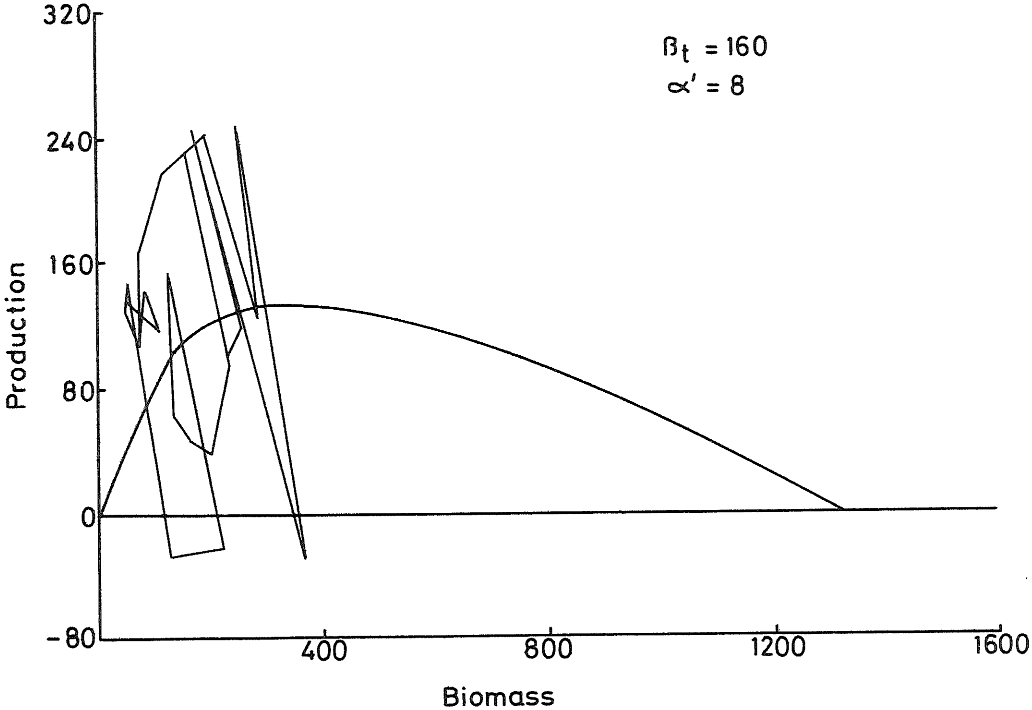


Figure 2.3.15 S. HORSE MACKEREL

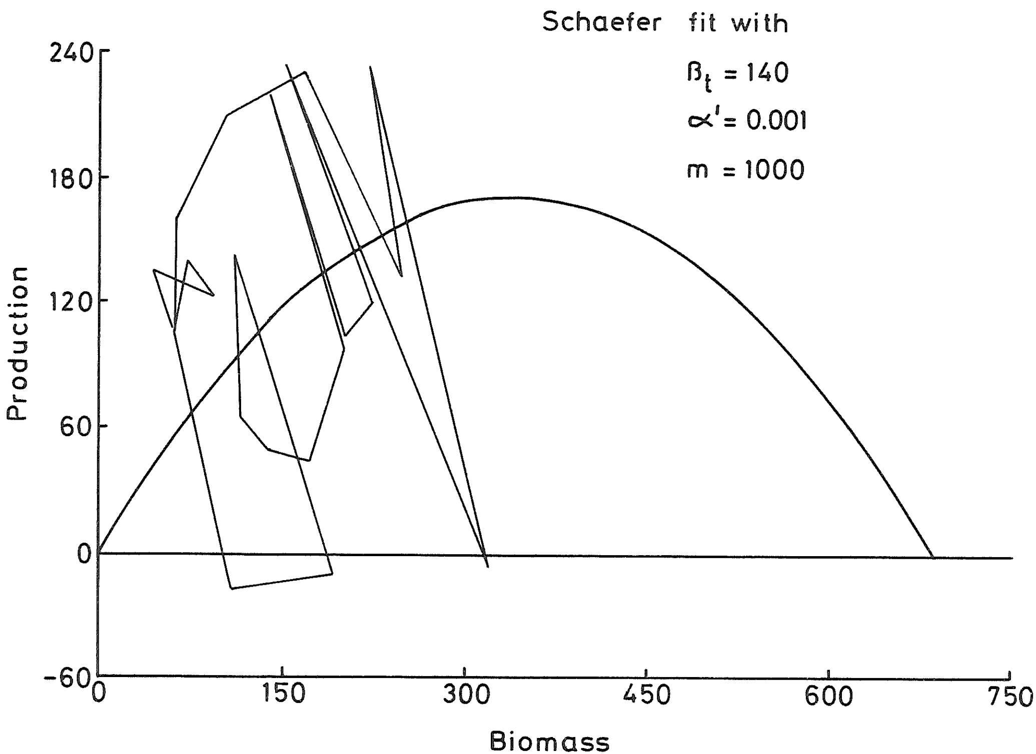
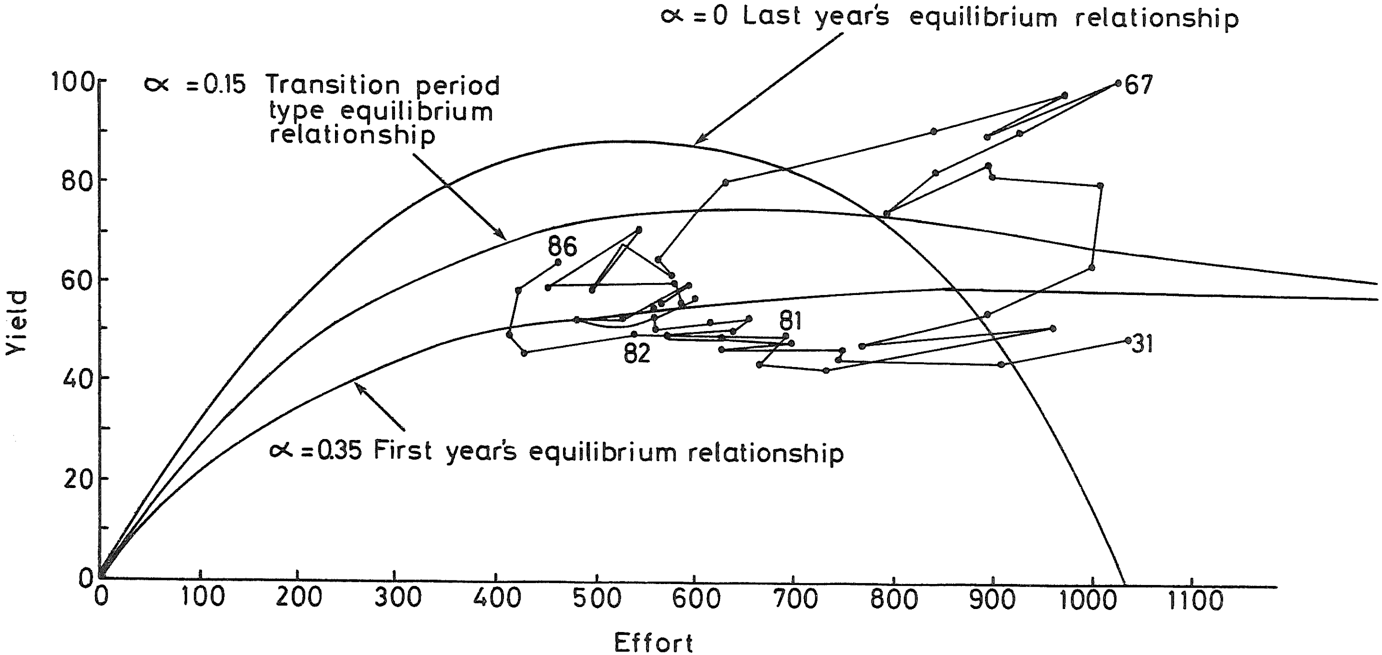


