RIVO-Netherlands Institute for Fisheries Research

P.O. Box 68
NL 1970 AB Ymuiden
The Netherlands
Phone: +31 255 564646
Fax: +31 255 564644
Internet:postmaster@rivo.wag-ur.nl

P.O. Box 77
NL 4400 AB Yerseke
The Netherlands
Phone: +31 113 672300
Fax: +31 113 573477

RIVO report

Number: C033/03

Analysis of the ICES short-term forecasts of North Sea plaice and sole: dealing with the "current year" assumption.

Sarah B. M. Kraak, Martin Pastoors and Adriaan Rijnsdorp

Commissioned by: Ministerie van Landbouw, Natuurbeheer en Visserij
T.a.v. de directeur Visserij
De heer drs. R.J.T. van Lint
Postbus 20401
2500 EK DEN HAAG

Project number: 324-12470-02

Contract number: 02.036

Approved by: Dr. J.J. de Leeuw
Deputy Head Department Biology and Ecology

Signature: ______________________

Date: 3rd of July 2003

Number of copies: 20
Number of pages: 34
Number of tables: 4
Number of figures: 18

Table of Contents:

Summary

The F1 working package of the F-project is concerned with the improvement of stock assessment of plaice and sole. The full range of problems of uncertainty and bias in the stock assessment will be analysed through a series of small investigations of single problems. The present report deals with the “current year” assumption in the short-term forecast.

Every year ICES Working Groups produce assessments of fish stocks as well as forecasts for the future of these stocks, which serve as a basis for advice. The short-term forecasts consist of a forward projection based on estimates of the numbers-at-age at the beginning of the current year. The weights-at-age and the relative exploitation-at-age are usually assumed to equal the average of the last three years. An assumption has to be made about the catch of the current year. Usually it is assumed that the catch of the current year corresponds to the catch that would be taken under a *status quo* F (that is the F estimate of the previous year). The alternative assumption is that the catch taken would equal the TAC that was set for the current year.

In this study on North Sea plaice and sole, we investigate under which of the two assumptions for current year catch the predictions more closely approximate “reality” as estimated by the most recent assessment, and also which of the two assumptions produces more precautionary predictions for the stocks. We also investigate how much each of the input estimates – stock numbers-at-age, weights-at-age, and relative exploitation-at-age – contributes to the inaccuracy of the forecasts.

The comparison of historical forecasts based on alternative catch assumptions shows that, for both plaice and sole, the *status quo* F assumption leads to less frequent and less severe prediction errors, especially overestimates, of SSB than the TAC assumption. Underestimates as well as overestimates of SSB occurred. In some years the TAC assumption leads to a more accurate forecast, but this is believed to be a spurious result linked to the overestimation of stock size. We can therefore not give any recommendations in what circumstances to use the TAC assumption instead of the usual *status quo* F assumption.

We simulated the consequences of different catch assumptions in three different situations: one where the assessment of the current year is correct, and two where the assessment is biased, i.e. where last year's fishing mortality has been overestimated and where last year's F has been underestimated. Such biases occur frequently, especially underestimates of F (and, consequently, overestimates of stock size). We found that in the case of bias, the highest catch assumption always results in the lowest catch forecast and the highest surviving SSB; in some years this is the *status quo* F assumption but in other years it is the TAC assumption. However, the accuracy of the forecasts depends only marginally on the catch assumption; by contrast, it is strongly affected by the inaccuracy of the assessment. For example, overestimation of the stock size leads to the situation that the F corresponding to the advised TAC exceeds the F that was intended and the surviving SSB is lower than predicted; this effect is marginally stronger with the lower catch assumption.

We found that better estimates of number-at-age would improve the forecasts, but better estimates of recruitment, weights-at-age, and relative exploitation-at-age would not improve the quality of the forecasts. The problem of time trends in weights-at-age is therefore of less concern than the estimation of accurate stock numbers. The analyses imply that the inaccuracy of the short-term forecasts is mainly caused by the large error of the number-at-age estimates given by the VPA each year. This finding focuses our concern again on the quality of the outcome of the VPA.

From this study we conclude that in order to minimise the prediction error in the short-term forecast, given the uncertainty and bias in the VPA, the *status quo* F assumption should be preferred. We also found that the quality of the outcome of the VPA itself plays a large role in the quality of the forecast. Further study of the uncertainty and bias in the stock assessment, as planned in the F-project, is clearly necessary.

Nederlandse samenvatting

In het F1 werkpakket van het F-project houden we ons bezig met verbetering van de toestandsbeoordeling van schol en tong. Het hele scala van problemen betreffende de onzekerheid en bias in de toestandsbeoordeling wordt onderzocht in een serie van kleinere deelstudies die elk een probleem bestuderen. In dit rapport komt de “lopende jaar aanname” in de korte termijn prognose aan de orde.

Elk jaar worden door de werkgroepen van de International Council on the Exploration of the Sea (ICES), bestandsschattingen en vangstprognoses gegevens voor de commercieel geëxploiteerde visbestanden zoals schol en tong. De vangstprognoses voor de korte termijn gaan uit van de meest recente schatting van de aantallen vis per leeftijdsgroep aan het begin van het lopende jaar. Op het moment dat de prognoses worden gedaan, zijn er echter ook nog een aantal andere onbekenden. Bijvoorbeeld: als een vangstprognoses moet worden gegeven voor het jaar 2004, dan zal dat worden gedaan in het jaar 2003. Er zijn dan echter pas vangstgegevens beschikbaar tot en met 2002. Daarom moeten er aannames worden gedaan over de vangsten in 2003 (het zgn. “lopende jaar”) en de visserijsterfte waar die vangsten mee worden genomen. De gangbare methode is om een gemiddelde waarde aan te nemen voor zowel het gewicht als de relatieve visserijdruk per leeftijdsgroep. Voor de vangst zelf kan een keuze worden gemaakt uit twee aannames: (1) de visserijsterfte komt overeen met de visserijsterfte van de vorige jaren (F *status quo* aanname), (2) de vangst is gelijk aan de TAC die voor dat jaar is vastgesteld (TAC aanname)

In deze studie aan Noordzee schol en tong, onderzoeken we welke van deze twee aannames het beste aansluit bij de achteraf geobserveerde ontwikkelingen in de visbestanden. Verder onderzoeken we hoeveel de verschillende invoergegevens voor de korte termijn prognoses (bestandsaantallen per leeftijd, gewichten per leeftijd en relatieve visserijsterfte per leeftijd) bijdragen aan de uiteindelijke onnauwkeurigheden van de prognoses.

De vergelijking van de historische prognoses (d.w.z. de prognoses die in het verleden zijn uitgevoerd) laten zien dat voor zowel schol als tong de F *status quo* aanname beter scoort dan de TAC aanname. Om precies te zijn, de F *status quo* aanname leidt tot minder talrijke en minder grote voorspelfouten, vooral overschattingen, van de paaistand. Zowel onder- als overschattingen van de paaistand kwamen voor. Er zijn echter ook gevallen waar de TAC aanname leidt tot een nauwkeuriger prognose, maar dit is waarschijnlijk een toevallig indirect effect van de bestandsoverschatting. Om die reden kunnen geen sterke aanbevelingen worden gedaan in welke situaties de TAC aanname boven de F *status quo* aanname verkozen zou moeten worden.

Vervolgens simuleerden we de hypothetische gevolgen van de verschillende vangstaannames in drie verschillende situaties: de situatie dat de toestandsbeoordeling zelf geen systematische afwijking had, en twee situaties met systematische afwijking namelijk onderschatting en overschatting van F in het vorige jaar. Deze afwijkingen komen veelvuldig voor, en wel speciaal onderschatting van F (en daarmee verbonden overschatting van de bestandsgrootte). De resultaten laten zien dat, in het geval van een systematische afwijking, de hoogste vangstaanname in het lopende jaar altijd resulteerde in de laagste vangstprognose en de hoogste paaistand na afloop van het prognosejaar. In sommige gevallen was de hoogste vangstaanname de F status quo aanname en in andere gevallen de TAC aanname. Verder toonden we aan dat de nauwkeurigheid van de prognose slechts in geringe mate afhankelijk was van de vangstaanname; de onzekerheid van de toestandsbeoordeling zelf had een veel groter effect. Een overschatting van de bestandsgrootte leidde er bijvoorbeeld toe dat de F die bij de geadviseerde TAC hoorde hoger was dan de beoogde F en dat de overlevende paaibiomassa lager was dan voorspeld; dit effect was enigszins sterker naarmate de gekozen vangstaanname lager was.

De grootste verbetering van de korte termijn prognose zou kunnen worden bereikt door de schatting van de bestandsaantallen per leeftijd te verbeteren. Verbeteringen in de schattingen van gewichten per leeftijd, relatieve visserijsterfte en jonge aanwas hebben minder invloed op de kwaliteit van de prognose. De analyse wijst er op dat de bestandsaantallen per leeftijd de belangrijkste oorzaak zijn voor onnauwkeurige prognoses. Deze uitkomst wijst op de noodzaak om veeleer de toestandsbeoordeling te verbeteren dan zoveel nadruk te leggen op welke vangstaanname het meest nauwkeurig is.

Uit deze studie kan geconcludeerd worden dat de *status quo* F aanname gebruikt moet worden wil men de voorspelfout in de korte termijn prognose zo klein mogelijk houden, gegeven de huidige onzekerheid en bias in de VPA. Bovendien blijkt dat de kwaliteit van de uitkomst van de VPA zelf een belangrijke rol speelt bij de kwaliteit van de prognose. Verder onderzoek aan de onzekerheid en bias in de toestandsbeoordeling, zoals gepland in het F-project, is daarom zeker noodzakelijk.

