Validating three methods for making probability statements in fisheries forecasts.

K.R. Patterson[1], R.M. Cook[1], C.D. Darby[2], S. Gavaris[3], B. Mesnil[4], A.E. Punt[5], V.R. Restrepo[6], D. W. Skagen[7] , G. Stefánsson[8],   M. Smith[2]

[1]FRS Marine Laboratory, P.O. Box 101, Victoria Road, Aberdeen AB11 9DB, Scotland, UK. email : pattersonkr@marlab.ac.uk. [2] CEFAS Laboratory, Lowestoft, UK. [3] Department of Fisheries & Oceans, Biological Station, St. Andrews, NB, Canada . [4]B. Mesnil: IFREMER, Nantes, France. [5]CSIRO, Hobart, Australia. [6]ICCAT, Madrid, Spain. [7]IMR, Bergen, Norway. [8]MRI and University of Iceland, Reykjavik, Iceland.

ABSTRACT

We report on a test of three methods that have been used by ICES in making probability statements about fisheries forecasts to management agencies (XSA/WGMTERM, ICA/ICP, and a stochastic projection method that was first applied to North Sea herring). The principle of the test is to calculate many stock assessments and forecasts using historic data, and to compare probability statements about some quantities (forecast biomass, relative biomass, catch etc.) with the latest and presumably more reliable estimates. For most ICES and NAFO stocks where a sufficiently long time-series exists, we calculate retrospective (8 year before present) assessments and then calculate medium term forecasts 5 years forwards. By comparing the frequencies of expected and actual outcomes, we conclude that some methods can be used to make reasonably reliable probability statements about relative biomass (e.g. biomass relative to biomass at some time in the past), but that for most other quantities and most methods the accuracy of probability estimates is very poor. In general, the methods tended to underestimate uncertainty and there was a relatively large proportion of "unexpected" outcomes in the forecasts.

INTRODUCTION

In principle, one could explore statistical properties of medium-term projection methods by comparing the frequency of occurrence of some events with the frequency with which they had been predicted to occur. However, this would rely on eventually knowing the true state of nature at some time after the forecasts had been made. An exploration similar in principle can be performed if one cares to make the assumption that when a significant number of new data are obtained, a new analysis will provide estimates of much higher precision and accuracy. This is essentially a test of the consistency of an assessment and forecasting protocol rather than a validation using precise and direct observations. The background and technical basis for this comparison is detailed below.

The use of medium-term predictions has become commonplace within ICES and other arenas. These are conducted using a variety of methods and have been developed by a number of groups within ICES and elsewhere. The purpose of the methodological development has been to define the techniques used, with an emphasis on what aspects of uncertainty to include and how to include it technically. Evaluation of the predictive capabilities of these simulations has, however not usually been addressed. Thus, even if confidence intervals include most sources of uncertainty thought to be important, it is not clear whether the

resulting confidence intervals cover the corresponding quantities with the given probability, or even whether the median is median-unbiased.

Evaluation of the performance of methods used to make probability statements does not seem feasible for single stocks in a limited time-scale because there are only a few outcomes that can be compared to predictions. However, we use a library of long time-series of data in order to examine whether 'unusual' outcomes (which prediction methods indicate to have low probability) do occur unexpectedly often.

The purpose of the following analysis is therefore to evaluate whether some confidence statements are correct and to identify how they fail and how badly they fail. We do not at this stage address the question of why prediction failures occur.

Given an uncertainty estimation procedure, one of the outputs is a set of confidence (or probability) intervals for several quantities. These are used for the purpose of predicting what the true values are or will be. Thus, the intent of the uncertainty estimation is to be able to make statements of the form

$$P[X<b]=\alpha$$

where α is an estimated probability, X is typically some future value of biomass or catch, and b represents the range of values that the future biomass or catch (X) may take. Obviously as b tends to zero, α also tends to zero, and as b tends to infinity, α tends to unity.

It is possible to wait until the actual event occurs and can be measured with higher precision and then record the probability with which it had been estimated that the observation of this value would have been exceeded. This procedure can be repeated many times using data for different fish stocks and different time periods. If the distribution of outcomes were accurately predicted by the uncertainty estimation method, one would expect that half of the outcomes would occur below the estimated 50th percentile and half above. Similarly, 10% of the outcomes would occur below the 10th percentile and 90% would occur above, and so for all other percentiles. One may therefore simply compare the distribution of the outcomes with the corresponding α_X values, and expect to find a linear correspondence between the two. Deviations from such linear correspondence can be used to detect deficiencies of the uncertainty estimating method.

METHODS

For the purpose of forecasting fish stock dynamics in the medium term, projection methods are typically used to estimate populations some 5 to 10 years ahead of the most recently estimated stock size. Here we take a 5 year time horizon as a period of interest for reasons of data availability, so that there are sufficient data to compare forecasts with outcomes. Performance of the methods has been investigated by comparing the predicted outcome conditioned on the observed fishing mortalities during the projection period with the 'observed' outcome. This is because it is not a part of the test to forecast the activity of fishing fleets. We denote by 'observed' outcome an estimate which is calculated with at least three years of additional data, and is believed to be estimated much more precisely.

The process for a data set of T years was as follows:

1. A stock assessment was calculated using the full time series of available information, the choice of assessment model formulation being guided by the choices made most recently by the relevant experts. Except for the last three years, estimates from this assessment were treated as precise "observations" against which the forecasting methods were tested.
2. A 'retrospective' stock assessment was calculated on T-8 years of data, with the same assessment model. In some cases, entire data subsets are removed because insufficient observations remain to allow estimation of some parameters, and minor modifications to the model structure become necessary on an *ad hoc* basis.
3. Beginning from the retrospective assessment, a five-year forecast was calculated with the estimated exploitation pattern in each year, scaled to the "observed" overall annual F in 1.

Where sufficient data were available, this procedure was repeated by making further retrospective replicates and repeating steps 2. and 3. above. For each dataset up to five such "replicates" were performed, subject to data availability.

Stock-recruit models are an essential feature of any projection method. In cases where no obvious stock-recruit dependency existed, an assumption of stock - independent recruitment over the range of observed stock sizes and an assumption of linear dependence of recruitment on stock size below this range was made- the so-called "Ockham" assumption , following e.g. Anon. (1997). Otherwise a Beverton-Holt model was fitted, or in some cases where there was significant precedent a Ricker model was used. Variability in recruitment about the stock-recruit relationship was in all cases modelled by nonparametric bootstrap with no autocorrelation, following Anon. (1993).

Evaluations have been calculated for those stocks for which a sufficiently long time-series of data of appropriate quality were available, and for which the data could be obtained conveniently in electronic form. The thirty-nine stocks used in the analyses are listed in Table 1, grouped by geographical areas. Most of these are from the ICES area, but four stocks from the Western Atlantic did have sufficiently long time-series to be included in the analysis. This list of stocks is not exhaustive, but is believed to include a sufficiently wide representation of fish species and areas to base general conclusions about the statistical behaviour of forecasting methods when confronted with real data.

We consider performance of the assessment models in terms of the accuracy with which uncertainty in the following seven parameters is characterised:

- Spawning stock size in the last year of the catch-at-age data (SSB_x)
- Spawning stock size in the first forecast year (SSB_{x+1})
- Spawning stock size in a five-year forecast (SSB_{x+5})
- The ratios of above three parameters to the estimated stock size five years before the last year of catch at age data: SSB_x/ SSB_{x-5} ; SSB_{x+1}/ SSB_{x-5} ; SSB_{x+5}/ SSB_{x-5}
- The catch in the fifth year of the forecast, C_{x+5}

We have included three methods in our tests. These are described below, together with specific details of their implementations.

Separable Models; ICA/ICP Implementation

The population model used is a separable catch-at-age model (Deriso et al 1985) as implemented by Patterson and Melvin (1992). Population parameters are found by nonlinear minimisation of terms of general form:

$$\sum_{y=Y}^{y=Y-6} \sum_{a=A}^{a=a_{max}} \left(\lambda_a \ln \left(\frac{C_{a,y}}{F_y S_a N_{a,y}} \right)^2 \right) +$$

$$\sum_{y=y_{min}}^{y=Y} \sum_{a=A}^{a=a_{max}} \sum_f \left(\lambda_{f,a} \ln \left(\frac{U_{a,y,f}}{\hat{U}_{a,y,f}} \right)^2 \right)$$

where C represents catches at age; U, survey indices of abundance; F, fishing mortality year effect, S, relative fishing mortality at age (ie selection pattern or partial recruitment) and N represents the mean annual population abundance. Subscript a represents age in the range from the youngest age amin to the

oldest age amax; subscript y represents years of the available data set (from ymin to ymax), and subscript f represents age-structured abundance indices. λ represents a "weighting factor" related to the reciprocal of the standard deviation of the observations with respect to the model-predicted values. Treatment of these weighting factors is explained below.

Weighting Factors

Although under maximum-likelihood considerations, the weighting factor λ represents a value proportional to the reciprocal of the standard deviation of the observations, in many cases in fisheries the available data series are short and the precision with which the standard deviation can be estimated is believed to be extremely low. In order to improve the stability of the estimates of the parameters of interest, these values are usually either replaced with conventional values or else estimated subject to more or less stringent constraints. Where a separable model had previously been used by a working group, then we have retained that working groups' assumptions about the values of λ. In other cases, we have applied the following constraints as a general rule:

Weighting factors for the catches at age were in most cases assigned fixed values: : $\lambda_{c,amin} = 0.1$, $\lambda_{c,amin+1} = 0.5$, $\lambda_{c,a>amin+1} = 1$. This represents the assumption that recruiting year-classes are poorly sampled in the catches. However, for cod and haddock in the North Sea, a choice of all $\lambda_c = 1$ was made to represent the belief that all ages are well-sampled in the catches.