Figure 2.3.16



∝ = 0 Last year's equilibrium relationship
∝ = 0.15 Transition period type equilibrium relationship
∝ = 0.35 First year's equilibrium relationship
Yield
Effort
0
20
40
60
80
100
0
100
200
300
400
500
600
700
800
900
1000
1100
67
86
81
82
31

FIG 4.3.1 PACIFIC HALIBUT

RELATIVE FISHING MORTALITY (1974=100%)

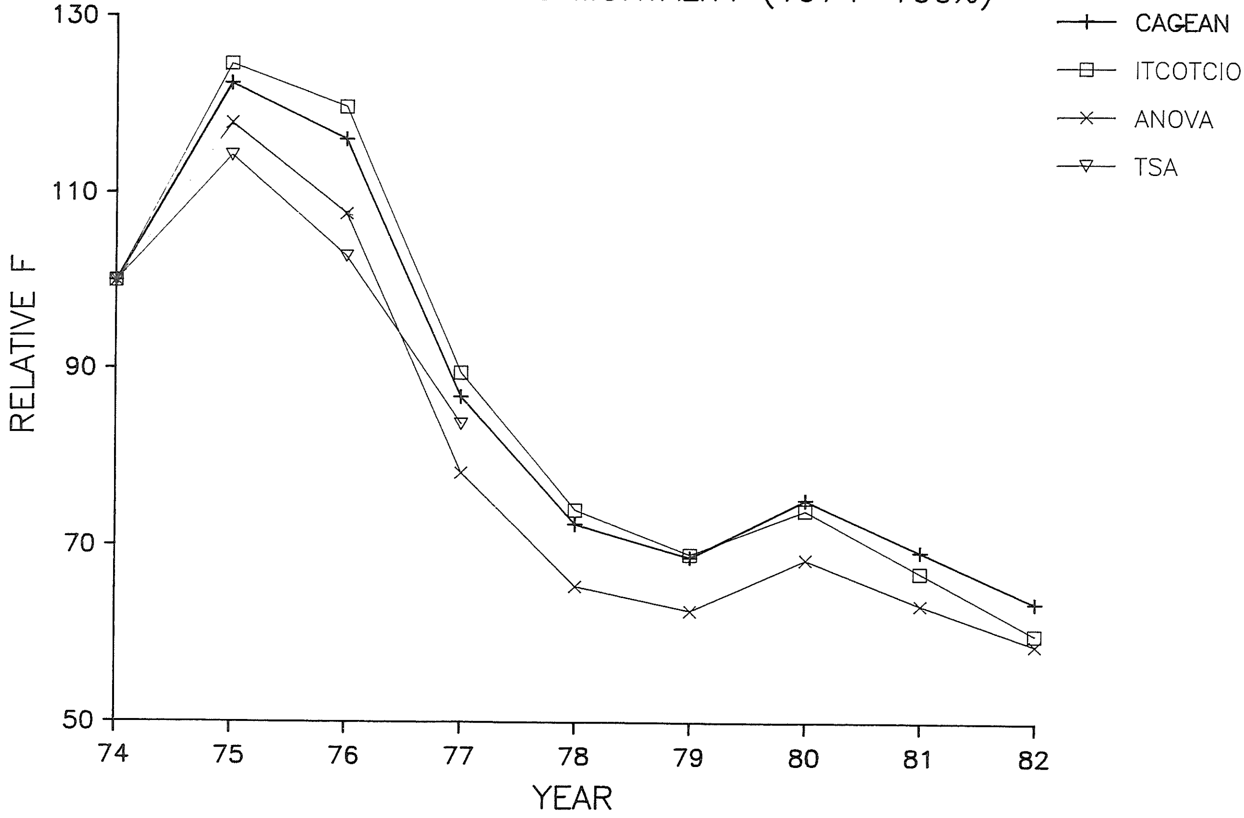


FIG 4.3.2 PACIFIC HALIBUT

RELATIVE YEAR CLASS STRENGTH AT AGE 8 IN YEAR (1974=100%)