1. Introduction

The F-project is a 4-year research project with the objective to improve the mutual understanding between fishermen, scientists and fisheries managers, by stimulating communication and collaboration between fishermen and fisheries scientists. One of the three working packages of the F-project is concerned with the improvement of stock assessment of plaice and sole. The results of the annual stock assessments of plaice and sole by ICES have raised serious criticism on the transparency of the methodology, the quality of the input data and the quality of the stock assessment models used. The objectives of the F-project are to prepare for comprehensive fisheries evaluations of North Sea flatfish by analysing and seeking improvements of the following points:

- Representativity of the input data
- Uncertainty and bias in the stock assessment
- Uncertainty and bias in the short-term prognosis
- Biological reference points
- Produce a manual on quality assurance
- Explore alternative methods

These issues will be investigated in several smaller studies of which a total of 13 separate reports and 4 other products will be produced, which, taken together, represent an extensive analysis of the problem. One of the issues that raises concern in the fishing sector is the "current year assumption" in the short-term forecast, which is the subject of investigation in this report.

Total Allowable Catches (TACs) have been applied world wide to regulate fisheries. The establishment of TACs often relies on catch predictions. In the North East Atlantic, the International Council for the Exploration of the Sea (ICES) is responsible for the catch predictions that underpin the decisions on TACs. Consistent with the "Precautionary Approach" (PA), ICES provides yearly advice intended to keep the stock (or bring it back) within safe biological limits by keeping the fishing mortality (F) below F_{pa} and the spawning stock biomass (SSB) above B_{pa}. The advice is presented in terms of recommended fishing mortalities, which can be translated into corresponding landings or catches.

ICES Working Groups are responsible for the historic assessments of the stocks where abundance and exploitation rates are estimated and for the forecasts on the likely future of these stocks. The stock assessments are usually based on a backward Virtual Population Analysis (VPA) (Darby and Flatman 1994; Shepherd 1999) or statistical catch at age models (Patterson 1998; Quinn and Deriso 1999). Each yearly assessment provides updated estimates of abundance and fishing mortality per age group over the whole historic period.

The assessment is typically done using data up to the year prior to that during which the assessment is made and advice for management is required for the year following that during which the assessment is made. For the assessment completed in 2003, commercial catch at age and weights at age data up to the end of 2002 will be used (survey results for 2003 may available and used, however) to formulate management advice for 2004. The short-term forecasts take the results from the historic assessments and project the surviving population forward for two years using different fishing mortality scenarios. The consequences of the fishing mortality scenarios can be evaluated in terms of projected landings (or catches) in the TAC year (2004) and in terms of spawning stock biomass surviving the TAC year (i.e. SSB at the beginning of 2005). However, an assumption has to be made about the catch (landings) in the current year. Usually a choice is made between two options:

- "F-constraint" (or "F *status quo* assumption"): the catch in the current year corresponds to the catch that would be taken under a *status quo* fishing mortality.
- "TAC constraint": the catch taken in the current year equals the TAC that was set for that year.

Moreover, assumptions are made about the average weight per age group and the relative exploitation per age group. These are usually assumed to equal the average of the last three years.

In order to avoid confusion let us at this point clarify the terminology regarding years used in the literature on historical short-term forecasts. The year in which a particular VPA assessment is performed and the advice is formulated is called the “current year”, the “assessment year”, the “mid-year”, the “median year”, or the “intermediate year”. The year before that is called the “last-data-year” (because this is, at the time, the last year about which catch-at-age and weights-at-age data are available), or the “previous year”. The year following the “current year” is called the “next year”, or the “TAC year” (because this is the year for which a TAC advice is being formulated).

Several authors (Van Beek and Pastoors 1999, Brander 1987, Cook et al. 1991, Darby 2002a,b, Sparholt 2002, Sparholt and Bertelsen 2002) have evaluated the quality of short-term forecasts. For the evaluation of the historical short-term forecasts, one can make use of the common assumption that in the present we have better knowledge of the past than at the time when the forecast was produced, because in the meantime the actual catch-at-age and weights-at-age and other data were acquired. Therefore, the latest stock assessments provide our best estimate of reality. Due to the convergence properties of VPA models, the estimates of historical trends in the more distant past are no longer dependent on the calibration process, which is used to estimate the terminal populations. For convenience, some authors refer to the estimates from the most recent VPA assessment as the “true”, “actual”, or “realised” values (Van Beek and Pastoors 1999, Sparholt 2002, Sparholt and Bertelsen 2002), but we have to keep in mind that they are nothing more than the best estimates we have at present.

Sparholt and Bertelsen (2002) present an overview of the quality of 341 short-term forecasts made in the period 1988-1999 for 33 stocks. They compared the SSB predicted to survive the TAC year with the “true” SSB of that year as retrospectively given by the most recent (2001) assessment. They showed that on average stocks are estimated 14% too high, and that in 5% of the cases stocks are estimated more than 3 times higher (up to 8 times higher) than their “true” size. They also showed that forecasts did not improve over the years. Van Beek and Pastoors (1999) examined the relationship between the expected fishing mortality and the “realised” fishing mortality as estimated by the 1998 stock assessment. They compared 5 stocks and observed no positive relationship between expected and realised fishing mortality. Apparently, short-term forecasts produce large errors. It is the aim of this paper to investigate the possible source(s) of these errors.

One problem we address here is the controversy of the catch assumption for the current year. ICES often used the F-constraint option in the forecasts it presents. The choice of the *F status quo* assumption is controversial because in cases where the calculated catch under that assumption is higher than the TAC set for that year, the fishing sector claims that it implies an assumption of underreporting or black landings. From an assessment perspective the F *status quo* assumption is thought to be more robust, because the errors in fishing mortality (F) and stock numbers (N) are linked in the assessment models (Jakobsen and Sparholt 2002). An overestimation in stock numbers is associated with an underestimation of fishing mortality and vice versa. A projection carried out with an F *status quo* assumption will thus balance the relative errors in F and N. If the TAC assumption (or “TAC constraint”) would be used, the errors in F and N would be de-coupled in the forecast, because the fishing mortality will be directly derived from the stock numbers and the TAC instead of from the calibrated assessment. If for example, the stock numbers are overestimated and fishing mortality is underestimated in the assessment, then using a TAC constraint may give estimates of fishing mortality which are even lower than the underestimates that were derived from the historic assessment. Therefore, using the TAC constraint may give an F that is inconsistent with the stock numbers. This could lead to an overly optimistic estimate of stock number at the start of the TAC year (Jakobsen and Sparholt 2002).

The short-term forecast is also based on other estimates and assumptions. Four estimated vectors are used in the forecasts:
(1) stock numbers-at-age at the beginning of the current year,
(2) relative exploitation-at-age (selection pattern) for the current year,
(3) weights-at-age in the stock for the current year, and
(4) weights-at-age in the catch for the current year.
Stock numbers-at-age are survivors after the last year of data, and thus at the beginning of the current year, estimated by the VPA run in the current year. The procedure to estimate the number of recruits may vary from year to year and is explained in the ICES Working Group reports, but it generally involves using information from surveys. The relative exploitation-at-age is usually taken to be the average of the relative exploitation-at-age over the three previous years, as determined by the VPA of the current year. The weight-at-age vectors are usually taken to be the averages of the respective vectors over the three previous years although other assumptions are also frequently made. Therefore, the accuracy of the short-term forecasts depends heavily on the accuracy of the historical assessment and on the validity of using three-year averages for exploitation patterns and weights at age. Recently, Darby (2002a,b) has shown that trends in weights-at-age over time may adversely influence the quality of the forecasts if three-year average weights are used.

The goal of the present study is to gain insight in the causes of the low quality of the forecasts. Special emphasis will be devoted to the influence of the catch assumption for the current year on the quality of the forecasts. The study focuses on North Sea plaice and sole and aims to:

- Compare the consequences of using the TAC constraint versus the *F status quo* constraint for the current year. We want to determine under which of the two assumptions the predictions more closely approximate “reality” as estimated by the most recent assessment and/or are more precautionary for the stocks .
- Investigate how much each of the input parameters contributes to the inaccuracy of the forecasts: stock numbers-at-age (including recruits), weights-at-age in the catch and in the stock, and relative exploitation-at-age. Again, we compare with “reality” as estimated by the most recent assessment.

2. Material and methods

2.1 Comparison between TAC assumption and *F status quo* assumption.

Two studies are performed.
(A) is a comparison between the prediction error of historical forecasts, usually based on the *F status quo* assumption, and the prediction error that would have been made with the TAC assumption.

(B) is a simulation study that investigates the consequences of one or the other assumption if the VPA assessment produced an under- or overestimation of F.

<u>A. Comparison of the prediction errors of historical forecasts based on alternative catch assumptions.</u>

We reviewed the historical short-term forecasts for North Sea plaice and sole made in the assessment years from 1982 to 1999. The forecasts, and the data and procedures used, are published in the annual ICES reports (1983-2000) of the Working Group (WG) that deals with North sea plaice and sole. Before 1982, the assessment methods were considered to be too different for comparison. The WG report for the assessment year 1983 was missing. For plaice, no traditional forecasting had been done in 1985 – 1987, and for sole no forecasting had taken place in 1985 – 1988. These years are therefore left out of the analyses.

We consider the SSB surviving the current year (i.e. SSB of the 1^{st} of January of next year) as predicted by the respective forecast, usually based on the *status quo* F assumption. We then calculate an alternative SSB surviving the current year, through forward projection, using the same data and procedure, but this time with the alternative catch assumption, usually the TAC assumption. For both assumptions we calculate the difference between the predicted SSB surviving the current year and the “actual” SSB for that year, as a percentage of the “actual” SSB. The “actual” SSB for sole is taken from the assessment run in 2001 (ICES 2002). For plaice the “actual” SSB is taken from the assessment run in 2000 (ICES 2001), because in 2001 the procedure to arrive at the international catch-at-age has undergone a major change[1] which could be expected to have an influence on the analysis. The relative differences are measures for the extent of overestimation or underestimation of SSB by the forecasts. They reflect the prediction error under the *status quo* F assumption and the TAC assumption which can therefore be compared.

<u>B. Simulation of the consequences of different catch assumptions.</u>

The issue under study is the question “what happens if the estimates for the stock numbers and fishing mortality from the VPA assessment are biased?”. If the fishing mortalities of the last data year (y) have been underestimated, then the stock on the 1^{st} of January of the current year (y+1) will be overestimated. In that case, a catch in year y+2 that is equal to the set TAC will in fact generate a fishing mortality that is higher than the intended F because it will be taken from a smaller stock. This will result in a lower than predicted SSB surviving the TAC year. The reverse holds true as well. Here we study these effects quantitatively.