Weighting factors for age-structured surveys were estimated subject to strong constraints:

$$\lambda_{f,a} = \begin{cases} \dfrac{1}{\sigma_{f,a}} & if \quad \dfrac{1}{\sigma_{f,a}} < 1 \\ 1 & otherwise \end{cases}$$

Where a separable model had previously been used by a working group, then that working groups' assumptions were retained. In other cases, some arbitrary selection constraint was chosen by visual inspection of the F-matrix as calculated by the relevant assessment Working Group.

Variability in the weights at age and in maturity were represented using a parametric bootstrap on lognormal and arcsine-transformed normal scales respectively, following Patterson and Melvin (1996). The estimates of mean and variance for the parametric bootstrap were derived from historic observations on the whole time-series of available observations for each stock.

Computation of the distributions of parameters of interest in the forecasts was made by a numerical delta method (Kennedy and Gentle, 1980), using the methodology described by Richards *et al.* (1998). One thousand random draws of vectors of population dynamics parameters (logarithms of population abundance at age, selection at age, fishing mortality) were taken with mean and covariance computed at the minimum of the weighted least squares term. Each of these was used to forecast stock and catch subject to fishing mortalities estimated in the assessment computed including the full time-series of data.

If a working group had a strong reason to use a particular stock-recruit model, then this choice was also used to model the dependency of recruitment on stock size in the projections. Where no conventional model could be used, a simple 'Ockham' model was chosen. Variability in recruitment was represented by nonparametric bootstrap on a log scale with no attempt to model autocorrelation.

Survivors Analysis; XSA/WGMTerm implementation

The XSA technique (Shepherd, 1999; Darby and Flatman 1994) is very widely used within ICES for stock assessment, and the WGMTerm methodology described by Anon. (1993) is also widely used as a stochastic

forecasting tool. As use of the methods is so widespread, we have used a slightly different approach to testing this methodology. Instead of applying a standard model structure to many datasets, we have instead adopted the model choices most recently made by the relevant experts in each case. Hence, where the previous test is a test of a rather standardised model application, here we are effectively exploring the statistical characteristics of both the model and its users.

Otherwise, the procedure outlined as above (Steps 1 to 3) was repeated in the same way as for the ICA/ICP assessments and forecasts. More technical detail of the forecasting approach is given below.

In order to investigate the cumulative distribution function of the interest parameters, 1000 replicates of the cohort terminal population numbers (Nmc) were drawn from

$$\text{Ln Nmc} \sim \text{Ln N} + D\,(0,\sigma^2)$$

where N is the XSA estimate of the survivors for a cohort. $\mathbf{D}(0,\sigma^2)$ represents a number drawn from the normal distribution with unit mean and variance σ^2 as derived from the XSA model. Within XSA two estimates of standard error are identified - "internal" and "external". To be conservative the larger of the two values at each age was adopted as the estimate of variance.

The Monte Carlo replicates of terminal population size were used to begin a backwards cohort analysis and initiate a forward projection. Weight and maturity at age for the back calculations were the data used within the appropriate assessment. Weight at age etc. for the forward projections were derived as the average of the final three years of assessment data, no stochastic variation is assumed in weights, maturities or fishing mortalities.

Recruitment at the youngest age of the forecast populations was simulated from the SSB estimated for the appropriate year and a parametric stock and recruit function fitted to the stock and recruit data pairs from the XSA assessment. Where the parametric model was considered inappropriate a Butterworth- Bergh (Ockham) model was fitted and used. Stochastic variation was introduced by sampling the residuals from the fitted model.

Separable Models with Harvest Control Rules; STPR Implementation

STPR is a pure prediction program, which takes input data from a separate assessment. The origin of the program was previous work on the effect of recruitment variation on the future development of the stock, when applying simple harvest control rules (Skagen, 1991). The present version (Skagen 1999) was primarily intended to explore harvest control rules in the medium term with respect to the robustness to uncertainty in present and future assessment and adherence to quotas as well average yield and year to year variation in the yield.

The STPR methodology has been used in providing advice on medium-term stock development for North Sea herring but is in principle easily applicable to other resources.

The input data are stock numbers at age, selection at age for the fishery and parameters for a parametric stock recruitment functions. The initial stock numbers at age can be stochastic assuming lognormal distributions, provided a variance - covariance matrix is included in the input, while the selection pattern is fixed. In the present runs, the data were taken from ICA runs, identical to those presented above. This included the variances-covariances for the stock numbers, but not for the other parameters (fishing mortality and selection at age).

The recruitment variation at given SSB around the value from the parametric function value is expressed by a multiplicative noise term, i.e. $R = R(SSB)*\exp(\varepsilon)$. The term ε can be expressed as an autoregressive process: $\varepsilon_y = \alpha_1\varepsilon_{y-1} + \alpha_2\varepsilon_{y-2} + \ldots\xi$, where ξ has a normal distribution with variance σ^2. The stock-recruitment functions used were generated on a spreadsheet by non-linear regression on the stock-recruit

pairs form the assessment, taking these values as error-free. In some cases, where the most recent recruitments seemed unreliable, those pairs were not used. Autoregressive terms were not included in the runs presented here. A truncation was applied: If $|\varepsilon_y|$ was larger than a maximum value, corresponding approximately to the largest value observed, this ε was discarded and a new one drawn.

Weights and maturities at age were taken randomly from historic values. This is done by drawing a year each time such data are needed, and using data at all ages from that year, to preserve correlations between ages.

The program projects stock trajectories forwards, with randomly drawn initial numbers, recruitments, maturities and weights as described above. The mortalities are computed according to the harvest control rule defined for the run. This is a rule for decision according to the SSB, where three different intervals of SSB with different rules are allowed. The rule can be either a catch constraint, an F constraint or a combination of these, and two fleets with separate harvest control rules are allowed.

The program allows for accounting for uncertainty in future assessments as decision makers see them, by allowing the SSB to which the rule is applied to deviate from the model value by a stochastic multiplier. Likewise, the realised catch can deviate from that determined by the rule. The multipliers are input by the user, and these options are only intended for testing the robustness of a HCR to errors in future assessments and in adherence to management regulations. However, these options were not used in this instance.

The present calculations were made to estimate probability coverage and used a simplified harvest control rule. The initial numbers were stochastic numbers at the start of the last assessment year (year 0). This was done because the variance-covariance matrix corresponded to these data. The initial number of the recruiting year class was substituted by a stochastic recruitment according to the stock-recruit function. Only one fleet was used, and no error in future assessment or adherence to assumed catches were included. A catch constraint, corresponding to the realised catch, was used for the first year, as is the standard procedure in this program. For the next 5 years, yearly F-constraints irrespective of SSB were applied. These were the realised F's, according to the most recent assessment. The SSB and catch in year 0 and year 5 for each replication were assembled on a separate file for each simulation, and results in fractiles were extracted from this file. A thousand replicates were made for each simulation.

RESULTS

Separable Models; ICA/ICP Implementation

The proportion of outcomes is compared with the ICA-ICP model predictions is tabulated in Table 2 and presented in cumulative plots in Figure 1. In order to remove the possible effects of correlation introduced by repeating assessments and forecasts on the same data sets, data including only one forecast from each stock are given in Table 3 and in Figure 2. We interpret these figures as follows :

(1) Figures 1 a, c, and g show a close linear correspondence between the distribution of outcomes and the distribution of the forecasts of those outcomes, for relative stock sizes. The probability coverage appears accurate for probability statements made about relative stock size, whether this is for the current year, one year ahead, or five years ahead.

(2) However, for statements about absolute stock sizes, Figures 1 b and d indicate that many of the outcomes occurred in the tails of the distributions of the forecasts. For example, Figure 1b indicates that 40% of stocks will be found to be smaller than a stock size which had been estimated as corresponding to a 20% probability level. In the upper tail about 10% of stocks will be found to be larger than a stock size which had been estimated as corresponding to a 99% probability level. This indicates that the method substantially underestimates uncertainty, and particularly underestimates the probabilities of low stock biomasses ocurring.

(3) There is also an indication of bias, and a tendency to overestimate stock size. More outcomes occurred in the lower tail of the distributions than the upper in figures 1b and 1d.

(4) In five-year forecasts of absolute biomass the probability coverage is more accurate, with fewer outcomes occurring in the tails of the distributions of the forecasts (Figure 1f). This indicates that the method performs better at characterising variability in recruitment (which will be a large part of the uncertainty in most five-year estimates of biomass) than at estimating uncertainty in stock biomass in the short term.

(5) Uncertainty in five-year catch forecasts is also underestimated (Figure 1g).

(6) Removing within-stock replicates does not change these perceptions very much (Figure 2). The graphs retain the same overall shape, though the patterns are less smooth.

(7) There is no obvious grouping of such pathological cases by area or species (Table 2).

Survivors Analysis; XSA/WGMTerm implementation

It has not yet proved possible to complete assessments and forecasts for all the stocks given in Table 1 using the XSA/WGMTerm implementation, and this is therefore a report of work in progress. Estimates are only available for a smaller subset of stocks in the North Sea and in the North-Western area (Table 4). This may be sufficient however to draw some preliminary conclusions about the performance of the forecasts. The proportion of outcomes is compared with the forecasts in cumulative plots in Figure 3.

We interpret that :

(1.) For all parameters, uncertainty is greatly underestimated. For both absolute estimates of stock size and for relative stock sizes, very many outcomes occur far in the tails of the forecast distributions.

(2.) The method appears to generate overly high forecasts of stock size and catches, as more outcomes fall in the lower percentiles of the forecasts. This effect appears especially strongly in the five-year forecasts, where about 50% of the outcomes fell in the lower 10th percentile of the forecasts.

