ICES COOPERATIVE RESEARCH REPORT

RAPPORT DES RECHERCHES COLLECTIVES

NO. 191

Recommended format for purposes of citation:
ICES 1993. Reports (3) of the working group on "Methods of fish stock assessment". Report No. 191. pp. 249.
https://doi.org/10.17895/ices.pub.4607

REPORTS OF THE WORKING GROUP ON METHODS OF FISH STOCK ASSESSMENTS

Copenhagen, 9-16 June 1987
Reykjavik, 6-12 July 1988
Nantes, 10-17 November 1989

ISSN 2707-7144
ISBN 978-87-7482-531-9

International Council for the Exploration of the Sea
Palægade 2-4, DK-1261 Copenhagen K
DENMARK

January 1993

THE METHODS
GEYSIR

CONTENTS

REPORT OF THE WORKING GROUP ON METHODS OF FISH STOCK ASSESSMENTS

Copenhagen, 9-16 June 1987

REPORT OF THE WORKSHOP ON METHODS OF FISH STOCK ASSESSMENTS

Reykjavik, 6-12 July 1988

REPORT OF THE WORKING GROUP ON METHODS OF FISH STOCK ASSESSMENTS

Nantes, France, 10-17 November 1989

REPORT OF THE WORKING GROUP ON METHODS OF FISH STOCK ASSESSMENTS

Copenhagen, 9-16 June 1987

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 9-16June 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 auxiliary information based on biological studies. This point is important because it appears that the models are somewhat under-determined when only catch and effort data are used for estimation. The delay-differential models proposed by Walter (1973) and expanded by Marchesseault *et al.* (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. Auxiliary 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 φ (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:

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

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:

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

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 behaviour 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 parameters, 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 explicitly 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) is 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) present 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 used 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 unacquainted 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 (though 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 Figures 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 Attempts to fit halibut (1932-1986) catch/effort data with a model with uncatchable quantities of biomass

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:

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

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

where i being the year from 1932 (i = 1) to 1986 (i = 55). α 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 sub-routine EO4FDF 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

$$SC=\sum_{i=1}^{55} [(P_i-P)/P]^2$$

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 difference 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 decrease 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.

2.3.7 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 auxiliary 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 auxiliary 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 occurred. 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 analyze 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 occurrence 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.

<u>Point 1</u>

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 committed 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 through 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 worthwhile 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 age-age-year-class factors, where the age factors and year-class factors were common to both data sets and where 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.

Correlations 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 multi-fleet 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	λ1	λ3	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 λ3 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 undersampling, 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 Results of Case Studies

5.3.1 Sensitivity analysis

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 include 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 pre-empting 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 <u>or survey</u>. 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 <u>not</u> 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 fitting 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 correspondence, 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. **In** 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 cohort 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. Inst. Français 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.

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

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.

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 $\times 10^{-2}$	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 $\times 10^{-2}$	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 $x10^{-3}$	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.	239.	.000632	.155	-9.
1965	969.	98.	104.	211.	.000654	.137	-6.
1966	889.	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.	.000738	.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									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	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								North Sea Roundfish WG 1987
	1	2	3	4	5	6	7	8	
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).

Scenario	$F_{0.1}$		F_{max}		F_{last}		$\frac{dY}{dE} \times \frac{E}{Y}$
	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

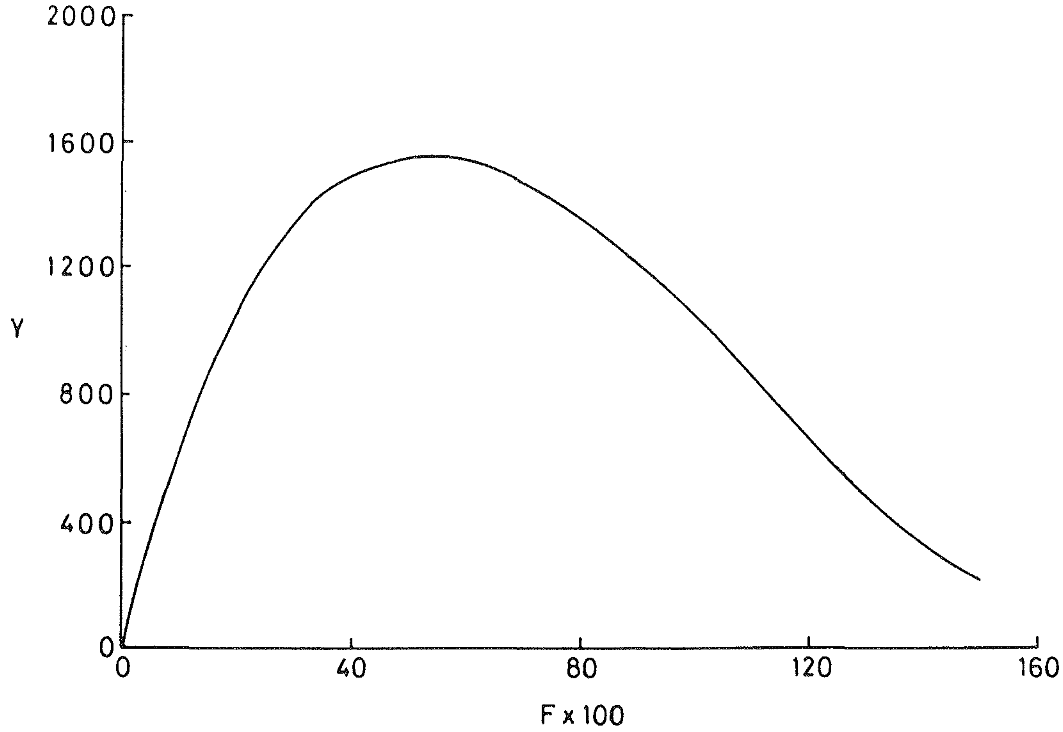


Figure 2.3.2 Equilibrium Stock-Recruit-Simulated data

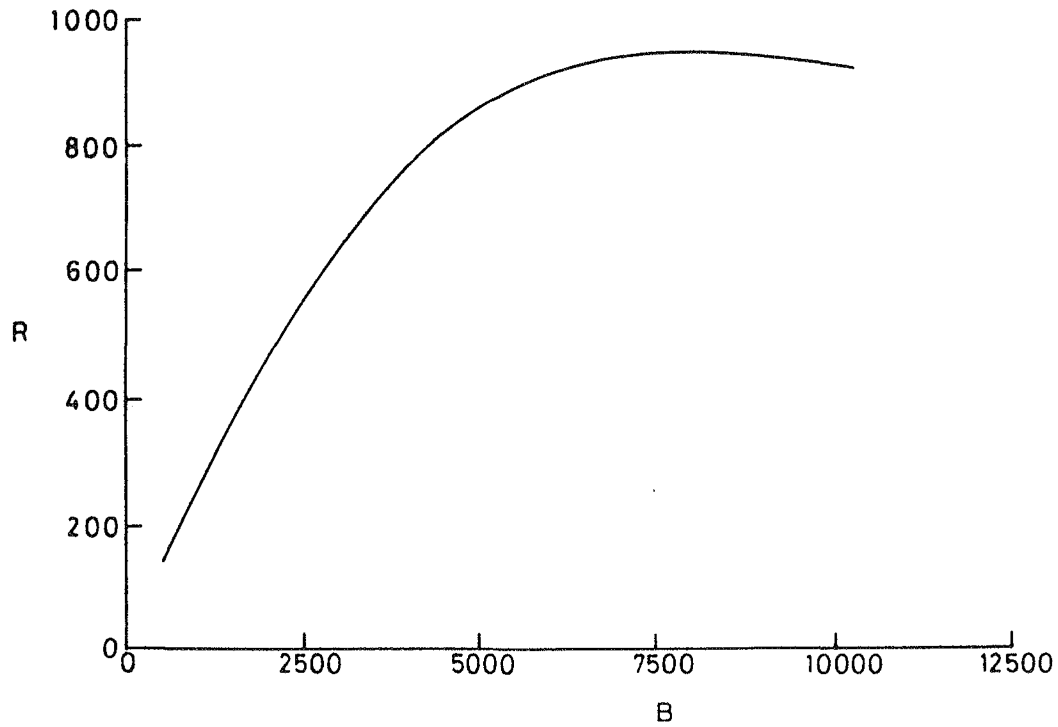


Figure 2.3.3 SIMULATED DATA — NO NOISE

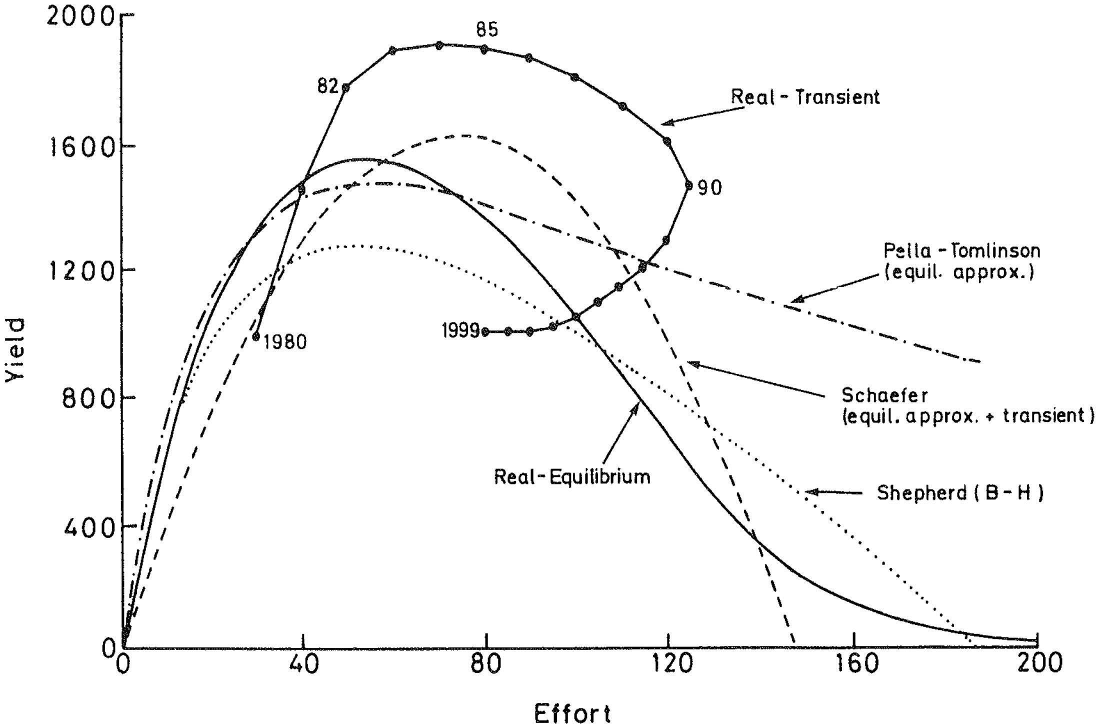


Figure 2.3.4 SIMULATED DATA — MEASUREMENT NOISE

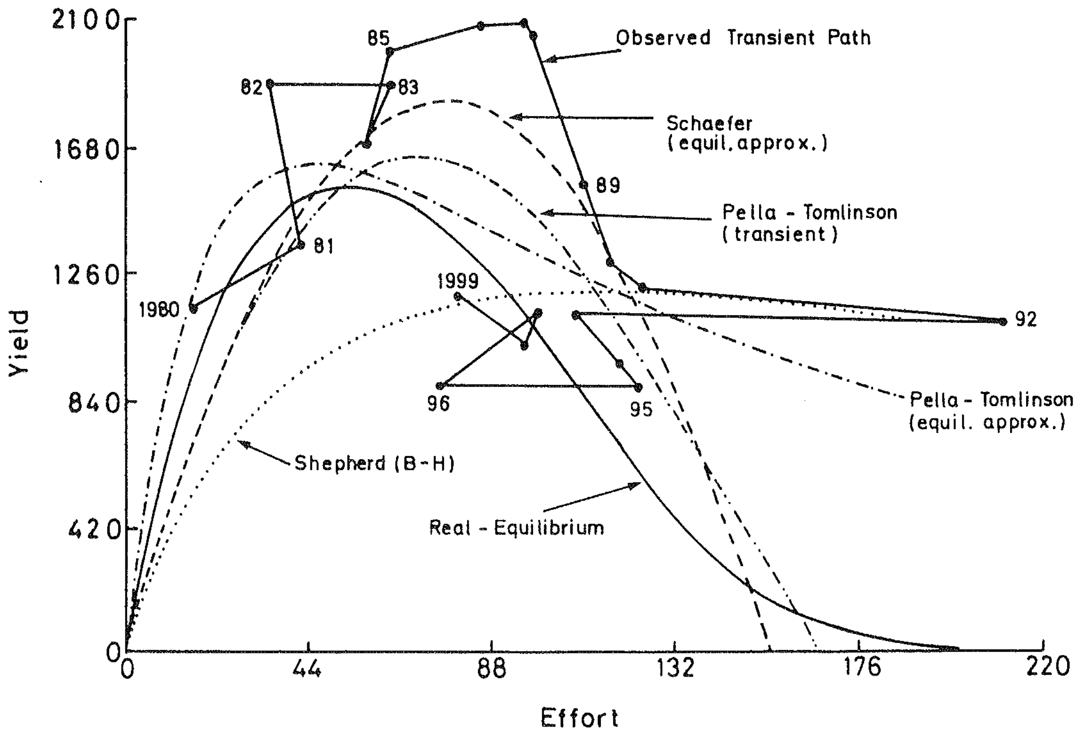


Figure 2.3.5 SIMULATED DATA — PROCESS ERROR

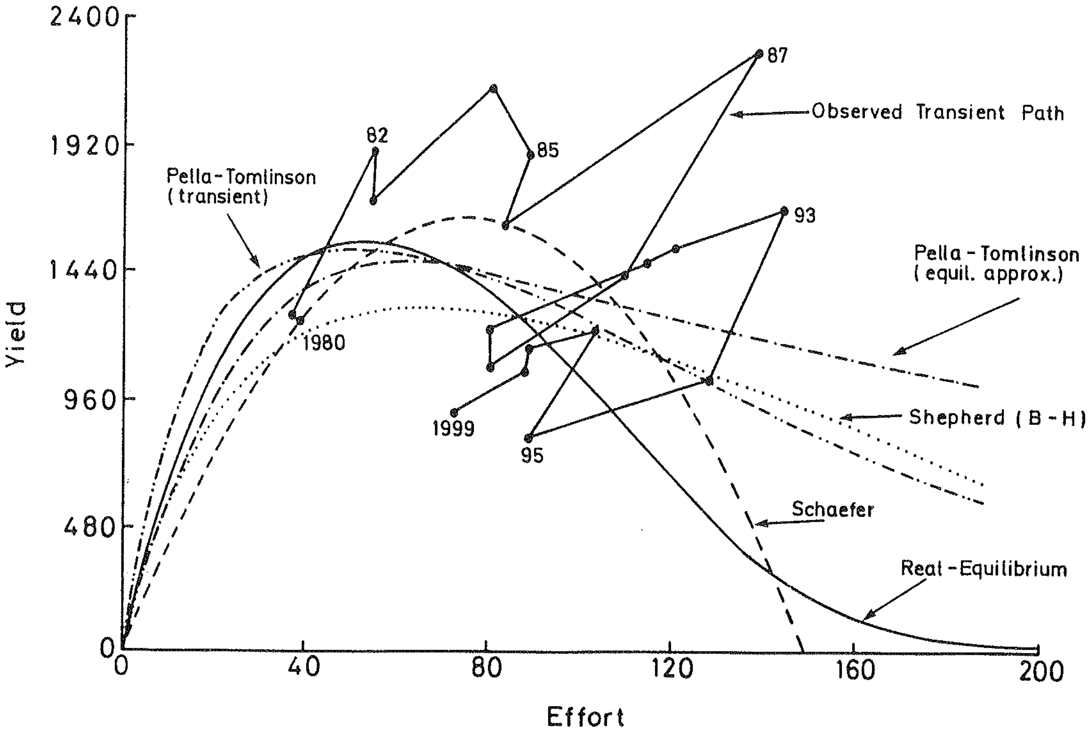


Figure 2.3.6 NORTH SEA COD

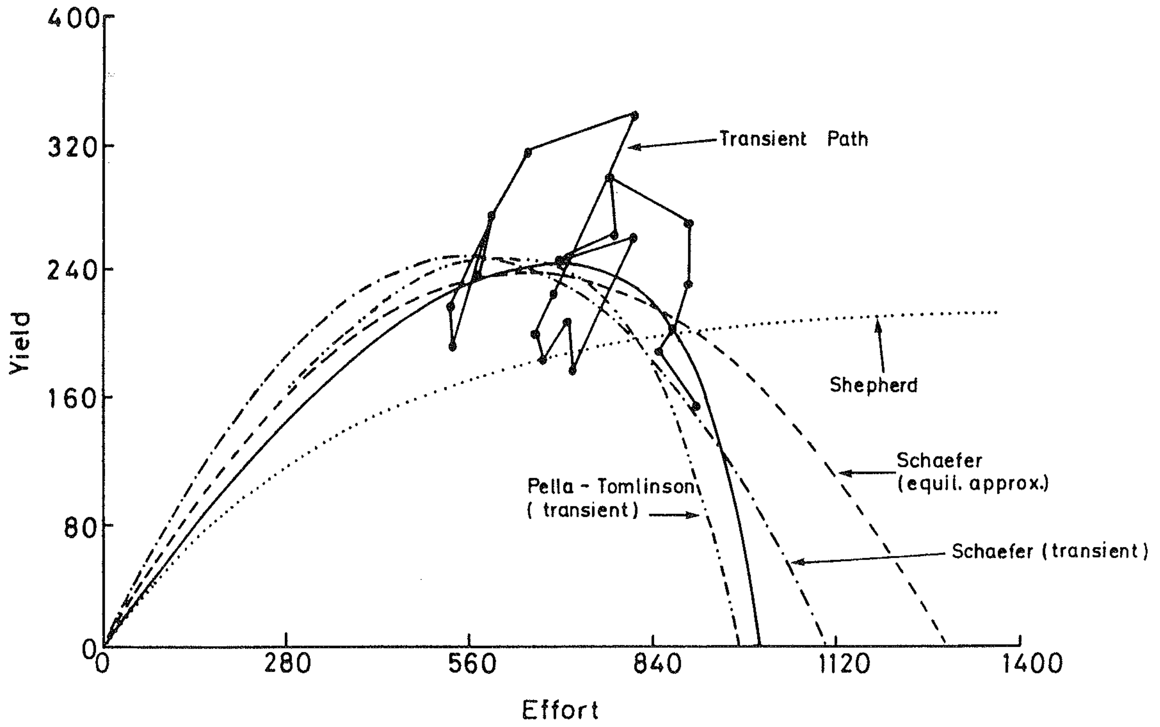


Figure 2.3.7 S. HORSE MACKEREL

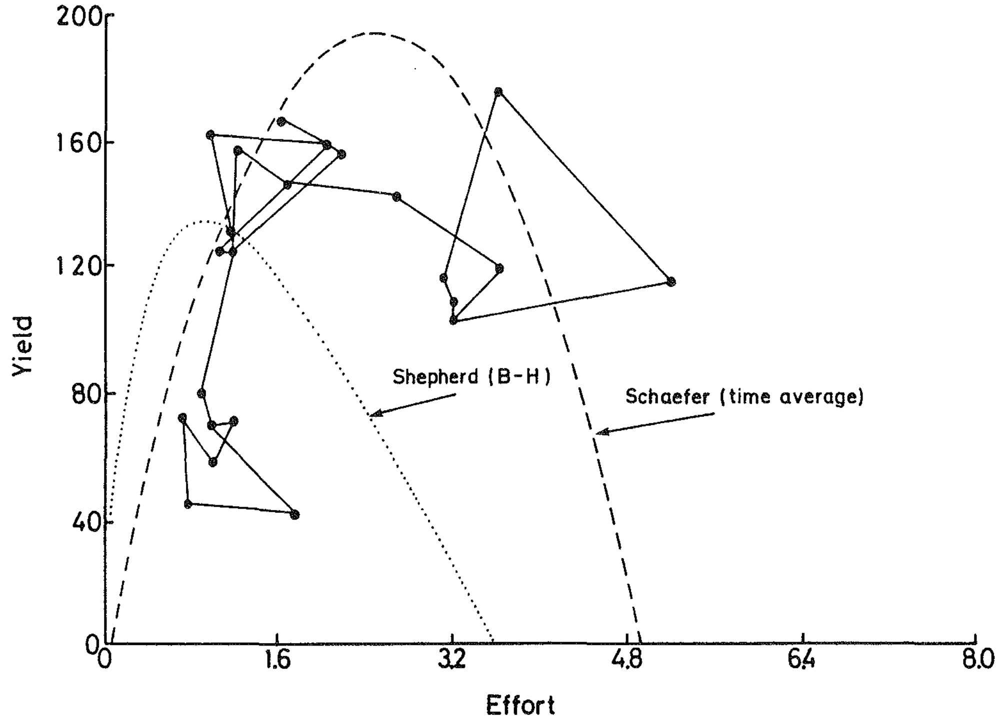


Figure 2.3.8 PACIFIC HALIBUT

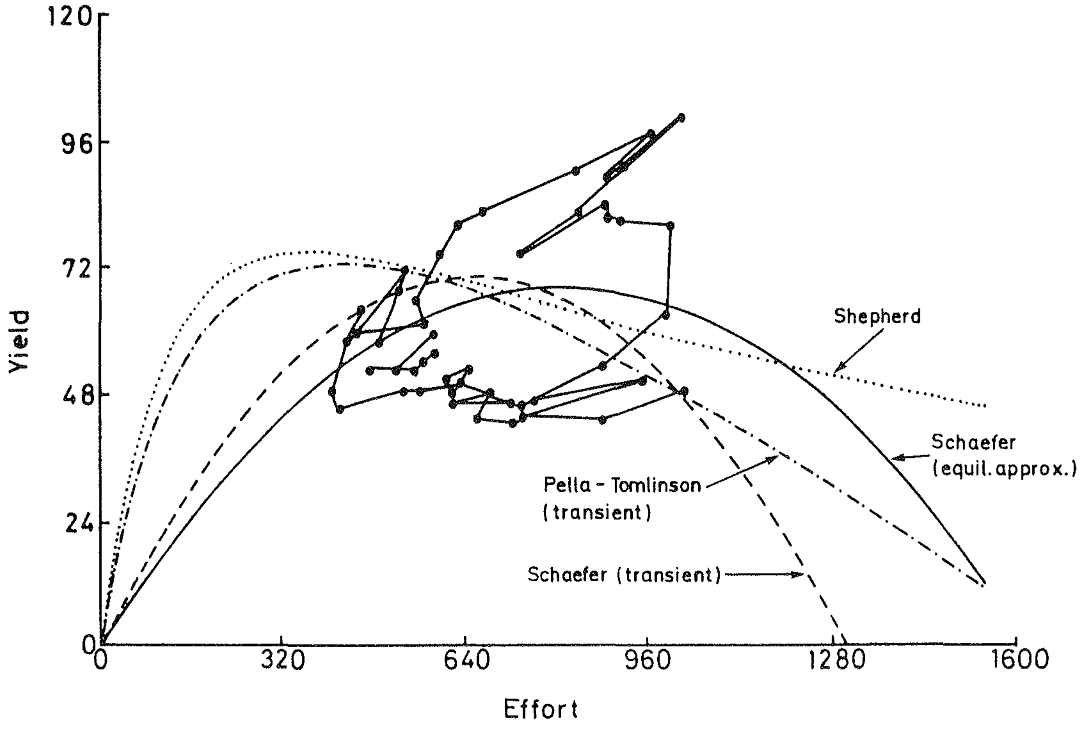


Figure 2.3.9 SIMULATED DATA NO NOISE

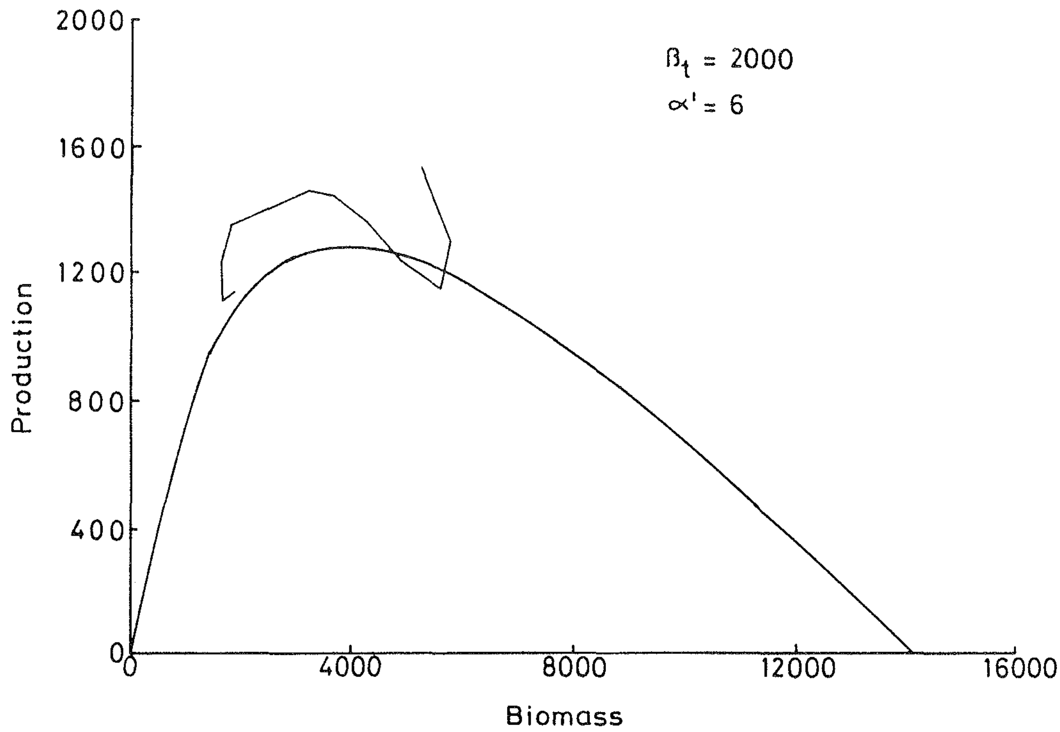


Figure 2.3.10 SIMULATED DATA MEASUREMENT NOISE

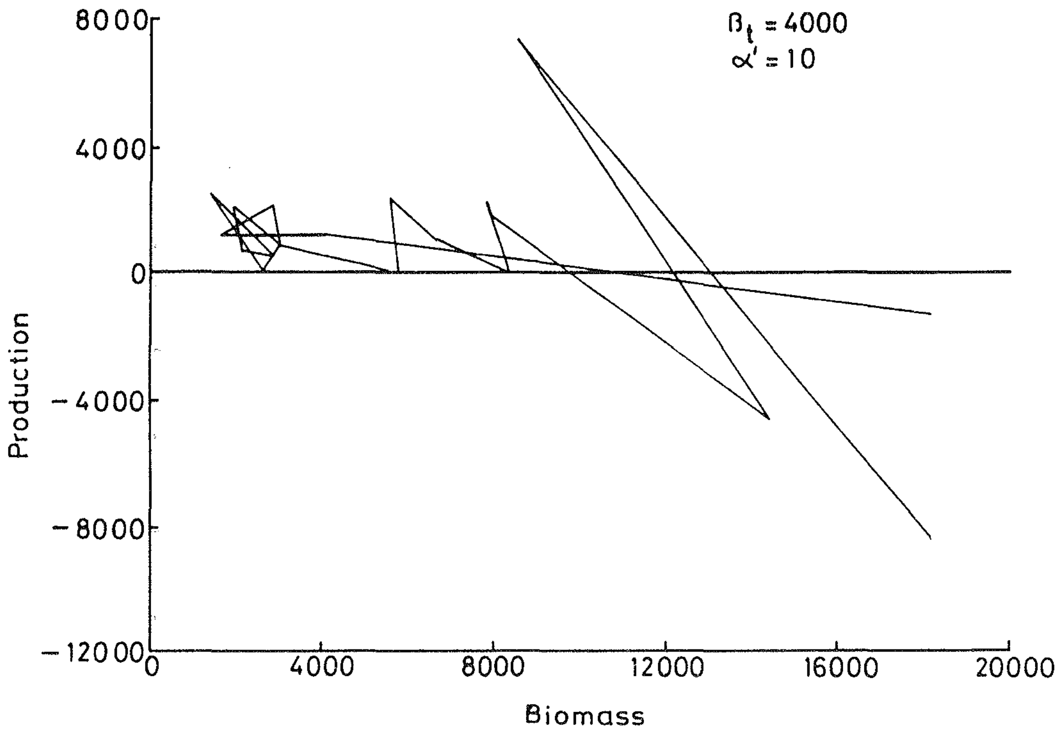


Figure 2.3.11 SIMULATED DATA PROCESS NOISE

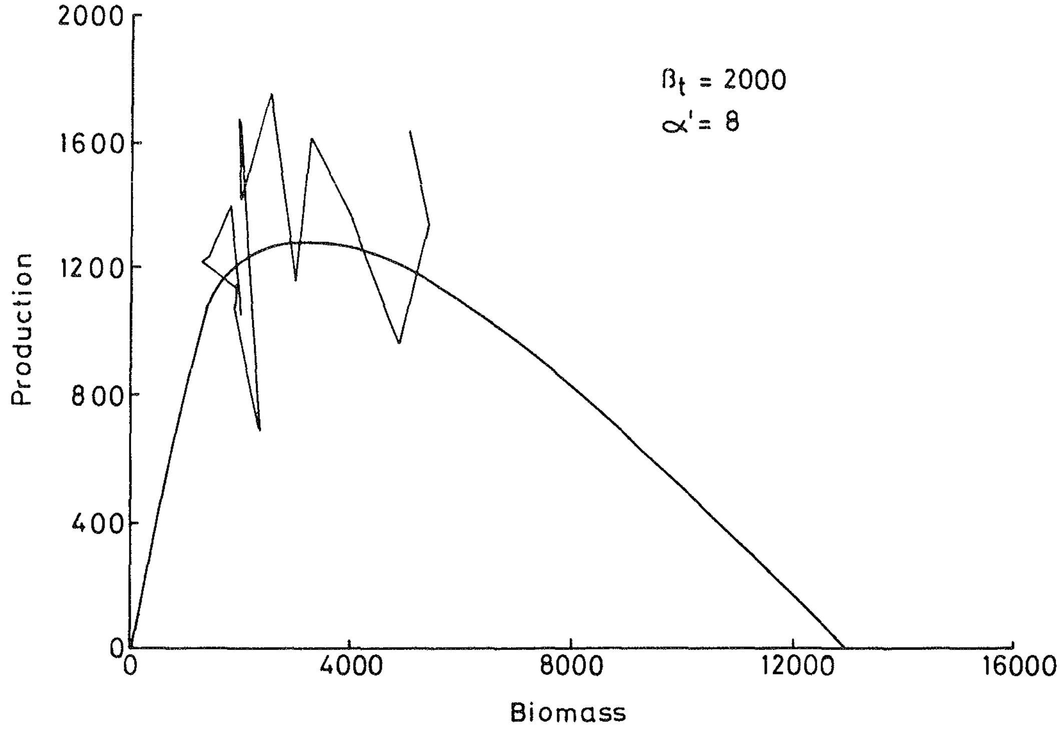


Figure 2.3.12 COD

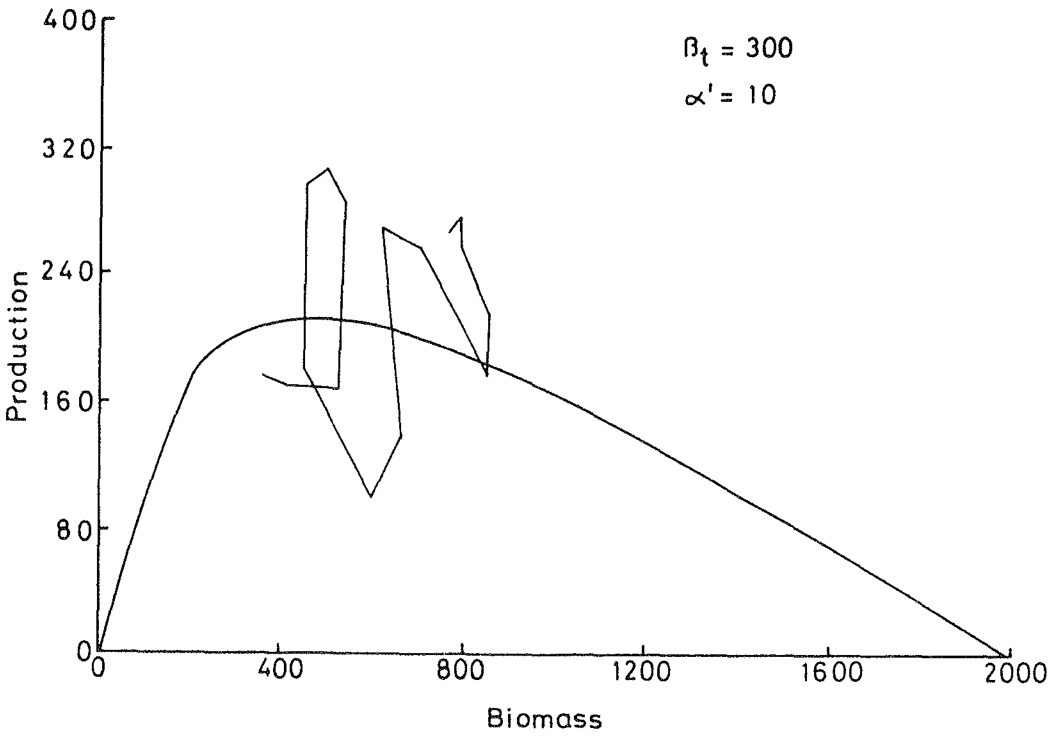


Figure 2.3.13 HALIBUT

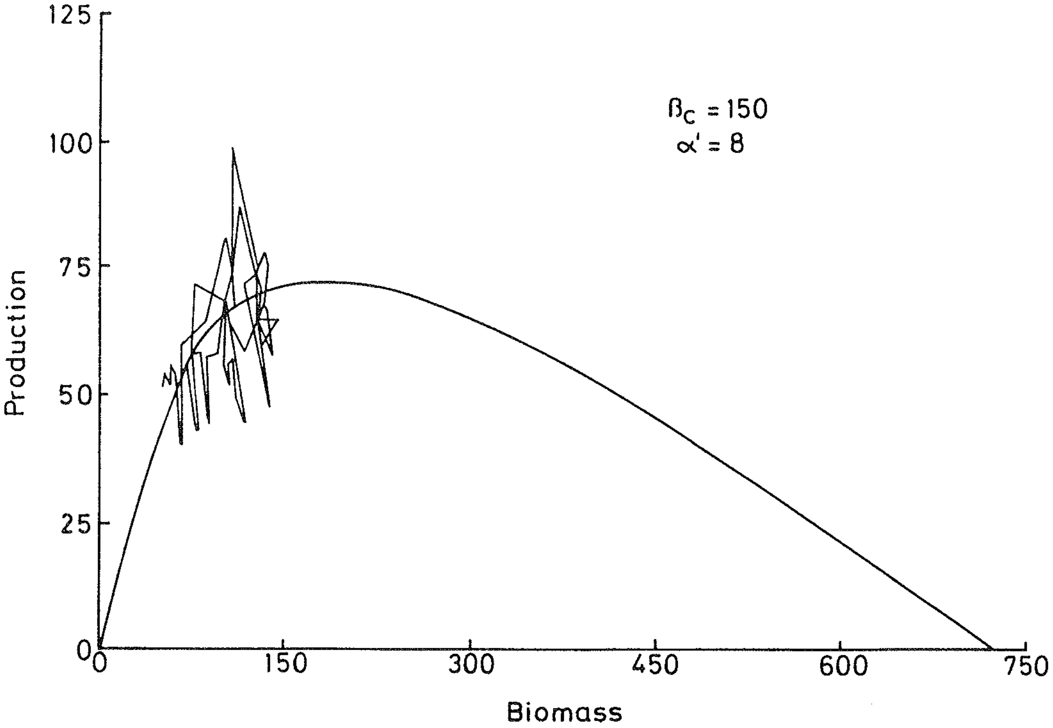


Figure 2.3.14 S. HORSE MACKEREL

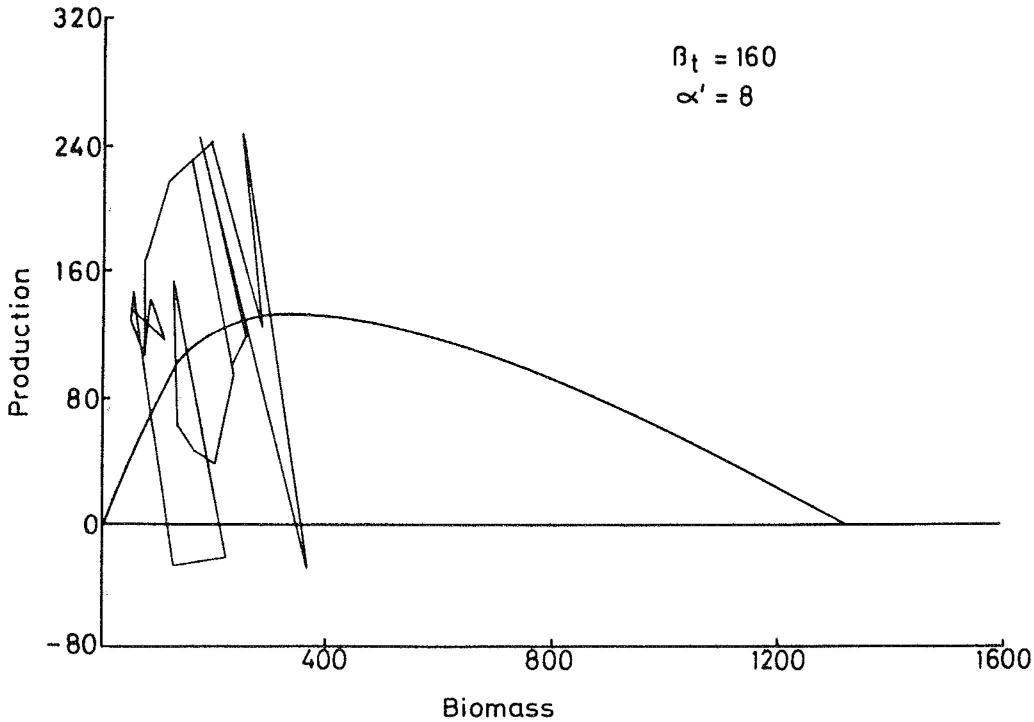


Figure 2.3.15 S. HORSE MACKEREL

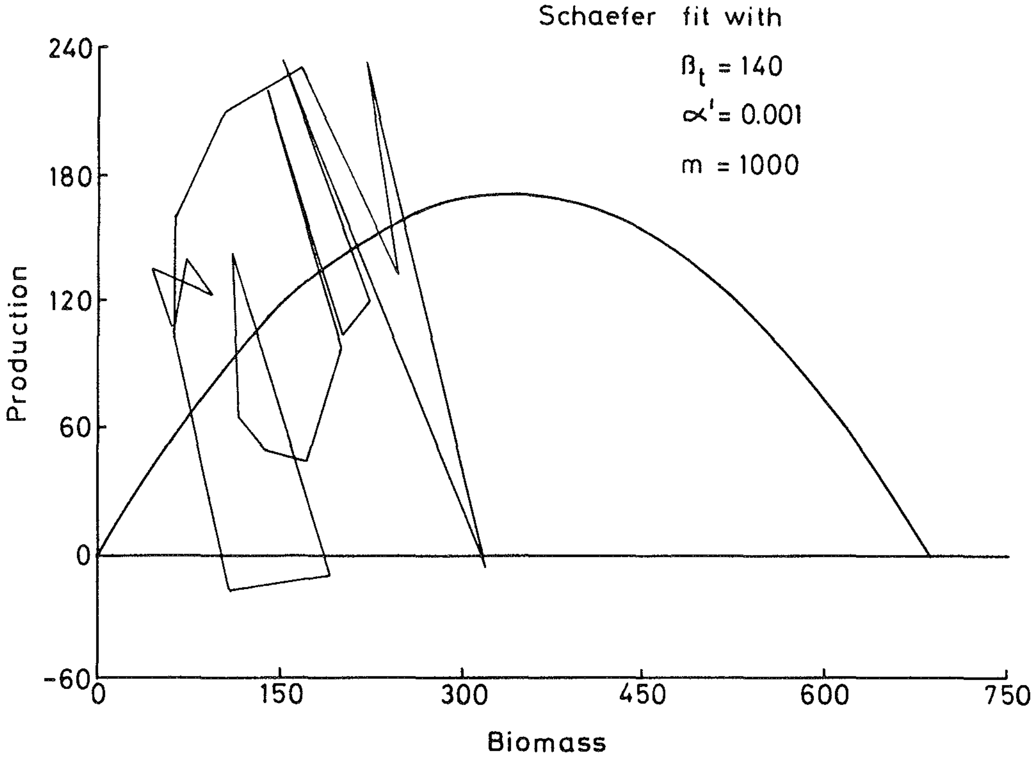


Figure 2.3.16

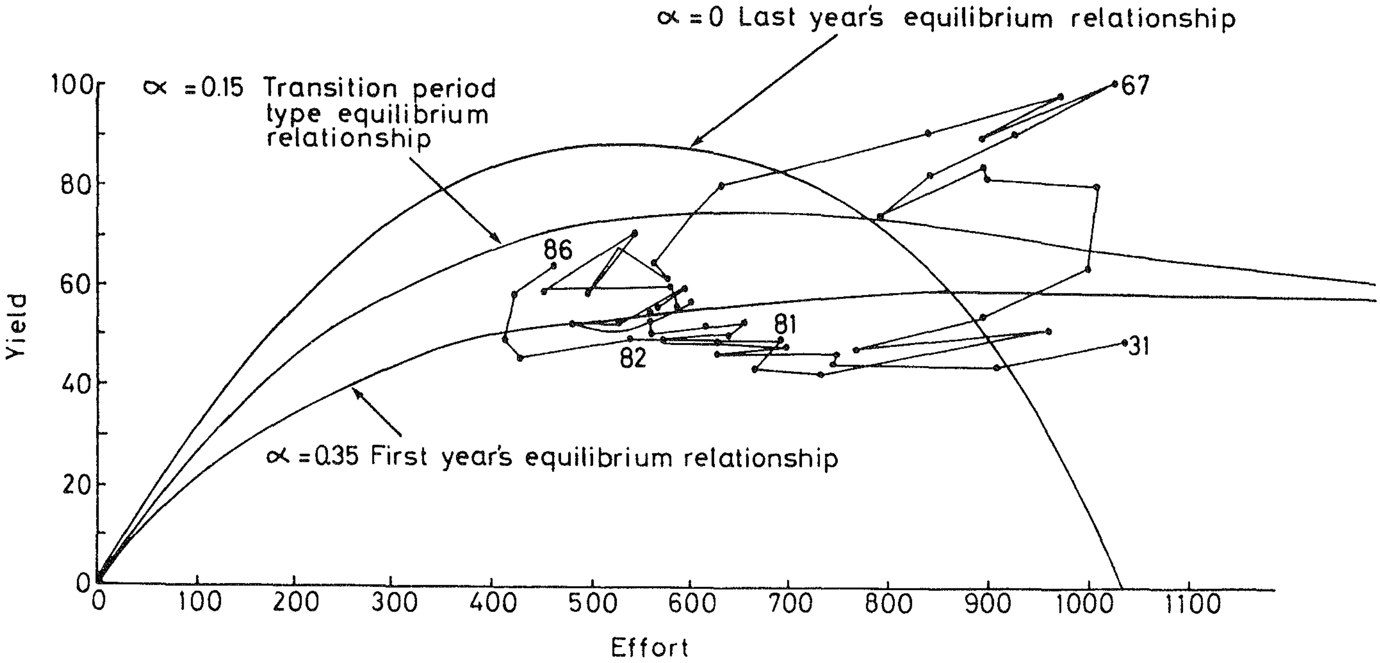


FIG 4.3.1 PACIFIC HALIBUT

RELATIVE FISHING MORTALITY (1974=100%)

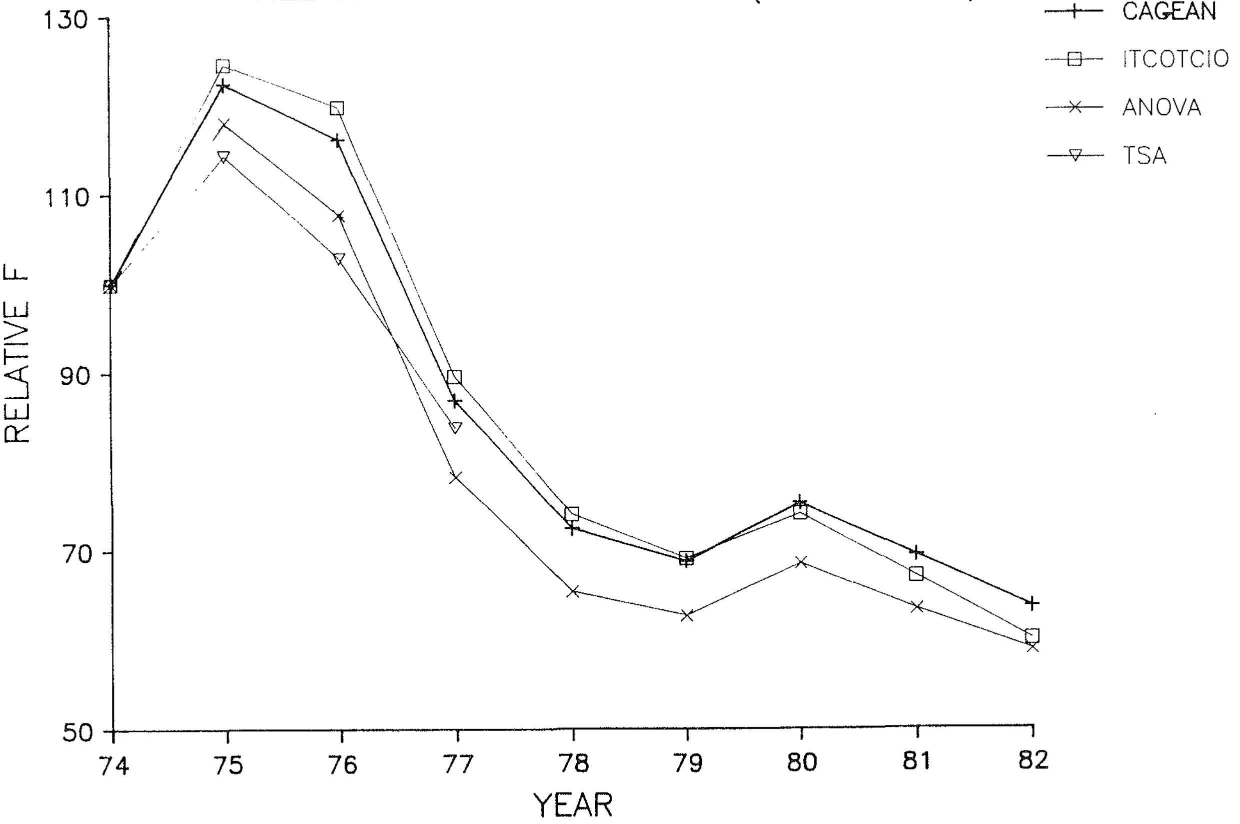


FIG 4.3.2 PACIFIC HALIBUT

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

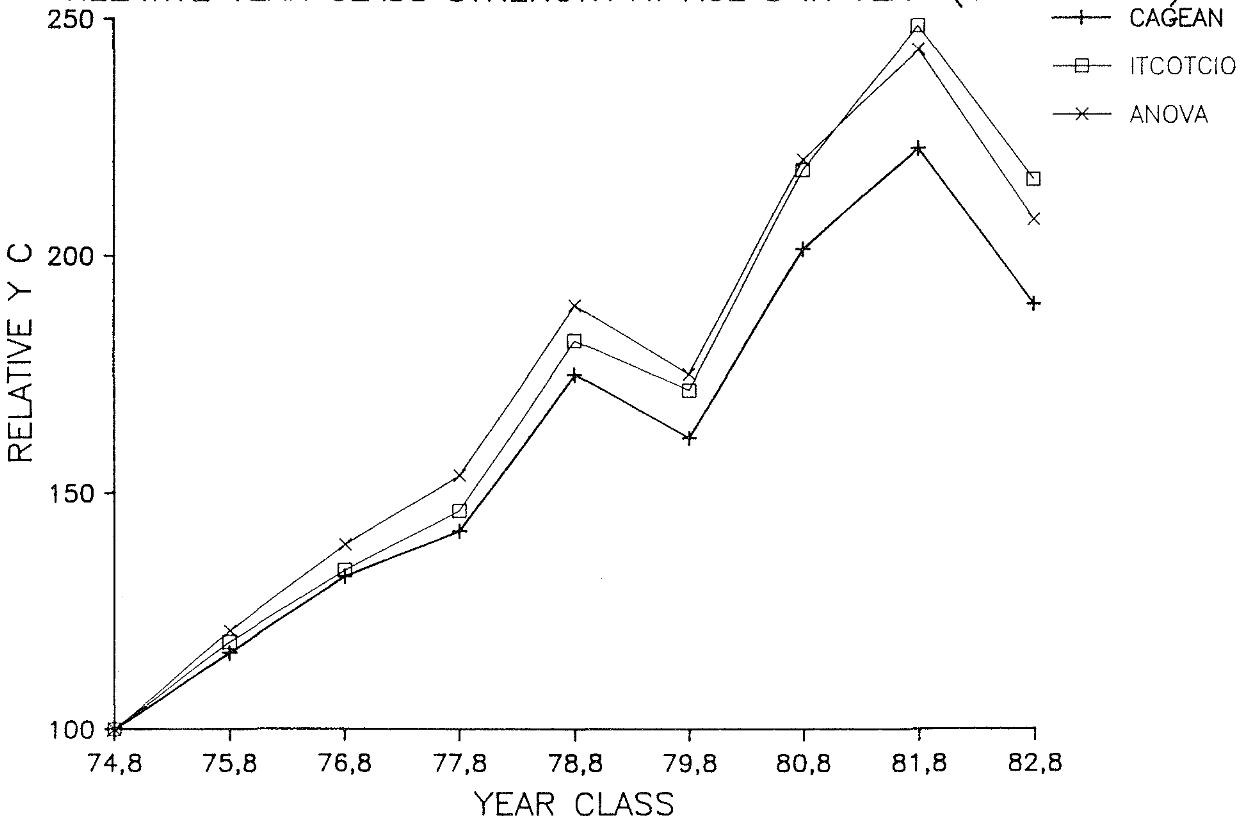


FIG 4.3.3 PACIFIC HALIBUT

RELATIVE EXPLOITATION PATTERN,AGE 14=100%

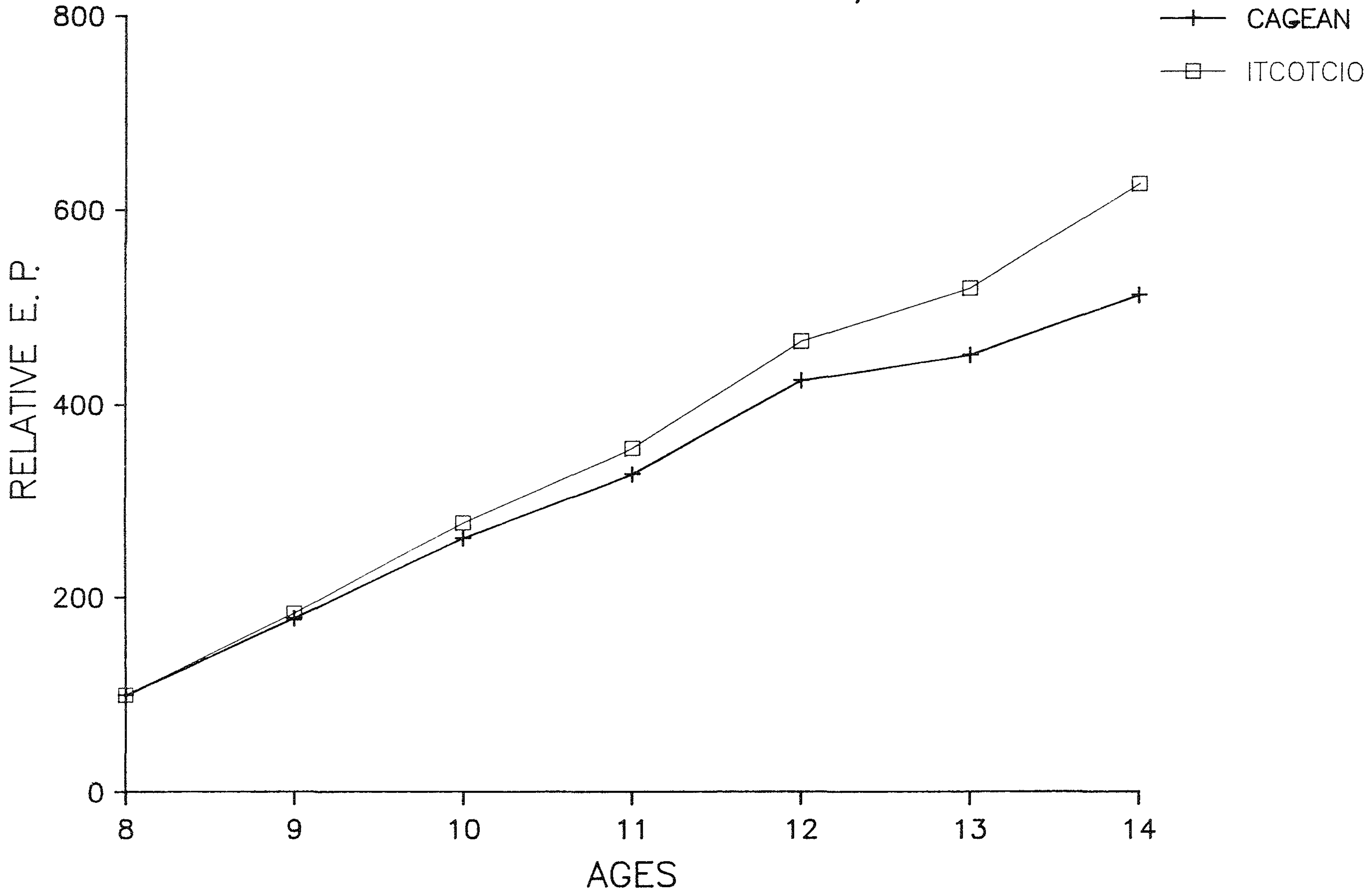
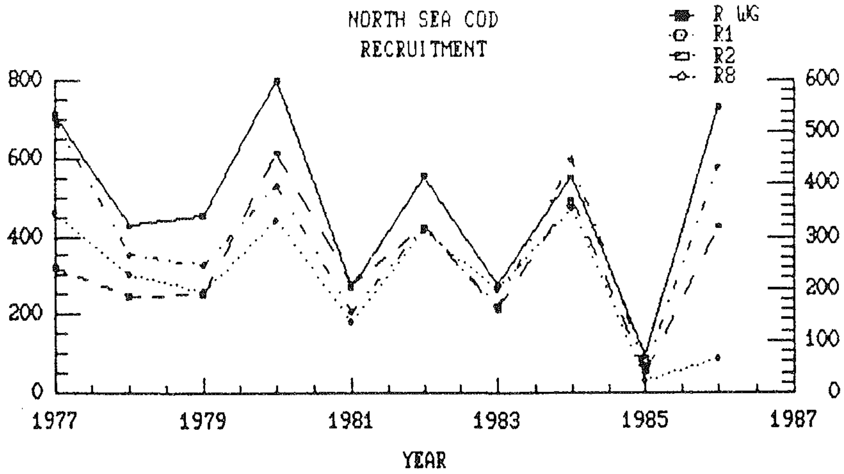
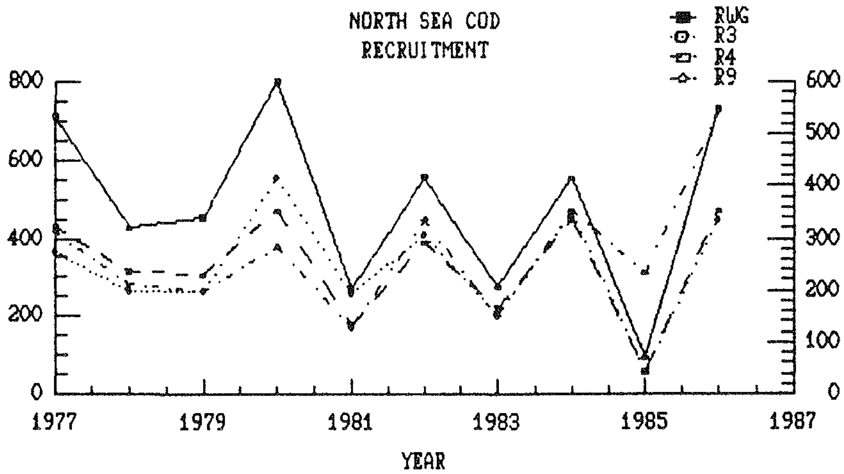


Figure 4.3.4


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

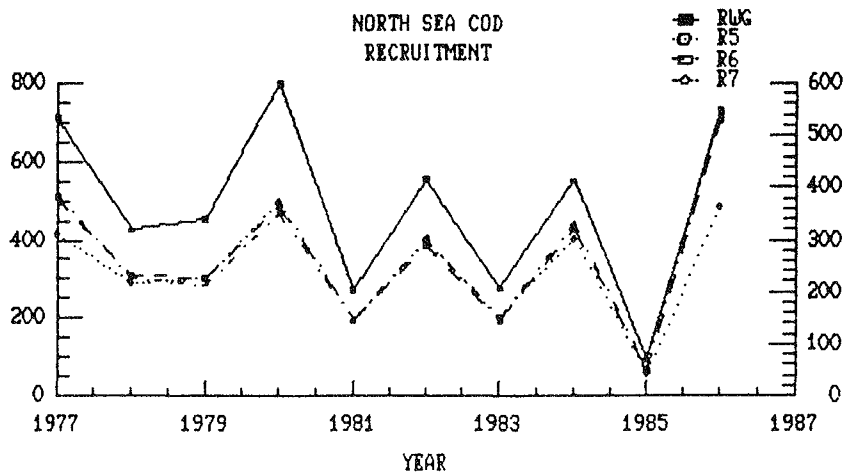
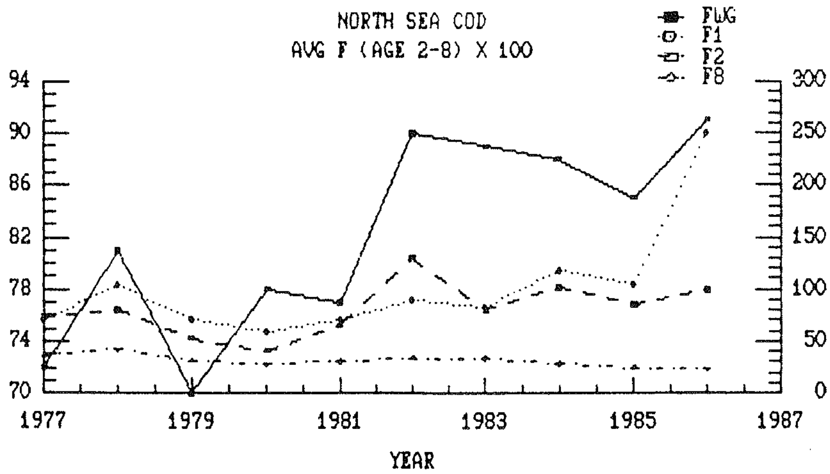
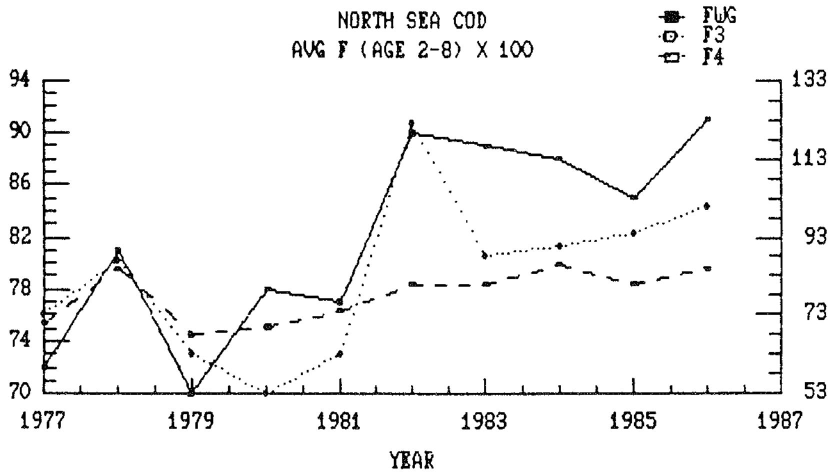

NORTH SEA COD
RECRUITMENT
RWG
R5
R6
R7
800
600
400
200
0
600
500
400
300
200
100
0
1977
1979
1981
1983
1985
1987
YEAR

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

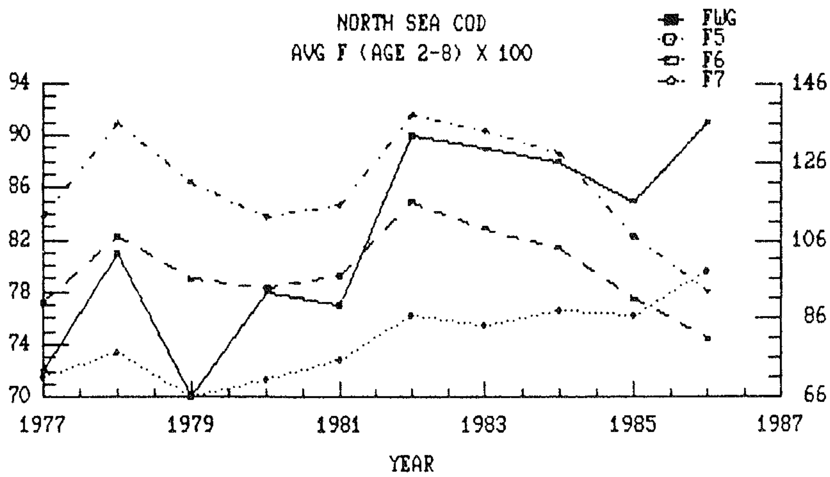
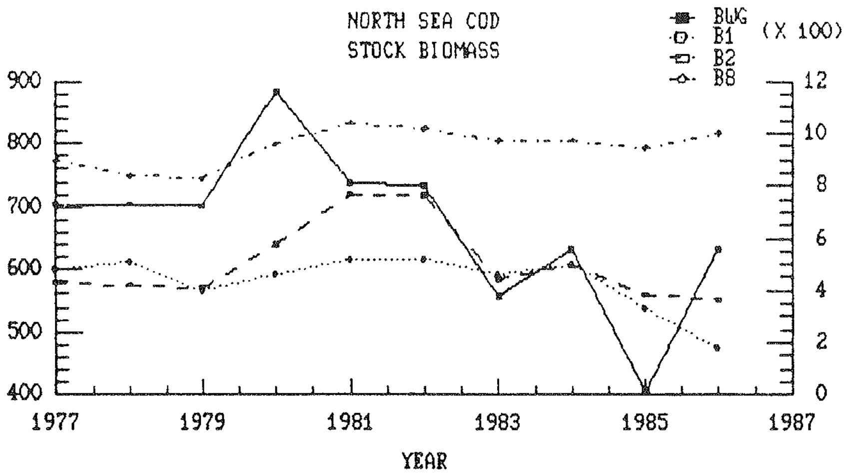
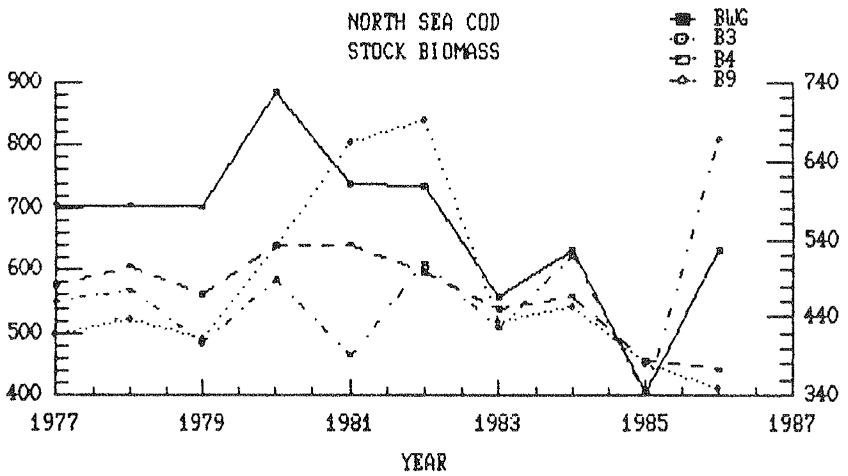
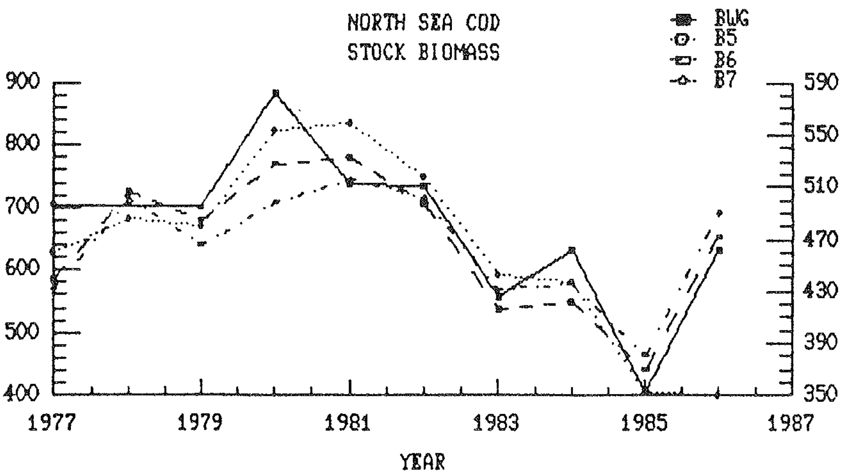

NORTH SEA COD
AVG F (AGE 2-8) X 100
FWG
F5
F6
F7
94
90
86
82
78
74
70
146
126
106
86
66
1977
1979
1981
1983
1985
1987
YEAR

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

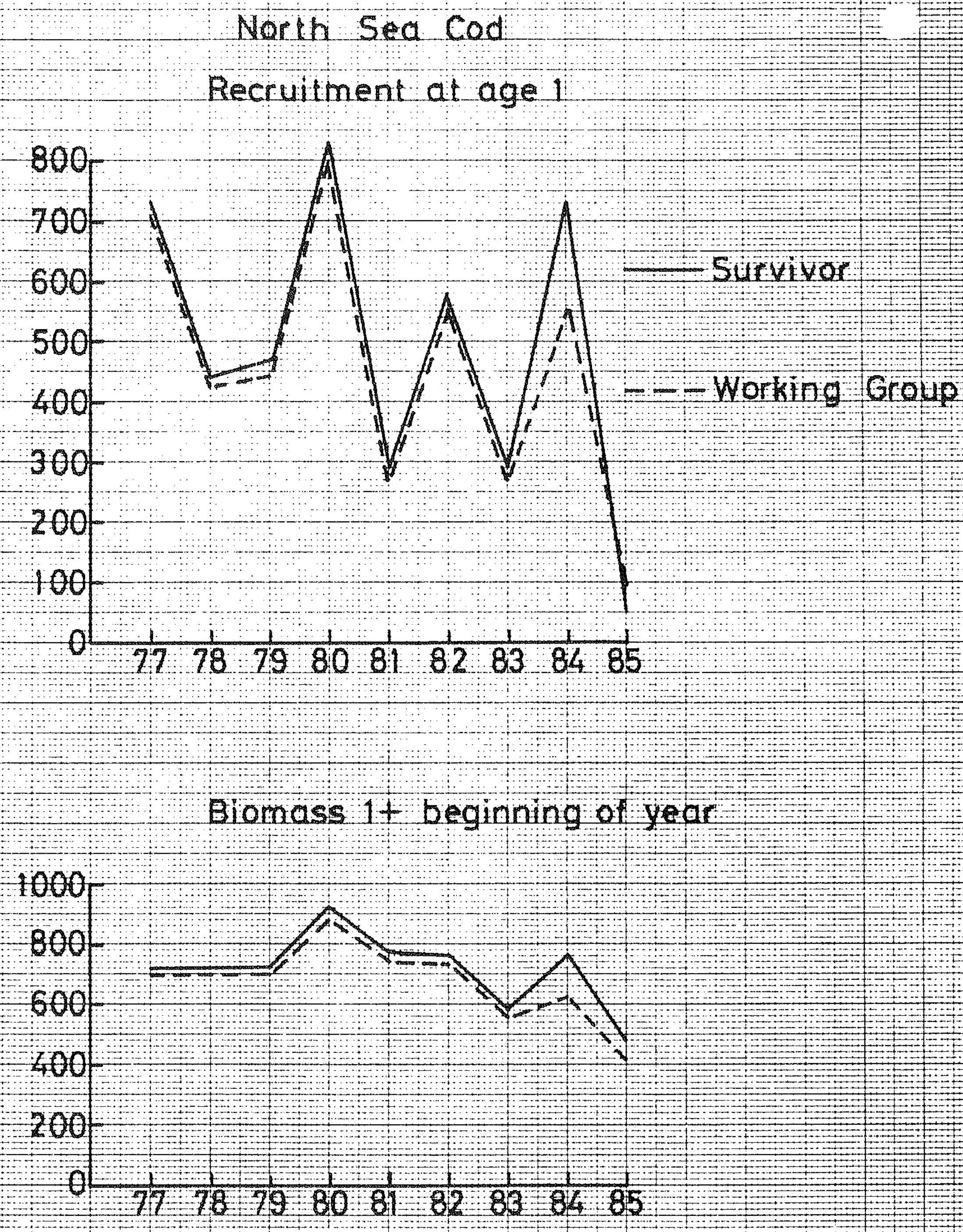

North Sea Cod
Recruitment at age 1
800
700
600
500
400
300
200
100
0
77 78 79 80 81 82 83 84 85
Survivor
Working Group
Biomass 1+ beginning of year
1000
800
600
400
200
0
77 78 79 80 81 82 83 84 85

Figure 4.3.7

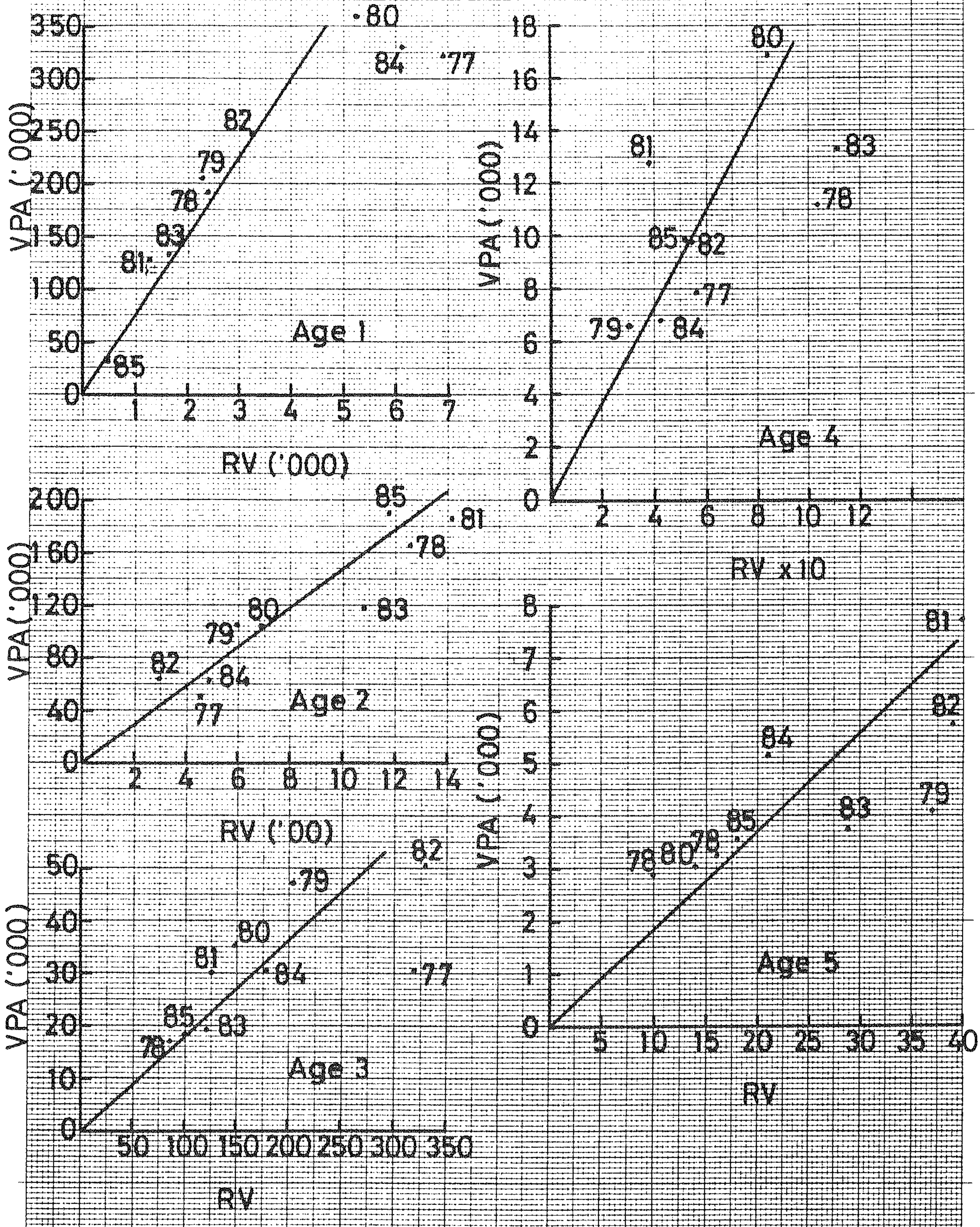


Figure 4.3.8 North Sea cod. Results of SURVIVOR analysis.
Population numbers vs reseach vessel estimates.