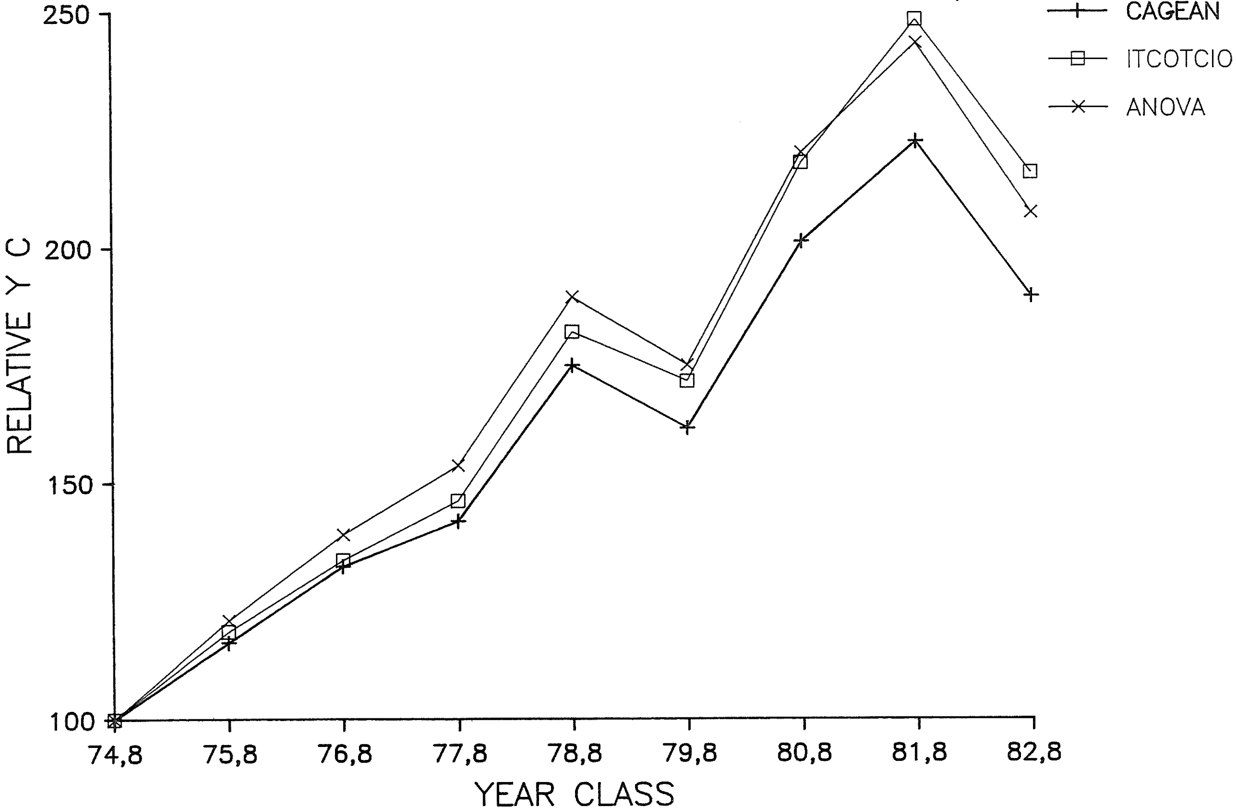


FIG 4.3.3 PACIFIC HALIBUT

RELATIVE EXPLOITATION PATTERN,AGE 14=100%

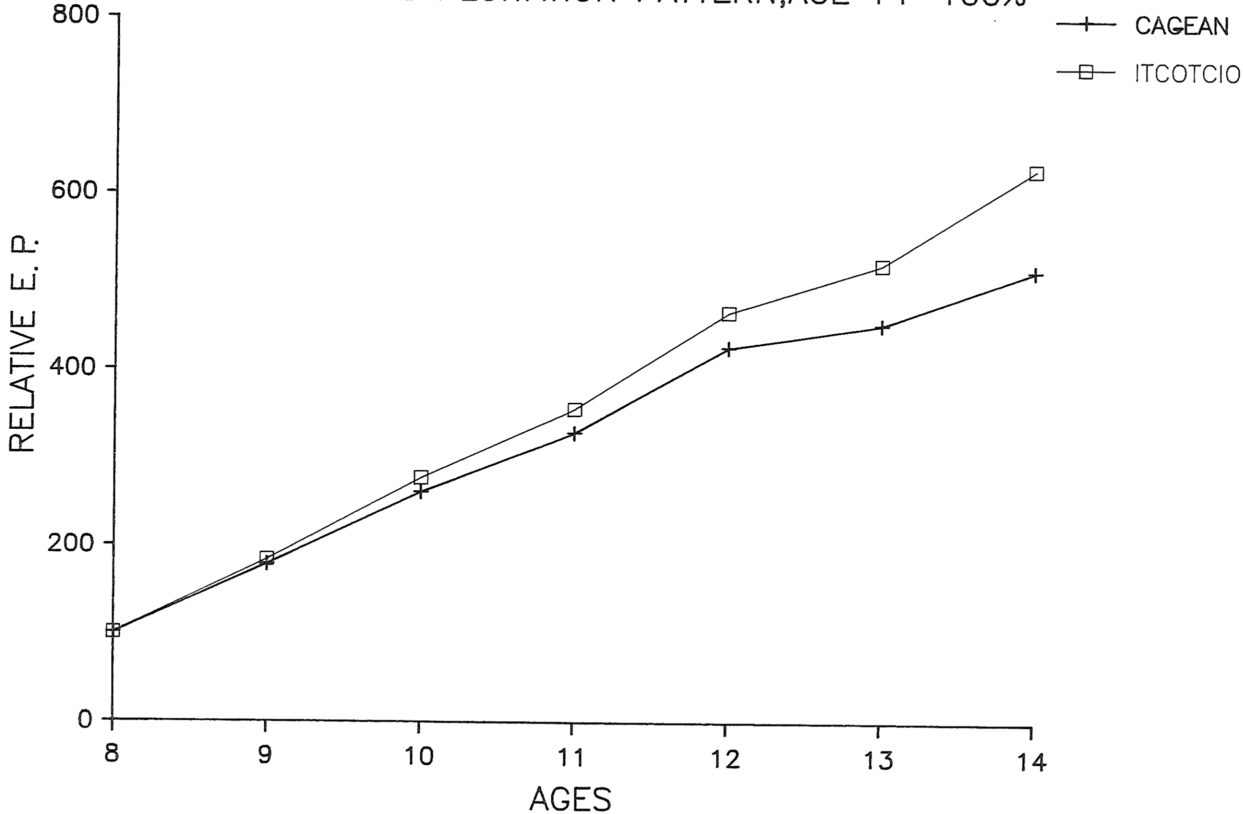
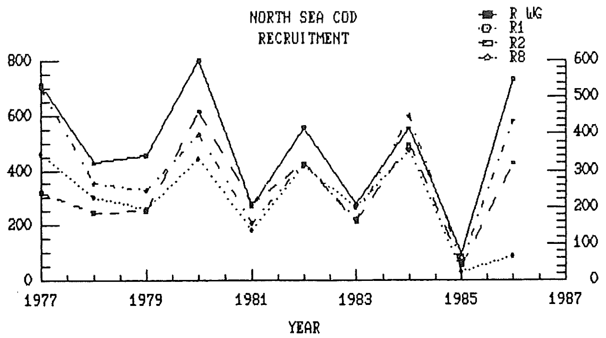
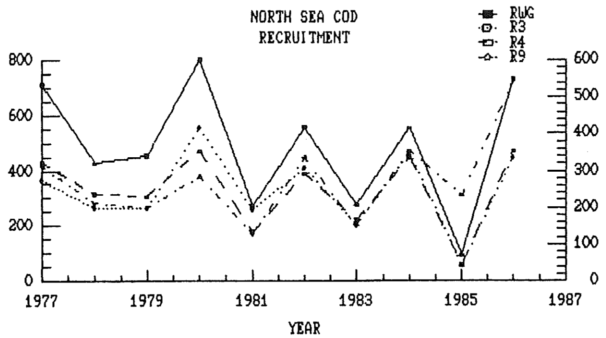


Figure 4.3.4


NORTH SEA COD
RECRUITMENT
R WG
R1
R2
R8
800
600
400
200
0
600
500
400
300
200
100
0
1977
1979
1981
1983
1985
1987
YEAR


NORTH SEA COD
RECRUITMENT
RWG
R3
R4
R9
800
600
400
200
0
600
500
400
300
200
100
0
1977
1979
1981
1983
1985
1987
YEAR

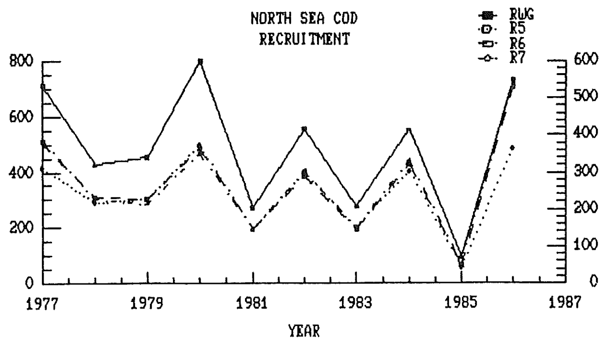
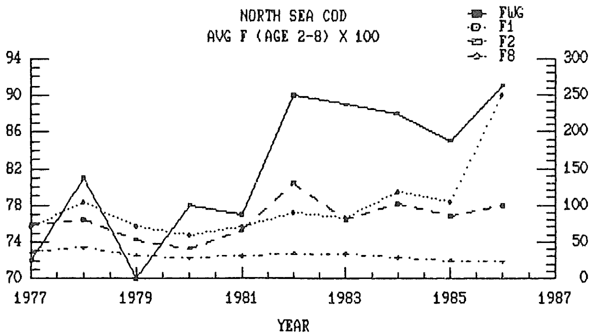
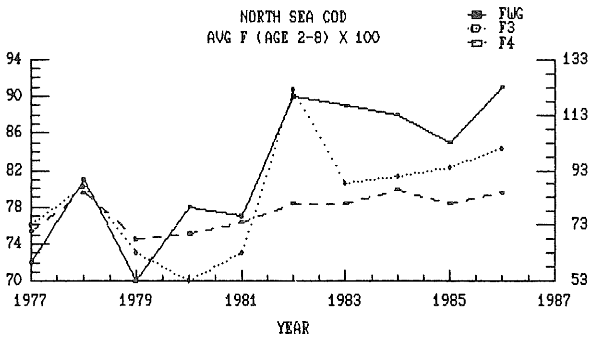

NORTH SEA COD
RECRUITMENT
RWG
R5
R6
R7
800
600
400
200
0
600
500
400
300
200
100
0
1977
1979
1981
1983
1985
1987
YEAR

Figure 4.3.5


NORTH SEA COD
AVG F (AGE 2-8) X 100
FWG
F1
F2
F8
94
90
86
82
78
74
70
300
250
200
150
100
50
0
1977
1979
1981
1983
1985
1987
YEAR