---

[1] The raising procedure for North Sea plaice was changed by the 2001 ICES working group. The change was due to the treatment of so-called flag vessels, i.e. vessels that fly the flag of England but that are owned by Dutch companies. In former years these flag vessels were treated as English vessels and the age compositions derived from the English market sampling was applied to these vessels. However, the flag vessels have a very different fishing pattern from the English vessels. Therefore, for the years 1991 and after they were removed from the English raising procedure and instead were raised with an average age-length key for the international fleet.

We investigate two scenarios:
(1) a scenario in which the **true** F of the previous year was in fact 10% higher than that estimated by the VPA, and
(2) the opposite scenario in which the **true** F of the previous year was in fact 10% lower than that estimated by the VPA.
The chosen magnitude of the biases of 10% is arbitrary to show the sensitivity. In both scenarios we assume the same arbitrary value for the realised catch for the current year. We then calculate what the realised F for the TAC year and the realised SSB surviving the TAC year will be if the advised TAC is taken in the TAC year.

North Sea plaice was chosen as the example stock for this simulation study. The year 1999 was chosen as the current year. At that time, the size of the stock of the 1st of January of 1999, and the fishing mortality of 1998 are thought to be known (through estimation by the VPA). We assume that a TAC has been set for 1999. Based on one of the catch assumptions for the current year, the SSB surviving the current year is calculated (assuming that, as usual, the weights-at-age and relative exploitation-at-age estimates are given by the averages over the last three years). Then the TAC advice for the next year (2000) is calculated given the objective of fishing at $F_{pa} = 0.3$. Subsequently, the SSB that would survive this TAC is calculated. For North Sea plaice, mean fishing mortality is averaged over the ages 2 to 10. Each different catch assumption for the current year will lead to a different TAC advice for the next year and a different SSB predicted to survive the TAC year. We investigate three catch assumptions:
(1) the assumption that the TAC will be taken (the TAC assumption),
(2) the assumption that the catch taken will correspond to the *status quo* F (the *status quo* F assumption), and
(3) the assumption of an arbitrary catch. The third assumption is investigated to get a short series of decreasing catch assumptions.

The simulated input data for the short-term forecast are generated as follows. For each of the ages 1 to 15+ the F-at-age for the previous year was increased or decreased by 10%. Given the original catch-at-age data of the previous year, new stock-at-age data for the 1st of January of the previous year were calculated from the new F-at-age, using equation 1. The catch-at-age data of the previous year were assumed to be correct.

$$N_{a,y} = C_{a,y} * (F_{a,y} + M) / F_{a,y} (1 - e^{(-M - Fa,y)}) \qquad (1)$$

Subsequently, from the new F-at-age and N-at-age of the previous year, the new stock-at-age of the 1st of January of the current year was calculated, using equation 2.

$$N_{a+1,y+1} = N_{a,y} * e^{(-M - Fa,y)} \qquad (2)$$

N represents stock in numbers. C represents catch in numbers. M represents natural mortality (which is set to 0.1 for all years and all ages). The indices a and y are for age and year respectively.

The mean exploitation pattern at age was re-scaled to the mean fishing mortality in 1998 (contrary to the working group, which did not re-scale the exploitation pattern).

2.2 Do better estimates of stock-at-age numbers (including recruits), weights-at-age in the catch and in the stock, and relative exploitation-at-age improve the quality of the forecasts?

For these analyses the same historical forecasts of North Sea plaice and sole as described above in part 2.1A (ICES 1983-2000) are used. In order to investigate the question above, we repeat the calculations of the relative differences between predicted and "actual" SSB (as described above in 2.1A). However, this time we substitute (one by one):
- "actual" stock-at-age for the respective current year,
- "actual" relative exploitation-at-age for the respective current year,
- "actual" weight at age for the respective current year.

The "actual" data were taken from the 2000 assessment for plaice and 2001 assessment for sole (ICES 2001, 2002). The "actual" values are the best estimates we have (given the assessment procedures) and, therefore, represent the maximum improvement that could be attained. We evaluate the change in prediction error when these substitutions are made. Here we do not consider the difference between the two catch assumptions, but only the *status quo* F assumption.

For North Sea plaice, we also analysed the influence of a better estimate of 1-year-old fish (recruits) versus a better estimate of stock-at-age for all ages. We calculated, through forward projection, the SSB surviving the current year as well as the SSB surviving the TAC year, assuming a *status quo* F for the current year and the TAC year. We calculated the SSBs predicted if the "actual" recruitment data are substituted for the current and next year, and the SSBs predicted if the "actual" stock-at-age data of the current year as well as "actual" recruitment of the next year are substituted. As before, the relative differences of the predicted SSBs with the respective "actual" SSBs (i.e. the prediction errors) are judged and compared with those derived from the original forecasts. "Actual" values are taken from the assessment run in 2000 (ICES 2001).

Since concern exists in particular about possible trends in weights-at-age over time (Darby 2002a,b) we evaluated these trends for North Sea plaice, looking at the weights-at age data reported in the most recent assessment (ICES 2002). Apart from the question whether time trends in weights-at-age adversely influence the quality of the short-term forecasts, the question arises whether the weight estimates should rather be smoothed using regressions of weight against age. We carried out regressions for each year class for which 8 or more data points were available, and found that linear regressions have higher R-square values than loglinear regressions in 10 out of 13 cases. We decided to investigate whether the quality of the forecasts would improve when smoothed values were used for weight at age, using the predicted values from linear regressions. We substituted the averages over the last three years of these "smoothed" values in the calculations of the forecasts. In addition we substitute the "smoothed" actual weight of the current year (actual data taken from ICES 2001). In both cases we evaluated the change in prediction error of the SSB surviving the current year.

3. Results

3.1 Comparison between TAC assumption and *status quo* F assumption.

A. Comparison of the prediction errors of historical forecasts based on alternative catch assumptions.

Figures 1 and 2 present the relative differences between the plaice and sole SSBs predicted to survive a particular “current year” and the “actual” SSB of that year. The relative difference is represented as a percentage of the “actual” SSB. These relative differences are the prediction errors; a positive value designates an overestimation, whereas a negative value stands for an underestimation. The year on the x-axis represents the year for which the SSB is predicted on the 1st of January, i.e. the next year. As an example: the SSB in 1996 is predicted from data up to and including 1994 and then surviving the “current year” 1995 to the 1st of January 1996. A comparison can be made between the prediction errors under the *status quo* F assumption and under the alternative assumption of the TAC constraint.

The plaice SSBs appear to have been overestimated in most years, up to almost 50% (figure 1). The “actual” SSB estimated from the 2000 assessment is always (much) smaller than we thought at the time. In recent years the TAC assumption resulted in stronger overestimation, whereas in the more distant past the *status quo* F assumption gave the higher overestimates. Regarding sole (Figure 2), the SSB in recent years has been overestimated, while in the more distant past the SSB has been underestimated. The TAC assumption always gave the higher SSB estimates, which means that in the case of overestimation the TAC assumption results in bigger prediction error. Especially for sole it appears to be more cautious to use the *status quo* F assumption than the assumption of a TAC constraint because the first leads to less severe overestimates of SSB than the last.

The same results can be presented in a slightly different way. Figures 3 and 4 present the gain in accuracy when using the TAC assumption versus the *status quo* F assumption. The inaccuracy is here taken to be the absolute value of the prediction error. The gain in accuracy under the TAC assumption is the difference between the inaccuracy under the *status quo* F assumption and the inaccuracy under the TAC assumption. If the inaccuracy is higher than zero than the TAC assumption is an improvement to the F *status quo* assumption; if it is below zero, then the TAC assumption makes things worse.

In the case of plaice (Figure 3), the differences in accuracy between the two assumptions are often small. In the assessment years 1993 – 1995 the TAC assumption would have improved the quality of the forecast, but in the assessment years 1984, 1996, and 1998 the TAC assumption would have led to a lower quality of the forecast. In the case of sole (Figure 4), the TAC assumption would have lowered the accuracy of the forecast in most years. The TAC assumption would have been an improvement only in the assessment years 1989 – 1991 when the forecasts underestimated the SSB.

Trends in “actual” SSB, actual catch, the catch under the *status quo* F assumption, and the agreed TAC are shown in figures 5 (plaice) and 7 (sole). The time trends in “realised” or “actual” F, *status quo* F, and the implied F under the TAC assumption are shown in figures 6 (plaice) and 8 (sole). If we compare the trends in Figure 1 and Figure 5 for plaice, it can be seen that the prediction errors are particularly large when SSB is particularly low (below Bpa = 300,000). However, these are precisely the years in which a reliable forecast is of the utmost importance. The large improvements that would have been made by use of the TAC assumption in the assessment years 1993 – 1995 (Figure 3) coincide with the period when the “actual” SSB was decreasing to values below the SSB_{lim} (= 210,000) (Figure 5) and that the TAC was well above the actual catch.

At the end of that period the TAC was brought down considerably, as a corrective measure, from over 150,000 tonnes to below 100,000 tonnes (the actual catch had already gone down gradually in the period before). There are two possible explanations: (1) When the stock is dangerously low, the TAC assumption tends to lead to more accurate predictions. (2) When the TAC is well above the actual catch and/or the F *status quo* catch, the TAC assumption tends to lead to more accurate predictions and can compensate the overestimation of stock size. In the assessment year 1994, when the prediction under the TAC assumption was very accurate (1995, Figure 1), the difference between the actual catch and the TAC year is large, whereas the catch under the *status quo* F assumption resembles the actual catch very closely (Figure 5). Apparently the assumption that renders the forecast most accurate is the one that overestimates the realised catch. This finding may be fortuitous. The converse is of course also true. Comparing Figure 1 and Figure 6, it can be seen that prediction errors for plaice are particularly large when the *status quo* F overestimates the "realised" F. In the assessment years in which the TAC assumption would have been a large improvement (1993 – 1995), the F implied by the TAC assumption overestimates the "realised" F to a large extent.