(3.) The survivors method performs about equally poorly for absolute values of stock size as for relative values. This is quite different from the separable model where the distributions of relative stock sizes appear to be well estimated. We interpret this as due to differences in the structural model: in a VPA-based analysis, biomasses in earlier years become constrained due to 'VPA convergence' (Pope, 1972) Hence, forecasts of biomasses and of ratios of biomasses will acquire similar statistical properties. In a separable model, the estimates of biomass in earlier years are not so constrained (Kimura, 1990).

Separable Models; STPR Implementation

The proportion of outcomes is compared with the MC-Project-STPR forecasts on cumulative plots in Figure 4. The results for the years x and x+1 essentially show how this model receives the data from ICA, and would be expected to be similar from those obtained by ICA/ICP. STPR seems to perform quite well for the SSB in years x and x+1, and better than the ICA/ICP. The main differences between these two methods, in which the reason for the difference in performance between the methods may be sought, include:

− Fishing mortality in Year X is predicted with a catch constraint in STPR, but estimated from the data in ICA/ICP
− Weights are drawn by referring to a random year in STPR
− STPR does not take the covariances between stock numbers and selection parameters into account.
− STPR does not estimate uncertainty in year x-5. For calculation of ratios, the point estimate from ICA was used.

For the five year forecasts, STPR tends to underestimate the uncertainty, both with respect to overestimation and underestimation of SSB and catches. Again, this is in contrast to the ICA/ICP, where the uncertainty estimate improves at five year compared to the initial year.

DISCUSSION

In considering uncertainty estimation in fisheries, it should be remembered that any calculation of uncertainty is conditioned (or predicated) on a number of biological assumptions about mortality, recruitment, growth, catchability, selection and many other processes. Some of these assumptions are about the values that some parameters may take and other assumptions are about the structures and relationships of various processes. Whilst uncertainty in some of these features may be routinely measured using statistical tools or fish stock assessment software, it is clear that the measured uncertainty is only a part of the real uncertainty. "Real uncertainty" are the words we choose to describe the uncertainty in the eventual development of a fish population.

We have argued elsewhere (Patterson et al. in prep) that the appropriateness of conditioning choices is unknown and untested, and that the extent to which estimates of uncertainty made in stock assessments is an adequate representation of real uncertainty is presently unknown. This study represents a first step in examining whether the uncertainty estimates proposed by stock assessment scientists do represent the real uncertainties, and therefore risks, attached to exploiting fish stocks. In this evaluation we examined three practices which all used the delta estimation method but made different structural assumptions for the assessment model.

Our conclusion from this preliminary study is that in general terms the real risks are not well represented : the estimated uncertainty does not well reflect the eventual outcome in terms of the development of future stock size. In addition, the forecasts tend to be too optimistic and do not well represent the likelihood that very small stock sizes may be outcomes.

We summarise our views of the performance of the methods in estimating uncertainty in the various parameters in the text table below:

	SSBx/ SSBx-5	SSBx+1/ SSBx-5	SSBy/ SSBx-5	SSBx	SSBx+1	SSBy	Catch in 5th year
ICA/ ICP	Slight under-estimation of stock size	Acceptable	Acceptable	Under-estimates uncertainty	Under-estimates uncertainty	Slight over-estimation of stock size	Some under-estimation of uncertainty
XSA/ WGM	Under-estimates uncertainty	Under-estimates uncertainty Over-estimates stock size	Under-estimates Uncertainty Over-estimates stock size	Under-estimates uncertainty	Under-estimates uncertainty Over-estimates stock size	Under-estimates uncertainty Over-estimates stock size	Under-estimates uncertainty Over-estimates catch
STPR	Slight over-estimation of stock size	Slight over-estimation of stock size	Under-estimates uncertainty	Slight over-estimation of stock size	Slight over-estimation of stock size	Under-estimates uncertainty	Under-estimates uncertainty

Often in ICES, probability statements of the form

Probability that stock size two years ahead will exceed 800 000t = XX %

are provided in advisory documents. Unfortunately this sort of statement appears poorly supported by the methods at hand. Apparently, estimation of uncertainty in relative changes in stock size appears to be more consistently estimated.

We have not yet explored why the estimated uncertainty fails adequately to characterise real uncertainty in future stock sizes. This is a complex issue, as there is a large number of conditioning assumptions underlying each stock assessment, and it may be the case that a different conditioning assumptions may be inappropriate in different cases. It is also possible that factors exogenous to single-species assessment models may not be represented adequately in the stochastic simulation processes, such as long-term climatic effects introducing autocorrelations, multispecies effects introducing ecosystem changes, or data biases caused by inadequate catch reporting. However, it is extremely difficult to identify case-specific effects because of the few observations on any specific case.

The methods tested in this analysis involve a two stage estimation process followed by a simulation phase. Firstly, estimates of stock numbers, recruitment and fishing selectivity are obtained from a "stock assessment" model with their associated uncertainties. Secondly these estimates are used to derive a stock-recruitment model. Finally, the stock-recruit model and fishing selectivities are used to run a forward projection. This means that the performance of the projection phase is heavily dependent on the stock assessment used to derive the inputs. If the stock assessment phase results in biased initial conditions with poor estimates of uncertainty, the projection phase will obviously perform badly. In addition, of the methods considered here, the projection phases do not all use the same sources of uncertainty. ICA/ICP and STPR for example allow for many more sources of uncertainty, such as in weight at age and some parameter covariance estimates, while XSA/WGMterm does not. Thus it is perhaps not surprising to find that the latter method underestimates uncertainty in the medium term simulations.

All methods tested in this analysis used the numerical Delta method of estimating uncertainty in assessments but employed different structural assumptions for the assessment, stock-recruitment and projection models. Restrepo et al. (2000) found that results obtained with the numerical Delta, when not adjusted for bias, did not give accurate probability coverage while Gavaris et al. (2000) found that the magnitude of this bias was smaller for relative change in SSB with separable structural models. The use of methods which did not adjust for bias coupled with the observation from Gavaris et al. (2000) may explain some of the discrepancies.

On the other hand, taking into account the large number of stocks that were analysed, it may be expected that the specification of the model may not have been optimal in all cases. Thus, is should not be surprising that the final result in some cases are further away from the first estimate than might be anticipated.

We note however that the performance of relative measures of stock size using separable models performs well in our test of uncertainty estimation. If at least one method appears to perform well for some parameters, then this suggests that there exists no widespread failure of conditioning assumptions, but rather that the assessment models may not be parameterised in an appropriate way for the purpose of estimating uncertainty in stock biomass. Indeed, the focus of the development of stock assessment models has usually been to develop very robust short-term forecasting. We speculate that assessment model formulations specifically designed for this latter purpose (e.g. XSA) may in consequence perform poorly at uncertainty estimation.

CONCLUSIONS

1. Calculations of uncertainty in current stock size estimates, short-term stock forecasts and in medium-term stock forecasts appear to be markedly underestimated with methods in common use by ICES. Estimates of uncertainty in changes in stock size appeared to be more consistent in the cases in which separable models had been used. Probability statements of the form

Probability that stock size in two years' time will exceed 800 000t = XX %

are poorly supported by the assessment and forecasting implementations tested here.

2. In the cases where the separable models considered had been used, and only for current estimates and one year-ahead forecasts, probability statements of the form

Probability that stock size next year will exceed stock size five years ago = XX%

appear to perform acceptably.

3. In the cases where the separable model implementation was used, management agencies could be provided with probability statements of acceptable quality about changes in stock size. However, probability statements about absolute stock size – as are presently presented to management agencies by ICES – appear poorly to represent eventual outcomes. Therefore, it is suggested that probability statements about stock development should preferably be formulated in terms of changes in stock size and using the separable model implementation, or else other implementations with better performance should be developed.

4. In cases where probability statements about medium term forecasts are provided to management agencies, it is recommended that these be presented as comparative only, in that they indicate some relative measure of the risks associated with different harvesting strategies. They should not be taken as representing the actual probabilities of eventual outcomes.

5. We have not evaluated the overall performance of the various implementations. The test we have applied is only a test of the consistency with which uncertainty is estimated. Clearly, other issues such as the accuracy with which other parameters such as stock size or catch forecasts are estimated are also important, and would need to be considered in evaluating performance of methods for management purposes.

ACKNOWLEDGEMENTS

REFERENCES

Note: Acronyms in square brackets refer to Table 1.

Anonymous, (1993). Report of the working group on the assessment of demersal stocks in the North Sea and Skagerrak. ICES C.M. 1993/Assess: 6.

Anonymous, (1997a). Report of the working group on the assessment of mackerel, horse mackerel, sardine and anchovy. ICES CM 1997/Assess: 3

Anonymous, (1997b) [SARC]. 24th Northeast Regional Stock Assessment Workshop (24th SAW): Stock Assessment Review Committee SARC) consensus summary of assessments. Northeast Fisheries Science Centre Reference Document 97-12; 291 p.

Anonymous, (1999) [HAWG] Report of the Herring Assessment Working Group for the area South of 62 °N. ICES CM 1999/ACFM: 12

Anonymous, (1999) [AFWG] Report of the Arctic Fisheries Working Group. ICES CM 1999/ ACFM: 3. 312pp.

Anonymous, (1998) [WGNSSK] Report of the Assessment Working Group on the Assessment of Demersal stocks in the North Sea and Skagerrak. ICES CM 1999/ACFM:8. 596pp.

Anonymous, (1999) [WGNPBW] Report of the Northern Pelagic and Blue Whiting Fisheries Working Group. ICES CM 1999/ACFM:18. 238pp.

Anonymous, (1999) [NWWG]Report of the North-Western Working Group. ICES CM 1999/ ACFM: 17. 327pp.

Anonymous, (1999) [WGSSDS]Report of the Working Group on the Assessment of Southern Shelf Demersal Stocks. ICES C.M. 1999/ ACFM: 4, 669pp.