NORTH SEA COD
AVG F (AGE 2-8) X 100
FWG
F3
F4
94
90
86
82
78
74
70
133
113
93
73
53
1977
1979
1981
1983
1985
1987
YEAR

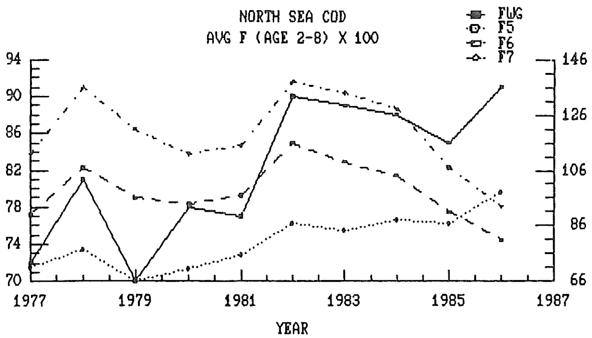
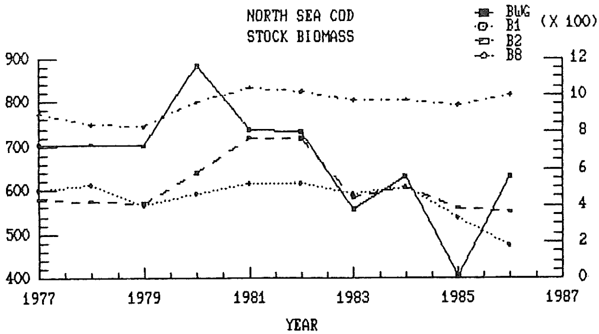
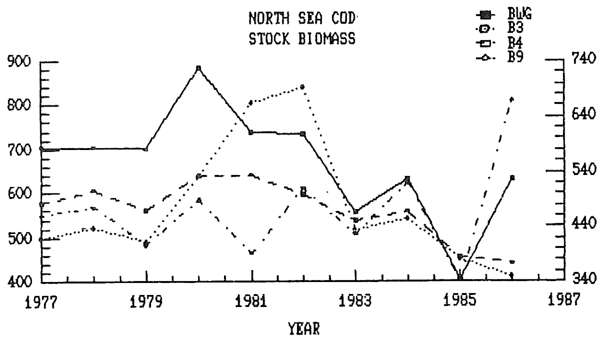
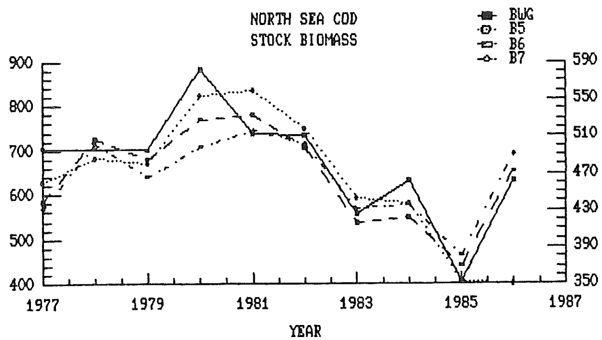

NORTH SEA COD
AVG F (AGE 2-8) X 100
FWG
F5
F6
F7
94
90
86
82
78
74
70
146
126
106
86
66
1977
1979
1981
1983
1985
1987
YEAR

Figure 4.3.6


NORTH SEA COD
STOCK BIOMASS
BWG
B1
B2
B8
(X 100)
900
800
700
600
500
400
12
10
8
6
4
2
0
1977
1979
1981
1983
1985
1987
YEAR


NORTH SEA COD
STOCK BIOMASS
BWG
B3
B4
B9
900
800
700
600
500
400
740
640
540
440
340
1977
1979
1981
1983
1985
1987
YEAR


NORTH SEA COD
STOCK BIOMASS
BWG
B5
B6
B7
900
800
700
600
500
400
590
550
510
470
430
390
350
1977
1979
1981
1983
1985
1987
YEAR

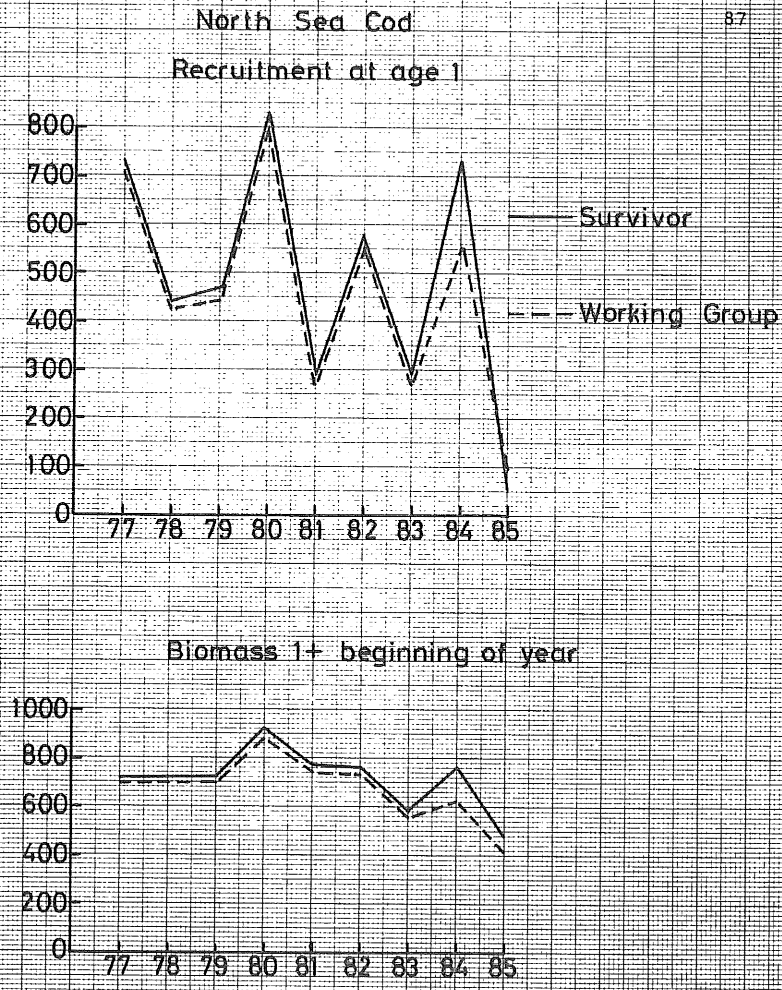

North Sea Cod
Recruitment at age 1
800
700
600
500
400
300
200
100
0
77
78
79
80
81
82
83
84
85
Survivor
Working Group
Biomass 1+ beginning of year
1000
800
600
400
200
0
77
78
79
80
81
82
83
84
85