For sole, Figures 2 and 7 show that those years for which SSB was underestimated are those years in which SSB had risen to particularly high levels due to the strong 1987 year class. Apparently the forecasts are not very well able to follow such developments. In the same years the TAC assumption would have improved the quality of the forecasts (Figure 4), but the TAC was much lower than the actual catch (Figure 7). Here the TAC assumption makes the forecast more accurate when the TAC is an underestimation of the realised catch. This is in opposition to what is found for plaice, strengthening our suspicion that these associations are spurious (and that interactions with other factors must be present).

Looking at Figures 2 and 8 for sole, it is clear that the severest cases of overestimation of SSB coincide with underestimation of F under both assumptions. The assessments had not been able to track the steep decline of the stock from 1995 onwards (Figure 7) and the rise in F (Figure 8). The only years in which the TAC assumption would have been an improvement, are years in which the F implied by the TAC assumption underestimates the "realised" F (Figure 8). This observation implies that when the assumption for the current year underestimates F, a better forecast is reached. This is directly opposite to what was found for plaice, which makes it more likely that these findings are spurious.

B. Simulation of the consequences of different catch assumptions.

In Table 1, three different forecasts are presented, based on three catch assumptions for North Sea plaice in 1999 ("the current year"): the TAC assumption, the *status quo* F assumption, and an arbitrary assumption of a catch of 90,000 tonnes. The three TACs advised for the TAC year (2000) are based on the assumption that F2000= F_{pa}= 0.3. The SSBs surviving if the respective TACs are taken are also presented in the table.

Table 1. North Sea plaice. Implied TAC advice for 2000 given different catch assumptions in 1999 ("the current year"). Unit: tonnes.

Catch assumption (current year)	TAC advice	SSB surviving TAC year
102,000 (TAC)	53,500	202,800
90,000	57,500	211,800
77,900 (*status quo* F)[1]	61,500	220,900

[1] In this paper we use the numbers of 2- and 3-year olds taken from the VPA run in 1999, whereas the WG in 1999 used different numbers in the forecast. Moreover, in contrast to the forecast of the WG in 1999, we use the mean values of F-at-age of the last three years scaled to the average F for 1998. Therefore, the TAC advice presented here differs from the one given by the WG in 1999, and the catch under the *status quo* F assumption reported here differs from the one implied in the forecast of the WG in 1999.

Table 2 presents the results of the simulation of a set of fictive worlds where the **real** F-at-age in the previous year was 10% higher than the values perceived by the assessment of 1999 (ICES 2000) and where the stock differed accordingly (see methods section). We compare three fictive worlds, which differ among themselves only in the choice of the catch assumption that was chosen in 1999 and the TAC advice for 2000 following from that. In each case the same arbitrary catch of 90,000 tonnes is taken in the current year. We calculated the fishing mortalities and the SSBs surviving each of the TACs given the **real** stock numbers, and present these in the Table. Table 3 presents the results of a similar simulation where the **real** F-at-age in the previous year was 10% lower than that given by the assessment of 1999 (ICES 2000).

Table 2. North Sea plaice. Fs implied by the three TAC advices and SSBs surviving those TACs given the **real** stock numbers. In this scenario **real** F-at-age is 10% higher than assessed. Unit: tonnes.

Catch assumption	TAC advice	F in TAC year	SSB surviving TAC year
102,000 (TAC)	53,500	0.33	187,750
90,000	57,500	0.36	184,000
77,900 (*status quo* F)	61,500	0.39	180,300

Table 3. North Sea plaice. Fs implied by the three TAC advices and SSBs surviving those TACs given the **real** stock numbers. In this scenario **real** F-at-age is 10% lower than assessed. Unit: tonnes.

Catch assumption	TAC advice	F in TAC year	SSB surviving TAC year
102,000 (TAC)	53,531	0.23	248,350
90,000	57,494	0.25	244,500
77,894 (*status quo* F)	61,510	0.27	240,600

In case of a 10% higher F, the resulting mean F in the TAC year is higher than the intended F_{pa} of 0.3 and the SSB surviving the TAC year is lower than predicted. In case of a 10% lower F, the reverse is true: the resulting mean F in the TAC year is lower than the intended F_{pa} and the SSB surviving the TAC year is higher than predicted. In the first case the precautionary approach would have failed and in the latter case the approach would have been too cautious and the TAC advice unnecessarily stringent.

Moreover, in the case of an underestimated F (Table 2), a lower catch assumption resulted in a larger prediction error of mean F and a lower SSB surviving the TAC year. It would, therefore, be more precautionary for the stock to choose the higher catch assumption for the current year, resulting in a lower advised TAC whereby mean F will exceed F_{pa} to a lesser extent. On the other hand, in the case of an overestimated F (Table 3), the lowest catch assumption is associated with the smallest prediction error. Here the advice would be too cautious, and the more so with a higher catch assumption for the current year.

It should also be noted that the prediction error of the SSB surviving the TAC year differs only marginally (about 4% or less) between the three catch assumptions – even though the three catch assumptions themselves differ to a rather large extent (about 15%). The prediction error depends much more strongly on the inaccuracy of the assessment. The impact of the inaccuracy of the assessment (of only 10%) is much greater than the impact of the large difference in catch assumption.

3.2 Do better estimates of stock-at-age numbers (including recruits), weights-at-age, and relative exploitation-at-age improve the quality of the forecasts?

A. Stock-at-age and recruitment

Figures 9 and 10 present the prediction errors of the original plaice forecasts, together with those for which the “actual” recruitment or the “actual” stock numbers-at-age were substituted[1]. Figure 9 shows the SSB surviving the current year and Figure 10 shows the SSB surviving the TAC year. *Status quo* F is assumed for both the “current year” and the “TAC year”[2]. For sole, we only analysed the prediction errors in SSB surviving the current year (figure 11), and only when the complete “actual” stock-at-age vector is substituted[3].

For plaice, the accuracy of the short-term forecasts would have been greatly improved if the “actual” stock-at-age had been known (Figures 9 and 10). By contrast, if only the “actual” recruitment had been known, improvements are only small. The improvement with substitution of the “actual” recruitment is smaller for the SSB surviving the current year (Figure 9) than for the SSB surviving the TAC year (Figure 10); this reflects that in the former forecast only one “actual” recruitment number is used versus two in the latter forecast. Figure 11 shows that also for sole the predictions become more accurate if the “actual” stock-at-age data were known. However, the effect is not as strong as in plaice.

B. Weights-at-age in the catch and in the stock.

Figures 12 and 13 present for plaice and sole respectively, the error in the prediction of the SSB surviving the current year, in the original forecast with *status quo* F assumption, and a similar forecast in which the actual weights-at-age was used. In the case of plaice, there is almost no difference in accuracy of the forecast, whether the actual weights-at-age data are substituted or not (Figure 12). The same holds true for sole (Figure 13), except for the SSB of 1995, in which case substitution of actual weight data yields a very accurate prediction. Apparently, the forecasts are not sensitive to the weight data.

In the literature, concern has been raised about possible time trends in weights-at-age (Darby 2002a,b). Figure 14 shows the plaice weights-at-age in the stock over time. Especially the weight of older fish fluctuates strongly from year to year. Table 4 shows that for seven age groups (1-, 8-, 9-, 10-, 11-, 12-, and 13-year old fish) the trend is significant: fish of those ages have become lighter in recent years.

Table 4. Correlations of plaice weights-at-age with year.

Age	Pearson *r*	significance
1	-0.89	<0.001
2	-0.04	>0.1
3	0.30	>0.1
4	0.03	>0.1
5	-0.14	>0.1
6	-0.23	>0.1
7	-0.48	<0.1
8	-0.58	<0.05

[1] Note that the year on the x-axis is the “current year” and not the year whose SSB was referred to as in the previous figures.

[2] Note that for the forecasts in Figure 10, “actual” recruitment is substituted in both years, the current year and TAC year.

[3] Note that in Figure 11 the year on the x-axis again refers to the year whose SSB is referred to.

9	-0.77	<0.01
10	-0.79	<0.001
11	-0.57	<0.05
12	-0.69	<0.01
13	-0.80	<0.001
14	-0.19	>0.1
15	-0.53	<0.1

The fluctuations in weights-at-age over time might be ground for concern: one may ask whether the estimates are accurate enough for their use as a basis in the forecasts. The average over three years is taken in order to average out random fluctuations (but this would not be appropriate if a clear trend exists). Figure 15 shows the change in weight of fish of the different year classes over time, i.e. of fish growing older (Figure 15a displays all ages, and Figure 15b displays only the ages 10 – 15+). For all year classes it is the case that sometimes the average weight of the fish in the stock decreases when the fish grow one year older! This is clearly an artefact due to sampling, implying that the weight in the stock as estimated from the fish samples is not very accurate. We investigated whether the quality of the forecasts improves when “smoothed” values (see MATERIAL & METHODS) are used, using the predicted values from linear regressions of weight on age. We substituted the averages over the last three years of these “smoothed” values in the calculations of the forecasts. In addition we substituted the “smoothed” actual weight of the current year.

Figure 16 compares the prediction errors of the conventional forecasts with the prediction errors when the 3-year-averages of the “smoothed” values are substituted, and when the “smoothed” values of actual weight are substituted. There is a small but consistent improvement for plaice. For sole the substitutions do not always lead to an improvement. Apparently the forecasts are not very sensitive to the weight data.

C. Exploitation-at-age.

Finally we analysed whether the forecasts improve when the “actual” relative exploitation-at-age is substituted. Figures 17 and 18 show, for plaice and sole respectively, the prediction errors in SSB surviving the current year in the original forecasts with *status quo* F assumption, and when the “actual” relative exploitation-at-age was substituted. The substitutions do not lead to any improvement, implying that the forecast of SSB is not sensitive to the relative exploitation-at-age data.

4. Discussion

This study looks at the possible sources of the inaccuracy of the short-term forecast. It addresses the question of which of the two “current year” assumptions leads to a more accurate forecast (and should therefore be preferred on scientific grounds) and also what are the relative contributions of the various sources to the overall inaccuracy of the forecast. These issues are part of the complex problem of the uncertainty and bias in the stock assessments, which is under investigation in the F-project.