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

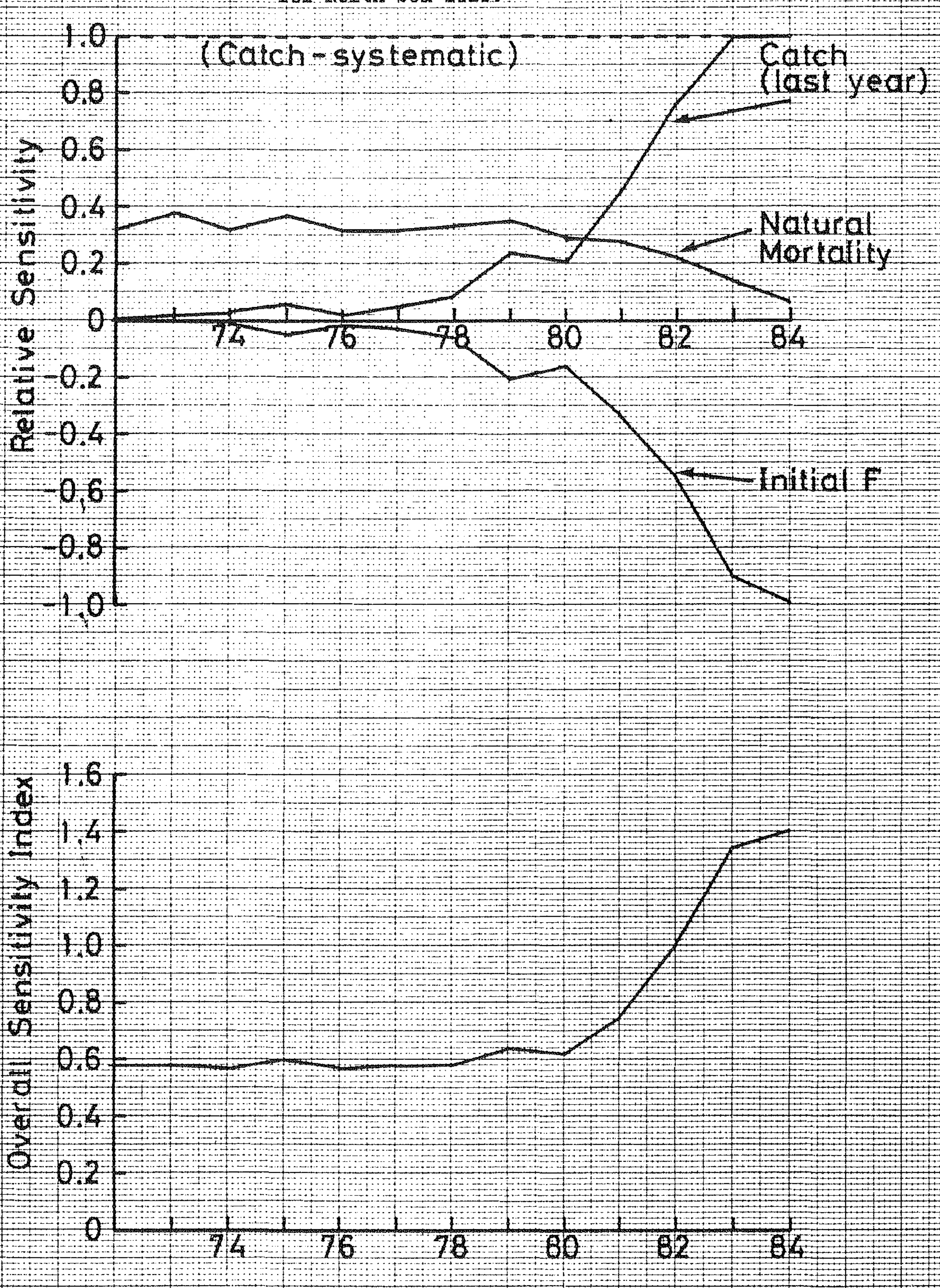
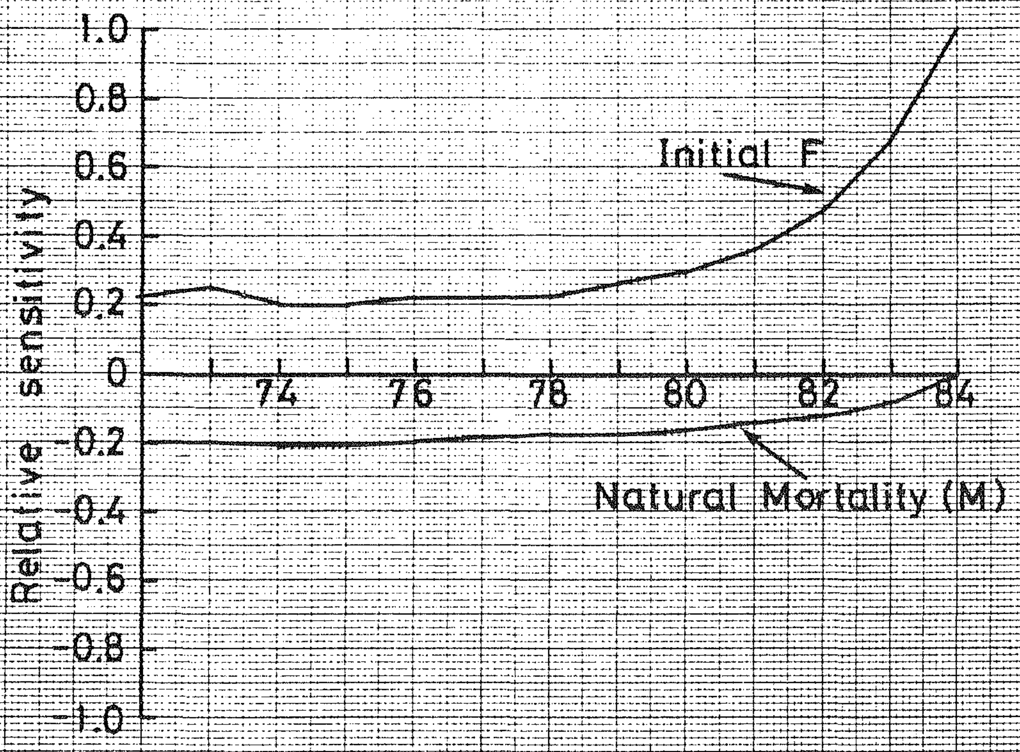


Figure 5.3.2 Relative sensitivity coefficients of fishing mortality (F) for all age groups to initial F and M: North Sea sole.



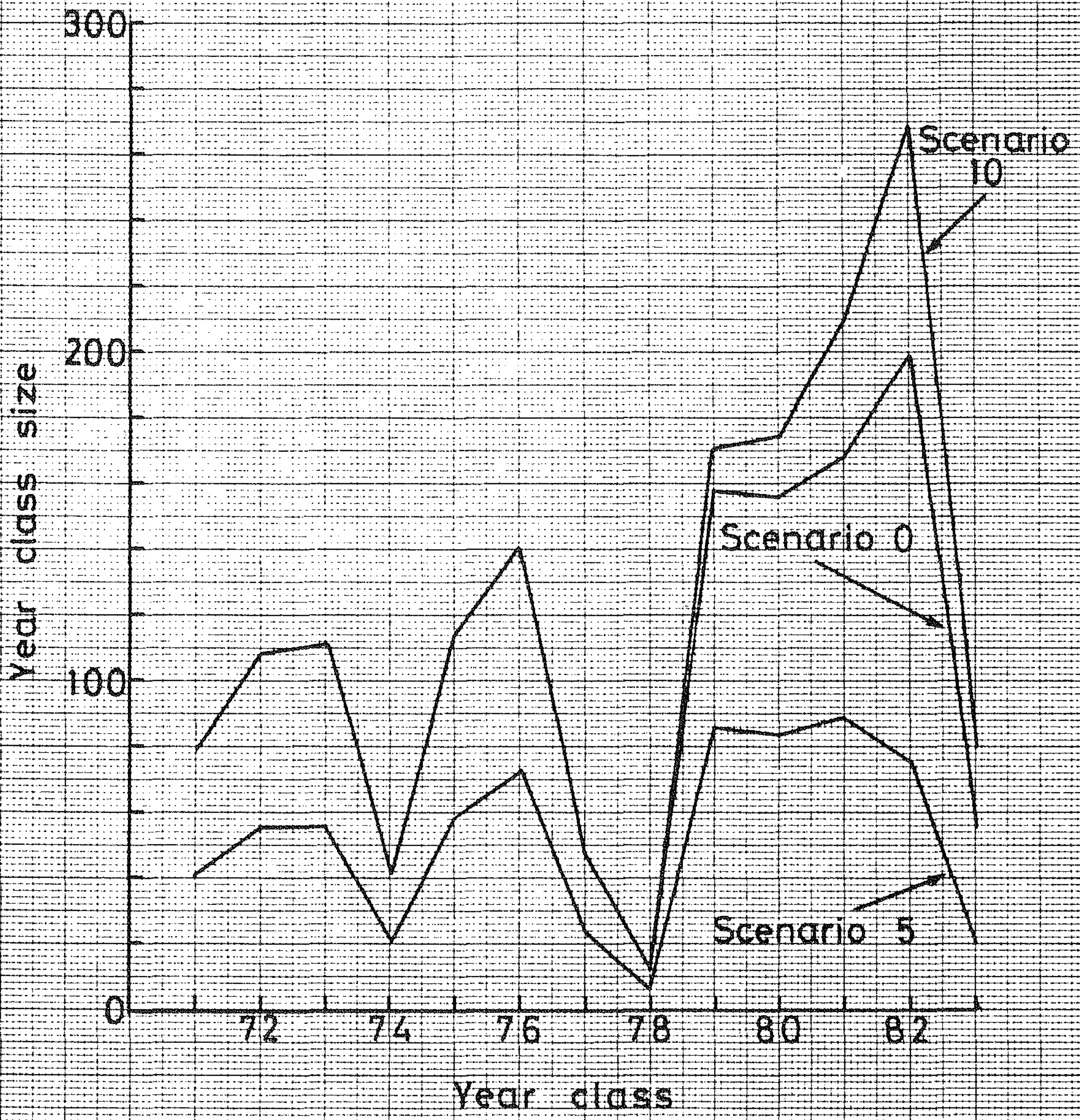
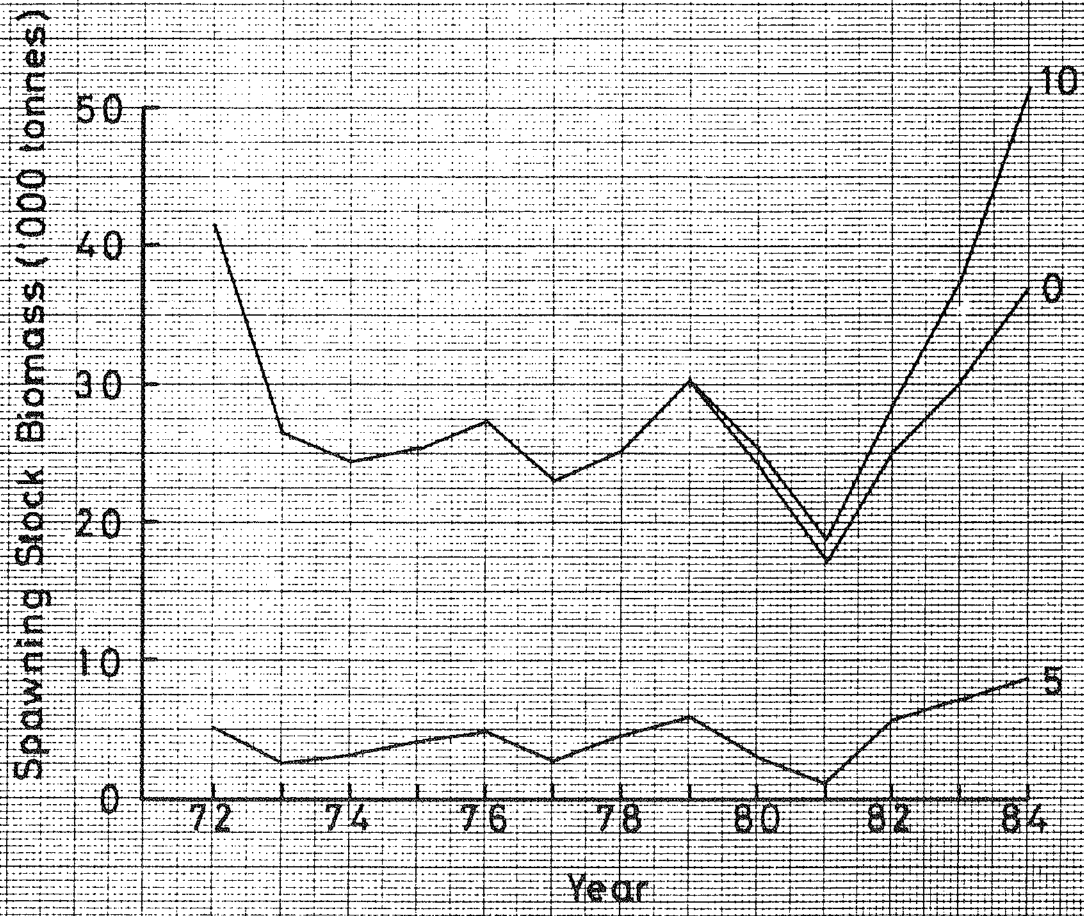


Figure 5.3.3 Effect of various sources of error on the estimation of annual recruitment (year class size) for the North Sea sole.

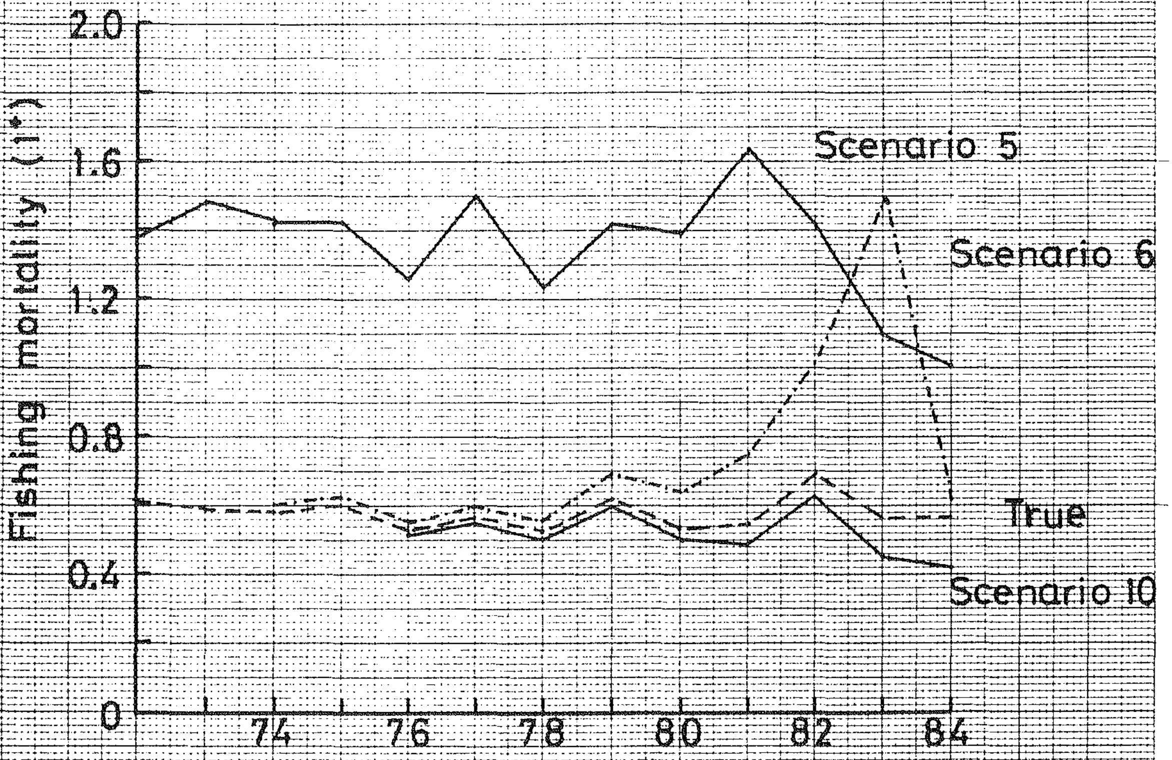
Scenario	Description
0	True catches
7	Increasing trend in misreporting over years and and ages, effort misreported proportionally
9	Effort misreported by 20% in last year

Figure 5.3.4



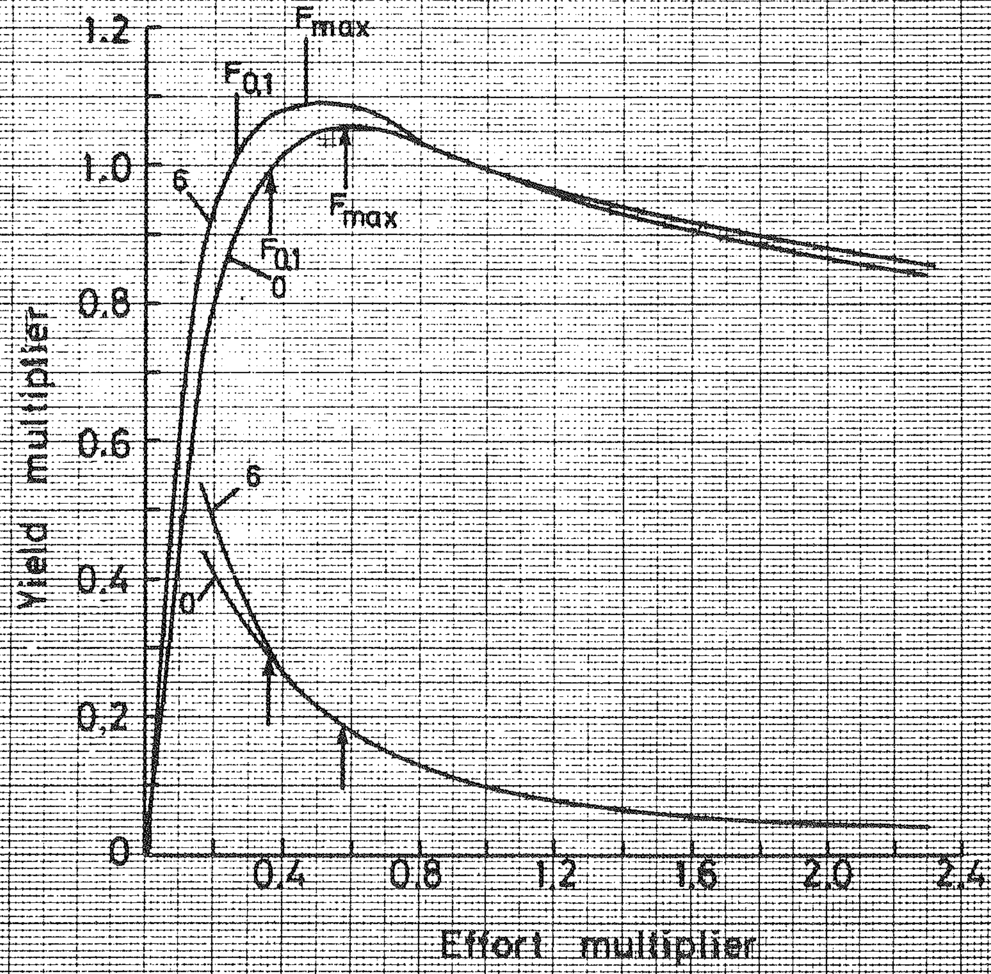
Spawning Stock Biomass ('000 tonnes)
50
40
30
20
10
0
72
74
76
78
80
82
84
Year
10
0
5

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


Yield multiplier
Effort multiplier
F_{max}
$F_{0.1}$
F_{max}
$F_{0.1}$
6
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

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 multi-fleet 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 vari- ations 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

NOTE: 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
$l\infty$	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 method of assessment	-	-	M	M	i	-	-
3.	Measures of overall fising 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 assessmemts	-	-	-	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-at-age analysis with auxiliary information	Deriso *et al.* (1985)
4. ITCOTCIO	GLIM Macro	Wide implementations	Catch-at-age analysis	J.G. Pope and T.K. Stokes (WP 4)
5. Non-linear fitting	Fortran Genstat	NORD	Fitting production models with Marquadt algorithms	F. Laloe[3]
6. PROD	APL	Mainframe	Production models: transient forms	S. Gavaris[4]
7. PRODAFIT	APL	IBM-OC	Production models: equilibrium approximations	Rivard (1982)
8. RCRTINY	Fortran 77	MSDOSN NORD HP, etc.	Weighted average combination of recruitment indices	J.G. Shepherd[5]
9. SHAEFER	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	MORD	Analysis of residuals	Gudmundsson (1986)
13. TSM	Fortran 77	VAX	Catch-at-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, Suffull NR 33 OHT, UK.
[2]A. Laurec, IFREMER, rue de l'ile d'Yeu, B.P. 1049, 44037 Nantes Cédex, France.
[3]F. Laloe, C.R.O.D.R/. 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, Duffolk NR33 OHT, UK.
[6]R.K. Mohn, Dept. of Fisheries & Oceans, P..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:

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

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

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

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:

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

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:

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

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.

REPORT OF THE WORKSHOP ON METHODS OF FISH STOCK ASSESSMENTS

Reykjavik, 6-12 July 1988

1 PARTICIPANTS AND TERMS OF REFERENCE

1.1 Participants

David Armstrong (Chairman)	UK (Scotland)
Armando Astudillo	Spain
Vladimir Babayan	USSR
M. Fatima Borges	Portugal
Ghistain Choinard	Canada
Ray Conser	USA
Robin Cook	UK (Scotland)
Yury Efimov	USSR
Eduardo Ferrandis	Spain
Dominique Gascon	Canada
Stratis Gavaris	Canada
Asta Gudmundsdóttir	Iceland
Gudmundur Gudmundsson	Iceland
Thorkell Helgason	Iceland
Vidar Helgason	Iceland
Mikael Hildén	Finland
Holger Hovgård	Greenland
Tore Jakobsen	Norway
Hans Lassen	Greenland
Alain Laurec	France
Peter Lewy	Denmark
Qun Liu	UK (Wales)
Robert Mohn	Canada
Steen Munch-Petersen	Denmark
Ransom A. Myers	Canada
Phillip R. Neal	USA
Gunnar Petersson	Iceland
John G. Pope	UK (England)
Terrance Quinn	USA
Denis Rivard	Canada
Andrew A. Rosenberg	UK (England)
John G. Shepherd	UK (England)
Arnauld Souplet	France
Gunnar Stefánsson	Iceland
Bjørn Steinarsson	Iceland
Man Sun	UK (England)

1.2 Terms of Reference

At the 75th Statutory Meeting of ICES (1987) it was decided (C.Res.1987/2:11) that:

"As part of the preparatory process for the next meeting of the Working Group on Methods of Fish Stock Assessments, a Workshop will be held in Reykjavik from 6-12 July 1988 (Chairman: Mr A. Laurec) for the purpose of testing software methods which perform statistical integrated analysis of catch-at-age data and auxiliary information, and constructing and implementing appropriate test data sets. Results of these methods will be contrasted with the output from equivalent *ad hoc* VPA tuning methods. Local arrangements for the Workshop will be coordinated by Dr G. Stefánsson."

Following this resolution, Mr Laurec found that, because of other commitments, he could not act as Chairman and it was decided at the November 1987 meeting of ACFM to offer the chairmanship to Mr D.W. Armstrong.

2 INTRODUCTION

2.1 Interpretation of "Stock Assessment"

For the purpose of this report, the meaning of "fish stock assessment" is restricted to any procedure by which the historical and current state of a fish stock is estimated.This definition includes no reference to prediction of possible future states of the stock and no attention was given to prediction in the course of this meeting.

It should also be noted that, in real-life assessments, recruitment estimates for the most recent data years are often obtained by techniques additional to those used to analyze the catch-at-age and auxiliary data. No attention was given to such methods at this meeting.

2.2 Requirements for Testing Methods of Assessment

Particularly during the past 4-5 years, considerable development of new methods for fish stock assessment has occurred. In many instances, the new methods have not been extensively tested and the first application of any of them has often taken place during stock assessment working group meetings when the results are of material importance to non-scientists. In some instances, use of different methods to assess the same stock has produced considerably different results leading to confusion.

Furthermore, development of new techniques has taken rather different routes in Europe and North America. In North America, the focus has been on fitting formal mathematical models by standard statistical techniques (minimization of an objective function). In Europe, much more attention has been given to developing so-called *ad hoc* "tuning" methods in which non-standard techniques are used to find a solution for the last data year which is consistent with historical parameter estimates.

Given this background, it was felt essential that the various methods should be tested at least to identify those which produce unacceptably poor results. Ultimately, the aim of the testing procedure should be to identify an overall best method or a best method contingent on the nature of the stock being assessed.

2.3 Methods Tested

The 18 methods listed below were tested.

Number	Name of Method	Acronym
1	Hybrid	HYBRID
2	Laurec-Shepherd	LS
3	Armstrong-Cook 1	AC1
4	Armstrong-Cook 2	AC2
5	Armstrong-Cook 3	AC3
6	Armstrong-Cook 4	AC4
7	Alternative Estimation of Fishing Mortalities	AEFM
8	Corrected Catch per Unit Effort	CCPUE
9	Survivors	SURVIV
10	Extended Survivor Analysis	XSA
11	Catch at Age Analysis	CAGEAN
12	Adaptive Approach	ADAPT
13	General Linear Model	GLM
14	Collie-Sissenwine	COLSIS
15	Time Series 1	TSER1
16	Time Series 2	TSER2
17	Separable VPA	SVPA
18	Conventional VPA	CONVEN

A description of each of these methods together with details of the way in which they were applied, an account of the ease (or otherwise) of application, and references to further descriptions in the scientific literature are given n Annex 2.

Methods 1-8 in the list above are *ad hoc* tuning methods. Methods 11-14 are the integrated methods. Methods 9 and 10 incorporate some features of both the *ad hoc* and the integrated approach. Methods 17 and 18, unlike the others, cannot make use of auxiliary data (CPUE) and were tested to indicate the improvement which may be obtainable by the appropriate use of such data.

The methods are listed in the order in which they appear in the tabulations included in this report. The acronyms listed above are used to indicate the methods in these tables.

The assumptions inherent in each of the methods are summarized in Table 2.1. It should be noted that the assumptions listed are those incorporated to produce the results presented in this report. Within many of the methods these assumptions can be modified. The various tuning methods can be regarded as the same method run under different assumptions. Similarly, the difference between the two Time Series methods is that TSER1 analyzes only the total catch-at-age data, whereas TSER2 also analyzes CPUE data from one of the research vessels. The adaptive approach is specifically designed to allow modification of assumptions and incorporation or exclusion of various data sets.

3 PROCEDURE FOR TESTING METHODS

3.1 Simulated Data Sets

The basic approach adopted was to investigate how well each method estimated certain parameters employed in creating simulated data sets. Details of the simulation method and the input parameters for each simulation are provided in Annex 1. By appropriate choice of the values of the input parameters, it is possible to simulate different types of fisheries exploiting different types of stocks and hence, for each combination of fishery and stock, to produce data of the type commonly analyzed by stock assessment.

The output from the simulation process consisted of estimates of catch at age for each of seven fleets, four of which were commercial fisheries (two trawler fleets, one liner fleet, and one fleet of fixed nets), and the other three were research vessels. Estimated fishing effort was provided for the research vessels, for liners, and for one of the trawler fleets. Catch-at-age data were provided for ages 3-12 for a period of 30 years for all fleets.

Noise was added to the output data sets in the form of process error and measurement error as described in Annex 1. These errors were different for different age groups and fleets.

Mean weight at age and proportion mature at age were assumed to be constant and known. Natural mortality rate was assumed to be 0.2 for all ages and years and known.

Six data sets were assessed the main features of which are described below (see Annex 1 for full details).

Data Set 1: No trends in catchability in any fleet. Total international F about 0.4 for the whole of the 30-year period. Process and measurement errors log-normal. Separable F at age for each fleet.

Data Set 2: No trends in catchability in any fleet. Total international F about 1.0 for the whole of the 30-year period. Process and measurement errors log-normal. Separable F at age for each fleet.

Data Set 3: Catchability trends in the two commercial fleets for which effort data are available. No catchability trends in other commercial fleets or in research vessels. Total international F around 0.4, but with steadily increasing trend. Process and measurement errors log-normal. Separable F at age for each fleet.

Data Set 4: Catchability trends in all fleets for which effort data are available (including research vessels). Total international F around 0.8 in year 1 increasing to about 1.2 in year 30. Process and measurement errors log-normal. Separable F at age for each fleet.

These four data sets were sent to the assessors in advance of the meeting. Having carried out their assessments, all of the assessors considered that the data were too "clean". In particular and when the method of simulation and the precise nature of these data sets was revealed, it was suggested that:

i) the research vessel data should have higher variances,

ii) separability assumptions for each fleet may be violated in reality,

iii) errors in catch-at-age data may be gamma-distributed rather than log-normally distributed,

iv) some methods assumed exponential trends in catchability and since this assumption is incorporated in those data sets where catchability is allowed to change, these methods would be in an advantageous position when assessing data of the type provided,

v) research vessel effort data varied considerably from year to year.

Accordingly, during the meeting, two other data sets were prepared in an attempt to overcome these criticisms.

Data Set 5: Same as Data Set 3 except that gamma-distributed process noise used on F-at-age and catch-at-age data (log-normal noise retained on fishing effort). Level of noise increased compared to Data Sets 1-4.

Data Set 6: Noise treated in the same way as Data Set 5. F at age not separable for any fleet for the whole of the simulated time period.

It should be stressed that, ideally, the assessors would have carried out extensive exploratory analysis of the data sets prior to producing their results. Many of the methods routinely produce diagnostic statistics (HYBRID, LS, CAGEAN, ADAPT, TSER) and some methods (especially ADAPT) actively encourage intervention by the operators. However, in the time available, only cursory reference to diagnostics was possible. Because of this, the results from these methods presented in this report may not be the best attainable.

These data sets are large, and it has been decided that they will not be tabulated in this report. Copies of them can be obtained on IBM-formatted disk from:

D.W. Armstrong,
DAFS Marine Laboratory, P.O.Box 101
Torry, Aberdeen AB9 8DB, UK

or

G. Stefánsson
Marine Research Institute
P.O. Box 1390, Skúlagata 4
121 Reykjavik, Iceland

3.2 Estimation of Parameters of the Last Data Year in Simulated Data Sets

One of the most important results arising from a stock assessment is an appreciation of the state of the stock in the last data year since short-term conservation measures (TACs, effort and mesh regulations, etc.) are highly dependent on the current state of the stock. The current state of the stock is describable by estimating the parameters for the last data year of an appropriate fisheries model.

3.2.1 Procedure for comparison of methods

Because the simulation method incorporates stochastic processes, it is possible to produce many different realizations of the outputs for any constant set of input parameters. In principle, this property could have been used in a Monte Carlo test of each assessment method in which a large number of realizations of a data set could be analyzed to obtain the mean value (expectation) and variance of each parameter. These quantities could be used to compare the efficiency of the methods.

In advance of or during the meeting, a single realization of each of the six data sets was supplied to a number of nominated stock assessors. Each stock assessor was requested to apply a method which he had originated or which he is accustomed to using to each of the data sets. The true input parameter values were not provided to the assessors at this stage.

The assessors were asked to:

i) apply their method to data for years 2-21 and estimate parameter values for year 21

ii) apply the method to years 3-22 and estimate parameters of year 22,

iii) repeat for years 4-23, 5-24,......., 11-30.

The assessors were asked to record their estimates of:

i) number at age,

ii) F at age and mean F for ages 5-9,

iii) total and spawning biomass,

iv) catchability at age for each fleet for which effort data were provided.

(It should be noted that, in the time available, it was not possible to analyze estimates of catchability.)

The estimates were then compared to the true values used in producing the data sets supplied to the assessors. (In this context, the true values are the "realized" values referred to in Annex 1.) Two comparisons were made:

i) The percentage discrepancy between estimate and truth was calculated as:

$$PD = 100\,[(\text{Estimate}/\text{Truth})-1]$$

For each of the parameters listed above, ten discrepancies can be calculated (e.g., for each data set, there are ten estimates of F at age 4 to be compared with corresponding true values). The discrepancies are presented as frequency distributions in Tables 3.1, et seq.

It should be noted that in some of the frequency distributions of percentage discrepancies, the frequencies do not add to 10. There are reasons for this:

a) True values of N at age were truncated to the nearest million by the program producing the frequency distributions. In simulations incorporating high mortality rates, the true number in the sea sometimes becomes less than 0.5 million at high age. In this case, the truncated value is zero and it is, therefore, not possible to calculate a percentage discrepancy.

b) Some of the assessment methods estimated values of zero or infinity for fishing mortality rates (and associated catchabilities). Such values were not included in the frequency distributions.

c) In the case of the Collie-Sissenwine and Time Series methods, it was possible in the time available only to make estimates of parameters in one last data year. The frequency distributions in these cases, therefore, consist of only one frequency of unity.

Some assessors found it impossible in the time available to apply their allocated method to some of the data sets and in these cases the associated table of histograms is blank. Estimates which were ignored or non-computable for the reasons described above were also excluded when calculating mean logarithmic ratios and associated root mean square deviations referred to below.

ii) Indicators of bias and precision of the estimates were calculated.

The mean of the logarithms of the ratio of estimate to truth was calculated as a measure of bias in the estimates.

The logarithmic transformation was adopted to reduce the effect of estimates which departed widely from truth. Lower absolute values indicate less biassed results.

$$MLR = 1/10\Sigma\,[\ln(\text{Estimate}) - \ln(\text{Truth})]$$

The root mean square of the logarithms of the ratio of estimate to truth was calculated as an indicator of the precision of the estimates. Lower values indicate more precise results.

$$RMS = \left[1/10\Sigma[\ln(\text{Estimate}) - \ln(\text{Truth})^2]\right]^{1/2}$$

Values of 1000MLR and 100RMS are presented in Tables 3.2, 3.3, *et seq.*

In the time available, it was not possible to perform the above-mentioned analyses on estimates of catchability.

To present the true values required to carry out the calculations indicated above would require a prohibitively large number of tables. Copies of the true values can be obtained on IBM-formatted disk from D.W. Armstrong or G. Stefánsson at the addresses shown in Section 3.1.

3.2.2 Problems with the simplified procedure

The procedure adopted is, from the statistical point of view, less satisfactory than the full Monte Carlo approach in that the successive data sets are not statistically independent even though they are analyzed separately and the number of estimates achieved (10) is too small for precise statistical conclusions to be drawn. However, since the important factor to be investigated is the relative performance of the methods, statistical independence between trials is probably not a crucial point.

3.3 Estimation of Historical Trends in Simulated Data Sets

The description of the current state of the stock is a very important product of stock assessment techniques but the utility of this information is greatly enhanced by the perspective on the historical state of the stock which assessment methods also provide. If the current state of the stock can be observed in relation to previous states, conservation advice intended to rectify immediate and longer-term problems can be provided more readily.

It is, of course, important to be confident that an assessment is not providing an erroneous impression of historical states, i.e., assessment methods should be capable of detecting changes when they exist and should not suggest the existence of changes which have not occurred. This aspect is particularly important for results for years close to the last data year because of the greater influence which they will exert in deciding on changes required in the future in the state of the stock.

To investigate this aspect of assessment methodology, the assessors were also requested to present an assessment for the whole of the 30-year period of Data Sets 4 and 6. From these outputs, time series for the last 10 years of estimates of recruitment (N at age 3), spawning biomass, and F for ages 5-9 were plotted. True values of these quantities were plotted on the same graphs to allow comparison between estimates and truth. In addition, the estimate of each quantity obtained as a last-data-year value, as described in Section 3.2, was also plotted.

3.4 Estimation of Parameters in Last Data Year for Real Data Sets

As stated in Section 2.2, application of different methods to the same data set has, on some occasions, produced rather different and confusing results. It was, therefore, decided to apply the methods implemented at this Workshop to real data sets to demonstrate the kind of differences which can arise.

The assessors were provided with real data sets for North Sea cod and haddock comprising catch at age for commercial and research vessels, associated mean weight at age, fishing effort where available, and estimates of natural mortality rate and proportion mature at age.

The assessors were requested to carry out an assessment using each of these data sets and to record their estimates for 1986 (the last data year) of N at age, mean F at ages 5-9, spawning biomass and total biomass.

A summary of the data available for each stock is given in the text table below. As with the simulated data, no tabulation of the data sets are included in this report. Copies may be obtained from D.W. Armstrong or G. Stefánsson at the addresses indicated in Section 3.1.

Fleet	Cod	Haddock
England Seine	*	
England Trawl	*	
Scotland Seine	*	*
Scotland Trawl	*	*
Scotland Light Trawl	*	*
Scotland *Nephrops* Trawl	*	*
Other nations all gears	*	*
Int. Young Fish Survey	*	*
English Groundfish Survey	*	*
Dutch Groundfish Survey	*	
Scottish Groundfish Survey	*	*

4 INTERPRETATION OF RESULTS

Because it was necessary to analyze Data Sets 5 and 6 during the meeting, relatively little time could be spent discussing the results of the analyses. The interpretation presented below is an attempt to reflect the points raised in discussion, but also includes other suggestions received by correspondence or which became apparent during the writing of the report.

4.1 Estimates of Parameters in the Last Data Year of Simulated Data Sets

4.1.1 Frequency distribution of percentage deviations from truth

Data Sets 1-4

For Data Sets 1-4, most of the methods performed well. Most of the estimates of N at age and F at age are within 30% and many of them are within 10% of the true values. This result is to be expected given the low variance of the data in these sets. In addition, many of the methods assume log-normal errors and/or changes in catchability following an exponential function, and both of these properties are included in these data sets.

However, even on these excellent data, all of the methods can produce estimates which depart widely from truth, especially at the higher ages. Greater attention to any available diagnostics would probably have resulted in improved results, but careful handling of F and/or catchability at high age is clearly indicated.

Results for the current version of Extended Survivor Analysis (XSA) demonstrate trends with age in Data Sets, 1, 2, and 4. A similar problem exists with results from the General Linear Model (GLM) for Data Sets 3 and 4. Both of these methods are still under development and problems of this type may be resolved in the future.

A note of caution should be given about the results of the CAGEAN analysis of Data Sets 1-3. As explained more fully in Annex 2, these results are possibly better than they should be since they are conditioned by prior knowledge obtained by running the method on the full 30-year data set. The results presented for Data Set 4 are perhaps more typical of possibilities which can occur. It appears that, in this case, CAGEAN was initiated with levels of F far lower than the true values and subsequently failed to converge towards the true values.

Conventional VPA and Separable VPA, neither of which employ auxiliary data, both performed poorly on Data Sets 1-4 and failed to track changes in fishing mortality rate or numbers at age as well as the other methods. This confirms the desirability of obtaining and using auxiliary data to allow improved estimation of mortality rate and stock size in the most recent years.

However, the Time Series method applied only to total catch-at-age data and ignoring auxiliary information (TSER1) also performed well. Unfortunately, only one set of parameters was estimated by this method for these data sets, but the results suggest that this method may be worth considering if auxiliary data are not available. The performance of the Time Series method appears to be improved if auxiliary data are included in the analysis (TSER2).

Estimates of total biomass, spawning biomass, and mean F tended to cluster closer around true values than did the estimates of N at age and F at age. This is probably because the biomass and mean F values are aggregates over age groups and errors at age tend to cancel.

Data Sets 5 and 6

Estimates of N at age and F at age are much less closely clustered around the true values as expected given imprecise data which do not comply with the assumptions of the analytical methods.

Trends in the results for N and F at age are still evident for the Extended Survivors and General Linear Model methods (XSA and GLM). CAGEAN performed better on these data sets than on Set 4 perhaps because the initiating value of F used was reasonably close to the true value.

Comparison of the results from the Armstrong-Cook methods indicates a possible advantage in using a logarithmic transform in that AC1 and AC2, which use log-transformed data, performed better than AC3 and AC4 which use untransformed data.

4.1.2 Bias and precision indicators (MLR and RMS)

Because of limited time, no interpretation was attempted at the meeting of MLR and RMS of the N- and F-at-age data, but subsequent inspection of these results revealed nothing that has not already been referred to in Section 4.1.

During the meeting, a preliminary attempt was made to rank the methods in order of performance. This procedure was confined to results from Data Sets 5 and 6 since these were considered to be the most realistic sets. Within the results from each data set, the methods were ranked according to the values of bias and precision indicators calculated for mean F for ages 5-9 and for spawning biomass. The latter quantities were selected since they are formed by aggregating over age groups and thus may represent a more reasonable representation of the overall performance of the methods than analogous rankings on an age-by-age basis. The rankings are shown in Table 4.1.

Subsequent to the meeting, the ranking procedure was modified and extended to all data sets. A 2-way classification is presented in which methods are assigned to intervals of both MLR and RMS. The results of the modified procedure are shown in Tables 4.2-4.13.

Methods listed in the top left-hand area of the tables exhibit better performance.

For Data Sets 1-4, the 2-way tables confirm the generally poor performance of Separable and Conventional VPA, although for Data Set 3, both of these methods would be judged good performers according to the criteria adopted. The problems mentioned above with extended Survivors Analysis, the General Linear Model, and CAGEAN are also reflected in these tables.

For Data Sets 5 and 6, Extended Survivors Analysis and CAGEAN are among the highest ranked performers in estimating spawning stock biomass, but perform less well in estimating mean F. Overall, the Laurec-Shepherd method exhibits the least erratic high rankings for these data sets.

It should be added that many of the participants expressed severe reservations over attempting to rank the methods in the manner indicated. It should be recalled that it was not possible to implement the full diagnostic features associated with many of the integrated methods and that these may, therefore, have performed less well than could otherwise be possible. In addition, it is by no means certain that the criteria for rankings are the most appropriate or valid.

4.2 Estimates of Historical Trends in Simulated Data Sets

4.2.1 Data Set 4: Tuning methods (Figures 4.1-4.8)

The advantage of using tuning methods when catchabilities are changing is obvious in these results. All tuning methods produced quite similar results as may be expected since the methods employed at this meeting are all variations on the same theme.

HYBRID, AC1, and AC2 performed best because the trend in catchability assumed by HYBRID corresponds exactly to that used in the data simulation model, while the catchability trend assumed in AC1 and AC2 is sufficiently flexible to take a shape close to the true one. For AC3 and AC4, the assumed trend in catchability approximates less well to truth, and these methods exhibited a poorer performance.

Techniques which assume local constancy in catchability also performed less efficiently on this data set. The Laurec-Shepherd method produced biassed results, in that it tended to underestimate fishing mortality and overestimate spawning biomass. Results from AEFM and CCPUE do not exhibit this consistent bias.

4.2.2 Data Set 4: Survivors and Extended Survivors (Figure 4.9)

Survivors reproduced the major feature of the data set for early years, but underestimated fishing mortality and overestimated spawning biomass in later years.

Extended Survivors Analysis, as applied to this data set, overestimated fishing mortality and underestimated spawning biomass in the later data years.

4.2.3 Data Set 4: Integrated methods (Figures 4.10-4.13)

It was not possible to run the Time Series and Collie-Sissenwine methods on the full 30-year data set during the meeting.

All the other integrated techniques appear to have performed less efficiently than the tuning methods. CAGEAN failed to reproduce both the historical trends and the last-data-year values which perhaps implies that considerable care should be taken in choosing the quantities used to initiate this method.

ADAPT produced better results when a trend in catchability was taken into account, but even in this case, the results were poorer than those produced by tuning methods. The GLM method reproduced the early years' historical trend reasonably well, but underestimated mean F and overestimated spawning biomass in the later years.

4.2.4 Data Set 4: Conventional and Separable VPA (Figures 4.14-4.15)

In both cases, the effects of convergence of the VPA can be observed, in that the estimates correspond well to truth in the earlier data years, but less well in the later years. In fact, true catchabilities (and hence fishing mortalities) were increasing. These methods tended to underestimate the fishing mortality in the last data year and hence overestimated biomass.

4.2.5 Data Set 6: Tuning methods (Figures 4.16-4.21)

None of the methods produced really satisfactory results. The main features of the time series are reproduced by AC1, LS, and, to a lesser extent, CCPUE, but these and all other tuning methods erroneously estimated a sharp reduction in F in the last data year. This was because, by chance, the CPUE estimates in the last data year for three of the fleets which had, until then, provided the most reliable data were subject to large positive measurement error which resulted in the underestimation of fishing mortality. Such a result would be very unfortunate in a real assessment since it would indicate a better situation than that which actually exists.

Techniques such as HYBRID, which permit catchability changes in all fleets will probably always perform poorly on data sets such as this where the level of noise is high and, consequently, the estimation of the parameters descriptive of trends is difficult. Difficulties are also encountered when the assumptions implicit in the analytical method (e.g., probability distribution of errors, functional form of catchability trends, assumption of separability) do not conform to truth. This is the case for all of the tuning methods applied to this data set.

Probably the safest approach in these circumstances is to employ one of the more constrained techniques. If it is thought (or if diagnostics can indicate) that changes in catchability are not important for any fleet in recent years, methods such as LS seem appropriate. If recent years' catchability can be assumed constant only for some fleets, mixed methods such as AC1 and AC2 may provide a reasonable approach.

4.2.6 Data Set 6: Survivors and Extended Survivors (Figure 4.22)

Survivors tended to overestimate fishing mortality and underestimate spawning biomass. (Reference to diagnostics on the results obtained identified this problem and indicated that one of the research vessel surveys had produced data of very high variance which should be excluded from the analysis.) The Extended Survivors Analysis gave good results for this data set.

4.2.7 Data Set 6: Integrated methods (Figure 4.23-4.25)

It was not possible to apply the Collie-Sissenwine method to this data set and, of the time series methods, only TSER1 (omitting the use of auxiliary data) could be implemented.

TSER1 performed efficiently on this data set and estimated fishing mortality and biomass in the last data year with no important discrepancy from the true values. This is, at least partly, because TSER1 does not use auxiliary data and was, therefore, not affected by the misleading CPUE values for the last data year which created problems for the tuning methods. All other integrated methods, which make use of auxiliary data, underestimated fishing mortality in the last data year.

4.2.8 Data Set 6: Separable VPA (Figure 4.26)

This method produced satisfactory results purely because the arbitrarily chosen inputs to initiate the computations happened to approximate close to truth.

4.3 Applications to Real Data Sets

Estimates of numbers at age, F at age, total and spawning stock biomass, and mean F for 1986 for North Sea cod and haddock are given in Tables 4.3.1 and 4.3.2, respectively. (No estimates are available for seven of the methods tested at this meeting - see tables for details.)

Estimates of these parameters made by the 1988 North Sea Roundfish Working Group are also included in the tables for comparison. The North Sea Roundfish Working Group's data base included data for 1987, and estimates of F at age and associated N at age for that year were obtained for fish of ages greater than 1 by the Laurec-Shepherd method. The results shown in the tables for 1986 are derived by VPA from the estimates for 1987.

The Collie-Sissenwine method produced implausible results. Estimates of F for cod were either very high (age 2) or very low (other ages) when compared with recent historical values obtained by the Roundfish Working Group. No estimate of F was obtained for many age groups of haddock because this method estimates values of N at age less than the observed catch.

Results for CAGEAN and Survivors were more plausible and it would be difficult to demonstrate that they were not correct. However, the results are in many cases, very different from those obtained by the Roundfish Working Group both for 1986 and for other recent years. This is particularly the case for the results from CAGEAN for haddock where the estimated values of F are low and corresponding values for N are high. It is doubtful that the Roundfish Working Group would accept such estimates.

The range of results from the *ad hoc* method exemplified the difficulties encountered by the Roundfish Working Group in deciding on final estimates of F and N at age in the last data year. In many cases, the estimates obtained are in reasonable agreement. However, occasional "wild" values occur (e.g., high estimates of F at age 3 and 4 for haddock when using AC2) and it is difficult to select the results of any one of these methods as being the best.

Estimates of F and N at age are most variable for the youngest age groups (0 and 1 for haddock, 1 for cod). This indicates the continued requirement mentioned in Section 2.1 to use additional methods to estimate these values.

4.4 General Comments

None of the variants of *ad hoc* tuning is obviously preferable in all circumstances to any of the others. This is not surprising since, as stated previously, all the variants tested are closely related. The Laurec-Shepherd and Hybrid methods are the longest established of the tuning variants and diagnostic outputs are well developed for these methods. The Laurec-Shepherd method generally has lower prediction error (RMS) and higher bias (MLR) than the Hybrid method when there are strong changes in catchability for some fleets and generally appears to be more robust, in line with theoretical expectations. In practice, however, examination of diagnostics often leads to reformulation of the method. An example of this is referred to in the last paragraph of Section 1 of Annex 2 where an analysis was initiated using the Laurec-Shepherd method, but the final formulation incorporated a mixture of that method and the Hybrid method allowing for trends in catchability in some fleets and constant catchability in others. Where such procedures are required, there would be considerable benefit from obtaining good standardized commercial effort data or survey data so that catchability can unambiguously be held constant for as many fleets as possible in a mixed analysis.

The integrated methods have a more respectable statistical basis than the *ad hoc* methods in that integrated methods utilize standard and generally accepted statistical methods for parameter estimation. The properties of these estimators are understood, at least asymptotically, and some approximations for their precision are available. Furthermore, most of the integrated methods produce copious diagnostic statistics and, especially in the case of the adaptive framework, users are encouraged to modify their model specification in the light of diagnostic outputs.

Judging by their performance at this meeting, the integrated methods seem to be intermediate in performance among the tuning variants and no major advantage in using integrated methods was demonstrated. However, as previously, in the time available, it was not possible to make full use of diagnostic features. In all cases, it was necessary to choose a model specification *a priori* and to produce results dependent on this specification. For this reason, many of the applications of the integrated methods incorporated misspecified models (e.g., assuming constant catchability, separability, etc. for data sets where such assumptions were not valid). In these circumstances, it is perhaps surprising that integrated methods did well at all.

The integrated methods are computationally much more demanding than the VPA-based methods and lengthy run times may not be able to be accommodated in the ICES working group environment unless some means can be found for extending the time available to carry out the required assessments. The main difference between integrated and *ad hoc* methods is that the former are capable of allowing for errors in the total catch-at-age data. For stocks where these errors are smaller than the errors in the commercial CPUE and survey series, the extra complexity and effort involved in implementing integrated methods may not be worthwhile in terms of parameter estimation.

At present, therefore, there is no indication that any of the methods which use auxiliary data clearly and consistently performs much better than any of the others. It has yet to be demonstrated that full implementation of integrated methods produces enhanced results. Equally, it has not yet been demonstrated that, except on the grounds of computational speed, it is preferable to use *ad hoc* methods. Further testing of both types of method against realistic data sets (e.g., Data Sets 5 and 6) is clearly required before decisions can be made on which type of method is preferable. Finally, it was suggested that modifications of some of the integrated methods may be desirable. In particular, CAGEAN may perform better if initial parameter estimates are obtained using an *ad hoc* method.

5 FUTURE TESTING OF ASSESSMENT METHODS

Testing of methods, as performed at this meeting, was based on studying how estimation procedures behave on simulated data sets. This procedure could serve as the general approach to verifying new methods before they are applied for assessment of real fish stocks.

The approach which has been taken when simulating data sets is:

a) define a plausible underlying deterministic model to describe the fishery;

b) stochastically perturb (some of) the parameter values incorporated in this model, i.e., add process error to the underling parameter values to produce realized parameter values;

c) produce catch-at-age and effort data associated with the realized parameter values;

d) add measurement error to catch-at-age and effort data.

The realized parameter values are regarded as "truth". The efficiency of an assessment method is tested by how well it estimates a subset of the realized parameters.

When applying an assessment method to a data set, it is believed, at least temporarily, that the underlying fisheries model is known and that the method is appropriately specified with respect to process and measurement error (or perhaps to the combination of both types of error). However, even if this is the case, increased errors will increase the difficulty in obtaining good parameter estimates. Futhermore, within an assessment method the specfication of the underlying fisheries model or of the probability density functions of the errors may be incorrect. If this is the case, the estimation of parameters may also be adversely affected.

One possibility for quantifying the effects of the factors referred to above is to test each method against a set of simulated data organized as a factorial design. One such design is indicated in the text table below.

	Test no.											
	1	2	3	4	5	6	7	8	9	10	11	12
Measurement error												
None	*	*					*	*				
Correct specification			*	*					*	*		
Incorrect specification					*	*					*	*
Underlying model												
Correct specification	*		*		*		*		*		*	
Incorrect specification		*		*		*		*		*		*
Process error absent	*	*	*	*	*	*						
Process error present							*	*	*	*	*	*

Such an approach is attractive, but it should be recognized that it could be very labour-intensive since multiple runs would be required within those tests incorporating measurement or process error so that the effects of increasing level of error could be evaluated. In addition, since no method can be expected to perform well in all circumstances, it would probably be necessary to subject each method to the tests above for each of a number of types of fishery.

Furthermore, with such an approach, it is difficult to define a single incorrectly specified underlying model. This is because the model for simulating the data and the model implicit in an assessment method are both comprised of various sub-models. The specification of any of these sub-models in the simulation and in the assessment may or may not differ.

Similarly, it is also difficult to define an appropriate "incorrect" probability density function for measurement and/or process errors. (Most assessment methods assume that the measurement errors are normally or log-normally distributed, and it was suggested that the gamma distribution could be used as the incorrect specification.) Further thought needs to be given to these problems by the Methods Working Group.

An alternative suggestion on the future testing of methods was that a number of standard data sets could be created against which new and existing methods could be tested so that a preliminary ranking of methods can be obtained. The Group recognized that Data Sets 1-4 produced for this meeting are not suitable for this purpose. Data Sets 5 and 6 offer a more stringent test and may serve in the immediate future as standard sets. However, more thought needs to be given to producing appropriate data sets against which to test assessment methods. One possibility in this context is that the simulated data might be based on the fishery for which the method is intended. Few, if any, fisheries have been modelled with respect to creating a realistic error structure in the observations (as compared to adding errors derived from some conventional probability density function). In particular, it might be advantageous to produce the estimated catch-at-age data by simulating the biological sampling procedures used on that fishery. This should add measurement error of more or less the correct statistical form.

One of the major aspects of a good method is its ability to detect, by means of good diagnostics, when unreliable parameter estimates are being produced. Whatever method of testing is finally decided upon, the Group suggests that, wherever possible, the estimated variance-covariance matrix of the parameter estimates should be presented as the basis for an efficient set of diagnostics. In addition, serial correlations in the differences between the observations and their fitted values should also be made available along with the variances of the residuals for each age group. (It is recognized that this may be difficult in the case of *ad hoc* methods.) Variances of residuals for each year and for each fleet should also be made available to provide the user with hints, e.g., of badly sampled fleets, the data for which can then be down-weighted. These outputs should be arranged as a year-by-year table for each fleet.

In future testing, it would be useful to categorize methods according to their two components, i.e., estimation procedure and model specification, and to test these separately. With respect to estimation procedure, the methods examined fall into two broad categories, i.e., statistically-founded approaches and *ad hoc* approaches. It is possible that certain *ad hoc* estimation procedures correspond to realizations of statistically-founded procedures and clarification of this possibility is required. With respect to model specification, there is a varying degree of flexibility among the methods tested, and opinions ranged from advocating complete flexibility to specifying a single model. The success of a flexible approach hinges on the adequacy of diagnostics to define appropriate models, while a single model approach relies on the robustness of the specified model. Attempts should be made to determine whether, given the same underlying model, the statistically-founded approach works better or worse than the *ad hoc* approach and thereby discriminate between estimation procedure and model formulation.

The Group is also of the opinion that, since there is already a proliferation of new methods, authors should restrain themselves from publicizing new methods until they can demonstrate that some real advantage can be gained from their use.

Finally, it should not be forgotten that the ability to estimate the current and historical state of the stock is only one part of the assessment process. The desired end product of an assessment is often advice on an appropriate total allowable catch and this requires methods to predict how changes in fisheries will affect stock size and yield. This aspect of assessment was not dealt with during the meeting. It is, however, of considerable importance and should be the topic of future meetings of the Methods Working Group.

Table 2.1 Assumptions of the Methods

Method	HYBRID	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVIV	XSA	CAGEAN	ADAPT	GLM	COLSIS	TSER1	TSER2	SVPA	CONVEN
Separable model for fishing mortality	No	Each fleet	Some fleets	Some fleets	Some fleets	Some fleets	Each fleet	Each fleet	Yes	Some fleets	Each fleet	No	Each fleet	No	No	No	Sum of fleets	No
Time trends - catchability											Step-wise				Markov	Markov	n.a.	n.a.
Regress log(q) vs year	Yes																	
Regress log(q) vs log(yr)			Yes	Yes														
Regress q vs year					Yes	Yes												
Assumed absent		Yes					Yes	Yes	Yes	Yes		Yes	Yes	Yes				
Assumes existence of process error	<-------------- Not explicit ------------>						Not explicit		No	No	No	No but can do	Not explicit	Yes	Yes	Yes	Yes-added to catch error	
Assumes no Error in Catch-at-Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	log-normal	log-normal	log-normal	log-normal	log-normal	No	log-normal	log-normal	log-normal	Yes
Weight for fleets										No					n.a.	n.a.	n.a.	n.a.
1 / Var(q)	Yes	Yes	Yes	Yes	Yes	Yes												
1 / Var(N)								Yes										
1 / Var(residuals)											Yes	Yes	Yes					
1 / Var(survivors)									Yes									
Externally specified														Yes				
F Constrained for Older Ages	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes	No	Yes	Yes	Yes	Yes
Can handle multiple fleets CPUE data	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	n.a.	n.a.
Down-weights early-years data	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	n.a.	Yes	n.a.	n.a.	No	n.a.	No	No	n.a.	n.a.
Estimate of CPUE in last data year assumed exact	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No	n.a.	n.a.
Estimate of catch-at-age in last data year assumed exact.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes in VPA calcns	Yes in VPA calcns	No	Yes	n.a.	No	No	No	Yes in VPA calcns	Yes

Table 3.1 : Simulated Data Set 1 : Frequency Distributions of Percentage Deviation of Estimates of N at age from True Values

Methd: HYBRID										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50									1	
30	1							3	1	2
<10	8	10	10	10	9	9	8	3	5	4
-30	1				1	1	2	4	3	3
-50										1
-70										
< -70										

Methd: LS										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50									1	
30								1	1	1
<10	10	10	10	10	10	10	10	8	8	6
-30								1		3
-50										
-70										
< -70										

Methd: AC1										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50									1	
30	1						1			1
<10	8	10	10	10	10	10	9	8	8	5
-30	1							2	1	4
-50										
-70										
< -70										

Methd: AC2										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50									1	
30	1						1		1	1
<10	8	10	10	10	10	10	9	8	7	5
-30	1							2	1	4
-50										
-70										
< -70										

Methd: AC3										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50									1	
30	1						1		1	2
<10	9	10	10	10	10	10	9	9	8	4
-30								1		4
-50										
-70										
< -70										

Methd: AC4										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50									1	
30	1						2	1	1	1
<10	9	10	10	10	10	10	8	8	7	5
-30								1	1	4
-50										
-70										
< -70										

Methd: AEFM										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50									2	2
30	5							1	2	3
<10	4	9	9	9	8	9	8	5	3	
-30	1	1	1	1	2	1	2	3	3	5
-50								1		
-70										
< -70										

Methd: CCPUE										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										1
30									1	2
<10	9	10	10	10	10	10	8	8	5	3
-30	1						2	2	4	4
-50										
-70										
< -70										

Methd: SURVIV										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30	2								2	5
<10	7	10	10	10	10	10	10	10	8	4
-30	1									1
-50										
-70										
< -70										

Methd: XSA										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10	3	1								
-30	7	9	10	10	10	8	2	1		
-50						2	8	8	8	2
-70								1	2	8
< -70										

Methd: CAGEAN										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										1
<10	10	10	10	10	10	10	10	10	8	7
-30									2	2
-50										
-70										
< -70										

Methd: ADAPT										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30	2						2		1	3
<10	6	10	10	10	10	9	7	9	7	2
-30	2					1	1	1	2	4
-50										1
-70										
< -70										

Methd: GLM										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30	2		1	1				1	6	1
<10	7	10	9	9	10	10	9	7	4	5
-30	1						1	2		4
-50										
-70										
< -70										

Methd: COLSIS										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10								1		1
-30		1		1	1	1	1		1	
-50										
-70										
< -70			1							

Methd: TSER1										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10			1		1					
-30		1		1		1	1	1	1	
-50										
-70										
< -70										

Methd: TSER2										
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10		1	1	1						
-30					1	1	1	1	1	
-50										
-70										
< -70										

Methd: SVPA										
Age	3	4	5	6	7	8	9	10	11	12
> 70	3	4	2	1	2	2	1	3	5	2
70		3	3	4	2	3	4	3		1
50	3		2	2	3	2	1			2
30							1	1	1	2
<10	1	1	1	2	2	1	2	2	3	2
-30	2	2	2	1	1	2	1			1
-50	1							1	1	
-70										
< -70										

Methd: CONVEN										
Age	3	4	5	6	7	8	9	10	11	12
> 70	4	5	4	4	3	5	3	4	5	4
70	1		1		2		3	1		1
50	1	2	1	3	2	1		1		
30	1		1			1				1
<10			1			1	1	1	2	1
-30		1		1	1		1	1	1	1
-50	1	1	1	1	1	1	1	1	1	1
-70	1	1	1	1	1	1	1	1	1	1
< -70	1									

Table 3.2 : Simulated Data Set 1 : Mean Log Ratio of Estimates of N at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	-1	-1	-1	-1	-2	-2	-2	-4	-1	-7
LS	1	1	0	0	-0	-1	1	-3	4	-6
AC1	0	0	0	0	-0	-1	2	-3	3	-9
AC2	0	0	0	0	-0	-1	2	-3	4	-9
AC3	2	1	0	1	0	1	3	-0	6	-6
AC4	1	1	1	1	-0	-0	3	-1	5	-7
AEFM	6	-3	-2	-2	-4	-3	-2	-10	4	1
CCPUE	1	-2	-2	-2	-2	-4	-2	-8	-3	-0
SURVIV	3	2	2	2	2	2	4	2	5	5
XSA	-13	-14	-17	-21	-26	-32	-40	-54	-71	-87
CAGEAN	0	-0	-0	-0	-2	-2	-2	-3	-4	-2
ADAPT	-1	-1	-2	-1	-2	-2	-2	-3	-5	-7
GLM	-1	1	-1	-0	-1	-0	-2	-1	10	-6
COLSIS		-12	-241	-17	-16	-13	-13	-7	-19	0
TSER1		-31	-11	-15	-6	-14	-14	-17	-13	
TSER2		-11	-5	-11	-12	-14	-14	-17	-13	
SVPA	28	35	31	28	27	29	28	29	30	28
CONVEN	29	37	28	28	25	26	24	23	18	18

Table 3.3 : Simulated Data Set 1 : Root Mean Square Log Ratio of N at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	8	6	6	6	7	7	7	14	14	19
LS	4	3	2	3	3	3	4	9	11	15
AC1	6	3	2	2	2	3	6	8	11	16
AC2	6	3	2	2	2	3	6	8	12	16
AC3	6	3	3	3	3	4	6	6	12	18
AC4	6	3	3	3	3	3	7	8	13	17
AEFM	12	6	5	5	8	7	7	19	18	22
CCPUE	5	3	4	4	4	5	6	11	11	19
SURVIV	9	5	4	3	3	3	5	6	8	11
XSA	14	15	17	21	26	33	40	55	72	89
CAGEAN	4	2	2	2	3	3	3	5	7	8
ADAPT	9	6	6	4	5	7	11	7	10	20
GLM	9	5	6	5	5	6	6	11	13	14
COLSIS		12	241	17	16	13	13	7	19	0
TSER1		31	11	15	6	14	14	17	13	
TSER2		11	5	11	12	14	14	17	13	
SVPA	61	49	43	42	39	43	40	47	49	40
CONVEN	89	71	63	63	60	62	60	66	66	61

Table 3.4 : Simulated Data Set 1 : Frequency Distributions of Percentage Deviation of Estimates of F at age from True Values

Methd: HYBRID — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70	1									
50										2
30	1	1	2	2	2	5	2	4	2	4
<10	7	9	7	8	8	5	8	4	7	1
-30	1		1					2	1	3
-50										
-70										
< -70										

Methd: LS — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50	1							1		
30	1		1			1	1	1	1	5
<10	5	10	8	10	10	9	8	7	7	2
-30	3		1				1	1	2	3
-50										
-70										
< -70										

Methd: AC1 — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70	1									
50										
30	2		1			1	1	4	2	6
<10	4	10	8	10	10	9	8	5	5	3
-30	3		1				1	1	3	1
-50										
-70										
< -70										

Methd: AC2 — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70	1									
50										
30	2		1			1	1	4	2	6
<10	4	10	8	10	10	9	8	5	5	3
-30	3		1				1	1	3	1
-50										
-70										
< -70										

Methd: AC3 — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50	1									
30			1				1	2	1	5
<10	6	10	8	10	10	10	8	7	6	3
-30	3		1				1	1	3	2
-50										
-70										
< -70										

Methd: AC4 — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50	1							1		
30	1		1			1	1	2	2	5
<10	5	10	8	10	10	9	7	6	5	3
-30	3		1				2	1	3	2
-50										
-70										
< -70										

Methd: AEFM — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70								2		1
50							1	1		2
30	3	2	3	1	1	4	1	2	3	1
<10	3	8	7	9	9	6	8	4	4	4
-30	4							1	3	
-50										2
-70										
< -70										

Methd: CCPUE — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70	1									1
50								2		1
30	1		2		1	4	2	1	2	2
<10	5	10	8	10	9	6	8	6	6	4
-30	3							1	2	1
-50										1
-70										
< -70										

Methd: SURVIV — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70	1									
50										
30	1		1					2	1	
<10	4	9	9	9	10	10	10	7	6	8
-30	4	1		1				1	3	2
-50										
-70										
< -70										

Methd: XSA — Age	3	4	5	6	7	8	9	10	11	12
> 70							2	8	10	10
70	1					5	8	2		
50		1	1	2	9	5				
30	5	8	9	8	1					
<10	4	1								
-30										
-50										
-70										
< -70										

Methd: CAGEAN — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50	1									
30	5	1	1	1			3	2	3	2
<10	1	9	9	9	10	10	7	5	5	8
-30	3							3	2	
-50										
-70										
< -70										

Methd: ADAPT — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70	1									
50							1		1	1
30	1	2	1	2	2	4	1	5	1	2
<10	5	8	9	8	8	6	7	3	8	7
-30	3						1	2		
-50										
-70										
< -70										

Methd: GLM — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70	1									
50										
30	3	1	2	3	3		3	2	2	4
<10	3	7	6	5	7	10	7	5	1	6
-30	3	2	2	2				3	7	
-50										
-70										
< -70										

Methd: COLSIS — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50				1						
30		1	1		1	1	1			
<10								1	1	1
-30										
-50										
-70										
< -70										

Methd: TSER1 — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50		1							1	
30			1	1	1	1		1		
<10							1			
-30										
-50										
-70										
< -70										

Methd: TSER2 — Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50									1	
30		1	1	1	1	1	1	1		
<10										
-30										
-50										
-70										
< -70										

Methd: SVPA — Age	3	4	5	6	7	8	9	10	11	12
> 70	1									
70								1		
50	2	1	1	1	1	1			1	
30		1	1	1		1	1			1
<10		1	1	1	2	1	2	2	3	2
-30	3				2	2	2	1	1	2
-50	1	6	6	6	5	5	5	6	3	5
-70	3	1	1	1					2	
< -70										

Methd: CONVEN — Age	3	4	5	6	7	8	9	10	11	12
> 70	2	1	2	1	1	1	1	1	1	1
70		1		1	1	1	1	1		1
50	1								1	
30			1	1	1	1	1	1	2	1
<10		1					1			1
-30	2		1	1	1	2		1	1	1
-50	2	3	3	3	3	2	4	3	2	3
-70	2	4	3	3	3	3	2	3	3	2
< -70	1									

Table 3.5 : Simulated Data Set 1 : Mean Log Ratio of Estimates of F at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	3	2	2	2	3	5	4	6	1	8
LS	1	0	1	0	0	4	-1	6	-4	6
AC1	1	1	0	0	0	4	-1	6	-3	7
AC2	1	1	0	0	0	3	-1	6	-3	7
AC3	-1	-0	-0	-1	-0	2	-3	3	-6	5
AC4	0	0	0	-0	0	3	-2	4	-5	6
AEFM	-5	4	2	2	4	6	4	14	-3	2
CCPUE	1	3	3	2	3	8	3	11	2	6
SURVIV	-1	-1	-1	-2	-2	0	-2	-0	-5	-3
XSA	15	16	20	24	30	40	48	66	83	117
CAGEAN	8	-3	-0	3	4	3	5	-3	-1	3
ADAPT	3	2	3	2	2	5	4	5	5	9
GLM	7	-4	2	1	2	1	4	-3	-12	6
COLSIS		17	11	28	26	25	16	6	10	-9
TSER1		36	23	21	22	20	6	19	33	
TSER2		22	19	21	18	20	10	19	36	
SVPA	-26	-37	-34	-31	-29	-29	-29	-31	-33	-28
CONVEN	-28	-38	-31	-30	-27	-26	-23	-23	-20	-19

Table 3.6 : Simulated Data Set 1 : Root Mean Square Log Ratio of F at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	18	7	11	7	8	12	9	17	12	21
LS	15	4	8	4	3	7	8	14	11	15
AC1	17	3	7	4	3	7	9	13	13	13
AC2	17	3	7	4	3	7	9	13	14	14
AC3	16	4	8	4	3	7	8	12	14	12
AC4	17	4	7	4	3	7	9	13	15	13
AEFM	15	7	9	6	9	9	11	25	16	31
CCPUE	16	5	7	5	5	10	9	18	13	27
SURVIV	20	7	6	5	4	6	6	9	11	12
XSA	23	17	20	25	31	41	48	67	84	119
CAGEAN	21	8	8	7	6	4	8	13	14	7
ADAPT	19	9	5	7	6	11	13	12	13	15
GLM	22	10	12	11	8	6	9	12	21	10
COLSIS		17	11	28	26	25	16	6	10	9
TSER1		36	23	21	22	20	6	19	33	
TSER2		22	19	21	18	20	10	19	36	
SVPA	57	52	51	47	43	45	42	47	52	38
CONVEN	87	74	74	71	66	67	65	69	71	60