Figure 4.3.7

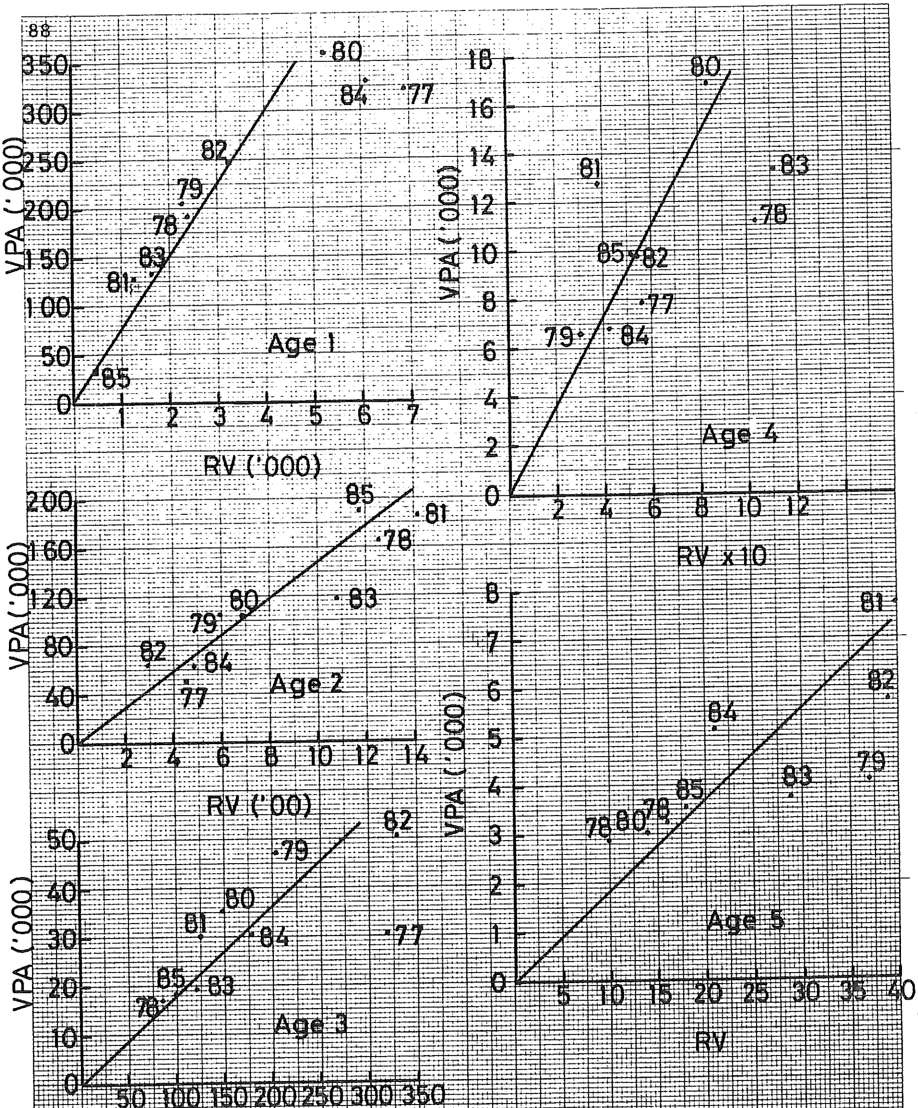


Figure 4.3.8 North Sea cod. Results of SURVIVOR analysis. Population numbers vs reseach vessel estimates.

Figure 5.3.1 Relative sensitivity coefficients of annual recruitment estimates to input parameters for North Sea sole.

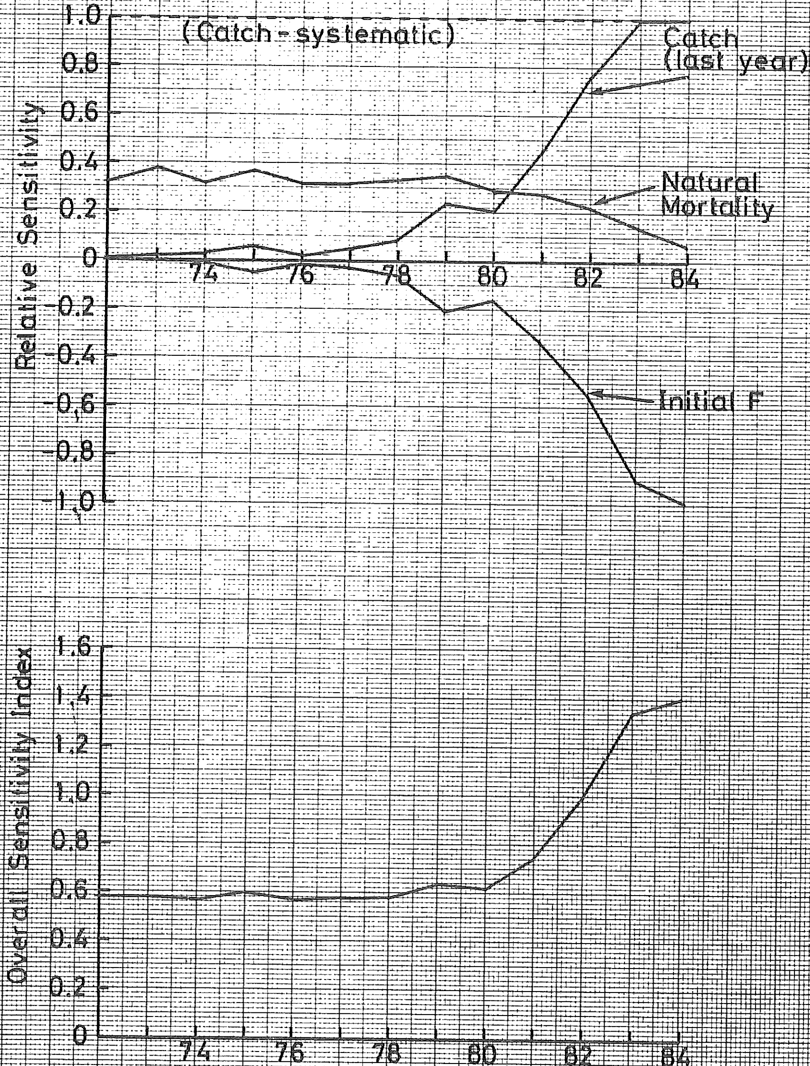
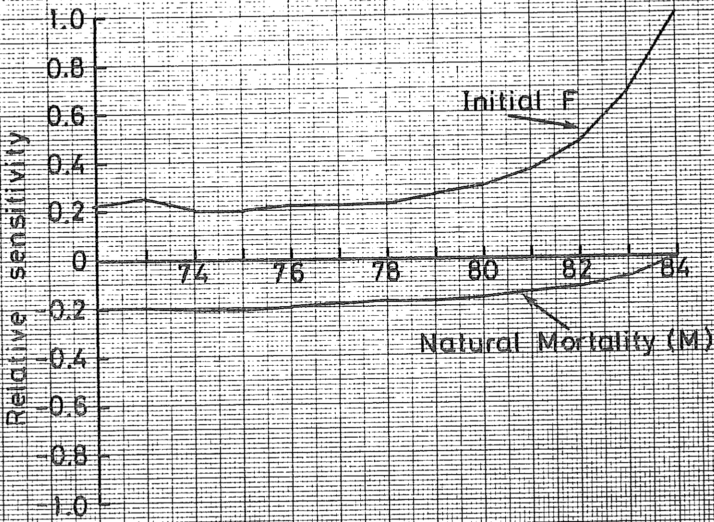


Figure 5.3.2 Relative sensitivity coefficients of fishing mortality (F) for all age groups to initial F and M: North Sea sole.



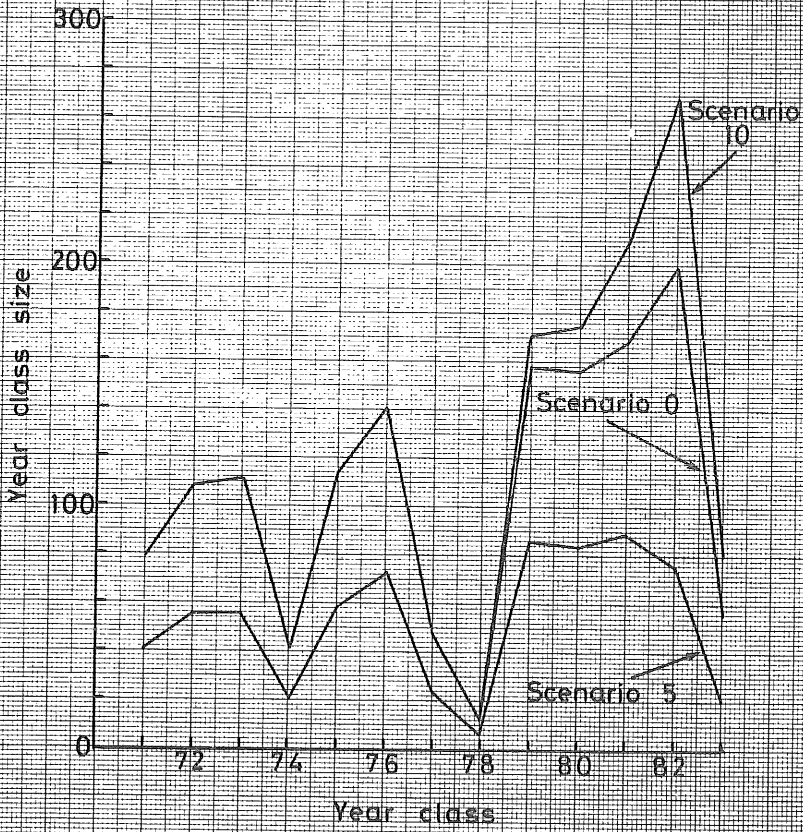


Figure 5.3.3 Effect of various sources of error on the estimation of annual recruitment (year class size) for the North Sea sole.