It is understood by the fishing sector that the VPA models used for the stock assessment yield only estimates of stock size and fishing mortality, and that estimates are by definition uncertain, and possibly biased, reflection of the true state of the stock, which cannot be known. It is the non-transparency of the methodology that raises concern in the fishermen. In the case of the current year assumption, they believe that by using a catch assumption higher than the TAC set for the current year, the scientists imply underreporting and/or black landings. In this light it is difficult for the fisheries managers to justify the scientific advice towards the sector. It would be ideal, and much more transparent, if the VPA would assess the stock size – with some uncertainty – and subsequently the TAC assumption would be used in the forecast, from which the catch option table would be derived. The uncertainty of the stock assessment could then be used when choosing from the option table and/or be incorporated in the PA reference points. However, the situation is far from ideal. Stock assessment is not only subject to uncertainty but also to systematic bias. Often a retrospective pattern occurs, meaning that with each addition of a year’s data the perception of the stock is adjusted in the same direction (i.e. the stock size appears to have been consistently over- or underestimated). The F-project aims at improving the stock assessment by addressing possible problems with the input data, such as the age composition of the landings (product A1), the CPUE series (product A4), and the survey indices (product A5), as well as the problem of discards (product A3). Moreover, it is intended to assess whether the use of more realistic biological data will improve the assessment (product A6). The uncertainty and bias caused by the VPA model itself will be investigated (product A10). An integral analysis will calculate the relative contributions to the overall uncertainty and bias by the input data and the model assumptions (product A11), and how the uncertainty and bias are translated into the forecast (products A14 and A15). The biological reference points and possible alternative reference points are under study as well (products A16 and A17 respectively). For the time being, this report shows from a scientific point of view which current year assumption yields the more accurate prediction given all the imperfections with the assessment model that we are presently confronted with.

The comparison of historical forecasts based on the two alternative catch assumptions shows that, for sole, it is clearly better to use the *status quo* F assumption than the TAC assumption because the former leads to less frequent and less severe overestimates of SSB. For plaice the results are less clear but here too the *status quo* F assumption tends to yield more accurate predictions. In some individual years the TAC assumption leads to a more accurate forecast, but this is believed to be a spurious result linked to the overestimation of stock size. It is recommended to use a generic simulation framework in order to investigate the details of how these phenomena occur and how they could be remedied. Therefore, we cannot at present recommend under what circumstances the TAC assumption should be preferred over the usual *status quo* F assumption.

The results of the simulations of the consequences of different catch assumptions and the bias in the assessment have shown that when the VPA is inaccurate, a higher catch assumption for the current year results in a higher SSB surviving the advised TAC. In the case that the assessment has overestimated the stock size, the fishing mortality in the TAC year will exceed F_{pa} – unknown – to a lesser extent when the higher catch assumption is used. It is therefore more cautious to choose the higher catch assumption, resulting in a lower TAC advice and a higher surviving SSB.

However, since we do not know beforehand whether the assessment under- or overestimates F, such a policy could as well lead to unnecessarily restrictive TACs. It should be noted that the simulations show that the higher catch assumption is more cautious, but that in some years the *status quo* F assumption assumes the highest catch and in other years the TAC assumption.

The more interesting result of the simulation study is that the quality of the forecasts (in terms of prediction error of SSB surviving the TAC year) depends only marginally on the catch assumption, but is strongly affected by the inaccuracy of the assessment. This finding should shift our focus away from the controversy of the catch assumption towards improvement of the assessment.

Our historic investigation shows that better estimates of stock-at-age would improve the forecasts, but better estimates of recruitment, weights-at-age, and relative exploitation-at-age would improve only marginally the quality of the forecasts of SSB. The problem of time trends in weights-at-age is therefore of less concern than the estimation of accurate stock numbers. The analyses imply that the inaccuracy of the short-term forecasts is mainly caused by the large error of the stock-at-age estimates given by the VPA each year. This finding focuses our concern again on the quality of the outcome of the VPA

Similar results have been reached through other studies. Jakobsen and Sparholt (2002) showed that the prediction error of the short-term forecast, as a function of error in the assessment, is usually larger under the TAC assumption than under the *status quo* F assumption. They recommended that "there needs to be strong evidence that the F in the assessment is not underestimated before a TAC constraint should be preferred". Van Beek and Pastoors (1999) also found that the input stock-at-age was one of the most crucial factors responsible for the differences between predicted and "realised" fishing mortality. The present study corroborates that finding.

Five important conclusions come forward from this study.
(1) The *status quo* F assumption yields a more accurate forecast more often than the TAC assumption. Prediction errors can be overestimates as well as underestimates; the *status quo* F yields fewer and less severe overestimates than the TAC assumption.
(2) The highest catch assumption for the current year, which in some years may be the *status quo* F assumption and in other years the TAC assumption, yields the most cautious TAC advice.
(3) Inaccuracy of the F and N estimation by the VPA has a bigger impact on the development of the stock than the choice of the catch assumption.
(4) The quality of the forecast is much more strongly affected by the accuracy of the VPA than by the catch assumption used (this corresponds with conclusion 3).
(5) The quality of the forecast is more strongly affected by the accuracy of the stock-at-age data than by the accuracy of recruitment, weight, and relative exploitation estimates.

Conclusions 3, 4, and 5 compel us to shift our attention to the quality of the stock assessment.

All of the results presented in this report are conditional on a number of strict assumptions. It is likely that the most stringent assumption is that the most recent stock assessment for the stocks considered reflects the true development of the stock. Although the most recent years were excluded from the analysis – circumventing problems of reliance on calibration data from commercial CPUE series and survey data – the converged part of the VPA is still uncertain because it relies solely on catch at age data and initial population estimates derived from calibration. This can only be remedied by application of a simulation approach in which both the underlying population dynamics and the different management procedures are simulated. Work of this description is envisioned to be carried out in 2003 and 2004 within the F-project.

5. Acknowledgements

Funding for this work came from the Dutch Ministry of Agriculture, Nature Management and Fisheries and from the fishery sector (via “Produktschap Vis”) under the “F-project”, which aims to analyse and improve the stock assessments of North Sea flatfish.

6. Literature

Van Beek, F.A. and Pastoors, M. A. 1999. Evaluating ICES catch forecasts: the relationships between implied and realised fishing mortality. ICES C.M. 1999/R:04

Brander, K. (1987). "How well do working groups predict catches?" Journal du Conseil / Conseil International pour l'Exploration de la Mer **43**: 245-252.

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

Darby, C. 2002a. Over-estimation bias in the North Sea cod short-term forecasts. Appendix B – Working Document WA2. WGMG. ICES C. M. 2002/D:01

Darby, C. 2002b. Estimation bias in the North Sea short-term forecasts. Appendix C – Working Document WA3. WGMG. ICES C. M. 2002/D:01

Darby, C. D. and S. Flatman (1994). Virtual population assessment: version 3.2 (windows/dos) user guide. Information technology series, Ministry of agriculture, fisheries and food Directorate of fisheries research. **1**.

ICES 1983. Report of the North Sea Flatfish Working Group (Copenhagen, 20 -25 September 1982). ICES C.M. 1983/Assess:1

ICES 1985. Report of the North Sea Flatfish Working Group (Copenhagen, 3 -8 September 1984). ICES C.M. 1985/Assess:1

ICES 1985. Report of the North Sea Flatfish Working Group (Copenhagen, 27 March -2 April 1985). ICES C.M. 1985/Assess:13

ICES 1986. Report of the North Sea Flatfish Working Group (Copenhagen, 27 February -5 March 1986). ICES C.M. 1986/Assess:13

ICES 1987. Report of the North Sea Flatfish Working Group (Copenhagen, 15 -17 October 1986). ICES C.M. 1987/Assess:7

ICES 1988. Report of the North Sea Flatfish Working Group (Copenhagen, 15 -21 October 1987). ICES C.M. 1988/Assess:9

ICES 1989. Report of the North Sea Flatfish Working Group (Copenhagen, 20-26 October 1988). ICES C.M. 1989/Assess:6

ICES 1990. Report of the North Sea Flatfish Working Group (Copenhagen, 20 -26 October 1989). ICES C.M. 1990/Assess:6

ICES 1991. Report of the North Sea Flatfish Working Group (Copenhagen, 18-24 October 1990). ICES C.M. 1992/Assess:5

ICES 1992. Report of the North Sea Flatfish Working Group (Copenhagen, 17-23 October 1992). ICES C.M. 1992/Assess:6

ICES 1993. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (Copenhagen, 6-14 October 1992). ICES C.M. 1993/Assess:5

ICES 1994. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (Copenhagen, 7-15 October 1993). ICES C.M. 1994/Assess:6

ICES 1995. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (Copenhagen, 6-14 October 1994). ICES C.M. 1996/Assess:8

ICES 1996. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (Copenhagen, 2-10 October 1995). ICES C.M. 1996/Assess:6

ICES 1997. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (Copenhagen, 7-15 October 1996). ICES C.M. 1997/Assess:6

ICES 1998. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (Copenhagen, 6-15 October 1997). ICES C.M. 1998/Assess:7

ICES 1999. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (Copenhagen, 5-14 October 1998). ICES C.M. 1999/ACFM:8

ICES 2000. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (Copenhagen, 11-20 October 1999). ICES C.M. 2000/ACFM:7

ICES 2001. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (Copenhagen, 3-12 October 2000). ICES C.M. 2001/ACFM:7

ICES 2002. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (Hamburg, 19-28 June 2001). ICES C.M. 2002/ACFM:01

Jakobsen, T. and Sparholt, H. 2002. Definition and use of *status quo* F. Working paper for SGWP and SGPA. ACFM May 2002.

Patterson, K. R. (1998). Integrated catch at age analysis, version 1.4. Aberdeen, Marine Laboratory.

Quinn, T. J. and R. B. Deriso (1999). Quantitative fish dynamics, Oxford University Press.

Sparholt, H. 2002. Quality of ACFM advice: How good have forecasts been since 1988? Appendix A – Working Document WA5. WGMG. ICES C. M. 2002/D:01

Sparholt, H. and Bertelsen, M. 2002. Quality of assessments. Working paper for SGPA and WGMG. ACFM May 2002.