Table 3.7 : Simulated Data Set 1 : Frequency Distributions of Percentage Deviations of Estimates of Total Biomass from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70																	3	4
70																	1	1
50																	3	1
30																		1
<¦10	8	10	10	10	10	10	10	10	10		10	10	10				1	
-30	2									10							2	1
-50																		1
-70																		1
< -70																		

Table 3.8 : Simulated Data Set 1 : Frequency Distributions of Percentage Deviations of Estimates of Spawning Biomass from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70																	3	3
70																	2	2
50																	1	1
30																	1	1
<¦10	9	10	10	10	10	10	9	10	10		10	9	10				2	
-30	1						1			7		1					1	1
-50										3								1
-70																		1
< -70																		

Table 3.9 : Simulated Data Set 1 : Frequency Distributions of Percentage Deviations of Estimates of Mean F (Ages 5-9) from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70										8								1
70										2								1
50																	1	
30	4	1	1	1		1	4	3				2		1	1	1		1
<¦10	5	9	9	9	9	8	6	7	10		10	8	8				2	
-30	1				1	1							2				2	2
-50																	5	2
-70																		3
< -70																		

Table 3.10: Simulated Data Set 1 : Mean Log Ratio of Estimates of Biomass and Mean F to True Values

Method	TSB	SSB	FBAR
HYBRID	-2	-2	4
LS	0	-0	1
AC1	-0	-0	2
AC2	-0	-0	2
AC3	1	1	-1
AC4	1	1	0
AEFM	-1	-2	6
CCPUE	-2	-3	6
SURVIV	2	2	-2
XSA	-27	-36	56
CAGEAN	-1	-2	1
ADAPT	-3	-3	5
GLM	-0	-0	-2
COLSIS			17
TSER1			20
TSER2			21
SVPA	30	28	-31
CONVEN	28	24	-24

Table 3.11 : Simulated Data Set 1 : Root Mean Square Log Ratio of Biomass and Mean F to True Values

Method	TSB	SSB	FBAR
HYBRID	7	8	10
LS	3	3	6
AC1	3	3	6
AC2	3	3	6
AC3	4	4	6
AC4	3	4	7
AEFM	4	6	10
CCPUE	4	5	9
SURVIV	3	3	3
XSA	27	37	56
CAGEAN	2	2	4
ADAPT	4	5	6
GLM	5	5	8
COLSIS			17
TSER1			20
TSER2			21
SVPA	45	41	45
CONVEN	66	62	67

Table 3.12 Simulated Data Set 2 : Frequency Distributions of Percentage Deviation of Estimates of N at age from True Values

Methd	HYBRID									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50							1	2		
30	1				1			1		
<10	7	10	10	10	9	10	8	5	7	5
-30	2						1	1		
-50									1	
-70										
< -70										

Methd	LS									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50							1	2	1	
30	2					1	1	1		
<10	7	10	10	10	10	9	7	5	7	5
-30	1						1	1		
-50										
-70										
< -70										

Methd	AC1									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50							1	2	1	
30	2				1	1	1	1		
<10	6	10	10	10	9	9	7	5	7	4
-30	2						1	1		
-50										1
-70										
< -70										

Methd	AC2									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50							1	2	1	
30	2					1		1		
<10	6	10	10	10	10	9	8	5	7	4
-30	2						1	1		
-50										1
-70										
< -70										

Methd	AC3									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50							1	2	1	
30	2		1		1	2	1	1		
<10	8	10	9	10	9	8	8	6	7	4
-30										
-50										1
-70										
< -70										

Methd	AC4									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50							1	2	1	
30	2					1	1	1		
<10	8	10	10	10	10	9	8	5	7	4
-30								1		
-50										1
-70										
< -70										

Methd	AEFM									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50								3	2	
30	2									
<10	6	8	9	7	7	6	8	6	5	4
-30	2	2	1	3	3	3	1			
-50						1	1		1	1
-70										
< -70										

Methd	CCPUE									
Age	3	4	5	6	7	8	9	10	11	12
> 70								1		
70										
50							1	1		
30								1		
<10	6	8	9	9	9	10	7	5	7	4
-30	4	2	1	1	1		2	1		
-50									1	1
-70										
< -70										

Methd	SURVIV									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50	1						1	1	1	
30	2	1						1		
<10	6	9	10	10	10	10	9	6	7	5
-30	1							1		
-50										
-70										
< -70										

Methd	XSA									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50							1	1	1	
30	1							1		
<10	5	8	6	8	8	7	7	7	6	5
-30	4	2	4	2	2	3	1			
-50							1		1	
-70										
< -70										

Methd	CAGEAN									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50								2		
30	2					2		1		
<10	7	10	10	10	10	8	10	6	8	5
-30	1									
-50										
-70										
< -70										

Methd	ADAPT									
Age	3	4	5	6	7	8	9	10	11	12
> 70						1	3	2	1	
70						1	1			
50			1		2	2		1	1	
30	2	1		1	4			1		
<10	6	8	9	8	2	4	3	3	4	5
-30	2	1		1	2	2	2			
-50							1	2	2	
-70										
< -70										

Methd	GLM									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50							1			
30	1	1		1						
<10	5	7	8	8	10	9	7	6	6	2
-30	4	2	2	1		1	2	2		
-50								1	2	1
-70										
< -70										

Methd	COLSIS									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10			1				1			
-30		1		1	1					
-50						1				
-70										
< -70										

Methd	TSER1									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50								1	1	
30			1	1	1	1				
<10		1					1			
-30										
-50										
-70										
< -70										

Methd	TSER2									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10		1	1	1	1	1	1	1	1	
-30										
-50										
-70										
< -70										

Methd	SVPA									
Age	3	4	5	6	7	8	9	10	11	12
> 70	6	2	3	3	1	1	2	1	2	
70	1	4	2	1	1	1	1			
50		3	3	3	5	4	3	3	1	
30	2	1	2	3	2	4	4	1		
<10	1				1			4	5	5
-30										
-50										
-70										
< -70										

Methd	CONVEN									
Age	3	4	5	6	7	8	9	10	11	12
> 70	3						1		1	
70							1			
50	2	6	5	3	3	3	1	4	1	
30	2	4	5	6	6	5	1	1		
<10	2			1	1	2	6	4	6	4
-30	1									
-50										1
-70										
< -70										

Table 3.13: Simulated Data Set 2 : Mean Log Ratio of Estimates of N at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	-2	1	3	2	0	0	3	7	-9	0
LS	-1	2	2	1	-1	2	5	7	5	0
AC1	-1	1	3	2	2	3	4	7	5	-14
AC2	-1	1	3	1	0	3	3	7	5	-14
AC3	1	3	5	3	4	5	8	10	5	-14
AC4	-0	2	3	1	0	2	7	7	5	-14
AEFM	-1	-4	-3	-7	-6	-8	-8	11	1	-14
CCPUE	-7	-5	-3	-5	-4	-0	-0	13	-9	-14
SURVIV	5	1	1	-3	-1	2	5	4	5	0
XSA	-5	-5	-5	-7	-6	-6	-4	7	-4	0
CAGEAN	-1	-0	1	-0	1	3	0	10	0	0
ADAPT	0	-2	3	-1	12	13	22	22	1	0
GLM	-4	-2	0	-1	-4	-4	1	-9	-17	-23
COLSIS		-14	-11	-24	-25	-40	0			
TSER1		7	13	17	16	17	10	29	29	
TSER2		4	3	7	4	5	0	0	0	
SVPA	54	44	45	39	32	32	37	22	22	0
CONVEN	33	27	26	22	18	20	18	19	14	-14

Table 3.14: Simulated Data Set 2 : Root Mean Square Log Ratio of N at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	10	6	5	4	5	0	15	19	25	0
LS	9	5	5	3	4	5	15	19	14	0
AC1	10	5	5	3	4	6	15	19	14	31
AC2	9	5	5	3	2	6	15	19	14	31
AC3	9	6	6	5	5	8	15	18	14	31
AC4	9	5	5	4	2	6	14	19	14	31
AEFM	12	9	6	9	11	16	23	20	32	31
CCPUE	13	7	5	7	7	4	15	33	25	31
SURVIV	15	8	5	4	5	4	13	17	14	0
XSA	14	9	9	11	9	10	26	15	28	0
CAGEAN	10	6	4	3	2	6	0	18	0	0
ADAPT	12	9	12	9	20	33	56	57	54	0
GLM	12	9	7	7	5	6	16	17	35	40
COLSIS		14	11	24	25	40	0			
TSER1		7	13	17	16	17	10	29	29	
TSER2		4	3	7	4	5	0	0	0	
SVPA	65	46	49	42	36	35	44	32	38	0
CONVEN	46	28	28	24	21	21	30	26	28	31

Table 3.15 : Simulated Data Set 2 : Frequency Distributions of Percentage Deviation of Estimates of F at age from True Values

HYBRID

Age	3	4	5	6	7	8	9	10	11	12
> 70	1									1
70									2	
50							1	2		
30	1	1	1		1	2	2	1	2	3
<10	2	6	9	10	8	8	7	4	4	4
-30	6	3			1			3	1	1
-50									1	1
-70										
< -70										

LS

Age	3	4	5	6	7	8	9	10	11	12
> 70	1								1	
70										
50									1	1
30	1	1	1		1	2	3	3	2	1
<10	2	7	9	10	8	8	7	3	4	3
-30	6	2			1			4	2	4
-50										1
-70										
< -70										

AC1

Age	3	4	5	6	7	8	9	10	11	12
> 70	1								1	
70										
50							1		1	
30	1	1	1		1	2	2	2	2	2
<10	2	6	9	10	8	7	7	5	4	7
-30	6	3			1	1		3	2	1
-50										
-70										
< -70										

AC2

Age	3	4	5	6	7	8	9	10	11	12
> 70	1								1	
70										
50							1	2	1	
30	1	1	1		1	2	2		2	3
<10	2	7	9	10	8	7	7	5	4	6
-30	6	2			1	1		3	2	1
-50										
-70										
< -70										

AC3

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70	1									
50									1	
30	1		1				2		2	2
<10	2	7	9	10	9	8	8	3	4	7
-30	6	3			1	2		7	3	1
-50										
-70										
< -70										

AC4

Age	3	4	5	6	7	8	9	10	11	12
> 70	1									
70									1	
50							1	1	1	
30	1	1	1		1	2	1	1	2	2
<10	2	7	9	10	8	8	8	4	3	8
-30	6	2			1			4	2	
-50									1	
-70										
< -70										

AEFM

Age	3	4	5	6	7	8	9	10	11	12
> 70	1					1	2			
70									1	1
50		1	1	4	2		1		1	2
30	1	1	5	1	2	6		2	1	1
<10	3	7	4	5	5	3	6	4	3	3
-30	4	1			1		1	2	2	1
-50	1							2	2	2
-70										
< -70										

CCPUE

Age	3	4	5	6	7	8	9	10	11	12
> 70	1								1	1
70							2	1		
50			1	1	1		1		1	1
30	1	1	7	4	1	4	2	2	3	3
<10	4	9	2	5	8	5	3	3	3	4
-30	4					1	2	2	1	
-50								2	1	1
-70										
< -70										

SURVIV

Age	3	4	5	6	7	8	9	10	11	12
> 70	1								1	1
70										1
50			1				1	1		
30		1	3	4	1	2	2	3	2	1
<10	3	7	5	6	7	7	4	3	5	2
-30	5	2	1		2	1	3	2	2	4
-50	1							1		1
-70										
< -70										

XSA

Age	3	4	5	6	7	8	9	10	11	12
> 70	1			1						
70			1				2	1	1	
50		1	1	1	2	1	1		1	
30		2	3	3	2	3	1	4		1
<10	4	7	5	5	5	6	4	2	4	3
-30	4				1		2	3	3	2
-50	1								1	3
-70										1
< -70										

CAGEAN

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30			1						2	3
<10	7	9	9	10	10	10	10	10	7	7
-30	3	1							1	
-50										
-70										
< -70										

ADAPT

Age	3	4	5	6	7	8	9	10	11	12
> 70	1						1	3	3	
70						1	1	1	2	
50			1			1	1			
30		2	2	2	3	1				1
<10	4	6	5	7	1	2	3	2	2	2
-30	4	2	2	1	4	2		2		6
-50	1				2	2			1	1
-70						1	4	1	1	
< -70								1	1	

GLM

Age	3	4	5	6	7	8	9	10	11	12
> 70									2	6
70									6	1
50						1		1		2
30	1	3	4	4	2	2	3	7	2	1
<10	6	5	5	5	7	6	6	2		
-30	2	2	1	1	1	1	1			
-50	1									
-70										
< -70										

COLSIS

Age	3	4	5	6	7	8	9	10	11	12
> 70				1		1				
70					1					
50			1				1			
30										
<10		1								1
-30										
-50										
-70										
< -70										

TSER1

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10					1					
-30		1	1	1		1	1	1	1	
-50										
-70										
< -70										

TSER2

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10		1	1	1	1	1	1			
-30								1	1	
-50										
-70										
< -70										

SVPA

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10									1	1
-30	2	2	4	4	4	5	3	5	6	4
-50	4	5	3	3	3	3	6	3	3	4
-70	4	3	3	3	3	2	1	2		1
< -70										

CONVEN

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30									1	1
<10	1					1	1	1	3	1
-30	2	7	9	8	8	8	8	7	6	7
-50	7	3	1	2	2	1	1	2		1
-70										
< -70										

Table 3.16 : Simulated Data Set 2 : Mean Log Ratio of Estimates of F at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	-4	-2	4	0	-3	2	7	-1	6	6
LS	-5	-3	5	1	-3	0	2	-5	8	-9
AC1	-5	-3	4	1	-3	-1	4	-5	4	4
AC2	-5	-2	5	1	-3	0	4	-4	5	5
AC3	-7	-4	2	-2	-7	-5	-2	-12	-2	-1
AC4	-5	-3	5	1	-3	-0	3	-5	1	4
AEFM	-5	4	13	14	8	13	14	-9	-6	-2
CCPUE	2	5	13	9	5	6	16	-8	6	10
SURVIV	-11	-2	7	7	-2	1	5	-1	5	-5
XSA	-1	5	16	15	8	10	11	3	-5	-31
CAGEAN	-6	0	3	-0	-1	-1	-2	1	1	4
ADAPT	-6	1	6	5	-18	-14	-20	6	10	-16
GLM	-7	1	6	2	2	6	-0	15	46	54
COLSIS		4	31	59	44	75	33			-6
TSER1		-13	-13	-15	-9	-15	-14	-26	-29	
TSER2		-7	-2	-4	0	-4	-3	-13	-16	
SVPA	-64	-51	-47	-47	-46	-45	-45	-43	-32	-42
CONVEN	-41	-32	-25	-27	-27	-25	-21	-23	-9	-18

Table 3.17 : Simulated Data Set 2 : Root Mean Square Log Ratio of F at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	25	10	7	5	9	8	13	21	28	24
LS	25	8	8	5	8	8	12	20	26	21
AC1	25	9	7	6	8	8	13	20	25	11
AC2	25	9	7	6	8	8	13	22	26	11
AC3	24	8	6	5	8	8	9	16	20	11
AC4	25	9	8	5	8	8	11	20	27	11
AEFM	31	11	16	21	19	21	29	26	33	33
CCPUE	27	10	15	15	12	11	27	33	37	34
SURVIV	33	12	13	12	10	8	18	21	26	34
XSA	32	12	22	23	17	15	27	23	31	43
CAGEAN	11	6	6	6	4	5	5	4	8	12
ADAPT	26	11	16	12	29	42	67	94	121	24
GLM	18	12	13	11	12	15	10	17	48	59
COLSIS		4	31	59	44	75	33			6
TSER1		13	13	15	9	15	14	26	29	
TSER2		7	2	4	0	4	3	13	16	
SVPA	69	55	51	52	51	49	49	48	37	47
CONVEN	44	32	26	28	29	26	24	25	16	23

Table 3.18: Simulated Data Set 2 : Frequency Distributions of Percentage Deviations of Estimates of Total Biomass from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70																	4	
70																	2	
50																	3	4
30												1					1	6
<10	10	10	10	10	10	10	10	10	10	10	10	9	10					
-30																		
-50																		
-70																		
< -70																		

Table 3.19: Simulated Data Set 2 : Frequency Distributions of Percentage Deviations of Estimates of Spawning Biomass from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70																	2	
70																	2	
50												1					3	1
30					1							1					3	9
<10	10	10	10	10	9	10	10	10	10	10	10	8	10					
-30																		
-50																		
-70																		
< -70																		

Table 3.20: Simulated Data Set 2 : Frequency Distributions of Percentage Deviations of Estimates of Mean F (Ages 5-9) from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70												2						
70												1						
50								1		1			1					
30	3	2	1	2		2	6	3	1	1		2	7					
<10	7	8	9	8	9	8	3	6	8	8	10	4	2	1		1		3
-30					1		1		1			1			1		5	7
-50																	4	
-70																	1	
< -70																		

Table 3.21 : Simulated Data Set 2 : Mean Log Ratio of Estimates of Biomass and Mean F to True Values

Method	TSB	SSB	FBAR
HYBRID	2	2	4
LS	2	2	2
AC1	2	3	1
AC2	2	2	2
AC3	4	5	-5
AC4	2	2	0
AEFM	-3	-3	7
CCPUE	-4	-3	8
SURVIV	1	1	3
XSA	-4	-4	8
CAGEAN	1	1	-0
ADAPT	2	8	22
GLM	-2	-2	15
COLSIS			3
TSER1			-19
TSER2			-7
SVPA	45	39	-42
CONVEN	26	22	-21

Table 3.22: Simulated Data Set 2 : Root Mean Square Log Ratio of Biomass and Mean F to True Values

Method	TSB	SSB	FBAR
HYBRID	3	3	10
LS	3	3	9
AC1	3	3	8
AC2	3	3	8
AC3	5	6	7
AC4	3	3	8
AEFM	4	5	13
CCPUE	5	4	14
SURVIV	3	2	10
XSA	6	6	12
CAGEAN	2	2	3
ADAPT	5	12	38
GLM	5	5	19
COLSIS			3
TSER1			19
TSER2			7
SVPA	47	42	47
CONVEN	28	23	23

Table 3.23: Simulated Data Set 3 : Frequency Distributions of Percentage Deviation of Estimates of N at age from True Values

Methd	HYBRID										LS										AC1										AC2									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70																																								
70																																								
50																			1																					
30	2				1				1		2				1		3	2	3		2								2		2								2	
<10	7	8	9	9	7	6	5	2	5		7	10	10	10	9	10	6	7	4	6	7	10	10	10	10	10	9	5	4	3	7	10	10	10	10	10	9	5	4	3
-30	1	2	1	1	2	4	4	8	3	2	1						1	1	2	1	1						1	5	4	2	1						1	5	4	2
-50							1		1	8										3										5										4
-70																																								1
< -70																																								

Methd	AC3										AC4										AEFM										CCPUE									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70																													1											
70									1										1										1											
50									1										1									1		1						1			3	1
30	4		1	1	1		1	3	3		4	1		1	1		1	3	3		4	1		1		1	1	2	1		3			1	2	2	2	2		1
<10	6	10	9	9	8	9	8	4	2	4	5	9	10	9	9	9	8	4	2	4	4	9	10	9	10	6	7	3	4	4	6	10	10	9	6	4	6	2	3	3
-30					1	1	1	3	3	2	1					1	1	3	3	2	2					3	2	4	3	1	1				2	3	2	6	4	
-50										4										3										3										4
-70																				1																				1
< -70																														1										

Methd	SURVIV										XSA										CAGEAN										ADAPT									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70																																							3	
70										1	1																										1	1	1	
50	1								3	3	1	1																			1					2	2	5	2	1
30	3	2	1	1	2	2	3	4	4		1	1	2	2	2	1					3	2							1		2	2	2	3	6	5	3	1	1	
<10	4	6	9	9	8	8	7	5	3	5	7	8	8	8	8	8	9	5	3		7	8	10	10	10	10	10	8	7	8	5	7	8	6	3	2	2	2	1	6
-30	2	2						1								1	1	4	6	3								2	2		2	1		1	1	1	1	1	2	
-50										1								1	1	7										2							1			3
-70																																								
< -70																																								

Methd	GLM										COLSIS										TSER1										TSER2									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70																																								
70	1																																							
50	1								2																															
30	5	4	3	3	2			2																																
<10	3	6	7	7	8	9	5	5	1	1												1	1	1	1	1	1					1	1	1	1	1	1	1	1	
-30						1	5	3	3	1																		1	1											
-50									4	6																														
-70										2																														
< -70																																								

Methd	SVPA										CONVEN									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70	2								1		1								1	
70	1	1							1		1									
50	1		1					3			2	1							1	
30	3	6	4	5	4	3	2			1	2	2	2		1		1	2	1	
<10	2	3	5	5	6	6	6	5	4	2	1	6	7	10	8	5	5	3	2	2
-30	1					1	2	2	3	2	2	1	1		1	5	4	3	2	2
-50									1	5	1							2	2	6
-70																			1	
< -70																				

Table 3.24 : Simulated Data Set 3 : Mean Log Ratio of Estimates of N at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	-0	-3	-3	-3	-6	-9	-13	-18	-12	-52
LS	5	3	3	4	3	-0	4	-1	7	-18
AC1	1	0	-0	1	-1	-5	-3	-13	-5	-34
AC2	1	0	-0	1	-1	-5	-3	-13	-5	-41
AC3	8	2	1	3	-0	-2	-0	-4	8	-27
AC4	6	2	1	3	-0	-2	1	-4	9	-34
AEFM	5	4	3	5	-0	-4	-3	-6	8	-28
CCPUE	3	-0	2	3	0	-2	-1	-12	-2	-29
SURVIV	4	2	4	5	5	7	9	7	17	12
XSA	11	10	8	7	4	1	-3	-16	-21	-48
CAGEAN	2	2	2	1	2	0	-0	-4	-1	-8
ADAPT	1	2	2	5	9	14	12	20	28	-11
GLM	15	9	7	6	3	-2	-12	-5	-24	-59
COLSIS										
TSER1		4	-1	1	-6	0	-8	-22	-29	
TSER2		7	4	2	-4	0	0	-7	0	
SVPA	25	16	10	12	9	2	-2	1	-1	-32
CONVEN	12	5	2	2	-0	-5	-7	-14	-13	-38

Table 3.25 : Simulated Data Set 3 : Root Mean Square Log Ratio of N at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	10	7	6	7	11	10	18	21	23	55
LS	10	5	4	6	6	2	11	9	18	30
AC1	10	5	4	4	4	6	9	16	18	44
AC2	10	5	3	4	4	6	9	16	18	54
AC3	11	5	6	7	8	6	10	16	25	38
AC4	12	6	5	7	8	6	11	16	26	50
AEFM	13	8	4	7	4	9	12	21	25	53
CCPUE	10	5	6	6	9	17	14	22	26	52
SURVIV	15	8	7	6	6	7	10	12	21	28
XSA	19	13	10	8	8	8	9	20	26	51
CAGEAN	10	6	4	4	3	3	3	7	8	18
ADAPT	18	11	8	10	14	23	27	31	44	31
GLM	22	12	10	9	9	6	15	15	42	66
COLSIS										
TSER1		4	1	1	6	0	8	22	29	
TSER2		7	4	2	4	0	0	7	0	
SVPA	37	19	15	14	12	9	13	23	35	44
CONVEN	32	14	10	6	7	10	15	30	43	45

Table 3.26: Simulated Data Set 3 : Frequency Distributions of Percentage Deviation of Estimates of F at age from True Values

Methd	HYBRID										LS										AC1										AC2									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70							1		1	2																			1	1									1	1
70	2						1	1	1	1																	1			1	1						1	1		1
50					1		2	3	3	2	2						1		1	1	3						1	2		1	2						1	1		1
30	2	2	2	1	7	7	4	4	1	3	1	2			1	1	1	1			1	2			3	4	2	4	5	4	1	2	1		3	4	1	3	6	4
<10	3	7	8	9	1	3	1	1	3	2	3	6	9	8	8	8	6	7	6	4	3	7	9	10	7	6	5	2	1	2	3	7	8	10	7	6	6	3		2
-30	3	1			1		1	1			3	2	1	2	1	1	2	2	2	5	2	1	1				1	2	3	1	2	1	1				1	2	2	1
-50									1		1								1		1										1								1	
-70																																								
< -70																																								

Methd	AC3										AC4										AEFM										CCPUE									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70																			1										1	1										3
70									1	2							1			2							1	1											1	
50	1						1			1	2									1	1						2				1					1	2	2	2	1
30	2	1			3	1	3	4	1	3	1	2			3	1	3	3	1	3	4	1	1	1	3	2	1	1	1	3	5	2		1	5	3	4	4	2	1
<10	3	8	9	9	6	8	5	4	3	2	3	7	9	9	6	8	5	6	4	2	1	8	9	5	7	8	5	5	2	4	1	7	10	6	4	5	1	3	2	3
-30	3	1	1	1	1	1	1	2	4	2	3	1	1	1	1	1	1	1	3	2	3	1		4			1	3	4		3	1		3	1	1	3	1	1	1
-50	1								1		1								1		1								2	2									2	1
-70																																								
< -70																																								

Methd	SURVIV										XSA										CAGEAN										ADAPT									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70									1								1	1	1	2																			1	
70											1								3	4											1						1			
50	2						1										1	2		1										2	1				1	1	1			1
30	2	2			1		1				2				2	2	2	2	2	2		1	3			1	2		4	3	3	3		1	1	1	1	2		1
<10	3	7	7	7	8	7	4	6	1	1	3	6	6	7	7	5	5	4	3		6	7	5	8	10	8	7	5	2	4	2	5	8	6	5	2	1	2	1	3
-30	2	1	3	3	1	3	4	3	5	4	3	4	4	3	1	3	1	1	1	1	4	2	2	2		1	1	5	4	1	2	2	2	3	3	4	5	2	2	4
-50	1							1	3	5	1																				1					2	1	3	2	1
-70																																						1	4	
< -70																																								

Methd	GLM										COLSIS										TSER1										TSER2									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70									2	1																														
70										3																														
50							2	2	1	2																			1										1	
30					1	4	6		3	3																1		1												
<10	4	4	5	8	8	3	2	3	3	1												1	1	1	1		1						1	1	1	1	1	1		
-30	4	6	5	2	1	3		3	1																							1								
-50	2							2																																
-70																																								
< -70																																								

Methd	SVPA										CONVEN									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70																		1	1	1
70										1								1	2	1
50									3	1							1	1	1	1
30			1			1	5	1	1	5	2		2		2	4	5	1	1	4
<10	4	3	3	5	8	7	5	7	3	1	4	8	7	9	8	6	4	5	3	2
-30	5	6	6	5	2	2		2	3	2	3	2	1	1				1	1	1
-50	1	1									1								1	
-70																				
< -70																				

Table 3.27 : Simulated Data Set 3 : Mean Log Ratio of Estimates of F at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	6	4	4	3	15	12	21	21	18	30
LS	0	-2	-3	-4	3	2	1	-1	-5	-7
AC1	4	1	0	-2	8	7	9	10	9	20
AC2	4	1	0	-2	8	7	9	10	9	20
AC3	-2	-1	-2	-3	7	4	6	1	-5	14
AC4	-1	-0	-1	-4	7	4	6	1	-5	14
AEFM	0	-3	-3	-5	7	5	9	5	-8	6
CCPUE	2	2	-2	-3	7	4	7	13	2	16
SURVIV	1	-1	-4	-6	0	-6	-4	-11	-20	-42
XSA	-5	-8	-9	-8	1	-1	10	14	25	37
CAGEAN	-11	-3	-0	-6	1	2	4	-11	-6	8
ADAPT	5	-0	-3	-6	-4	-15	-8	-25	-29	-11
GLM	-20	-13	-8	-8	1	2	19	-9	20	32
COLSIS										
TSER1		-9	6	3	7	11	5	15	39	
TSER2		-17	1	0	1	6	-1	10	36	
SVPA	-20	-15	-12	-13	-4	-2	9	-4	2	13
CONVEN	-6	-4	-2	-2	7	6	14	14	18	22

Table 3.28: Simulated Data Set 3 : Root Mean Square Log Ratio of F at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	26	11	7	7	19	16	29	27	37	36
LS	22	8	7	6	8	8	17	13	23	19
AC1	25	9	6	4	11	10	19	20	27	30
AC2	25	9	6	5	11	10	19	20	28	31
AC3	22	8	6	7	11	10	19	14	28	28
AC4	24	9	6	8	12	11	20	14	32	29
AEFM	25	9	7	10	9	10	21	20	38	47
CCPUE	23	10	6	10	14	17	23	20	30	48
SURVIV	25	10	8	9	7	9	17	19	38	46
XSA	25	13	11	10	10	12	22	26	42	43
CAGEAN	15	11	12	9	6	9	9	18	20	21
ADAPT	29	13	7	14	17	29	31	40	71	25
GLM	25	17	13	12	10	13	21	27	32	37
COLSIS										
TSER1		9	6	3	7	11	5	15	39	
TSER2		17	1	0	1	6	1	10	36	
SVPA	25	18	16	16	8	10	13	16	22	25
CONVEN	18	10	11	7	9	9	17	26	37	33

Table 3.29 : Simulated Data Set 3 : Frequency Distributions of Percentage Deviations of Estimates of Total Biomass from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70																		
70																		
50																		
30					1	1		1	1			4	2				6	
<10	9	10	10	10	9	9	10	9	9	10	10	6	8				4	10
-30	1																	
-50																		
-70																		
< -70																		

Table 3.30 : Simulated Data Set 3 : Frequency Distributions of Percentage Deviations of Estimates of Spawning Biomass from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70																		
70																		
50																		
30		1			1	1		2				10					1	
<10	9	9	10	10	9	9	10	6	10	10	10		10				9	9
-30	1							2										1
-50																		
-70																		
< -70																		

Table 3.31 : Simulated Data Set 3 : Frequency Distributions of Percentage Deviations of Estimates of Mean F (Ages 5-9) from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70												1						
70										1								
50	3		1	1			1	1										2
30	5	2	5	5	3	3	1	5	1	6			5		1		1	2
<10	1	7	3	3	6	6	8	2	6	3	10	5	5			1	9	6
-30	1	1	1	1	1	1		2	3			2						
-50												2						
-70																		
< -70																		

Table 3.32 : Simulated Data Set 3 : Mean Log Ratio of Estimates of Biomass and Mean F to True Values

Method	TSB	SSB	FBAR
HYBRID	-5	-9	19
LS	3	2	1
AC1	-1	-3	10
AC2	-1	-3	10
AC3	1	-1	3
AC4	1	-0	4
AEFM	2	-0	6
CCPUE	0	-1	8
SURVIV	6	7	-6
XSA	5	1	14
CAGEAN	1	0	-3
ADAPT	10	14	-11
GLM	4	-1	7
COLSIS			
TSER1			15
TSER2			10
SVPA	11	5	-0
CONVEN	1	-3	12

Table 3.33: Simulated Data Set 3 : Root Mean Square Log Ratio of Biomass and Mean F to True Values

Method	TSB	SSB	FBAR
HYBRID	8	11	24
LS	4	4	10
AC1	2	4	15
AC2	2	4	15
AC3	5	6	13
AC4	5	7	14
AEFM	4	3	14
CCPUE	7	10	17
SURVIV	7	8	12
XSA	5	3	20
CAGEAN	3	2	5
ADAPT	10	15	29
GLM	7	5	10
COLSIS			
TSER1			15
TSER2			10
SVPA	12	7	8
CONVEN	4	6	17

Table 3.34: Simulated Data Set 4 : Frequency Distributions of Percentage Deviation of Estimates of N at age from True Values

Methd	HYBRID										LS										AC1										AC2									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70																	1																							
70																																								
50						1	1				1					2	3	2								1	1	1								1	1	1		
30	1				1	2	2				4	7	8	6	5	5	2				3	3	2		1	3	2				3	3	2		1	3	2			
<10	8	10	10	10	9	6	7	6	5	1	5	3	2	4	5	3	4	5	5	1	7	7	8	10	9	6	7	6	5	1	7	7	8	10	9	6	7	6	5	1
-30	1					1																																		
-50								1																																
-70																																								
< -70																																								

Methd	AC3										AC4										AEFM										CCPUE									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70									1										1								1													
70																																								
50	1					1	2	2			1					1	2	1								2	3	2								1	1	1		
30	4	5	3	3	3	5	3				5	6	3	1	3	2	1				2	2	3	1	2	2					3	1	2	1	2	2	1			
<10	5	5	7	7	7	4	5	5	5	1	4	4	7	9	7	7	7	6	5	1	3	8	6	7	7	4	5	5	6	1	5	9	8	9	7	6	8	6	6	1
-30																					5		1	2	1	2					2				1	1				
-50																											1													
-70																																								
< -70																																								

Methd	SURVIV										XSA										CAGEAN										ADAPT									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70							1										1	2	1		3					2	5	4	2											
70																					4	3	1	4	3	4	3													
50	1					2	1	2								1	6	2			3	7	8	6	7	4	1	3												
30	4	3	4		4	6	3							1	5	6							1																	
<10	5	7	6	10	6	2	5	5	6	1	2	2	7	7	3	2	3	3	5	1							1		4	1										
-30											6	7	3	2	2	1																								
-50											1	1																												
-70											1																													
< -70																																								

Methd	GLM										COLSIS										TSER1										TSER2									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70							1																																	
70																																								
50						1	3																																	
30			2	5	5	5	2																1			1														
<10	4	3	8	5	5	3	4	6	4	1														1	1		1	1	1				1	1	1	1		1	1	
-30	4	6				1																1										1					1			
-50	2	1						1																																
-70																																								
< -70																																								

Methd	SVPA										CONVEN									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70						1	2	1	1		1						2	2	1	
70	1	1																		
50	1	2	3	3	3	1	2				1	2		1		2	2	1		
30	4	2	4	4	3	6	2				2	3	7	6	6	4	2			
<10	1	4	3	3	4	2	3	5	4	1	4	5	3	3	4	4	4	3	4	1
-30	3	1									2									
-50							1	1										1		
-70																				
< -70																				

Table 3.35: Simulated Data Set 4 : Mean Log Ratio of Estimates of N at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	-2	1	3	1	1	7	8	-10	0	0
LS	10	13	13	11	10	16	21	12	0	0
AC1	8	7	6	4	4	9	8	6	0	0
AC2	7	7	6	3	4	9	8	6	0	0
AC3	13	11	9	8	7	12	13	12	12	0
AC4	13	10	9	7	7	8	10	6	12	0
AEFM	-7	3	5	-3	3	7	11	12	0	0
CCPUE	2	4	5	3	2	6	6	6	0	0
SURVIV	11	9	9	1	8	16	17	12	0	0
XSA	-26	-21	-8	-3	4	14	27	31	12	0
CAGEAN	47	37	36	38	41	48	51	57	23	0
ADAPT										
GLM	-20	-16	4	11	9	11	21	-10	0	0
COLSIS										
TSER1		-14	19	3	0	16	0	0	0	
TSER2		-14	3	-6	-6	-8	-22	0	0	
SVPA	13	15	20	17	14	21	22	0	14	0
CONVEN	8	13	12	13	11	15	25	16	14	0

Table 3.36 : Simulated Data Set 4 : Root Mean Square Log Ratio of N at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	9	4	4	3	6	16	16	26	0	0
LS	14	14	14	12	11	20	30	22	0	0
AC1	12	8	7	5	6	15	16	15	0	0
AC2	12	8	7	5	6	15	16	15	0	0
AC3	17	12	10	9	9	17	19	22	28	0
AC4	17	11	10	7	9	15	18	15	28	0
AEFM	16	8	12	8	9	18	37	22	0	0
CCPUE	11	6	8	5	8	16	13	15	0	0
SURVIV	14	10	9	3	9	20	28	22	0	0
XSA	37	28	12	11	12	21	35	43	28	0
CAGEAN	48	37	37	38	41	49	55	59	40	0
ADAPT										
GLM	26	18	7	13	11	19	30	26	0	0
COLSIS										
TSER1		14	19	3	0	16	0	0	0	
TSER2		14	3	6	6	8	22	0	0	
SVPA	26	23	23	22	20	28	39	37	31	0
CONVEN	24	16	15	16	12	20	36	48	31	0

Table 3.37: Simulated Data Set 4 : Frequency Distributions of Percentage Deviation of Estimates of F at age from True Values

Methd	HYBRID										LS									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70									1											
70																			1	
50	1						1	1	2	1									1	
30	2	1	1	1			1	1	2	4	1						1	1	3	2
<¦10¦	3	7	9	9	10	10	5	5	4	2	3	3	5	4	1	4	2	3	3	3
-30	3	2					3	3	1	3	3	7	5	6	9	6	6	5	2	5
-50	1										3						1	1		
-70																				
< -70																				

Methd	AC1										AC2									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70									1										1	
70									1										1	
50								1									1	1		
30	1						1		3	2	1			1				1	3	2
<¦10¦	3	7	10	8	6	9	3	4	4	8	3	7	10	7	6	9	4	3	4	8
-30	4	3		2	4	1	6	4	1		4	3		2	4	1	5	4	1	
-50	2							1			2							1		
-70																				
< -70																				

Methd	AC3										AC4									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70																				
70																			1	
50																	1	1	1	
30							1	1	3										1	1
<¦10¦	3	3	8	6	4	7	2	3	5	10	3	4	8	7	5	7	3	3	5	9
-30	4	7	2	4	6	3	7	6	2		4	6	2	3	5	3	6	5	1	
-50	3										3							1	1	
-70																				
< -70																				

Methd	AEFM										CCPUE									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70							2		1	1							1		1	1
70	1		1															1		
50	1	1		2		1			1	1			1						1	1
30	1		1	1	1	2	1	4	1	1	2	1		1	1	1	2	1	3	3
<¦10¦	3	7	3	5	7	5	2	1			3	6	5	7	7	8	4	3	2	2
-30	3	2	5	2	2	2	3	3	6	6	5	3	4	2	2	1	3	4	1	1
-50	1						2	2	1	1								1	2	2
-70																				
< -70																				

Methd	SURVIV										XSA									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70											1	2	1							
70											1		1	1						
50								1	2		3	1	2	2	1		1			
30	2			2			1			1	2	4	2	1	1	1			1	
<¦10¦	2	6	6	7	2	1	3	1	2	1		3	3	4	3	2	2	3	1	
-30	4	4	4	1	8	9	4	7	5	6	3		1	2	5	7	3	1	4	1
-50	2						2	1	1	2							3	5	2	5
-70																	1	1	2	4
< -70																				

Methd	CAGEAN										ADAPT									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70																				
70																				
50																				
30																				
<¦10¦																				
-30			1																	
-50	9	10	9	10	9	9	9	5	4											
-70	1				1	1	1	5	6	9										
< -70										1										

Methd	GLM										COLSIS									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70																				
70																				
50																				
30	1								2	3										
<¦10¦	7	6	6	3	1	1		1	8	7										
-30	2	4	4	7	9	9	8	9												
-50							2													
-70																				
< -70																				

Methd	TSER1										TSER2									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70																				
70																				
50																				
30																		1	1	
<¦10¦			1	1	1	1	1	1	1				1	1	1	1	1			
-30																				
-50		1										1								
-70																				
< -70																				

Methd	SVPA										CONVEN									
Age	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
> 70																				
70																				
50																				
30							1		2	2									1	
<¦10¦	3	3	3	3	3	3	3	3	4	3	2	3	5	4	1	1	1	1	8	4
-30	5	7	5	4	4	4	3	3	4	2	8	7	5	6	9	9	9	9	1	6
-50	2		2	3	3	3	3	4		3										
-70																				
< -70																				

Table 3.38 : Simulated Data Set 4 : Mean Log Ratio of Estimates of F at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	-5	-1	3	2	-3	-2	-4	-3	13	3
LS	-17	-16	-12	-13	-17	-15	-16	-14	7	-7
AC1	-14	-9	-2	-2	-8	-5	-8	-7	15	3
AC2	-14	-8	-1	-1	-8	-5	-7	-7	16	4
AC3	-20	-13	-6	-7	-13	-10	-13	-12	1	-4
AC4	-20	-12	-5	-5	-11	-7	-10	-10	2	-1
AEFM	1	-4	0	9	-6	3	1	-10	-10	-8
CCPUE	-8	-5	-1	1	-5	-1	1	-5	5	7
SURVIV	-18	-11	-5	2	-15	-17	-18	-17	-9	-17
XSA	21	27	18	11	-4	-13	-22	-37	-39	-67
CAGEAN	-57	-52	-49	-55	-60	-61	-64	-72	-71	-100
ADAPT										
GLM	-3	-5	-8	-15	-21	-21	-28	-19	5	6
COLSIS										
TSER1		-63	-6	-8	-6	-7	-6	-4	-2	
TSER2		-48	4	9	10	9	10	11	13	
SVPA	-20	-18	-20	-20	-23	-25	-20	-25	-9	-15
CONVEN	-15	-15	-10	-15	-20	-18	-19	-20	-0	-11

Table 3.39: Simulated Data Set 4 : Root Mean Square Log Ratio of F at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	21	10	7	7	5	5	17	17	27	17
LS	26	18	14	14	18	16	23	21	24	17
AC1	24	13	5	7	10	7	19	21	27	8
AC2	25	13	6	8	9	6	19	22	30	9
AC3	27	15	9	10	14	11	19	19	19	7
AC4	27	15	10	9	12	9	20	20	25	7
AEFM	27	15	19	19	12	17	38	29	33	32
CCPUE	19	12	13	11	10	8	24	26	49	49
SURVIV	28	15	10	8	16	17	26	27	27	23
XSA	35	39	28	23	18	22	37	46	51	74
CAGEAN	58	53	50	56	61	61	64	72	72	101
ADAPT										
GLM	9	9	10	16	22	23	29	20	10	11
COLSIS										
TSER1		63	6	8	6	7	6	4	2	
TSER2		48	4	9	10	9	10	11	13	
SVPA	25	23	24	26	30	32	29	33	21	25
CONVEN	16	17	13	17	21	19	21	22	9	13

Table 3.40 : Simulated Data Set 4 : Frequency Distributions of Percentage Deviations of Estimates of Total Biomass from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70																		
70											7							
50											3						3	1
30		7	1	1	5	5	1	1	4								2	4
<10	10	3	9	9	5	5	9	9	6	5			8				5	5
-30										4			2					
-50										1								
-70																		
< -70																		

Table 3.41 : Simulated Data Set 4 : Frequency Distributions of Percentage Deviations of Estimates of Spawning Biomass from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70																		
70											6							
50											4						2	
30		8			3	2			2	4			3				5	7
<10	10	2	10	10	7	8	10	10	8	4			7				3	3
-30										2								
-50																		
-70																		
< -70																		

Table 3.42: Simulated Data Set 4 : Frequency Distributions of Percentage Deviations of Estimates of Mean F (Ages 5-9) from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70																		
70																		
50								2										
30	1		1	1			2									1		
<10	9	6	8	8	6	7	6	6	4	2			4		1		4	3
-30		4	1	1	4	3	2	2	6	8			6				3	7
-50											9						3	
-70											1							
< -70																		

Table 3.43 : Simulated Data Set 4 : Mean Log Ratio of Estimates of Biomass and Mean F to True Values

Method	TSB	SSB	FBAR
HYBRID	1	2	2
LS	12	12	-10
AC1	6	5	-1
AC2	6	5	-1
AC3	10	9	-9
AC4	9	7	-6
AEFM	1	2	-1
CCPUE	4	4	2
SURVIV	8	8	-14
XSA	-10	2	-20
CAGEAN	41	41	-65
ADAPT			
GLM	-3	6	-16
COLSIS			
TSER1			-5
TSER2			10
SVPA	15	16	-20
CONVEN	11	12	-15

Table 3.44: Simulated Data Set 4 : Root Mean Square Log Ratio of Biomass and Mean F to True Values

Method	TSB	SSB	FBAR
HYBRID	3	3	6
LS	13	12	11
AC1	7	6	7
AC2	7	5	8
AC3	11	9	10
AC4	10	8	9
AEFM	7	5	12
CCPUE	7	5	15
SURVIV	9	8	16
XSA	17	10	23
CAGEAN	41	41	66
ADAPT			
GLM	7	8	18
COLSIS			
TSER1			5
TSER2			10
SVPA	20	21	28
CONVEN	14	14	17

Table 3.45 : Simulated Data Set 5 : Frequency Distributions of Percentage Deviation of Estimates of N at age from True Values

HYBRID

Age	3	4	5	6	7	8	9	10	11	12
> 70	2									2
70		1	1	1	1				1	
50		1		1			1		1	1
30	1	1	3	1	3	3	2	2	1	
<¦10¦	2	4	2	2	3	1	3	2	1	3
-30	4	2	4	4	3	5	3	3	4	
-50	1	1		1		1	1	3	2	3
-70										1
< -70										

LS

Age	3	4	5	6	7	8	9	10	11	12
> 70	1		1							1
70	1	2	1	1	2					
50	1		3	1	1	1	2			1
30	2	3	1	1	2	3	1	3	1	
<¦10¦	4	4	4	5	5	2	2	4	3	4
-30	1			2		4	5	2	5	
-50		1						1	1	3
-70										1
< -70										

AC1

Age	3	4	5	6	7	8	9	10	11	12
> 70	1	1			1					
70		1	1	1			1			
50		1	4		3					1
30	3	1	1	2	1	1	1	3	1	1
<¦10¦	3	2	3	4	4	5	3	4	4	4
-30	3	3	1	1		4	5	2	2	1
-50				2	1			1	3	2
-70		1								1
< -70										

AC2

Age	3	4	5	6	7	8	9	10	11	12
> 70	1	1	1		1					
70		1	1	1	1		1			
50		1	3		2		1			1
30	3	1	1	2	1	1		3	1	1
<¦10¦	3	2	3	3	4	5	3	4	4	4
-30	2	3		2		4	5	2	1	1
-50	1		1	2	1			1	4	2
-70		1								1
< -70										

AC3

Age	3	4	5	6	7	8	9	10	11	12
> 70	2	1			1				2	
70	1	1	2		1		1			
50	1	1	2	3	2	1	2		1	5
30	1	1	4	2	3	4	2	8	1	
<¦10¦	4	4	1	4	2	3	2		3	3
-30	1	2	1	1	1	1	2	1	3	
-50						1	1	1		1
-70										1
< -70										

AC4

Age	3	4	5	6	7	8	9	10	11	12
> 70	2	2	1		2				2	
70	1		2			1	1			
50	1	1	2	3	4		2	1	1	5
30		2	3	2	1	2	1	5		
<¦10¦	4	3	1	3	2	4	2	2	4	2
-30	2	2	1	2	1	2	4	2	2	1
-50						1			1	1
-70										1
< -70										

AEFM

Age	3	4	5	6	7	8	9	10	11	12
> 70	2	1			1		2	1		
70	1		1							
50		2	1		3	2	1			1
30			3	4	1			3		1
<¦10¦	5	6	3	2	2	2	1	1	2	1
-30	2			1	2	4	5	2	2	
-50		1	2	2	1	1		1	5	6
-70				1		1	1	2	1	1
< -70										

CCPUE

Age	3	4	5	6	7	8	9	10	11	12
> 70	2				1	1	1	1	1	2
70			1							
50	1	2	1	1	3	1	1	1		
30	1	5	4	3	1	1	1	1	1	
<¦10¦	3		3	2	3	2	2	4	3	2
-30	3	2	1	4	2	4	4	1	3	1
-50		1				1	1	2	2	4
-70										1
< -70										

SURVIV

Age	3	4	5	6	7	8	9	10	11	12
> 70	2									
70			1							
50	2	1	2		1					
30	2	5	2	3		1		4	1	
<¦10¦	3	2	1	2	2	2	7	4	4	5
-30	1		2	3	6	6	3	1	3	2
-50		2	2	2	1	1		1	2	3
-70										
< -70										

XSA

Age	3	4	5	6	7	8	9	10	11	12
> 70	1									
70	1									
50		2	1	2	1			1		
30	2	2	5	1	1					
<¦10¦	5	4	3	6	6	8	6	4	3	2
-30	1	2	1	1	2	2	4	4	4	
-50									2	7
-70								1		1
< -70									1	

CAGEAN

Age	3	4	5	6	7	8	9	10	11	12
> 70	1									
70									1	
50	3	2	1	1	1	1	1	1	1	4
30	2	3	4	4	3	4	3	5	5	2
<¦10¦	3	4	4	3	6	3	6	3	2	4
-30		1	1	2		2		1	1	
-50	1									
-70										
< -70										

ADAPT

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<¦10¦										
-30										
-50										
-70										
< -70										

GLM

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70	2	1			1					
50		1	1	1						
30	2	1	3	2	1		1			
<¦10¦	3	2	5	2	4	3	1	1		
-30	2	4	1	4	4	5	6	1	1	
-50	1	1		1		2	2	6	8	7
-70								2	1	
< -70										3

COLSIS

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<¦10¦										
-30										
-50										
-70										
< -70										

TSER1

Age	3	4	5	6	7	8	9	10	11	12
> 70									1	
70										
50		1	1	1	1	1	1	2	1	
30		3	1	1	1	1	2		1	
<¦10¦		2	3	5	4	3	2	2		
-30		2	4	3	3	2	1	2	2	
-50		2	1		1	3	4	3	4	
-70								1	1	
< -70										

TSER2

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<¦10¦										
-30										
-50										
-70										
< -70										

SVPA

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<¦10¦										
-30										
-50										
-70										
< -70										

CONVEN

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<¦10¦										
-30										
-50										
-70										
< -70										

Table 3.46 : Simulated Data Set 5 : Mean Log Ratio of Estimates of N at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	4	-1	0	-3	3	-11	-5	-19	-9	-10
LS	18	8	23	7	17	2	-3	-5	-14	-21
AC1	6	-1	17	-3	12	-8	-7	-5	-16	-19
AC2	6	-1	20	-4	15	-8	-7	-4	-18	-19
AC3	23	14	21	10	20	1	6	8	18	8
AC4	20	17	27	10	25	2	5	7	15	2
AEFM	22	4	2	-10	10	-16	-6	-18	-42	-44
CCPUE	16	5	13	1	14	1	-0	1	-9	-16
SURVIV	26	1	-1	-13	-16	-20	-12	-4	-15	-24
XSA	14	8	10	6	3	-4	-8	-14	-35	-58
CAGEAN	15	9	8	7	8	7	11	12	14	20
ADAPT										
GLM	3	-4	3	-6	-5	-22	-21	-55	-58	-94
COLSIS										
TSER1		-10	-8	-6	-8	-13	-16	-23	-27	
TSER2										
SVPA										
CONVEN										

Table 3.47: Simulated Data Set 5 : Root Mean Square Log Ratio of N at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	42	31	24	27	22	25	26	31	32	65
LS	35	28	30	22	25	15	21	21	23	58
AC1	35	38	29	27	29	17	25	23	26	47
AC2	36	39	33	28	32	17	25	23	28	47
AC3	37	28	28	20	31	21	29	21	53	42
AC4	39	31	35	23	37	24	27	20	57	48
AEFM	61	30	30	40	38	41	46	46	50	62
CCPUE	34	30	24	21	33	32	31	37	48	75
SURVIV	42	30	36	27	24	26	17	21	28	36
XSA	28	19	20	17	16	13	15	32	54	66
CAGEAN	29	21	18	17	14	15	16	19	24	28
ADAPT										
GLM	27	29	18	23	21	26	27	60	63	102
COLSIS										
TSER1		29	24	18	24	29	34	40	53	
TSER2										
SVPA										
CONVEN										

Table 3.48 : Simulated Data Set 5 : Frequency Distributions of Percentage Deviation of Estimates of F at age from True Values

HYBRID

Age	3	4	5	6	7	8	9	10	11	12
> 70	1							1	2	3
70	1	1		2	1	1		1		
50			2	1	1	2	3		4	1
30	2	2	1	1		2	2			
<10	1	2	2	3	5	2	3	5		
-30	4	2	3	1	2	2	2	3	2	4
-50		3	2	2	1	1			1	1
-70	1								1	1
< -70										

LS

Age	3	4	5	6	7	8	9	10	11	12
> 70	1								2	2
70								1	1	
50	1	1		1			1	1	1	2
30				3	1	1	3		2	1
<10	1	2	3		2	4	3	3	1	
-30	3	3	2	4	5	4	3	3	2	4
-50	3	4	5	2	2	1		2	1	1
-70	1									
< -70										

AC1

Age	3	4	5	6	7	8	9	10	11	12
> 70	1	1		1	1			1	3	
70				1			1			1
50	1			1		3	1	1	2	2
30	2	1	2	1	1		2	1		4
<10	2	3		3	3	5	4	2	2	1
-30	1	2	5	2	2	2	2	2	1	2
-50	2	2	3		3			3	1	
-70	1	1		1					1	
< -70										

AC2

Age	3	4	5	6	7	8	9	10	11	12
> 70	1	1		1	1			1	3	
70				2			1			1
50	1		1			2	1	1	2	2
30	2	1	1	1	1	1	2	1		4
<10	2	3		3	2	5	4	2	2	1
-30	1	2	5	2	3	2	2	2	1	2
-50	2	2	3		3			3	1	
-70	1	1		1					1	
< -70										

AC3

Age	3	4	5	6	7	8	9	10	11	12
> 70									1	
70	1									
50				1		2				
30	1			1	1		3	1	1	2
<10	1	3	2	3	3	3	3	3	4	4
-30	2	2	5	3	3	4	3	2	1	3
-50	4	4	2	2	2	1	1	3	1	1
-70	1	1	1		1			1	1	
< -70									1	

AC4

Age	3	4	5	6	7	8	9	10	11	12
> 70	1								1	
70				1		1		1		
50	1				1	1				1
30			1	1			3	2	1	2
<10	1	3	1	4	3	4	3	1	4	3
-30	2	3	2	2	2	3	3	1	1	3
-50	4	3	5	1	3	1	1	4	1	1
-70	1	1	1	1	1			1	1	
< -70									1	

AEFM

Age	3	4	5	6	7	8	9	10	11	12
> 70	1		1	3		1	2	3	4	4
70		1			1	2		1	1	2
50	2		1		1		1		1	1
30	1		1	1	2	2	2		1	1
<10	1	1	1	1	1	3	1	1	3	1
-30	1	6	4	2	3		2	1		1
-50	1	2	2	3	1	2	2	3		
-70	2				1			1		
< -70	1									

CCPUE

Age	3	4	5	6	7	8	9	10	11	12
> 70	2							1	3	4
70		1		1	1		1	1	1	
50						1				1
30			1	3	1	4	4	1	1	2
<10	2	2	2	2	3	1	1	2	3	1
-30	1	3	4	3	3	2	3	2		
-50	4	4	3	1	1	1	1	2		
-70	1				1	1		1	1	1
< -70									1	1

SURVIV

Age	3	4	5	6	7	8	9	10	11	12
> 70							1	1	2	1
70	1	1	3	4	2	3		1		1
50				1	3	1	2	1	2	2
30	1	1			2	3	3	1		3
<10	4		2	3	1	2	1		3	2
-30		7	3	1	2	1	3	3	2	
-50		1	2	1				3	1	1
-70	4									
< -70										

XSA

Age	3	4	5	6	7	8	9	10	11	12
> 70	1						1	3	5	4
70						1				2
50				1	2			1	1	1
30	1		1	1		1	4			2
<10	3	2	3	3	4	5	2		2	
-30	1	6	4	4	4	3	2	4	2	1
-50	3	2	2	1			1	1		
-70	1							1		
< -70										

CAGEAN

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30					1					
<10	1	4	4	2	1	1	3	1		
-30	3	3	6	8	7	8	6	5	3	1
-50	6	3			1		1	4	6	8
-70						1			1	1
< -70										

ADAPT

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

GLM

Age	3	4	5	6	7	8	9	10	11	12
> 70							1	2	1	9
70					1	1	1	4	3	
50				2	1	3	2	2	2	
30		3	2	2	2	3	2	1	3	1
<10	2	3	3	3	4	2	3	1	1	
-30	5	2	5	2	2	1	1			
-50	3	2		1						
-70										
< -70										

COLSIS

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

TSER1

Age	3	4	5	6	7	8	9	10	11	12
> 70						2		1	2	
70					3		2	1	1	
50		1		1		2	3	1		
30		2	4	2	1	3	1	2	2	
<10		3	4	4	3	1	1	3	3	
-30		3	1	3	2	1	2	1		
-50		1	1		1	1	1	1	2	
-70										
< -70										

TSER2

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

SVPA

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

CONVEN

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

Table 3.49 : Simulated Data Set 5 : Mean Log Ratio of Estimates of F at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	-6	-14	-5	2	-4	9	6	6	9	4
LS	-20	-23	-30	-10	-20	-7	3	-11	16	12
AC1	-8	-14	-24	3	-14	6	8	-9	16	12
AC2	-8	-14	-27	4	-17	5	8	-10	17	12
AC3	-26	-30	-28	-13	-24	-6	-7	-26	-25	-9
AC4	-22	-33	-35	-13	-29	-6	-7	-23	-21	-6
AEFM	-25	-19	-7	12	-11	16	8	14	54	63
CCPUE	-18	-21	-19	-3	-17	-5	1	-15	11	19
SURVIV	-29	-16	-4	14	18	19	14	-10	12	18
XSA	-16	-24	-16	-9	-5	0	9	6	35	53
CAGEAN	-37	-24	-16	-15	-18	-26	-16	-30	-48	-55
ADAPT										
GLM	-30	-10	-5	0	8	19	17	40	30	79
COLSIS										
TSER1		-4	-1	2	8	16	15	11	11	
TSER2										
SVPA										
CONVEN										

Table 3.50 : Simulated Data Set 5 : Root Mean Square Log Ratio of F at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	49	34	30	32	28	27	24	31	59	52
LS	47	36	39	29	30	19	19	33	44	44
AC1	50	40	35	39	37	21	22	37	54	25
AC2	50	42	39	41	42	22	24	40	55	25
AC3	46	39	35	29	36	29	20	38	62	20
AC4	54	43	43	34	44	32	24	45	69	23
AEFM	74	33	34	58	44	52	45	84	83	92
CCPUE	49	36	31	28	41	37	27	53	71	76
SURVIV	55	31	39	36	30	27	33	45	48	32
XSA	41	28	26	25	21	17	36	57	55	72
CAGEAN	40	31	21	17	24	31	21	34	50	56
ADAPT										
GLM	36	28	18	24	23	27	30	45	35	83
COLSIS										
TSER1		24	25	18	32	39	35	33	38	
TSER2										
SVPA										
CONVEN										

Table 3.51 : Simulated Data Set 5 : Frequency Distributions of Percentage Deviations of Estimates of Total Biomass from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70																		
70								1										
50	1	2	1	1	1	3	2	1										
30	2	3	3	3	7	5	2	3	1	1	3		2					
<10	3	4	2	3			3	2	5	9	7		5					
-30	3	1	3	2	2	2	2	2	3				2					
-50	1		1	1			1	1	1				1					
-70																		
< -70																		

Table 3.52 : Simulated Data Set 5 : Frequency Distributions of Percentage Deviations of Estimates of Spawning Biomass from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70								1										
70																		
50					2	3	2											
30	3	4	2	3	4	4	1	4			4							
<10	1	5	6	5	2	1	4	2	4	8	6		4					
-30	6	1	2	2	2	2	2	3	6	2			5					
-50							1						1					
-70																		
< -70																		

Table 3.53 : Simulated Data Set 5 : Frequency Distributions of Percentage Deviations of Estimates of Mean F (Ages 5-9) from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70							4						1					
70							2	1		1			1		3			
50	4					1		2	2									
30	2	4	6	6	2	1	1	2	7	6			6		3			
<10	1	4	2	2	1	2	2	2	1	3			1		2			
-30	3	2	2	2	6	5	1	2			9		1					
-50					1	1					1				2			
-70								1										
< -70																		

Table 3.54: Simulated Data Set 5 : Mean Log Ratio of Estimates of Biomass and Mean F to True Values

Method	TSB	SSB	FBAR
HYBRID	-4	-5	9
LS	9	6	1
AC1	0	-0	7
AC2	1	1	7
AC3	14	11	-14
AC4	17	14	-11
AEFM	1	-3	35
CCPUE	5	4	0
SURVIV	-10	-16	17
XSA	1	-4	18
CAGEAN	4	4	-27
ADAPT			
GLM	-10	-17	24
COLSIS			
TSER1			12
TSER2			
SVPA			
CONVEN			

Table 3.55: Simulated Data Set 5 : Root Mean Square Log Ratio of Biomass and Mean F to True Values

Method	TSB	SSB	FBAR
HYBRID	23	20	26
LS	19	14	14
AC1	20	13	19
AC2	21	14	20
AC3	21	20	22
AC4	24	23	24
AEFM	26	26	46
CCPUE	26	28	37
SURVIV	17	19	19
XSA	7	7	22
CAGEAN	11	9	28
ADAPT			
GLM	19	22	30
COLSIS			
TSER1			34
TSER2			
SVPA			
CONVEN			

Table 3.56: Simulated Data Set 6 : Frequency Distributions of Percentage Deviation of Estimates of N at age from True Values

Methd	HYBRID									
Age	3	4	5	6	7	8	9	10	11	12
> 70	1	1	1	2				2	1	1
70		2		1	2	1	1			
50	1		2		1	2	2		1	
30	3	2	2	2		1	2	1	1	
<10	3	1	3		4	3	1	2	4	9
-30	2	4	2	3	2	2	3	5	2	
-50				2	1	1	1		1	
-70										
< -70										

Methd	LS									
Age	3	4	5	6	7	8	9	10	11	12
> 70		1								
70		2	1	2						
50	1	1	2	2	1	1	1	1	1	3
30	1	5	5	1	3	4	3	1	2	1
<10	4	1	2	3	3	3	4	3	2	5
-30	4			2	3	1	2	4	5	1
-50						1		1		
-70										
< -70										

Methd	AC1									
Age	3	4	5	6	7	8	9	10	11	12
> 70										1
70		1								
50			1						1	1
30	2	1	3	3			1	2	1	
<10	3	3	2	3	4	3	3	1	1	2
-30	3	5	4	1	5	5	5	6	5	
-50	2			3	1	2	1	1	2	5
-70										1
< -70										

Methd	AC2									
Age	3	4	5	6	7	8	9	10	11	12
> 70										1
70		1								
50			1						1	1
30	2	1	3	4	1		1	2	1	
<10	3	3	2	1	3	3	3	1	1	2
-30	3	2	4	2	5	5	5	6	5	
-50	2	3		3	1	2	1	1	2	5
-70										
< -70										1

Methd	AC3									
Age	3	4	5	6	7	8	9	10	11	12
> 70	6	3	2	2	3	2	3	1	3	4
70		1		2	1	2		1		
50	1	2	2	2			1			2
30	2	1	3	1	2	3	1	3		
<10	1	3	3	3	4	2	4	3	3	1
-30						1	1	1	3	2
-50								1	1	1
-70										
< -70										

Methd	AC4									
Age	3	4	5	6	7	8	9	10	11	12
> 70	4	3	2	4	3	2	3	1	3	3
70	2	2				1				1
50	1	1	3	1	1	1	1	3		2
30	3	1	3	2	2	3	3	1	1	
<10		2	2	2	3	2	3	3	2	1
-30		1		1	1	1		1	3	1
-50								1	1	2
-70										
< -70										

Methd	AEFM									
Age	3	4	5	6	7	8	9	10	11	12
> 70	3	1	2			2		2		2
70				1	1				1	1
50		2	1			1	2			
30	1	1	1	1	1	1	1	1	1	
<10	4	2	3	3	2	1	1	4	3	3
-30	1	2	1		3		1		3	1
-50	1	1	2	5	1	3	5	2	2	2
-70					2	2		1		1
< -70		1								

Methd	CCPUE									
Age	3	4	5	6	7	8	9	10	11	12
> 70		2			1	1				2
70	1			1	1					
50	2	2	1		1					1
30		1		1	2	2	1			
<10	2	2	5	4	3	2	2	4	3	2
-30	4	3	4	2	2	3	6	3	3	
-50	1			1		1	1	3	3	3
-70				1		1			1	2
< -70										

Methd	SURVIV									
Age	3	4	5	6	7	8	9	10	11	12
> 70	3									
70										
50		1								
30	3	2				1		1	1	
<10	2	3	5	4	4	3	4	2	2	4
-30	2	3	4	3	2	2	5	3	4	2
-50		1	1	3	4	4	1	4	2	3
-70									1	1
< -70										

Methd	XSA									
Age	3	4	5	6	7	8	9	10	11	12
> 70	1									
70		1	1							
50	2	1		1	1	1	1		1	
30	2	3	4	3	2	2	1	3	2	1
<10	3	3	5	5	6	5	7	4	3	7
-30	2	2		1	1	2	1	3	4	1
-50										1
-70										
< -70										

Methd	CAGEAN									
Age	3	4	5	6	7	8	9	10	11	12
> 70	1									1
70	3	1					1		1	
50		2	1			1		1		5
30	3	2	4	5	4	3	4	8	6	2
<10	1	3	5	5	6	5	5	1	2	1
-30		2				1			1	
-50	2									1
-70										
< -70										

Methd	ADAPT									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

Methd	GLM									
Age	3	4	5	6	7	8	9	10	11	12
> 70	1	1	1				1			
70		1		1	1		1			
50	2			1		1				
30		1	1		1	1	1			
<10	1	3	7	4	2	2			2	2
-30	5	4	1	3	6	4	6	4	6	1
-50	1			1		2	1	5	2	6
-70								1		1
< -70										

Methd	COLSIS									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

Methd	TSER1									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70		1								
50			1				1			
30		1	2	4	1	1		2		
<10		6	5	3	4	5	3	1	2	
-30		2	2	2	5	2	4	4	5	
-50				1		2	2	3	3	
-70										
< -70										

Methd	TSER2									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

Methd	SVPA									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

Methd	CONVEN									
Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

Table 3.57: Simulated Data Set 6 : Mean Log Ratio of Estimates of N at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	11	15	11	8	2	5	6	3	5	11
LS	1	27	21	12	4	3	4	-6	-5	10
AC1	-14	-8	2	-14	-12	-24	-17	-14	-16	-23
AC2	-14	-13	2	-14	-13	-26	-16	-15	-16	-30
AC3	62	39	33	36	29	31	31	16	27	26
AC4	65	43	36	37	29	31	33	16	23	22
AEFM	23	-3	6	-19	-31	-12	-23	-5	-13	-3
CCPUE	-2	19	-5	-8	16	-9	-17	-20	-32	-17
SURVIV	29	-7	-13	-23	-23	-25	-18	-25	-33	-30
XSA	11	9	11	8	4	4	3	-0	-2	-5
CAGEAN	15	13	12	8	7	10	15	16	17	25
ADAPT										
GLM	-1	3	4	-5	-7	-9	-5	-49	-29	-54
COLSIS										
TSER1		2	2	-1	-8	-12	-19	-23	-29	
TSER2										
SVPA										
CONVEN										

Table 3.58: Simulated Data Set 6 : Root Mean Square Log Ratio of N at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	30	43	29	37	30	29	29	32	31	31
LS	16	33	25	27	16	22	18	22	21	23
AC1	26	23	18	33	20	32	24	25	31	56
AC2	27	30	20	34	23	33	24	25	31	70
AC3	76	57	49	47	45	53	53	40	74	54
AC4	78	62	52	51	48	52	49	37	67	59
AEFM	76	56	37	40	57	75	43	55	31	62
CCPUE	29	40	19	34	36	36	25	30	42	72
SURVIV	51	26	19	32	30	33	23	32	45	43
XSA	28	20	18	15	14	16	16	16	19	19
CAGEAN	40	26	18	12	12	18	22	17	24	36
ADAPT										
GLM	37	30	20	28	24	25	37	56	35	63
COLSIS										
TSER1		17	16	19	16	23	31	37	38	
TSER2										
SVPA										
CONVEN										

Table 3.59 : Simulated Data Set 6 : Frequency Distributions of Percentage Deviation of Estimates of F at age from True Values

HYBRID

Age	3	4	5	6	7	8	9	10	11	12
> 70									1	2
70		1	1		1	1			1	1
50		2		2	1	2		1		1
30		1	1	3	2	1	3	2		
<10	1	2	2	1	3	2	3	1		
-30	5	2	4	2	1	2	2	3	1	3
-50	2	1	1	1	2	2	2	2	6	1
-70	2	1	1	1				1	1	1
< -70										1

LS

Age	3	4	5	6	7	8	9	10	11	12
> 70									1	2
70						1				
50				1	1	2	1	1		1
30		1	1	2	3		1	3		
<10	4	1	2	2	1	2	4	2	2	1
-30	3	7	3	4	5	3	2	2	3	2
-50	1		4			2	2	2	4	3
-70	2	1		1						1
< -70										

AC1

Age	3	4	5	6	7	8	9	10	11	12
> 70				3	1	3	1	1	1	3
70		1		1			2			2
50	2	2	2		2		2	2		
30	3	3	3	1	2	5	1	2	2	1
<10	1	3		3	4	2	1	3	3	2
-30	2		3	1	1		3	2	3	2
-50	1	1	2	1						
-70									1	
< -70	1									

AC2

Age	3	4	5	6	7	8	9	10	11	12
> 70				3	1	3	1	1	1	3
70		4		1	1		2			2
50	2	2	2		2	1	2	2		
30	3		3	2	1	4	1	2	2	1
<10	1	3		2	4	2		2	3	2
-30	2		3	1	1		4	3	3	2
-50	1	1	2	1						
-70									1	
< -70	1									

AC3

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50						1	1	1		1
30					1	2		1		
<10		4	1	1	2			1	1	2
-30	1	2	5	5	3	3	2	3	5	3
-50	3	2	2	2	2	2	5	2	1	3
-70	3	2	2	2	2	1	2	2	2	1
< -70	3					1			1	

AC4

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70							1	1		
50					1	1				1
30				1				1	1	
<10		4	2	1	3	2		1	2	2
-30	1	2	4	4	3	2	2	2	1	4
-50	3	1	2	1		3	5	3	3	2
-70	2	2	2	3	3	1	2	2	2	1
< -70	4	1				1			1	

AEFM

Age	3	4	5	6	7	8	9	10	11	12
> 70		2	2	5	3	4	2	2	1	1
70		1			2	1	1			
50					1		1		1	1
30	2	1	2	1	1		2	1	1	2
<10	2	2	1	3			2	4	1	2
-30	2	3	2		2	3	2		4	1
-50	1		2		1	1		1	1	2
-70	1	1	1	1				2	1	1
< -70	2					1				

CCPUE

Age	3	4	5	6	7	8	9	10	11	12
> 70				2		2	2	1	2	2
70			1		1		1		1	1
50			2	1		1	1	3		1
30	2	3	2	2		2		1	1	1
<10	2	3	2	3	4	2	3	2	2	2
-30	2	3	2	1	2	2	3	3	3	
-50	3		1		2				1	3
-70	1	1		1	1	1				
< -70										

SURVIV

Age	3	4	5	6	7	8	9	10	11	12
> 70		1		1	2	2	2	2	2	3
70			1	4	2	3	1	1		1
50		3	1	1		1	1	1	3	3
30	2	2	3	3	3	1	2	2		1
<10	2	1	5		1	2	2	3		1
-30		3		1	2	1	1	1	3	1
-50	2						1		2	
-70	2									
< -70	2									

XSA

Age	3	4	5	6	7	8	9	10	11	12
> 70									1	
70						1	1			1
50			1		1		1			3
30	1	2		2	1	2	2	2	1	
<10	1	7	2	5	4	3	1	2		2
-30	4		6	2	4	3	2	5	3	4
-50	2	1	1	1		1	3	1	5	
-70	2									
< -70										

CAGEAN

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30				1						
<10		3	3	3	1	1	1	3		
-30	1	5	7	6	6	5	7	4		2
-50	5	2			3	3	2	3	8	8
-70	4					1			2	
< -70										

ADAPT

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

GLM

Age	3	4	5	6	7	8	9	10	11	12
> 70								6		4
70					1	2	1			
50				2				1		3
30	1	3	3	2	3	3	3	1	2	1
<10	1	4	5	3	3	3	2	1	3	
-30	2	2	1	2	1	1	2	1	4	1
-50	5	1	1	1	2	1	1		1	1
-70	1						1			
< -70										