Scenario	Description
0	True catches
7	Increasing trend in misreporting over years and and ages, effort misreported proportionally
9	Effort misreported by 20% in last year

Figure 5.3.4

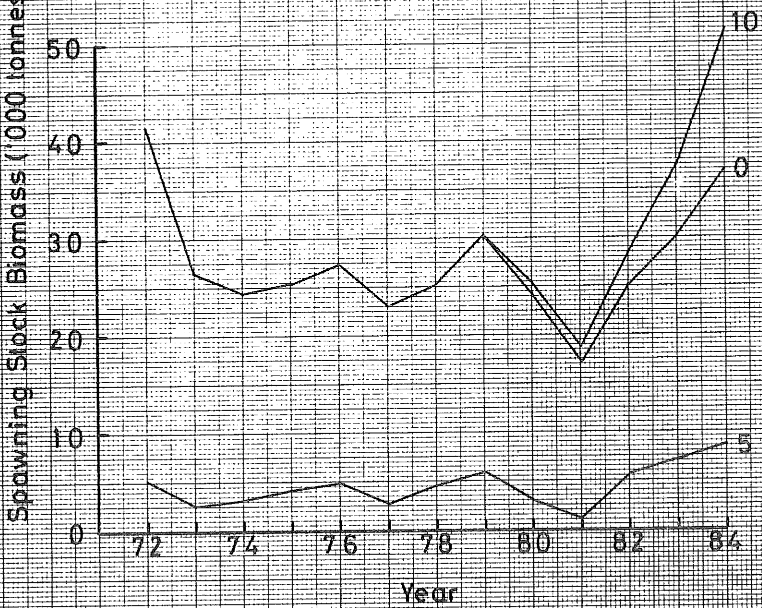
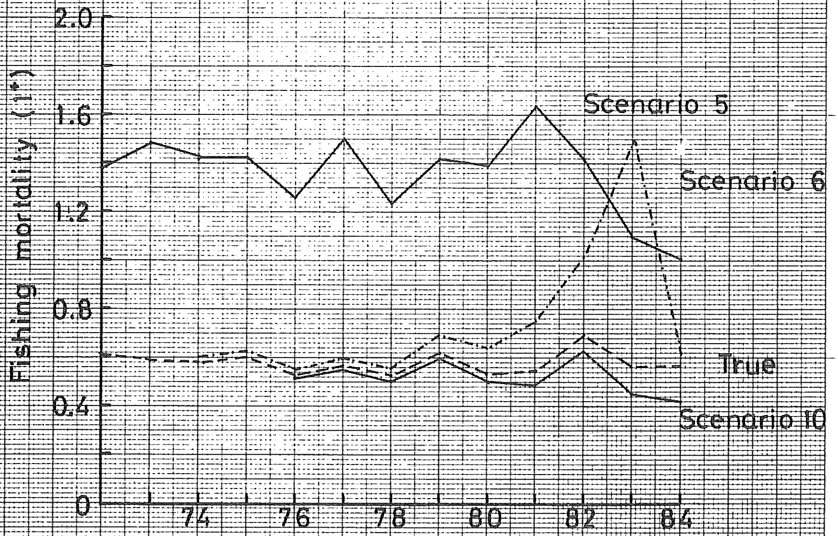
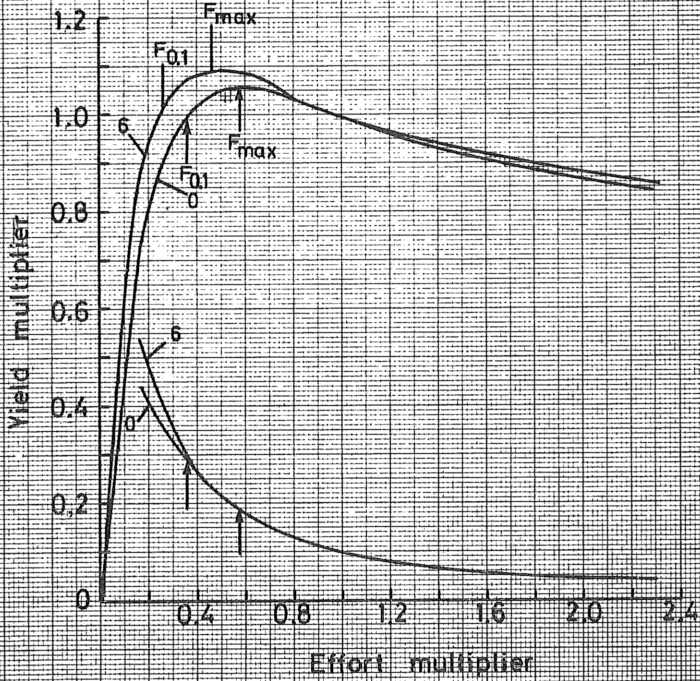

Spawning Stock Biomass ('000 tonnes)
50
40
30
20
10
0
72
74
76
78
80
82
84
Year
10
0
5

Figure 5.3.5 Effect of various sources of error on the estimation of weighted annual F (1+) for North Sea sole.



Footnote: Scenario 5: misreporting is age-selective for all years
Scenario 10: effort is misreported by 20% in the last year
Scenario 6: % age-selective misreporting in the last year of the analysis

Figure 5.3.6



Fmax
F0.1
6
Fmax
F0.1
0
6
0
1.2
1.0
0.8
0.6
0.4
0.2
0
0.4
0.8
1.2
1.6
2.0
2.4
Yield multiplier
Effort multiplier

APPENDIX A

WORKING PAPERS

1. "Contribution à l'étude du modèle global pour la dynamique des populations marines exploitées. Formulation, ajustement et sensibilité à certaines sources d'erreurs" by F. Laloe.

   A description of different fitting methods is presented.

   A discussion on the precision of parameters estimators is made and the shape of the confidence region (MSY-fMS) is presented in a case study.

   An approach using minimization of catchability variation is discussed.

   An introduction of environmental effects is also presented.

   A simulation with some errors in data and parameters is made.

2. "A simple production model with unaccessed quantity of biomass" by F. Laloe.

   A Schaefer model is presented in which it is assumed that an unmatchable proportion of the virgin biomass exists.

   This model leads to equilibrium catch-effort relationships which are analogous to those that can be obtained from a generalized (Pella and Tomlinson) model.

   Two examples are studied in which the unmatchable quantity of biomass has changed during the history of the fishery.

   This modelization may take into account change in stocks underlying dynamics during the history of the fishery.

3. "The use of multiplicative models for separable VPA, integrated analysis and the general VPA tuning problem" by J.G. Pope and T.K. Stokes.

   Describes the methods of integrated analysis of catch-at-age data and CPUE or effort data. The use of the CCIM model was a central theme as was the development of models which help to promote insight into the tuning problem.

4. "Understanding the structures of catch-at-age and effort data: the value of GLIM" by J.G. Pope and T.K. Stokes.

   Extended the work of the previous paper on ANOVA interpretations of multifleet separable data and/or linearized multifleet interpreted separable analysis. The important messages of this paper were:

a) Catch-at-age data often have a nearly linear structure.

b) Having data on a number of fleets does not alter the nature of the estimation.

c) Permitting catchability to vary freely on all fleets means effort data will fail to specify terminal F uniquely.

5. "Combination of recruitment indices using weighted averages" by J.G. Shepherd.

This Working Paper describes in more detail the method for combination of recruit indices using weighting averages which were briefly described in Anon. (1986a). This has now been implemented as a Fortran program (RCRIZNX) and has been used by some ICES working groups. The method uses a log-log calibration regression as recommended by the Working Group (Anon., 1984) and combines estimates in accordance with their estimated prediction errors. It gives very low weight to poorly correlated data sets in practice, and generally finds slopes less than 1 (VPA less extreme than index) even using the calibration method.

6. "Towards improved stock-production models" by J.G. Shepherd.