Anonymous, (1999) [WGBFAS] Report of the Baltic Fisheries Assessment Working Group. ICES CM 1999/ACFM:15

Anonymous, (2000). [WGNSDS] Report of the Working Group on the Assessment of Northern Shelf Demersal stocks. ICES CM 2000/ ACFM:1

Darby, C.D. and Flatman, S. (1994) Virtual Population Analysis: version 3.1 (Windows/Dos) user guide. Information Technology Series, MAFF Directorate of Fisheries Research, Lowestoft No. **1**, 85pp.

Deriso, R.B., T.J. Quinn II, and P.R. Neil (1985). Catch-age analysis with auxiliary information. Canadian Journal of Fisheries and Aquatic Science 42: 815-824.

Gavaris, S., K.R. Patterson, C.D. Darby, P. Lewy, B. Mesnil, A.E. Punt, R.M. Cook, L.T. Kell, C.M. O'Brien, V.R. Restrepo, D.W. Skagen and G. Stefánsson (2000). Comparison of uncertainty estimates in the short term using real data. ICES C.M. 2000/V:03

Kennedy, W.J. Jr. and J.E. Gentle (1980). Statistical computing. Marcel Dekker, New York, 591pp.

Kimura, D. (1990). Approaches to age-structures separable sequential population analysis. Canadian Journal of Fisheries and Aquatic Science 47: 2364-2374.

Mayo, R.K. L. O'Brien, and S.E. Wigley (1998). Assessment of the Gulf of Maine Atlantic cod stock for 1998. Northeast Fisheries Science Centre Ref. Doc. 98-13; 88 p.

Patterson, K.R. and G.D. Melvin (1996). Integrated catch at age analysis, version 1.2. Scottish Fisheries Research Report 58: 60pp.

Pope, J.G. (1972) An investigation into the accuracy of virtual population analysis using cohort analysis. ICNAF Research Bulletin 9: 65-74

Patterson, K.R., R.M. Cook, C. Darby, S. Gavaris, L. Kell, A. Punt, V. Restrepo, D.W. Skagen and G. Stefánsson (In prep.). Estimating uncertainty in fisheries. Submitted to "Fish and Fisheries", 22nd August 2000.

Restrepo, V.R., K.R. Patterson, C.D. Darby, S. Gavaris, L.T. Kell, P. Lewy, B. Mesnil, A.E. Punt, R.M. Cook, C.M. O´Brien, D.W. Skagen, and G. Stefánsson (2000) Do different methods provide accurate probability statements in the short term? ICES CM 2000/V:08

Richards, L.J., J.T. Schnute and N. Olsen. (1998). A statistical framework for the analysis of limit reference points. In: Funk, F., Quinn, T.J. II, Heifetz, J., Ianelli, J.N., Powers, J.E., Schweigert, J.F., Sullivan, P.J. and

Chang, C.-I. eds. Fishery Stock Assessment Models, Alaska Sea Grant College Program Report No. AK-SG-98-01, University of Alaska Fairbanks, pp. 185-197.

Shepherd, J. G. (1999) Extended survivors analysis: An improved method for the analysis of catch-at-age data and abundance indices. ICES Journal of Marine Science 56, 584-591.

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

Skagen, D.W. (1999) Programs for stochastic prediction and management simulation (STPR). In: A review of some methods for estimating uncertainty in fisheries (Fisheries Research Services Report No. 7/99) (eds K.R. Patterson, R.M. Cook, C.D. Darby, S. Gavaris, L.T. Kell, P. Lewy, B. Mesnil, A.E. Punt, V.R. Restrepo, D.W. Skagen, and G. Stefánsson. *Available from the author.*

Table 1. List of fish stocks by species and area for which data have been used, and abbreviations used in other tables. The acronym listed under "reference" can be used to identify the report where data were obtained.

Species		Area	Code	Reference
Herring	*Clupea harengus*	North Sea	her_47d3	HAWG
Herring	*Clupea harengus*	West of Scotland	her_vian	HAWG
Cod	*Gadus morhua*	North Sea +	cod_347d	WGNSSK
Haddock	*Melanogrammus aeglefinus*	North Sea +	had_34	WGNSSK
Saithe	*Pollachius virens*	North Sea +	saiiv	WGNSSK
Plaice	*Pleuronected platessa*	Celtic Sea +	plaice	WGNSSK
Whiting	*Merlangius merlangus*	North Sea	whiiv	WGNSSK
Plaice	*Pleuronectes platessa*	English Channel	plai7d	WGNSSK
Sole	*Solea solea*	North Sea	soleiv	WGNSSK
Sole	*Solea solea*	Western English Channel	sole_7d	WGNSSK
Herring	*Clupea harengus*	Off Iceland	her_vasu	WGNPBW
Blue whiting	*Micromesistius poutassou*	North-East Atlantic	whb_comb	WGNPBW
Cod	*Gadus morhua*	Faroes Plateau	cod_farp	NWWG
Haddock	*Melanogrammus aeglefinus*	Faroes Plateau	had_faro	NWWG
Greenland Halibut	*Reinhardtius hippoglossoides*	East Greenland	ghl_gm	NWWG
Cod	*Gadus morhua*	Off Iceland	cod_iceg	NWWG
Saithe	*Pollachius virens*	Faroes Plateau	sai_faro	NWWG
Sole	*Solea solea*	Eastern English Channel	sole_echw	WGSSDS
Plaice	*Pleuronectes platessa*	Eastern English Channel	plai_echw	WGSSDS
Whiting	*Merlangius merlangus*	Western English Channel +	whg_7e_k	WGSSDS
Cod	*Gadus morhua*	Celtic Sea +	cod_7fgh	WGSSDS
Sole	*Solea solea*	Celtic Sea +	sole_celt	WGSSDS
Megrim	*Lepidorhombus whiffiagonis*	Bay of Biscay	mgw_78	WGSSDS
Hake	*Merluccius merluccius*	Northern European	hke_nrtn	WGSSDS
Hake	*Merluccius merluccius*	Southern European	hke_soth	WGSSDS
Haddock	*Melanogrammus aeglefinus*	North-East Arctic	nea_had	AFWG
Greenland Halibut	*Reinhardtius hippoglossoides*	North-East Arctic	ghl_12	AFWG
Saithe	*Pollachius virens*	North-East Arctic	nea_sai	AFWG
Cod	*Gadus morhua*	North-East Arctic	Nea_cod	AFWG
Cod	*Gadus morhua*	Kattegat	cod_kat	WGBFAS
Cod	*Gadus morhua*	Baltic Areas 22-24	cod_2224	WGBFAS
Cod	*Gadus morhua*	Baltic Areas 25-32	cod_2532	WGBFAS
Flounder	*Platichthys flesus*	Baltic Areas 24-25	fle_2425	WGBFAS
Herring	*Clupea harengus*	Baltic Areas 25-29 and 32	her_2532	WGBFAS
Haddock	*Melanogrammus aeglefinus*	West of Scotland	had_via	WGNSDS
Cod	*Gadus morhua*	Georges Bank	cod_5z6	SARC
Haddock	*Melanogrammus aeglefinus*	Georges Bank	had_5z6	SARC
Yellowtail Flounder	*Limanda ferruginea*	Georges Bank	yel_5Z6	SARC
Cod	*Gadus morhua*	Gulf of Maine	cod_5Y	Mayo et al.

+ and adjacent waters

Table 2. "ICA/ICP" Separable model. Summary of assessments and forecasts by stock and by year.

Assessment WG Acronym	Species	Stock	Last Data Year (T)	Last Assessment Year (X)	Forecast Year Y	Stock-Recruit Model	Fractiles of			Fractile of			Reported Catch in Forecast Year Y	
							SSBx/SSBx-5	SSBx+1/SSBx-5	SSBy/SSBy-5	SSBx	SSBx+1	SSBy	Tonnes	Fractile
HAWG	Herring	North Sea, IIIa	1998	1990	1995	BH	0,26	0,28	0,44	0,01	0,04	0,32	640794	0,14
				1989	1994		0,43	0,37	0,35	0,11	0,13	0,23	562619	0,10
				1988	1993		0,75	0,61	0,34	0,69	0,42	0,22	671155	0,12
				1987	1992		0,66	0,60	0,49	0,38	0,34	0,33	716903	0,20
				1986	1991		0,63	0,63	0,65	0,32	0,39	0,51	654147	0,46
HAWG	Herring	W. Scotland	1998	1990	1995	BH	0,04	0,02	0,02	0,00	0,00	0,00	29575	0,00
				1989	1994		0,05	0,04	0,03	0,00	0,00	0,00	24474	0,00
				1988	1993		0,05	0,05	0,06	0,00	0,00	0,01	31778	0,00
				1987	1992		0,16	0,17	0,13	0,00	0,01	0,02	28992	0,00
				1986	1991		0,25	0,19	0,21	0,01	0,01	0,03	28527	0,00
WGNSSK	Cod	North Sea, IIIa	1997	1989	1994	Ricker	0,37	0,10	0,16	0,00	0,00	0,32	110618	0,17
				1988	1993		0,66	0,32	0,38	0,04	0,03	0,26	121749	0,14
				1987	1992		0,56	0,33	0,31	0,00	0,01	0,18	71988	0,90
				1986	1991		0,24	0,29	0,20	0,00	0,13	0,16	102478	0,04
				1985	1990		0,27	0,03	0,35	0,01	0,00	0,31	125314	0,28
WGNSSK	Haddock	North Sea, IIIa	1997	1989	1994	BH	0,91	0,03	0,25	0,90	0,02	0,25	153864	0,27
				1988	1993		0,90	0,03	0,23	0,89	0,02	0,22	174296	0,26
				1987	1992		0,68	0,57	0,08	0,61	0,51	0,07	138001	0,10
				1986	1991		0,86	0,17	0,00	0,80	0,11	0,00	97021	0,01
				1985	1990		0,82	0,38	0,02	0,83	0,37	0,02	92720	0,01
WGNSSK	Saithe	North Sea, IIIa	1997	1989	1994	Ricker	0,50	0,33	0,45	0,01	0,03	0,24	102593	0,06
				1988	1993		0,63	0,27	0,47	0,09	0,01	0,23	104606	0,13
				1987	1992		0,41	0,35	0,37	0,00	0,05	0,16	92499	0,02
				1986	1991		0,32	0,14	0,40	0,00	0,01	0,22	98552	0,11
				1985	1990		0,26	0,09	0,56	0,00	0,00	0,33	88080	0,14
WGNSSK	Plaice	North Sea	1997	1989	1994	Ockham	0,06	0,04	0,19	0,00	0,00	0,11	110392	0,13
				1988	1993		0,08	0,04	0,34	0,00	0,00	0,20	117113	0,00
				1987	1992		0,22	0,20	0,47	0,02	0,04	0,29	125190	0,14
				1986	1991		0,41	0,54	0,66	0,09	0,23	0,42	148004	0,55
WGNSSK	Whiting	North Sea	1997	1989	1994	Ricker	0,04	0,01	0,46	0,01	0,01	0,44	92606	0,07
				1988	1993		0,11	0,01	0,39	0,03	0,01	0,32	116165	0,10
				1987	1992		0,27	0,00	0,76	0,03	0,01	0,65	109704	0,32
				1986	1991		0,21	0,00	0,96	0,01	0,01	0,85	124975	0,45
				1985	1990		0,00	0,00	0,95	0,00	0,00	0,92	153452	0,49
WGNSSK	Plaice	Channel	1997	1989	1994	Shepherd	0,69	0,86	0,54	0,68	0,92	0,43	6121	0,33
				1988	1993		0,74	0,85	0,50	0,78	0,85	0,40	5331	0,24
				1987	1992		0,61	0,60	0,47	0,50	0,51	0,36	6337	0,33
WGNSSK	Sole	North Sea	1997	1989	1994	Ockham	0,55	0,68	0,86	0,41	0,62	0,84	33002	0,91
				1988	1993		0,91	0,81	0,81	0,91	0,79	0,79	31491	0,88
				1987	1992		0,99	0,99	0,93	1,00	0,99	0,93	29341	0,93
				1986	1991		0,82	1,00	0,93	0,87	1,00	0,94	33513	0,96
				1985	1990		0,78	0,93	0,93	0,83	0,93	0,93	94335	0,95
WGNSSK	Sole	VIId	1997	1989	1994	Ockham	0,25	0,19	0,38	0,01	0,03	0,18	4643	0,76
				1988	1993		0,61	0,47	0,60	0,06	0,04	0,33	4511	0,60