COLSIS

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

TSER1

Age	3	4	5	6	7	8	9	10	11	12
> 70						1	1	1	2	
70				1	1		1	1		
50		1	1	2		1	1	1	2	
30			1		4	4	4	2	5	
<10		7	6	7	3	3	3	4	1	
-30		2	2		2	1		1		
-50										
-70										
< -70										

TSER2

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

SVPA

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

CONVEN

Age	3	4	5	6	7	8	9	10	11	12
> 70										
70										
50										
30										
<10										
-30										
-50										
-70										
< -70										

Table 3.60: Simulated Data Set 6 : Mean Log Ratio of Estimates of F at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	-39	-10	-14	-4	-2	-5	-11	-15	-23	-10
LS	-30	-24	-25	-9	-5	-3	-8	-5	-17	-8
AC1	-15	14	-3	22	14	31	17	7	4	25
AC2	-15	20	-4	23	16	33	17	8	5	27
AC3	-93	-37	-38	-36	-33	-34	-40	-29	-56	-29
AC4	-93	-41	-41	-36	-33	-35	-41	-28	-51	-28
AEFM	-52	11	-7	29	43	24	25	-1	-1	1
CCPUE	-27	-14	4	15	-18	13	18	13	24	26
SURVIV	-58	14	13	32	28	31	19	21	15	36
XSA	-39	-5	-15	-5	-6	-5	-8	-12	-20	2
CAGEAN	-62	-18	-19	-12	-27	-34	-23	-24	-58	-44
ADAPT										
GLM	-39	-9	-4	-1	0	6	-7	52	-12	35
COLSIS										
TSER1		-4	3	8	9	15	22	18	23	
TSER2										
SVPA										
CONVEN										

Table 3.61 : Simulated Data Set 6 : Root Mean Square Log Ratio of F at age to True Values

Method	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12
HYBRID	51	46	36	40	35	37	24	36	63	69
LS	48	33	33	29	21	31	25	25	54	51
AC1	58	32	28	45	28	42	36	29	56	43
AC2	58	39	30	48	32	44	39	29	57	46
AC3	102	57	53	47	50	61	53	51	78	44
AC4	102	63	57	52	54	60	54	50	74	42
AEFM	86	69	46	55	75	100	46	73	52	53
CCPUE	41	38	28	46	39	50	46	36	71	76
SURVIV	89	31	23	41	42	43	43	33	65	47
XSA	55	21	26	20	20	24	36	26	58	30
CAGEAN	67	24	22	16	31	38	27	28	59	45
ADAPT										
GLM	51	25	23	28	26	30	38	64	25	56
COLSIS										
TSER1		17	14	20	18	29	28	31	30	
TSER2										
SVPA										
CONVEN										

Table 3.62 : Simulated Data Set 6 : Frequency Distributions of Percentage Deviations of Estimates of Total Biomass from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70	1				4	3	1											
70					1	2		1					1					
50	3	2			1	1	1			1	2		1					
30	1	5			2	2		2	1	1	5							
<10	3	2	7	7	2	2	4	6	2	8	3		3					
-30	2	1	2	2			2	1	7				5					
-50			1	1			2											
-70																		
< -70																		

Table 3.63 : Simulated Data Set 6 : Frequency Distributions of Percentage Deviations of Estimates of Spawning Biomass from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70					2	2	1											
70	1				1	1												
50	3				2	2	1	1			2		2					
30	1	5			3	3				2	4							
<10	3	4	5	5	2	1	2	7		8	4		2					
-30	1	1	4	4		1	4	2	9				6					
-50	1		1	1			2		1									
-70																		
< -70																		

Table 3.64: Simulated Data Set 6 : Frequency Distributions of Percentage Deviations of Estimates of Mean F (Ages 5-9) from True Values

Methd	HYBRI	LS	AC1	AC2	AC3	AC4	AEFM	CCPUE	SURVI	XSA	CAGEA	ADAPT	GLM	COLSI	TSER1	TSER2	SVPA	CONVE
> 70			1	1			4	3	2									
70			1	2			2						1		2			
50	2	1	2	1		1			1	1			2					
30	2	2	2	2	1		1	1	4				3		4			
<10	1	2	4	4	1	1	2	5	3	5			2		4			
-30	3	4			3	4				4	7		1					
-50	2	1			4	3		1			3		1					
-70						1	1											
< -70					1													

Table 3.65 : Simulated Data Set 6 : Mean Log Ratio of Estimates of Biomass and Mean F to True Values

Method	TSB	SSB	FBAR
HYBRID	11	8	-8
LS	12	8	-4
AC1	-10	-13	19
AC2	-12	-14	20
AC3	41	33	-36
AC4	43	34	-34
AEFM	2	-10	32
CCPUE	4	-2	18
SURVIV	-13	-22	27
XSA	6	4	-6
CAGEAN	15	15	-33
ADAPT			
GLM	-2	-8	11
COLSIS			
TSER1			17
TSER2			
SVPA			
CONVEN			

Table 3.66 : Simulated Data Set 6 : Root Mean Square Log Ratio of Biomass and Mean F to True Values

Method	TSB	SSB	FBAR
HYBRID	30	28	29
LS	19	15	19
AC1	18	18	30
AC2	19	19	32
AC3	52	47	51
AC4	53	47	48
AEFM	47	38	57
CCPUE	17	16	39
SURVIV	17	24	37
XSA	12	7	22
CAGEAN	20	18	34
ADAPT			
GLM	24	23	29
COLSIS			
TSER1			25
TSER2			
SVPA			
CONVEN			

Table 4.1: Ranking of Methods - Data Sets 5 and 6

Ranking on MLR(Mean F)

	Set 5		Set 6	
Rank	Method	MLR	Method	MLR
1	CCPUE	0	LS	-4
2	LS	1	XSA	-6
3	AC1	7	HYB	-8
4	AC2	7	GLM	11
5	HYBRID	9	TSER1	17
6	AC4	-11	CCPUE	18
7	TSER1	12	AC1	19
8	AC3	-14	AC2	20
9	SURVIV	17	SURVIV	27
10	XSA	18	AEFM	32
11	GLM	24	CAGEAN	33
12	CAGEAN	-27	AC4	34
13	AEFM	35	AC3	36
?	ADAPT	*	ADAPT	*
?	COLSIS	*	COLSIS	*
?	TSER2	*	TSER2	*
?	SVPA	*	SVPA	*
?	CONVEN	*	CONVEN	*

Ranking on MLR(SSB)

	Set 5		Set 6	
Rank	Method	MLR	Method	MLR
1	AC1	0	CCPUE	-2
2	AC2	1	XSA	4
3	AEFM	-3	LS	8
4	CCPUE	4	HYBRID	8
5	XSA	-4	GLM	8
6	CAGEAN	4	AEFM	-10
7	HYBRID	5	AC1	-13
8	LS	6	AC2	-14
9	AC3	11	CAGEAN	15
10	AC4	14	SURVIV	-22
11	SURVIV	-16	AC3	33
12	GLM	-17	AC4	34
?	TSER1	*	TSER1	*
?	ADAPT	*	ADAPT	*
?	COLSIS	*	COLSIS	*
?	TSER2	*	TSER2	*
?	SVPA	*	SVPA	*
?	CONVEN	*	CONVEN	*

Ranking on RMS(Mean F)

	Set 5		Set 6	
Rank	Method	RMS	Method	RMS
1	LS	14	LS	19
2	SURVIV	19	XSA	22
3	AC1	19	TSER1	25
4	AC2	20	HYBRID	29
5	AC3	22	GLM	29
6	XSA	22	AC1	30
7	AC4	24	AC2	32
8	HYBRID	26	CAGEAN	34
9	CAGEAN	28	SURVIV	37
10	GLM	30	CCPUE	39
11	TSER1	34	AC4	48
12	CCPUE	37	AC3	51
13	AEFM	46	AEFM	57
?	ADAPT	*	ADAPT	*
?	COLSIS	*	COLSIS	*
?	TSER2	*	TSER2	*
?	SVPA	*	SVPA	*
?	CONVEN	*	CONVEN	*

Ranking on RMS(SSB)

	Set 5		Set 6	
Rank	Method	RMS	Method	RMS
1	XSA	7	XSA	7
2	CAGEAN	9	LS	15
3	AC1	13	CCPUE	16
4	LS	14	CAGEAN	18
5	AC2	14	AC1	18
6	SURVIV	19	AC2	19
7	AC3	20	GLM	23
8	HYBRID	20	SURVIV	24
9	GLM	22	HYBRID	28
10	AC4	23	AEFM	38
11	AEFM	26	AC3	47
12	CCPUE	28	AC4	47
?	TSER1	*	TSER1	*
?	ADAPT	*	ADAPT	*
?	COLSIS	*	COLSIS	*
?	TSER2	*	TSER2	*
?	SVPA	*	SVPA	*
?	CONVEN	*	CONVEN	*

Table 4.2 : Simulated Data Set 1

MLR and RMS of Mean F for each Method

100*RMS \ 100*MLR	0-9	10-19	20-29	30-39	40-49	>50
0-9	LS AC1 AC2 AC3 AC4 CCPUE SURVIV CAGEAN ADAPT GLM					
10-19	HYBRID AEFM					
20-29						
30-39						
40-49				SVPA		
>=50			CONVEN			XSA

Not included : COLSIS : TSER1 : TSER2

Table 4.3 : Simulated Data Set 1

MLR and RMS of SSB for each Method

100*RMS \ 100*MLR	0-9	10-19	20-29	30-39	40-49	>50
0-9	HYBRID LS AC1 AC2 AC3 AC4 AEFM CCPUE SURVIV CAGEAN ADAPT GLM					
10-19						
20-29						
30-39				XSA		
40-49			SVPA			
>=50			CONVEN			

Not included : COLSIS : TSER1 : TSER2

Table 4.4 : Simulated Data Set 2

MLR and RMS of Mean F for each Method

100*RMS	100*MLR 0-9	10-19	20-29	30-39	40-49	>50
0-9	LS					
	AC1					
	AC2					
	AC3					
	AC4					
	CAGEAN					
10-19	HYBRID					
	AEFM					
	CCPUE					
	SURVIV					
	XSA					
20-29		GLM	CONVEN			
30-39			ADAPT			
40-49					SVPA	
>=50						

Not included : COLSIS : TSER1 : TSER2

Table 4.5 : Simulated Data Set 2

MLR and RMS of SSB for each Method

100*RMS	100*MLR 0-9	10-19	20-29	30-39	40-49	>50
0-9	HYBRID					
	LS					
	AC1					
	AC2					
	AC3					
	AC4					
	AEFM					
	CCPUE					
	SURVIV					
	XSA					
	CAGEAN					
	GLM	ADAPT				
10-19						
20-29			CONVEN			
30-39					SVPA	
40-49						
>=50						

Not included : COLSIS : TSER1 : TSER2

Table 4.6 : Simulated Data Set 3

MLR and RMS of Mean F for each Method

100*MLR

100*RMS	0-9	10-19	20-29	30-39	40-49	>50
0-9	CAGEAN SVPA					
10-19	LS AC3 AC4 AEFM CCPUE SURVIV GLM	AC1 AC2 CONVEN				
20-29		HYBRID XSA ADAPT				
30-39						
40-49						
>=50						

Not included : COLSIS : TSER1 : TSER2

Table 4.7 : Simulated Data Set 3

MLR and RMS of SSB for each Method

100*MLR

100*RMS	0-9	10-19	20-29	30-39	40-49	>50
0-9	LS AC1 AC2 AC3 AC4 AEFM SURVIV XSA CAGEAN GLM SVPA CONVEN					
10-19	HYBRID CCPUE	ADAPT				
20-29						
30-39						
40-49						
>=50						

Not included : COLSIS : TSER1 : TSER2

Table 4.8 : Simulated Data Set 4

MLR and RMS of Mean F for each Method

100*RMS	100*MLR 0-9	10-19	20-29	30-39	40-49	>50
0-9	HYBRID AC1 AC2 AC4					
10-19	AC3 AEFM CCPUE	LS SURVIV GLM CONVEN				
20-29			XSA SVPA			
30-39						
40-49						
>=50						CAGEAN

Not included : COLSIS : TSER1 : TSER2

Table 4.9 : Simulated Data Set 4

MLR and RMS of SSB for each Method

100*RMS	100*MLR 0-9	10-19	20-29	30-39	40-49	>50
0-9	HYBRID AC1 AC2 AC3 AC4 AEFM CCPUE SURVIV GLM					
10-19	XSA	LS CONVEN				
20-29		SVPA				
30-39						
40-49				CAGEAN		
>=50						

Not included : COLSIS : TSER1 : TSER2

Table 4.10 : Simulated Data Set 5

MLR and RMS of Mean F for each Method

100*RMS \ 100*MLR	0-9	10-19	20-29	30-39	40-49	>50
0-9						
10-19	LS AC1	SURVIV				
20-29	HYBRID AC2	AC3 AC4 XSA	CAGEAN			
30-39	CCPUE	TSER1	GLM			
40-49				AEFM		
>=50						

Not included : COLSIS : ADAPT : TSER2 : SVPA : CONVEN

Table 4.11 : Simulated Data Set 5

MLR and RMS of SSB for each Method

100*RMS \ 100*MLR	0-9	10-19	20-29	30-39	40-49	>50
0-9	XSA CAGEAN					
10-19	LS AC1 AC2 SURVIV					
20-29	HYBRID AEFM CCPUE	AC3 AC4 GLM				
30-39						
40-49						
>=50						

Not included : COLSIS : ADAPT : TSER1 : TSER2 : SVPA : CONVEN

Table 4.12: Simulated Data Set 6

MLR and RMS of Mean F for each Method

100*RMS \ 100*MLR	0-9	10-19	20-29	30-39	40-49	>50
0-9						
10-19	LS					
20-29	HYBRID XSA GLM					
30-39		AC1 CCPUE TSER1	AC2 SURVIV	CAGEAN		
40-49				AC4		
>=50			AEFM	AC3		

Not included : COLSIS : ADAPT : TSER2 : SVPA : CONVEN

Table 4.13: Simulated Data Set 6

MLR and RMS of SSB for each Method

100*RMS \ 100*MLR	0-9	10-19	20-29	30-39	40-49	>50
0-9	XSA					
10-19	LS CCPUE	AC1 AC2 CAGEAN				
20-29	HYBRID GLM		SURVIV			
30-39		AEFM				
40-49				AC3 AC4		
>=50						

Not included : COLSIS : ADAPT : TSER1 : TSER2 : SVPA : CONVEN

Table 4.14: Real Data Set : HADDOCK in North Sea

No estimates available for XSA, ADAPT, GLM, TSER1, TSER2, SVPA, CONVEN

Estimates of Number at Age in 1986

	Age										
Method	0	1	2	3	4	5	6	7	8	9	10
HYBRID	41095	3444	403	525	39	13	2	4	1	0	0
LS	40148	3602	448	513	40	15	2	4	1	0	0
AC1	17638	3726	391	469	36	14	2	3	1	0	0
AC2	17683	3733	380	405	33	14	3	3	1	0	0
AC3	85518	4346	460	524	41	14	2	4	1	1	0
AC4	86927	4346	452	481	39	14	2	4	1	1	0
AEFM	76601	4874	466	514	40	14	2	3	0	0	0
CCPUE	50182	3820	432	537	38	13	2	3	0	0	0
SURVIV	5584	2290	338	405	31	19	4	3	0	0	0
CAGEAN	na	6989	739	777	50	18	3	6	na	na	na
COLSIS	na	1924	249	202	39	*	*	*	*	*	*
WG(88)	36956	2959	322	485	39	16	3	4	1	0	0

Estimates of F at Age in 1986

	Age										
Method	0	1	2	3	4	5	6	7	8	9	10
HYBRID	3	99	735	1136	1457	1402	1055	757	428	496	485
LS	3	94	635	1184	1283	1105	887	734	503	748	871
AC1	8	91	768	1408	1682	1187	740	765	300	553	534
AC2	8	91	801	2013	2208	1219	597	753	257	557	522
AC3	2	78	612	1139	1283	1181	920	681	353	242	425
AC4	2	78	628	1341	1418	1256	1031	661	331	212	402
AEFM	2	69	602	1182	1331	1183	1057	993	960	865	865
CCPUE	3	89	667	1090	1441	1365	1091	912	959	864	864
SURVIV	24	151	959	2007	2529	757	411	1280	1280	1280	1280
CAGEAN	na	94	456	998	529	527	526	526	na	na	na
COLSIS	na	181	1710	*	1425	*	*	*	*	*	*
WG(88)	4	115	1033	1305	1493	1053	820	722	767	971	971

* indicates that catch bigger than estimated number in sea

Estimates of Total and Spawning Biomass and Mean F (ages 2-6) in 1986

Method	TSB	SSB	Mean F
HYBRID	1803	223	1156
LS	1808	226	849
AC1	787	208	1157
AC2	758	188	1368
AC3	905	232	1027
AC4	885	218	1134
AEFM	964	227	1071
CCPUE	1047	228	1131
SURVIV	570	184	1160
CAGEAN	1410	337	607
COLSIS	402	99	na
WG(88)	682	207	1141

Table 4.15: Real Data Set : COD in North Sea

No estimates available for XSA, ADAPT, GLM, TSER1, TSER2, SVPA, CONVEN

Estimates of Number at Age in 1986

Method	Age 1	2	3	4	5	6	7	8	9	10	11
HYBRID	732	33	48	6	4	1	1	0	0	0	0
LS	786	35	54	7	5	1	1	0	0	0	0
AC1	666	31	50	8	5	1	1	0	0	0	0
AC2	657	30	52	7	3	1	1	0	0	0	0
AC3	924	36	50	7	5	1	1	0	0	0	0
AC4	1010	37	52	7	5	1	1	0	0	0	0
AEFM	1029	38	55	7	5	1	1	0	0	0	0
CCPUE	615	33	53	6	5	1	1	0	0	0	0
SURVIV	595	28	58	7	3	1	1	0	0	0	0
CAGEAN	1271	50	54	6	4	1	1	na	na	na	na
COLSIS	na	23	121	18	15	6	5	3	3	3	3
WG(88)	581	37	52	8	5	1	1	0	0	0	0

Estimates of F at Age * 1000 in 1986

Method	Age 1	2	3	4	5	6	7	8	9	10	11
HYBRID	167	1184	1098	1176	1003	1353	1926	1801	1222	1409	3672
LS	154	1065	894	1025	850	1132	1840	1341	930	938	1066
AC1	184	1325	1014	789	800	777	873	804	819	790	1009
AC2	188	1496	967	848	1541	1084	2746	1368	1147	1417	1571
AC3	130	1063	1030	851	823	903	883	1024	970	1039	2035
AC4	118	990	971	888	902	1205	1161	1154	1026	1165	2347
AEFM	200	966	881	933	821	1107	881	937	869	613	609
CCPUE	201	1192	930	1042	841	1217	1481	1081	714	611	606
SURVIV	208	1677	804	930	1641	910	910	910	910	910	910
CAGEAN	102	1665	1380	974	974	974	974	na	na	na	na
COLSIS	na	3336	313	262	195	99	104	115	115	115	115
WG(88)	216	854	920	840	761	817	711	676	731	937	1528

Estimates of Total and Spawning Biomass and Mean F (ages 3-8) in 1986

Method	TSB	SSB	Mean F	
HYBRID	600	80	1393	
LS	651	88	1180	
AC1	583	95	842	
AC2	561	80	1425	
AC3	729	93	919	
AC4	778	90	1046	
AEFM	798	95	925	
CCPUE	550	87	1093	
SURVIV	533	80	1018	
CAGEAN	510	81	1055	(omits estimates for ages 8 and older)
COLSIS	539	313	181	(omits 1-group in biomass)
WG(88)	553	100	797	

Figure 4.1

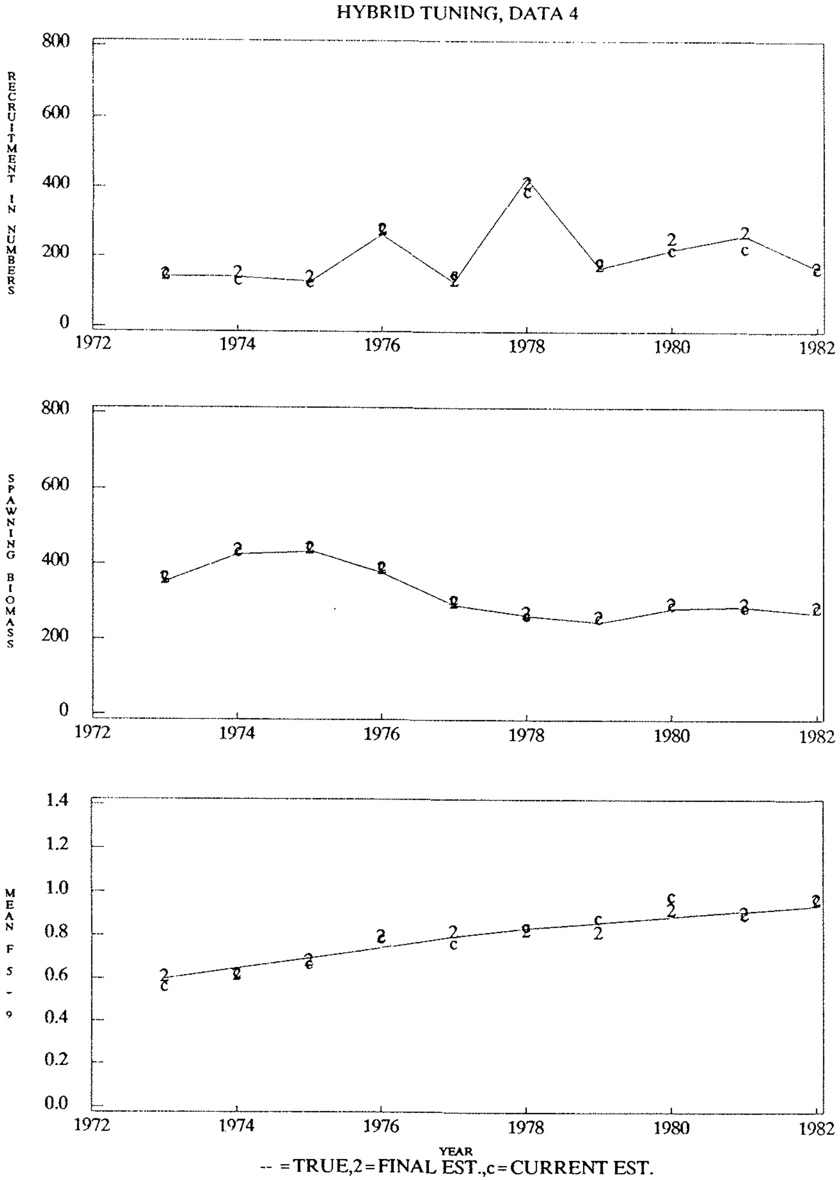


Figure 4.2

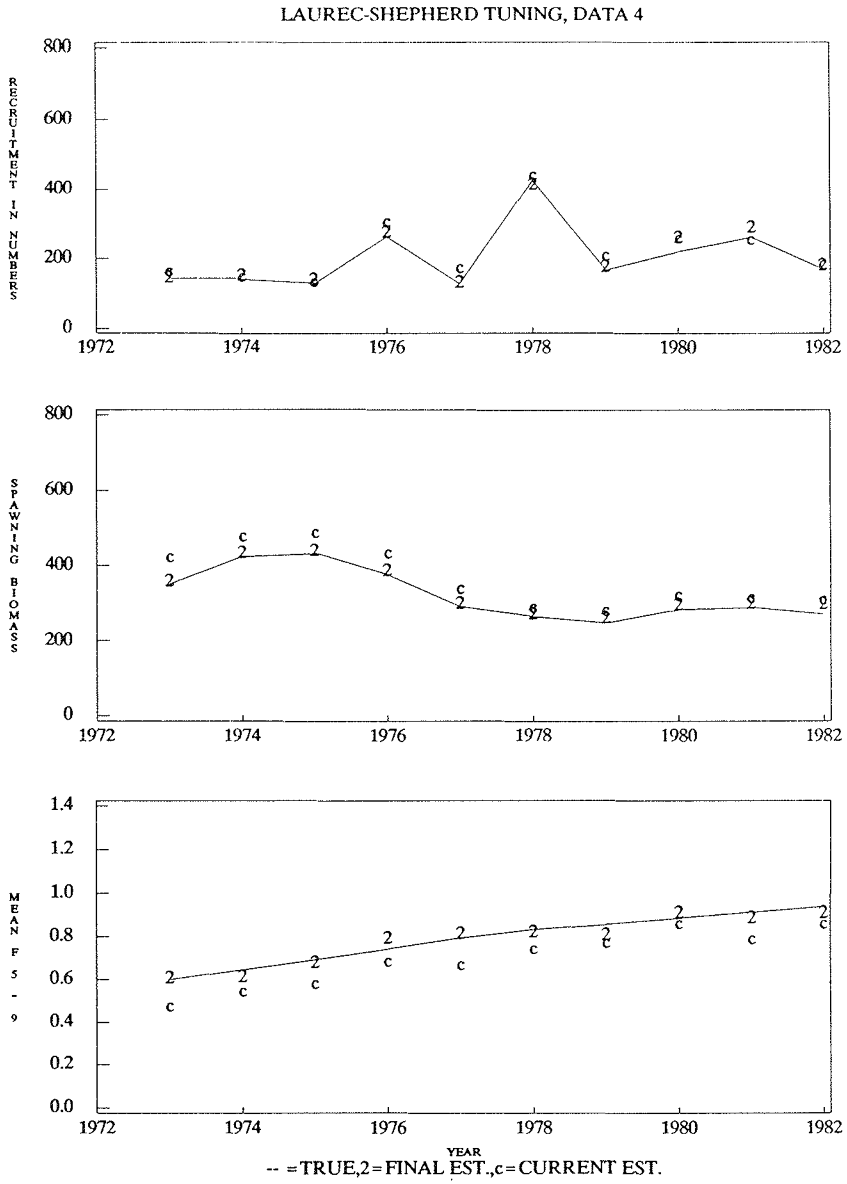


Figure 4.3

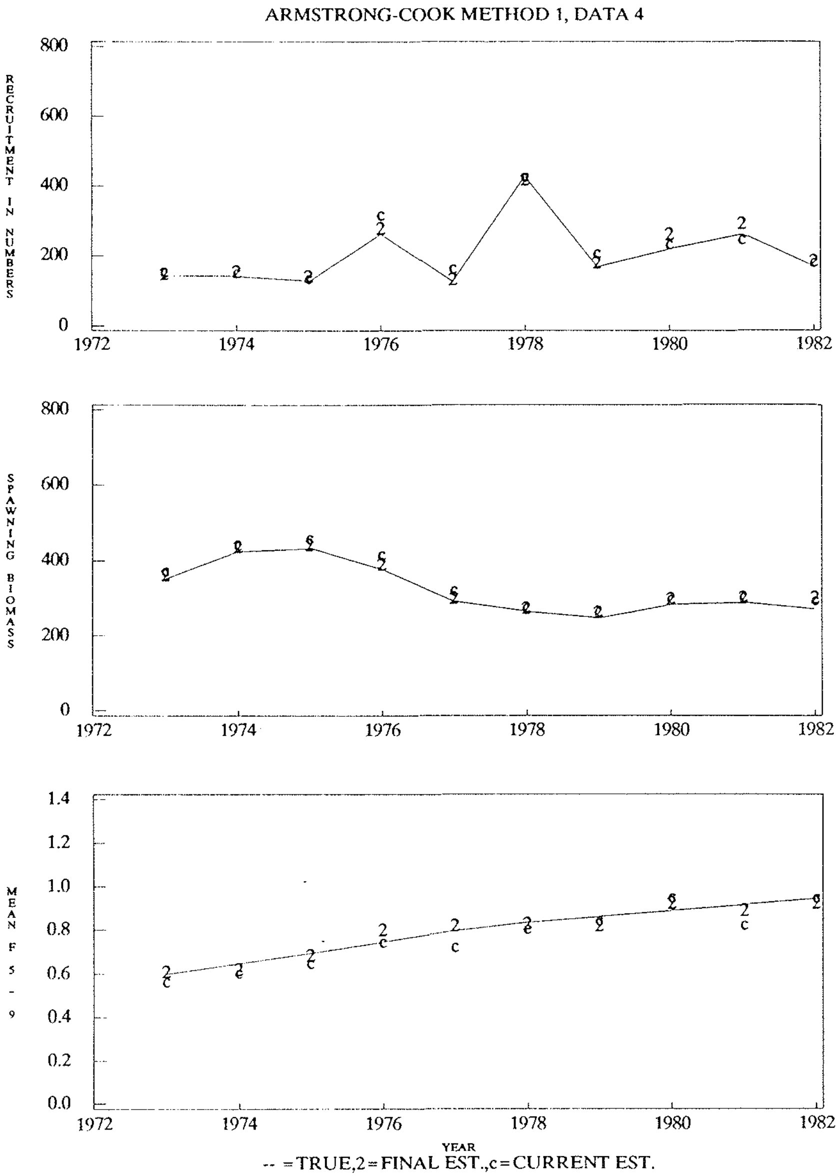


Figure 4.4

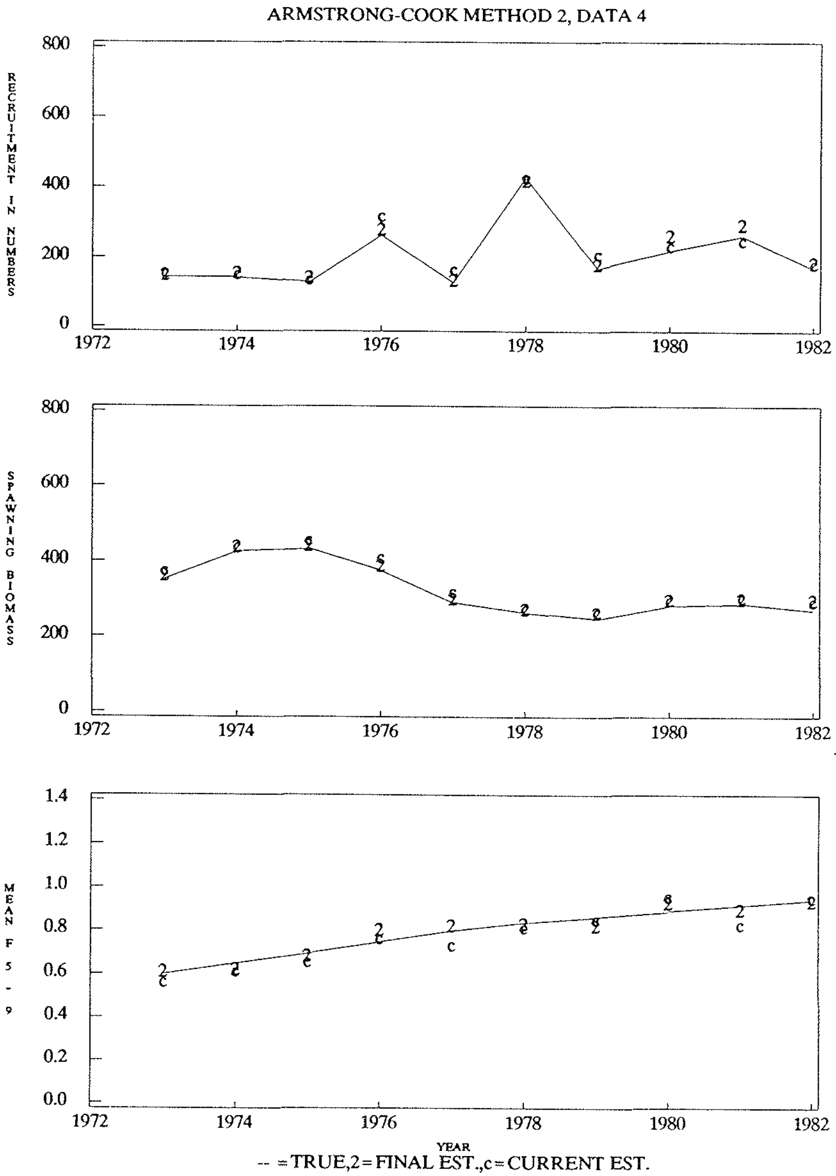


Figure 4.5

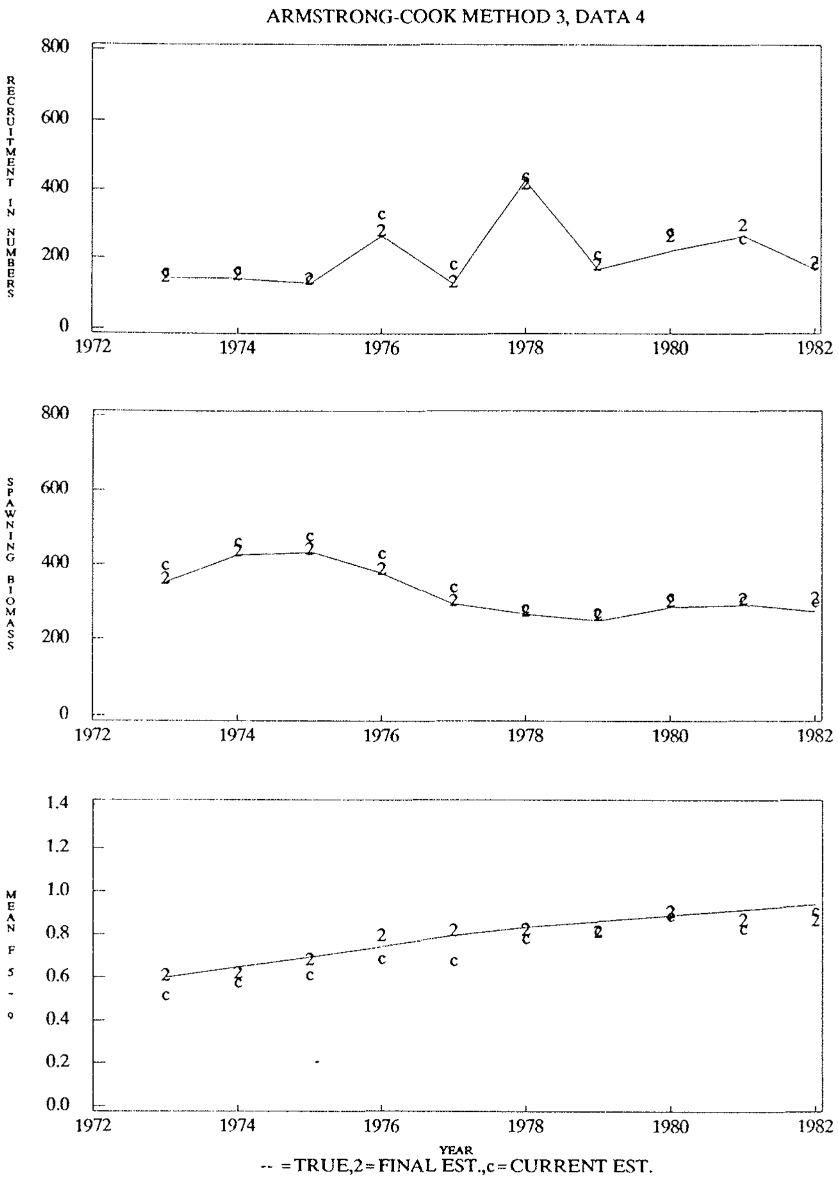


Figure 4.6

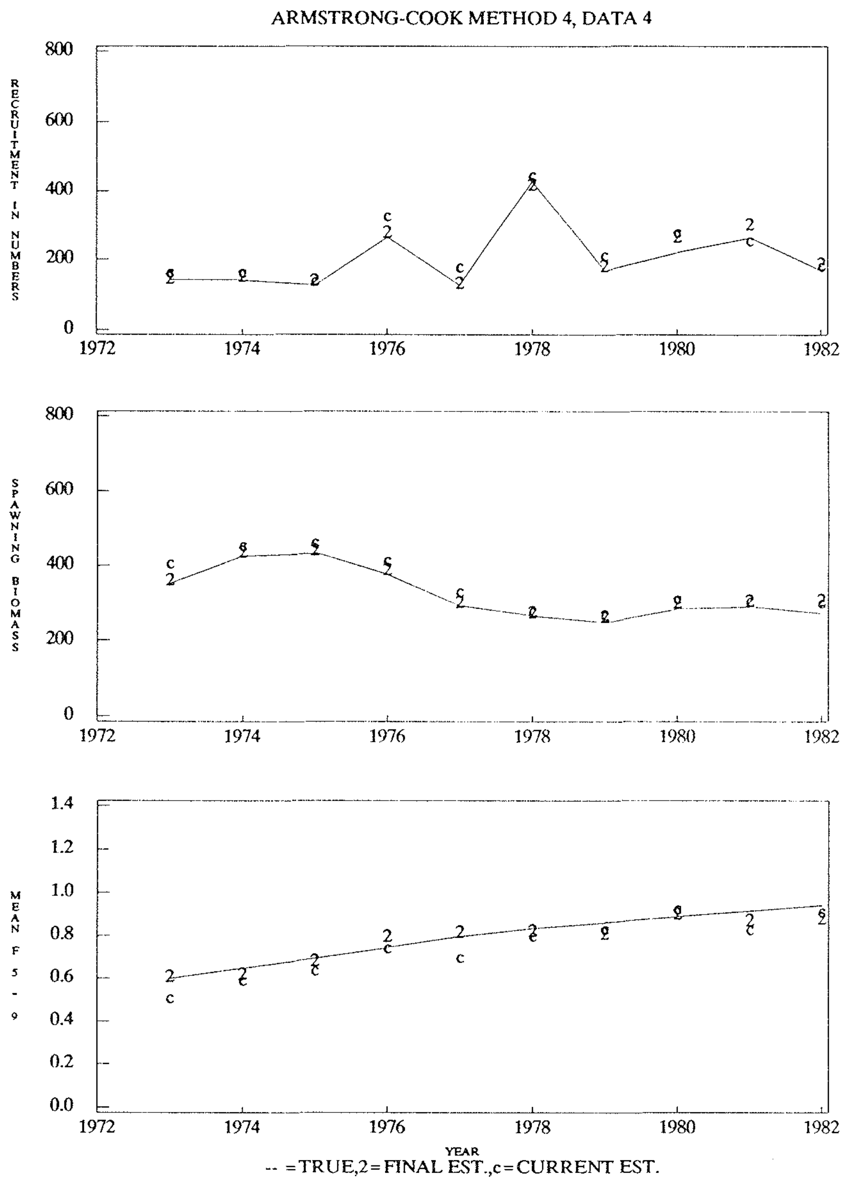


Figure 4.7

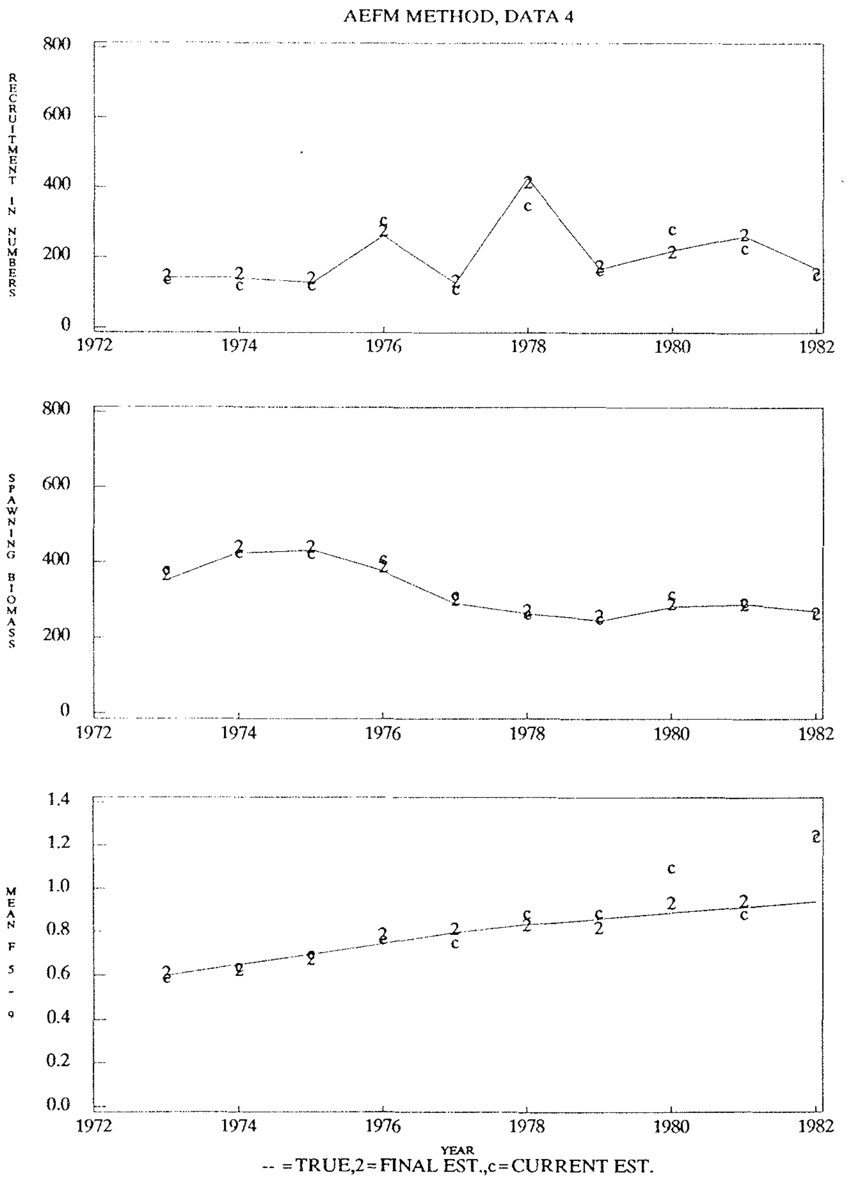


Figure 4.8

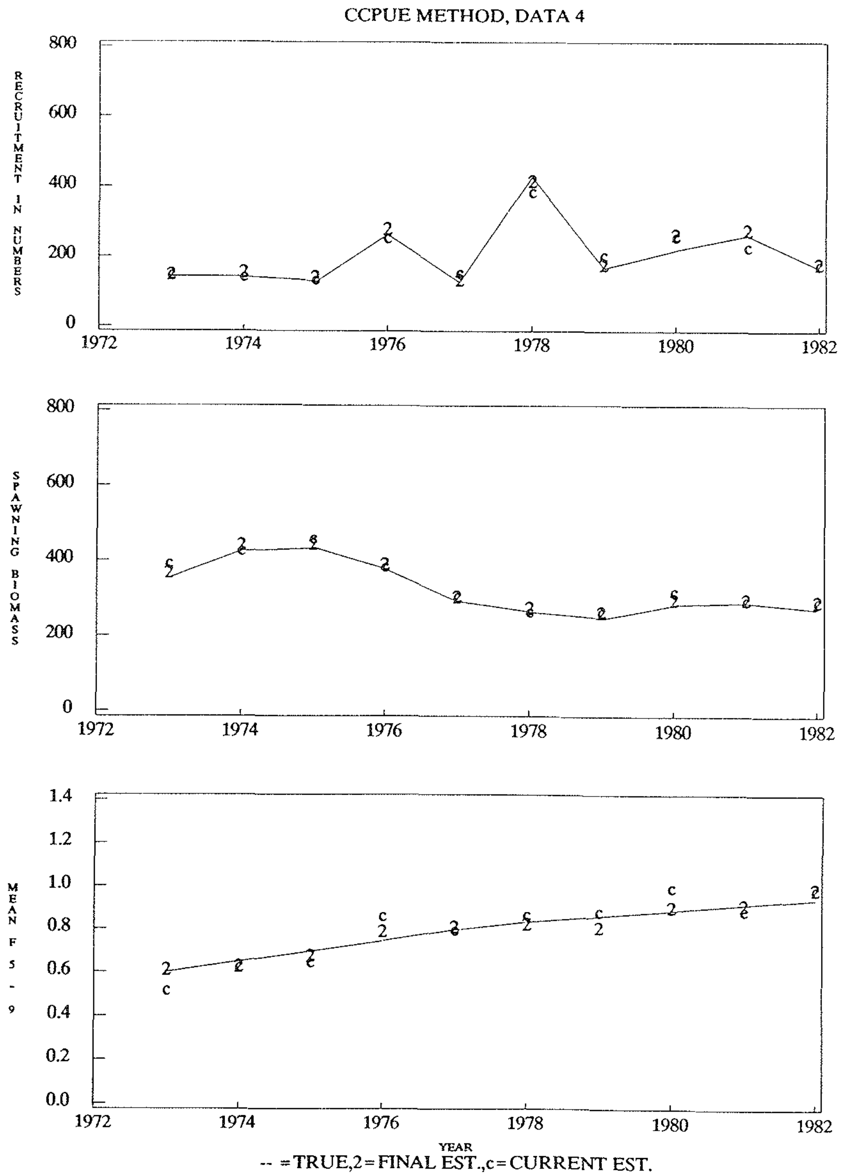
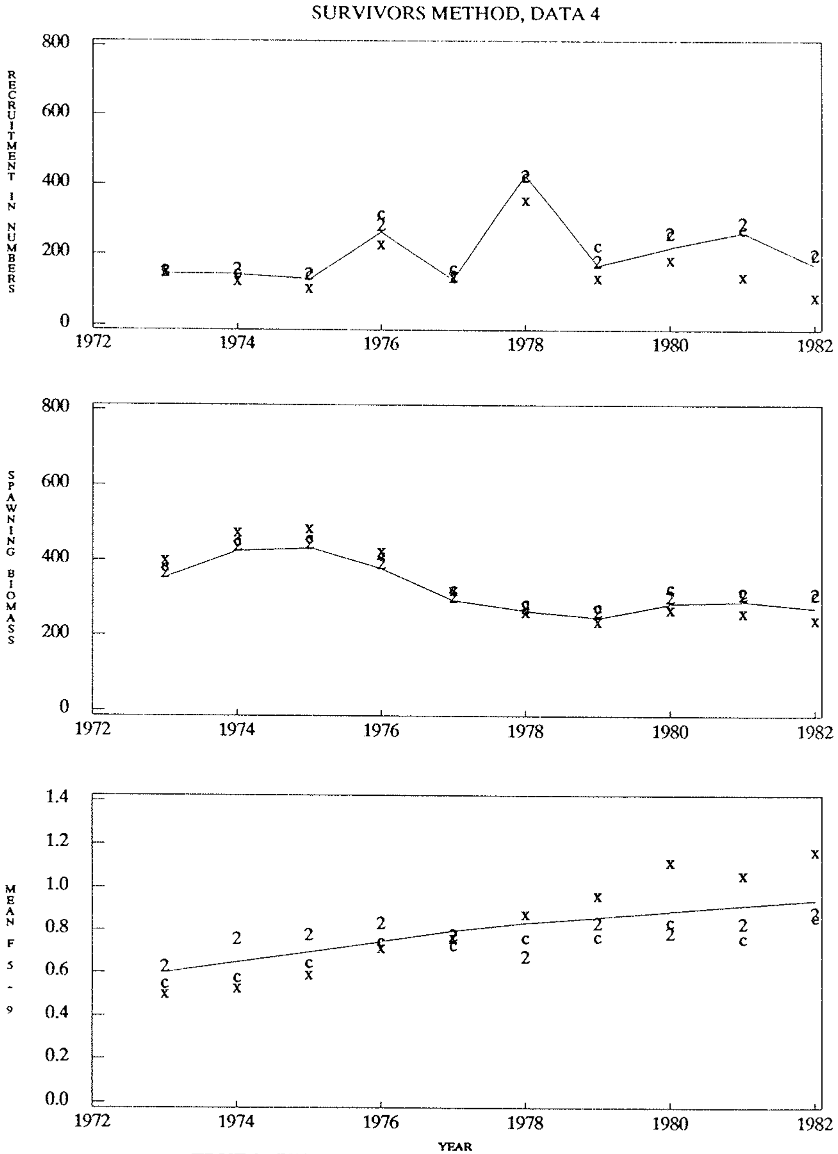
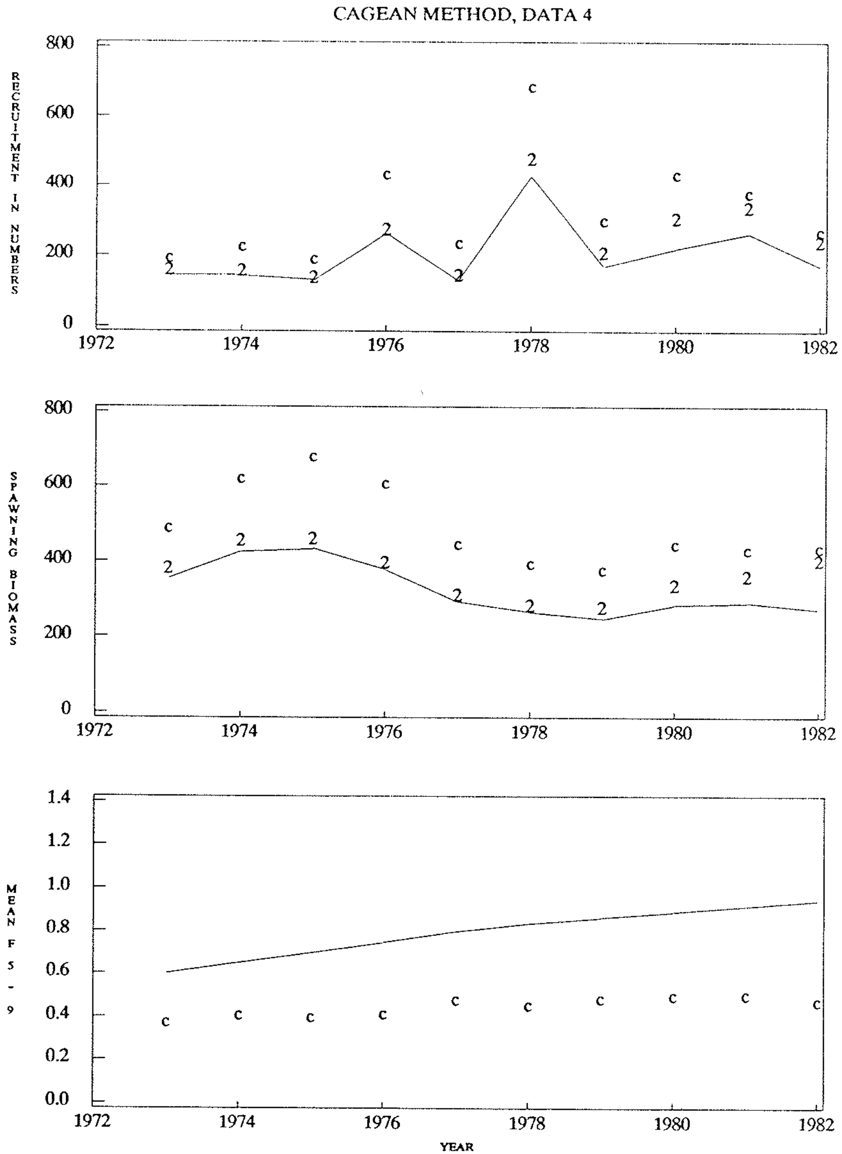


Figure 4.9



-- =TRUE,2=FINAL EST.,c=CURRENT EST.,x=EXTENDED SURV.

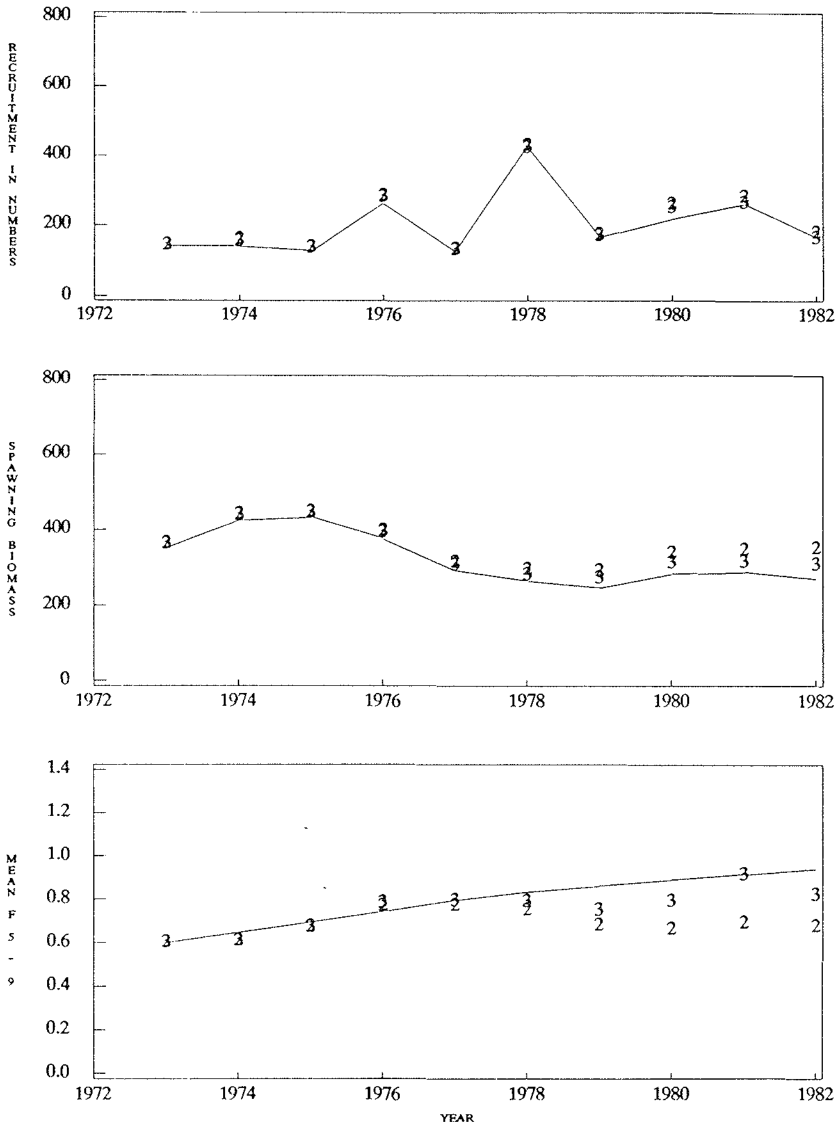
Figure 4.10



-- =TRUE,2=FINAL EST.,c=CURRENT EST.

Figure 4.11

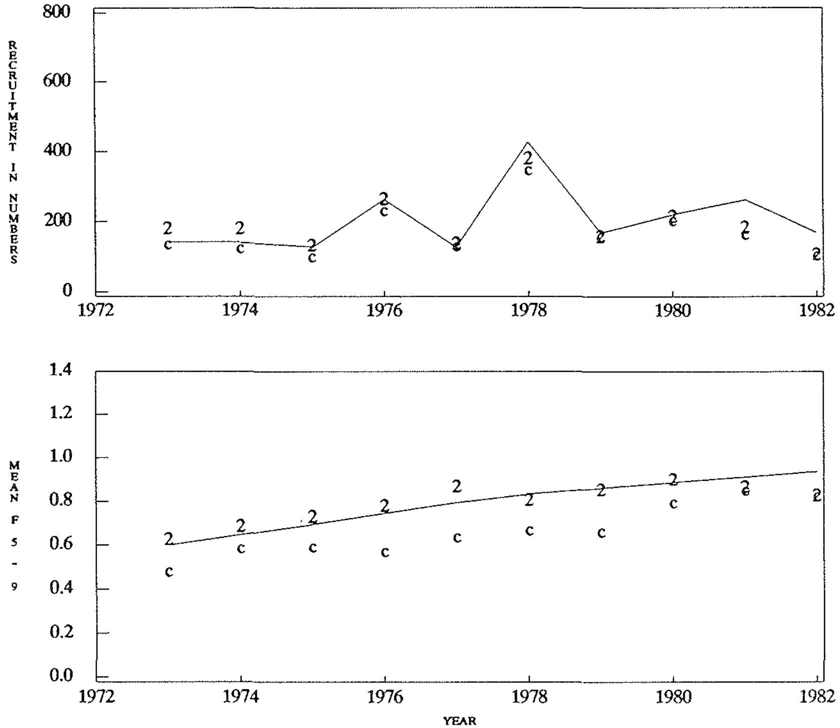
ADAPTED METHOD, DATA 4



-- = TRUE, 2 = FINAL EST., 3 = FINAL EST. W. TREND

Figure 4.12

GLM METHOD, DATA 4



-- = TRUE, 2 = FINAL EST., c = CURRENT EST.

Figure 4.13

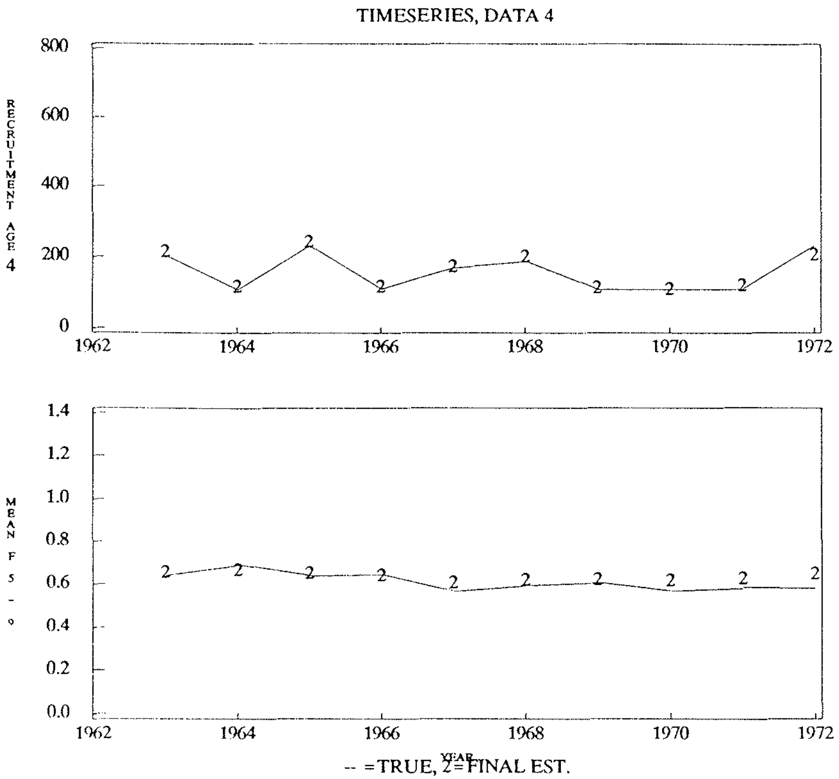


Figure 4.14

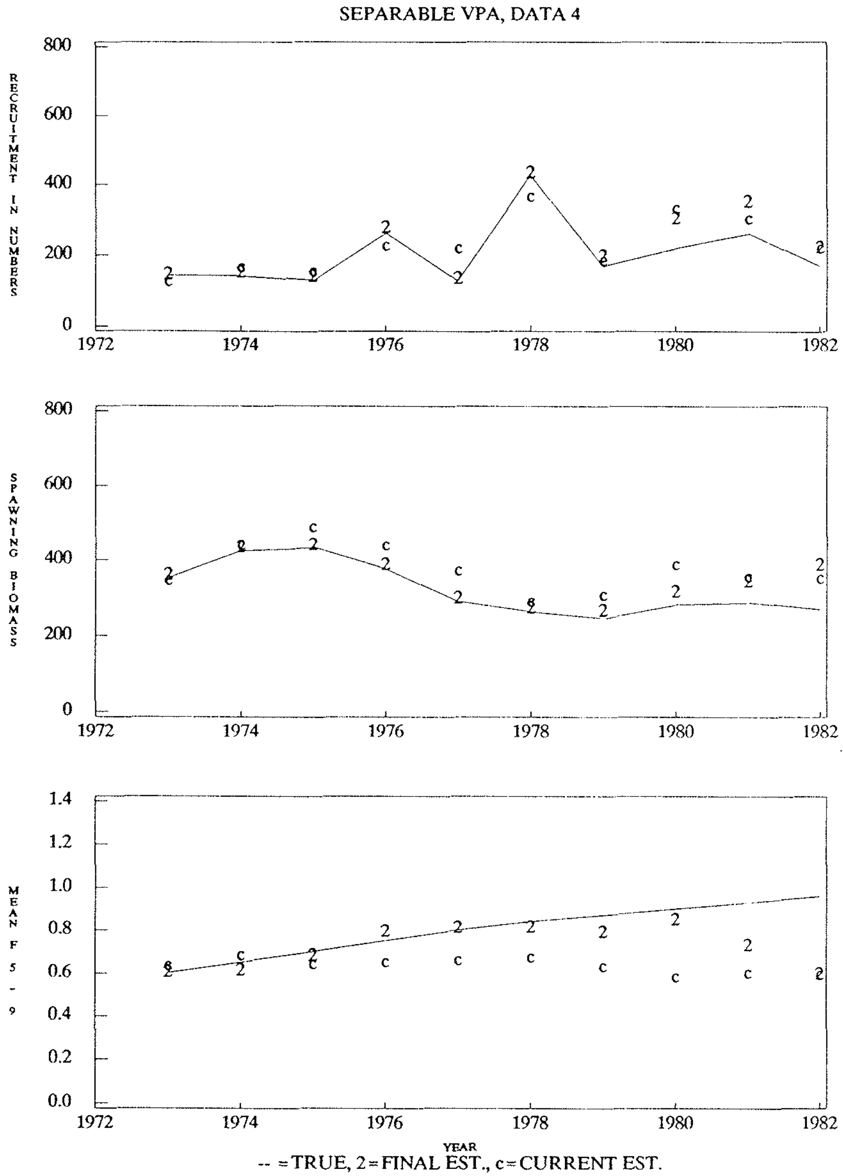


Figure 4.15

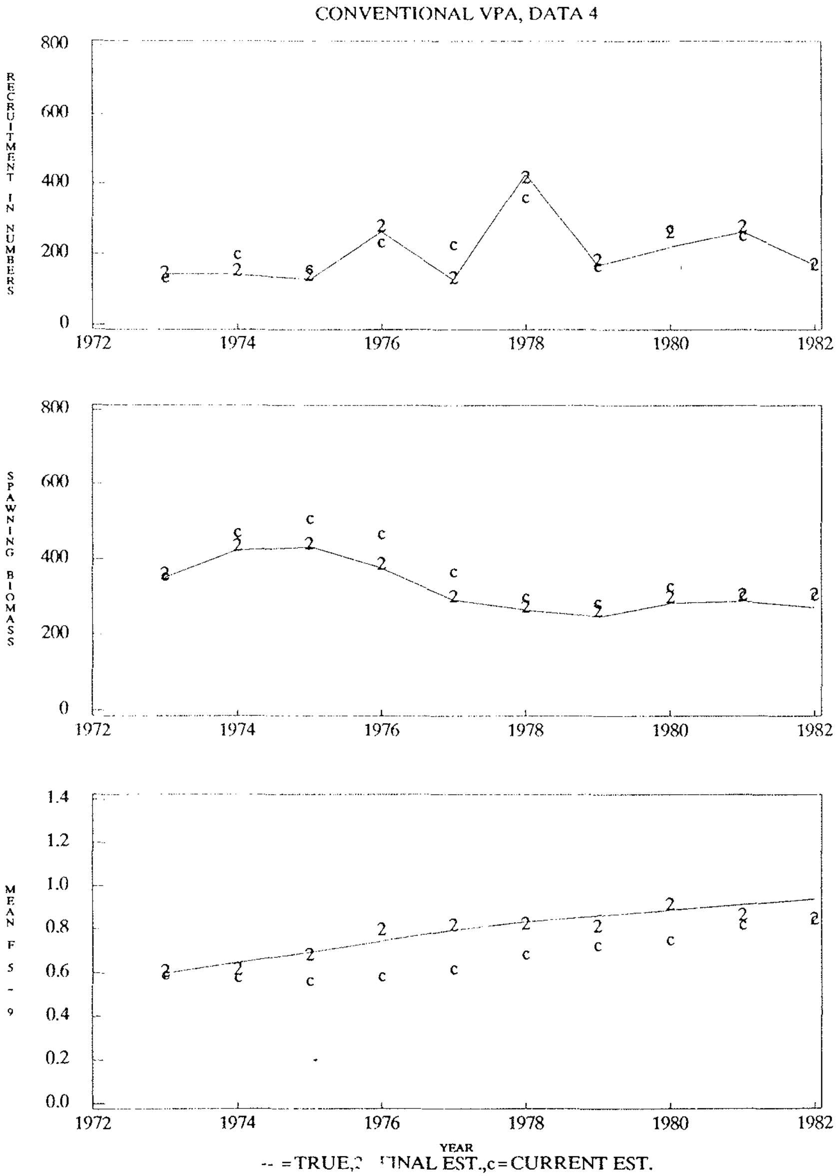


Figure 4.16

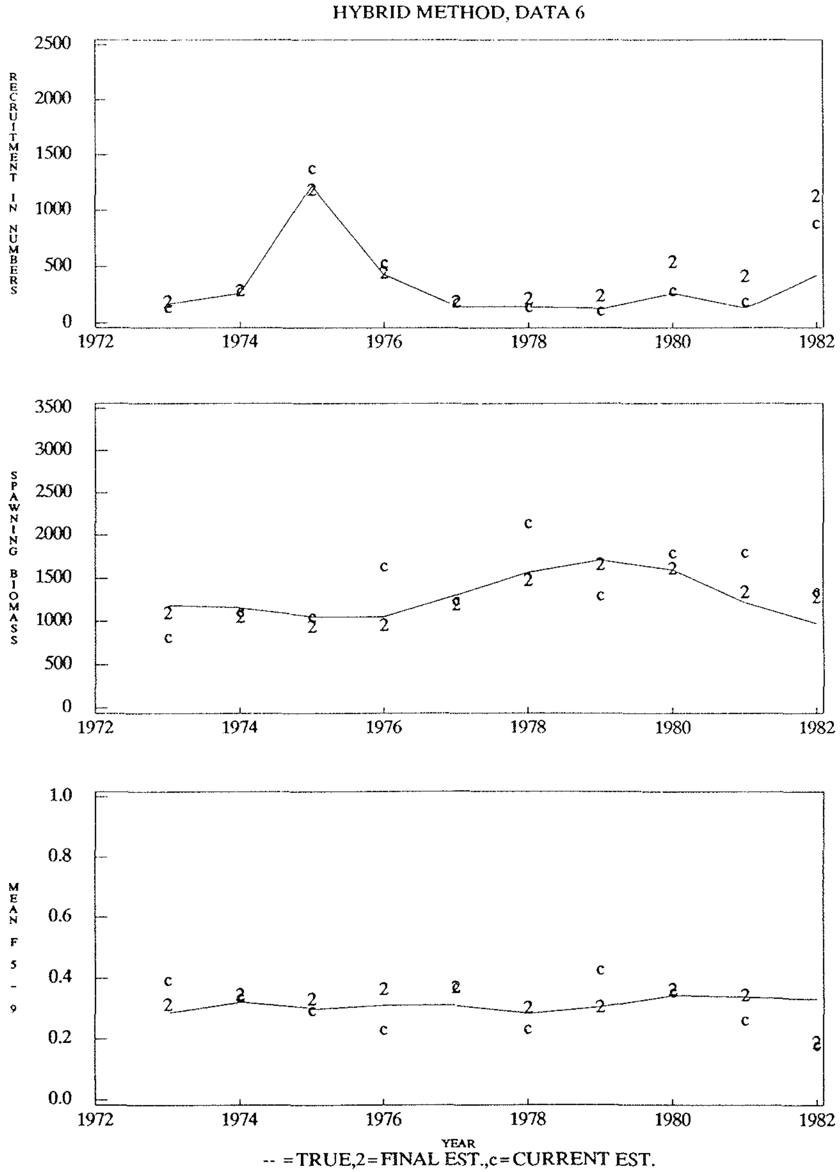


Figure 4.17

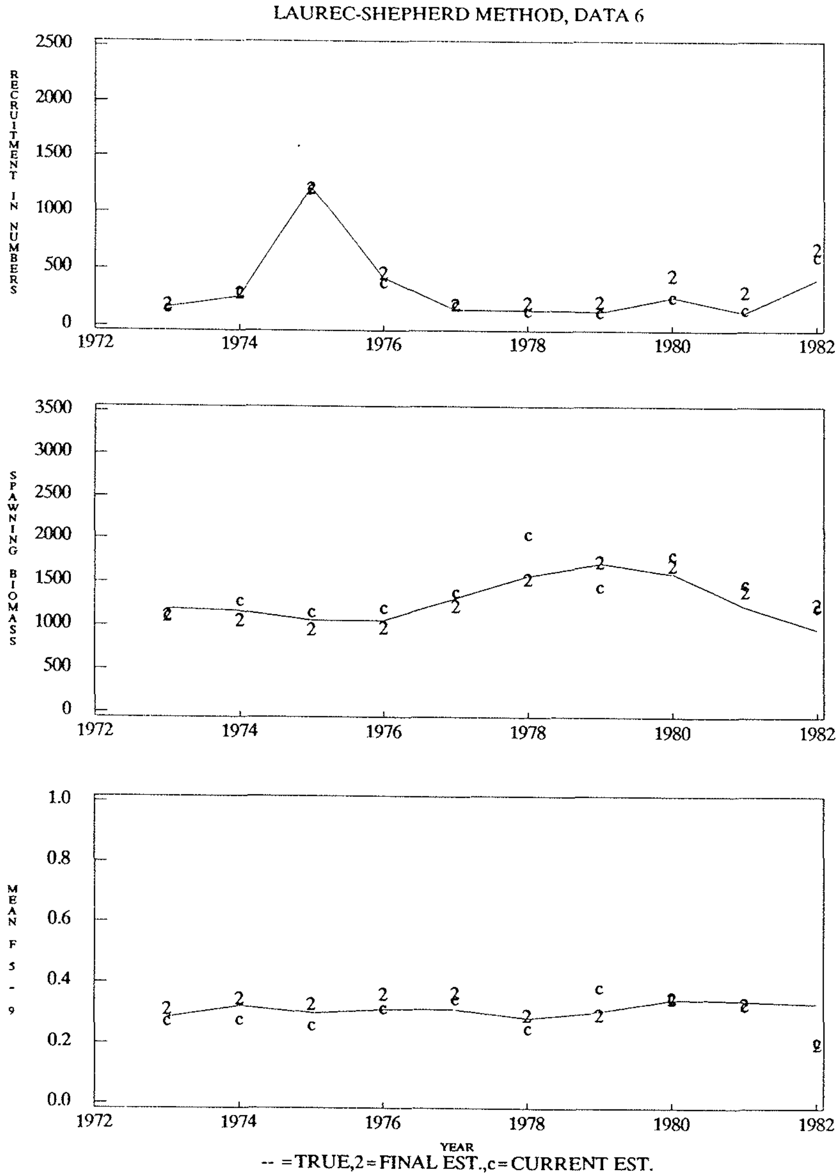


Figure 4.18

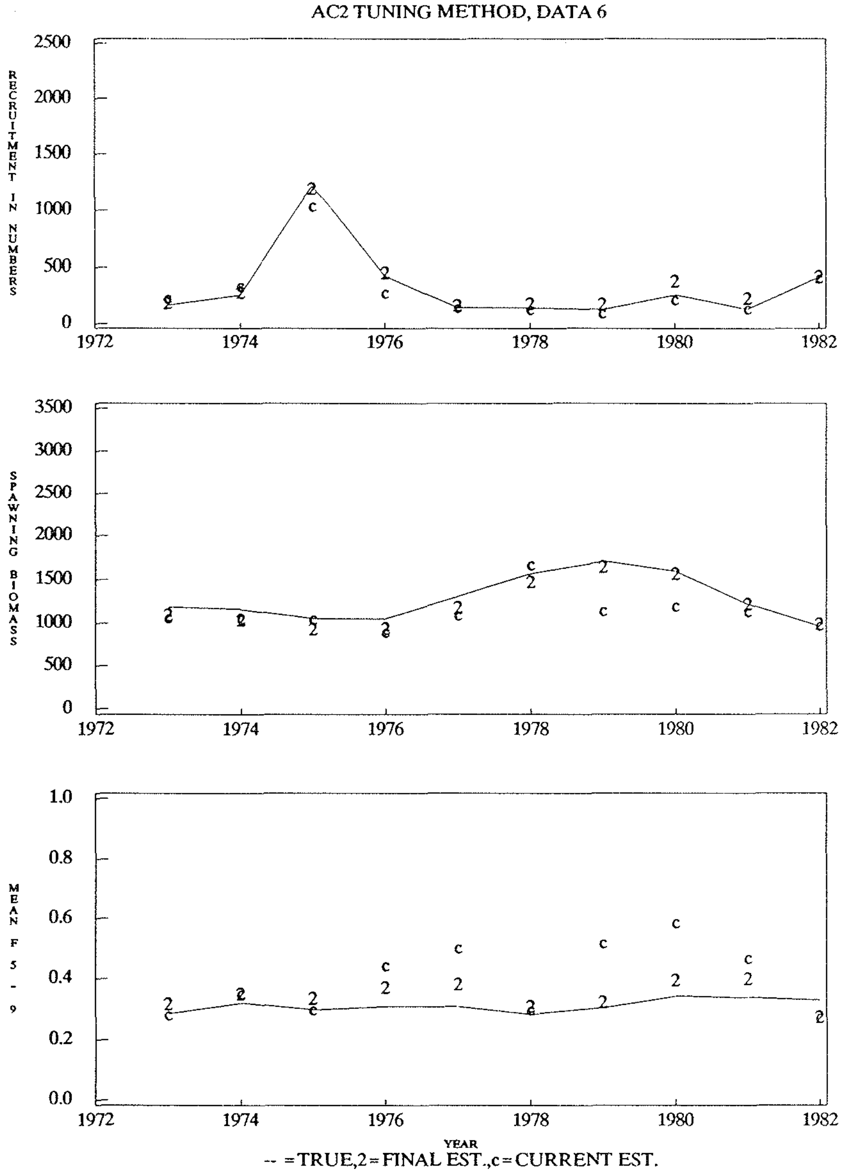
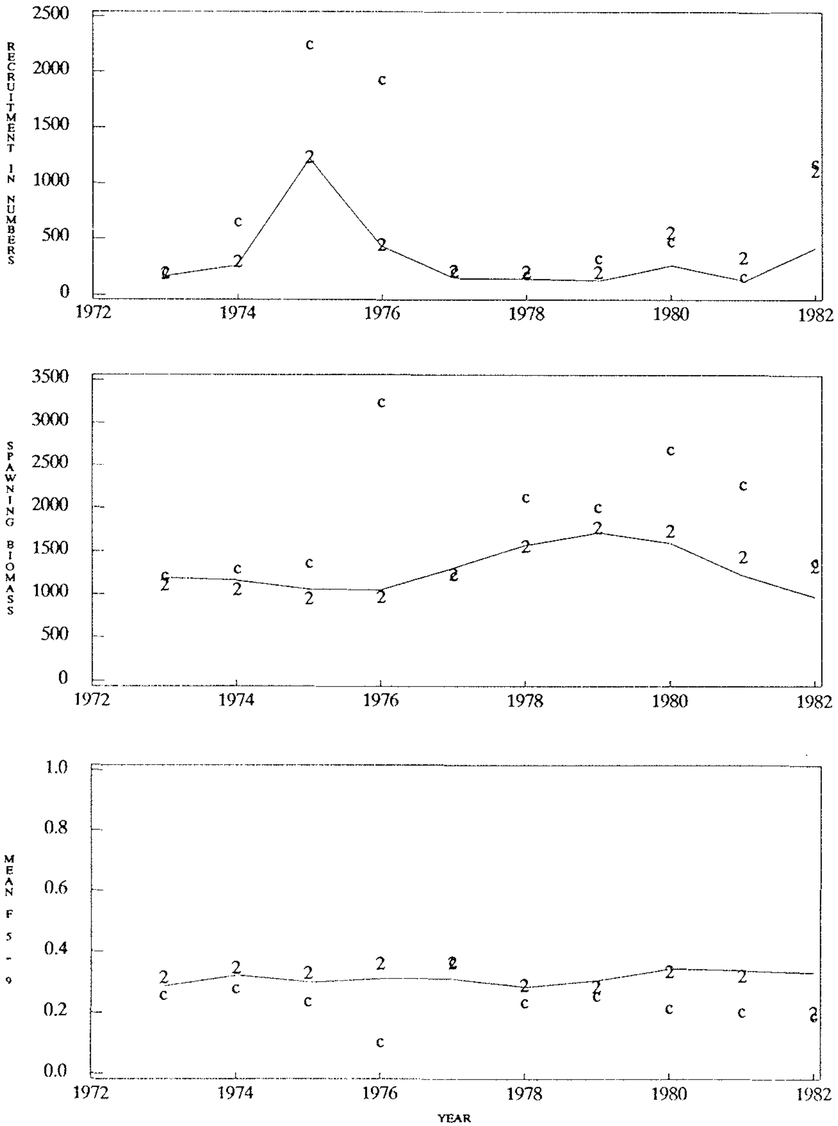


Figure 4.19

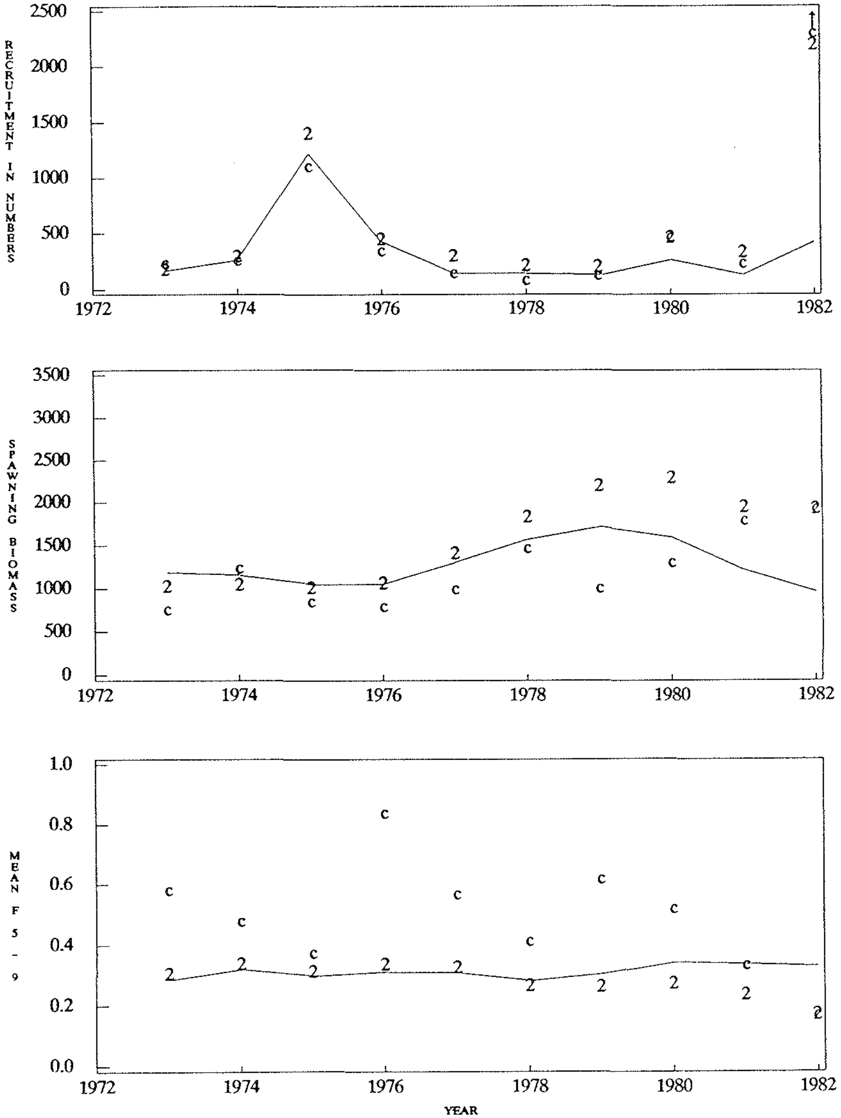
AC3 TUNING METHOD, DATA 6



-- =TRUE,2=FINAL EST.,c=CURRENT EST.