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.

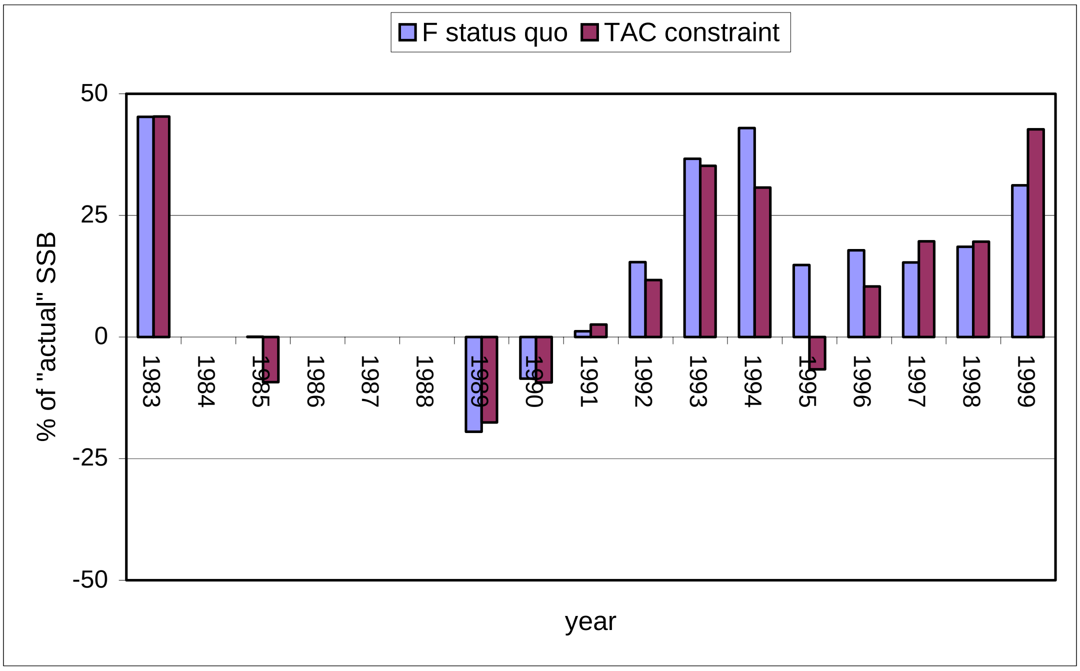


Figure 1. North Sea plaice. Relative difference between predicted and "actual" SSB.

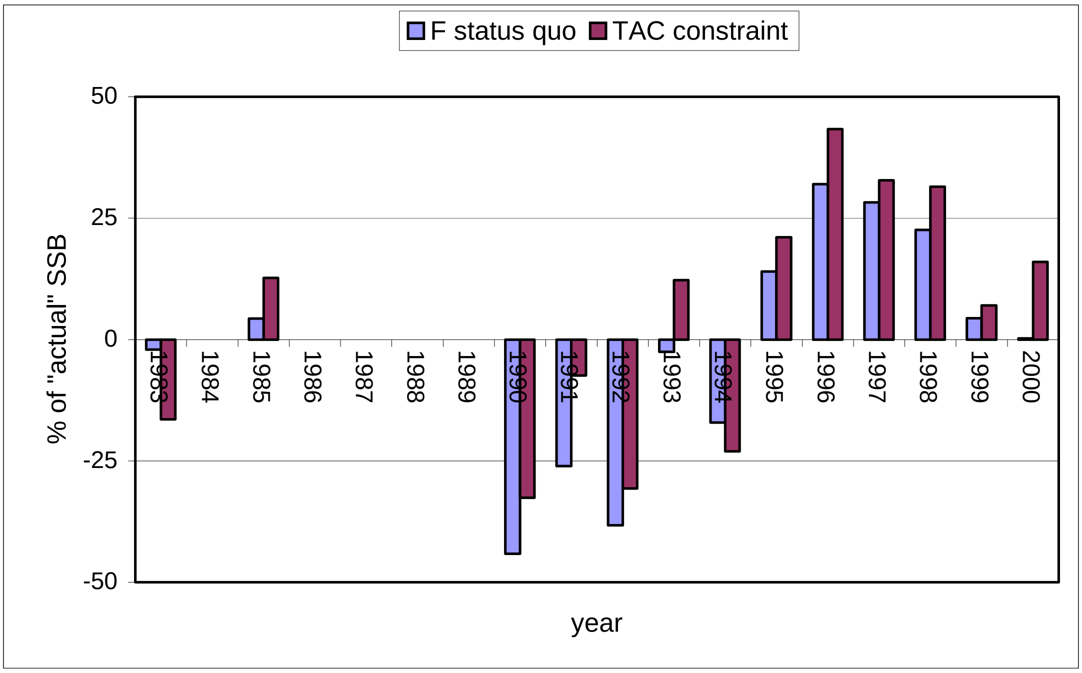


Figure 2. North Sea sole. Relative difference between predicted and "actual" SSB.

Figure 3. North Sea plaice. Gain in accuracy of the SSB prediction under the TAC assumption versus the *status quo* F assumption.

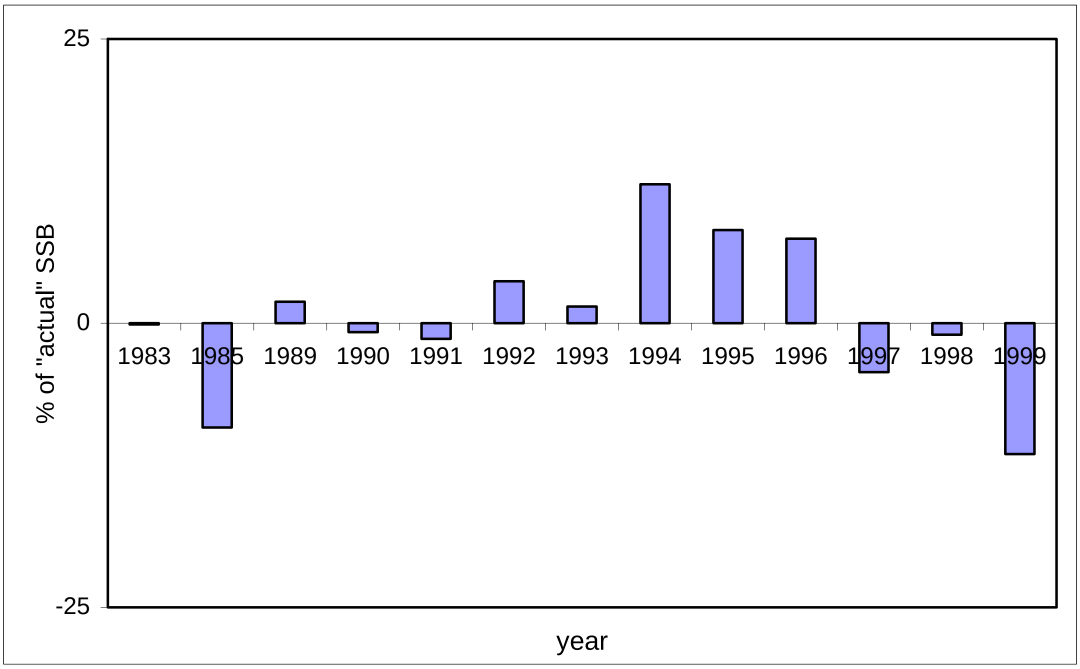
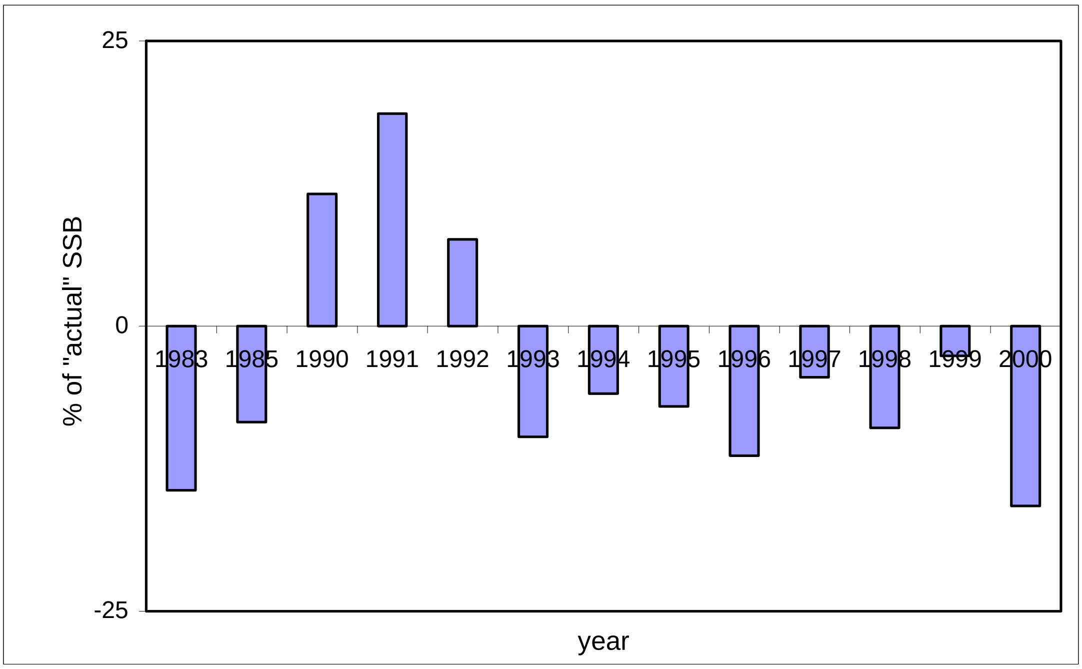


Figure 4. North Sea sole. Gain in accuracy of the SSB prediction under the TAC assumption versus the *status quo* F assumption.