Table 2 (Contd.). "ICA/ICP" Separable model. Summary of assessments and forecasts by stock and by year.

Assessment WG Acronym	Species	Stock	Last Data Year (T)	Last Assessment Year (X)	Forecast Year Y	Stock-Recruit Model	Fractiles of			Fractile of			Reported Catch in Forecast Year Y	
							SSBx/SSBx-5	SSBx+1/SSBx-5	SSBy/SSBy-5	SSBx	SSBx+1	SSBy	Tonnes	Fractile
WGNPBW	Herring	Icelandic, SS	1998	1990	1995	BH	0,66	0,34	0,50	0,32	0,07	0,26	125851	0,38
				1989	1994		0,93	0,73	0,85	0,75	0,44	0,69	134003	0,63
				1988	1993		1,00	1,00	1,00	0,96	0,99	1,00	102802	0,99
				1987	1992		0,91	0,99	1,00	0,75	0,95	0,99	106825	0,98
NWWG	Cod	Faroes Plateau	1998	1990	1995	Ricker	0,29	0,05	0,38	0,00	0,00	0,15	22494	0,50
				1989	1994		0,61	0,49	0,50	0,28	0,24	0,29	8818	0,05
NWWG	Haddock	Faroes Plateau	1998	1990	1995	Ockham	0,97	0,76	0,12	0,92	0,67	0,10	4948	0,13
				1989	1994		0,97	0,80	0,04	0,92	0,74	0,04	4252	0,06
NWWG	Greenland Halibut	V and XIV	1998	1990	1995	Ockham	0,54	0,60	0,55	0,24	0,45	0,31	36300	0,30
				1989	1994		0,45	0,38	0,19	0,09	0,10	0,01	36958	0,06
				1988	1993		0,59	0,52	0,32	0,29	0,26	0,06	40817	0,30
				1987	1992		0,75	0,67	0,11	0,81	0,63	0,01	35423	0,03
NWWG	Cod	Icelandic	1998	1990	1995	Ricker	0,56	0,47	0,82	0,50	0,35	0,86	169424	0,30
				1989	1994		0,49	0,68	0,74	0,08	0,50	0,62	178809	0,09
				1988	1993		0,40	0,63	0,40	0,03	0,45	0,20	251979	0,09
				1987	1992		0,73	0,74	0,50	0,89	0,81	0,44	267714	0,41
				1986	1991		0,76	0,94	0,57	0,98	0,99	0,64	308560	0,96
NWWG	Saithe	Faroes Area	1998	1990	1995	Ockham	0,93	0,39	0,42	0,98	0,19	0,28	27222	0,49
				1989	1994		0,91	0,46	0,37	0,98	0,34	0,26	33193	0,68
				1988	1993		0,88	0,58	0,64	0,84	0,42	0,51	33554	0,90
				1987	1992		0,45	0,51	0,47	0,22	0,34	0,30	36487	0,81
				1986	1991		0,02	0,07	0,09	0,00	0,04	0,03	54863	0,82
WGNPBW	Blue Whiting	NE Atlantic	1998	1990	1995	Ockham	0,29	0,35	0,13	0,06	0,15	0,03	578693	0,12
				1989	1994		0,25	0,25	0,32	0,05	0,07	0,21	459414	0,38
WGSSDS	Sole	VIIe	1997	1989	1994	BH	0,88	0,97	0,88	1,00	1,00	0,98	678	0,98
				1988	1993		0,75	0,84	0,56	0,95	0,96	0,56	762	0,67
				1987	1992		0,73	0,88	0,46	1,00	1,00	0,68	769	0,91
				1986	1991		0,65	0,79	0,38	1,00	1,00	0,69	731	0,94
				1985	1990		0,70	0,87	0,44	1,00	1,00	0,85	1082	0,95
WGSSDS	Plaice	VIIe	1997	1989	1994	Ricker	0,90	0,91	0,42	0,94	0,92	0,27	1156	0,36
				1988	1993		0,75	0,50	0,06	0,84	0,43	0,02	1417	0,02
				1987	1992		0,53	0,50	0,12	0,28	0,31	0,04	1624	0,02
				1986	1991		0,76	0,74	0,87	0,63	0,58	0,72	1848	0,68
				1985	1990		0,95	0,98	0,97	0,98	0,99	0,90	2593	0,88
WGSSDS	Whiting	VIIe-k	1997	1989	1994	Ockham	0,99	1,00	0,99	1,00	1,00	0,99	20032	0,96
				1988	1993		0,98	1,00	1,00	1,00	1,00	1,00	16320	1,00
WGSSDS	Cod	VIIfgh	1997	1989	1994	BH	0,64	0,58	0,56	0,39	0,40	0,47	8790	0,37
				1988	1993		0,69	0,64	0,62	0,60	0,56	0,56	8971	0,62
				1987	1992		0,25	0,43	0,65	0,01	0,21	0,58	7605	0,70
				1986	1991		0,20	0,18	0,73	0,00	0,02	0,67	6883	0,50
				1985	1990		0,30	0,40	0,89	0,03	0,19	0,87	10093	0,73
WGSSDS	Sole	VIIfg	1997	1989	1994	Ockham	0,95	0,99	0,93	1,00	1,00	0,94	1009	0,97
				1988	1993		0,85	0,85	0,98	0,89	0,83	0,97	928	0,95
				1987	1992		0,84	0,96	0,99	0,63	0,94	0,98	981	0,98
				1986	1991		0,94	0,93	0,99	0,82	0,84	0,98	1107	1,00
				1985	1990		0,96	0,99	0,99	0,62	0,92	0,98	1189	0,91

Table 2 (contd.). "ICA/ICP" Separable model. Summary of assessments and forecasts by stock and by year.