Figure 4.20

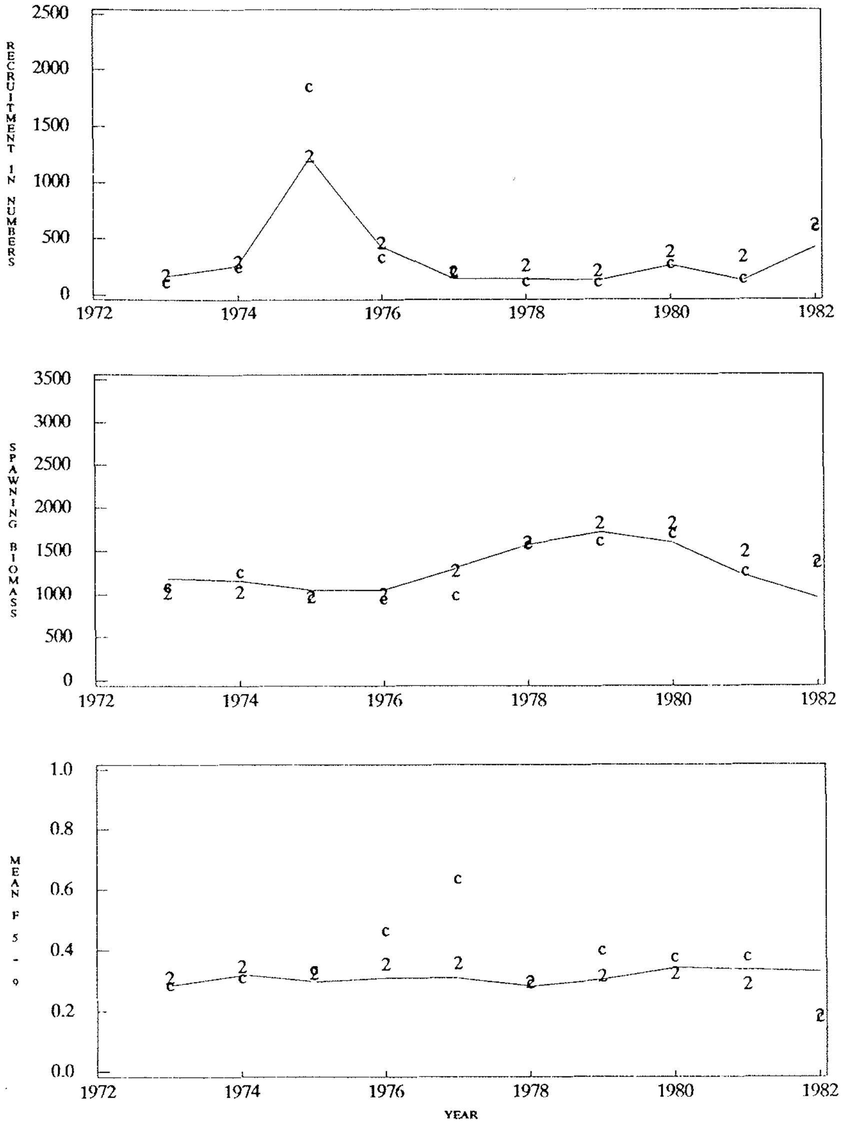
AEFM TUNING METHOD, DATA 6



-- =TRUE,2=FINAL EST.,c=CURRENT EST.

Figure 4.21

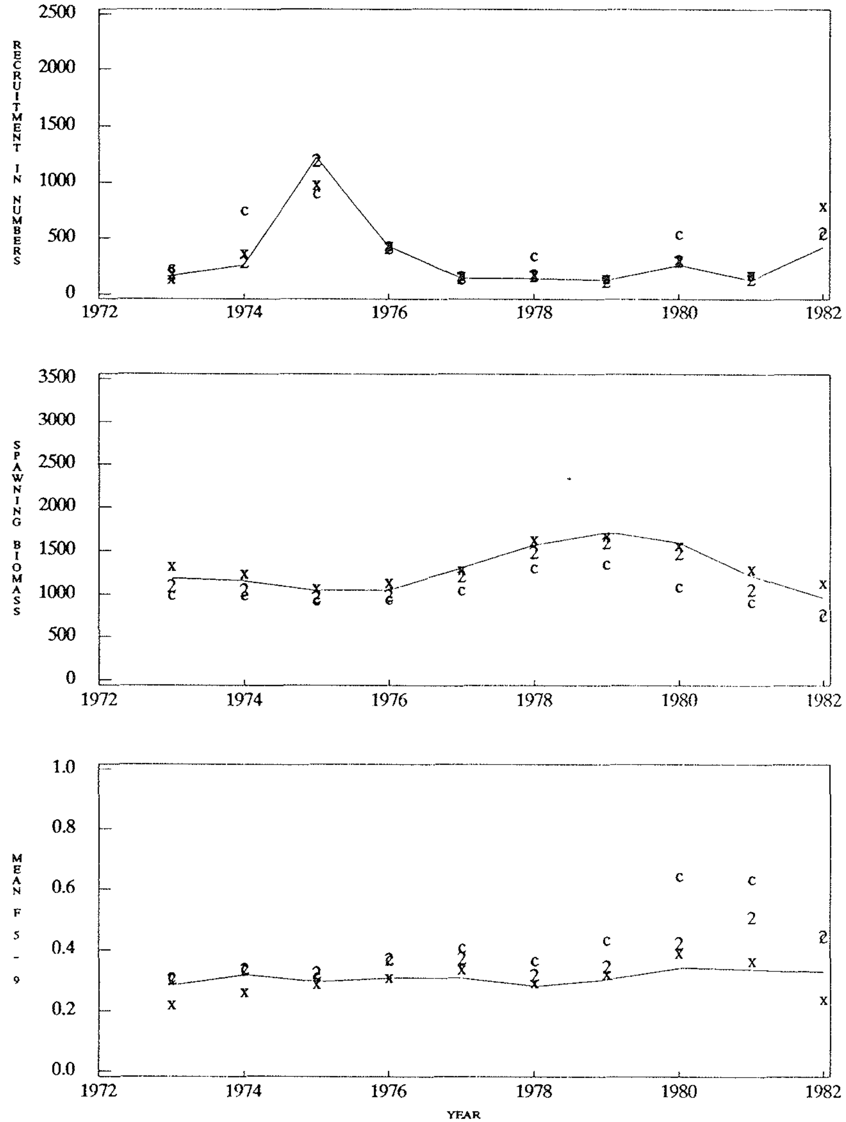
CCPUE TUNING METHOD, DATA 6



-- =TRUE,2=FINAL EST.,c=CURRENT EST.

Figure 4.22

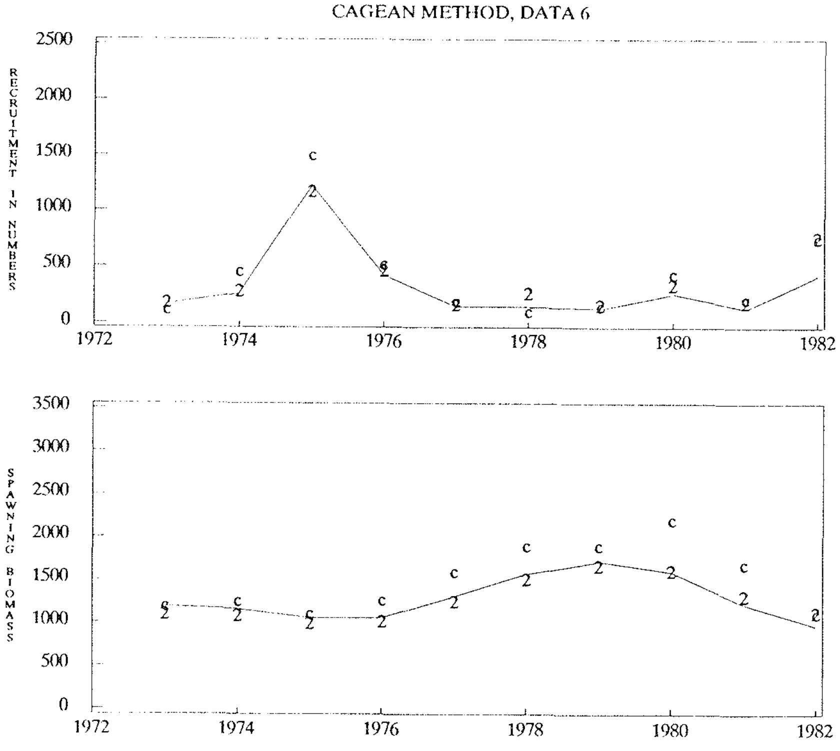
SURVIVORS METHOD, DATA 6



-- =TRUE,2=FINAL EST.,c=CURRENT EST.,x=EXTENDED SURV.

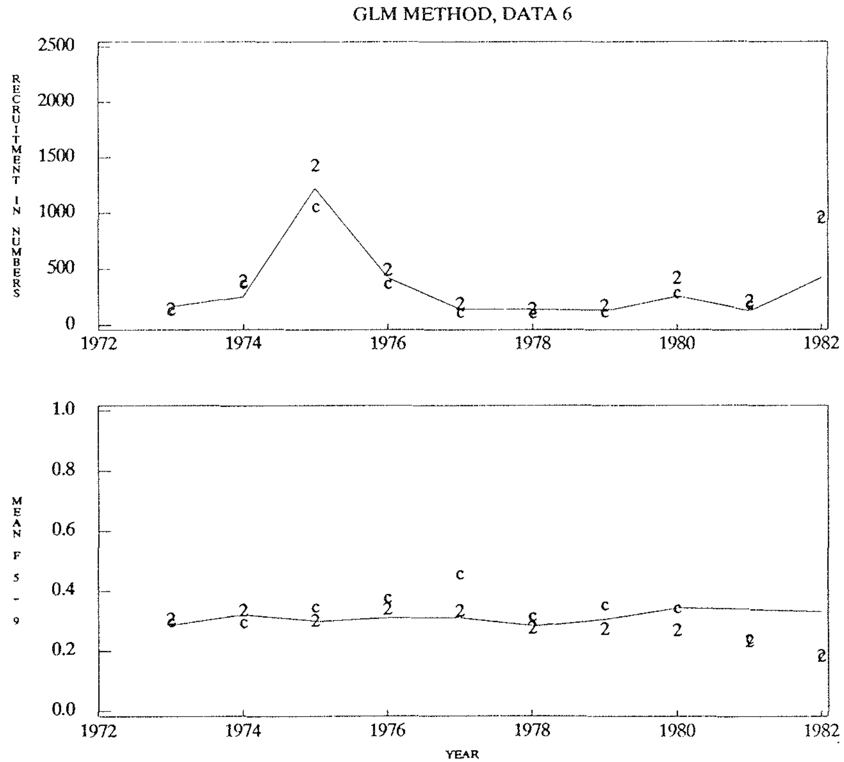
Final extended survivor estimate (XSA) was not computed.

Figure 4.23



-- =TRUE,2=FINAL EST.,c=CURRENT EST.

Figure 4.24



-- =TRUE,2=FINAL EST.,c=CURRENT EST.

Figure 4.25

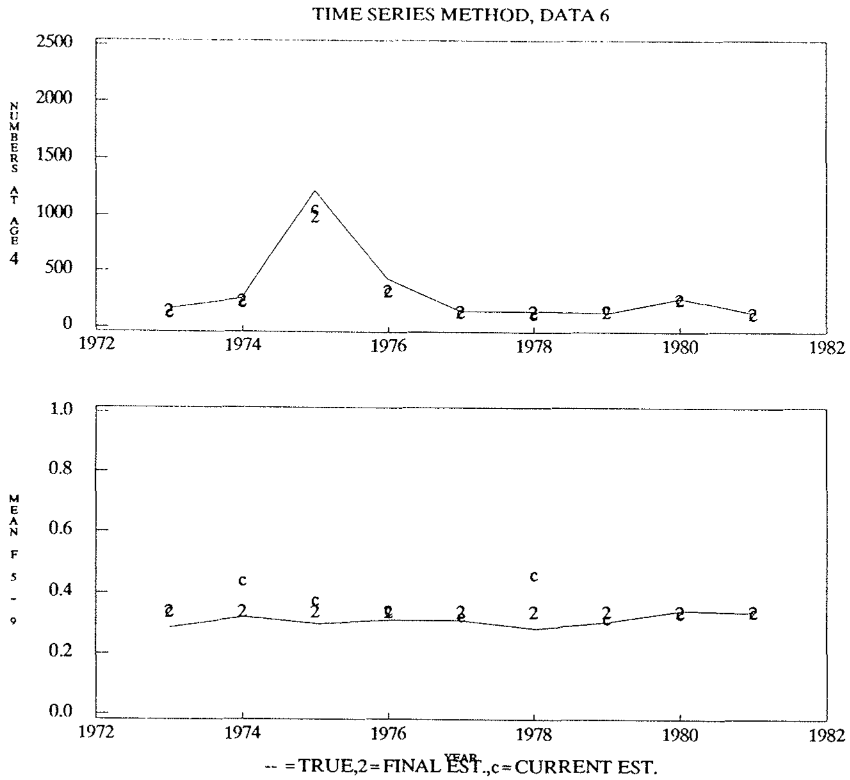
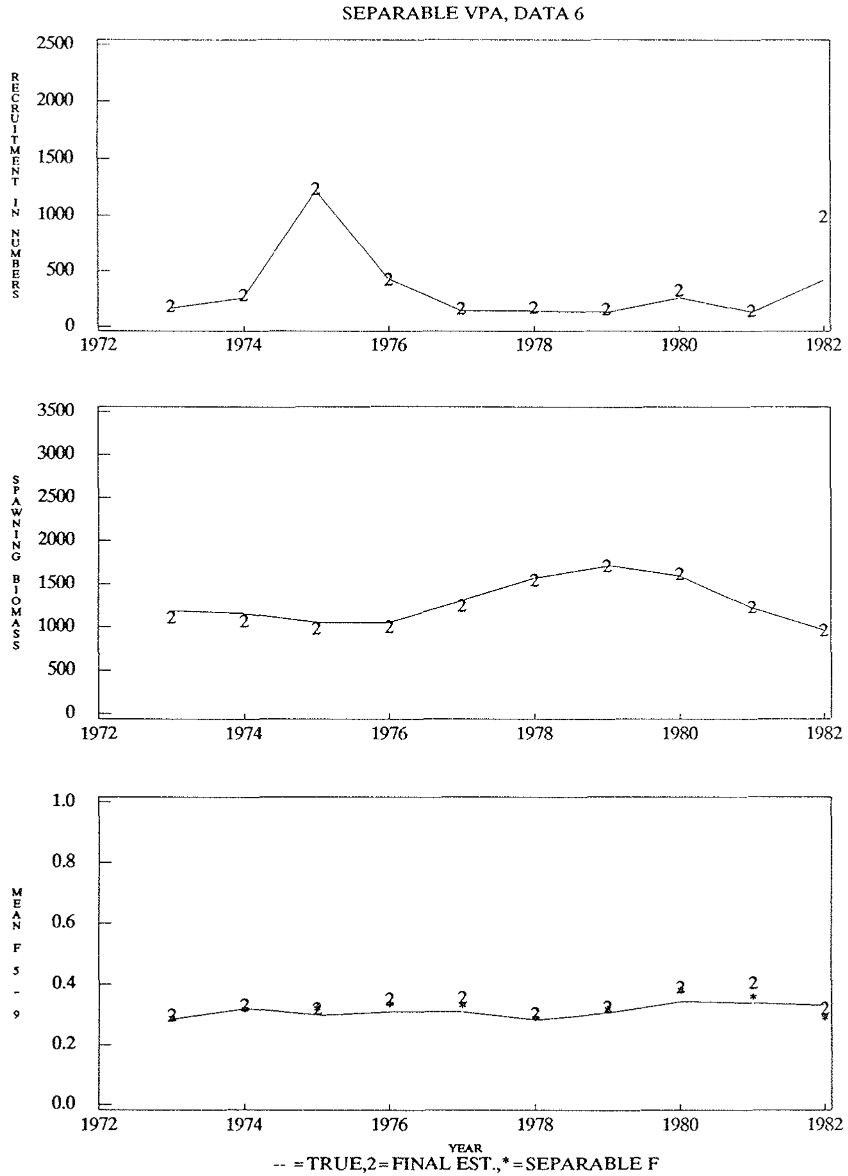


Figure 4.26



ANNEX 1

SIMULATION OF DATA

1 INTRODUCTION

Six data sets were produced either before or during the meeting, and a description of the type of data generated is provided in Section 3.1. Assessment methods were applied to these data to estimate the "true" values of the parameters used to generate the data. Comparison of estimate with truth was used to judge the viability of the methods.

Because of the very large number of tables involved, reproduction of the true values in this report is not possible. Copies of the true parameter values can be obtained on IBM-formatted disk from D.W. Armstrong or G. Stefansson at the addresses cited in Section 3.1.

2 UNDERLYING (NON-STOCHASTIC) MODEL

The underlying model is the conventional fisheries model. If there were no errors involved, the following equations would hold true:

Let a = age (3-12)
y = year (30 years: 1953-1982)
f = fleet (7 fleets: 2 trawlers, 1 liner, 1 fixed net, and 3 research vessels)
C = catch in numbers
F = fishing mortality rate
M = natural mortality rate
N = stock size in numbers

Catches

$$C(a,y,f) = \frac{F(a,y,f)\{1-\exp[Z(a,y)]\}[N(a,y)]}{Z(a,y)}$$

where F(a,y,f) is the mortality induced by fleet f and

$$Z(a,y) = \text{total mortality rate} = \sum_f F(a,y,f) + M(a)$$

Stock

$$N(a+1,y+1) = N(a,y)\exp[-Z(a,y)]$$

Separability

The fishing mortality rate for each fleet is assumed to follow the separable model, so that

$$F(a,y,f) = F(A,y,f)S(a,f)$$

for some overall level of $F(A,y,f)$. For convenience, we take selection to be 1 at the maximum, or equivalently,

$$F(A,y,f) = \max F(a,y,f)$$

(For Data Set 6, we violated the assumption of separability for the commercial fleets. A detailed description of how this was done is provided in Section 3.1.)

Relationships between fishing effort and fishing mortality

The effort data for each fleet are related to fishing mortality in some simple fashion.

To simulate fleets in which catchability changes, we write

$$\ln E(y,f) = c(f) + d(f)y + \ln[F(A,y,f)]$$

To simulate a fleet which exhibits no change in catchability, we set $d(f)y = 0$ and hence

$$\ln E(y,f) = c(f) + \ln[F(A,y,f)]$$

(For Data Set 6, we altered the model relating effort and fishing mortality to the follwing form:

$$E(y,f) = F(A,y,f)[c(f) + d(f)y]$$

This corresponds to a trend in catchability described by the function

$$1/[c(f) + d(f)]$$

the convexity of which is opposite to the exponential funtion assumed in all other data sets.)

We refer to the above as the UNDERLYING model and, in particular, we refer to values of $F(a,y,f)$ as the underlying (nonstochastic) fishing mortalities.

This underlying model is assumed for all fleets including research vessels.

3 STOCHASTIC ADDITIONS

Process error of fishing mortality rates, realized values of F, N, and C

We introduce errors directly into the fishing mortalities.

$$\ln F'(a,y,f) = \ln F(a,y,f) + e(1,a,y,f)$$

This is equivalent to saying that a fleet has "decided" to induce a given level of fishing effort, but the target value has not been achieved due to random variations in weather and other factors.

For convenience, we have taken the errors $e(1,a,y,f)$ from a normal distribution (with different variances for different data sets and fleets). These errors are termed the PROCESS ERROR with variance $v(1,a,f)$.

The values $F'(a,y,f)$ are those which the fleet actually induces and are termed the REALIZED fishing mortalities.

The realized total mortality rate is, therefore

$$Z'(a,y) = \sum_f F'(f,a,y) + M(a)$$

The corresponding realized stock sizes are given by

$$N'(a+1,y+1) = N'(a,y)\exp[-Z'(a,y)]$$

The associated realized catches are given by

$$C'(a,y,f) = \frac{F'(a,y,f)\{1-\exp[-Z'(a,y)]\}N'(a,y)}{Z'(a,y)}$$

Note that an assessment method attempts to estimate the realized values (or some subset of them). It is the realized values that are, therefore, referred to as "truth" in the main body of this report.

Measurement error of catch at age, estimated catches

The realized catches $C'(a,y,f)$ are the quantities which are actually landed. These catches are sampled to produce ESTIMATED catches which incorporate MEASUREMENT ERRORS.

$$\ln\bar{C}(f,a,y) = \ln C'(f,a,y) + e(2,a,y,f)$$

The measurement error $e(2,a,y,f)$ is assumed to follow a normal distribution with variance $v(2,a,f)$ for Data Sets 1-4. For Data Sets 5 and 6, a gamma distribution parameterized to have a mean of 1 and a coefficient of variation between 0 and 1 was used to generate measurement errors in catch at age and process errors in the fishing mortalities.

Measurement error of effort data, estimated effort

It is unlikely, in reality, that effort data are exact. Errors will be incorporated as effort data are collected. To simulate this, a stochastic element is added to the relationship between effort and overall fishing mortality to produce the ESTIMATED effort data.

$$\ln\bar{E}(y,f) = c(f) + d(f)y + \ln F(A,y,f) + e(3,y,f)$$

For all data sets, the effort errors $e(3,y,f)$ are drawn from a normal distribution with variance $v(3,f)$ and are different for each fleet. This procedure was applied to all of the data sets.

4 TECHNICAL NOTES

(i) Random number generation was carried out using the Tausworthe shift-register generator.

(ii) Normal errors were generated using the Box-Mueller transform.

(iii) Gamma-distributed errors were generated by encoding an algorithm due to Knuth.

(iv) The program for simulating data sets can optionally generate log-normal or gamma-distributed errors and can include linear or exponential trends in effort.

5 GENERAL NOTES

(i) Changes in catchability are modelled by introducing a bias in the fishing effort data.

(ii) The estimated effort data are generated from the underlying fishing mortalities not from the realized fishing mortalities.

(iii) After analysis of Data Sets 1-4, it was found that the variances v(1,a,f) and v(2,a,f) included in the simulations were far too small for the research vessels. Caution is, therefore, required in interpreting the results from these data sets since many of the methods will perform better than they would on more realistic data.

(iv) For simulated Data Sets 1-4, the variance v(3,f) for research vessels was set at zero. Some higher value should have been used to allow simulation of the fact that research vessel catchabilities vary considerably from year to year.

(v) The model described above is used for all fleets including the research vessels. Differences between fleets are created by the choice of underlying fishing mortality rate, variances associated with the error terms, and the choice of changes in catchabilities reflected in c(f) and d(f)y. Stock numbers at the youngest age and for each age in the first simulated year were based on data for Icelandic cod and were not generated by a simulation process.

6 OVERVIEW OF THE CHARACTERISTICS OF DATA SETS 1-6

Process error and measurement error - general comments

An analysis of variance of log-catch data for North Sea and Icelandic cod indicated that the effects of process errors and measurement errors are almost additive into log-catch. However, no information is available on the degree to which the variance in log-catch is divisible between the two types of error. For this reason, the relative dimension of process and measurement error in each data set is arbitrary.

Process and measurement errors were given the highest values for the youngest and the oldest age groups.

Date Set 1

No bias in effort data [i.e., no trends in catchability, d(f)=0]. The level of the underlying fishing mortality rates for all fleets combined is about 0.4. Process error equal to measurement error. No effort error for research vessels.

Data Set 2

No bias in effort data for any fleet. Overall level of underlying fishing mortality is about 1.0. Process error = 0.5 x measurement error. No effort error for research vessels.

Data Set 3

Bias in effort data for two of the commercial fleets. Overall level of underlying fishing mortality about 0.4, but with a steadily increasing trend.

Data Set 4

Bias in effort data for all fleets. Overall level of underlying F about 0.8 in early years to about 1.2 in the last data year. No measurement error, only process error.

Data Set 5

Same underlying structure as Data Set 3, but process error on fishing mortalities and measurement error on catch at age derived from gamma distribution rather than log-normal distribution. Log-normal distribution retained for effort errors. Higher levels of noise used than in Data Sets 1-4. Catch measurement error coefficients of variation range from 14-70%, with higher values on the youngest and oldest age groups and on the research vessels' data. Process error coefficient of variation of 20% on all ages and fleets. Strong year class recruited in year 24 (1977) of abundance (1.2 billion) about an order of magnitude greater than the weakest year class.

Data Set 6

Based on Data Set 5, but some aspects of the underlying model altered. Changes to funtional form for trends of catchability with time explained in Section 2.

In addition, separability in commercial fleets no longer valid. For one of the commercial fleets, catchability increases on the two youngest age groups between years 14 and 20. Beyond year 21, catchability increases further on the young age groups and decreases on ages 9-12. This procedure simulates a progressive shift by this fleet towards fishing of younger fish.

For another of the commercial fleets, a shift towards fishing on older fish from year 18 onwards was simulated. This was achieved by increasing realized fishing mortality on ages 7-12 by the quantity 1+[0.2(age 6)].

Finally, it was assumed that all commercial fleets increased their catchability on a very large 1972 year class. The realized fishing mortalities at the appropriate years and ages were multiplied by 1.2 to simulate this effect.

ANNEX 2

DESCRIPTION OF METHODS

1 AD HOC TUNING OF VPA

The basic *ad hoc* tuning algorithm is outlined in the pseudocode below.

```
Guess F in last data year
Do VPA
Calculate catchability for each age and fleet
   For each age
      For each fleet
         Fit model to catchabilities
         Estimate terminal catchability and associated variance
         Calculate terminal F
      Next fleet
      Combine estimates of terminal F as weighted average value
   Next age
Iterate
```

The methods iterate to find a solution consistent with historical parameter estimates and do not seek to minimize any statistical objective function. For this reason, these methods are not regarded as being based on a formal statistical model.

The methods estimate catchablities for each age group and fleet separately. Some plausible model is then fitted to these estimates to allow estimation of catchability in the last data year. This value is then used in conjunction with the appropriate CPUE value to estimate population size. The population size is then used in conjunction with total catch-at-age data to estimate fishing mortality. The CPUE data and the total catch-at-age data are treated as exact. Errors in CPUE, therefore, affect both the population and F estimates while errors in the total catch-at-age data affect only the F estimates.

Ad hoc methods are simple to implement, computationally fast (run times of 1-2 minutes are typical) and rarely crash or give infeasible results.

Some of the *ad hoc* methods analyze the logarithm of catchability. In these cases, it makes no difference whether one analyzes the relationship between CPUE and abundance or that between fishing mortality and fishing effort (Laurec and Shepherd, 1983). Use of a logarithmic tranformation is also consistent with the non-negative, but highly skewed distributions of catch-at-age and CPUE-at-age data.

There is a family of *ad hoc* methods generated by choice among the following options:

(a) Use log-transform or not.

(b) Assume constant catchability or fit a regression (usually against time, but could also be against stock abundance, etc.).

(c) Combine estimates of terminal F using inverse variance weighting (usual procedure in recent years) or some other rule (becoming less popular).

(d) In addition, further variants may be generated by use of various procedures for down-weighting data for distant years and for shrinking estimates of terminal F (or N) towards some historical prior value.

The following eight methods were tested at this meeting:

(i) Laurec-Shepherd (Laurec and Shepherd, 1983; Pope and Shepherd, 1985). This uses a logarithmic transformation, applies a 20-year tricubic taper to down-weight historical data, assumes no linear trends in the log-catchabilities (locally constant catchability) and F on the oldest age group was iteratively reset to the average over the five next youngest ages.

(ii) Hybrid (Pope and Shepherd, 1985). This is identical to the Laurec-Shepherd method except that a linear time trend is fitted to the (down-weighted) log-catchabilities.

(iii) Armstrong-Cook methods. These are basically a mixture of the Laurec-Shepherd and Hybrid methods. Catchability is regressed against time for commercial fleets, but is assumed constant for research vessels. A 20-year tricubic taper with maximum weight applied 3 years before the last data year is used to down-weight. Estimates of terminal F are combined by inverse variance weighting. An additional option of shrinking estimates of terminal F towards the historic mean from VPA is also available.

Four variants of this method were tested:

AC1: Log-transformed catchabilities, shrink towards historical F

AC2: Log-transformed catchabilities, no shrinkage towards historical F

AC3: Untransformed catchabilities, shrink towards historical F

AC4: Untransformed catchabilities, no shrinkage towards historical F

(iv) Lewy's (1988) methods. These methods estimate stock numbers in the last data year by regressing numbers on corrected CPUE (CCPUE). No transformation of the data is used and catchability is assumed constant for the last 10 data years. Fishing mortality on the oldest age group is set equal to the average for the three next youngest age groups.

The CCPUE method combines predicted N values using inverse variance of the predicted Ns.

The AEFM method uses a different weighting procedure. Fitted values of fishing mortality and stock numbers are obtained for the last 10 years. These are used, via the conventional catch equation, to produce corresponding estimates of "fitted" catch. The inverse variance of the fitted and observed catches is used to weight the last data year estimates of N.

All the ad hoc tuning methods were run with no major problems on all six simulated data sets and the two real data sets for both the multiple realizations and the 30-year analysis. All of the methods recovered the main features of the data sets, especially in the case of Data Sets 1-4. The only computational difficulty encountered was that the AEFM method does not converge if the period when catchability is assumed constant includes the last data year. This method converges rapidly if the last two data years are excluded from the above-mentioned period.

The software developed to run the Armstrong-Cook methods was intended to run automatically without user intervention. If these methods are to be further developed, more attention needs to be given to diagnostic output. In the case of Data Set 6, examination of the slopes of the regressions through commercial catchablity estimates indicated that many of them did not appear significant.

The more highly-developed diagnostic features of the Laurec-Shepherd and Hybrid methods were particularly useful in analyzing Data Set 6. Large standard errors and significant conflicting trends in catchability were indicated and the Hybrid method indicated highly significant trends in catchability at all ages for all commercial fleets except one of the trawlers. A mixed analysis was, therefore, carried out by specifying catchability on this fleet. This indicated strong and highly consistent commercial catchability trends for almost all ages, especially for Fleet 3 and relatively weak but sometimes significant trends for the survey fleets. It was considered likely that it was the commercial rather than the survey fleets which exhibited real trends. A second mixed analysis was then run with fixed q for Surveys 1 and 2 (since the diagnostics for Survey 3 had indicated rather variable trends). This analysis revealed a weak but statistically significant negative trend for Survey 3, no significant trend for Commercial Fleet 1 and strongly significant positive trends for Commercial Fleet 3. This analysis was accepted even though it is probable that a mixed analysis with fixed q on all fleets except Commercial Fleet 3 would be preferable. (This level of confusion and inconsistency of results is considered by the assessor to be fairly typical of real life!)

2 SURVIVORS AND EXTENDED SURVIVORS

Survivor analysis

Survivor analysis combines catch-at-age information and a research vessel abundance index at age to produce estimates of stock size for each age at the end of the current year (i.e., survivors). The method is described by Doubleday (1981), and a computer implementation is provided by Rivard (1982).

Underlying assumptions specify that

(a) catch is taken uniformly throughout the year,

(b) the research vessel abundance index is a mid-year estimate of numerical stock abundance,

(c) the natural mortality rate is a "known" constant applicable to all years and age groups represented in the catch-at-age data.

The research vessel abundance index is calibrated against VPA population numbers by defining calibration constants [say k(i)] within a pre-defined calibration block which correspond to the ages and years for which the VPA has converged. Within that block, the survey index at age [say A(i+0.5,t+0.5), where i+0.5, t+0.5 is used to identify the mid-year] is related to mid-year population abundance [say N(i+0.5,t+0.5)) as follows:

$$N(i+0.5,t+0.5) = k(i)A(i+0.5,t+0.5)e \qquad (1)$$

The calibration constant can thus be estimated as

$$\ln[k(i)] = \frac{\sum_{t=t0}^{t1} \ln N(i+0.5,t+0.5) - \ln A(i+0.5,t+0.5)}{t1 - t0 + 1} \qquad (2)$$

where t0 and t1 are the first year and the last year in the calibration block, respectively.

The mid-year population abundance is obtained from a generalized method of sequential population analysis in which the survivors appear explicitly as input parameters. This formulation allows estimation of the variance of the survivors, which is input to the catch projections, i.e.

$$N(i+0.5,t+0.5) = fS[i,t(f)] \qquad (3)$$

Consequently, from an initial estimate of survivors for the last year and for the oldest age-groups, we can estimate

$N(i+0.5,t+0.5)^1$ from equation (3)

and the calibration constants $k(i)^1$ are calculated from equation (2), where the superscript 1 identifies the first step of the iteration process. Then j independent estimates of the survivors in the final year, for age groups i, can be obtained from each survey index which provides an independent measure of stock size along a cohort, i.e.

$$S[i,t(f),j]^1 = \{k(j)^1 A[j+0.5,t(f)-i+j+0.5] - f[C(i,t)]\}\exp(-M(i-j+0.5) \qquad (4)$$

The j independent estimates of the survivors along a cohort are then averaged as follows:

$$S[i,t(f)]^{1} = \sum_{j} w[i,t(f),j]^{1} S[i,t(f),j]^{1} \qquad (5)$$

where $w[i,t(f),j]$ is a function of the variance of the estimated survivors.

$S[i,t(f)]^{1}$ becomes a new starting value for (3) and the calculations represented by (2), (4), and (5) are repeated in an iterative manner until the relative difference between the successive estimates of survivors is small (say <0.001).

This iterative process provides estimates of the survivors for the oldest age group in each cohort in the catch matrix together with corresponding variance estimates.

In practice, the method works well when the calibration block is extended to all years available. For the analysis of the simulated data sets, the calibration block was defined to include all years except the last data year and ages 3-9. Separate calibration constants were obtained for ages 3, 4, and 5, and a common calibration constant was estimated for ages 6-9. No attempt was made to evaluate the effect of the number of calibration constants on the results.

The Survivors Analysis was initially designed to accommodate the situation where no auxiliary information is available except that from a single survey estimate of abundance. The application of the methops to the simulated data (which provided the results of three independent surveys) required some pre- or post-processing.

(i) The commercial catch rate data were not utilized.

(ii) For Data Sets 1-3, where the survey data exhibited similar trends, the three survey indices were standardized and averaged to produce a single data set.

(iii) For Data Set 4, divergent trends were observed in the research vessel data. The analysis was applied using each data set and the results were averaged _a posteriori_.

(iv) For Data Sets 5 and 6, _a posteriori_ averaging of results derived by using each survey series separately was also used. Diagnostics revealed that the assumption of log-normality of errors was incorrect for these data sets (large number of outliers in residuals and large proportion of residuals of the same sign in results obtained using Surveys 2 and 3, estimates of fishing mortality less variable than expected). For Data Set 5, the coefficients of variation (CVs) for survivor estimates for ages 4-7 were calculated.

Survey number	CV (%)
1	30-40
2	55-75
3	90-150

These estimates are inflated since they assume (actually non-existent) log-normality.

Survivor Analysis was also applied to the full 30-year data series for Data Sets 4 and 6.

For Data Set 4, comparison of stock abundance estimates and survey indices indicated an increasing catchability trend in Surveys 1 and 3 and a decreasing trend in Survey 2, and the survey indices were not tracking the trends in stock size. Also, in order to assess the effect of the changes that took place for the second research vessel in the 27th year, an analysis of catchability at age was made for that vessel. This led to the estimation of a conversion factor of 1.2 for the last four years of that series. Finally, a retrospective analysis (Rivard and Foy, 1987) was applied to the last 10 years of the time series. That analysis indicated that combining the three survey estimates led to a systematic overestimation of stock size. In view of these observations, Survivor Analysis was applied using Survey 2, multiplied by 1.2, for the last 4 years to account for the change in vessel efficiency, and Survey 1. Combining Surveys 1 and 2 had the same effect as removing the trend in catchability for each series. The retrospective analysis was applied again and indicated that a systematic overestimation of stock size was still present, but was reduced compared to the previous analysis.

For the analysis of the 30-year series of Data Set 6, the three survey series and stock abundance estimates were normalized and plotted against time. No obvious trends in catchability in any of the surveys were apparent. A retrospective analysis applied to the last 10 years of data indicated that combining the three surveys led to a systematic underestimation of stock biomass for older fish of 15-20%. Also, the coefficients of variation of Survivors for ages 6 and older estimated using Survey 3 were extremely high (120-180%). The logical step following from these observations would have been to re-analyze the data with Surveys 1 and 2 only and to apply diagnostic tools again to the new results. Lack of time prevented this, and the results referred to in Section 4.2 correspond to the application of the Survivor Analysis for the last 20 years by combining all three surveys. Thus, these results contain a bias of 15-20% which could have been eliminated by further analysis.

For the real data sets (North Sea cod and haddock), only one survey provided estimates for a sufficient range of ages and years under present implementation. The other sets could not be utilized.

Extended Survivor Analysis

Work in progress by Sun (pers. comm.) suggests that a major source of error in assessment calculations is sensitivity to errors in the data for the final year. Many of the assessment methods treat these data as being exact, but this is not necessary except in the VPA calculations of VPA-based techniques. The Survivors method of Doubleday (1981) allows estimates of terminal populations based on all data for each cohort to be used, which should reduce the sensitivity to final-year errors. The original method, however, allows auxiliary data for only one fleet to be analyzed and uses an estimation procedure for Survivors which is inconsistent with that used for catchability. In addition, the algorithm frequently produces negative estimates of survivors which are censored and replaced by zeroes.

Shepherd and Sun (pers. comm.) have recently developed an extended version of the same general procedure. This allows use of auxiliary data from multiple fleets and employs an exponential decline algorithm (rather than the original subtractive algorithm) which is consistent with the use of logarithmic mean catchability and avoids negative estimates.

A preliminary implementation of this method was available, although this did not include certain desirable features such as inverse variance weighting. By mistake, the method was run on Data Sets 1-3 with no constraint on catchability at the oldest ages, which leaves the solution ill-determined. For data sets 4-6, catchability was assumed to be constant on ages 10-12.

3 CAGEAN - CATCH-AT-AGE ANALYSIS

A well-documented description of CAGEAN can be found in Deriso et al. (1988) and references therein.

Some problems were identified in the approach taken to the estimation of last-data-year parameters for the period 1973-1972 (as specified in Section 3.2.1) for Data Sets 1-3. This work was carried out prior to the meeting. Essentially, the assessors conditioned the analysis of each 20-year data set by prior knowledge obtained from detailed analysis of the corresponding 30-year data sets. The final results from analysis of any 20-year data set was accepted only if estimated biomass agreed fairly closely with that obtained by analyzing the full 30 years data.

The original intention had been to perform an independent assessment on each 20-year data series. Because of lack of time, the assessors could not recompute the results for Data Sets 1-3, but Data Sets 4-6 were analyzed. The analysis was, in many ways, less rigorous than that which would be carried out given more time. It was only possible to analyze 10-year data sets. Some up-to-date software was not available at the meeting, and not enough time could be spent examining diagnostics and hence appropriately modifying the analyses. The comments in Section 4.1.1 on the apparent performance of CAGEAN should be read with these qualifications in mind.

Overall, it appears that the relative weighting given to each type of data and also the values used to initiate the computations need to be handled with considerable care. Different weightings can lead to substantially different results and careful consideration of diagnostics is required to obtain an acceptable assessment.

4 ADAPTIVE FRAMEWORK

Model

The basic framework is simply a mathematical expression for the application of a common statistical technique, least squares, to examine the discrepancy between observations of variables and the values of those variables estimated as functions of a population matrix, in order to determine the most appropriate estimate of that population matrix. That is, we require to find

$$\min \sum_{i} \{W(i)[O(i) - f(P,G)]\}^2 \qquad (6)$$

where W(i) = weight for observed variable set i
O(i) = observed variable set i
P = population matrix
G = matrix of any other required parameters

Note that O(i) and W(i) may be matrices of vectors (series). The W(i) are needed to accommodate differences in the reliability of the elements within an observed variable set as well as any differences in reliability between variable sets. Lacking such measures, transformations may be employed in attempting to stabilize variance. The summation is taken over all sets (i) as well as within each set.

The framework is adaptive in the sense that any observed variable which is a function of the population matrix can be accommodated by equation (6). Furthermore, various formulations of the structural relationships and statistical error models which link these observed variables with the population matrix may be invoked. This flexibility is considered essential given the wide range of situations encountered in stock assessment. Common statistical diagnostics, e.g., residual plots, standard errors, and correlation matrices of the parameters estimated, are used to select from among the formulations those which are most suitable for the particular conditions experienced. To elucidate the basic framework and to demonstrate the flexibility in the types of relationships which may be employed, two hypothetical scenarios are described.

Scenario A

The commercial catch has been sampled using a double sampling design and the estimated catch at age C(a,y) is available with the associated standard error CS(a,y). It is known that age determination for older ages is variable; therefore, ages 1-5 are treated individually, while ages 6 and older are aggregated. There are no reliable data on effort from the commercial fishery. A research vessel survey index of abundance at age, I(a,y), is available. The survey was conducted at the beginning of the year using a stratified random design, and the appropriate standard error for the index, IS(a,y), has been derived. There are no other relevant observed variables.

The expression to be minimized is:

$$\sum_{a=1}^{6+}\sum_{y=1}^{20}\left[\frac{1}{CS(a,y)}[C(a,y) - \hat{C}(a,y)]\right]^2 + \sum_{a=1}^{6+}\sum_{y=1}^{20}\left[\frac{1}{IS(a,y)}[\hat{I}(a,y) - I(a,y)]\right]^2 \qquad (7)$$

a = index for age
b = index for year (20 years of data)

Note that results from the beginning-of-the-year survey are available at the time the assessment is done.

In order to ensure that population size decreases along cohorts with time, the parameter set P is replaced by R, an estimate of the year-class size for each cohort, and F, the fishing mortality matrix.

The associated population matrix can then be calculated using the relationship:

$$Q(a,y) = Q(a+1,y+1)\exp[F(a+1,y+1) + M] \qquad (8)$$

where natural mortality rate, M, is assumed constant for all ages and years. The appropriate cohort year-class size, R, is substituted into Q as required.

The predicted catch can then be obtained using the conventional catch equation:

$$\hat{C}(a,y) = F(a,y)Q(a,y)\{1-\exp[-F(a,y) - M]\}/[F(a,y) + M] \qquad (9)$$

A linear relationship through the origin can be assumed between the abundance index and population size. Therefore, the predicted index is obtained from:

$$\hat{I}(a,y) = k(a)P(a,y) \qquad (10)$$

where k(a) = calibration coefficient for age a. The parameter set G consists of only k(a) in this scenario. Equations 7-10 can be used to solve for the least squares estimates of R, F, and k.

Scenario B

The commercial catch has been sampled, as in Scenario A above; however, the errors in the estimates of catch at age are considered negligible. A combined catch rate series, U(y), has been derived with a multiplicative model, and its associated standard error is US(y). There are two research survey abundance indices, I(1) and I(2), and their standard errors, IS(1) and IS(2), were computed on the basis of the respective survey designs. Survey I(2) is considered a recruitment index suitable for the first two ages only and is only available for the most recent 6 years. Both surveys are related to the beginning of year population.

The expression to be minimized is:

$$\sum_{a=1}^{10}\sum_{y=1}^{21}\left[\frac{1}{IS(1,a,y)}[I(1,a,y) - \hat{I}(1,a,y)]\right]^2 +$$

$$\sum_{a=1}^{2}\sum_{y=16}^{21}\left[\frac{1}{IS(2,a,y)}[I(2,a,y) - \hat{I}(2,a,y)]\right]^2 +$$

$$\sum_{y=1}^{21}\left[\frac{1}{US(y)}[U(y) - \hat{U}(y)]\right]^2 \tag{11}$$

Since errors in the catch at age are considered negligible, the parameter set P is reduced to R, the year-class size of each cohort. The last year and the oldest age are used as the designate age for the year-class size. The population matrix can then be derived using:

$$Q(a,y) = C(a,y)\exp(M/2) + Q(a+1,y+1)\exp(M) \tag{12}$$

where the appropriate cohort year-class size is substituted into Q as required.

Linear relationships are assumed for both survey indices. However, intercepts are accepted for survey index I(2) even though the mechanism to generate such a relationship has not been established. Therefore:

$$\hat{I}(1,a,y) = k(1,a)q(a,y) \tag{13}$$

and

$$\hat{I}(2,a,y) = k'(2,a) + k(2,a)Q(a,y) \tag{14}$$

A fishing mortality matrix is calculated from:

$$F(a,y) = \ln[Q(a,y)/Q(a+1,y+1)] - M \tag{15}$$

The partial fishing mortality rate matrix for the otter trawl fleet was obtained as:

$$F(T,a,y) = F(a,y)C(T,a,y)/C(a,y) \tag{16}$$

The annual fully-recruited fishing mortality for all trawlers was derived from:

$$F'(T,y) = \sum_{a=5}^{10} Q(a,y)F(T,a,y)/\sum_{a=5}^{10} Q(a,y) \tag{17}$$

The annual partial recruitment for the trawler fleet is then obtained:

$$PR(T,y) = F(T,a,y)/F'(T,y) \tag{18}$$

and used to calculate the average annual exploitable biomass:

$$\bar{B}'(T,y) = \bar{W}(a,y)(Q(a,y)\{1-\exp[-F(a,y) - M]\}/[F(a,y) + M]PR(T,y) \tag{19}$$

A linear relationship through the origin is hypothesized for the otter trawl catch rate and the exploitable biomass:

$$\hat{U}(y) = k(3)\bar{B}'(Y,y) \tag{20}$$

We now have the quantities required for minimization of expression (11).

Application of simulated data

Data Set 1

Errors in the catch-at-age data were assumed negligible. The three survey indices were used for individual ages 3, 4, and 5 and aggregated for ages 6 and older. The two commercial fleets for which effort data were available were employed by deriving a total catch rate in numbers for each fleet, i.e.

$$U(T,y) = \sum_{a=1}^{10} C(T,a,y)/E(T,y) \tag{21}$$

No standard errors were provided and, therefore, logarithmic transformation of the survey indices and commercial CPUE was applied.

The expression minimized was:

$$\sum_{i=1}^{3} \sum_{a=3}^{6+} \sum_{y=1}^{20} [\ln I(i,a,y) - \ln I(i,a,y)]^2 + \sum_{T=1}^{2} \sum_{y=1}^{20} [\ln U(T,y) - \ln U(T,y)]^2 \tag{22}$$

The population matrix was calculated using equation (12). However, because older ages appeared fully recruited, the population size for the oldest age was not included in the parameters R. Instead, the population was derived using catch equation (9), and a fully-recruited fishing mortality calculated as the weighted average for ages 6-9 inclusive.

With the population matrix available, relationships of the form of equation (13) were used to obtain predicted survey indices. The predicted catch rate indices were computed as for Scenario B [omitting the weights in equation (19) since the catch rates are in numbers].

A total of 23 parameters require to be estimated (9 year-class strengths at the end of year 20, catchability coefficients for ages 3, 4, 5, and 6+ for each of the three survey series, and catchability coefficients for each of the two commercial catch rate series.

The number of residuals calculated was 80 for each survey (4 age groups, 20 years), and 20 for each catch rate (20 years), giving a total of 280 residuals.

Convergence was rapid in all runs and no obvious problems were detected from analysis of residuals. Coefficients of variation for population size in the final year were of the order of 5-10%.

Data Set 2

The same formulation was used as for Data Set 1 except that the fully-recruited fishing mortality was calculated as the weighted averages of ages 7-10. Parameter estimation was difficult in the last few blocks of 20 years and in fact no suitable convergence criteria were obtained.

Coefficients of variation for the population size in the final year were 20-40% for the younger ages and higher for the older ages. The residuals revealed disturbing patterns suggesting that at least one of the indices did not conform to the model equations. Furthermore, the assumption of flat-topped exploitation pattern was questionable, especially in the later years. In conclusion, refinement of the model equations was indicated if the analysis of this data set was to be extended.

Data Set 3

The same model formulation as that for Data Set 1 was used with fully-recruited fishing mortality calculated as the average for ages 6-9. Convergence was not as rapid as for Data Set 1 (usually 7 iterations being required as compared to 3 for Data Set 1), but no basic problems in convergence were encountered.

Correlation between parameters was low, in the range 0.01-0.1. Coefficients of variation for population size in the last data year were 9-15% for ages 4-9 and higher for older ages. Residuals were not examined for trends.

Data Set 4

Only the analysis of the full 30-year data set was carried out. Initially, the same model formulation as that used for Data Set 2 was employed. Analysis of residuals revealed very strong patterns with time. Surveys 1 and 3 exhibited increasing catchability, while catchability in Survey 2 decreased. The model was modified to include a linear trend for catchability.

The coefficients of variation for the final year population estimates were lower (about 10% for younger ages and 20% for others) under the revised model. The slopes for the linear trends were highly correlated with the associated intercepts, but their coefficients of variation were only about 20%. There still remained, however, a significant trend in the residuals for the linear catch rates indicating that increases in catchability in some commercial gears may not have been adequately accounted for.

5 GENERAL LINEAR MODEL

This is a new method still in a state of development and testing. The method fits the General Linear Model (GLM):

$$\ln C(a,y,f) = A(a,f) + Y(y,f) + I(a,y) + e(a,y,f) \tag{23}$$

$$\ln E(y,f) = Y(y,f) + n(y,f) \tag{24}$$

where a, y, f are age, year, and fleet indices, respectively, A, Y, I are age X fleet, year X fleet, and age X year effects, and e(a,y,f), n(a,y,f) are error terms.

In the current implementation, the fit is done in the GLIM package of Baker and Nelder (1978) which allows the error structure to be any member of the exponential family of distributions (Normal, Poisson, Gamma, or Binomial). At present, the model is fitted assuming log-normality, but this could be easily modified. The parameter estimates obtained by the GLM described in equations (23) and (24) are adapted so that the fit to the data is unaffected, but the terms are reinterpreted in relation to the conventional fisheries catch and stock equations.

$$\ln C(a,y,f) = \ln q(a,f) + \ln E'(y,f) + \ln \bar{N}(a,y) \tag{25}$$

$$\ln E(y,f) = \ln E'(y,f) \tag{26}$$

where q, E', and $\bar{N}$ are catchability, effort, and average population terms, respectively.

This is done using factors d(a) and p(y) such that:

$$\ln q(a,f) = A(a,y) + d(a) \tag{27}$$

$$\ln E'(y,f) = Y(y,f) + p(y) \tag{28}$$

$$\ln \bar{N}(a,y) = I(a,y) - d(a) - p(y) \tag{29}$$

The values of d(a) and p(y) are chosen such that a GLM of ln N(a,y)

$$\ln N(a,y) = \text{Year-class Effect}(y-a) + \text{Age Effect}(a) + \text{Year Effect}(y) - k \times \text{CUMZ}(a,y) + \text{error}(a,y) \tag{30}$$

has k = 1 and the age and year effects equal to zero. The age effects are fitted to ages up to 3 less than the oldest age in order to preserve constant values of q(a,f) on the last four ages of the last fleet. This fleet should, therefore, be chosen as one using a gear likely to have an exploitation pattern which is flat over these ages.

$$\text{CUMZ}(a,y) = \sum_{i<a} Z(i,y-a+i) + \ln \frac{1-\exp[-Z(a,y)]}{Z(a,y)} \tag{31}$$

where $$Z(a,y) = M(a) + \sum_{\text{all } f} q(a,f)E'(y,f)$$

Suitable values of d(a) and p(y) are estimated by progressive iterations based on the GLIM fit. At each step j, we have:

$$d'(a,j) = 0.6k \times \text{Age effect}(a) \qquad (32)$$

$$p'(y,j) = 0.6 \times \text{Year effect}(y) \qquad (33)$$

where d(a,j) and p(y,j) sum to d(a) and p(y), respectively.

Preliminary runs of the model have been made giving uniform weighting to each component of Data Sets 1-4. The method would, therefore, probably give better results using appropriate weightings based upon the prior information provided and that gleaned from a study of the residuals.

Implementation time is about 20 minutes on an HP 9000-318 with 10 ages, 10 years, and 7 fleets included in the data sets.

The diagnostics which can be applied to the model results potentially comprise anything that can be done within the GLIM package and are, therefore, open-ended. The method routinely outputs tables, plots, and histograms of residuals with estimates of residual variation by fleets, ages, and years.

Most attention was given to the diagnostics when analyzing Data Set 6, which was one of the most difficult prepared for this meeting. Considerable departures from the assumed within-fleet separability were indicated, raising questions about the applicability of this manifestation of the method for analyzing this data set.

Work carried out for this meeting indicated that the present implementation could be improved in three important ways:

(i) Make into a tidy package.

(ii) Make fleet and age weighting automatic.

(iii) The means by which selection is fixed on the older ages could be arranged better.

Until these points are put into effect, it would be inappropriate to use this method to carry out a real assessment.

6 COLLIE-SISSENWINE METHOD

Collie and Sissenwine (1983) developed a modified DeLury method (DeLury, 1947; Allen, 1966) for estimating fish population size using a single relative abundance index and total catch data from the fishery. The method estimates a catchability coefficient for the index of abundance using non-linear regression techniques. In addition, it accounts for measurement error in the index by estimating an index of abundance for each year and age. Two models were proposed. One requires data on the age structure of the catch, while the other is a non-age-structured model. The age-structured model is of interest here.

Collie and Sissenwine fit the age-structured model to data from haddock populations (Georges Bank and NAFO Division 4X). The estimated population size at age agreed closely with results from VPA analyses. Despite these results, the method has not been widely used in practice.

One major reason for the lack of application is the assumptions and restrictions imposed by the model. In particular, the model assumes that the catchability (q) is constant over time and age, that natural mortality is constant for all ages, and allows only one index of abundance. Only minor modifications are required to account for age-specific natural mortality, but incorporating age- and/or year-specific qs and multiple indices of abundance requires fundamental changes to the model's structure.

In each of the real and simulated data sets considered at this meeting, multiple indices of abundance were available, catchability was thought to vary with age and/or time, and for the real data sets, natural mortality rates are age-specific. To examine the utility of the Collie-Sissenwine model structure, the method was extended to incorporate all of the above-mentioned aspects.

In extending the model, the Collie-Sissenwine concept of separating the process (or equation) error from the measurement error was maintained. The Collie-Sissenwine process error was generalized to incorporate age-specific qs and age-specific natural mortality. The measurement error term, a measure of the variability within an index of abundance, is essentially the same as that of the Collie-Sissenwine model except that log-normally distributed error was not assumed. This change allowed all terms in the objective function to be in the same units. A new consistency error term was developed which provides a measure of the variability between indices of abundance. Retention of the basic DeLury model, in which catch is assumed to be taken instantaneously at the start of the year, may induce bias in the estimates of N and F.

As with the Collie-Sissenwine model, parameters were estimated using a Levenberg-Marquadt algorithm an finite-difference Jacobian. An option to constrain all parameter estimates to be positive was also incorporated. All calculations were carried out using high-precision arithmetic.

The model estimates age-specific qs for each index and predicted indices for each year and age. The quotient of these estimates provides stock size numbers at age for each year. Using the mean stock size numbers, N(y,a), the catch, C(y,a), and the natural mortality rate, M(a), fishing mortality, F(y,a), is calculated from the conventional catch equation via a Newton-Raphson iteration. Total and spawning biomass are also calculated using N(y,a) in conjunction with input data on mean weight and proportion mature at age.

Implementation of the new model requires the estimation of a large number of parameters, and computer run time becomes a constraining factor (9 hours CPU time on a VAX 8800 is typical), but the use of some minimization method other than that of Levenberg-Marquadt may overcome this problem.

A single run was made on Data Set 1 using all ages and 20 years of data. Auxiliary data from the three research vessel surveys were used in fitting the model (commercial CPUE was ignored). Catchability was assumed constant with time, but age-specific estimates were made for ages 3, 4, 5, and 6+. All indices of abundance were given equal weight as were the three error types (measurement, process, and consistency errors). Measurement and process error residuals appeared to be well behaved and estimated q at age appeared to be similar for the three survey fleets.

A single run was made on Data Set 2 using all ages, 10 years of data, and three with the same assumptions on catchability as made for Data Set 1. The second research vessel index was given twice the weight applied to the others on the basis of "anecdotal" information supplied with the data set. Systematic patterns in the measurement and process error residuals indicated that this specification of the model may have been inappropriate.

Runs similar to those on Data Sets 1 and 2 were attempted on Data Sets 3 and 4, but no solution was obtained after extensive run times.

The model was applied to data on North Sea cod for the period 1971-1986 and for ages 1-8+. Four research vessel indices were used. For the first three indices, a single q was estimated for all ages for which data were available. For the last index, age-specific qs were estimated for ages 1, 2, and 3+. Run time was about 20 minutes.

Application to North Sea haddock data used data for years 1971-1986 and ages 0-8+. Three research vessel indices were used. Age-specific qs were estimated for ages 0, 1, and 2+. The Marquadt algorithm was constrained to providing only positive estimates by the implementation of a penalty function. Run time was about 20 minutes.

7 TIME SERIES METHOD

Full details of the estimation and application of this model are given in Gudmundsson (1987).

The main feature of this methods is that fishing mortality rates are modelled as time series, as follows:

$$\log F(a,y) = U(a,y) + V(a,y) + n1(a,y) \qquad (35)$$

where

$$U(a,y) = U(a,y-1) + n2(a,y) \qquad (36)$$

$$V(y) = V(y-1) + T1 + n3(y) \qquad (37)$$

$$\sum_{\text{all } a} U(a,y) = \text{constant} \qquad (38)$$

The residuals n1, n2, n3 are assumed to be serially uncorrelated and normally distributed with mean zero and covariances var1 x Q1, var2 x Q2, and var3, where Q1 and Q2 are given matrices.

The residuals n1 represent transient random variations. Equation (36) is associated with changes in selectivity, and equation (37) describes equal proportional changes in F at all ages.

Recruitment is represented by the equation:

$$N(1,y) = NO + T2(\text{recruitment index}) + n4(y) \qquad (39)$$

(or NO alone if no suitable recruitment index is available). The residuals have variance var4.

The measurement errors of catch-at-age observations are assumed to be serially uncorrelated with covariances s1 x H1, where H1 is a given matrix.

Initial values of the fishing mortality rates are represented by a function of three parameters, and the first year's observations are used to calculate corresponding stock estimates. The next year's Ns and Fs are predicted by means of the equations above and used to calculate catch predictions. The latter are compared to the actual catches, and the predictions of N and F updated by means of the Kalman filter before proceeding to predict the third year's values, etc.

Apart from the initial values, the unknown parameters in this model are var1, var2, var3, var4, T1, T2, and NO. These are estimated by maximizing the likelihood function of the catch prediction errors. Extensive diagnosis of residuals is performed.

Given the natural mortality rate, the estimation can be carried out with no further observations.

However, observed catch per unit effort can also be included in the estimation. Catch per effort is given as:

$$CPUE(a,y) = S(a)Cb(y)f[F(a,y)] + e2(a,y) \qquad (40)$$

f[F(a,y)] is a given function which depends on whether CPUE is obtained from a research vessel survey or a commercial fleet. S(a) describes variation of catchability with age, and is assumed constant. The residuals in this equation [e2(a,y)] represent measurement errors and irregular variations of CPUE. The residuals are assumed to be N(0,s2 x H2), where H2 is a given matrix.

Variations in catchability affecting all ages are modelled as

$$Cb(y) = W(y) + n5(y) \qquad (41)$$

$$W(y) = W(y-1) + n6(y) \qquad (42)$$

The residuals are assumed normally distributed, serially uncorrelated with zero mean and variances var5 and var6, respectively. In equation (41), the residuals represent transient variation, whereas each of the values of n6(y) affects all subsequent values of Cb(t).

With the present programs, estimation of the parameters for 10 years of data and 8 ages and ignoring CPUE data takes about 20 minutes on a VAX 8250. With 4 ages of CPUE data as well as total catches, the computational time increases to more than 1 hour.

The model was run on 10 years of data for ages 4-11 on Data Sets 1-4. Two runs were made on each data set, one run including and the other run excluding research vessel CPUE data. Only one set of CPUE data, selected by the assessor as the "best" set on the basis of trial runs, was used. For Data Sets 5 and 6, no CPUE data were included. In the latter case, it was found that much of the error in the last-year estimates was produced by T1, which is estimated with a high standard error. Addition of CPUE data should improve this situation.

8 CONVENTIONAL AND SEPARABLE VPA

The main purpose of the Workshop was to test the performance of various methods which utilize both total catch-at-age data and auxiliary (catch-per-effort) data. Conventional and separable VPA do not make use of auxiliary data, but were applied to Data Sets 1-4 mainly to demonstrate how they would perform in comparison to other methods as a basis for estimating the improvement which may be gained by the appropriate use of auxiliary data. Furthermore, work in progress (Man Sun, pers. comm.) shows that results from conventional and separable VPA can form the basis for reasonably accurate short-term catch predictions, and this might naively be taken to imply that there is no need to collect auxiliary data. However, conventional and separable VPA have no basis for estimation of true fishing mortality rates and stock size in recent years, and these quantities are important when formulating advice on conservation measures.

The conventional VPA was applied by iteratively replacing F in the last data year by average F computed for the previous 5 years and F at the highest age by average F computed for the 5 younger age groups. (This method is referred to as the JAM method; the acronym is variously expanded as the Judicious Averaging Method or Just Another Method.)

The separable VPA (Pope and Shepherd, 1982) was also applied by iteratively replacing F in the last data year by that obtained for four years previously. Terminal S was set equal to that obtained at age 7 (with unit selection at age 5).

In considering the results from these methods, it should be remembered that they are not tuning methods and should not be judged by the same criteria.

9 REFERENCES

Anon. 1986. Report of the ad hoc Multispecies Assessment Working Group. ICES, Doc. C.M.1986/Assess:9.

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

Anon. 1987b. Report of the Working Group on Methods of Fish Stock Assessment. ICES, Doc. C.M.1987/Assess:24.

Allen, K.R. 1966. Some methods for estimating exploited populations. J. Fish. Res. Board Can. 23:1553-1574.

Baker, R.J. and Nelder, J.A. 1978. The GLIM system, release 3, generalized linear interactive modelling manual. Numerical Algorithms Group, Oxford.

Collie, J.S. and Sissenwine, M.P. 1983. Estimating population size from relative abundance data measured with error. Can. J. Fish. Aquat. Sci. 40:1871-1879.

DeLury, D.B. 1947. On the estimation of biological populations. Biometrics 3:145-167.

Deriso, R.B., Neal, P.R., and Quinn, T.J. II. 1988. Further aspects of catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. In press.

Doubleday, W.G. 1981. A method of estimating the abundance of survivors of an exploited fish population using commercial catch at age and research vessel abundance indices. In W. G. Doubleday and D. Rivard (Ed.), Bottom trawl surveys. Can. Spec. Publ. Fish. Aquat. Sci. 58, 273 p.

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

Hogg, R.V. 1979. An introduction to robust estimation. In R.L. Launer and G.N. Wilkinson (Ed.), Robustness in statistics. Academic Press, New York.

Laurec, A. and Shepherd, J.G. 1983. On the analysis of catch and effort data. J. Cons. int. Explor. Mer 41:81-84.

Lewy, P. 1988. Assessments of North Sea cod and haddock based on two tuning methods applying commercial catch, effort and research vessel data. ICES, Doc. C.M.1988/D:20.

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. 1985. A comparison of the performance of various methods for tuning VPAs using effort data. J. Cons. int. Explor. Mer 42:129-151.

Pope, J.G. and Stokes, T.K. 1988. The use of multiplicative models for separable VPA, integrated analysis and the general VPA tuning problem. Am. Fish. Soc. Spec. Publ. In press.

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

Rivard, D. and Foy, M.G. 1987. An analysis of errors in catch projections for Canadian Atlantic fish stocks. Can J. Fish. Aquat. Sci. 44:967-981.

REPORT OF THE WORKING GROUP ON METHODS OF FISH STOCK ASSESSMENTS

Nantes, France, 10-17 November 1989

1 INTRODUCTION

1.1 Participants

V. Babayan	USSR
F. Borges	Portugal
R. Chevalier	France
R.J. Conser	USA
R. Cook	UK (Scotland)
S. Fryer	UK (Scotland)
S. Gavaris	Canada
H. Heessen	Netherlands
M. Hildén	Finland
J. Hoenig	Canada
I. Horbowy	Poland
T. Jakobsen	Norway
M. Kizner	USSR
A. Laurec (Chairman)	France
R. Mayo	USA
B. Mesnil	France
J. Modin	Sweden
R. Mohn	Canada
R. Myers	Canada
P. Neal	USA
D. Pelletier	France
R. Rivard	Canada
J. Shepherd	UK (England)
H. Sparholt	Denmark
G. Stefánsson	Iceland
K. Stokes	UK (England)
S. Turner	USA

1.2 Terms of Reference

At the 76th Statutory Meeting of ICES (1988), it was decided (C.Res.1988/2:4:16) that the Working Group on Methods of Fish Stock Assessment (Chairman: Dr A. Laurec) should meet in Nantes from 10-17 November 1989 to:

a. Consider the construction of CPUE and survey indices by detailed analysis of spatially disaggregated data.

b. Advise on the implications of the timing of Working Group advice on the precision of TACs and on what implications a change in the TAC year to 1 April - 31 March would have on the precision of advice.

c. Consider the report of the July 1988 Workshop in Reykjavik and any matter arising on methods of VPA tuning.