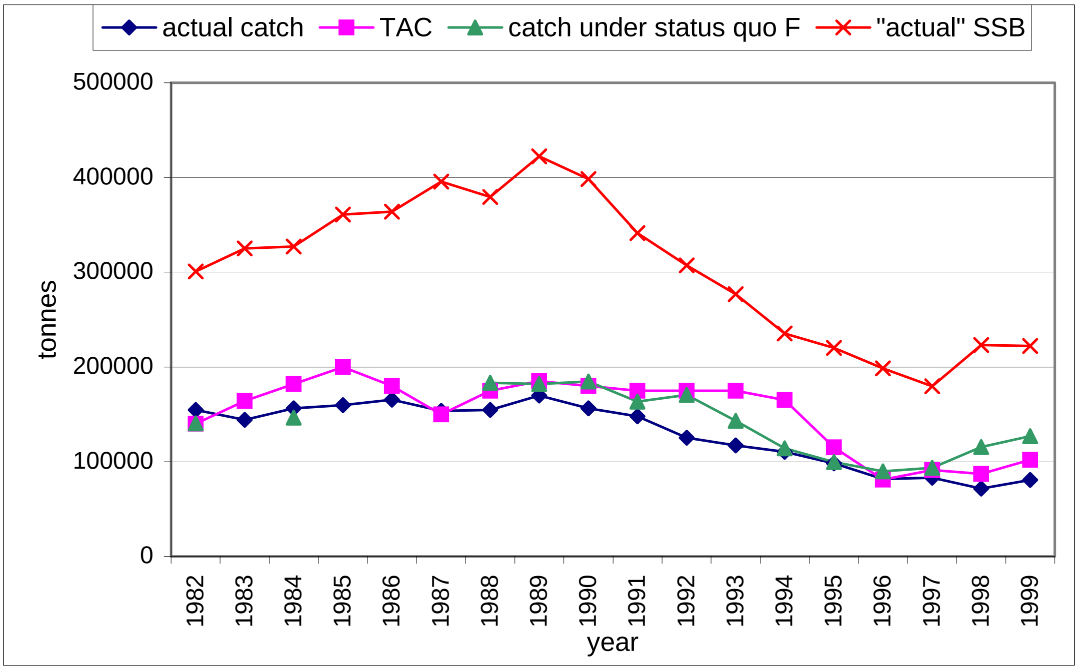


Figure 5. North Sea plaice. “Actual” SSB, actual catch, TAC, and catch under *status quo* F.

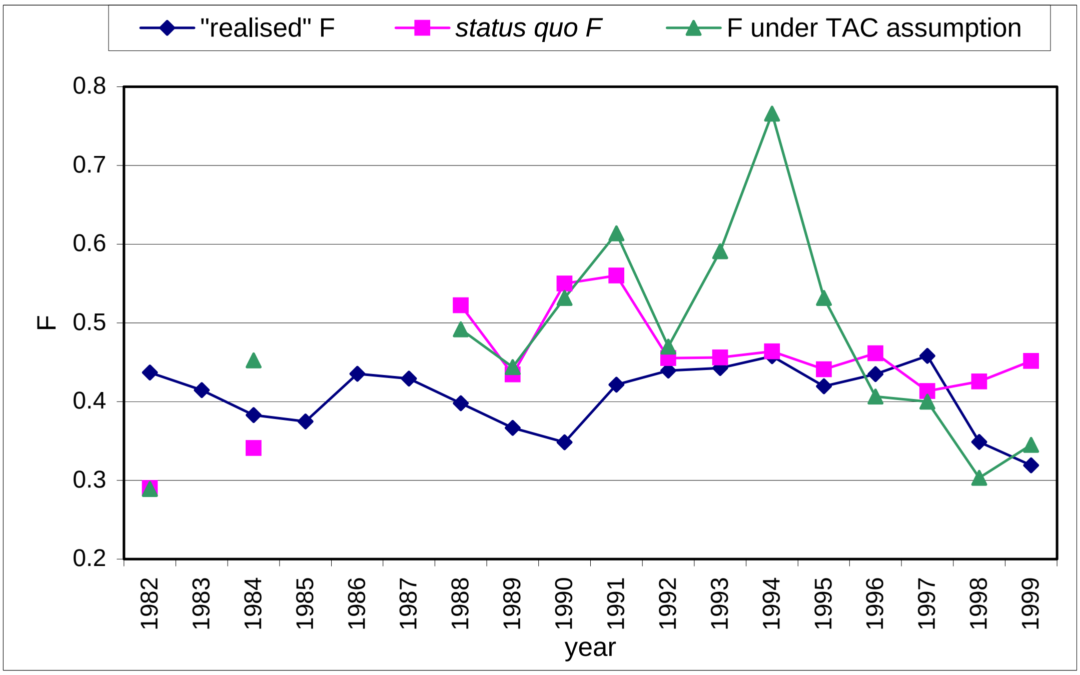


Figure 6. North Sea plaice. “Realised” F, *status quo* F, and F under TAC assumption.

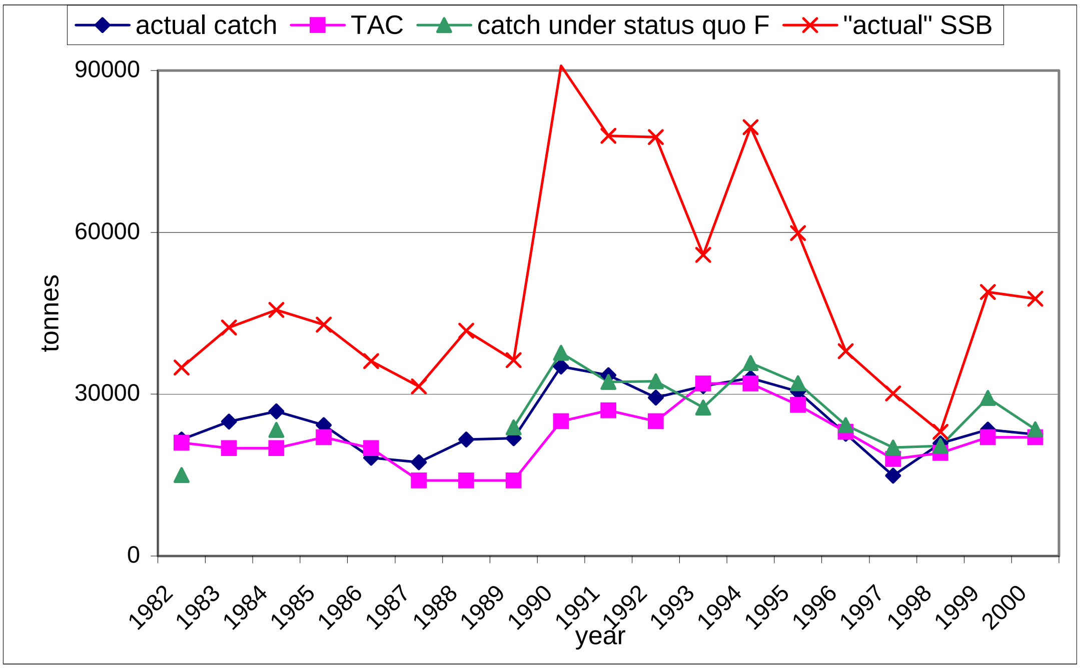


Figure 7. North Sea sole. “Actual” SSB, actual catch, TAC, and catch under *status quo* F.

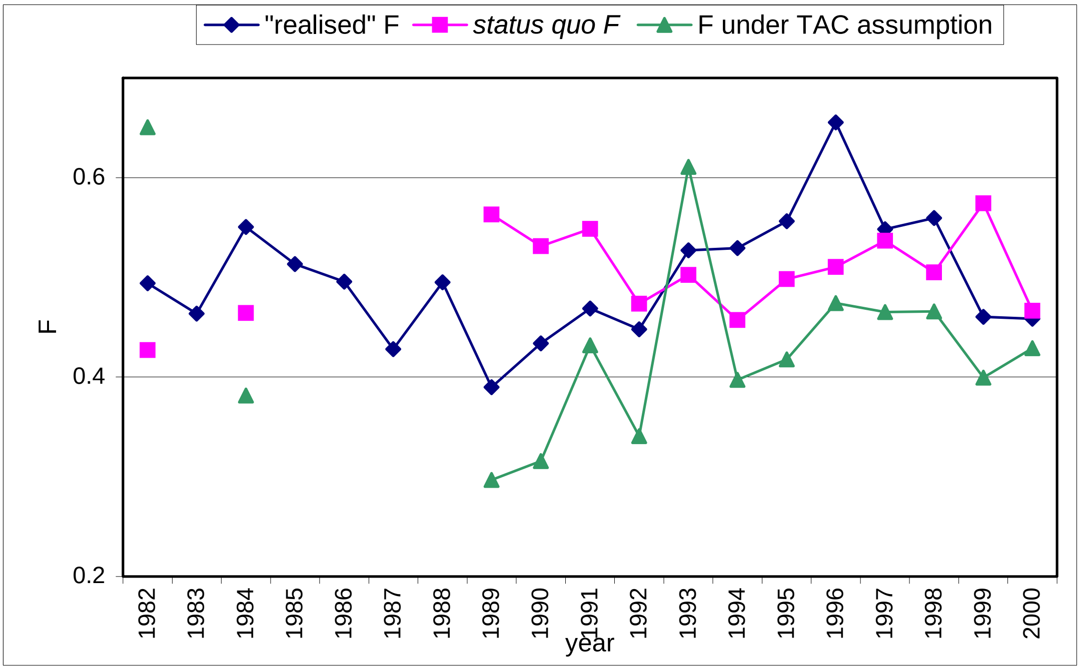


Figure 8. North Sea sole. “Realised” F, *status quo* F, and F under TAC assumption.

Figure 9. North Sea plaice. The error in the prediction of the SSB surviving the current year, (1) in the original forecast with *status quo* F assumption, (2) when the "actual" recruitment would have been known, and (3) when the "actual" stock-at-age would have been known. (Note that the year on the x-axis is the current year and not the year whose SSB is referred to, as in the previous graphs).