Assessment WG Acronym	Species	Stock	Last Data Year (T)	Last Assessment Year (X)	Forecast Year Y	Stock-Recruit Model	Fractiles of SSBx/SSBx-5	SSBx+1/SSBx-5	SSBy/SSBy-5	Fractile of SSBx	SSBx+1	SSBy	Reported Catch in Forecast Year Y Tonnes	Fractile
WGSSDS	Megrim (*L. whiffiagonis*)	VII VIII	1997	1989	1994	Ockham	0,29	0,16	0,50	0,01	0,02	0,11	16385	0,81
WGSSDS	Hake	Northern	1997	1989	1994	Ockham	0,33	0,15	0,29	0,02	0,01	0,02	53140	0,05
				1988	1993		0,34	0,17	0,18	0,14	0,05	0,04	52637	0,07
				1987	1992		0,31	0,13	0,05	0,13	0,03	0,00	58290	0,00
WGSSDS	Hake	Southern	1997	1989	1994	Ockham	0,47	0,66	0,76	0,04	0,24	0,42	9542	0,41
				1988	1993		0,56	0,31	0,98	0,08	0,09	0,29	10944	0,20
AFWG	Haddock	North-East Arctic	1998	1990	1995	BH	0,81	0,82	0,93	0,47	0,51	0,86	138423	0,95
			1998	1989	1994	BH	0,68	0,54	0,69	0,21	0,09	0,41	121365	0,81
			1998	1988	1993	BH	0,56	0,27	0,70	0,23	0,04	0,59	77619	0,42
			1998	1987	1992	BH	0,58	0,17	0,43	0,32	0,03	0,21	53886	0,38
			1998	1986	1991	BH	0,61	0,16	0,23	0,30	0,04	0,05	33605	0,11
AFWG	Greenland Halibut	North-East Arctic	1998	1990	1995	BH	0,99	1,00	0,49	1,00	1,00	0,63	11665	0,31
			1998	1989	1994	BH	1,00	1,00	0,75	1,00	1,00	0,75	9318	0,52
			1998	1988	1993	BH	1,00	1,00	0,88	1,00	1,00	0,86	12462	0,93
			1998	1987	1992	BH	1,00	1,00	0,90	1,00	1,00	0,82	8602	0,86
AFWG	Saithe	North-East Arctic	1998	1990	1995	BH	0,67	0,41	0,98	0,01	0,01	0,87	168174	0,95
			1998	1989	1994	BH	0,79	0,41	0,96	0,13	0,02	0,76	146544	0,86
			1998	1988	1993	BH	0,32	0,40	0,79	0,00	0,06	0,50	113627	0,67
			1998	1987	1992	BH	0,11	0,29	0,60	0,00	0,01	0,25	127606	0,53
			1998	1986	1991	BH	0,06	0,02	0,38	0,00	0,00	0,11	107326	0,41
AFWG	Cod	North-East Arctic	1998	1990	1995	BH	0,61	0,74	0,98	0,05	0,44	1,00	739999	0,91
			1998	1989	1994	BH	0,54	0,75	0,81	0,03	0,67	0,83	771086	0,95
			1998	1988	1993	BH	0,60	0,79	0,78	0,26	0,77	0,82	581611	0,86
			1998	1987	1992	BH	0,67	0,92	0,80	0,30	0,96	0,76	513234	0,62
			1998	1986	1991	BH	0,69	0,82	0,58	0,17	0,73	0,20	319158	0,02
WGBFAS	Cod	Kattegat	1998	1990	1995	BH	0,30	0,71	0,94	0,00	0,83	0,93	8165	1,00
			1998	1989	1994	BH	0,48	0,68	0,97	0,13	0,62	0,99	7802	1,00
			1998	1988	1993	BH	0,28	0,66	0,95	0,00	0,58	0,97	7013	0,89
			1998	1987	1992	BH	0,35	0,07	0,66	0,00	0,00	0,59	6271	0,16
			1998	1986	1991	BH	0,54	0,69	0,12	0,41	0,73	0,06	6834	0,04
WGBFAS	Cod	Baltic 22-24	1998	1990	1995	BH	0,75	0,02	0,87	0,74	0,00	0,80	30737	0,83
			1998	1989	1994	BH	0,79	0,86	0,89	0,82	0,88	0,82	29067	0,90
			1998	1988	1993	BH	0,81	0,97	0,84	0,71	0,97	0,71	19953	0,48
			1998	1987	1992	BH	0,53	0,44	0,24	0,24	0,24	0,10	15257	0,03
			1998	1986	1991	BH	0,48	0,41	0,37	0,10	0,13	0,09	15257	0,01
WGBFAS	Cod	Baltic 25-32	1998	1990	1995	BH	0,54	0,81	0,09	0,79	0,97	0,03	107717	0,01
			1998	1989	1994	BH	0,63	0,91	0,02	0,86	0,99	0,00	93354	0,04
			1998	1988	1993	BH	0,78	0,77	0,00	0,98	0,94	0,00	45183	0,01
			1998	1987	1992	BH	0,95	0,99	0,00	1,00	1,00	0,00	54886	0,00
			1998	1986	1991	BH	0,99	1,00	0,01	1,00	1,00	0,01	122893	0,00
WGBFAS	Flounder	Baltic 24-25	1998	1990	1995	BH	0,63	0,76	0,76	0,49	0,66	0,65	12603	0,74
			1998	1989	1994	BH	0,62	0,80	0,72	0,61	0,83	0,69	8493	0,91
			1998	1988	1993	BH	0,60	0,78	0,77	0,44	0,72	0,63	2745	0,11

Table 2 (contd.). "ICA/ICP" Separable model. Summary of assessments and forecasts by stock and by year.

Assessment WG Acronym	Species	Stock	Last Data Year (T)	Last Assessment Year (X)	Forecast Year Y	Stock-Recruit Model	Fractiles of SSBx/SSBx-5	SSBx+1/SSBx-5	SSBy/SSBy-5	Fractile of SSBx	SSBx+1	SSBy	Reported Catch in Forecast Year Y Tonnes	Fractile
WGBFAS	Herring	Baltic 25-29, 32	1998	1990	1995	BH	0,12	0,04	0,00	0,59	0,21	0,01	221000	0,82
			1998	1989	1994	BH	0,62	0,43	0,21	1,00	0,91	0,52	244000	0,99
			1998	1988	1993	BH	0,98	0,83	0,71	1,00	0,98	0,89	231000	1,00
			1998	1987	1992	BH	1,00	1,00	0,52	1,00	1,00	0,78	210000	1,00
			1998	1986	1991	BH	0,85	0,41	0,15	0,94	0,66	0,24	213000	0,89
WGNSDS	Haddock	VIa	1998	1990	1995	Ockham	0,44	0,42	0,55	0,08	0,16	0,45	21167	0,42
			1998	1989	1994	Ockham	0,11	0,05	1,00	0,00	0,01	1,00	25435	1,00
			1998	1988	1993	Ockham	0,05	0,01	0,36	0,00	0,00	0,26	35971	0,34
			1998	1987	1992	Ockham	0,19	0,32	0,17	0,00	0,07	0,09	20752	0,33
			1998	1986	1991	Ockham	0,45	0,30	0,10	0,15	0,07	0,01	19752	0,21
NAFO	Cod	5Y	1998	1990	1995	Ockham	0,93	0,59	0,60	0,79	0,25	0,31	6788	0,23
			1998	1989	1994	Ockham	0,85	0,95	0,24	0,94	0,98	0,16	7887	0,19
			1998	1988	1993	Ockham	0,68	0,96	0,33	0,91	0,99	0,34	8299	0,28
			1998	1987	1992	Ockham	0,83	0,93	0,89	0,90	0,95	0,96	10890	0,95
NAFO	Cod	5Z6	1997	1989	1994	BH	0,83	0,72	0,11	0,73	0,69	0,02	15169	0,04
			1997	1988	1993	BH	0,91	0,78	0,34	0,95	0,77	0,20	23113	0,25
			1997	1987	1992	BH	0,90	0,77	0,43	0,98	0,77	0,35	28567	0,57
			1997	1986	1991	BH	0,78	0,76	0,59	0,41	0,71	0,40	37630	0,43
			1997	1985	1990	BH	0,70	0,61	0,49	0,57	0,57	0,46	42503	0,57
NAFO	Haddock	5Z6	1997	1989	1994	BH	0,59	0,78	0,22	0,46	0,79	0,11	2629	0,16
			1997	1988	1993	BH	0,50	0,63	0,12	0,70	0,77	0,08	4414	0,20
			1997	1987	1992	BH	0,73	0,86	0,25	0,76	0,88	0,14	6063	0,35
			1997	1986	1991	BH	0,81	0,88	0,40	0,57	0,80	0,18	6841	0,24
			1997	1985	1990	BH	0,59	0,72	0,39	0,13	0,48	0,24	5284	0,13
NAFO	Yellowtail Flounder	5z	1996	1988	1993	BH	0,55	0,94	1,00	0,11	0,99	1,00	3980	1,00
			1996	1987	1992	BH	0,48	0,59	1,00	0,08	0,31	0,98	4760	0,88
			1996	1986	1991	BH	0,67	0,62	0,74	0,33	0,30	0,53	2100	0,26
			1996	1985	1990	BH	0,61	0,69	0,65	0,16	0,39	0,40	3560	0,33
			1996	1984	1989	BH	0,33	0,40	0,63	0,01	0,05	0,41	1170	0,14

Table 3. "ICA/ICP" Separable model assessments and forecasts. Stocks by species, area and ICES Working Group ordered by fractile of relative stock size in the last assessment year.

Assesst. WG Acronym	Species	Stock	Stock-Recruit Model	Fractiles of SSBx/SSBx-5
HAWG	Herring	W. Scotland	BH	0,04
WGNSSK	Whiting	North Sea	Ricker	0,04
WGNSSK	Plaice	North Sea	Ockham	0,06
WGBFAS	Herring	Baltic 25-29, 32	BH	0,12
WGNSSK	Sole	VIId	Ockham	0,25
HAWG	Herring	North Sea, IIIa	BH	0,26
NWWG	Cod	Faroes Plateau	Ricker	0,29
WGNPBW	Blue Whiting	NE Atlantic	Ockham	0,29
WGSSDS	Megrim (*L. whiffiagonis*)	VII VIII	Ockham	0,29
WGBFAS	Cod	Kattegat	BH	0,30
WGSSDS	Hake	Northern	Ockham	0,33
WGNSSK	Cod	North Sea, IIIa	Ricker	0,37
WGNSDS	Haddock	VIa	Ockham	0,44
WGSSDS	Hake	Southern	Ockham	0,47
WGNSSK	Saithe	North Sea, IIIa	Ricker	0,50
NWWG	Greenland Halibut	V and XIV	Ockham	0,54
WGBFAS	Cod	Baltic 25-32	BH	0,54
WGNSSK	Sole	North Sea	Ockham	0,55
NAFO	Yellowtail Flounder	5z	BH	0,55
NWWG	Cod	Icelandic	Ricker	0,56
NAFO	Haddock	5Z6	BH	0,59
AFWG	Cod	North-East Arctic	BH	0,61
WGBFAS	Flounder	Baltic 24-25	BH	0,63
WGSSDS	Cod	VIIfgh	BH	0,64
WGNPBW	Herring	Icelandic, SS	BH	0,66
AFWG	Saithe	North-East Arctic	BH	0,67
WGNSSK	Plaice	Channel	Shepherd	0,69
WGBFAS	Cod	Baltic 22-24	BH	0,75
AFWG	Haddock	North-East Arctic	BH	0,81
NAFO	Cod	5Z6	BH	0,83
WGSSDS	Sole	VIIe	BH	0,88
WGSSDS	Plaice	VIIe	Ricker	0,90
WGNSSK	Haddock	North Sea, IIIa	BH	0,91
NAFO	Cod	5Y	Ockham	0,93
NWWG	Saithe	Faroes Area	Ockham	0,93
WGSSDS	Sole	VIIfg	Ockham	0,95
NWWG	Haddock	Faroes Plateau	Ockham	0,97
AFWG	Greenland Halibut	North-East Arctic	BH	0,99
WGSSDS	Whiting	VIIe-k	Ockham	0,99

Table 4. XSA-WGMTerm. Summary of assessments and forecasts by stock and by year.