1.3 Structure of the Report

Preliminary work had been conducted on a variety of subjects, as reflected in the list of Working Papers given in Appendix A. Some of these papers are unpublished (interested readers should contact the authors directly). The most important findings have nevertheless been taken into account in the report.

During the meeting, most of the work has been concentrated on item a, further calculations being conducted on three data sets (EGFS data for cod age 1 and 2, Icelandic cod survey data and Icelandic commercial data). The discussions are reported in Section 2.

Concerning item c, it was decided not to try new calculations, but to reconsider the details of the results obtained in Reykjavik, taking account of the various remarks about the existing report. The results of these reconsiderations are given in Section 3. It also appeared useful to give further details for users of the Lowestoft tuning package. These are given in Appendix B.

Prior to the meeting, the General Secretary of ICES submitted a request to the Chairman of the Group, concerning the assessment of North East Arctic cod. It appeared useful to discuss this issue since it gives an illustration of potential problems when using tuning techniques and is thus relevant to previous work of this Working Group. The conclusions are given in Appendix C.

Section 4 is devoted to item b. The general conclusions and recommendations are given in Section 5.
In addition to Appendices B and C mentioned above, five appendices have been included. Some correspond to theoretical considerations and potential methods that can be developed but cannot be recommended at present (Appendices D and E). They just suggest possible paths for future works.

Three other appendices have been written for readers not familiar with basic issues such as the various meanings of randomness (Appendix F), the use of estimators other than a simple arithmetic mean for estimating a spatial mean (Appendix G), or the consideration of interactions in multiple-factor models (Appendix H).

2 PRE-PROCESSING

2.1 Introduction

2.1.1 Background

The quality of the results given by all tuning techniques depends directly on the quality of the relationships between fishing mortalities and fishing effort, or (and often equivalently) between stock abundance and CPUE or survey indices. It is thus crucial to use a satisfactory definition of the fishing effort and the best possible indices of abundance. Such attempts are a necessary complement to the efforts developed throughout the years to get the most efficient tuning techniques.

It is difficult within a Working Group to deal with the large amount of disaggregated data necessary for an efficient comparison of the possible pre-processing methods. Such data correspond to logbook items or landing results by trip. Most of the work during the meeting has been concentrated on survey data. As stated by Shepherd (*Anon.*, 1986):

"*The long-standing problem of estimating abundance from research surveys is essentially the same as estimating the quantity of coal in a heap. One needs to integrate under the surface and this may be done from either regularly- or irregularly-spaced observations of the height of the surface. The individual observations are subject to sampling error and, therefore, follow some probability distribution (may be log-normal, Poisson or negative binomial for samples of fish or plankton).*

The distribution of the ensemble of observations is, however, determined by the location of the observations and the shape of the heap. There is no reason whatsoever to expect them to conform to any recognizable probability distribution, and any resemblance must be regarded as fortuitous and unpredictable. For example, the distribution would change if the location of the observations were changed. The spatial distribution is, in fact often quite systematic and repeatable from year to year, and adjacent observations are highly correlated.

This high spatial autocorrelation is a further reason for not regarding the observations as drawn randomly from a population, but may be utilized in analyzing the data using appropriate methods. Techniques which make use of spatial autocorrelation are commonly used in geostatistics (Kriging), meteorology and physical oceanography (objective analysis), and in algorithms for contouring. Developments of these for use in analysis of marine survey data are probably feasible.

There are essentially two valid approaches to the analysis of survey data, namely integration under the surface (possibly using the techniques noted above, but presently usually by constructing appropriate summations over the observations), and multiplicative modelling. The latter technique relies on systematic repeatability of the spatial pattern from year to year (whilst the former does not), but it is better able to cope with incomplete and error-prone data. Which method has the better performance must depend on the extent to which the spatial pattern is indeed repeatable from survey to survey, and this needs to be further investigated."

In complement to this summary, it must be pointed out that beyond strictly multiplicative models, which can be linearized by a logarithmic transformation, more complicated ones may be fitted, taking into account interactions between years and other factors. The use of such models may, however, be dangerous. This question was discussed during the meeting. The only possible way for models to include interactions with years, appears to be to use the fitted values for interpolating at any place before integrating the corresponding estimated densities over space. This will, in fact, correspond to a very intensive smoothing procedure while direct interpolations may lead to much less smooth interpolated spatial patterns.

Stratified sampling schemes have also been commonly used for constructing abundance indices. They have been considered by the Working Group. They may, in fact, be connected with the previous discussion about the various levels of smoothing intensities. The construction of a total index of abundance requires the estimation of the average density within each stratum. Spatial variations within a stratum are absorbed by the averaging. With spatially large strata, this corresponds to a very intense smoothing. The smoothing intensity is, in fact, related to the size of the strata.

2.1.2 General comments

A - For a clearly defined purpose

Research surveys can be used for several purposes. Two of them will be considered here:

- the estimation of annual changes in total abundance (year effect),
- the mapping of an average (over years) spatial distribution (space effect).

The results of a series of surveys conducted over several years, at the same period, will depend on the combination of these two effects, and to interaction terms, corresponding to changes in the spatial distribution from year to year.

The main purpose considered here will be the estimation of changes in abundance from year to year. Some comments on this issue will be presented. The differences with studies aiming at the mapping of resources will be stressed.

B - Absolute abundance/relative abundance

Constant absolute and relative biases

The estimation of annual changes does not require unbiassed estimates of the total abundance for each year. It just requires a constant bias from year to year. More precisely, if absolute differences were looked at, the absolute (additive) bias would have to be constant. In most circumstances, however, relative changes will be the most important feature. This appears to be the case in tuning, when logarithms of abundance are required. Statistically this corresponds approximately to a constant relative bias on absolute numbers. So, when looking for unbiassed estimates of changes in abundance, one should keep in mind that three points of view can be considered: constant additive bias on the abundance, constant multiplicative bias on the abundance, constant additive bias on the logarithm of the abundance.

Catchability

It is just as well that an estimate of the absolute abundance is not necessary, because it is also impossible. The local catchability, sometimes called vulnerability (and corresponding to the efficiency of the trawl) is generally unknown. This does not make it impossible to get usable estimates of relative abundance, as long as the efficiency is constant. It has to be constant from year to year, but also from place to place. More precisely, if the spatial distribution of the fish is persistent from year to year, its vulnerability may vary over space. But it is much more difficult, if not at present impossible, to deal with situations combining a high level of spatial variation in vulnerability with strong changes from year to year in the spatial distribution of the stock.

Here lies the fundamental limitation. If for one year fish concentrate in an area where they are highly vulnerable, it will be difficult to compare the corresponding apparent density with another year where the stock is mainly located in an area where the gear has a poor efficiency.

C - Statistical inferences

Assumptions

Biases have been mentioned several times. Apart from biases, it will be essential to provide at least a rough estimate of the reliability of the estimated relative abundances, and preferably estimates of the corresponding variances. As well as for the estimation of possible biases, this will require a number of assumptions, corresponding to a more or less precisely constrained model. The development of resampling techniques (Jackknife, bootstrap, cross-validation) makes it possible to try making inferences without assumptions on the statistical distribution of the residuals. They do not cancel the problems associated with correlations between residuals obtained after a model has been fitted. Once more such a dependence will generally be associated with the space/year interaction questions. This emphasizes the need for a careful examination of the residuals.

From a more general point of view, the various hypotheses should be systematically checked. This includes the identity of the stock and may include a possible spatial pattern of vulnerability, interfering with a change in the spatial distribution. If a full correction seems impossible, indications can be available on changes in the gear efficiency associated with various factors (e.g., depth). In such a case, if for a given year the relative abundance appears much higher in deeper or shallower waters, biases are likely to occur.

The most constraining hypothesis is that of constant catchability from year to year. Changes will be very difficult to detect in real time. However, the correlations between apparent abundance and various environmental factors can be studied possibly within a year over the various stations, and (preferably) over years, for past years where the existence of VPA estimates makes it possible to calculate annual catchabilities. Any anomaly in the relevant factors for the current year would suggest a necessary correction in the apparent abundance.

The various variances

Estimating the variance of the estimation of the integrated abundance and of the year effect may require sophisticated treatments, but when a stratified sampling scheme is being applied, simple formulas are used. The intrastratum variance, due to space effects within a stratum, will affect the arithmetic average over the samples as an estimate of the mean density. If one is interested in the difference from year to year in this mean density, the calculation of variance can be misleading. If the spatial structure within a stratum is stable from year to year, the real uncertainty can be smaller than suggested by the sum of the annual within-stratum variances, as usually calculated, provided fixed stations are used.

What has been said here for stratified sampling is true for various techniques providing variance estimates, such as Kriging. Most existing formulas to estimate the variance of an abundance index, for a given year, refer to the discrepancy between this index and the real abundance. When, from year to year, the spatial pattern is totally or partially persistent when the sampling design is totally or partially constant, the error in year-to-year changes will be different from what is calculated by usual procedures, which refer to spatial changes within a year. Generally speaking, when the catchability is strictly constant, year-to-year variances will be overestimated by "spatial" variances, the overestimation being bigger when the spatial distribution is markedly persistent from year to year.

The comparison of the constructed indices and VPA estimates for past years has been mentioned. A logarithmic transformation is often useful for various reasons, including the statistical distribution of the real recruitments and the consistency with the tuning techniques.

A regression of the recruitment (survey) indices on VPA results makes it possible to calculate residuals and their variance. If VPA results can be considered as error-free, this variance corresponds to the "real" uncertainty to be taken into account when considering the index as an estimate of a year-class strength. Such a variance includes several components:

- Changes in catchability from year to year which do not depend on the sample sizes within each year.

- Spatial heterogeneity in fish abundance which creates a "sampling error" due to the fact that only a finite number of locations will be sampled.

It has been said previously that the formulas for estimating annual variances in the case of Kriging or stratification can be misleading and would tend to overestimate the real sampling variances. On the other hand, they do not take into account the component associated with the catchability. They can be, in fact, very difficult to use in practice. The same can be said about variances calculated for year effects when fitting a multi-year model since the assumptions about the residuals required by the simple formulas have little chance of being fulfilled.

D - Structure of year-to-year changes in the spatial pattern

If the spatial distribution were stable from year to year, it would be very simple to monitor year-to-year changes. A single haul per year at the same location would be sufficient. If this simple scheme was just perturbated by a white noise, replicates would be efficient.

Since, in most cases, changes in the spatial distribution from year to year would occur, showing some consistency in the sense that changes in neighbouring places are correlated, more sophisticated sampling designs are required.

The changes in the spatial distribution can occur at various spatial scales. Macrostructures can be affected, which implies that year after year the sampling design should cover the whole area occupied by the stock. Changes can also affect microstructures at a scale from a few (latitude) degrees to tens of miles (very small structures can just be considered as creating extra white noise).

Changes in macrostructures can be associated with various phenomena. It may happen that, within what is considered to be a single stock, various sub-stocks exist associated with different geographical areas. In such a case, if the abundance varies in a different way for each sub-stock from year to year, the relative abundance in the various areas will change from year to year. This does not necessarily affect the spatial distribution for a given sub-stock in its specific area: macrostructure can be affected but not necessarily the smaller scales.

Since space/year interactions create most of the difficulties, they should be studied carefully, and could imply important changes in the procedures of assessment and management of the stocks.

When the major phenomenon corresponds to the "sub-stocks model", it can be envisaged to apply a multiplicative model to each sub-area rather than to the whole area covered by the surveys.

Changes in spatial pattern can also be related to changes in stock size (Myers and Stokes, 1989). For example, as a stock increases in biomass, it may fill the best habitat available. This may lead to density increasing relatively more in selected areas.

Local inconsistencies seem more difficult to correct for simply.

E - Distributional form

Knowledge of the sampling distribution of catch per tow from research vessel surveys is essential to selecting the correct model to use. For each stratum in each year of the English and Iceland groundfish surveys, we calculated the sample mean and variance. In general, the log variance increased approximately linearly with the log mean with a slope of 1.8 to 2. There was a slight tendency for the log variance to be a convex function of the log mean [Figures 2.1.(3), 2.1.(4) and 2.1.(7)]. This type of relationship is consistent with a negative binomial sampling distribution with the k parameter approximately equal to one, because the variance of a negative binomial distribution is the mean (1 + mean/k).

Thus a reasonable model to use for the error structure is the negative binomial in which the variance is the square of the predicted mean. Such models can be fitted using the user-defined model in the GLIM statistical package. If there were not too many zeroes, e.g., less than 10%, a log transform with a constant added to the numbers per tow may also be acceptable. The Poisson distribution, or an extra Poisson distribution in which the variance increased in proportion to the mean, would not be good models for these data.

The choice of the constant added before the log transform is quite important. The smallest encountered non null value is commonly referred to, but it may be divided by a factor ranging from 1 to ten provided the required added constant. The variance stabilization is illustrated by Figure 2.1.(5).

F - Other approaches

Five additional approaches to constructing an index of abundance can be considered. These are:

1 - Principal component analysis

2 - Regression of difference estimation

3 - Bayesian estimation

4 - Empirical Bayes estimation

5 - Time-series analysis

Time-series analysis requires a long series of observations to be available, especially if the dynamics of the system being modelled are complex. The use of time-series analysis in fisheries research has been explored in a number of studies and consequently will not be considered further here.

Some comments on the remaining four approaches are given below.

Principal component analysis

The results from a series of surveys can be grouped into a matrix form, each year corresponding to a column, each station to a row. A principal component analysis can be performed on such a matrix using the covariance (or correlation) matrix between stations. A first factor associated with a year effect is expected.

A transformation will generally be necessary beforehand. When a logarithmic transformation is performed, the expected underlying model corresponds to the multiplicative model. The common factor is associated with the year effect. In the simplest situation, no extra significant factor should be found, the residuals being uncorrelated from station to station. A more complex situation can, however, be exhibited in which extra factors group "similar stations". It may even happen that instead of a clear first common axis, the examination of the eigen values result in identifying several factors grouping, for instance stations, associated with possible sub-stocks.

PCA can be useful. It should, however, be used only as an exploratory tool rather than as a technique which provides directly an annual index of abundance. It can even be combined with the fitting of a multiplicative model, the correlations being calculated on the residuals.

Regression or difference estimation

Difference and regression estimation require that there is an auxiliary variable X_i availale which:

a - is correlated with the variable of interest Y_i (Y_i = CPUE at station i)

b - is known for every observation (haul) i

c - has a known mean U_x.

The trick is to find such an auxiliary variable. One way to do this is to fit some sort of model to previous survey data. A prediction can then be made for each haul location (i) in the current survey. The mean of the predictions, U_x, is obtained by integrating or summing over all explanatory variables in the model, e.g., by integrating over space. The reduction in variance achieved by using the auxiliary variable can be shown to be equal to the correlation (r^2) between X and Y.

The estimation formula is given by:

$$Y^* = Y + C\,(U_x - X)$$

where Y^* is the estimated mean CPUE, Y is the mean of the current year's observations on CPUE, X is the mean of the model predictions corresponding to the locations sampled to obtain Y. This estimator, in effect, says that if X is below its mean U_x, then Y is probably below the true mean CPUE. Further, the amount of correction to be applied depends on C: C is the change in Y per unit change in X.

If C is fixed in advance, the result is a difference estimator which is unbiassed. The bias will be zero even if the model used to obtain the auxiliary variable (predictions) has lack of fit problems. For example, ignoring interaction terms in the development of the model will not cause bias (but will reduce the gain in precision achieved by using the difference estimator).

C can also be estimated. The optimal value of C is the coefficient of regression of Y on X. Note that if C is estimated from the data, the estimator is known as a regression estimator and has a statistical bias of order N-1 (where N is the number of hauls).

Bayesian and empirical Bayes estimator

Bayesian estimation attempts to combine the current information from the trawl survey, with "prior" information. The prior information can be based on past observations, intuition, etc. The user must specify a prior distribution (form, mean, variance). These pieces of information are then combined in an optimal manner using Baye's rule.

However, since the specification of the prior distribution is subjective, different people will get different answers. The results are, therefore, likely to be highly controversial unless a group consensus (on the nature of the prior distribution) can be achieved, or unless the results can be shown to be insensitive to the choice of the prior distribution.

Empirical Bayes estimation circumvents the problem of subjectivity by considering past survey results as a representative sample from the prior distribution. The estimator turns out to be a weighted mean of the current survey results and the results from prior years. Those years with low estimated variance of the index will receive higher weights than those with high variance (all other things being equal); also, those years that agree well with the current survey results will receive higher weights than those that agree less well.

Empirical Bayes estimators are biassed but they can be shown to reduce the mean squared error under very general conditions (Cassella, 1985).

The assumption that the past survey results provide a representative sample of the prior distribution will not be met if there is a trend over time. However, it turns out that the method is robust to this failure of assumption. Intuitively, it can be seen that, since the weighting depends on how well the observations agree with the current results, observations further back in time will (in this case) receive low weights and thus not affect the estimation very much.

This question has been indirectly addressed in a working document to the previous meeting of the Working Group (Laurec and Souplet, 1987), and is taken into account by the RCRTINX2 package.

2.1.3 Data sets

Three basic data sets have been considered to exemplify the methods: the indices for ages 1 and 2 of cod from the English groundfish surveys, the indices for age 1 cod in Icelandic surveys, and commercial CPUE for Icelandic trawlers.

A - English Groundfish survey data

The English Groundfish survey (EGFS) of the North Sea has been described in a working paper to the Roundfish Working Group (Harding and Macer, 1986). The survey has run since 1977 and was originally planned with a stratified random design (stratified by depth and bottom type). Since 1977, stations have been fixed on the original design but with numerous changes and deletions. There are now 46 "primary" stations and 30 "secondary" stations which are targetted each year.

For this Working Group, all years' CPUE data for 1- and 2-year-old cod, together with trawling latitude and longitude, have been made available. Additionally, data from 1977 to 1988, during which time replicate trawls were made at each station, have also been provided with a set of redefined stations (clusters). These clusters have been defined, using the CLUS procedure available in the SAS statistical package, so as to provide a set of unique, non-overlapping areas that divide the 570 trawls into 50 strata each covering a circular area of 450 nm^2. The mean log (CPUE + 0.1) for ages 1 and 2, by cluster, are shown in Figures 2.1 (1) and 2.1 (2).

The 1- and 2-year-old cod data from EGFS have previously been considered by Houghton (1987) who attempted to characterize the nature of different fish distributions in terms of their simplicity/complexity and persistence/changeability. He characterized the young cod distributions as complex and changeable, i.e., the distributions could not be represented by a simple surface and showed no annual consistency. Myers and Stokes (1989) have also considered EGFS data; they attempted to characterize how space is used as fish populations expand or contract. For young cod, they showed that there is no simple, overall response and that no habitat saturation occurs but rather, that some areas respond more than others to changes in overall population (i.e., that the distribution of young cod is changeable rather than persistent).

B - Icelandic survey data

An Icelandic research survey is carried out every year. The stations in this survey are fixed and number 580 per year. The survey grid is shown in Figure 2.1.(6). At the meeting, the Working Group had at its disposal the number of 1-group cod at each fully recorded station of the survey. Some missing values had been encountered during a previous analysis of the data (e.g., depth or temperature) and those stations have been dismissed for the purpose of the analysis. The number of 1-group cod at each station have been computed by analyzing the total length at age for each year, finding an accurate cut-off point for the 1-group and using that for each station.

C - Icelandic commercial data

This data set from commercial vessels has been previously analyzed by Stefánsson (1988). The raw data are recorded by trawler captains and comprise catch estimated by species for each tow, position (in terms of statistical square) and towing time. For the purpose of analysis, the basic data are aggregated to the level of trawler, month, square, i.e., catch and towing time are recorded only as a total for all tows that a trawler takes within a statistical square in each month.

It should be noted that there is a different stock composition in the Northern and Southern regions, due partly to seasonal migration. Thus, analysis of Icelandic cod data is normally done separately for the time periods January - May and June - December, as well as for the Northern/Eastern and Southern/Western regions. Only the northern area was considered [Figure 2.1.(8)].

2.2 Stratification

2.2.1 Introduction

Decreasing the size of the strata will decrease the intra-stratum variances. This suggests a potential decrease of the final calculated within-year sampling variances. On the other hand, creating more strata while holding the number of hauls constant implies smaller sample sizes within (at least some) strata. If the set of smaller strata is more homogeneous than the set of larger strata, then a variance reduction will result.

Otherwise, creating more strata can result in an increase in variance (Cochran, 1977). This is due to the smaller sample sizes for each stratum (the variance of the mean is s^2/n) and also to the fact that, if each stratum is allocated a minimum of two hauls, then there may be few hauls left to allocate optimally. As an example, Gavaris and Smith (1988) found that precision could be improved in the Scotian Shelf (Canada) groundfish survey by reducing the number of strata (i.e., combining strata).

This is the reason why on the EGFS data for age 1 cod, a series of stratification schemes using progressively smaller strata, has been tried. Since transformations are commonly used, this influence has also been checked. The corresponding results appear in Section 2.2.2. Section 2.2.3 gives a summary of the discussion which took place on post-stratification. Section 2.2.4 offers some concluding remarks.

2.2.2 Application to EGFS data

A - Description of the various schemes

The North Sea has been simplified to 12 squares as indicated in Figure 2.2.(1).

The first scheme treats the whole North Sea as a single stratum. The second one distinguishes three strata, associated, respectively, with Divisions IVa, IVb and IVc.

The third scheme considers a sub-division of Divisions IVa and IVb between three sub-areas, creating eight strata. The "eight strata" scheme groups blocks are indicated here between brackets: (3, 4, 5), (6, 7, 8), (9, 10), (11, 12), (13, 14). Blocks 1, 2 and 15 remain isolated. The fourth scheme takes into account fifteen strata using a further sub-division of the strata considered in the previous scheme. These fifteen blocks are numbered in Figure 2.2.(1).

The sampled locations are always considered as if they had been randomly chosen within a stratum although this is not strictly the case.

B - The estimators

For each stratification scheme, various options have been tried for deriving the overall survey index, depending on the stage at which a transformation (if any) was applied to the data. In all cases, the stratified mean is constructed by weighting the estimated stratum mean in proportion to the stratum area (i.e., by the number of rectangles) and eventually log-transforming for comparison with the log VPA estimates.

The simplest estimator makes use of the raw data, the stratum mean being the arithmetic average of results in each haul (untransformed estimates). The stratum mean can also be the approximate geometric mean of individual observations, in which case these are first transformed as $y = \log(x + 0.1)$ and the average back-transformed as $x = e^y - 0.1$ before integrating over the surface (transformed estimates).

Since the simple back-transformation may introduce a bias, a correction factor can be applied beforehand (see Appendix F).

Here the term $\hat{s}(\frac{n-1}{n})$ where n is the number of hauls and $\hat{s}$ the estimated variance within the stratum, has been added to the mean of the logged data before back-transforming to yield the stratum mean (corrected transformed estimates).

C - The results

The various estimated series appear in Table 2.2.(1) as well as the corresponding logarithms of VPA results. The correlation (r^2) with VPA results is given for each estimator. It must first be pointed out that if the results obtained without preliminary transformation are quite consistent from one sampling scheme to another, important changes appear when logarithmic transformations are used. Regardless of the +.1 additive constant, this corresponds to the classical differences between geometric and arithmetic means, the later one being systematically larger. The higher the dispersion of the individual values, the higher the discrepancies are.

When a single stratum is considered, this discrepancy is very high. When small strata are used, the between-strata variations being eliminated, the discrepancies will tend to decrease.

The overall estimated abundance obtained after transformations will tend towards more consistent values than those obtained through "non transforming" procedures. It can also be noticed that the r^2 values obtained for transformed uncorrected estimators are higher when more disaggregated sampling schemes are used. However, they remain less efficient, in terms of r^2, than those based on raw data. The latter are not very sensitive to the sampling schemes. A single stratum even appears to be a "good" solution. It must, however, be recalled that the considered sampling scheme is not a Simple Random Sampling one, and that it assumes that a satisfactory coverage of the whole North Sea at a macroscopic scale is provided. Otherwise stratifications would be likely to be necessary.

Coming finally to the corrected transformed estimators, it can be noticed that if a single stratum is considered, the correction seems to be useful. It can be related to the fact that within the whole North Sea, the statistical dispersion of the numbers per haul is high. When a detailed stratification is being used, correction factors become less necessary, at least when constant biases are accepted. The correction can even become dangerous, as suggested by the fifteen strata scheme, where the corrected series gives a poorer correlation than the uncorrected one. This may be related to the fact that, due to the limited number of hauls within a stratum, the s^2 values appearing in the correction factors are poorly estimated. They tend to introduce an additional noise. More sophisticated bias correction techniques can be used, including resampling techniques. The possibility of introducing extra noise through the correction could nevertheless remain.

2.2.3 Post-stratification

There are a number of reasons for considering strata, including statistical efficiency, logistic/administrative convenience and interest in the strata *per se* (e.g., a stratum may correspond to a political or administrative zone for which estimates are needed).

Sometimes it is not possible to delineate the desired strata beforehand. For example, one may wish to stratify by surface water temperature but a map of surface temperatures may not be available until after the survey. Also, different species may have different spatial distributions so that the sampling requirements for one species may conflict with those for another.

In these cases, one may wish to use simple random sampling or proportional allocation and then post-stratify into strata that are believed to provide greater precision. Post-stratification must be on the basis of auxiliary information not on the basis of the observed variable of interest (CPUE). Otherwise, the estimated variance can be made arbitrarily small and is meaningless.

In certain circumstances, post-stratification can be extremely effective. Note, however, that it uses prior information only to improve stratification but not in the estimation itself. Some techniques described in Section 2.1.1 explicitly use prior observations in the estimation of the current index.

2.2.4 Concluding remarks

The previous discussion does not cover all the problems related to the choice of a stratification scheme. The possibility of using, for a given year, the results from the previous ones has been explored by the Working Group on the EGFS data set. However, it did not give interesting improvements, due partially to the limited number of years for which this was possible.

The possibility of defining optimal clusters has been mentioned in Section 2.1.2, which indicates that the statistical packages such as SAS can offer this possibility. A comparison of the r^2 values obtained from such a scheme with those obtained as described in Section 2.2.2, using the same data, indicates that the gain would be small. But again, this may be due to the specificities of the data set.

If transformations are to be used due to the lack of robustness of formulas based upon log-normality assumptions, resampling techniques would be useful for simultaneously correcting biases and estimating variances.

When fixed stations exist, it would also be useful to consider year-to-year changes for each one, since it appears that considering the variances of these differences would make it possible to build a sampling variance really referring to the estimation of year effects.

When some stations are fixed, and others are randomnized every year, it would be very useful to compare the efficiency of the corresponding estimators for a given number of hauls. Resampling techniques would also be useful in this respect.

Although further investigations on this topic are warranted, a clear conclusion is: if preliminary transformations are to be used, it will be very dangerous to change the stratification scheme from year to year.

2.3 Generalized Linear Models

2.3.1 Introduction to general linear models

A GLM (General Linear Model) can be characterized as a model where each measured value has, after a specified transformation, if necessary, a mean which is a function of a linear combination of independent variables, with a specificed error structure. More specific definitions can be found in the GLIM manual. A good source reference is McCullagh and Nelder (1983). Such models have been found useful when computing indices of stock abundance (cf. Myers and Pepin, 1986).

They can also be used for mapping, although this has not been commonly done. One of the greatest virtues of the GLM models is the possibility of including variables which affect catchability (such as wind speed and direction, depth, etc.) when the higher level interaction does not exist. Missing strata in a given year can be estimated. Some relevant examples will be given in what follows.

When a GLM model is fitted, it is necessary to specify the structure of the underlying mean. This is usually done by either defining factors describing areas, or a combination of factors and regressors.

<u>1 - Log-transforms</u>

One of the simplest linear models for commercial or research vessel catch-per-unit effort data is obtained by using logarithmic transforms. In the following general model description, Y will refer to a catch measurement and X will refer to an auxiliary parameter (which may be multi-dimensional and the effect bx may also be a factor level). A basic model of the form:

$$Y = \exp(a + bx) \exp(E)$$

(where E is an error term) becomes

$$\ln(Y) = a + bx + E$$

after the simplest logarithmic transformation. If the latter model is fitted using least squares, it is implicitly assumed that the errors, E, have a common distribution, with a constant variance. Such a model fit will give optimal estimates of the parameters a and b if E is normally distributed, i.e., if the errors E in the original model are normally distributed and not autocorrelated.

All aspects of the model are optimal on a log scale as long as the assumptions of log-normality hold, but several problems appear when log-transforms are used, including potentially large biases on the untransformed scale unless corrected for the exact variance. When the variance correction uses a variance estimate, it should be noted that the corrected back-transformed estimate may be highly inaccurate (see Section 2.2.2). What has been said previously about the lack of robustness of variance-correction formulas for bias correction is also true for GLM.

The behaviour of the log-transform also depends quite heavily on which additive constant, c, is chosen. When a log-transform is used, different constant values should be tested.

The item of interest, namely biomass, is obtained by integrating the biomass surface. The analysis of the fitted values requires integrating the exponentiated fitted values in a log-transformed model. When a geometric mean is used, it is not at all clear how this corresponds to the intergration of transformed values. It is, therefore, strongly suggested that integration (or averaging or summation over stations/strata) be performed on the original scale.

<u>2 - Models</u>

Packages such as GLIM allow formal modelling of the mean and distribution, which is then used for fitting. It is thus possible to define that the expected CPUE at a given time and location is of the form exp(linear effect), with some variance (which may depend on the mean) and some parametric sampling distribution (e.g., Poisson or negative binomial distributions).

The linear effect may be of any of the following forms:

- list of factors describing areas
- simple representation of the CPUE surface (e.g., quadratic)
- additional variables such as time of the year or depth strata.

The distribution used can be taken as normal after a transformation or as other members of the exponential family, e.g., Poisson.

The Poisson distribution assumes that the mean is proportional to the variance. This is usually not true for CPUE data, as they tend to be overdispersed compared with the Poisson. The GLIM package allows the user to overcome this problem by defining different variance-to-mean relationships. Testing such models was not possible during the meeting, but the Working Group notes that defining appropriate models may have some benefits in terms of obtaining more accurate estimates.

Spatial and temporal interactions with year effects are often present in CPUE data., When linear models are fitted assuming no interactions, a potential problem can exist, e.g., density-dependent habitat utilization (Myers and Stokes, 1989). If the interactions are an important source of variation, it is essential to include them in the model, and to use integration over the fitted values (or summation over strata) to obtain an annual biomass index. It should be noted that when the model includes interactions, the year effects as estimated by linear models do not have any clear interpretation (see Appendix H). When the interactions are of little importance, it may not be of great consequence to omit them from the model, but the effects of this are not clear.

2.3.2 Application of the EGFS data

Indices for 1- and 2-year-old North Sea cod have been calculated using a variety of generalized linear models. The standard index (arithmetic mean weighted by roundfish area) exists for the years 1977 to 1988. A geometric mean index has also been calculated for all of these years. This index has been calculated as mean (log[CPUE+0.1]) where 0.1 is the minimum, non-zero CPUE in the data set. This index has been calculated from the clustered data (see Section 2.1.1) and uses only those 37 clusters that are common to all years. All other indices are calculated from the data set with replicates and, therefore, only cover the period 1977 to 1981.

The cod distributions have previously (Houghton, 1987) been characterized as complex and changeable. The ideas of complexity/simplicity and changeability/persistence may be formulated in a series of linear models.

$M_i + Øy + Iiy$ .... complex, changeable

$M_i + Øy$ .... complex, persistent

$Q(i) + Øy + IQ(i)_y$ .... simple, changeable

$Q(i) + Øy$ .... simple, persistent

where:

i — index rectangles

y — indexes years

M_i — rectangle (spatial effect)

Q(i) — parameterized functional form for spatial variation

Øy — year effect

I_{iy} — rectangle year interaction

$IQ(i)_y$ — space year interaction (interannual variation of parameters of Q)

Houghton (1987), for example, chose a quadratic for Q(i).

We have created indices for a complex, changeable model using both log(CPUE + 0.1) transformed data and a normal error structure, and untransformed data using a Poisson error structure and a log-link function (see Section 2.1.1). Additionally, a series of (easier to fit) complex, persistent models have been used to create indices. This set of models was run to test the effects of transforming data. An untransformed data set was fitted using a Poisson error structure, and four log-transformed data sets (with constants 0.1, 0.25, 0.5 or 1.0) were fitted using a normal error structure.

Indices created using all methods are presented in logarithmic form in Table 2.3.1. Also shown are the r^2 values and the residual standard errors for the regressions of each series on the VPA series. The VPA numbers at age have been taken from the 1989 Roundfish Working Group report.

Considering the 1-year-old cod indices first, all methods produce excellent results. For this species, especially at this age, even a simple (or weighted) arithmetic mean captures 97 to 99% of the variance and has a low residual standard error. It must, however, be recalled that the fitted series is very short (six points). The complex persistent models used to test the effects of the various data transformations and error structures, all perform well in terms of variance explained, even though the model is inappropriate. The residual standard errors from the regressions of these indices (or the geometric mean) on the VPA numbers are, however, relatively high.

By comparison, the complex, persistent model fitted with a Poisson error structure performs very well both in terms of variance explained and reduction in residual standard error. Concerning transformations, it is difficult, from the models applied to this data set, to make firm conclusions. What can be said is that if data are log-transformed, then care is needed in choosing the transform constant. When the underlying error structure is known it is so far safer (and better) to fit models that take acount of this.

The indices for the 2-year-old cod all perform less well than those for the 1-year-old fish. The most striking difference in their performance is not in that they consistently explain less variance when regressed on the VPA, but that the residual standard errors are always high. As with the 1-year-old indices, it is difficult to draw firm conclusions from the comparisons.

2.3.3 Indices of abundance from commercial catch and effort data

This topic was addressed during the first meeting of the Methods Working Group (*Anon.*, 1984). In particular, the multiplicative model (Fonteneau and Laurec, 1987; Gavaris, 1980) was examined and found to be a useful tool for developing indices of abundance from commercial catch and effort data. In principle, this model is the same as the general linear models (GLM) discussed in the previous sections for developing indices of abundance from research survey data. The primary difference lies in the use of year interaction terms, which were incorporated in the GLMs for survey indices (Section 2.3.1), but have not been employed generally in developing indices from catch-effort data. The utility of these terms and the practical problems associated with their use were not explored fully during the 1984 Methods Working Group meeting.

Catch and effort data from Icelandic trawler reports (Stefànsson, 1988) were used to examine the effect of incorporating year interaction terms into the GLM. These data are described in Section 2.1.2. Stefànsson (1988) partitioned the Icelandic cod data from these trawler reports into four components, reflecting two areas (North and South) and two seasons (Spring and Fall). The North-Spring component was examined by the Working Group. These data were available by statistical rectangles (one by one-half degree) and by month. The statistical rectangles were aggregated into four larger areas based on biological and distributional information as well as historical fishing patterns (Figure 2.1.(8)).

Following Stefànsson (1988), only records in which the cod catch (in weight) exceeded 50% of the total catch were used in the analysis. A ten-year subset of the available data, 1974-1983, was selected to facilitate comparison with the converged portion of the Icelandic cod VPA. The resulting data contained 7,611 catch-effort observations distributed over 10 years, 4 areas and 5 months. The number of observations from year-area-month-stratum is given in Table 2.3.(2A) and (2B).

Three GLMs were fitted to the Icelandic cod data. The multiplicative models incorporated year (Y), area (A), and month (M), main effects and varying degrees of interaction effects to predict catch-per-unit effort (U).

1 U = (K) (Y) (A) (M) (E) main effects only

2 U = (K) (Y) (A) (M) (AxM) (YxA) (E) main effects, area-month and year-area interactions

3 U = (K) (Y) (A) (M) (AxM) (YxA) (YxM) (E) main effects, area-month, year-area, year-month interactions

where K is a constant and E is the lognormal error term.

Following Stefànsson (1988), a curvilinear relationship between effort and fishing mortality was incorporated. Specifically, catch per effort was computed as:

U = C / (T**1.29)

where C is the catch (in weight), T is trawling time (in minutes) and 1.29 is the estimated exponent for the North-Spring component of the stock.

The GLMs were fitted and using the estimated coefficients for the various effects, the predicted CPUE for each stratum (relative to that in area 101 during January 1974) was calculated as:

UHAT (y,a,m) = exp [b(y) + b(a) + b(m) + X(y,a,am)]

where b(y), b(a) and b(m) are the estimated parameters for the main effects and C is the sum of the b's for all of the interaction effects, as appropriate for the stratum and type of GLM used. The UHATs for all strata are provided in Tables 2.3 (3A, 3B and 3C) for GLMs (1), (2) and (3), respectively. The UHATs were then integrated over area using the arithmetic mean to calculate predicted CPUE by year and month [Tables 2.3 (4A, 4B and 4C)].

These UHATs can be regarded as indices of abundance. When no interactions are incorporated [GLM1 - Table 2.3 (4A)], the year effect, b(y), shows the same trend as that in any given month [(Figure 2.3.(1)]. When AxM and YxA interactions are included [GLM2 - Table 2.3. (4B)], the monthly indices show the same trend, but this trend is not the same as that obtained by simply using the year effect from the GLM without interaction terms [Figure 2.3.(2)].

When AxM, YxA and YxM interactions are included [GLM3 - Table 2.3.3.(4C)], the monthly indices become more variable, but still exhibit differences in trend from the simple year effect index obtained from the GLM without interaction terms [Figure 2.3.(3)]. An alternative presentation of these same data by month is provided in Tables 2.3.4a-c. Parameter estimates for all of the main effects and interaction terms are provided in Tables 2.3.3a-2.3.3c for GLMs (1), (2) and (3), respectively.

Age-size keys from Stefánsson (1988) were used to estimate an index of abundance for age 4 stock size (in number) from the exploited stock biomass index described above. This index has been compared with the Icelandic cod VPA results. Coefficients of variation were computed in two ways: (1) VPA stock numbers vs. age 4 index, and (2) log (VPA numbers) vs. log (age 4 index). (r^2) values are higher for the linear model and are generally higher for the GLMs with interaction terms.

These results indicate that the incorporation of year interaction terms can affect the trend in the estimated index of abundance from the GLM approach. However, when year interactions are incorporated, it is important not to use the year effect from such a GLM as the index of abundance, as it may have no relation to stock abundance. The procedure outlined here for the Icelandic cod GLM provides a mechanism for incorporating year interactions and extracting meaningful indices of abundance from the fitted model parameters.

In many cases, a single annual index may be preferred to various monthly indices. While it is straightforward procedurally to integrate over months in a fashion similar to integration over area (i.e., by using the arithmetic mean), this is particularly true for stocks with high fishing mortality rates, where the stock size is changing appreciably from month to month. In such cases, it may be more appropriate to select an appropriate month based on biological and sampling considerations. If the index is to be used for VPA tuning, it may be appropriate to select a month at the beginning of the year (if tuning is carried out using Jan. 1 stock sizes) or in the middle of the year (if mean stock sizes over the year are used in the tuning). The various indices can also be used individually for the tuning.

Care should also be taken in integrating over area. Generally, the integration should be done using an arithmetic mean, weighted by the surface area of each sub-area. While this is fairly straightforward in the case of research surveys, it can be more complicated when dealing with commercial fisheries that may not fish the entire sub-area each year.

It may be noted that if all possible interactions are included in a GLM, then a resulting index will be no different than a stratified arithmetic mean. This would have been the case, for example, if a three-way interaction (YxAxM) had been added to the GLM (3) model, above. In practice, however, it is seldom possible to include all interactions in a GLM because of the lack of observations in one or more strata.

In all applications of GLM, the correction for the log bias that result from estimating the parameters on the log scale and then transforming back to the original scale should be considered. The robustness question has not, however, been addressed. The correction would have little effect on the Icelandic cod GLM because the mean squared error was much larger than the variance of any of the individual parameters. However, this may not be the case generally and the bias correction could affect the estimated trend.

When year interactions are employed, the type of normalization used to implement the GLM may affect the parameter estimates (e.g., use of the row and column means, choice of standard cells, etc.). In practice, the magnitude of this effect should be examined as an integral part of GLM analysis.

In developing indices of abundance (whether by GLM or other methods) for the purpose of VPA tuning, every attempt should be made to estimate age-specific indices (in number of fish rather than weight). However, if from logbooks it is possible to obtain area-disaggregated results, in most cases, only landings integrated over a whole trip can be sampled to obtain length- or age-disaggregated results.

Following the methods suggested by Laurec and Pérodou (1987), it is theoretically possible to build and to fit models providing area and age (or length) specific parameters. The technique suggested for commercial categories disaggregation can, in fact, be easily extrapolated. Other models than the one they suggest can be imagined adapted to the available data. It is likely that they would not be considered as GLMs but other statistically and numerically efficient techniques now exist.

2.4 Interpolating

2.4.1 Introduction

It is assumed that the quantity being estimated by a survey (e.g., the abundance of age 1 cod in the North Sea) really exists (i.e., is not itself a random variable), but that it can be observed only approximately by the survey process. We observe the catch Z at location (x_i, y_i) where x and y are the positions on a plane. Assuming that catchability is constant spatially, as seems to be essential for any present available method (see Appendix D for an attempt to overcome this difficulty), survey results may be regarded as point estimates of the density of fish. An abundance index may be obtained by integrating under the surface so defined (see Shepherd, 1986).

For equally spaced and equally precise observations, such an integration is well-approximated by a simple arithmetic summation over the survey data (multiplied by a constant for the area occupied by a "typical" station).

However, this is not true if the stations are not uniformly distributed, or have varying precisions. Both these complications normally arise in practice. They may be overcome by forming an appropriate average over stations within a stratum, and multiplying by the stratum area before summation. The estimation of an "appropriate" average is, however, not trivial or obvious: the precision of individual observations, their location and intercorrelation are all relevant. Furthermore, if the strata are large, there may be systematic spatial variation within strata which are not properly treated by this process.

An alternative procedure is to use numerical integration under some representation of the surface described by the data. This can be done through trend surface analysis. Such an approach will be optimum where the random component added to the trend which, in practice, has to be simple (described as a model with few parameters), is a white nose (no spatial autocorrelation). If this is not the case, a direct interpolation technique can be more efficient.

Visual representations of such surfaces are routinely presented by contouring and 3-D surface plotting packages. They generally rely on an interpolation of the data onto a regular rectangular grid, although some techniques may avoid this step (Watson and Philips, 1985). A simple summation of these interpolated values is a suitable procedure for the integration required (ignoring the numerical constant relation to the rectangular size and some small edge effects).

Various algorithms may be used for the interpolation onto the regular grid. Two techniques have been investigated. The first one called Kriging can be based on a firm statistical background and provides, in addition to estimators of interpolated densities, an estimation of the corresponding variances. For both concerns (optimality and variance estimations), this requires statistical assumptions which may be far from being satisfied in practice. In such a case, it just becomes a special case of the weighted mean estimators and may remain reasonable even though it may not be optimal. The second interpolation method, due to Shepherd and Nicholson (1986), is an intuitive, empirical one, which appeared especially interesting to compare with Kriging since it is not a weighted mean technique.

Mapping surfaces over a region can have many uses. Here we limit the discussion to the utility of these approaches for generating an aggregate measure of biomass indices over the entire region of interest.

For further discussions about the mapping problem, one should refer to the relevant ICES Working Group (*Anon.*, 1989). However, as pointed out in Section 2.1.2, mapping should always be recommended as a first stage in the analysis of data.

2.4.2 Kriging

The value of Z at any point (x_i, y_i) on the surface can be thought of as being composed of 2 parts, the mean (trend) and the error (random component) (see Appendix F for a discussion of the "randomness" concept).

$$Z(x_i,y_i) = m(x_i,y_i) + E(x_i,y_i)$$

The mean component generates the trend surface. Though there is a technique referred to as "universal kriging" which accounts for both the trend and the error structure, we did not evaluate it. In practice, simple kriging can be applied to the residuals after some procedure has been used to remove the trend, or it must be assumed that there is no trend in the data. De-trended data can be used to determine the correlation in the value between spatially adjacent points at various distances. Although direction can also be considered, this added complication was not examined The correlation is summarized by a semi-variogram which is a plot of the expected suarred differences against spatial distance:

$$E[Z(x_i,y_i)-Z(x_j,y_j)]^2$$

is plotted against

$$[(x_i-x_j)^2 + (y_i-y_j)^2\}^{0.5}$$

The mathematical expectation $E\ (Z(x_i,y_i) = Z\ (x_j,y_j)^2$ not being known, the observed values $(Z\ (x_i,y_i)$ are plotted [see examples on Figure 2.4.(1)].

Such plots for the English groundfish surveys were examined. Log-transformed data were used. Semi-variograms of raw data displayed more scatter. The effects of interpolating grid points on the log scale, then transforming to the arithmetic scale for integration are not known precisely but it is hoped that any bias will be a multiplicative factor, constant from year to year.

Construction of "smooth" semi-variograms by averaging the value differences [Figure 2.4.(2)] over coarse spatial intervals gave results which indicated some pattern, whereas semi-variograms of individual observations display a large amount of scatter [Figure 2.4.(1)].

Examination of semi-variograms for 1977-1981 clearly shows a group of points with lower value differences at a small spatial distance (near origin). These are likely due to "replicate hauls". The average value difference for these is often referred to as the "nugget" effect and reflects replicate variance. In all semi-variograms, 1977-1988, value differences appeared to asymptote at a distance of about 40 nautical miles. Some semi-variograms appeared horizontal beyond 40 nautical miles while others displayed some increase over the entire range. Most showed greater increase beyond 200 nautical miles.

The increase in value difference is symptomatic of a trend and further work should have been preceded perhaps by some de-trending process. From these observations a single semi-variogram "model" was chosen to apply to all years. Using different variograms from to year would imply varying smoothing intensities from year to year, which could introduce varying biases from year to year. The selected model included a "nugget" effect and had an asymptote starting at 40 nautical miles. This can be interpreted to mean that errors are not correlated for pairs of points further than 40 nautical miles. Subsequent to 1981, almost all stations were further than 40 nautical miles apart, and it can be anticipated that the "krigged" results would not be much different from arithmetic means.

We proceeded nevertheless to interpolate the grid points using the selected variogram. The grid-point values are a weighted average of surrounding points with weights being determined as a function of the selected semi-variogram. Beyond a distance of 200 miles from grid point, observed values are not taken into account. This choice reduces the possible consequences of existing and neglected large-scale spatial trends. The aggregate measure was obtained by retransforming the grid points back to the arithmetic scale and integrating for the region. Most kriging partages, including the one we used (MAGIK), provide not only an interpolated value on each grid point, but also an estimation of the corresponding error variance. This makes it possible to attempt some correction for bias. The two approaches used to retransform were:

a) $[\exp(z)] + 0.1$
and
b) $[\exp(z - \sigma^2/2)] + 0.1$

Two series of abundance indices were generated from the kriging results and these were compared with VPA and stratified means.

We also examined the semi-variogram for the Icelandic groundfish survey. A strong increasing trend was evident to about 300 nautical miles and a rapid decline thereafter, possibly due to the geometry of the ground surveyed, i.e., toroidal [see Figure 2.4.(3)]. De-trending would be necessary before proceeding. The conclusions are as follows:

1 Careful examination of the spatial structure should be completed (e.g., semi-variogram) before proceeding. Frequently, these studies will guide subsequent steps or even determine whether it is worth proceeding further.

2 If these studies show little structure other than a nugget effect and a trend, then it can be anticipated that integration of interpolated results will not do better than simple techniques.

3 For age 1 cod from the English groundfish surveys, the standard index (arithmetic mean) could not be improved upon by the interpolation/integration methods.

2.4.3 Empirical interpolation

2.4.3.1 The method

The method is in fact a method of reverse interpolation: it determines those values on a rectangular grid which, when subsequently used as a basis for interpolation by a local bicubic spline, give estimates at the survey locations which are as close as possible (in a weighted least squares sense) to the observed value at those locations. This is a good method for contouring because it involves little or no non-essential smoothing of the data, so that individual observations maintain correct positions relative to the contour lines to the maximum possible extent. Other common algorithms (e.g., that of McLain, 1974) may involve more or less arbitrary smoothing, which is usually undesirable for contouring (though not necessarily for the present application).

Some minimal smoothing is always required, however, because the interpolating spline used here is strictly local, so that the effects of spurious or incorrect observations are localized in space, and without smoothing the values on the grid would be undefined in unsampled regions, where one usually wishes to simply interpolate in some smooth fashion. By increasing the value of a smoothing parameter (a weight attached to the sum of squares of deviations between adjacent grid-point values), progressively smoother and smoother representations of the data can be constructed if so desired.

By working with either logarithmically-transformed or untransformed data, one can allow for the presence of log-normally distributed (multiplicative, constant cv) or normally-distributed (additive, constant variance) errors. The former is usually more appropriate, and also ensures that negative values are not computed for what are usually non-negative quantities. Since a least squares calculation is carried out, it would in fact be very easy to carry out a weighted least squares calculation, and thereby allow for other error structures (e.g., that proposed by Shepherd and Nicholson, 1986), but this has not been implemented so far. For integration, the results are in any case dominated by the large estimates (it does not matter if small numbers are badly estimated), and for these the constant cv assumption implicit in the log transformation is, in any case, the most appropriate, and down-weighting the influence of small observations would have little effect.

Once the gridded values have been obtained, a simple weighted summation of them provides an index of abundance. The weights allow for the partial rectangles adjacent to the coast, and are zero for those on land! Note that retransformation (without a bias correction) is required before summation, since integration is fundamentally an additive process. The size/spacing of the rectangular grid is a matter of choice: it is usual to have roughly the same number of grid points as actual sample locations. If fewer grid points are used, some smoothing (on the scales of the grid size) will result automatically. Points which are very close together (relative to the grid spacing) will be averaged together in any case.

The method was applied, using a log transformation, to:

a) the EGFS age-1 cod data for 1977 to 1988, with smoothing parameter values of 0.01 (minimal), 0.1, 0.3 and 1.0, and to the age-2 data with a smoothing parameter of 0.1.

   Contour plots may of course be obtained as a bonus when a contouring package is used for the calculation, although this is not an essential part of the procedure. The results for the years 1978, 1982, 1983 and 1985 are displayed in Figures 2.4.(4), 2.4.(5) and 2.4.(6), and they clearly show the varying degree of smoothing applied. With minimal smoothing, and the 1 degree by 1 degree grid used here, the surfaces are highly structured and appear to be reflecting considerable sampling noise (though this is a subjective judgement only).

b) the Icelandic cod survey data.

2.4.3.2 Results

a - <u>English Groundfish Survey</u>

The index series obtained are listed (together with the VPA estimates and the standard index values)in Table 2.4.(1) and plotted in Figure 2.4.(7). A regression analysis of these results (using the usual RCRTINX2 program) is given in Table 2.4.(2). It can be seen that the "rough" (minimally smoothed) estimate is clearly unsatisfactory, having a very low correlation with VPA. This seems to be due to occasional rough values in the gridded estimates, almost certainly because the problem is ill-conditioned. This estimate is not, therefore, plotted or considered further.

The other three estimates ("bumpy", "medium" and "smooth") are highly correlated with one another and with VPA. The medium and smooth estimates are progressively and systematically smaller than the bumpy estimate, reflecting a gradual progression towards the weighted geometric mean of the data as the smoothing is increased.

Interestingly, and rather surprisingly, the least smoothed (bumpy) estimate has the highest correlation with VPA. Figure 2.4.(7) shows clearly that the reason for this is that the low values in the time series are exaggerated, giving these series a higher dynamic range than the VPA, and, therefore, a lower slope in the calibration regression analysis. This may be due to excessive weight being given to the small observations, which have not been downweighted, but this requires further investigation.

The best performance by the "bumpy" estimate is, however, only marginally better than that of the conventional stratified mean estimate. The RMS residual error is 0.155 for the full series of 11 data points compared with 0.184 for the standard estimate. This indicates that such a method can perform better than conventional methods, but the improvement may not be worth the extra effort. It remains to be seen whether greater improvements are achievable on less well-behaved data sets, or by improving the method by downweighting small observations.

The danger of a change in the smoothing intensity, when a transformation is used, appears very clearly in Table 2.4.(1) and Figure 2.4.(7).

For age 2, the method was applied using the low level of smoothing (0.1) only. The method performs well (r^2 = 0.89, residual standard error = 0.21), but is not better than the standard index which is equally good.

b - <u>Icelandic Cod Survey Data</u>

The data were interpolated using a smoothing parameter of 0.1 onto a 19 x 11 grid spanning the space 63° to 68°N latitude, and 10° to 28°W longitude. The results yield an index series of 1118, 607, 98, 76, 64 for the years 1985 to 1989. This can only be compared with the multiplicative model estimates of Stefánsson (1988), i.e., 1000, 1218, 453, 404, 467. The overall picture is clearly similar (two good years followed by three poor ones) but the estimates are not directly comparable, and no further conclusion can yet be drawn.

2.4.4 Discussion

The method described is of interest because it is firmly based on the concept of integration under an imperfectly observed surface, representing a spatial pattern. The cause of the spatial pattern is not considered (i.e., the influence of depth or other causal factors is ignored). The pattern is simply taken for granted as observed.

The advantages of this method (and others of the same type) are:

a) It can be applied to almost any data, regardless of whether the station positions are uniform or not, or whether they are fixed or re-randomized.

b) It allows appropriately for non-uniform station locations.

c) It involves the fitting of a model surface so transformations and/or weighting procedures can be applied to allow for the error structure of the data, and it can cope gracefully with missing data.

d) An explicit smoothing procedure can be used to trade off bias and variance in the estimate obtained, ultimately (presumably) in some optimal way.

e) No (possibly incorrect) assumptions about systematic pattern in the surface fitted (compare with the multiplicative model) or the complication of parametric form of the surface (compare with simple trend surface fitting).

For these reasons, it seems quite possible that this procedure could be useful operationally, even though its basis tends to the heuristic rather than the rigorous. It should certainly go some way towards reducing the sensitivity of the indices derived to isolated high observed values, and towards averaging out the noise in the data in a sensible way.

2.5 Sampling Scheme Design

2.5.1 Random and fixed survey

Ideally, survey data would be collected according to a pre-specified design with known statistical properties. In practice, this ideal may be compromised.

Data from surveys may often be used for more than one purpose. Estimates of overall catch per unit of effort (CPUE) or of abundance may be required for several species. The spatial distribution of these species may also be of interest. This may make it difficult to optimise the survey design for a single objective. Also, surveys may have evolved from some preliminary survey, with particular features maintained for historical continuity. There will also be practical constraints such as the limits of the cruise program and bad weather. Some parts of the area may also be inaccessible, for example where gear damage is likely, or where fishing is denied, e.g., in the vicinity of pipelines.

A working paper presented to the meeting (Working paper number 1) showed that obtaining unbiassed estimates of differences in year effects depends upon the interaction between simplicity and persistence and the survey design. A persistent spatial pattern always leads to unbiassed trend estimates, regardless of the use of randomized or fixed station survey design. A simple spatial pattern may be fitted annually and extrapolated over the total area to yield unbiassed year-to-year estimates. In the case of a changeable and complex spatial pattern, a fixed stations design will lead to biassed estimates except in exceptional and unlikely circumstances. A randomized design results in unbiassed estimates even in this case.

Clearly, given the advantages of a randomized design in extracting unbiassed trends, such a design is desirable. In practice, however, a strictly randomized design is generally not a workable option. The exigencies of cruise timing and the weather, together with out-of-bounds areas, mitigate against randomization in favour of a fixed station design. But if a fixed design is used, the analysis may be more complicated. If problems can appear under very specific conditions (Hoenig, pers.comm.), the experience of this Working Group suggests that a fixed station design does not lead to difficulties in analysis, provided that appropriate analytical techniques are utilized.

2.5.2 Systematic and simple random sampling

It can be shown that a systematic grid will generally lead to more accurate estimations of the total numbers over an area than a Simple Random Sampling (SRS) scheme. This is due to the fact that within SRS schemes, large sub-areas may be unsampled, the corresponding local density being difficult to estimate from observations located too far away. On the other hand, some sampled points may appear quite close to each other so that they bring partially redundant information. This can be interpreted in terms of autocorrelation functions or equivalently through variograms. If no spatial pattern appears (no autocorrelation for distances different from zero, or variogram just limited to a nugget effect) the sampling design has no influence. Within such a "white noise" context only the sample size matters.

On the other hand, if strong spatial patterns exist, SRS will be much less efficient than systematic sampling.

In the fisheries survey context, macroscale patterns (scale = hundreds of miles) are quite common. It thus appears dangerous to implement a sampling scheme where large sub-areas remain unsampled. This leads to some systematic division into blocks, based upon some *a priori* knowledge about the fish distribution, taking, for instance, depth into account. If, within each "block", a SRS scheme is applied, this corresponds to the basic stratification scheme.

Within a block, a scheme other than SRS can also be applied, for instance a more or less systematic one. However, the importance of such choices can often be limited. At a scale of a few tens of miles, the nugget effect will often be markedly dominating. In such a context, the redundancy associated with neighbouring stations is limited, even if it still exists.

It must be pointed out that sampling variances are easily calculated within a stratified scheme. More sophisticated calculations, technically possible, are required in other schemes. However, in the present context, this has not a major importance, since the sampling variances calculated through the usual formulas are not pertinent with respect to year-to-year changes (see Section 2.1.2 C).

2.6 Discussion

The correlations between the North Sea cod, age 1, log abundance indices given by the various methods, and the corresponding VPA results are given in Table 2.6.(1).

2.6.1 Pre-processing of data-conclusions

1 - None of the methods tested yields significantly better results than the standard arithmetic mean index for the North Sea cod data set (although several methods perform slightly better). Interpolation and GLM methods should be more robust to missing data (using fixed stations), to variable precision of the observations, and to non-uniform distributions of the stations and they warrant further investigation.

2 - No advantage was obtained when using statistically optimal interpolation methods (Kriging). Complex interpolation methods are required primarily for mapping purposes and may not be essential for the estimation of integrated indices.

3 - Interpolation methods (including Kriging) should be used with care on commercial CPUE data, since fishermen may congregate on high abundances, and an inappropriately large choice of scale parameter for the interpolation may spuriously extend this high abundance over too large an area.

4 - The GLM models appear particularly useful if auxiliary information is to be included or if missing data are to be filled in.

5 - The GLM and other models are very useful for understanding factors affecting the distribution and population dynamics of fish, e.g., density-dependent habitat utilization.

6 - The results of the GLM models depend upon the transformation and distributional assumptions. The results of the interpolation and Kriging methods also depend upon these factors.

7 - Interaction terms in GLMs should be estimated and examined. If the year effect is to be taken as an index of abundance, no interactions terms with year should be included in the model. It is preferable to calculate fitted values when interactions are included, and construct the index from these by interpolation and integration.

8 - When an interpolation technique on a stratified scheme is combined with a transformation, the procedure should be kept strictly constant from year to year.

9 - Applicability of methods

	Fixed stations	Random stations
Stratified mean	+ (biassed?)	+ (increased variance?)
GLM	+	stratum effects
Interpolation	+	+

2.6.2 Need for further studies

Theoretical studies

Special attention should be paid within interpolation techniques to the fact that series of yearly surveys are available. This could help the estimation of possible spatial trends but it is also a specific development of variance estimation techniques. The cooperation of qualified statisticians would be necessary.

Numerical calculation

It seems almost certain that the performance of the various methods depends considerably on the nature of the data, and the Working Group was able to carry out reasonable comparative tests on only one data set, which is of relatively good quality, and most likely untypical in other respects.

The Working Group, therefore, recommends that more thorough comparative testing of the four main methods [stratified averages, simple interpolation, Kriging and GLM (multiplicative) models] should be carried out on a set of about four contrasting real data sets for which VPA calibration data exist.

These should probably include a random stratified data set (from the NW Atlantic, perhaps), at least one data set known to be troublesome (e.g., IYFS age 1 cod), one with a simple persistent structure (e.g., EGFS age 2 haddock) and one for a pelagic stock (IYFS 1-ring herring, perhaps).

These analyses could be carried out in national institutes, but it would be desirable if the results could be discussed and evaluated by the assessors and other interested parties, perhaps at a small, *ad hoc* working party meeting, with the aim of producing a report for consideration at the next full meeting of the Working Group in 1991.

3 STOCK ASSESSMENT

3.1 Workshop in Reykjavik

3.1.1 General comments

Results of the 1988 Workshop on Methods of Fish Stock Assessments (*Anon.*, 1988) were presented by G. Stefànsson and discussed by the Working Group.

It was felt that, although the Workshop was successful, it was not possible to recommend a single assessment method for universal application. On the basis of the limited simulations, however, certain methods do seem more promising for future application and research. These are discussed below.

It was further noted that the objectives of the Reykjavik meeting, namely to compare methods for the assessment of fishing mortality and population size, are only a part of the more general problem of comparing estimation and prediction methods. It would also be more useful to include the prediction models in the calculations so as to evaluate the impact of uncertainties in short- and long-term yield, population trends and catch trends under different management strategies.

3.1.2 Comparisons of methods

The Working Group noted that comparisons based on mean square errors could lead to different conclusions, depending on which tables were used. In fact, the tables giving ranks to methods (based on average F at ages 5-9) are at first sight inconsistent with tables giving individual sum of squares for each age group. This is primarily because the summary tables emphasize certain features of the results, but also because of the bias inherent in computation of averages of lognormally- distributed quantities and the cancellation of biases averaging.

It was, therefore, suggested that the histograms in the report should also be used for comparing the methods. The histograms of primary interest are given in Tables 3.45, 3.48 and 3.56 and 3.59 in the Reykjavik report. The Working Group acknowledges that these histograms should be treated with care and they do not give absolute reasons for accepting or rejecting methods, but they are felt to give strong indications as to which methods are worthy of further study. It should be noted that the ADAPT and TSER2 methods of analysis are quite promising, but histograms are not available for all data sets for these methods. Further, the time series method TSER1 did not use ages 3 or 12.

Based on the above tables, it is obvious that all methods tend to fail quite badly several times during the 10 runs performed. "Quite badly" can be defined as yielding an F-at-age which over- or under-estimates the true F-at-age by more than 50%. One interpretation of this result is that when advice is given on 10 stocks in a given year, one can expect to get quite wild results for at least one age group in one of the stocks. It was pointed out, however, that most current methods of prediction (*Status quo* TAC, constant or fixed percentage increase in biomass, etc.) are quite robust to such errors, as positive errors in F-at-age are often accompanied by corresponding negative errors in stock size, so the errors will cancel to some extent when prediction is performed.

A "minimax" index for analyzing these histograms was calculated as follows: a simple count of how often a method yields estimates more than 50% away from the true (counted across a relevant age range) is an indicator of the "badness" of a method and gives a simple non-parametric ranking, which is not affected by the biases involved in averaging. Since the most important age groups in the catches are ages 4-7, these are used in the following text table. The table thus gives the number of estimates which are 50% or more above or below the true values.

	Data set 5		Data set 6	
Method	N	F	N	F
Hybrid	8	15	15	16
LS	13	15	12	8
AC1	17	15	6	16
AC2	18	16	9	20
AC3	14	14	20	16
AC4	17	18	20	16
AEFM	16	17	20	22
CCPUE	10	13	11	13
SURVIV	12	18	10	16
XSA	6	8	5	5
CAGEAN	5	4	4	5
GLM	7	7	7	8
TSER1	8	8	3	6

The actual computation of this index depends somewhat on the age classes and percentages used. The following table is based on the same data sets, but shows the number of occurrences of estimates that deviate more than 70% from the true values, summed over the full age range, 3-12.

	Data set 5		Data set 6	
Method	N	F	N	F
Hybrid	10	17	16	18
LS	11	8	6	7
AC1	9	15	3	21
AC2	11	16	3	25
AC3	13	9	36	24
AC4	15	13	34	29
AEFM	16	32	23	40
CCPUE	11	24	14	21
SURVIV	3	25	5	28
XSA	5	19	3	6
CAGEAN	2	3	8	7
GLM	10	23	10	16
TSER1	-	-	-	-

3.1.3 Conclusions

From the above tables one may conclude that some of the methods would give unreliable results for stocks similar in structure to the ones in data sets 5 and 6. It is further observed that four methods emerge as data sets 5 and 6. It is further observed that four methods emerge as having the greatest potential in this setting, for estimating ages 4-7, which are the most prominent ages in the catch, namely XSA, CAGEAN, GLM and TTSER1. Integrated methods dominate this group. Further studies are needed to firmly establish the range of circumstances in which each method performs satisfactorily.

It should, however, be noted that the TSER1 variant of the Time Series method cannot strictly be regarded as a tuning or integrated method, since it makes no use of abundance indices, and, therefore, has no chance of detecting a sharp change of fishing mortality in the last year, which is a principal goal of tuning procedures. It is in fact likely to underestimate changes, including trends, in fishing mortality, unless a trend parameter is estimated, in which case the solution is likely to be ill-determined (Gudmundsson 1987a, p. 17, and 1987b pp.13 and 21).

Conversely, the inclusion of a prior assumption of modest changes of fishing mortality has the potential to improve the stability of the estimates considerably, and may well be desirable when auxiliary information is included as in the TSER2 variant.

When all age-groups are considered, the minimax index (based on 70% error) suggests that some methods work better for estimating F while others work better for N. Some methods worked well with one data set but not as well with the other. While CAGEAN did reasonably well in data sets 5 and 6, it showed a systematic bias in data set 4 with most estimates of N and F showing a systematic percentage deviation greater than 50% from the true values. No analysis of the simulation results has indicated that a single method outperforms the others for estimating F or N on all data sets studied.

While the Reykjavik meeting did not indicate a single method, there are a number of general observations that can be made from the results of the simulation.

1. Quality of data

The "quality" of the data is an important element in any assessment. As each method relates catch rates to biomass in some way, the stock size or fishing mortality estimates will be poor if the relationship cannot be measured with precision.

All methods are sensitive to errors in the data, but to varying extents. In particular, VPA-based techniques (i.e., *ad hoc* tuning, XSA and Survivors) treat catch-at-age data as exact and would be expected to work well only when the errors in it are small compared with those of CPUE data. In addition, *ad hoc* tuning methods are particularly sensitive to errors in CPUE in the final year, and if the CPUE data are of poor quality, an integrated method should be preferred. The position may be summarized in the table below, which indicates which class of method is appropriate in various situations.

		CATCH DATA	
		good quality	poor quality
CPUE data	good quality	any	integrated
	poor quality	Survivors or Integrated	None

In particular, it should be noted that successful application of *ad hoc* tuning methods requires the availability of good quality catch-at-age data, and at least one good quality CPUE index series. Even integrated methods would be degraded unless adequate catch and effort data are available.

2. Convergence

All methods depend upon the convergence of Virtual Population Analysis. The lower the fishing mortality in recent years, the weaker this convergence. Typically, weaker convergence of the VPA will translate into larger variances of stock size estimates for the last years. Very weak convergence of the VPA generally leads to a situation where the catchability (q) and terminal fishing mortality (or survivors) cannot be estimated simultaneously. In some integrated methods, this will translate into very high correlations between parameter estimates and the correlation matrix of parameters can be used to detect that such a problem exists. *Ad hoc* methods will tend to take more iterations to provide an answer and the resulting estimates may be ill-determined.

It should also be stressed that all available methods require some restrictive assumption about either fishing mortality or catchability (or selection pattern) on at least one fleet/survey for the oldest age group(s). This usually takes the form of assumed constancy of either the catchability or the total exploitation pattern on the oldest ages, but this assumption cannot usually be tested or verified. Such an assumption is, however, essential. It is, furthermore, dangerous to assume low F or reduced catchability on the oldest ages, since this may allow the method of analysis to converge to the ever-present trivial solution of the problem, with zero (or very small) fishing mortalities everywhere.

3. Parsimony in model specification and robustness

The methods which appear to perform reasonably well over most data sets are methods for which the number of parameters is limited. Parsimonious systems leave less possibility for the parameters to become aliased (structurally). The Working Group noted, however, that the observed robustness of certain models may be related to the structure or complexity of the data set studied. For instance, a method like Laurec-Shepherd performed reasonably well because all the data sets were based on a "multiple fleets" scenario where a relatively high number of indices is available. When the number of indices is reduced to one or two (as is often the case for stocks in the NW Atlantic or elsewhere), estimates from *ad hoc* tuning methods will tend to be strongly influenced by the uncertainties associated with the actual "realization" of the index in the last year.

Other methods, such as SURVIVORS, XSA, CAGEAN or ADAPT, provide estimates which are less dependent upon the realization of the index in the last year and would be expected to be more robust, particularly when the number of indices is limited. In short, while there may be little benefit to move beyond an *ad hoc* approach when several indices of good quality are available, the use of an integrated approach is very desirable in a situation where only one or two indices of moderate quality are available.

This explains, in essence, the difference between the approach used in the North East Atlantic (where multiple indices are often available) and that used in the NW Atlantic (where only one or a few indices of abundance are available). When only a single index is available, it is preferable to move to an integrated approach which treats the data more efficiently (by using information from all years), especially if its precision is less than excellent. In any case, it is desirable to use a method which provides diagnostic tools to evaluate the "quality" of the estimates or the degrees of deviation from underlying assumptions.

3.2 Directions in Further Development of Current Methods

3.2.1 Further development of current methods

As the Reykjavik Workshop indicated, it is felt that the more complex methods cannot realistically be used in their current version within the ICES assessment Working Group framework. Some statistical expertise is required and extensive analysis of residuals is needed to ascertain the specific models to be fitted.

The Working Group, therefore, suggests that these methods be augmented by adding program modules which guide the user in selecting initial parameter values and in specifying the underlying model. In some cases it may be possible to (partially) automate the model selection, although it is realized that an automatic model selection is usually inferior to one based on a careful scrutiny of output, including model residuals. However, automatic model selection is certainly preferable to the situation where a user chooses an incorrect model and does not realize this.

It is noted that the addition of such program modules is a step in the direction of expert systems, where as much information as possible is included in the programs, yet allowing the users to intervene at any stage.

3.2.2 Data sets for testing new methods

Data sets 5 and 6, as described in the Reykjavik report, are available on diskettes. The Working Group strongly recommends that new methods of fish stock assessments be initially tested on these data sets, using the same procedures as used by the Reykjavik Workshop.

It should be noted that in these two data sets, noise added to the data was such that it had mean zero on an untransformed scale (in data sets 1-4, however, no variance correction was used when adding noise, resulting in generated values which were unbiased on a log-scale).

Given the high variance of the estimates for individual ages noted above, however, it is not surprising that this is not noticeable in the table of residuals. As pointed out by Laurec and Perodou (1987), when the variance is high enough for the bias correction to be significant, the total prediction error is dominated by the variance rather than the bias.

3.2.3 Future techniques for testing methods

A given method may perform quite poorly in terms of estimating stock size, but still provide a good estimate of *status quo* TAC. A natural next step in testing methods is, therefore, to add the TAC computations to each method tested. Such computations must include some of the common management strategies, such as *status quo* TAC, TAC based on a target level of fishing mortality and a target spawning biomass.

Within such a framework, more methods can be tested, including stock production models.

A further development is to consider the longer-term trends in catches and stock sizes, given a particular management strategy and a particular TAC estimation method. Thus, for example, it is quite feasible to examine how a management strategy based on *status quo* TAC computed from the LS method will perform in a 15-year period. Items of interest include the total catches in the period, year-to-year variation of catches, minimum stock size in the period and terminal stock size. It should be noted, however, that the only methods which are testable in this framework are those which can be made fully automatic, since user intervention is not possible when a large number of simulations is performed.