In Working Paper 6, Shepherd proposed a simple non-equilibrium stock-production model based on explicit representation of natural mortality and growth plus recruitment. The latter process is modelled with a functional form based on a Beverton-Holt stock-recruitment relationship. The model itself is, therefore, not novel, and is, indeed, one of the general class described by Schnute (1985). However, the fitting procedure proposed is novel, based on mapping goodness-of-fit over feasible ranges of two of the three parameters, giving a hopefully more robust and informative analysis. The same fitting procedure can be applied to other models, and has been implemented for the Shepherd and Schaefer models in a Fortran program SPM.

7. "Time series models of fishing mortality rates" by G. Gudmundsson.

Stochastic models of fishing mortality rates, based on concepts from time series analysis, are estimated from catch-at-age data. The rate of natural mortality is supposed to be known. These models can be estimated with tolerable accuracy from actual data for all years and ages without any further observations. Trends in fishing mortality rates and variations in the pattern of selectivity, gradual or irregular, can be detected.

The estimation is carried out by an approximation to the Kalman filter. Unknown parameters in the models are obtained from the likelihood function of catch prediction errors.

Extension of this estimation procedure to a joint analysis with data from research vessel surveys is straightforward, but entails a substantial increase in computation.

8. "Analysis of icelandic trawlers reports" by G. Stefansson.

A preliminary analysis of Icelandic trawler reports was presented along with a method for using the resulting CPUE indices for cod in an integrated analysis with catch-at-age data.

The necessity of proper stratification and age disaggregation was emphasized.

The data are recorded by the fishermen as weight by species in each tow. Further, towing time and location of the tow are recorded. It was, therefore, possible to compute CPUE indices separately for small squares and then average over squares within the region of interest. Within squares, the index was computed as an unweighted average of indices for each trawler, where a trawler's index was computed as the sum of its catches divided by total towing time. This method of index construction is intended to let all trawlers weight equally in the index for each square and to let all squares weight equally in the overall abundance index for the year.

The need for age disaggregation was particularly obvious in that the aggregated indices do not indicate any relationship with usual biomass measures, but fairly high correlations are obtained between disaggregated indices and VPA biomass for the age groups of primary interest in the study (ages 4-6). Therefore, age composition by weight for the region was used to decompose the annual CPUE index into indices for each age group.

One way of estimating terminal F values is to first assume a fixed selection pattern, then try a particular terminal F value as input to a VPA run. This will yield biomass at age for each age group. For a fixed age group, a regression of log (CPUE) on log(B) can be performed to yield an error sum of squares, SSE (F,a). These can then be summed over relevant age groups to yield one sum of squares, SSE(F). The method proceeds by estimating F as the number which minimizes SSE(F) over F.

This approach is particularly easy to use, since it only requires a VPA program and a simple linear regression program. It is also easily extended to include a time trend in catchability. Further, confidence intervals for terminal F are easily obtained based on an F-test on the SSE values (cf. Halldorsson et al., 1986).

For the Icelandic cod data, ages 4-6, the preliminary results indicate fairly wide confidence intervals for the terminal F values. This would seem to point to the necessity for more accurate CPUE data, including more age groups.

9. "Utilisation des IYFS pour estimer le recrutement-utilisation de la distribution a priori - prise en compte de fonctions de perte" by A. Laurec and A. Souplet.

Considering that the unknown recruitment is coming from the same distribution as the previous ones, it is possible to build a maximum likelihood estimation that will offer a compromise between the historical geometric mean of recruitment and the estimation suggested by the usual calibration. When the recruitments are considered as corresponding to a log normal distribution, with no correlation from year to year, it is just equivalent to using the regression line, where VPA is predicted from survey indices.

In a second part, this paper discusses the possible use of non-quadratic loss functions, which would make it possible to take into account that underestimating recruitment may be more or less important that overestimating it, that the same level of estimation errors may be more important when the real recruitment is low. This paper will be developed and presented to the 1987 ICES Statutory Meeting.

10. "Multiplicative modelling of recruitment estimates" by R.M. Cook.

The problem of combining multiple indices of abundance from research vessel surveys to obtain a single "best" estimate is addressed via a multiplicative model. The model embodies a fleet effect and a year effect and allows for log-linear relationships between year effect and index. Historically, distant data are down-weighted using a tri-cubic function and data from different surveys are weighted by the inverse of a residual variance associated with each survey.

Trials of the model on simulated and real data are presented.

11. "Joint analysis of catch at age and CPU observations" by G. Gudmundsson.

Extends the models and estimation procedures of Working Paper 7 to also include observations of recruitment, groundfish surveys, or CPUE data from commercial fleets.

APPENDIX B

STANDARD NOTATION

<u>NOTE</u>: This standard (and largely mnemonic) notation is followed so far 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 group
t	"	last (terminal) year
g	"	oldest (greatest) age group
l	"	length
k	"	year class
$	"	summation over all possible values of index (usually fleets)
#	"	summation over all 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 number (including discards)
E (y,f)	Fishing effort
F (y,f,a)	Fishing mortality
F_s(y,f)	Separable estimate of overall fishing mortality
q	Catchability coefficient (in F = qE)
Y	Yield in weight
W	Weight of an individual fish in the catch
W_S	Weight of an individual fish in the (spawning) stock
B	Biomass
P	Population number (also fishing power)
E	Fishing effort
U	Yield or landings per unit of effort
C_W	Catch in weight of fish (including discards)
N	Stock in numbers of fish
F	Instantaneous fishing mortality rate
Z	Instantaneous total mortality rate
M	Instantaneous natural 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 number (excludes discards)
l	Length
l_∞	Von Bertalanffy asymptotic length
K	Von Bertalanffy "growth rate"
r	Recruit index
MSY	Maximum sustainable yield
F_{msy}	Fishing mortality rate associated with MSY
E_{msy}	Fishing effort associated with MSY
B_{max}	Pristine stock biomass
m	Shape parameter for various surplus production models

APPENDIX C

SUMMARY OF TOPICS

Topic	1981[1]	1983	1984	1985	1987	1988[2]	1989[3]
1. Application of separable VPA	-	M	r	-	-	M	m
2 Simpler methods of assessment	-	-	M	M	i	-	-
3. Measures of overall fishing mortality	-	-	-	-	-	-	-
4. Use of effort data in assessments	M	M	r	r	M	M	m
5. Need for two-sex assessments	-	-	-	-	-	-	-
6. Computation and use of yield per recruit	-	M	m	i	-	-	-
7. Inclusion of discards in assessments	-	-	-	M	-	-	-
8. Methods for estimation of recruitment	-	-	M	r	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	-
14. Use of less reliable fishery statistics	-	-	-	-	m	-	m
15. Construction of indices from disaggregated data	-	-	-	-	-	-	M

M = Major topic.
m = minor topic.
r = reprise.
i = incidentally considered.

[1] Meeting of ICES ad hoc Working Group on Use of Effort Data in Assessments.
[2] Special workshop suggested during this meeting.
[3] Next Methods Working Group meeting.

APPENDIX D

SOFTWARE ROUTINES AND PACKAGES USED BY THE WORKING GROUP

Name	Language	Hardware	Usage	Further info.
1. ANOVA	Glim	Wide implementations	Analysis of variance	J.G. Pope[1]/NAG
2. CALIB	Fortran 77	NORD	Calibration of recruitment using several survey indices	A. Laurec[2]
3. CAGEAN	Fortran 77	Cyper 7600 Prime 550 Burroughs NORD IBM-PC	Catch-age analysis with auxiliary information	Deriso et al. (1985)
4. ITCOTCIO	Glim Macro	Wide implementations	Catch-age analysis	J.G. Pope and T.K. Stokes (WP 4)
5. Non-linear fitting	Fortran Genstat	NORD	Fitting production models with Marquadt algorithm	F. Laloe[3]
6. PROD	APL	Mainframe	Production models: Transient forms	S. Gavaris[4]
7. PRODAFIT	APL	IBM-OC	Production models: Equilibrium approxim ations	Rivard (1982)
8. RCRTINY	Fortran 77	MSDOS NORD HP, etc.	Weighted average combination of recruitment indices	J.G. Shepherd[5]
9. SCHAEFER	APL	IBM-PC	Production models: Equilibrium approximations	Rivard (1982)
10. SCHNUTE	APL	IBM-PC	Production models: using Schnute (1977) linear form	R.K. Mohn[6]
11. SPM	Fortran 77	MSDOS NORD HP, etc.	Shepherd + Schaefer stock production models	J.G. Shepherd[5]
12.	Fortran 77	NORD	Analysis of residuals	Gudmundsson (1986)

Name	Language	Hardware	Usage	Further info.
13. TSM	Fortran 77	VAX	Catch-age analysis with or without auxiliary information	G. Gudmundsson (WP 7)
14. PRODFIT	Fortran 77	MSDOS	General production model fitting through approximated equilibriums	Fox (1975)
15. GENPROD	Fortran 77	MSDOS	General production model under transient situations	Pella and Tomlinson (1969)

[1] J.G. Pope, Fisheries Laboratory, Lowestoft, Suffolk NR33 OHT, UK.