Assessment WG Acronym	Species	Stock	Last Data Year (T)	Last Assessment Year (X)	Forecast Year Y	Stock-Recruit Model	Fractiles of SSBx/SSBx-5	SSBx+1/SSBx-5	SSBy/SSBx-5	Fractile of SSBx	SSBx+1	SSBy	Reported Catch in Forecast Year Y Tonnes	Fractile
WGNSSK	Cod	North Sea, IIIa	1997	1989	1994	Ricker	0,00	0,00	0,00	0,00	0,00	0,00	110634	0,01
				1988	1993		0,07	0,13	0,32	0,00	0,06	0,30	121749	0,10
				1987	1992		0,00	0,00	0,12	0,00	0,00	0,09	114020	0,01
				1986	1991		0,00	0,00	0,04	0,00	0,00	0,03	102478	0,00
				1985	1990		0,00	0,00	0,00	0,00	0,00	0,00	125314	0,02
WGNSSK	Haddock	North Sea, IIIa	1997	1989	1994	BH	0,956	0,086	0,789	0,955	0,083	0,789	153864	0,73
				1988	1993		1,000	0,988	0,765	1,000	0,986	0,764	174296	0,77
				1987	1992		1,000	0,996	0,415	1,000	0,996	0,414	138001	0,43
				1986	1991		0,997	0,828	0,007	0,997	0,824	0,007	97021	0,05
				1985	1990		1,000	0,583	0,032	0,999	0,569	0,032	92720	0,02
WGNSSK	Saithe	North Sea, IIIa	1997	1989	1994	Ricker	0,01	0,02	0,16	0,01	0,02	0,15	102593	0,33
				1988	1993		0,52	0,26	0,80	0,52	0,26	0,80	104606	0,93
				1987	1992		0,23	0,03	0,00	0,24	0,03	0,00	92499	0,01
				1986	1991		0,00	0,00	0,00	0,00	0,00	0,00	98553	0,03
				1985	1990		0,00	0,00	0,00	0,00	0,00	0,00	88080	0,00
WGNSSK	Plaice	North Sea	1997	1989	1994	Ockham	0,01	0,01	0,00	0,03	0,02	0,01	110392	0,07
				1988	1993		0,60	0,11	0,09	0,65	0,12	0,10	117113	0,15
				1987	1992		0,16	0,13	0,12	0,41	0,32	0,24	125190	0,09
				1986	1991		0,01	0,00	0,02	0,02	0,00	0,02	148004	0,21
WGNSSK	Whiting	North Sea	1997	1989	1994	Ricker	0,03	0,05	0,00	0,03	0,05	0,00	92606	0,00
				1988	1993		0,00	0,00	0,00	0,00	0,00	0,00	116165	0,01
				1987	1992		0,00	0,00	0,00	0,00	0,00	0,00	109704	0,02
				1986	1991		0,02	0,01	0,06	0,02	0,01	0,06	124975	0,04
				1985	1990		0,00	0,00	0,07	0,00	0,00	0,07	153453	0,05
WGNSSK	Plaice	Channel	1997	1989	1994	BH	0,90	0,85	0,02	0,95	0,89	0,02	6121	0,01
				1988	1993		0,93	0,33	0,66	0,95	0,36	0,67	5331	0,49
				1987	1992		0,96	0,50	0,78	0,97	0,52	0,79	6337	0,55
WGNSSK	Sole	North Sea	1997	1989	1994	Ockham	0,97	0,46	0,91	0,99	0,51	0,91	33002	0,76
				1988	1993		1,00	1,00	0,85	1,00	1,00	0,86	31491	0,72
				1987	1992		1,00	1,00	0,99	1,00	1,00	0,99	29341	0,83
				1986	1991		0,99	1,00	0,93	0,94	1,00	0,92	33513	0,82
				1985	1990		0,99	0,95	0,98	1,00	0,97	0,99	35120	0,85
WGNSSK	Sole	VIId	1997	1989	1994	Ockham	0,000	0,015	0,194	0,000	0,012	0,184	4643	0,96
				1988	1993		0,031	0,001	0,717	0,009	0,000	0,612	4511	1,00
NWWG	Cod	Faroes Plateau	1998	1990	1995	Ricker	0,03	0,00	0,26	0,00	0,00	0,12	22494	0,77
				1989	1994		0,00	0,00	0,08	0,07	0,06	0,57	8818	0,04
NWWG	Greenland Halibut	V and XIV	1998	1990	1995	Ockham	0,02	0,38	0,00	0,74	0,99	0,00	36300	0,00
				1989	1994		0,23	0,08	0,02	0,09	0,00	0,00	36958	0,00
				1988	1993		0,68	0,46	0,12	1,00	0,81	0,00	40817	0,00
				1987	1992		0,87	0,94	0,44	0,98	1,00	0,00	35423	0,00
NWWG	Cod	Icelandic	1998	1990	1995	Ricker	0,01	0,00	0,49	0,00	0,00	0,30	169424	0,03
				1989	1994		0,13	0,00	0,18	0,00	0,00	0,02	178809	0,00
				1988	1993		0,43	0,27	0,34	0,00	0,00	0,07	251979	0,00
				1987	1992		0,61	0,52	0,56	0,42	0,03	0,31	267714	0,00

a. Relative SSB in last Assessment Year(SSBx/SSBx-5)

b. SSB in last Assessment Year(SSBx)

c. Relative SSB in first Forecast Year(SSBx+1/SSBx-5)

d. SSB in first Forecast Year(SSBx+1)

e. Relative SSB in 5-year Forecast (SSBy/SSBx-5)

f. SSB in 5-yr Forecast (SSBy)
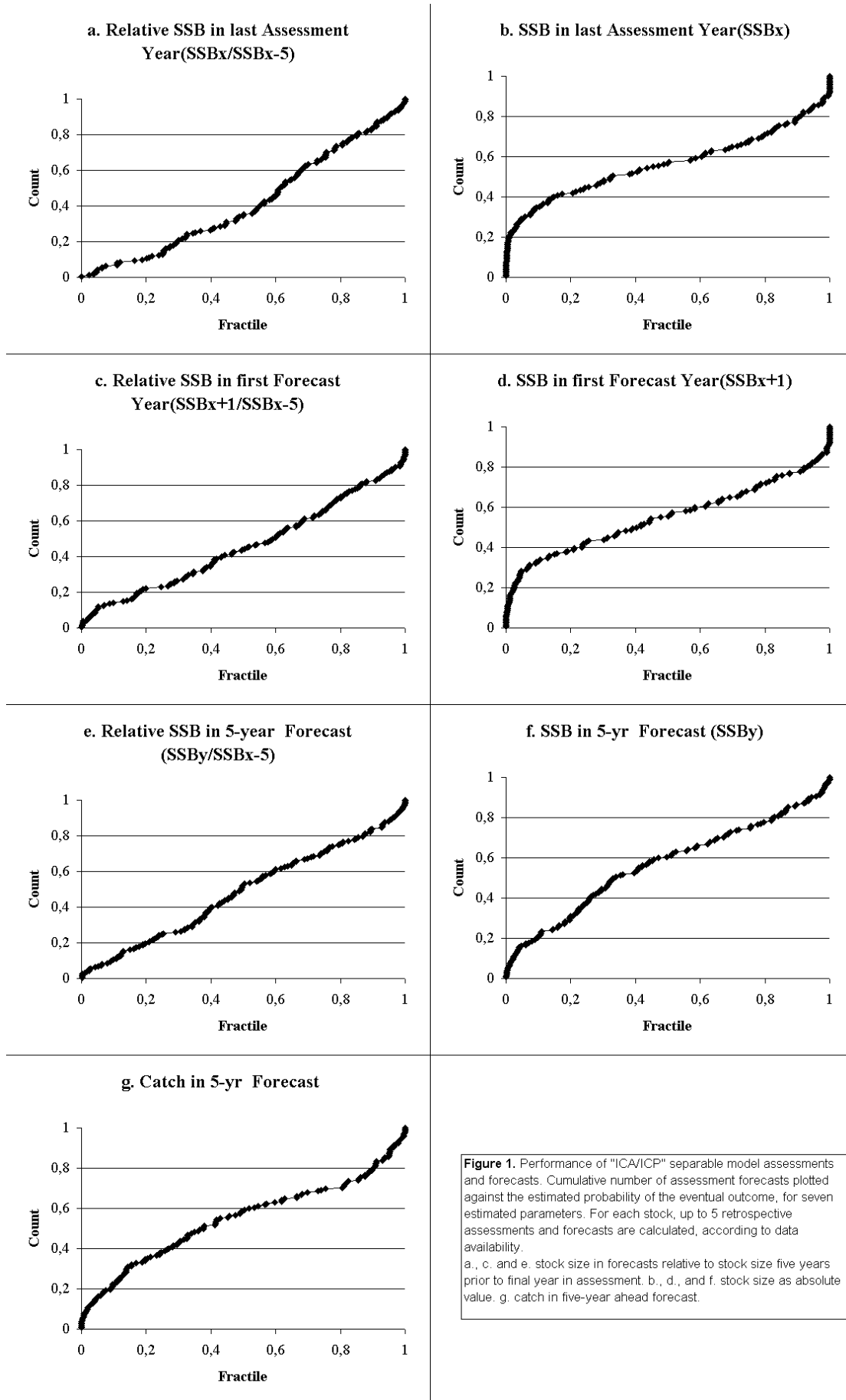
g. Catch in 5-yr Forecast