3.3 Incorporation of Recruitment Information

Some of the methods will allow inclusion of recruitment indices. These can be very valuable information, especially when the level of fishing mortality is high. The Working Group noted that recruitment estimation in some cases involves somewhat more complex procedures such as shrinkage toward the mean and estimations of non-linear relationships. These features are usually not incorporated in the stock estimation methods and must, therefore, be used outside the models.

However, in some cases where such anomalies do not seem to occur, it is the recommendation of the Working Group that tuning programs or integrated methods be used to incorporate this information, as this will provide consistency in the overall estimation procedure. For most of the methods, it would in any case be feasible to include these additional features in the tuning procedures and this would be desirable.

3.4 Dangers of Tuning Methods

Each method may fail if there are violations of the model's assumptions or if the user is inexperienced. Table 3.4.(1) gives a brief summary of the dangers of each of the methods described in the 1988 Methods Workshop Report under the available implementation. The methods are divided into the categories that are largely self-explanatory. The intermediate grouping describes the survivors and extended survivors methods, which are viewed as intermediate between the *ad hoc* and fully integrated methods.

This table should be read in conjunction with the report of the Reykjavik meeting.

3.5 Logistic Considerations

The Working Group has identified a need to move towards implementing more sophisticated methods when the available data are of poor quality, which is regrettably a common situation. In addition, it is apparent that many Working Groups already have difficulty in using existing relatively straightforward techniques.

The process of transmitting the necessary expertise to several members of all assessment Working Groups is already proving to be slow, expensive and inefficient. The Working Group considers that it would be highly desirable and appropriate for the ICES Secretariat to provide much more statistical guidance and assistance to assessment working groups and to ACFM than is possible at present.

This could probably be achieved by providing the ICES Statistician with (say) two additional staff with appropriate qualifications and experience, to provide a sustained source of expert advice, and to oversee the continuing development and improvement of the assessment tools available at ICES and the Working Group recommends that this be considered. It is not a major financial issue in the light of the great importance and central role of ICES in providing assessment advice, the existing staffing levels devoted to other activities, or the total investment in manpower and other costs to the assessment process.

4 IMPLICATIONS OF TIMING OF WORKING GROUP ADVICE AND CHANGE IN THE TAC YEAR

The timing of assessment working group meetings is closely linked to the timing of the ACFM meetings. Both managers and scientists want the most recent information from surveys and commercial fisheries to be included in the assessments before advice on TAC is given. However, this has the disadvantage that ACFM is given short time to consider some of the working group reports and the managers are given short time to consider the advice and carry out the framework of negotiations needed before the final TAC can be agreed.

The terms of reference to this Working Group reflect two possible solutions to this problem. One is to change the TAC year. The other is to ignore the more recent information, which might allow both working groups and ACFM to meet earlier in the year.

The Working Group observed that for a number of assessments, the data base only allows predictions to be made on a yearly basis and that a change in this procedure would require several years of notice. A change of TAC year, therefore, means that the TAC advice will be applied to a different period than that for which the calculations were made. It is also assumed that the TAC year must be the same for all stocks.

4.1 Change in the TAC Year

TAC years beginning during the first half of a year would be based on final advice given by ACFM in November with an appropriate lag. TAC years beginning in the second half of a year would be based on final advice given by ACFM in May. For example, advice given by ACFM in November 1990 for 1991 would be implemented in February or April 1991 while advice given by ACFM in May 1990 would be implemented in either July or September 1990.

The following scenarios for a TAC year were considered:

	TAC year	ACFM meetings	Working Groups
a	1 Jan-31 Dec	May, Nov	As present
b	1 Feb-31 Jan	May, Nov	As present
c	1 Apr-31 Mar	May, Dec	As present
d	1 Jul-30 Jun	Nov, May	Some moved to spring
e	1 Sep-31 Aug	Nov, May	Some moved to spring

The implications for the different scenarios are:

a. The present situation;

b. One month more for negotiations. TAC implemented one month later. Minor consequences for management;

c. One month more for ACFM. TAC implemented three months later, which means a corresponding time-lag in correcting for new advice next year;

d. More time for ACFM, as present for negotiations. Lower precision in advice for some stocks may be balanced by implementing the TAC six months earlier;

e. More time for ACFM, two months more for negotiations. Lower precision in advice for some stocks to some extent balanced by implementing the TAC four months earlier. Little new information available for most stocks between May and September.

It should be noted that a change in the TAC year may create problems for management of stocks with seasonal fisheries, especially if these coincide with the period when the TAC year starts.

It should also be noted that if the TAC year is changed, there will be a transitional period which may cause some practical problems.

A working paper (Working doc. No.11) was presented, estimating the precision of the advice for North Sea cod, haddock and whiting for TAC years starting at the beginning of each quarter. The exercise is based on quarterly data and assumes that the TAC advice given corresponds to the TAC year. It is also assumed that the most recent information is used in each case. It should be noted that the exercise is not considering all aspects of the problem. The calculated coefficients of variance are given in Table 4.(1). The results indicate that the most precise advice is given for the present TAC year for cod and haddock, whereas the TAC year 1 July - 30 June might be best for whiting. The differences between TAC years are, however, small. The results also show that the advice for cod in general is more precise than for haddock and whiting.

4.2 Timing of Working Group Advice

In principle, the working group assessments should be the basis for the advice given by ACFM. If there is a considerable lag between the assessment and the implementation of the advice, the most recent information will not be used in framing the advice. The Working Group could not assess the effect on the reliability of advice for all the different stocks, but in some cases, especially when recruitment indices are important, increasing the time between the advice and its implementation could severely impact the reliability of the advice. Even in the cases where the cost to precision may be tolerable, it will in practice be very difficult to avoid pressure to take into account new information if it is thought that this may significantly change the basis for the advice.

4.3 Advice

Apart from a possible improvement in the working conditions for ACFM and managers, there appears to be no argument for changing the TAC year. Because of the demand for using the most recent information, it is also unlikely that working conditions will be much improved. A change in the TAC year may create a lot of confusion and it is not known how management of the different stocks will be affected. The Working Group, therefore, does not advise a change in the TAC year. If a change is seriously considered, a thorough investigation of the implications for the different stocks is needed.

As long as there is a demand for using the most recent information, moving working groups to earlier periods means that ACFM more often will have to carry out revised assessments. Also moving ACFM to an earlier time will create a situation where there is no one to give advice on the basis of new information. The Working Group considers that the present situation is preferable and advises that the need for precision in the advice be the most important factor in deciding the timing of working group meetings.

The importance of the most recent information to a large extent reflects the fact that many stocks are heavily exploited and the catches are accordingly highly dependent on recruiting year classes. A reduction in fishing mortality will reduce this dependence and, therefore, the need for using the most recent recruitment indices. It will also improve the precision of the advice and give more stable yields.

5 CONCLUSIONS

5.1 Advice

- Careful analysis of the basic data, including mapping and the examination of residuals, should be conducted systematically when a yearly index of abundance is constructed.

- Whenever a transformation is used for constructing a yearly abundance index, the stratified sampling scheme in the interpolation procedure must be kept strictly constant from year to year.

- When a multiplicative model is used, if any interaction involving the years is included, the year effect should <u>not</u> be used as an estimate of yearly abundance. The fitted model should be used as an interpolating method, preceding an integration over space.

- Simple techniques may be almost as efficient as sophisticated ones, which will never create information which does not exist in the basic data, and can be misleading for unexperienced users.

- Comparisons of the various techniques used for constructing indices of abundance should be developed on real data, these indices being checked whenever it is possible against VPA output.

- Special attention should be paid to the development of processing techniques adapted to commercial disaggregated data, especially those from logbooks.

- No single tuning technique can be universally recommended at present. All methods require hypotheses that must be systematically checked. When the catch-at-age and the CPUE data are poor, no method can be really efficient. When both are of good quality the standard *ad hoc* tuning method (Laurec/Shepherd) performs satisfactorily. When all the CPUE data are of poor quality, integrated techniques may be necessary. They, however, require considerable statistical expertise and are not necessarily accessible at ICES.

- The Working Group therefore <u>recommends</u> that the ICES Statistician be provided with two additional staff with appropriate qualifications and experience to provide a sustained source of expert advice and to oversee the development and improvement of assessment tools available at ICES.

- The Group does not recommend at present any change in the TAC year.

5.2 Future Works

5.2.1 Strategical considerations

The Working Group considers that methodological research must be developed along various directions:

- Assessment techniques still need to be improved. The tuning problems still need more research so that the integrated techniques, potentially more powerful than *ad hoc* methods, are brought to a stage where the conditions for safe use are clearly guaranteed. This implies the development of diagnostics about the validity of the underlying hypothesis.

Among the basic assumptions common to all techniques, the dynamic pool hypothesis, neglecting spatial structure and migrations is a crucial one which would deserve special attention. The differences between sexes should also be taken into account. If integrated techniques are to be developed, in order to be used within assessment working groups, it appears that the existing software, which is highly interactive, should be replaced as many procedures as possible automated. An expert system approach may be fruitful and should be explored.

The methodological work should not be limited to assessment techniques deriving fishing mortalities and numbers-at-age estimates from given catch-at-age matrices, effort and apparent abundance series. The pre-processing question addressed during this meeting is still open. It is even necessary to go further backwards to the first step of the data collection. Sampling for ages, lengths and sexes requires a comprehensive examination due to the labour cost implied, and to influence on the accuracy of the derived catch-at-age matrices. The design, in space and time, of research cruises also needs to be optimized.

It is necessary to go beyond the estimation of numbers and fishing mortalities, in order to consider these estimation questions as part of a much wider problem. The analysis of the results obtained at the Workshop in Reykjavik shows the difficulties of building comprehensive measurements of the performance of the various tuning methods. Ideally, they should be considered as parts of a decision-making procedure in the context of decision under uncertainty. For a given fishery management strategy, considering well identified biological, economical (or political) targets, it would theoretically be possible to judge the relative performances of the various methods.

- Consideration of all questions, from sampling optimization and research surveys design, to management strategies is theoretically conceivable, but it corresponds to a very broad field.

The Working Group considers that it should contribute to its exploration in a realistic way, avoiding a dilution of its activity within too large a domain. It appears desirable to go one step further than fishing mortalities and numbers at age in order to evaluate the reliability of the various forecasts based on them, with special attention being paid to prognosis associated with classical management strategies (F_{max}, $F_{status\ quo}$...).

- Finally, since simulated data sets appear to play a very important part in the comparison of the various techniques, special attention should be paid to their construction.

5.2.2 Operational considerations

The previously-mentioned methodological developments cannot be covered by the members of the Working Group during their meetings.

A topic such as biological sampling does not appear mature enough yet to allow useful discussions within the Working Group. A symposium could be organized by the Statistics Committee in 1991, promoting scientific research in this field. From this symposium a limited number of key questions could appear that could be usefully addressed later on by the Methods Working Group.

It also appears important to take into account the work conducted outside ICES, to promote exchanges with these external organisations, and to take into account their plans. A special NAFO meeting in September 1990 will deal with "Management Under Uncertainty" covering simultaneously tuning techniques and management strategies. Several members of the Methods Working Group are going to participate. The conclusions of this meeting will be available before the terms of reference for the next meeting of the Methods Working Group are determined. They should be taken into account.

The number of items that the Working Group must deal with during a meeting is also very important. Too many items make it difficult to conduct calculations that may be necessary and can lead to the creation of various subgroups meeting simultaneously in a way which may be difficult to coordinate. On the other hand, a single precise topic can make it difficult to gather scientists from various countries within a forum in a way the Working Group has successfully managed to do up to now. For this latter reason, it appears necessary to have several topics for the future meetings. It also appears, once a precise question has been identified requiring a large amount of computer work, that a special workshop such as the Reykjavik one is extremely useful. In fact, the sequence of the Working Group meeting, identifying the need for a workshop, making intensive calculations during this workshop, and then reconsidering the workshop results after some months, appears very efficient. For the future, the combination of "plenary" meetings of the Working Group covering various items, and special workshops each devoted to a single technical problem, should be considered. The attention of potential participants to the special workshop should, however, be drawn to the fact that their participation will be useful, for both themselves and the workshop, if they are prepared to participate actively in the calculations.

5.2.3 Next meeting(s)

The Methods Working Group should meet in 1991 preferably before the Statutory Meeting and consider basically three items:

a) the influence of spatial structures, including migrations, on tuning techniques;

b) the accuracy of the prognosis derived from assessments based upon tuning techniques and corresponding to the classical management options (it should be evaluated on simulated and real data sets);

c) the validation or otherwise of the hypotheses upon which the various tuning techniques are based (use of diagnostics, etc.).

Depending on the conclusions of the NAFO 1990 Special Meeting, item b) could be modified.

It is hoped that prior to the next meeting of the Working Group, work will continue on the preprocessing of real survey and commercial data sets, including, if possible, an *ad hoc* working party (see Section 2.6.2). A discussion of the preliminary conclusions of such a workshop should be considered as a possible extra item for the next meeting of the Methods Working Group.

6. REFERENCES

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. 1987. Report of the Working Group on Methods of Fish Stock Assessments. ICES, Doc. C.M. 1987/Assess:24, 107 pp.

Anon. 1988. Report of the Workshop on Methods of Fish Stock Assessment, Reykjavik, July 1988. ICES, Doc. C.M.1988/Assess:26, 117 pp.

Anon, 1989. Report of the Workshop on Spatial Statistical Techniques. ICES, Doc. C.M.1989/K:3.

Armstrong, M., Renard, D., and Bertou, P. 1989. Applying geostatistics to the estimation of population of bivalves. ICES, Doc. C.M.1989/K:37.

Bretherton, F.P., Davis, R.E., and Fandry, C.B. 1976. A technique for objective analysis and design of oceanographic experiments applied to MODE-73. Deep-Sea Res., 23:559-582

Casella, G. 1985. An Introduction to Empirical Bayes Data Analysis. J. Amer. Statis. Ass. 39(2):83-87.

Cochran, W.G. 1977. Sampling techniques 3rd edition. Wiley and Son.

Davis, R.E. 1985. Objective mapping by least squares fitting. J. Geophys. Res. 90: 4773-4777.

Gavaris, S. 1980. Use of multiplicative model to estimate catch rate and effort from commercial data. Can. Journ. Fish. Aquat. Sci. 37:2272-22.

Gavaris, S., and Smith, S.J. 1988. Effect of allocation and stratification strategies on precision of survey abundance estimates for Atlantic cod (*Gadus movhua*) on the eastern Scotian Shelf. J. Northw. Atl. Fish Sci. Vol. 7:137-144.

Gudmundsson, G. 1987a. Time series models of fishing mortality rates. Univ. of Iceland, Science Institute RH-02-87, 36 p (mimeo).

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

Harding, D., Woolner, L., and Dann, J. 1986. The English Groundfish Surveys in the North Sea, 1977-1985. ICES, Doc. C.M.1986/G:13.

Harding, D., and Macer, C.T. 1986. The English Groundfish Survey of the North Sea. W.P. North Sea Roundfish Working Group, 1986, 9 pp.

Hardy, P.L. 1984. Universal kriging, least squares collocation and optimum interpolation (what's the difference). MS, 16th European Meeting of Statisticians, University of Marburg, Germany 3-7 September 1984, 10 pp.

Heimbuch, G.G., and Hoenig, J.M. 1989. Change in ratio estimators for habitat usage and relative population size. Biometrics: 45:439-451.

Houghton, R.G. 1987. The consistency of the spatial distributions of young gadoids with time. ICES, Doc. C.M.1987/D:15, 7 pp.

Hunton, J.K. 1986. Areal and temporal sources of variation in the English North Sea Groundfish Survey. ICES, Doc. C.M.1986/G:15, 7 pp.

Laurec, A., et Fonteneau, A. 1978. Estimation de l'abundance d'une classe d'âge. Utilisation des CPUE de plusieurs engins en différentes zones et saisons. Rec. Doc. Scient. CICTA, VIII(1): 79-100.

Laurec, A., et Pérodou, J.B. Considérations statistiques et informatiques sur l'analyse des puissances de pêche et des abondances apparentes. CIEM, Doc. C.M.1987/D:9.

Laurec, A., et Souplet, A. 1987. Calibration du recrutement par plusieurs indices d'abondance. Maximum de vraisemblance. Incorporation d'une distribution à priori. CIEM, Doc. C.M.1987/D:13.

Laurent, A.G. 1963. Lognormal distribution and the translation methods description and estimation problems. J. Amer. Statis. Ass., 58:231-235.

McCullagh, P., and Nedler, J.A. 1983. Generalized linear models. London: Chapman and Hull.

McLain, D.H. 1974. Drawing contours from arbitrary data points; Comp. J. 17(4):318-324.

Myers, R.A., and Pepin, P. 1986. The robustness of lognormal based estimators. Biometrics.

Myers, R.A., and Pepin, P. 1987. The robustness of lognormal based estimators of abundance. ICES, Doc. C.M.1986/D:2.

Nicholson, M.D. 1986. (pers comm).

Shepherd, J.G. 1986. *In*: Report of Theme Session on Design and Analysis of Trawl Surveys, 1986 Statutory Meeting. Proc. Verbaux de la Réunion, Cons. Int. Exp. Mer (1986), p. 147.

Shepherd, J.G. 1987. Combination of recruitment indices using weighted averages. W.G. ICES Methods Working Group, 1987, 4 pp.

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 35:221-227.

Stefánsson, G. 1988. A statistical analysis of Icelandic trawler reports. ICES, Doc. C.M.1988/D:13, Ref.G.

Topping, J. 1962. Errors of observation and their treatment (3rd edition). Chapman and Hall (London) for the Institute of Physics, 119 pp.

Watson, D.F., and Philips, G.M. 1985. A retirement of inverse distance weighted interpolation. Geoprocessing 2:315-327.

Table 2.2.1 Results obtained from the various stratification schemes and correlation (r^2) with VPA results.

	Untransformed				Transformed, no correction				Transformed correction included				VPA
No. of strata	1	3	8	15	1	3	8	15	1	3	8	15	
77	8.94	8.95	8.90	9.15	2.32	5.03	5.83	6.48	5.14	6.29	7.67	7.99	6.59
78	7.98	8.02	8.00	7.99	2.04	3.25	4.35	5.45	3.86	4.61	5.60	6.39	6.05
79	8.10	8.15	8.10	8.21	1.66	3.00	4.14	5.37	4.42	5.03	6.22	6.54	6.11
80	8.87	8.97	8.97	8.98	2.24	4.18	5.55	6.45	5.43	5.41	6.16	6.73	6.68
81	7.33	7.45	7.32	7.56	.02	1.11	2.32	3.44	2.42	3.76	5.09	6.42	5.61
82	8.33	8.38	8.41	8.46	2.10	3.32	4.42	5.75	4.39	5.17	6.28	7.39	6.32
83	7.58	7.56	7.54	7.56	1.01	2.10	3.33	4.61	3.88	4.58	5.62	6.34	5.60
84	9.01	8.97	8.97	8.97	3.14	4.34	5.49	6.42	5.18	5.85	6.91	7.96	6.29
85	6.33	6.35	6.29	6.26	-.51	.45	1.50	2.89	1.64	2.50	3.56	4.43	4.65
86	8.39	8.39	8.42	8.41	1.67	3.22	4.63	5.52	4.88	5.48	6.53	8.54	6.36
87	7.51	7.48	7.48	7.44	1.45	2.42	3.52	4.54	3.26	3.91	5.06	5.79	5.49
88	7.00	6.95	6.94	6.86	.65	2.20	3.36	3.94	2.64	4.51	4.97	5.79	5.06
r^2	.93	.95	.95	.96	.72	.78	.82	.87	.89	.78	.81	.71	

Table 2.3.1 Summary of LN(indices) and r^2 values of index *vs.* VPA estimates.

A) 1 YEAR OLDS

YEAR	VPA	GM	AR	SI	CCP	CCL (0.1)	CPP	CPL (0.1)	CPL (0.25)	CPL (0.5)	CPL (1.0)
1977	6.59	2.33	4.04	4.14	7.91	7.77	1.32	.51	.77	1.01	1.29
1978	6.05	1.90	3.10	3.13	6.93	6.79	.37	.18	.45	.68	.96
1979	6.11	1.65	3.16	3.19	6.91	6.79	.46	-.17	.16	.45	.78
1980	6.68	2.45	3.96	3.93	7.81	7.74	1.32	.53	.80	1.04	1.34
1981	5.61	.23	2.41	2.43	6.25	5.72	-.26	-1.43	-.97	-.57	-.11
1982	6.32	1.90		3.48							
1983	5.60	.30		2.73							
1984	6.29	2.43		4.11							
1985	4.65	-1.59		1.46							
1986	6.36	1.86		3.54							
1987	5.49	1.01		2.65							
1988	5.06	-.03		2.13							

R-SQUARES ON VPA SERIES

1977 TO 1981		.93	.99	.98	.99	.99	1.00	.91	.92	.93	.94

B) 2 YEAR OLDS

YEAR	VPA	GM	AR	SI	CCP	CCL (0.1)	CPP	CPL (0.1)	CPL (0.25)	CPL (0.5)	CPL (1.0)
1977	4.44	.16	1.63	1.50	5.51	5.13	-2.44	-2.35	-1.59	-.94	-.26
1978	5.66	1.30	2.36	2.53	6.18	5.98	-1.75	-1.17	-.56	-.06	.48
1979	5.16	.69	1.75	1.76	5.44	5.25	-2.35	-1.87	-1.17	-.59	.02
1980	5.19	.64	2.04	1.90	5.86	5.64	-2.03	-1.86	-1.15	-.57	.05
1981	5.77	.93	2.53	2.63	6.33	6.14	-1.53	-1.54	-.86	-.30	.30
1982	4.69	-.93		1.06							
1983	5.34	-.54		2.40							
1984	4.66	-.27		1.55							
1985	5.30	.06		2.48							
1986	3.74	-1.59		.18							
1987	5.32	.02		2.37							
1988	4.55	-.61		1.41							

R-SQUARES ON VPA SERIES

1977 TO 1981		.92	.84	.94	.84	.89	.90	.92	.93	.93	.94

VPA : VPA
GM : GEOMETRIC (CLUSTERED) MEAN
SI : STANDARD INDEX (WEIGHTED ARITHMETIC MEAN)
AR : ARITHMETIC MEAN
CC : COMPLEX CHANGEABLE - L INDICATES LOG(CPUE+X) TRANSFORM
- WHERE X IS SHOWN IN BRACKETS
CP : COMPLEX PERSISTENT - P INDICATES POISSON ERROR STRUCTURE

Table 2.3.2a Icelandic COD catch and effort data - Northern Area, Spring. Number of trawler-month observations by year, area, and month.

YEAR	AREA	JAN	FEB	MAR	APR	MAY	TOTAL
1974	101	77	53	89	15	43	277
	102	84	8	4	22	24	142
	103	6	16	3	2	7	34
	104	18	54	24	10	21	127
1975	101	180	173	105	20	66	544
	102	22	37	26	35	24	144
	103	1	2	0	7	0	10
	104	5	41	34	3	7	90
1976	101	164	57	93	7	40	361
	102	17	10	58	12	6	103
	103	10	3	1	5	2	21
	104	29	33	6	8	17	93
1977	101	159	91	116	125	67	558
	102	40	29	30	73	14	186
	103	9	36	11	9	39	104
	104	6	90	41	21	50	208
1978	101	180	34	14	40	79	347
	102	12	19	30	20	17	98
	103	34	22	20	22	11	109
	104	32	59	60	48	22	221
1979	101	127	98	86	41	65	417
	102	39	21	39	14	1	114
	103	12	24	7	13	6	62
	104	24	70	17	36	20	167
1980	101	197	203	152	22	26	600
	102	16	43	23	4	1	87
	103	9	12	7	20	11	59
	104	29	38	22	36	35	160
1981	101	126	8	87	68	51	340
	102	29	40	20	5	18	112
	103	42	22	15	2	6	87
	104	8	24	53	72	74	231
1982	101	28	20	37	10	21	116
	102	3	10	15	10	2	40
	103	15	32	14	38	1	100
	104	91	105	86	52	8	342
1983	101	48	28	36	42	39	193
	102	43	30	17	55	19	164
	103	98	30	29	14	9	180
	104	57	48	66	50	42	263
TOTAL		2126	1773	1593	1108	1011	7611

Table 2.3.2b Icelandic COD catch and effort data - Northern Area, Spring. Number of trawler-month observations by year and month.

YEAR	JAN	FEB	MAR	APR	MAY	TOTAL
1974	185	131	120	49	95	580
1975	208	253	165	65	97	788
1976	220	103	158	32	65	578
1977	214	246	198	228	170	1056
1978	258	134	124	130	129	775
1979	202	213	149	104	92	760
1980	251	296	204	82	73	906
1981	205	94	175	147	149	770
1982	137	167	152	110	32	598
1983	246	136	148	161	109	800

Table 2.3.3a Predicted Icelandic trawler catch rates relative to year 1974, January, and area 101 from the general linear model analysis which included effects for year, area, and month. For comparison, the year effect (prediction for the standard month and area) is shown.

RELATIVE CATCH RATES
Y+M+A MODEL

YEAR	AREA	FEB	MAR	APR	MAY	YEAR EFFECT
1975	102	0.994	1.069	1.065	0.922	1.004
	103	0.948	1.019	1.015	0.879	
	104	0.892	0.958	0.955	0.827	
1976	102	0.891	0.958	0.955	0.827	0.900
	103	0.849	0.913	0.910	0.788	
	104	0.799	0.859	0.856	0.741	
1977	102	0.984	1.057	1.054	0.913	0.994
	103	0.937	1.008	1.004	0.870	
	104	0.882	0.948	0.945	0.818	
1978	102	0.956	1.027	1.024	0.887	0.965
	103	0.911	0.979	0.976	0.845	
	104	0.857	0.921	0.918	0.795	
1979	102	1.096	1.178	1.174	1.017	1.107
	103	1.045	1.123	1.119	0.969	
	104	0.983	1.057	1.053	0.912	
1980	102	1.210	1.300	1.296	1.122	1.222
	103	1.153	1.239	1.235	1.070	
	104	1.085	1.166	1.162	1.006	
1981	102	1.492	1.604	1.598	1.384	1.507
	103	1.422	1.528	1.523	1.319	
	104	1.338	1.438	1.433	1.241	
1982	102	1.194	1.283	1.279	1.108	1.206
	103	1.138	1.223	1.219	1.056	
	104	1.071	1.151	1.147	0.993	
1983	102	0.981	1.055	1.051	0.910	0.991
	103	0.935	1.005	1.002	0.868	
	104	0.880	0.946	0.942	0.816	

Table 2.3.3b Predicted Icelandic trawler catch rates relative to year 1974, January, and area 101 from the general linear model analysis which included effects for year, area, month, year*area, and area*month. For comparison, the year effects from the analysis without interactions are shown.

		RELATIVE CATCH RATES Y+M+A+YA+AM MODEL				YR EFFECT Y+A+M
YEAR	AREA	FEB	MAR	APR	MAY	MODEL
1975	102	0.852	1.007	1.083	0.947	1.004
	103	0.483	0.357	0.517	0.402	
	104	0.859	0.820	0.884	0.810	
1976	102	0.820	0.969	1.042	0.911	0.900
	103	0.790	0.584	0.846	0.658	
	104	0.538	0.514	0.554	0.508	
1977	102	0.868	1.026	1.103	0.965	0.994
	103	1.090	0.805	1.168	0.907	
	104	1.003	0.958	1.032	0.946	
1978	102	0.834	0.985	1.060	0.927	0.965
	103	1.181	0.872	1.264	0.982	
	104	0.912	0.871	0.939	0.861	
1979	102	1.074	1.269	1.366	1.194	1.107
	103	1.101	0.814	1.180	0.916	
	104	0.902	0.861	0.928	0.851	
1980	102	1.086	1.283	1.381	1.208	1.222
	103	1.166	0.862	1.249	0.970	
	104	1.081	1.032	1.112	1.019	
1981	102	1.370	1.619	1.742	1.524	1.507
	103	1.648	1.218	1.765	1.371	
	104	1.957	1.870	2.014	1.846	
1982	102	0.999	1.181	1.270	1.111	1.206
	103	1.296	0.958	1.388	1.079	
	104	1.231	1.176	1.267	1.161	
1983	102	1.041	1.230	1.323	1.157	0.931
	103	1.232	0.910	1.319	1.025	
	104	0.917	0.876	0.943	0.865	

Table 2.3.3c Predicted Icelandic trawler catch rates relative to year 1974, January, and area 101 from the general linear model analysis which included effects for year, area, month, year*area, year*month and area*month. For comparison, the year effects from the analysis without interactions are shown.

		RELATIVE CATCH RATES Y+M+A+YA+YM+AM MODEL				YR EFFECT Y+A+M MODEL
YEAR	AREA	FEB	MAR	APR	MAY	
1975	102	1.023	0.695	0.901	0.791	1.004
	103	0.581	0.244	0.440	0.306	
	104	1.040	0.586	0.714	0.654	
1976	102	0.545	0.940	0.960	0.776	0.900
	103	0.537	0.574	0.813	0.521	
	104	0.414	0.593	0.569	0.480	
1977	102	0.773	1.008	1.168	1.012	0.994
	103	0.997	0.806	1.295	0.889	
	104	0.893	0.966	1.051	0.951	
1978	102	0.665	0.924	1.097	0.989	0.965
	103	0.983	0.848	1.396	0.997	
	104	0.762	0.880	0.981	0.922	
1979	102	1.064	1.409	1.376	1.140	1.107
	103	1.099	0.902	1.222	0.801	
	104	0.902	0.991	0.909	0.786	
1980	102	1.090	1.225	1.138	1.172	1.222
	103	1.256	0.875	1.128	0.919	
	104	1.170	1.091	0.952	1.023	
1981	102	1.179	1.647	2.142	1.594	1.507
	103	1.492	1.293	2.332	1.374	
	104	1.526	1.770	2.163	1.679	
1982	102	1.000	1.156	1.012	0.721	1.206
	103	1.275	0.915	1.110	0.626	
	104	1.149	1.103	0.907	0.674	
1983	102	0.921	0.928	1.320	1.118	0.991
	103	1.147	0.717	1.414	0.949	
	104	0.879	0.736	0.993	0.869	

Table 2.3.4a Arithmetic mean of area specific, predicted Icelandic trawler catch rates relative to year 1974, January, and area 101 from the general linear model analysis which included effects for year, area, and month. The year effect (the prediction for the standard month and area) is also shown.

YEAR	RELATIVE CATCH RATES Y+M+A MODEL FEB	MAR	APR	MAY	YEAR EFFECT
1975	0.945	1.015	1.012	0.875	1.004
1976	0.847	0.910	0.907	0.785	0.900
1977	0.934	1.004	1.001	0.867	0.994
1978	0.908	0.976	0.972	0.942	0.965
1979	1.041	1.119	1.116	0.966	1.107
1980	1.149	1.235	1.231	1.066	1.222
1981	1.417	1.523	1.518	1.315	1.507
1982	1.134	1.219	1.215	1.052	1.206
1983	0.932	1.002	0.998	0.865	0.991

Table 2.3.4b Arithmetic mean of area specific, predicted Icelandic trawler catch rates relative to year 1974, January, and area 101 from the general linear model analysis which included effects for year, area, month, year*area and area*month. For comparison, the year effects from the analysis without interactions are shown.

YEAR	RELATIVE CATCH RATES Y+M+A+YA+AM MODEL FEB	MAR	APR	MAY	YR EFFECT Y+A+M MODEL
1975	0.731	0.728	0.828	0.720	1.004
1976	0.716	0.689	0.814	0.692	0.900
1977	0.987	0.930	1.101	0.939	0.994
1978	0.976	0.910	1.088	0.923	0.965
1979	1.026	0.982	1.158	0.987	1.107
1980	1.111	1.059	1.247	1.066	1.222
1981	1.658	1.569	1.840	1.580	1.507
1982	1.176	1.105	1.308	1.117	1.206
1983	1.063	1.005	1.195	1.016	0.991

Table 2.3.4c Arithmetic mean of area specific, predicted Icelandic trawler catch rates relative to year 1974, January, and area 101 from the general linear model analysis which included effects for year, area, month, year*area, year*month and area*month. For comparison, the year effects from the analysis without interactions are shown.

	RELATIVE CATCH RATES Y+M+A+YA+YM+AM MODEL				YR EFFECT Y+A+M MODEL
YEAR	FEB	MAR	APR	MAY	
1975	0.881	0.508	0.685	0.584	1.004
1976	0.499	0.702	0.781	0.592	0.900
1977	0.887	0.926	1.171	0.951	0.994
1978	0.803	0.884	1.158	0.969	0.965
1979	1.021	1.101	1.169	0.909	1.107
1980	1.172	1.063	1.073	1.038	1.222
1981	1.399	1.570	2.212	1.549	1.507
1982	1.141	1.058	1.009	0.673	1.206
1983	0.982	0.794	1.239	0.979	0.991

Table 2.4.1 EGFS Cod age 1 : integrated indices.

Yearclass	Rough	Bumpy	Medium	Smooth	Standard	VPA
1976	84810	3623	1945	1290	627	726
1977	456100	1222	824	639	228	426
1978	22620	1442	947	656	242	449
1979	13110	3463	2549	1862	508	800
1980	14610	561	203	99	114	272
1981	57630	1621	1038	766	324	557
1982	6980	651	394	262	154	271
1983	16170	3386	2442	1942	612	528
1984	340400	156	95	62	43	105
1985	5911	1649	1154	768	344	576
1986	1799	626	468	352	142	250
1987	483	219	164	135	84	-11
1988	-11	-11	-11	-11	228	-11

Table 2.4.2

Analysis by RCRTINX2 of data from file integrat.dat
EGFS Cod age 1 : integrated indices

Data for 5 surveys over 13 years
REGRESSION TYPE = C
TAPERED TIME WEIGHTING APPLIED
POWER = 3 OVER 20 YEARS
PRIOR WEIGHTING NOT APPLIED
FINAL ESTIMATES SHRUNK TOWARDS MEAN
ESTIMATES WITH S.E.'S GREATER THAN THAT OF MEAN INCLUDED
MINIMUM S.E. FOR ANY SURVEY TAKEN AS .10
MINIMUM OF 5 POINTS USED FOR REGRESSION

Yearclass = 1984

Survey/ Series	Index Value	Slope	Inter-cept	Rsquare	No. Pts	Predicted Value	Sigma	Standard Error	Weight
rough	12.7379	1.133	-5.562	.0710	8	8.8722	1.57962	2.00586	.00439
bumpy	5.0562	.577	1.895	.8845	8	4.8135	.15772	.25335	.27546
medium	4.5643	.487	2.803	.8444	8	5.0271	.18740	.27203	.23893
smooth	4.1431	.441	3.288	.8001	8	5.1139	.21819	.30219	.19362
standa	3.7842	.696	2.190	.8259	8	4.8231	.20042	.31464	.17859
MEAN						6.1499	.40273	.40273	.10901

Yearclass = 1985

Survey/ Series	Index Value	Slope	Inter-cept	Rsquare	No. Pts	Predicted Value	Sigma	Standard Error	Weight
rough	8.6847	-1.354	20.326	.0982	9	8.5671	2.05592	2.37714	.00206
bumpy	7.4085	.617	1.592	.9466	9	6.1654	.16120	.17127	.39642
medium	7.0519	.578	2.160	.8987	9	6.2370	.22781	.24302	.19690
smooth	6.6451	.547	2.573	.8541	9	6.2086	.28041	.29847	.13053
standa	5.8435	.748	1.886	.9172	9	6.2563	.20391	.21788	.24496
MEAN						5.9729	.63189	.63189	.02912

Yearclass = 1986

Survey/ Series	Index Value	Slope	Inter-cept	Rsquare	No. Pts	Predicted Value	Sigma	Standard Error	Weight
rough	7.4955	-1.061	17.021	.1413	10	9.0707	1.60658	1.99136	.00296
bumpy	6.4409	.631	1.512	.9367	10	5.5785	.16942	.18272	.35208
medium	6.1506	.586	2.117	.8964	10	5.7237	.22160	.23566	.21166
smooth	5.8665	.557	2.527	.8505	10	5.7934	.27332	.28929	.14046
standa	4.9628	.757	1.845	.9161	10	5.6024	.19724	.21207	.26137
MEAN						6.0067	.61123	.61123	.03146

Yearclass = 1987

Survey/ Series	Index Value	Slope	Inter-cept	Rsquare	No. Pts	Predicted Value	Sigma	Standard Error	Weight
rough	6.1821	-1.861	24.707	.0356	11	13.2007	3.31622	4.29806	.00071
bumpy	5.3936	.637	1.469	.9384	11	4.9031	.16328	.19451	.34774
medium	5.1059	.603	1.989	.8885	11	5.0662	.22577	.25896	.19619
smooth	4.9127	.578	2.369	.8352	11	5.2071	.28315	.31573	.13198
standa	4.4427	.765	1.791	.9181	11	5.1914	.19044	.21413	.28693
MEAN						5.9510	.60080	.60080	.03645

Table 2.6.1 North Sea COD - Age 1.
Squared correlation coefficient of indices vs. VPA estimates.

Metod	Options	1977-1981 r^2	1977-1981 s.e.	1977-1988 r^2	1977-1988 s.e.
Stratification	Untransformed - 1 str.			0.93	
	" - 3 str.			0.95	
	" - 8 str.			0.95	
	" 15 str.			0.95	
	Transformed, uncorrected - 1 str.			0.72	
	" " - 3 str.			0.78	
	" " - 8 str.			0.82	
	" " - 15 str.			0.87	
	Transformed, corrected - 1 str.			0.89	
	- 3 str.			0.78	
	- 8 str.			0.81	
	- 15 str.			0.71	
Interpolation	Rough	0.00	30.61	0.02	4.22
	Bumpy	0.99	0.05	0.94	0.15
	Medium	0.96	0.10	0.89	0.22
	Smooth	0.92	0.15	0.84	0.27
	Kriging - uncorrected			0.92	
	" - corrected			0.91	
GLM	Geometric mean	0.90	0.17	0.89	0.22
(r^2 for 1977-1981)	Arithmetic mean	0.99	0.06	-	-
	Standard index	0.96	0.10	0.92	0.18
	Changeable - complex - poisson	0.98	0.08	-	-
	Complex - changeable - log	0.98	0.07	-	-
	Complex - persistent - poisson	0.98	0.08	-	-
	" " - log + 0.1	0.90	0.17	----n/a---	
	" " - log + 0.25	0.90	0.17	-	-
	" " - log + 0.5	0.91	0.16	-	-
	" " - log + 1.0	0.92	0.15	-	-

Table 3.4.1 Dangers of current implementations of the various methods.

Method	ad-hoc				intermediate				integrated			non-tune		
	Hybrid	LS	AC-1	AEFM	CCPUE	SURVIV	XSA	CAGEAN	ADAPT	GLM	TSER2	TSER1	SVPA	COVEN
Time trend in q	**	*	**	*	*	*	*	*	*	*	*	N.A.	N.A.	N.A.
No time trend in q	**	-	**	-	-	-	-	-	-	-	-	N.A.	N.A.	N.A.
Time trend in F	-	-	-	-	-	-	-	-	-	-	-	**	**	**
Single CPUE	**	*	**	*	*	-	-	-	-	-	-	N.A.	N.A.	N.A.
Inexperienced user	*	-	-	**	**	-	-	*	*	*	*	*	-	-
Test of fit provided	yes	yes	no	no	no	yes	yes	yes	lots	yes	yes	yes	few	no

Danger level:
** = very dangerous
* = dangerous
+ = somewhat dangerous
- = little danger
N.A. = Not applicable

Table 4.1 Coefficient of variation of the TAC for the average and worst case for North Sea cod, haddock and whiting for different TAC years.

	cod		haddock		whiting	
TAC year	average	worst	average	worst	average	worst
Jan-Dec	5.45	5.54	8.72	13.80	10.11	15.98
Apr-Mar	5.65	6.19	11.33	17.27	11.31	17.11
Jul-Jun	6.50	8.07	9.38	12.36	9.55	11.70
Oct-Sep	7.27	9.44	11.51	15.29	10.41	12.63

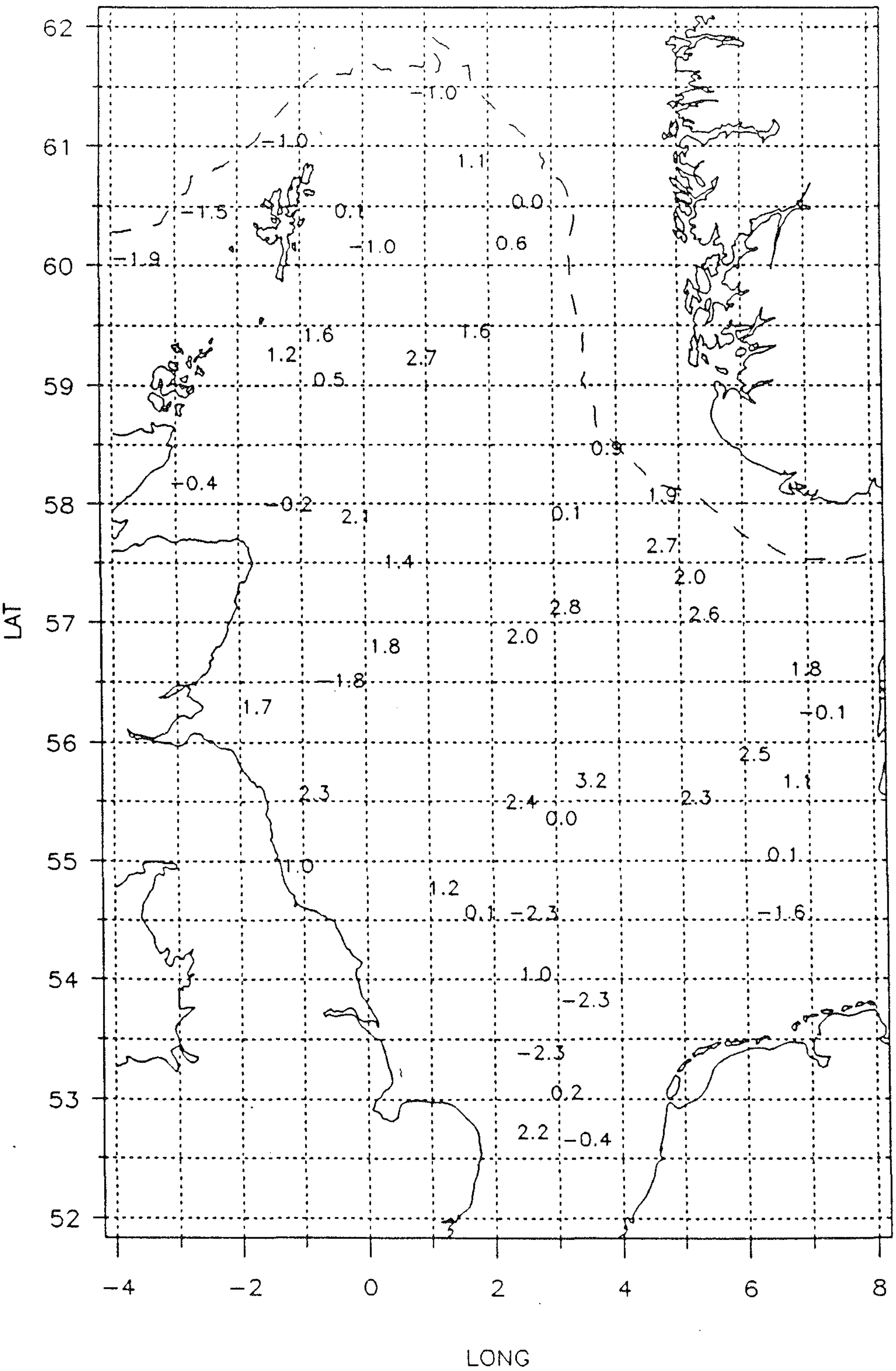


Figure 2.1.1 Mean log (CPUE + 0.1), age 2 COD, by station. 1977-1981 combined.

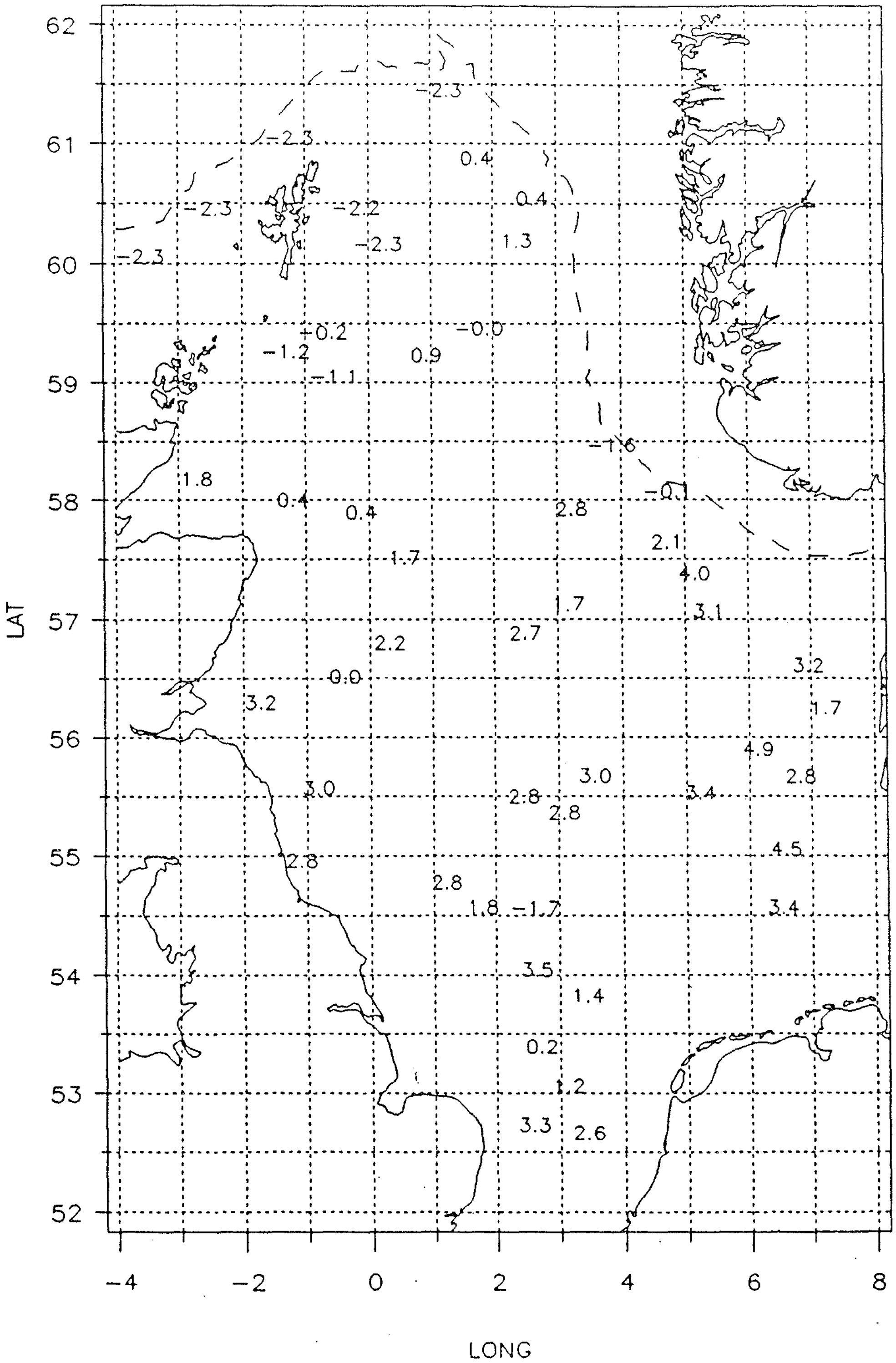


Figure 2.1.2 Mean log (CPUE + 0.1), age 1 COD, by station.
1977-1981 combined.

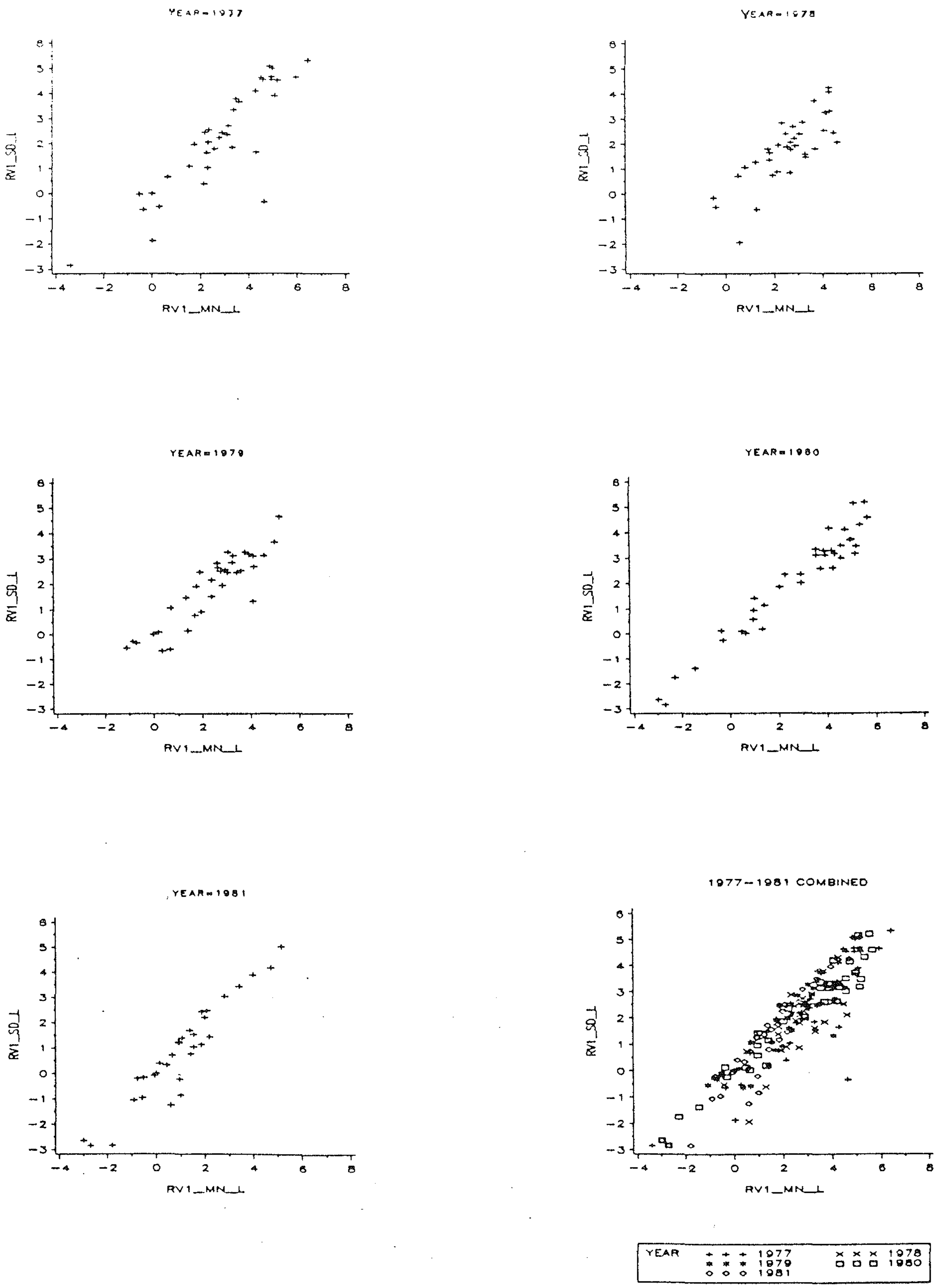


Figure 2.1.3 English groundfish surveys, age 1 COD.
Log standard errors vs. log means within clusters.

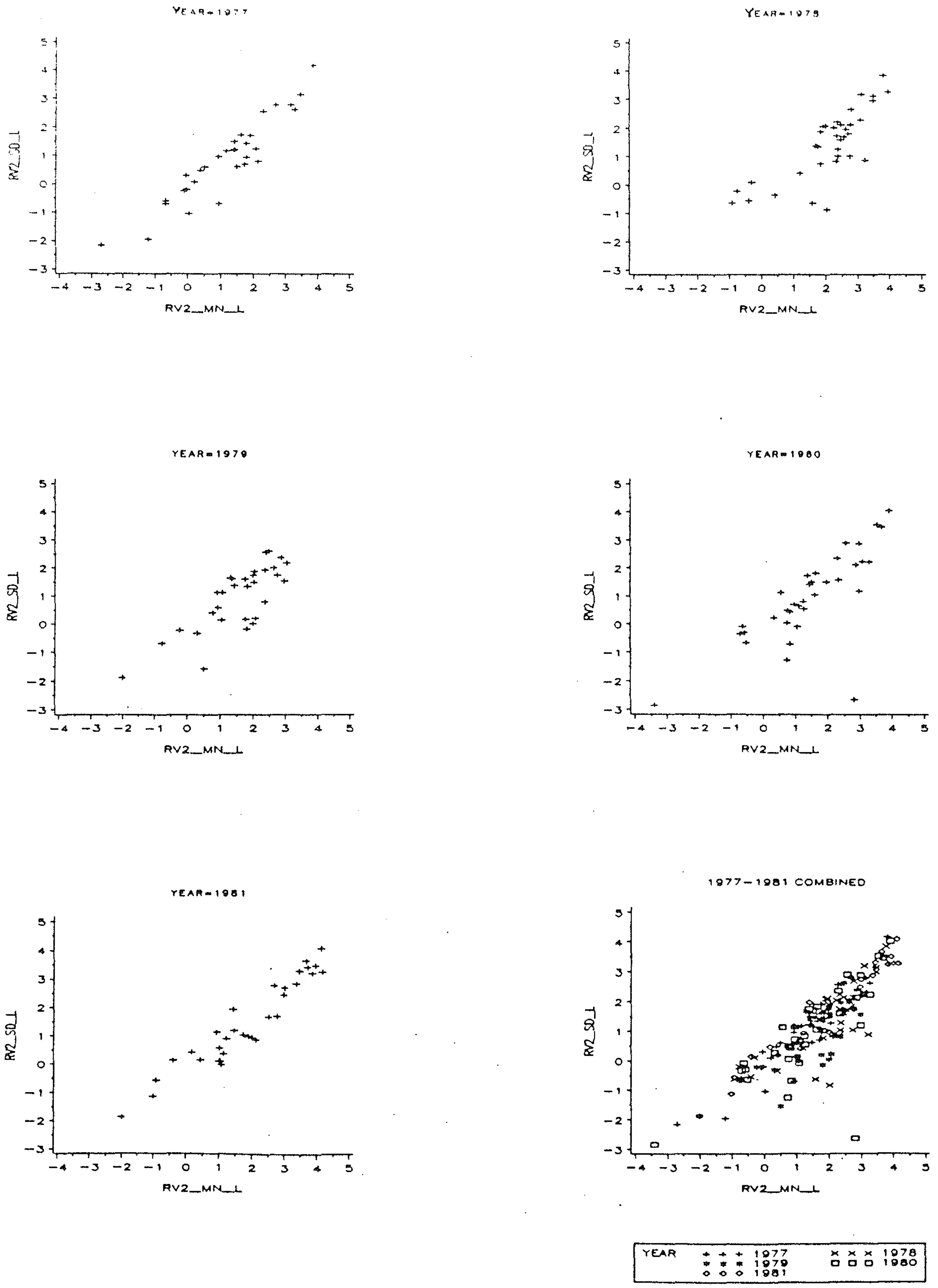


Figure 2.1.4 English groundfish surveys, age 2 COD.
Log standard errors vs. log means within clusters.

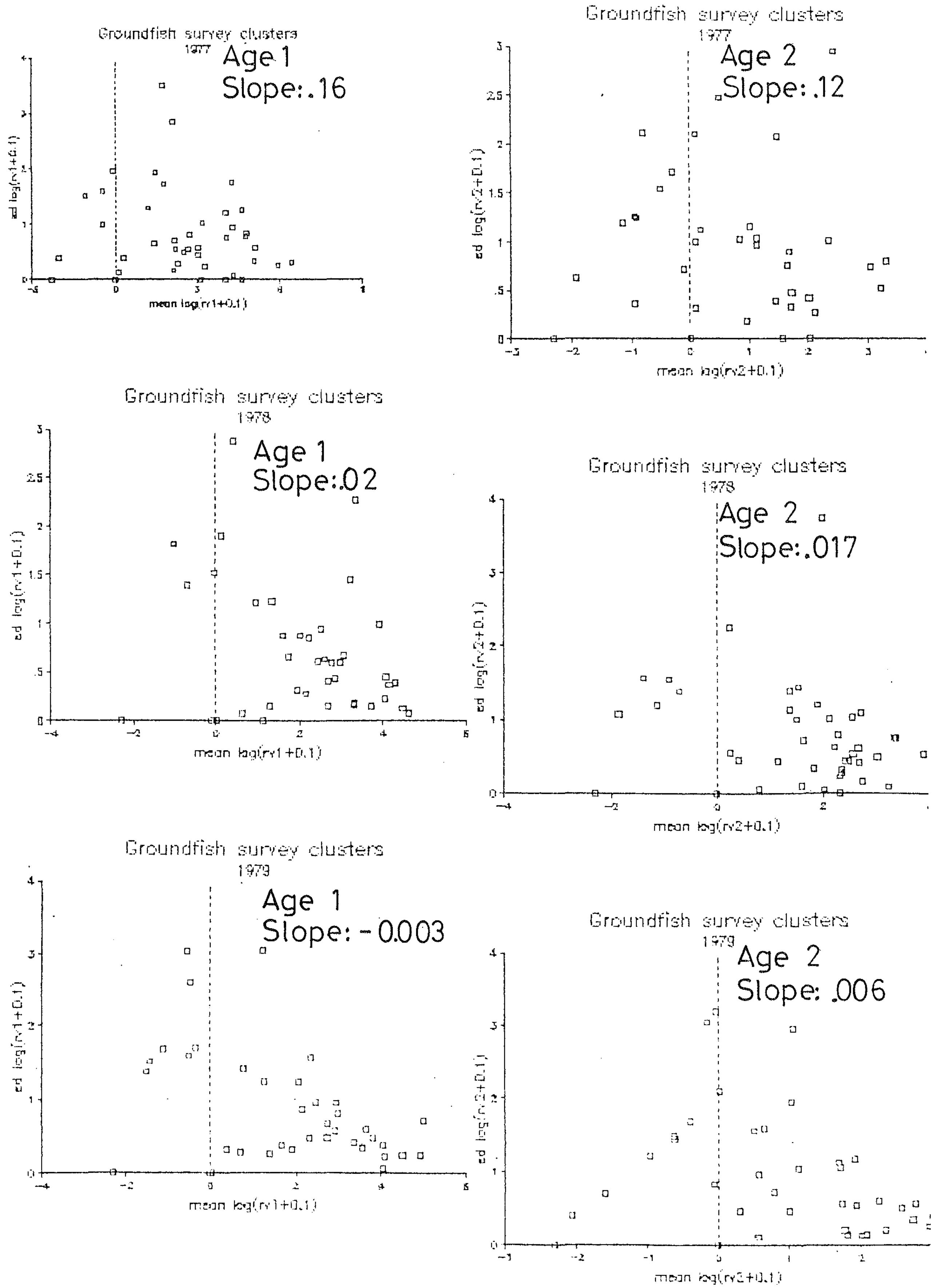


Figure 2.1.5 Relationship between sd and mean after log (x + 0.1) transformation for groundfish survey. Slope of regression line.

(cont'd)

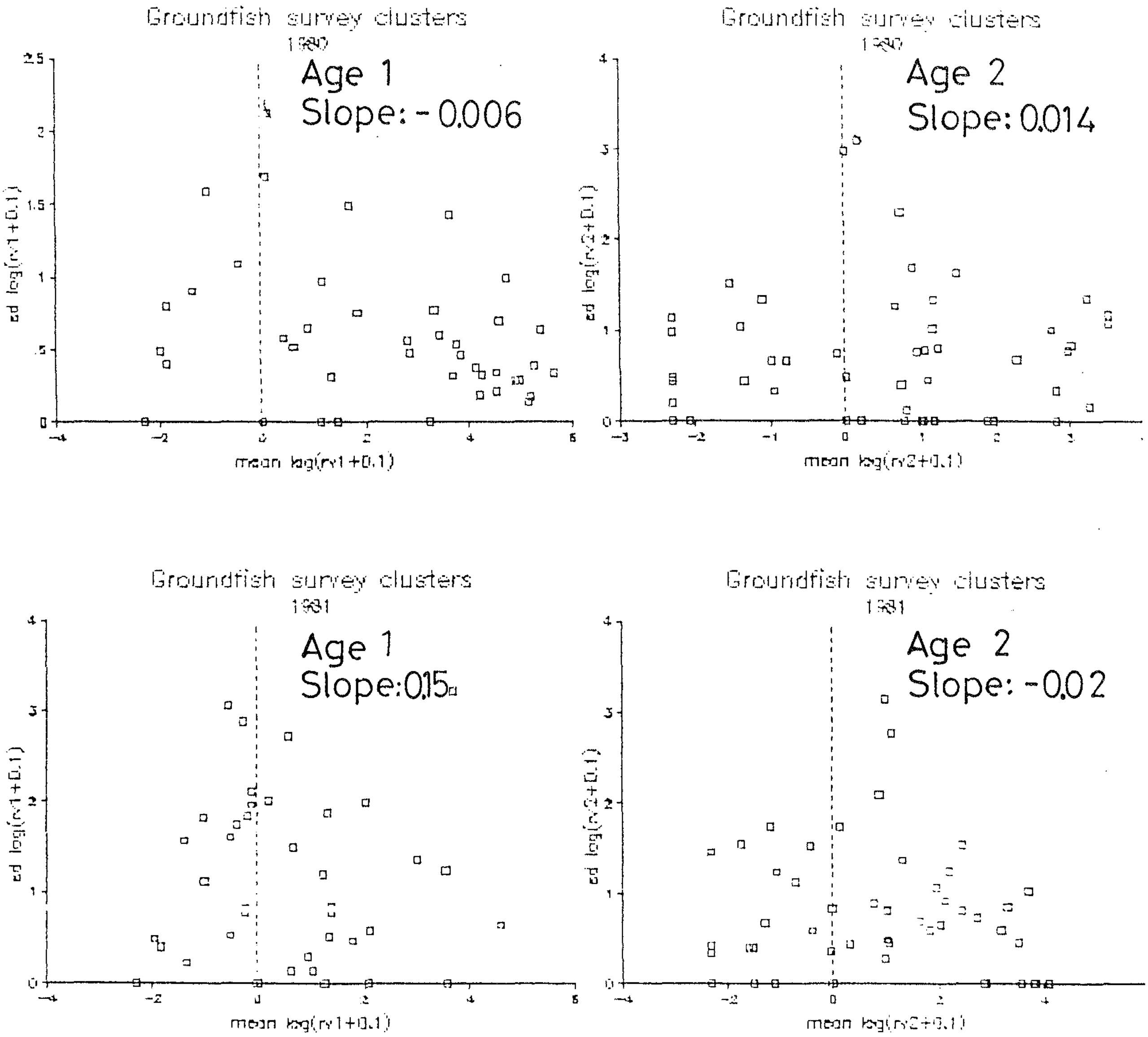


Figure 2.1.5 (cont'd)

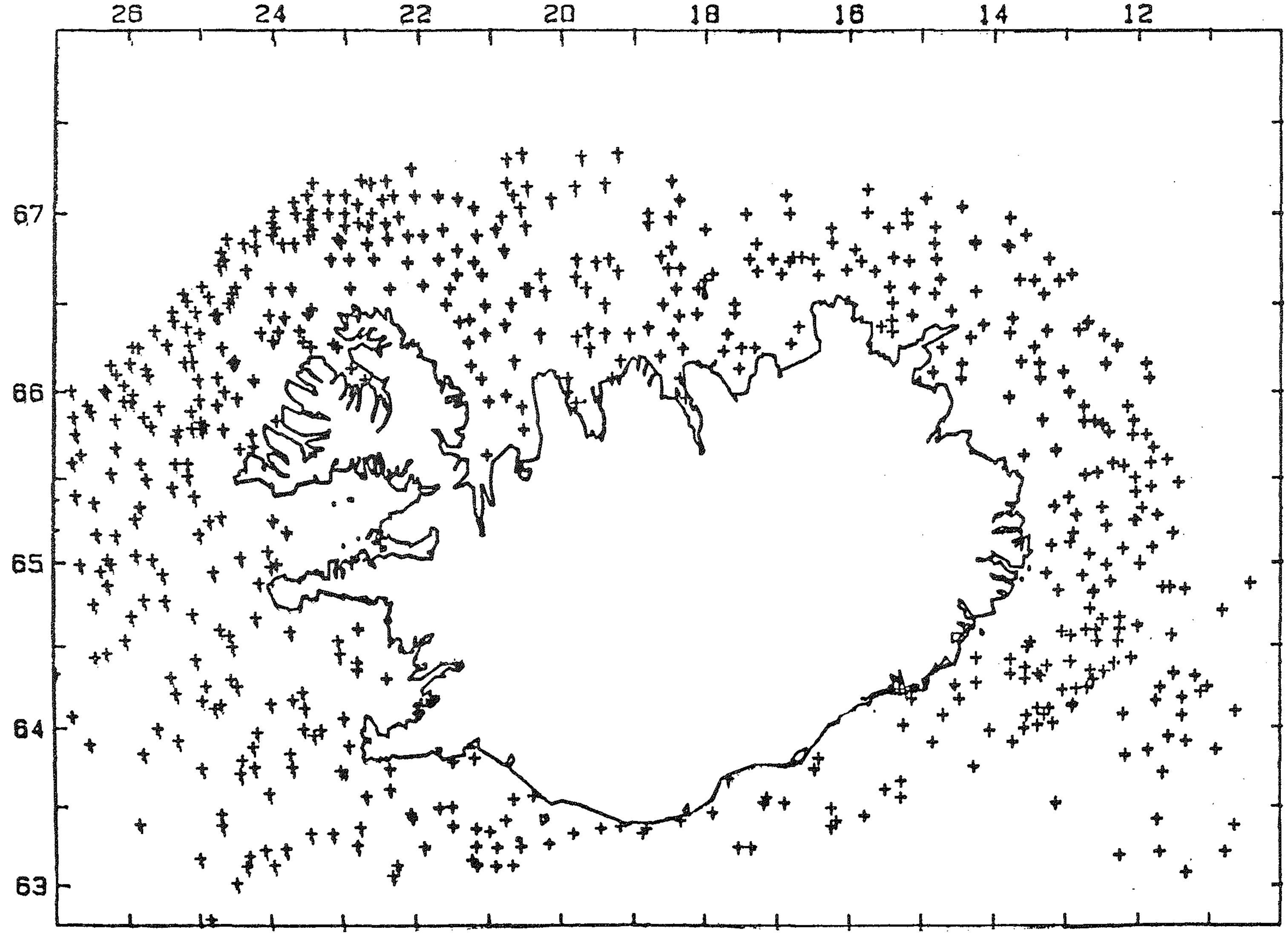


Figure 2.1.6 Stations in Icelandic groundfish survey.

Icelandic survey

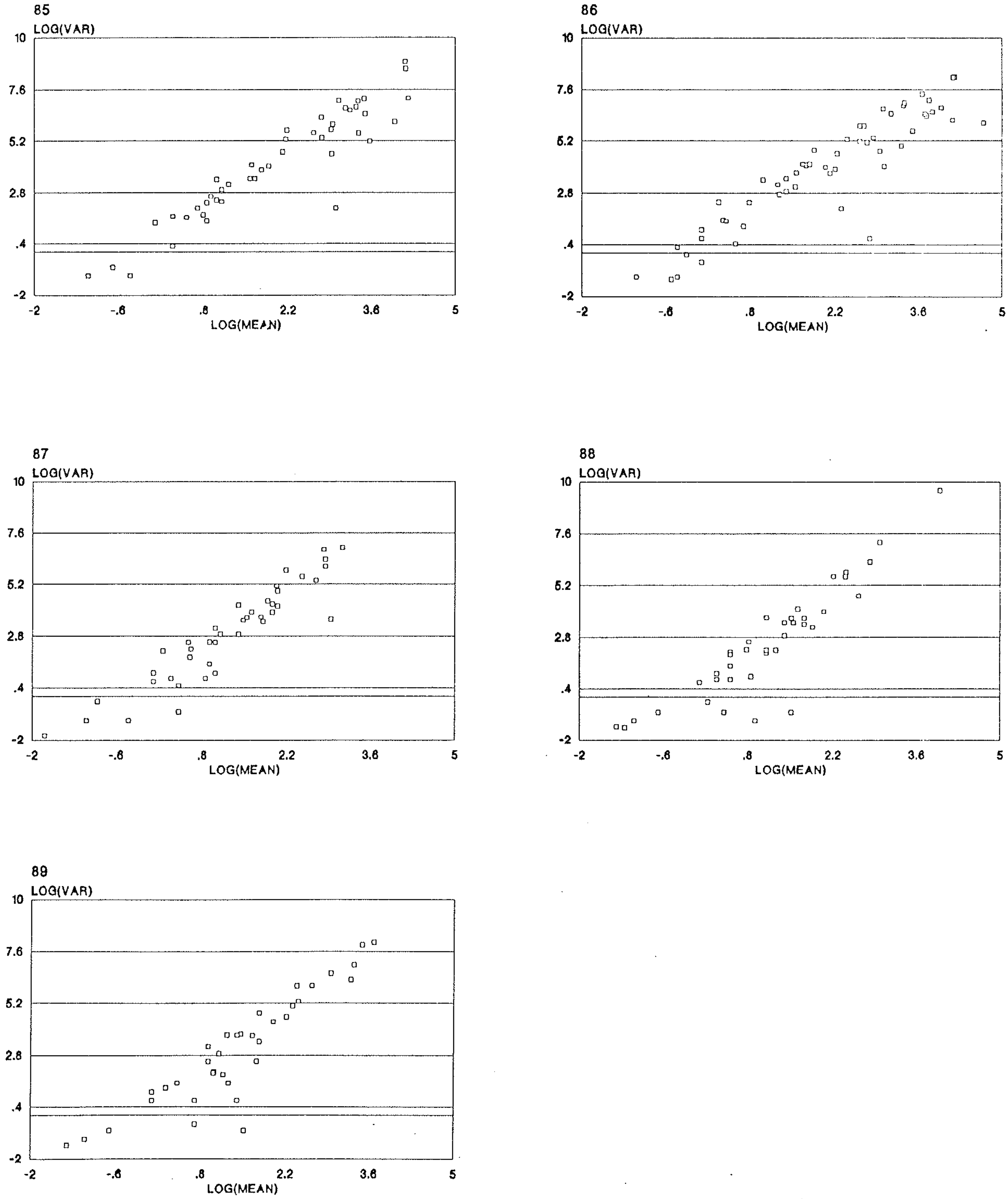


Figure 2.1.7 Icelandic surveys. Log variance vs. log mean.

Sub Areas marked with X were excluded.