[2] A. Laurec, IFREMER, rue de l'Ile d'Yeu, BP 1049, 44037 Nantes Cédex, France.

[3] F. Laloe, C.R.O.D.T., BP 2241, Dakar, Senegal.

[4] S. Gavaris, St. Andrews Biological Station, Dept. of Fisheries & Oceans, St. Andrews, N.B. E1A 3EO, Canada.

[5] J.G. Shepherd, Fisheries Laboratory, Lowestoft, Suffolk, NR33 OHT, UK.

[6] R.K. Mohn, Dept. of Fisheries & Oceans, P.O. Box 550, Halifax, N.S. B3J 2S7, Canada.

APPENDIX E

MULTICALIBRATION THROUGH MAXIMUM LIKELIHOOD

Notation/Assumptions

A data set covering Ny past years and Nf fleets will be considered. The logarithm of the abundance index for year y and fleet f is $u_{y,f}$. When this datum is available, the Kronecker symbol $\delta_{y,f}$ is equal to 1; otherwise $\delta_{y,f} = 0$.

The past recruitment for year y is Ry = exp(xy). This assumed to be known exactly (from VPA).

In addition to the set of past data, estimation of the recruitment for the current year will be based on the current abundance indices $u_{o,f}$. The same convention applies to the Kronecker symbols $\delta_{o,f}$.

The sum

$$\delta_{o,f} + \sum_{y=1}^{Ny} \delta_{y,f}$$

is denoted T_f. This is the total number of data points in the time series for fleet f.

Log-linear relationships will be assumed so that

$$u_{y,f} = a_f\, x_y + b_f + \varepsilon_{y,f}$$

so that curvature of the abundance/index relationships is permitted.

The residuals $\varepsilon_{y,f}$ are assumed to come from a normal distribution with zero mean and variance σ_f^2. They are assumed to be independent from year to year and fleet to fleet.

It will also be assumed in one case that the log recruitments themselves are drawn from a normal distribution, with a mean equal to $x_{\#}$ and a variance σ_x^2.

Log-likelihood functions

The basic multicalibration problem can be expressed in terms of the log-likelihood functions:

$$L = - \sum_f T_f \,\mathrm{Log}(\sigma_f) - \frac{1}{2\sigma_f^2} \sum_f \left[\sum_y (u_{y,f} - af x_y - b_f)^2 \delta_{y,f} + (u_{o,f} - af x_o - b_f)^2 \delta_{o,f} \right]$$

Maximizing this function is equivalent to minimizing some function

$$\varphi\ [(\sigma_f,\ a_f,\ b_f),\ x_o]$$

Differentiating L with respect to σ_f, then putting these derivatives to zero leads to the equation

$$(1) \qquad \sigma_f^2 = \frac{1}{T_f} \sum_y (u_{y,f} - a_f x_y - b_f)^2 \delta_{y,f} + (u_{o,f} a_f x_o - b_f)^2 \delta_{o,f}$$

This will lead to the concentrated likelihood function (Bard, 1974) by substituting σ as given by equation (1) in the likelihood function $- \varphi(fx_o)$ if

$$(2) \qquad \varphi = \sum_f T_f \, \mathrm{Log}(\sigma_f)$$

For a given x_o, the conditional maximum likelihood estimation will lead, as can be easily verified, to the usual empirical regression coefficients calculated for each fleet over the available couples $(u_{y,f}, x_f)$ and, if available, the final set $(u_{o,f}, x_o)$. This regression line relates to u as explained by x.

It is thus very easy for each value of x_o to calculate the conditional maximum likelihood estimations for the parameter a_f and b_f, and the corresponding maximum likelihood estimates for σ_f^2 through equation (1). From this, one deduces φ, which can be written as a function of x_o, which can easily be maximized by an iterative procedure.

The function φ (x_o) deserves careful consideration. The factor T_f leads to a weighting that increases the influence of long time series. On the other hand, the σ_f^2 deduced from equation (1) is biassed, as usual in maximum likelihood techniques. This bias may be considerable when T_f is not large compared to 2.

This basic procedure can be extended to include the previously mentioned hypothesis on the distribution of the recruitments. This will in fact add a term to the log-likelihood function, equal to:

$$- (Ny+1) \, \mathrm{Log}(\sigma_x) - \frac{1}{2\sigma_x^2} \left[\sum (x_y - x_\#)^2 + (x_o - x_\#)^2 \right]$$

The same concentration of the likelihood function will be possible since differentiating with respect to σ_x will lead to:

$$(3) \qquad \sigma_x^2 = \frac{1}{(Ny+1)} \left[\sum_y (x_y - x_\#)^2 + (x_o - x_\#)^2 \right]$$

The equivalent of the function φ will become:

$$(Ny+2) \, \mathrm{Log}(\sigma_x) + \sum T_f \, \mathrm{Log}(\sigma_f)$$

For each given value of x_o, the T_f will be calculated as previously mentioned, while $x_\#$ will be given by

$$x_\# = \frac{1}{(Ny+1)} \left(\sum_y x_y + x_o \right)$$

and σ_x^2 by equation (3).

Finally, it could be verified that these calculations can be easily adopted to situations where the slopes a_f are forced to 1, or to any weighting scheme assuming that the variance of $\varepsilon_{y,f}$ is $(w_{y,f})\sigma_f^2$ where $w_{y,f}$ is known.

APPENDIX F

PROPOSALS FOR A WORKSHOP

1. Purpose

Such a workshop should be strictly devoted to the practical application of selected existing methods, performing statistical integrated analysis of catch-at-age and auxiliary information.

It should arrive at some firm conclusions and recommend standard software that should be implemented in ICES as soon as possible after the workshop to become part of the standard assessment package. The Secretariat will be requested to make available the services of its staff and to invite an expert user to assist in the implementation of this software.

2. Time and location

It should take place in the second quarter of 1988 in a place where computer facilities are sufficient and correspond to standard procedures. Facilities should include the service of the necessary staff.

3. Participation

It should include members of the Methods Working Group, specialists of the stocks corresponding to the actual chosen data sets, and members of assessment working groups.

4. Software

The methods to be considered should strictly be selected by the Chairman of this workshop in consultation with the Chairman of the Methods Working Group and the Chairman of ACFM.

Since these methods will be existing ones, the corresponding software should be fully operational before the beginning of the meeting on the computers to be used by the workshop. User guides should systematically be available.

5. Data Sets

The following procedure is suggested to produce the data sets:

- Aberdeen selects two sets of actual data.

- Lowestoft, Reykjavik, and Seattle each produce a set of simulated data and a description of their properties.

- All these data sets are sent to the Chairman who determines whether they cover all aspects which ought to be considered and recommends to the authors changes that are needed to achieve this. After thus vetting the proposed sets, the Chairman distributes them to members of the Group, including only such prior information which the practitioners ought to have (such as natural mortality). This should be finished 6 months before the meeting, giving people ample time to carry out the analysis on their own machines before the meeting.