Figure 1. Performance of "ICA/ICP" separable model assessments and forecasts. Cumulative number of assessment forecasts plotted against the estimated probability of the eventual outcome, for seven estimated parameters. For each stock, up to 5 retrospective assessments and forecasts are calculated, according to data availability.
a., c. and e. stock size in forecasts relative to stock size five years prior to final year in assessment. b., d., and f. stock size as absolute value. g. catch in five-year ahead forecast.
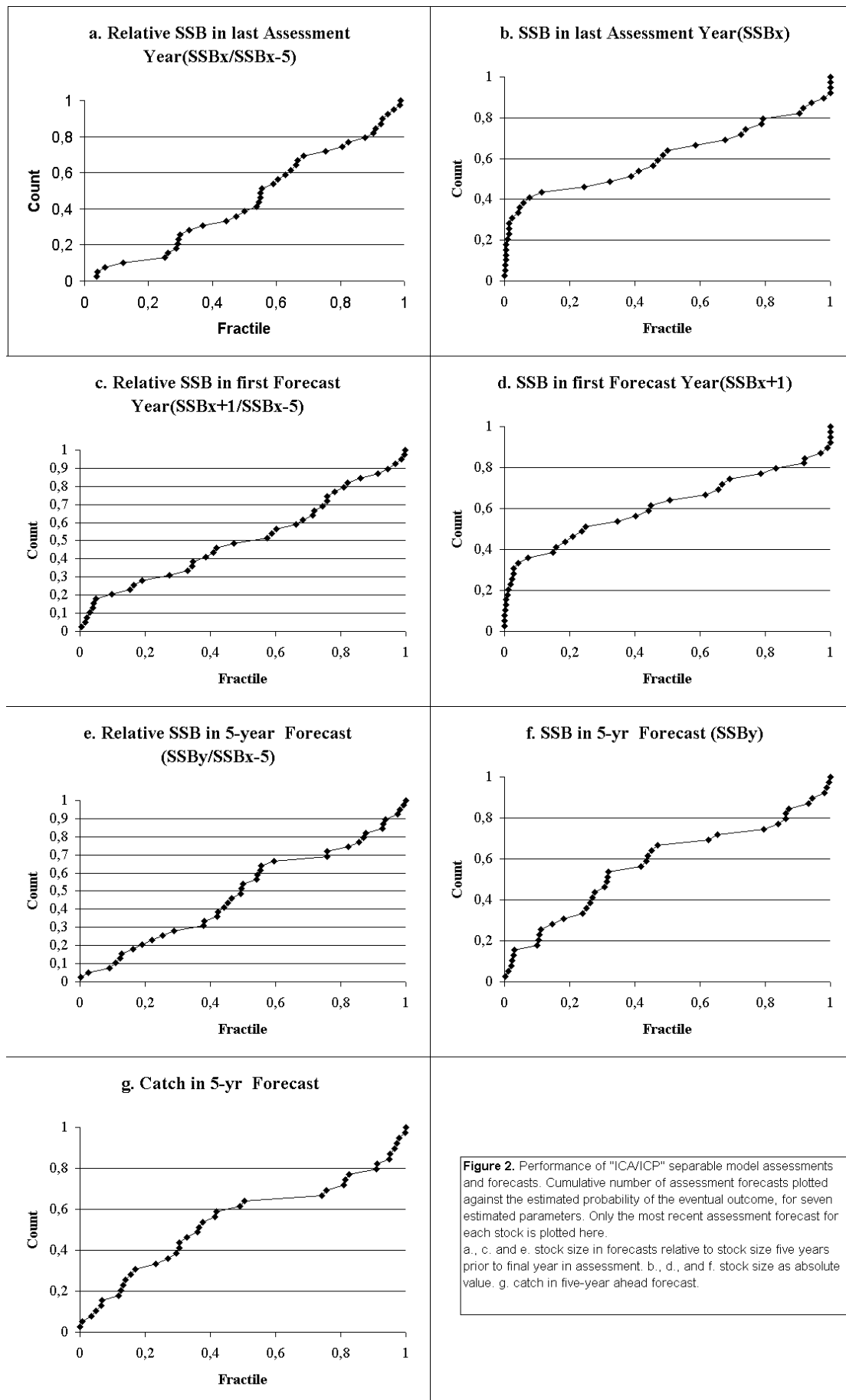
Figure 2. Performance of "ICA/ICP" separable model assessments and forecasts. Cumulative number of assessment forecasts plotted against the estimated probability of the eventual outcome, for seven estimated parameters. Only the most recent assessment forecast for each stock is plotted here.

a., c. and e. stock size in forecasts relative to stock size five years prior to final year in assessment. b., d., and f. stock size as absolute value. g. catch in five-year ahead forecast.

Figure 3. Performance of "XSA/WGM" assessments and forecasts. Cumulative number of assessment forecasts plotted against the estimated probability of the eventual outcome, for seven estimated parameters. For each stock, up to 5 retrospective assessments and forecasts are calculated, according to data availability.
a., c. and e. stock size in forecasts relative to stock size five years prior to final year in assessment. b., d., and f. stock size as absolute value. g. catch in five-year ahead forecast.
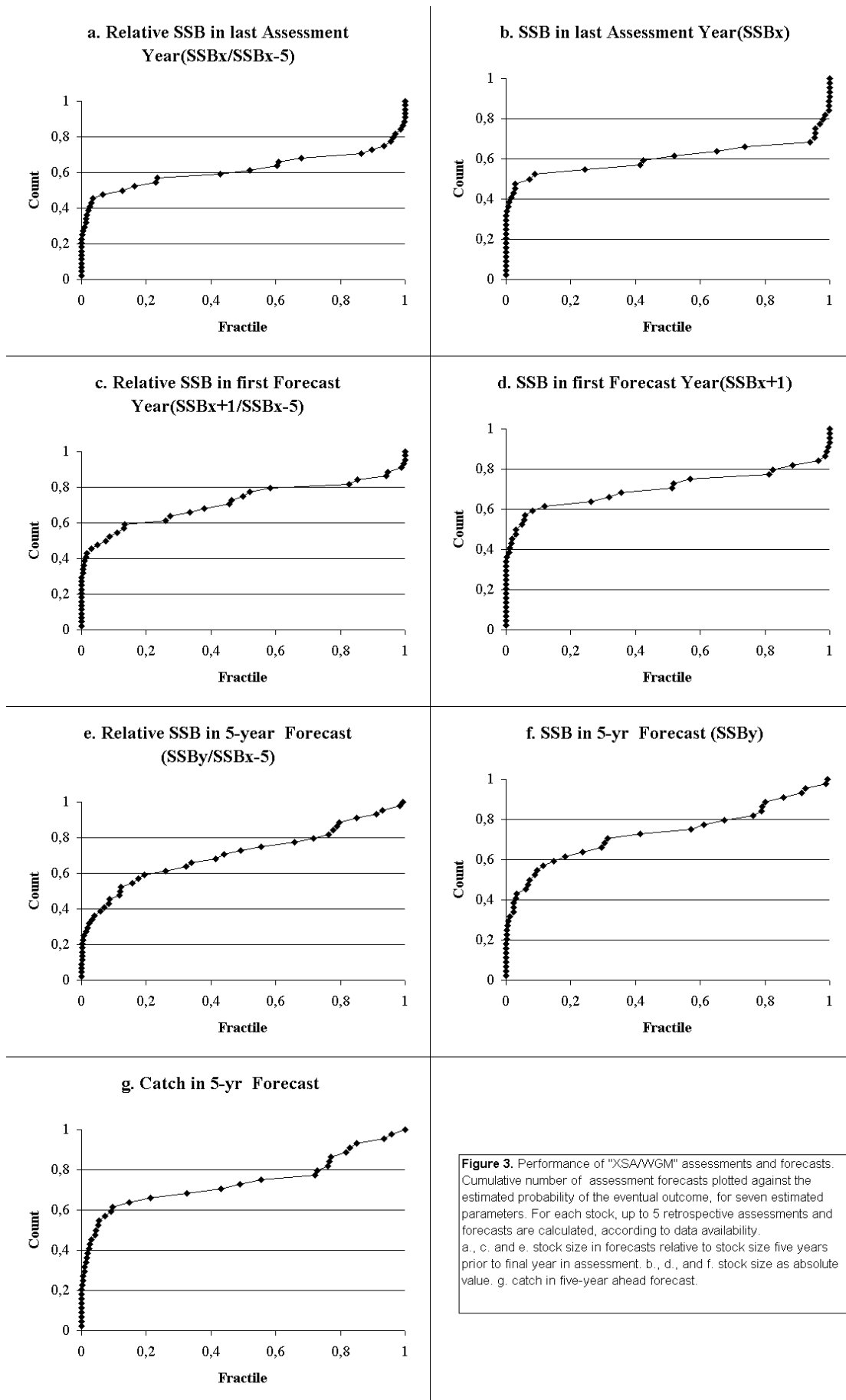
Figure 4. Performance of "STPR" forecast simulations based onseparable model assessments. Cumulative number of forecasts plotted against the estimated probability of the eventual outcome, for seven estimated parameters. For each stock, up to 5 retrospective assessments and forecasts are calculated, according to data availability.

a., c. and e. stock size in forecasts relative to stock size five years prior to final year in assessment. b., d., and f. stock size as absolute value. g. catch in five-year ahead forecast.