10 years: 1974 - 83
4 areas: 101 - 104
5 months: Jan - May

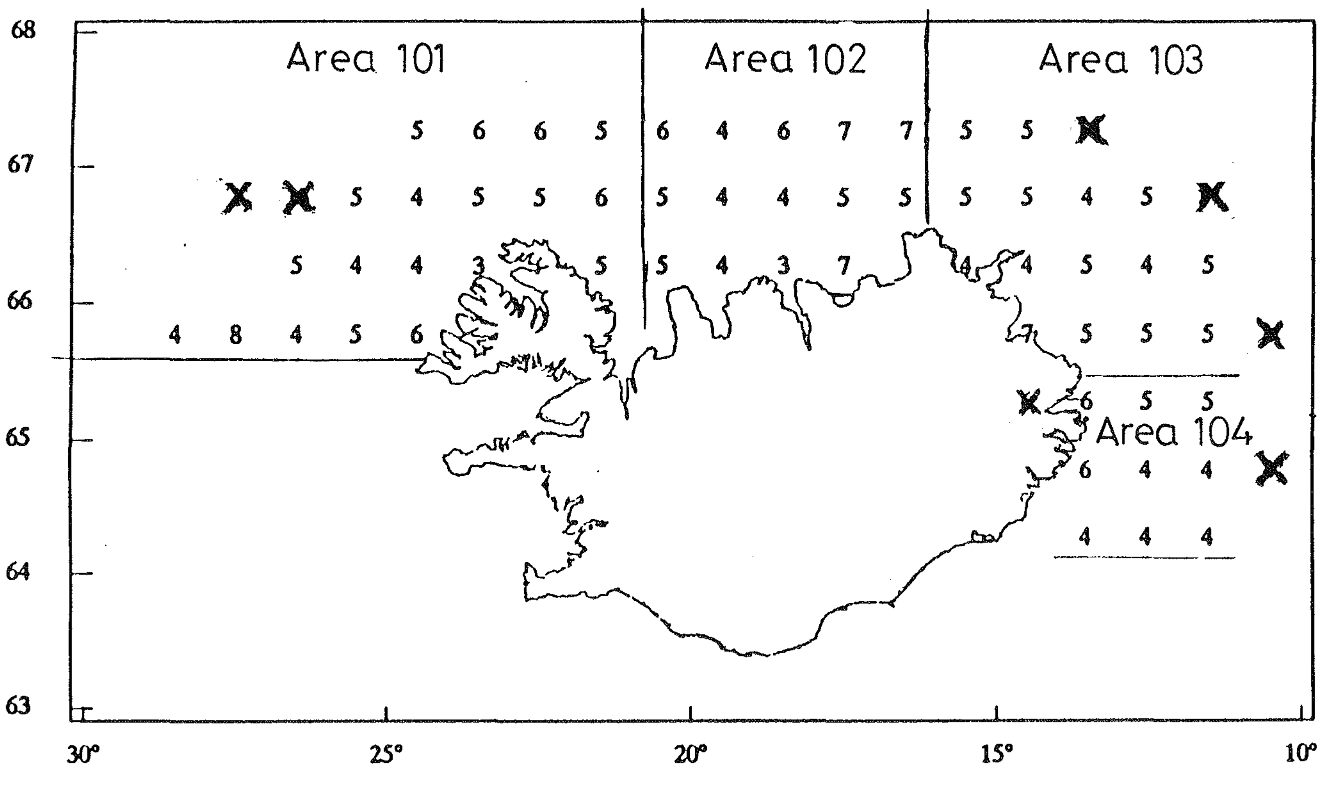


Figure 2.1.8 Areas used in GLM on Icelandic trawler cod catch per effort. Relative abundance per square, Northern region, January-May. In Stefanson, 1988 (ICES, Doc. C.M.1988/D:13).

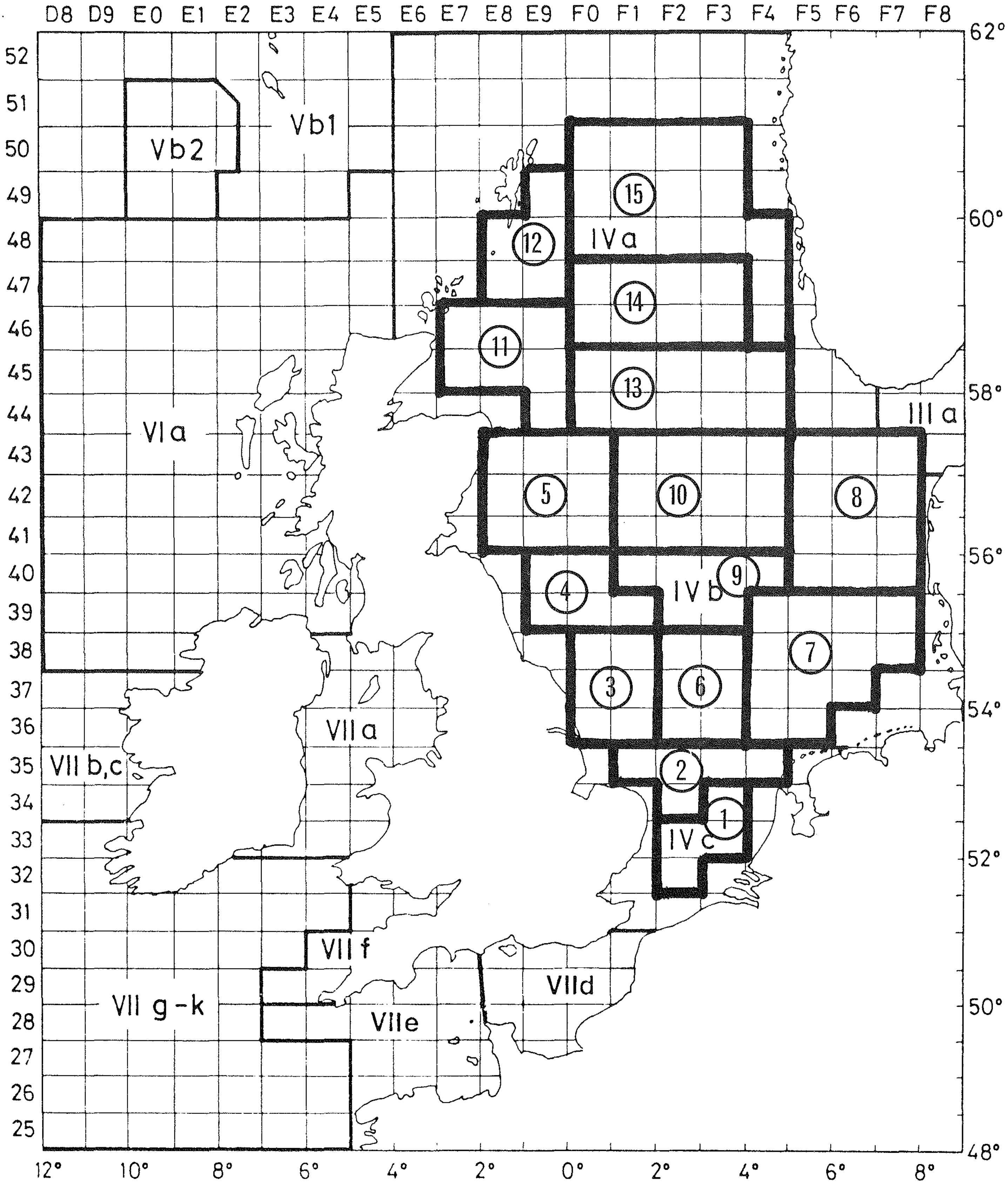


Figure 2.2.1 Various stratified sampling schemes. English groundfish surveys.

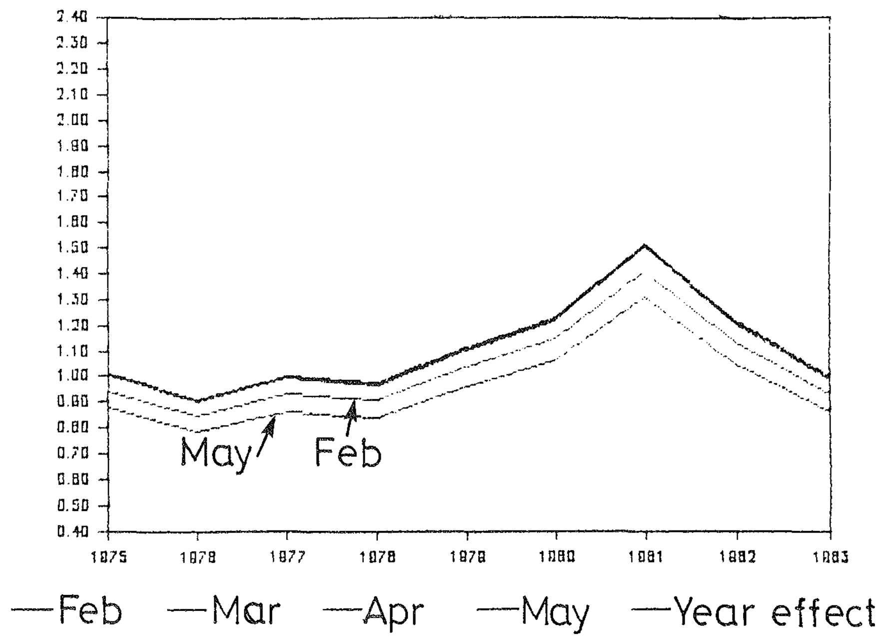


Figure 2.3.1 Yearly predicted relative catch rates of cod from Icelandic trawlers fishing in Jan-May in areas defined in Figure 2.1.8 from a GLM with year, area and month effects. The standard (bold line, pure year effect) and monthly values in a year are the arithmetic means of the values for areas in that year and month.

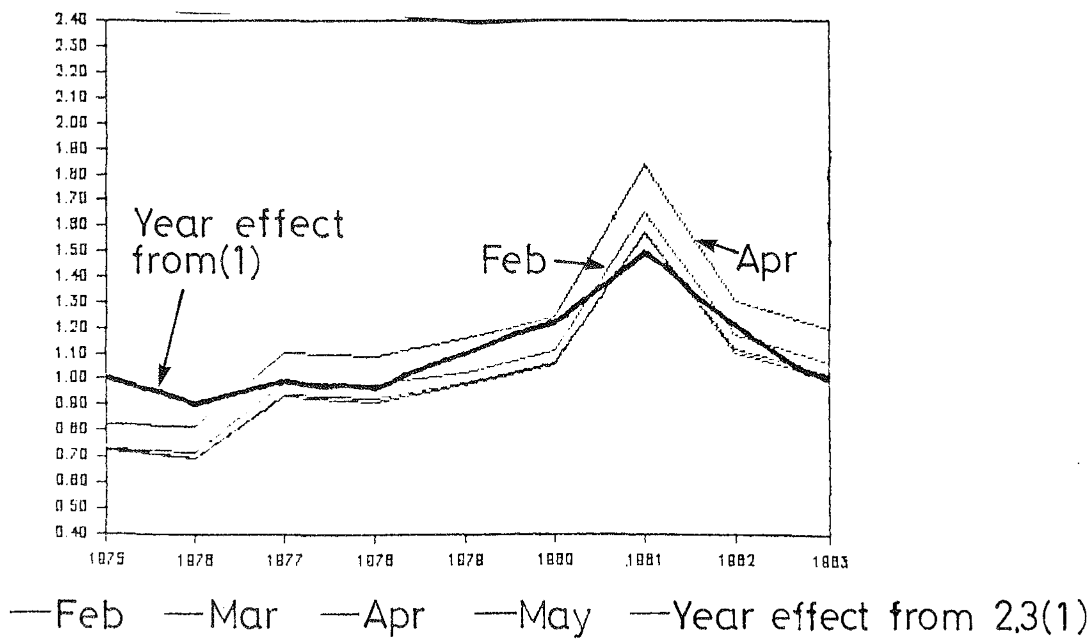


Figure 2.3.2 Yearly predicted relative catch rates of cod from Icelandic trawlers fishing in Jan-May in areas defined in Figure 2.1.8 from a GLM with year, area, month, year*area and area*month effects. The standard and monthly values in a year are the arithmetic means of the values for area in that year and month. The bold line is the prediction for the standard month from the GLM with only year, area and month effects.

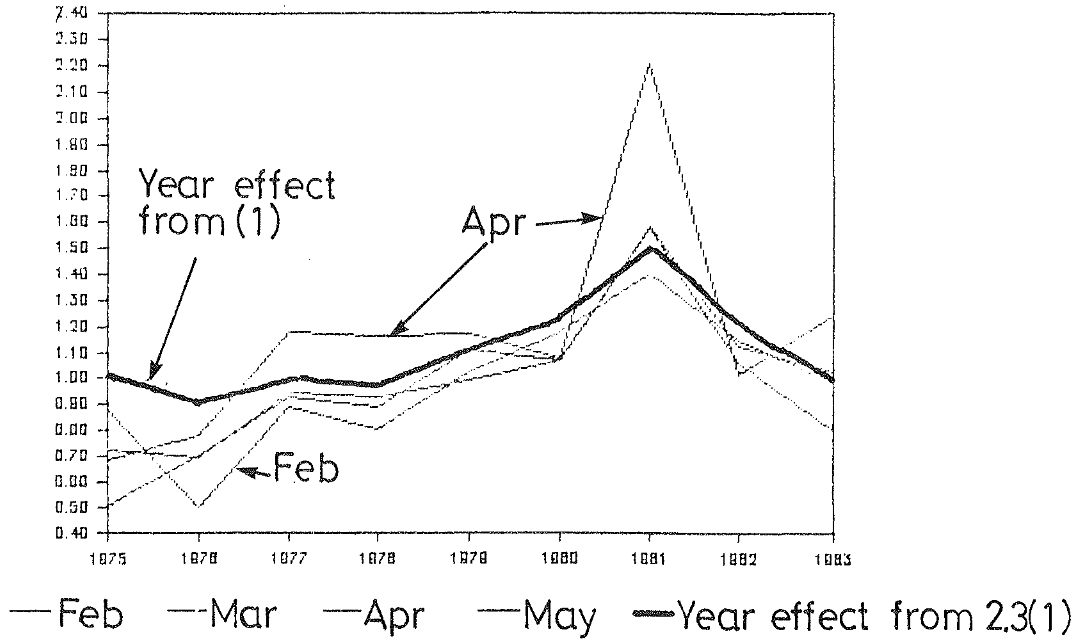


Figure 2.3.3 Yearly predicted relative catch rates of cod from Icelandic trawlers fishing in Jan-May in areas defined in Figure 2.1.8 from a GLM with year, area, month, year*area, year*month and area*month effects. The standard and monthly values in a year are the arithmetic means of the values for areas in that year and month. The bold line is the prediction for the standard month from the GLM with only year, area and month effects.

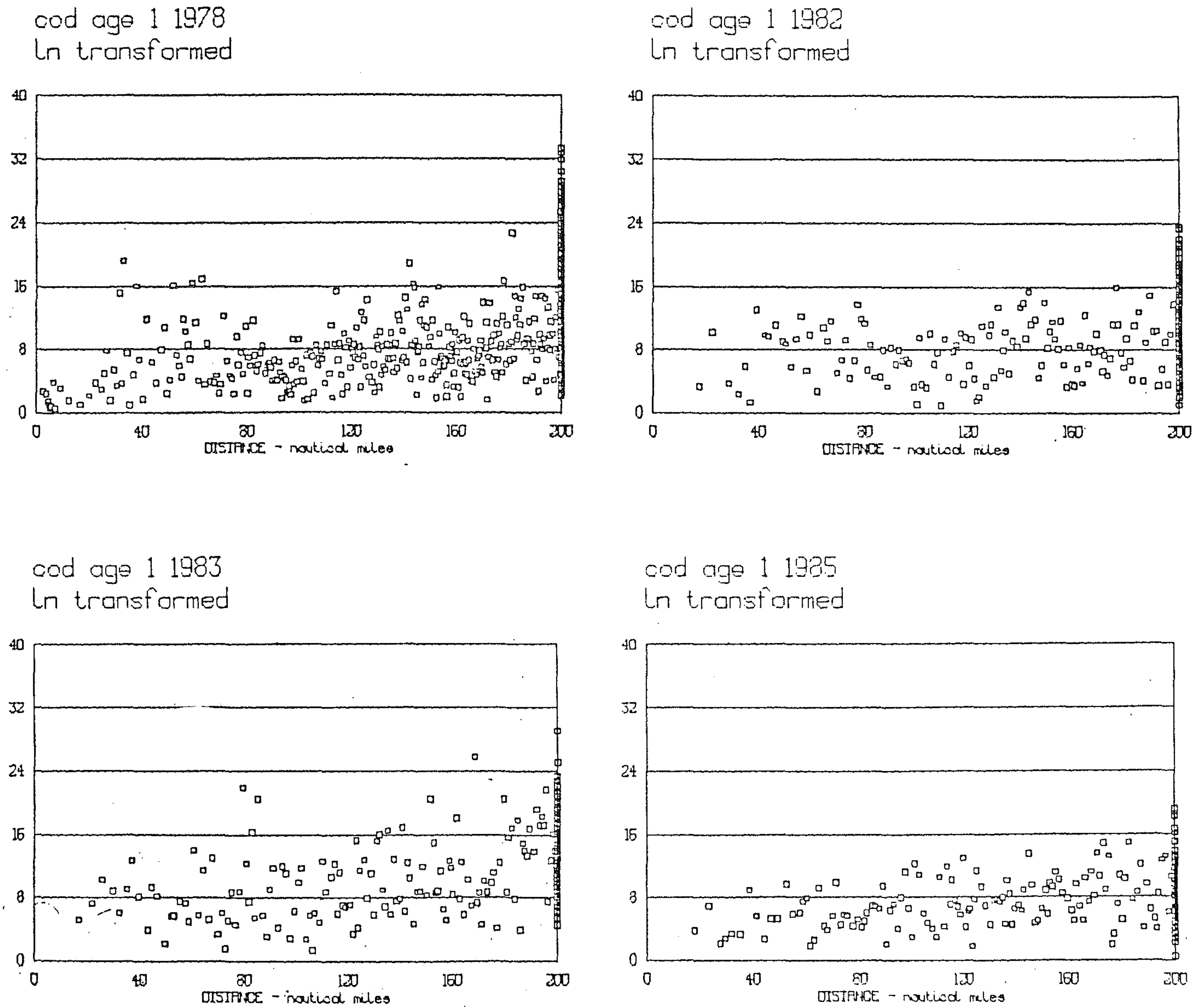


Figure 2.4.1 Empirical half year square difference of local log densities as a function of distances. English groundfish surveys, COD, age 1.

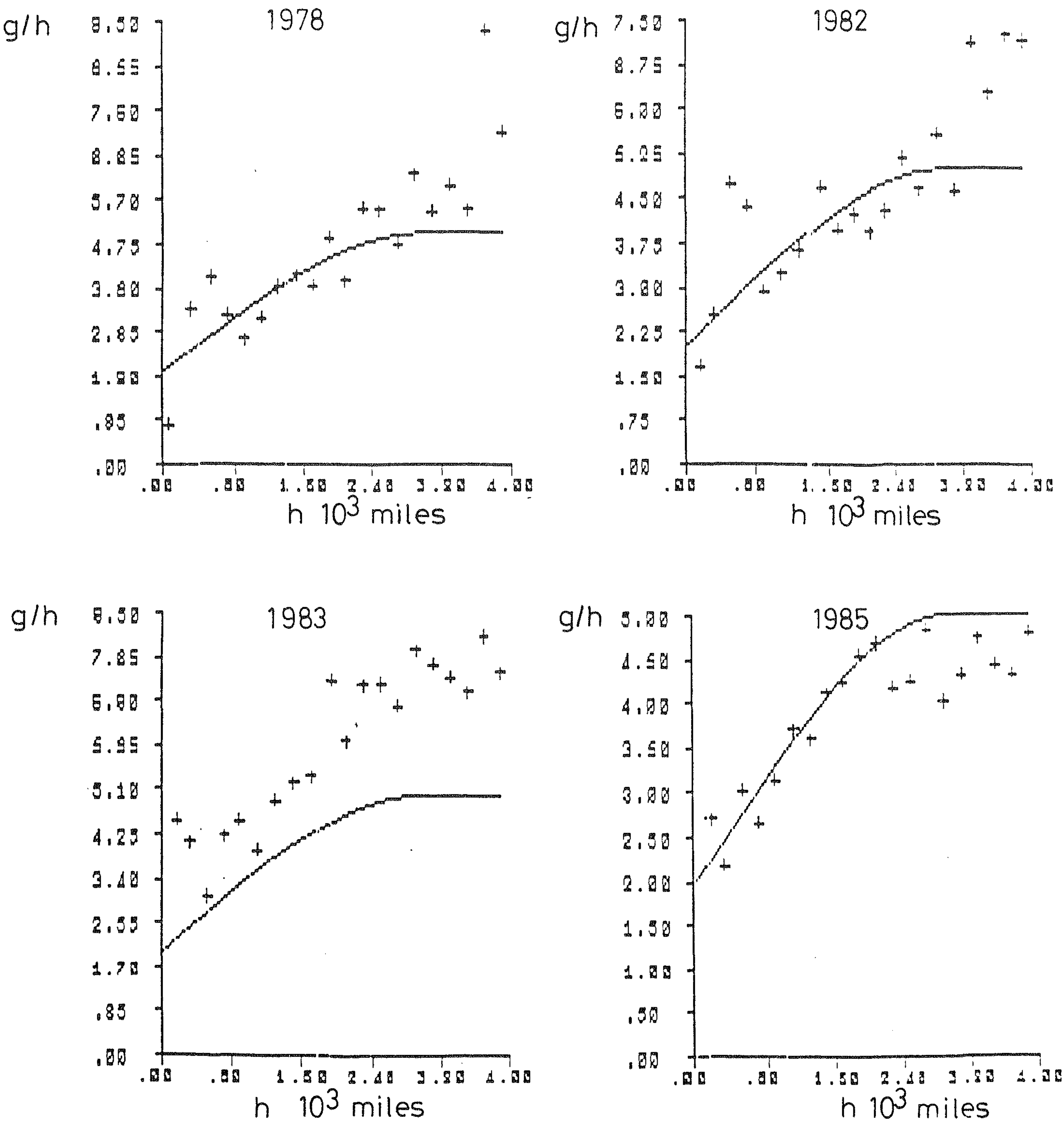


Figure 2.4.2 Theoretical semi-variogram g(h) against (averaged) half squared differences.

Spherical variogram $g(h) = 2 + 3\ (1.5\ \frac{h}{300} - 5(\frac{h}{300})^3)$ for $h < 300$.

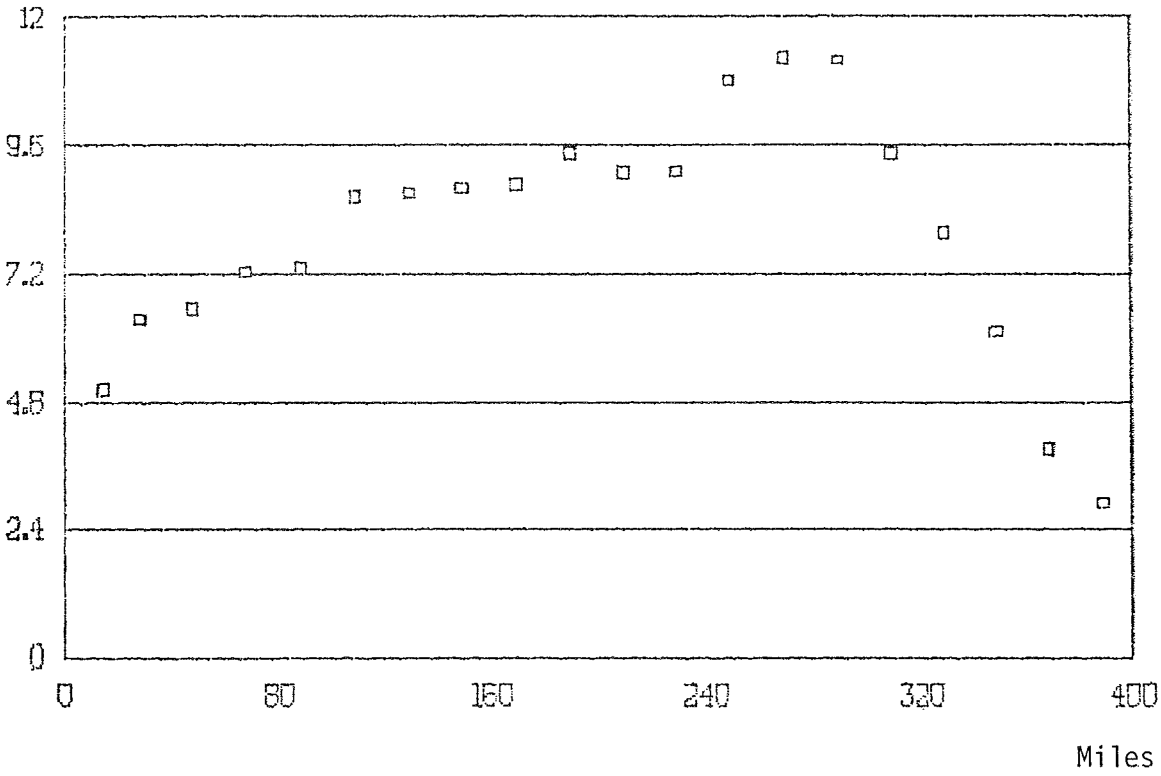


Figure 2.4.3 Observed half squared differences of log densities as a function of the distances.

Observations have been averaged over a distance interval of 20 miles.

Icelandic surveys - COD.

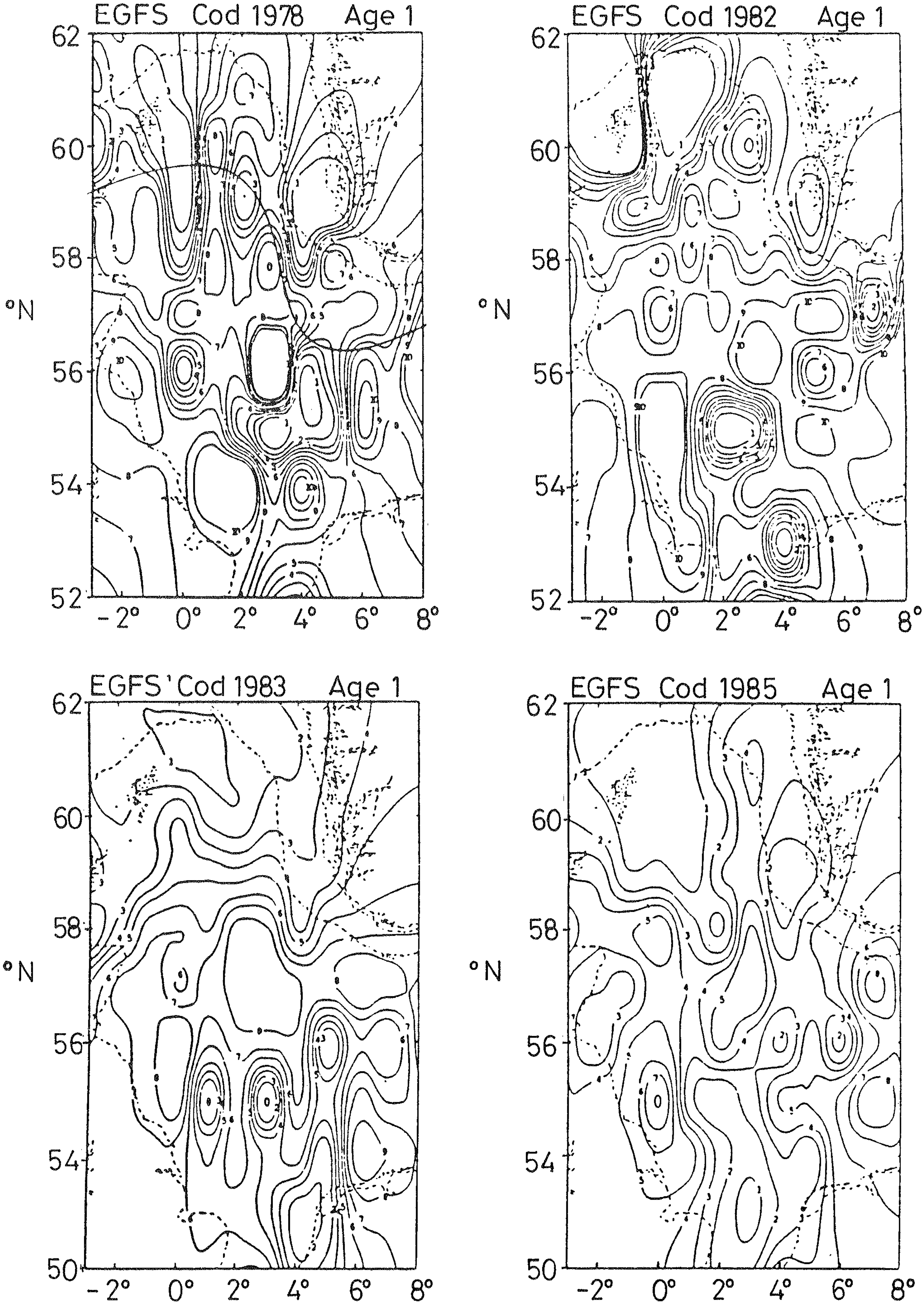


Figure 2.4.4 COD, age 1. Mapping of the interpolated log densities for years 1978, 1982, 1983, 1985. Parameter value 0.1.

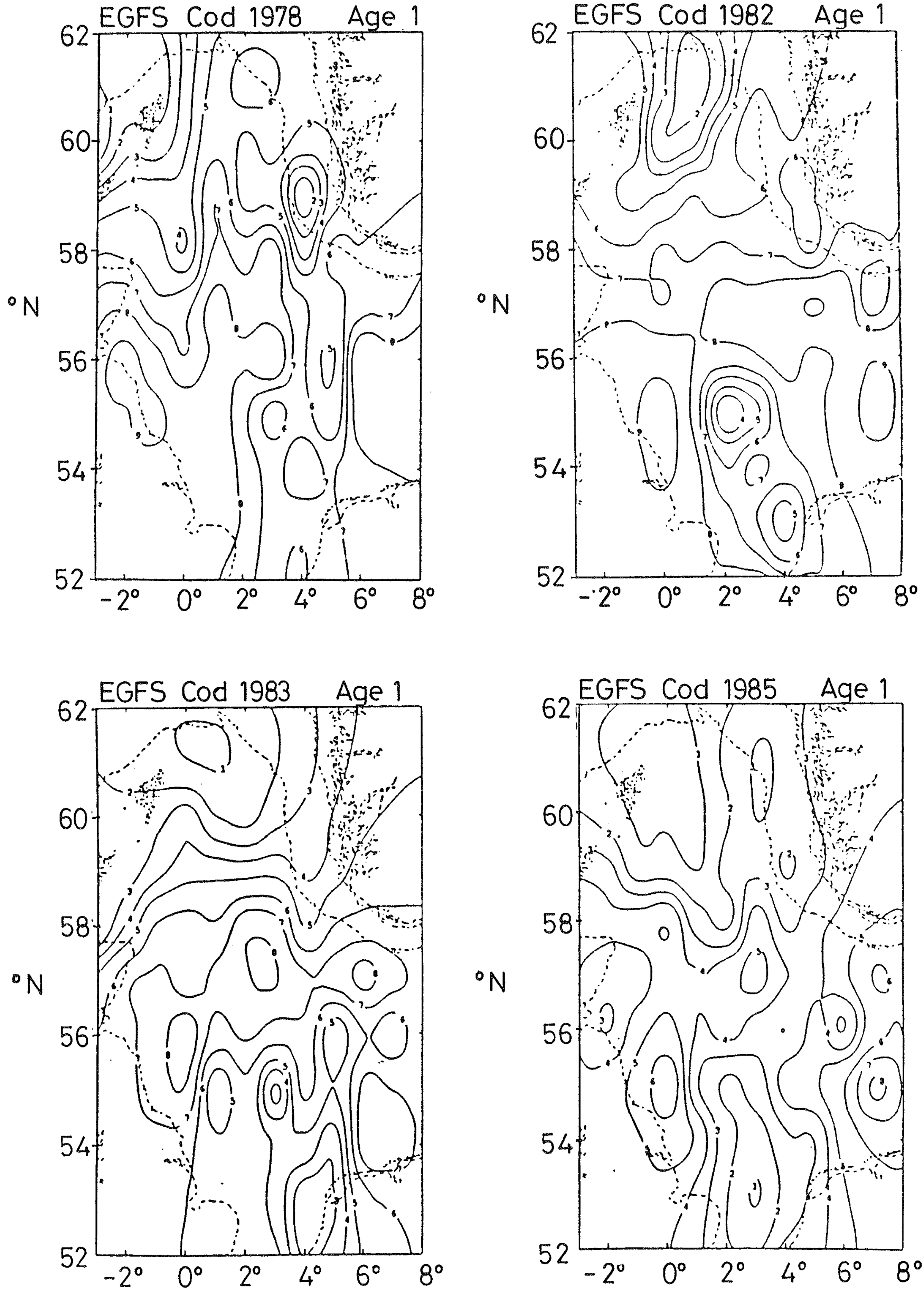


Figure 2.4.5 Same as Figure 2.4.4, but for parameter value equal to 0.3.

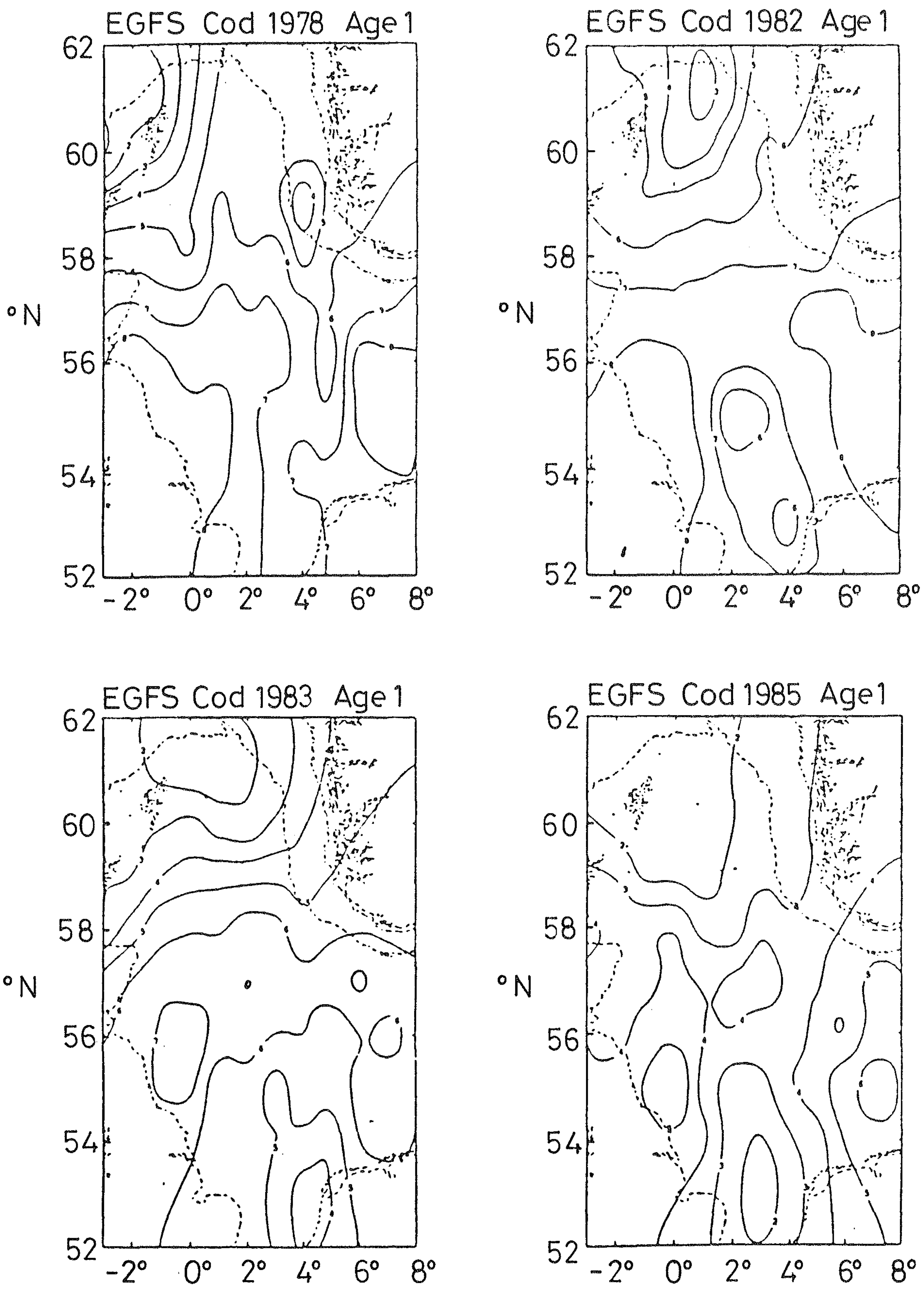


Figure 2.4.6 Same as Figures 2.4.4 and 2.4.5, but for parameter value equal to 1.0.

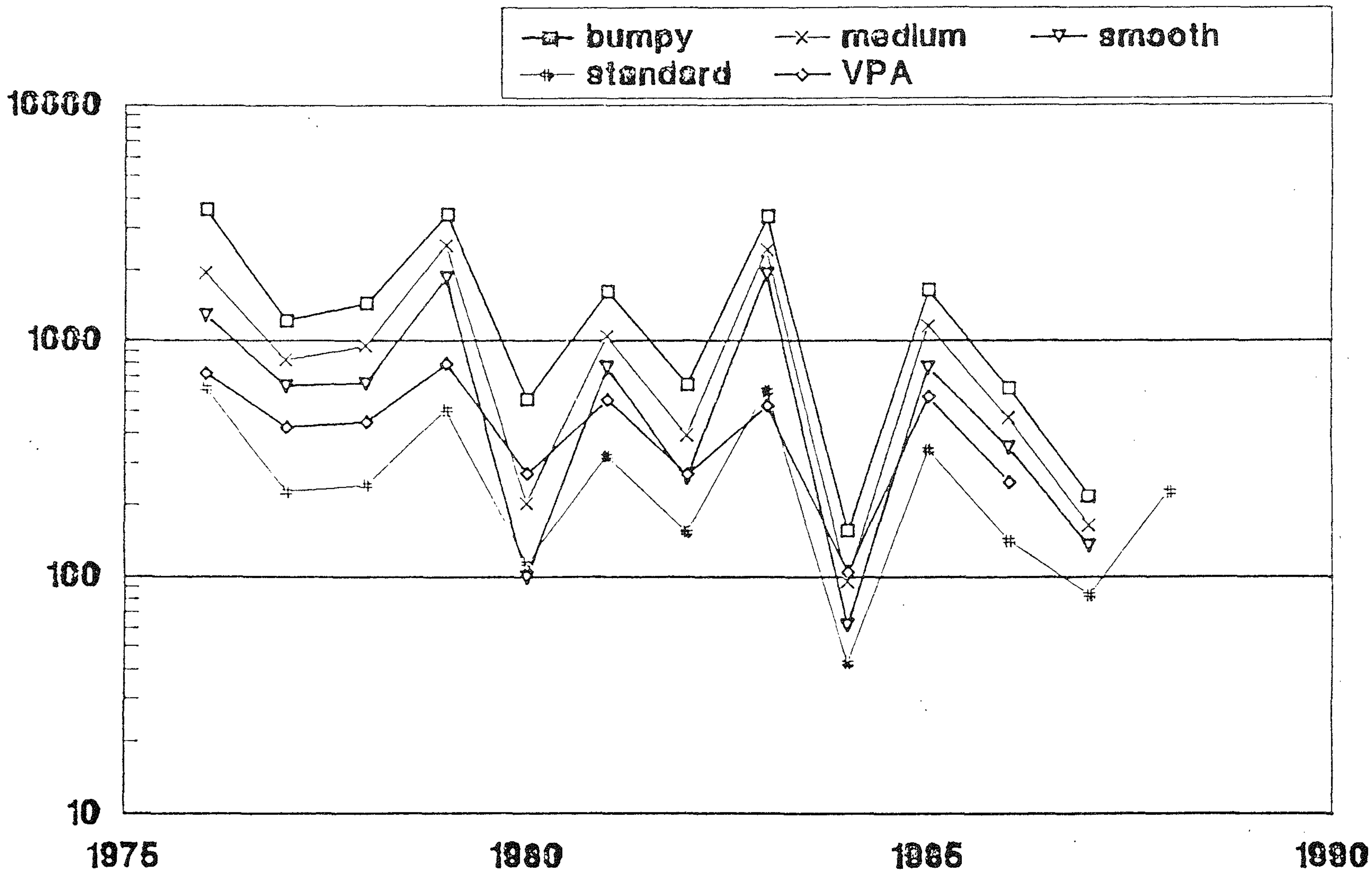


Figure 2.4.7 Estimation of the log abundance of COD through E.G.F.S. using various parameters in the empirical interpolation technique.

APPENDIX A

LIST OF WORKING PAPERS

1 - A. LAUREC, M.NICHOLSON, K.STOKES, J.G POPE, J.G SHEPHERD and C.DARBY - Processing of research survey data to build an annual index of abundance.

2 - J.G SHEPHERD - Indices of abundance derived from numerical integration over smoothed grid-point representations of survey data.

3 - R.A MYERS and K.STOKES - Density-dependent habitat utilization of groundfish and the improvement of research surveys - ICES C.M. 1989/D:15.

4 - R.J CONSER and J.E POWERS - Extensions of the ADAPT VPA tuning method designed to facilitate assessment work on tuna and swordfish stocks - ICCAT SCRS/89/43.

5 - Anon - Report of the workshop on spatial statistical techniques , Brest, May 1989 - ICES C.M. 1989/K:38.

6 - P. NEAL - Comparison of estimation methods under different levels of noise.

7 - G. GUDMUNDSSON - Comment on the report of the workshop in Reykjavik 1988.

8 - G.STEFANSSON - A statistical analysis of icelandic trawler reports, 1973-1987 - ICES C.M. 1988/D:13.

9 - M.D NICHOLSON, K.STOKES and A.B THOMPSON - Analysis of English groundfish survey data.

10 - H. SPARHOLT - Estimates of cod distribution in the Baltic based on YFS data.

11 - J.G POPE - Changing the timing of TAC's : the cost of precision.

12 - J.M HOENIG and C.M HEYWOOD - Using model predictions as an auxilliary variable to reduce variance in a fishery survey - ICES C.M. 1989/D:9.

13 - A. LAUREC and J.B PERODOU - Statistical and computational aspects of the analysis of fishing power and apparent abundance - translation of ICES C.M. 1987/D:9.

APPENDIX B

Use of diagnostics of ad-hoc Tuning

The Lowestoft + ICES VPA programs both produce an output file of diagnostic information, which should always be studied with care. The following notes may assist in the use of this information. The numbered points refer to the example output in Table B1. Note that the current version (2.1, May 1988) has a few minor peculiarities in the output labelling, which will be corrected in due course, but are explained so far as possible below :

1 - at 1a and 1b information on the options selected is printed. Note that :

a - the program prints 'Hybrid" if explanatory variate "Time" is chosen, but this may be modified by the decision to fix catchability on certain fleets. Thus, (as in this example) a "Mixed" method may be used, and the label "Hybrid" is misleading.

b - Conversely, the choice of no explanatory variate over-rides that to allow trends on some fleets. Thus the Laurec-Shepherd method is obtained (as indicated) and the "Terminal q estimated from trend" labels are misleading in this case.

c - This imperfection in the logic and labelling can be confusing. The definitive check is to examine the "Slope" column in the detailed Tables. If this is zero, no trend has been used; if it is non zero then a trend has been used.
d - The output "Fleets combined by variance" actually means that the inverse variance is used as the weighting factor.

2 - The VPA Table printed here is not the final VPA, but the last but one. If it is given to permit a quick evaluation of the results (e.g. stupid results indicating that something has gone wrong). It also permits a check on the convergence of the iterative procedure, since the values given for the final year can be checked against the final overall mean estimates (Fbar) in the detailed Tables, which are the ones passed to the final VPA output tabulation module.

Any significant difference (more than 0.01, say) indicates a probable failure to converge, so that the results are dubious.

3 - For each age, a complete set of log catchability estimates are printed, in a format designed to permit easy plotting. These should be examined carefully before accepting the results, preferably by plotting and any outliers or apparent trends should be examined.

This tabulation will probably be repaced by a Table of residuals in the future (this should be easier to understand). Note that it is possible (and desirable) to analyse these for time trends, even if no trend has been fitted, so that one may consider deliberately whether or not there is any evidence that the constant catchability assumption is untenable (c.f. methods used by the Irish Sea WG). The constant catchability assumption should, of course, be maintained for as many fleets as possible unless there is strong evidence against it.

Note also that it is possible (and desirable) to plot catchability against age : this may indicate whether or not the choice of the ratio determining the value of F on the oldest age is satisfactory or not (a strong trend with age on all fleets may be a counter-indication, but the question is difficult to decide).

4 - The final overall weighted mean estimate of F (Fbar) is printed together with two estimates of its log standard error (which is a good approximation to the fractional coefficient of variation). The SIGMA (overall) value is the final estimate of the precision of the analysis. If this standard error is large (say greater than 0.3, corresponding to a 30% coefficient of variation) for important age groups, then the results may be inadequate and the assessment should be considered dubious.

The other two estimates are the internal estimate SIGMA (int) and the external estimate SIGMA (ext). SIGMA (overall) is just whichever the larger of these. These quantities seem not to be well-known, but are well described by Topping (1962) pages 91-93.

The internal s.e. is based solely on the previous estimates of the standard errors i of the n individual estimates.

$$1/\sigma^2 \text{ (int)} = \sum_i 1/\sigma_i^2$$

It corresponds to the "Within samples" variance in a one-way ANOVA. The external s.e.

$$\sigma^2 \text{ (ext)} = \left(\sum_i (F_i - \text{Fbar})^2 / \sigma_i^2\right) / \left((n - 1) \sum_i 1/\sigma_i^2\right)$$

takes account of the actual scatter of the individual estimates Fi of log F about the weighted mean and corresponds to the between samples variance.

If these estimates differ very much, this indicates a discrepancy between the individual estimates (failure of the error bars to overlap). The variance ratio σ^2 (ext) / σ^2 (int) is output and may be tested as an F-statistic with n and n - 1 degrees of freedom. Values exceeding about 3 imply that the different series give conflicting results. Values less than 0.3 imply a suspicious degree of concordance (this may happen if trends are allowed for too many fleets).

5 - The individual estimates of total international F derived from each fleet (which are combined to give the final estimate) are given in the Table. Discrepant estimates may thus be identified . The log catchability estimate used (Pred.q) is also given. This is just the mean if catchability has been held constant. Its log standard error (SE(q)) is also given, and is the basic indicator of the quality and utility of the individual data series. Values greater than about 0.3 indicate substantial errors and values greater than 0.5 are seriously imprecise.

6 - When a trend has been fitted, the slope and its standard error are output. Only slopes which have the same sign and exceed (say) twice the standard error consistently on most age groups should be considered significant.

It is not suggested that the diagnostic output presently provided is ideal or complete, but it is sufficient for most serious problems to be identified and it should always be examined carefully.

VPA Version 2.1 - May 1988
N.SEA COD,INDEX FILE,UNSEXED,PLUSGROUP, H.C.LANDINGS
with cpue data from file cod4zefs.dat
DISAGGREGATED Qs
LOG TRANSFORMATION

Table B1

1a

Explanatory variate TIME
Fleet 1 ,SCOGFS , has terminal q estimated as the mean
Fleet 2 ,SCOTRL , has terminal q estimated from trend
Fleet 3 ,SCOSEI , has terminal q estimated from trend
Fleet 4 ,SCOLTR , has terminal q estimated from trend
Fleet 5 ,SCONTR , has terminal q estimated from trend
Fleet 6 ,ENGTRL , has terminal q estimated from trend
Fleet 7 ,ENGSEI , has terminal q estimated from trend
Fleet 8 ,INTGFS , has terminal q estimated as the mean
Fleet 9 ,NETGFS , has terminal q estimated as the mean
Fleet 10 ,ENGGFS , has terminal q estimated as the mean
FLEETS COMBINED BY ** VARIANCE **
Terminal Fs estimated using Hybrid method

1b

Regression weights
, 0.100, 0.200, 0.300, 0.400, 0.500, 0.600, 0.700, 0.800, 0.900, 1.000
Oldest age F = 1.000*average of 5 younger ages. Fleets combined by variance of predictions
Fishing mortalities

2

Age	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987
1	0.093	0.119	0.113	0.115	0.184	0.134	0.185	0.133	0.206	0.149
2	1.022	0.844	0.990	1.007	1.009	1.106	1.036	1.025	0.900	0.793
3	0.849	0.965	0.964	0.979	1.240	1.144	0.998	1.034	0.980	1.486
4	0.750	0.604	0.723	0.712	0.812	0.860	0.793	0.756	0.859	0.865
5	0.879	0.732	0.600	0.687	0.804	0.778	0.749	0.683	0.757	0.862
6	0.696	0.567	0.657	0.646	0.899	0.807	0.795	0.667	0.830	0.731
7	0.712	0.627	0.767	0.810	0.732	0.758	0.759	0.706	0.771	0.954
8	0.683	0.570	0.790	0.623	0.872	0.748	0.849	0.786	0.864	0.767
9	0.744	0.620	0.707	0.696	0.824	0.790	0.789	0.719	0.816	0.836

Log catchability estimates

Age 1

3

Fleet	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987
1					-17.87	-17.82	-17.54	-18.28	-17.68	-17.98
2	-18.65	-18.61	-19.06	-18.38	-17.30	-17.69	-17.43	-16.97	-17.56	-16.47
3	-17.80	-17.43	-18.80	-17.92	-17.07	-17.60	-17.12	-17.53	-17.21	-17.59
4	-17.21	-17.58	-18.37	-17.99	-16.76	-17.31	-17.21	-17.41	-17.07	-17.01
5	-19.43	-19.73	-20.58	-19.25	-18.86	-18.53	-19.13	-17.85	-19.07	-19.00
6	-18.17	-18.63	-18.63	-18.63	-17.43	-18.40	-17.37	-17.20	-16.31	-18.24
7	-17.97	-17.29	-17.73	-18.38	-17.60	-17.83	-17.65	-18.11	-17.88	-17.99
8	-16.90	-17.22	-17.26	-17.93	-17.46	-17.64	-16.92	-17.99	-16.93	-16.59
9			-14.98	-15.15	-15.27	-15.91	-14.84	-16.60	-15.05	-15.04
10	-16.33	-16.32	-16.14	-16.57	-16.21	-16.26	-15.53	-16.42	-16.22	-16.12

SUMMARY STATISTICS

6

5

Fleet	Pred. q	SE(q)	Partial F	Raised F	SLOPE	SE Slope	INTRCPT	SE Intrcpt
1	-17.87	0.252	0.0000	0.1668	0.000E+00	0.000E+00	-17.871	0.107
2	-16.72	0.376	0.0015	0.1160	0.244E+00	0.542E-01	-22.816	1.201
3	-17.32	0.367	0.0126	0.1959	0.618E-01	0.530E-01	-18.861	1.174
4	-17.01	0.321	0.0144	0.1490	0.914E-01	0.463E-01	-19.292	1.026
5	-18.61	0.557	0.0015	0.2201	0.108E+00	0.804E-01	-21.304	1.780
6	-17.22	0.701	0.0087	0.4108	0.148E+00	0.101E+00	-20.922	2.241
7	-17.97	0.228	0.0019	0.1522	-0.283E-01	0.330E-01	-17.260	0.729
8	-17.25	0.421	0.0000	0.0776	0.000E+00	0.000E+00	-17.246	0.165
9	-15.38	0.562	0.0000	0.1060	0.000E+00	0.000E+00	-15.383	0.226
10	-16.17	0.235	0.0000	0.1411	0.000E+00	0.000E+00	-16.173	0.092

4

Fbar	SIGMA(int.)	SIGMA(ext.)	SIGMA(overall)	Variance ratio
0.149	0.105	0.997E-01	0.105	0.735

APPENDIX C

North-east arctic cod assessment

A problem referred to the Working Group by the General Secretary relating to the recent assessment of the North-East Arctic cod stock and the revision to the assessment by ACFM was considered by a small sub-group.

The problem concerns the applicability of conventional methods for the tuning of VPA when a stock migrates between areas, and these are not fully covered by surveys and/or the commercial fisheries.

If the pattern of migration were consistent from year to year, and the tuning of surveys and fisheries were likewise consistent, there would be no difficulty and normal tuning methods may be applied.

If however the migration pattern changes from year to year, the survey / CPUE indices will not correlate so well with VPA, and substantial errors in the estimates for the final year may occur, particularly if an abnormal migration pattern occurs. This is clearly an extreme case of the space / time interaction problem discussed in section 2 which also affects the quality of the correlation between indices and VPA.

In this case, however, the Arctic Fisheries Working Group had attempted to overcome the problem by constructing a combined CPUE index. In principle, if such a combination can be carried out correctly, i.e. with proper standardisation before combination, and use of an integration-type method, this could overcome the problem.

However, the choice of proper standardisation is itself a difficult problem, particularly for commercial CPUE, and an inappropriate method may fail to give any improvement, or even make matters worse.

In general, therefore, continued use of separate data series in the tuning procedure is probably safer than ad hoc combination procedures, particularly since if the population shifts from one area to another in the final year, the errors in the conflicting survey results will partially cancel in the final weighted average estimates. The higher retrospective errors will also tend to result in more equal weights being given to the different index series, and not to excessive concentration on any one dataset.

The Working Group did not have all the necessary data available nor sufficient time to consider this particular problem in more detail, but in general considers that the procedure adopted by ACFM is preferable. The Arctic Fisheries Working Group should undertake a critical analysis of the tuning diagnostics, seeking in particular high variance ratios (indicating discrepant results from different data series), anti-correlated residuals and the performance (residual and prediction standard errors) obtained for various relative weightings in a combined index (compared with that obtained by the standard method), before deciding which procedure to use in future.

A first attempt at extending the theory to permit the determination of the differing catchabilities in different areas is described in Appendix D, but no operational method based on these ideas is yet available.

APPENDIX D

Estimating the proportion of a stock in each area (region) when the regional catchabilities are unknown but the stock redistributes in space over time : the case where mortality occurs.

J.M. HOENIG

INTRODUCTION

Suppose a stock is distributed over two regions and the gear efficiencies in the two regions are unknown but presumably different. Then interpretation of CPUE values is difficult. This is because the proportion of the stock in each region is unknown; hence, it is not known to how much of the stock an observed proportional change in CPUE applies.

If the population redistributes itself over space from time to time (i.e. if there is net movement between the regions), then it is theoretically possible to estimate the proportion of the population in each region at each time and also to estimate the ratio of gear efficiencies or catchabilities. The theory was developed by Heimbuch and Hoenig (1989) for the case where no mortality occurs between sampling times. In what follows I derive estimators for the case where mortality occurs between sampling times.

THE MODEL

We assume that the expected catch per tow is proportional to the population size in the region. Let Y_{ij} be the expected catch per unit of effort in region i at time j. $Y_{i,j}$ is assumed given by :

$$Y_{i,j} = q_i \; P_{i,j} \, N_j \qquad (1)$$

Where q_i = region - specific catchability

$P_{i,j}$ = proportion of population at time j which is present in region i

N_j = population size at time j

We note that $P_{2,j} = 1 - P_{1,j}$ so that two parameters can be eliminated; only $P_{1,1}$ and $P_{1,2}$ will be kept and noted respectively P_1 and P_2 for simplicity. Also, N_2 can be written as :

$$N_2 = N \, S \qquad (2)$$

Where N is the initial population size and S is the survival rate (proportion) between sampling times.

We also need to assume that between sampling dates the proportion in each region changes.

DEVELOPMENT OF THE ESTIMATORS

Let C_1 be the proportional change in CPUE in region 1 between times 1 and 2. Then C_1 is :

$$C_1 = \frac{Y_{12} - Y_{11}}{Y_{11}}$$

From (1) and (2), C_1 can be seen to be :

$$C_1 = \frac{q_1 \; P_2NS - q_1P_1N}{q_1P_1N}$$

$$C_1 = \frac{P_2S - P_1}{P_1} \qquad (3)$$

Similarly, C_2 is :

$$C_2 = \frac{Y_{22} - Y_{21}}{Y_{21}} = \frac{q_2 \; (1-P_2)NS - q_2(1-P_1)N}{q_2(1-P_1)N}$$

$$C_2 = \frac{(1-P_2)S}{1 - P_1} - 1 \qquad (4)$$

Hence, assuming S is known, we can solve (3) and (4) for P_1 and P_2 :

$$P_1 = \frac{S - C_2 - 1}{C_1 - C_2}$$

$$P_2 = \frac{P_1 \; C_1 + P_1}{S}$$

The ratio of catchabilities can be found by :

$$Y_{11} = q_1 \; P_1 \; N$$

$$Y_{21} = q_2 \; (1-P_1) \; N$$

Definition of Y_{11} and Y_{21}

Hence

$$\frac{Y_{11}}{Y_{21}} = \frac{q_1 \; P_1 \; N}{q_2 \; (1-P_1)N}$$

$$\frac{q_1}{q_2} = \frac{Y_{11} \; (1-P_1)}{Y_{21} \; P_1} \qquad (5)$$

A combined index of abundance, expressed in terms of the unknown but fixed catchability in region i, is given by :

$$R^{(i)} = \frac{Y_{ii}}{P_1}$$

This index gives the catch rate that would be expected in region i if the entire population were in region i. Note that if P_1 (the proportion in region 1) does not vary from year to year, then there is no need to adjust the observed catch rates by dividing by P_i.

DISCUSSION

If the externally obtained estimate of survival S is of maximum likelihood, then it can be shown that under quite general conditions the estimates of P_1, P_2 and q_1/q_2 are maximum likelihood estimates.

Preliminary results by Heimbuch and Hoenig (1989) suggest that the method is not sensitive to errors in S when the change in proportion in a region between times is large.

Further work on this methodological approach is under way.

APPENDIX E

COMBINING VARIOUS ESTIMATORS OF THE SAME QUANTITY

INTRODUCTION

Tuning methods often create situations where various estimators of the same quantity have to be combined. This is, for instance, the case with the ad-hoc tuning methods, when various fleets offer, for each of them, an estimate of the terminal F.

At present time, the combination is performed using a weighted mean, the weights being reciprocal to the estimated variances. Such a procedure is fully satisfactory when the estimation errors associated with the various fleets are not correlated.

When non negligible correlations between the estimation errors exists, other combinations may be more efficient. If some indices are highly correlated, they correspond to partially redundant information. Neglecting this fact will lead to an excessive weight being given to those indices. This is the question addressed in this Appendix.

I - NOTATIONS

N unbiased estimators $\hat{x}_i$ (i = 1,..N) of the same quantity x will be considered.

The variance-covariance matrix of the estimation errors

$(\hat{x}_i - x)$ is known and called C. $c_{i,i}$ is thus the variance of the estimator $\hat{x}_i$. D is the inverse matrix of C : $D=C^{-1}$.

II - MAXIMUM LIKELIHOOD APPROACH

The $(\hat{x}_i)$ vector will be considered as associated with a multivariate normal distribution, the variance-covariance matrix C being known. The log likelihood function of the vector (x_i) is equivalent to :

$$\emptyset = \sum_{i,j} (\hat{x}_i - x)\, d_{i,j}\, (\hat{x}_j - x) = \sum_{i,j} \hat{x}_i \hat{x}_j d_{i,j} - x \sum_{i,j} d_{i,j}\, (\hat{x}_i + \hat{x}_j)$$
$$+ x^2 \sum_{i,j} d_{i,j}$$

Differentiating $\emptyset$ with respect to x leads to :

$$\frac{d\emptyset}{dx} = 2x \sum_{i,j} d_{i,j} - \sum_{i,j} d_{i,j} (\hat{x}_i + \hat{x}_j)$$

since $\sum_{i,j} d_{i,j} (\hat{x}_i + \hat{x}_j) = 2 \sum_{i,j} \hat{x}_i \, d_{i,j}$

setting $\dfrac{d\emptyset}{dx} = 0$ is equivalent to :

$$x \sum_{i,j} d_{i,j} + \sum_i \hat{x}_i \left(\sum_j d_{i,j}\right) = 0 \quad \text{so that}$$

$$\hat{x} = \sum_i \hat{x}_i \frac{d_{i.}}{d_{..}} \quad \text{if } d_{i.} = \sum_j d_{i,j} \quad \text{and} \quad d_{..} = \sum_i d_{i.}$$

III - MINIMUM VARIANCE ESTIMATOR

Among the estimators defined as weighted means of the $\hat{x}_i$, the minimum variance one will be defined as :

$\sum_i \lambda_i \hat{x}_i$ with $\sum_i \lambda_i = 1$, the λ_i being the weighting coefficients.

The corresponding variance is trivially equal to : $\emptyset = \sum_{i,j} \lambda_i \lambda_j c_{i,j}$

In order to take into account the constraint $\sum_i \lambda_i - 1 = 0$ a Lagrange multiplier L will be introduced so that one will minimize :

$$= \emptyset - L \left(\sum_i \lambda_i - 1\right)$$

which is a function of the N (λ_i) and L

$$\frac{d\emptyset}{d\lambda_i} = 2\, c_{i,i} \lambda_i + \sum_{j \neq i} \lambda_j \, c_{i,j} - L \quad (\text{for } i = 1 \text{ to } N)$$

and $$\frac{d\emptyset}{dL} = \sum_i \lambda_i - 1$$

Putting the N first derivatives equal to zero leads to the system :

B.U = A expressed with matrices where $b_{i,j} = c_{i,j}$ for $i = j$ and

$b_{i,i} = 2\, c_{i,i}$ while $u_i = \lambda_i$ and $a_i = L$

The solution is given by : $V = B^{-1} A$

If $E = B^{-1}$ this leads to : $V = E\,A$

$$\lambda_i = L \sum_j e_{i,j} = L\, e_{i.} \text{ if } e_{i.} = \sum_j e_{i,j}$$

Because of the constraint $\sum_i \lambda_i = 1$, finally $\lambda_i = \frac{e_{i.}}{e_{..}}$

The optimum weighted mean will be given by $x = \sum_i \frac{e_{i.}}{e_{..}} x_i$

It can be noticed that this estimator will generally be different from the maximum likelihood one. However, they will coincide when no covariance exists, leading to the classical weighting by $\frac{1}{c_{i,i}}$

IV - POTENTIAL USE

The suggested estimators have not been used for any real calculation. Since, in practice, the error variance covariance matrix C will be generally unknown, it will have to be estimated. With the adhoc tuning techniques, this could be possible, using the empirical covariance between residuals, just in the same way as variances are presently being estimated. This should be tested for both estimation procedures (maximum likelihood and minimum variance), priority being given to the second one.

It would be dangerous, however, to recommend to replace the usual weighting procedures. The estimators suggested here are not likely anyway to permit major improvements unless important correlations exist. The residuals used for calculating the error covariance matrix correspond to a fitted model. If this model is unsatisfactory, the results of using the more complex weighting procedures are unpredictable. A careful examination of the residuals will always be necessary before a weighted procedure is used.

APPENDIX F

COMMENTS ABOUT RANDOMNESS

Within an area the real unknown fish density is a function ϕ of x and y. The total number of fish over the area S is $A = \iint_S \phi(x,y)\, dx\, dy$.

Setting aside discussion about catchability, the basic problem is the estimation of A from some observed values $\phi(x_o, y_o)$ (o = 1, ... No). All estimates are generally associated with an error. Many confusions arise from the fact that this error can be considered as random, without a clear view of what "random" means.

For a given density function $\phi(x,y)$, if the location of the sample is chosen with some level of randomness, all observed values $\phi_o = \phi(x_o, y_o)$ have a random component.

If all locations within area S have equal "chances" to be sampled, the probability distribution of the random variable ϕ_o just corresponds to the distribution of the existing (deterministic) densities. In other words, within a spatial discrete world, the histogram (descriptive statistics) of the density values over the area S would just coincide with the frequency distribution (probabilities statistics) of the random variable associated with an individual observation, in a location taken at random.

If some regions within the area S have a higher probability of being sampled than others, this implies that the random distribution associated with an individual observation will correspond to another probabililty density (or histogram) than the previously mentioned one.

If the stations are uniformly taken at random, but if the densities are not exactly measured, some errors being added, this will create another discrepancy between the distribution of the $\phi(x,y)$ and the probability density of the observed values. This will appear for instance if some subsampling is taking place in a survey, so that the exact number of fish caught in a haul for a given species (and maybe a given length or age) is not exactly known, but estimated through a sampling of the catches. A new random component will appear, adding a new variance, and distorting the probability density. Whenever one speaks of the random component, he should know exactly what he is refering to. This is also true for the associated variances.
Many formulas exist for calculating variances. If the user does not realize to which random component they refer, severe mistakes can be made.

The more difficult aspect concerning the various sources of randomness corresponds, may be, to the concept of underlying stochastic process referred to by kriging and other techniques. If for a given moment the fish density is a given deterministic function of latitude and longitude, these methods will consider an underlying stochastic process $\phi_{\omega}(x,y)$. In other words the existing spatial pattern corresponds to a situation (a realization of the stochastic process) which is just one among others that could have occured. At a given location (x,y) there exists an underlying stochastic variable, the whole set of the random variables associated with the various locations defining the previously-mentioned stochastic process. These techniques are constantly shifting from the random distributions at a given location over random events, to the distribution of the existing values over space. This is related with what statisticians call ergodicity. Some interesting relationships may exist between the two distributions and if, for a given realization, the location of an observation is taken at random (which corresponds to a special component of randomness), the two probability densities may even coincide (this corresponds also to what statisticians call strict stationarity). Such results are very important, delightful for theorists, but may disturb the users when they do not accept the fact that the word "random" may correspond to various phenomena, and often to combinations of them.

APPENDIX G

CONDITIONS OFFERING STATISTICAL OPTIMALITY TO THE ARITHMETIC MEAN.

Within an area S the fish density is a function Ø of x and y. Since it is impossible to measure this density at any place, a sample of stations {(xo,yo)}(O = 1, ...No) will be taken, from which an estimation of

$\iint_S$ Ø (x,y) dx dy will be built. The word "estimation" has so far no precise probabilistic meaning. It just implies that some error may be, and generally will be made.

Instead of estimating the integral over the area S, if the total surface of S is known, one can try to estimate the average density Ø.

If sample locations are taken according to a Simple Random Sampling (SRS) scheme this implies that all locations have equal probabilities of appearing in the sample, and that they are taken independently from one another. This case leads to a classical statistical problem : the estimation of the statistical mean (<u>over random events</u>) which corresponds to the average <u>over space</u> (see Appendix F if necessary).

The arithmetic mean over the observed values is the "best" possible estimation in various statistical meanings when :

- the distribution of the densities Ø(x,y) is normal,
- no measurement errors (see appendix F) interfere,
- sampling is by a S.R.S. scheme.

If this is not the case, other estimators than the arithmetic mean can show advantages.

NORMALITY

If the distribution is not normal, but corresponds to another (known) type of distribution, statistical methods can provide other estimators than the simple arithmetic mean. They can imply the use of some transformation of the observed apparent densities. However, due to the nature of the practical problems, it is always the arithmetic mean of the untransformed values which is being estimated.

For instance log normal distributions associated with transformations from Ø (x,y) to Ψ(x,y) = log Ø (x,y) are commonly refered to. The average theoretical value of Ø and Ψ are $\bar{Ø}$ and $\bar{\Psi}$.

If a SRS scheme is kept, the <u>probabilistic</u> distribution of an observed value Ψ(xo,yo) will be normal. The associated mathematical expectation is well known as different from log($\bar{Ø}$). In fact the variance of Ψ(xo,yo) being σ^2, it can be shown that

$$\bar{\emptyset} = \exp\left(\bar{\Psi} + \frac{\sigma^2}{2}\right)$$

Over a set of N observations, within a SRS scheme, an arithmetic empirical average $\hat{\bar{\Psi}}$ can be calculated, the variance of which

is equal to $\frac{\sigma^2}{N}$ so that the expectation of exp $(\hat{\bar{\Psi}})$ is equal to

$$\exp\left(\bar{\Psi} + \frac{\sigma^2}{2N}\right)$$

An unbiassed estimator of $\bar{\emptyset}$ is given by

$$\exp\left(\hat{\bar{\Psi}} - \frac{\sigma^2}{2\,N} + \frac{\sigma^2}{2\,N}\right) = \exp\left(\hat{\bar{\Psi}} + \frac{\sigma^2}{2}\left(\frac{N-1}{N}\right)\right)$$

Over a S.R.S. when σ^2 is not known, an estimator s^2 can be built based on the usual formulas, leading to the estimator

$$\exp\left(\hat{\bar{\Psi}} + \frac{s^2}{2}\left(\frac{N-1}{N}\right)\right)$$

The correction factor for the bias takes into account the estimated variance. It is sometimes called the variance correction, sometimes the bias correction (unfortunately!). Before using the above correction formula, one should consider three points :

1 - When s^2 is an unbiassed estimator of σ^2, $\exp\left(\frac{s^2}{2}\left(\frac{N-1}{N}\right)\right)$

will be a biassed estimator of $\exp\left(\frac{\sigma^2}{2}\left(\frac{N-1}{N}\right)\right)$

This can be however be corrected through some mathematical calculations (Laurent, 1963).

2 - The correction formulas, including the improved ones, hold only if the distribution of is really normal. Robustness problems will often occur in practice (Myers and Pepin, 1987). This question is much more important than the previous one.

3 - The estimation of the variance requires a high enough sample size. This corresponds to a well known statistical problem making any quantities refering to squared values more difficult to estimate that means. It is also related with the robustness problem mentioned above, since what matters is the distribution of $\hat{\bar{\Psi}}$, which due to the central limit theorem, may be close to normality when N is large enough.

ADDED (MEASUREMENT) ERRORS

Such errors arise from other sources of uncertainty than spatial variability. They may correspond to varying variances from one observation to another one. In such a case a weighting can improve the final estimator. However, these weighting factors must consider the overall variances (sampling + measurements), which can be difficult to evaluate, and ideally be independent from local densities. To illustrate this second question one may consider situations where high catches will be subsampled, while smaller ones will be exhaustively numbered. In such a case the measurement (= subsampling here) errors will be higher for higher densities. Down weighting the corresponding values will introduce biases. This does not imply that the benefits obtained in terms of reduced variance by the weighting should be neglected. But a great care is required. A fully satisfactory statistical treatment requires in fact a complete knowledge of the underlying stochastic model.

SAMPLED LOCATIONS

These locations will not be necessarily chosen according to a SRS scheme. In fact, in practice SRS will often be impossible. If the probability of any location being sampled is known a priori, a correction is possible. It may not be necessary if, within the area, the density does not show a strong pattern (if it is almost a white noise).

A departure from the SRS hypothesis will also occur when the haul locations are not chosen independently from one another. They can be either clustered, or systematically separated from each other. If the spatial distribution of the samples is really far from S.R.S., interpolating (see section 2.4) leads to better estimation than the sample mean. When using the simple arithmetic mean, it can be misleading to use the simple formulas for variances.

Applying the variance formulas corresponding to SRS will generally lead and to an underestimation in the first case and to an overestimation in the second one. Sampling variances are anyhow difficult to use in the context of monitoring year to year changes. It should be recalled, however, that if transformations are being used, realistic estimates of the variances may be necessary, if a bias correction is attempted.

APPENDIX H

Normalisation of the interaction

Consider data at three stations for 3 years giving differing results.

station	year 1	year 2	year 3
1	1	2	3
2	2	2	2
3	3	2	1

Station 1 shows a rising trend
Station 3 shows a falling trend
Station 2 shows a constant trend

Consider now linear modelling of these data, with station effects, year effects and the station/year interaction. Many normalisations are possible. Three common ones are :

1 - Normalisation to the mean (ANOVA)

Take row and Column Means

station	year 1	year 2	year 3	means
1	1	2	3	2
2	2	2	2	2
3	3	2	1	2
means	2	2	2	2

The model fitted is : constant = Grand Mean = 2
year effect = 0, 0, 0
Station effect = 0, 0, 0
Interaction = -1, 0, 1
0, 0, 0
1, 0, -1

2 - Normalisation to first effects (GLIM).

The first level effects are set to zero in both main effects and the interaction.

```
The model fitted is constant = 1
            year effect    = 0,  1,  2
            station effect = 0,  1,  2
            interaction    = 0,  0,  0
                             0, -1, -2
                             0, -2, -4
```

3 - Normalisation to last effects (SAS ?)

The last level effects are set to zero in both main effects and the interaction.

```
The model fitted is constant =  1
            year effect    =  3,  2, 1
            station effect =  3,  2, 1
            interaction    = -4, -2, 0
                             -2, -1, 0
                              0,  0, 0
```

note that :

a - The year effects are different in each case i.e

```
Model (1)   0, 0, 0
Model (2)   0, 1, 2
Model (3)   2, 1, 0
```

In the case of models 2 and 3, these correspond simply to the observed year effects at the first and last stations. All the rest of the data are thrown into the interaction term.

b - The interactions and station effects estimated are also different. Model 1 corresponds to the lowest mean square in the interaction term (this may be desirable).

c - If the interaction is supposed to be due to redistribution, the sum (integral) over space (stations) of the interaction should be constant (after retransformation if a transformation was used). This implies model (1) should be used.

In conclusion, if a spacetime interaction is to be fitted, the choice of normalisation of the interaction term is absolutely crucial and may lead to totally different results for the year effect estimated.

Therefore, one should either use minimum mean square interactions or space integral normalisation, or use the fitted model to estimate distributions for subsequent interpolation/integration.

The same applies to commercial CPUE data, when interactions of year with anything else have been included in the model.