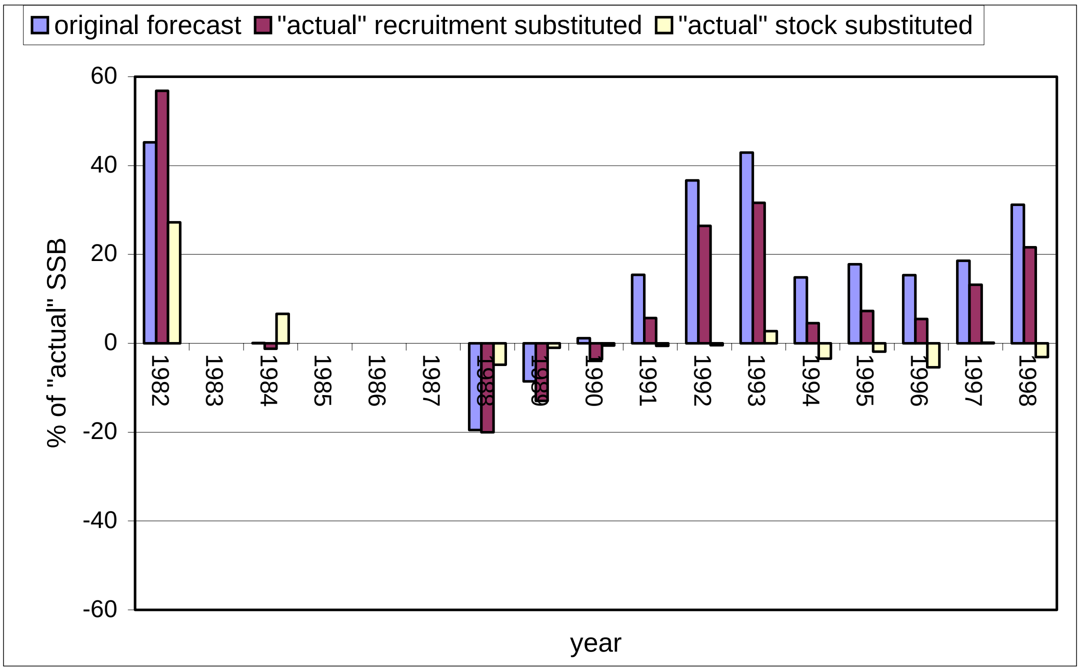


Figure 10. North Sea plaice. The error in the prediction of the SSB surviving the TAC year, (1) in the original forecast with *status quo* F assumption, (2) when the "actual" recruitment would have been known, and (3) when the "actual" stock-at-age would have been known. *Status quo* F is assumed for the TAC year as well. (Note that the year on the x-axis is the current year and not the year whose SSB is referred to).

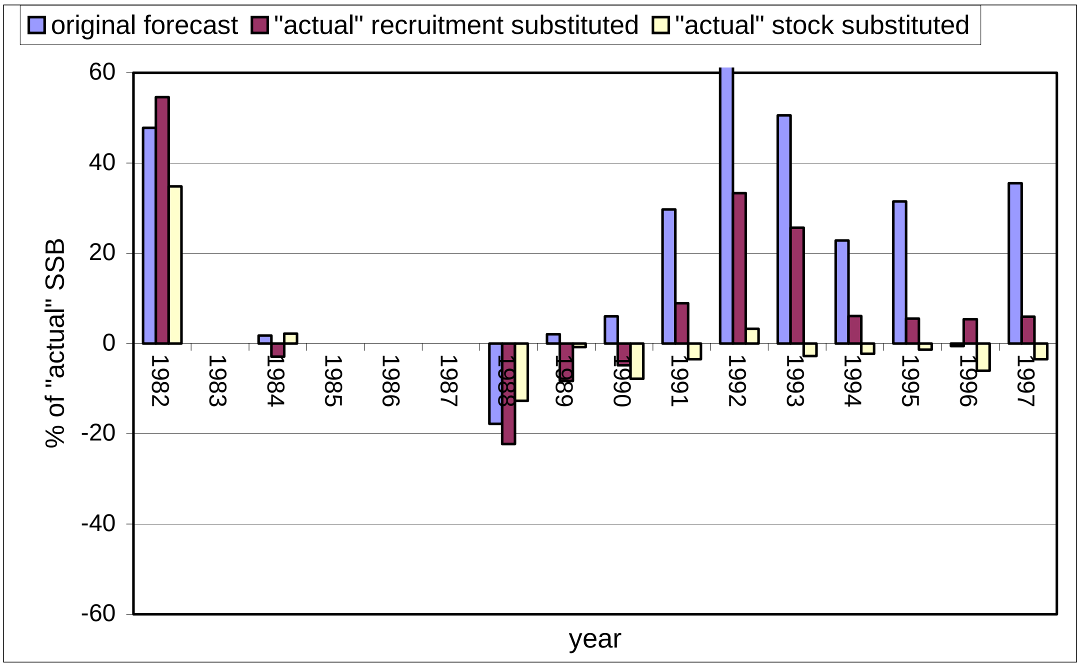


Figure 11. North Sea sole. The error in the prediction of the SSB surviving the current year, in the original forecast with *status quo* F assumption, and when the “actual” stock-at-age would have been known.

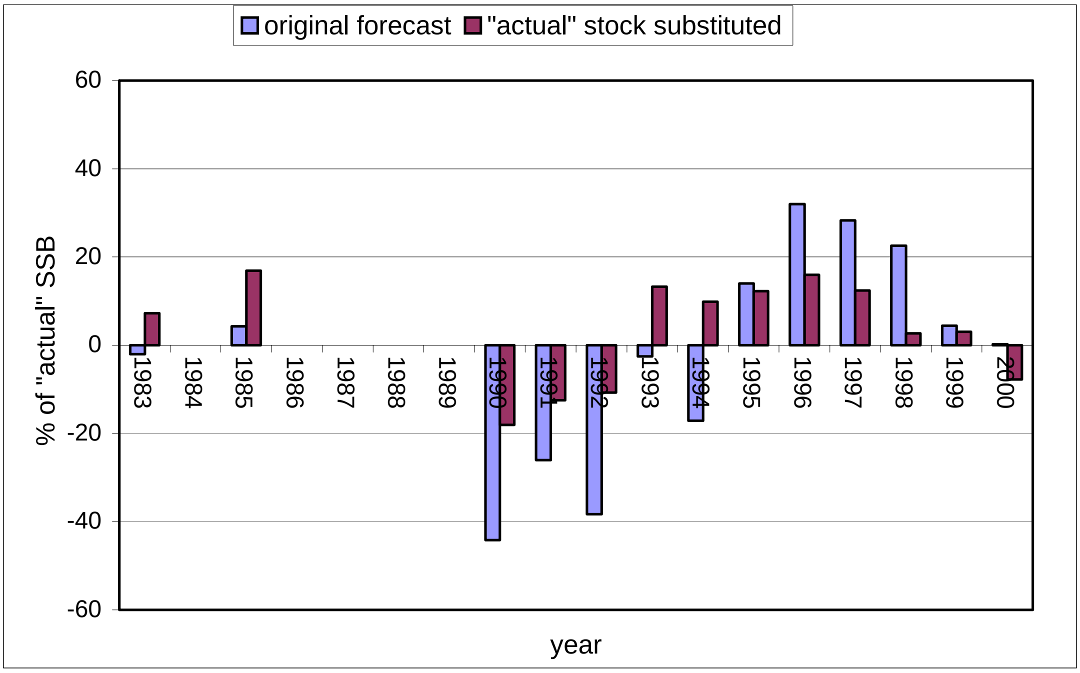


Figure 12. North Sea plaice. The error in the prediction of the SSB surviving the current year, in the original forecast with *status quo* F assumption, and when the actual weights-at-age in the stock and in the catch would have been known.

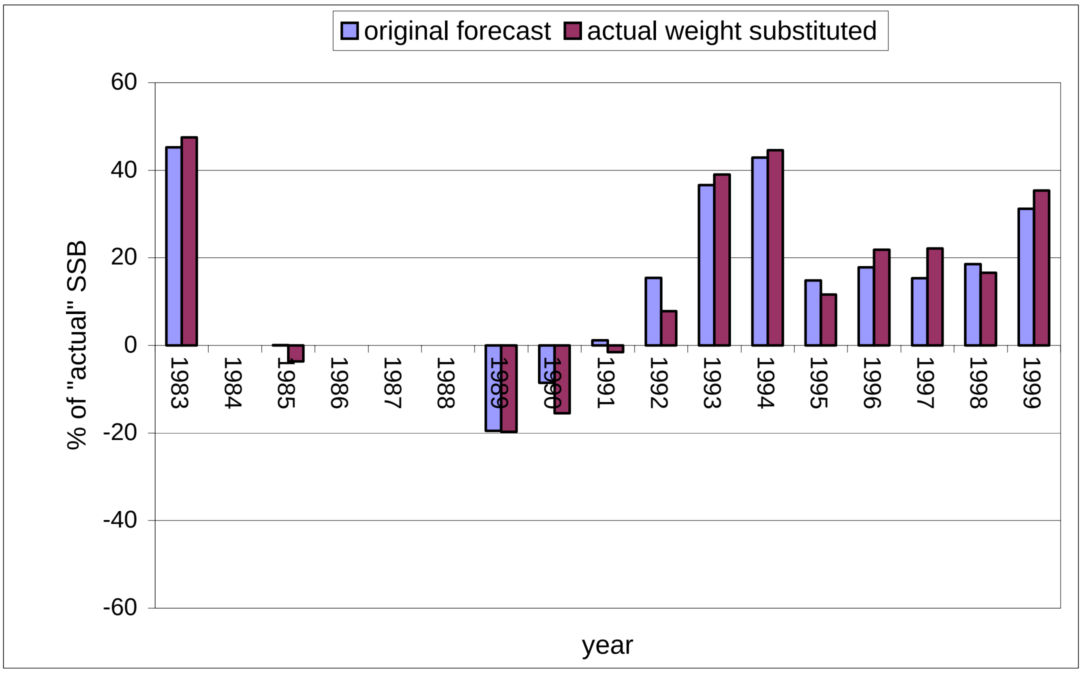


Figure 13. North Sea sole. The error in the prediction of the SSB surviving the current year, in the original forecast with *status quo* F assumption, and when the actual weights-at-age in the stock and in the catch would have been known.

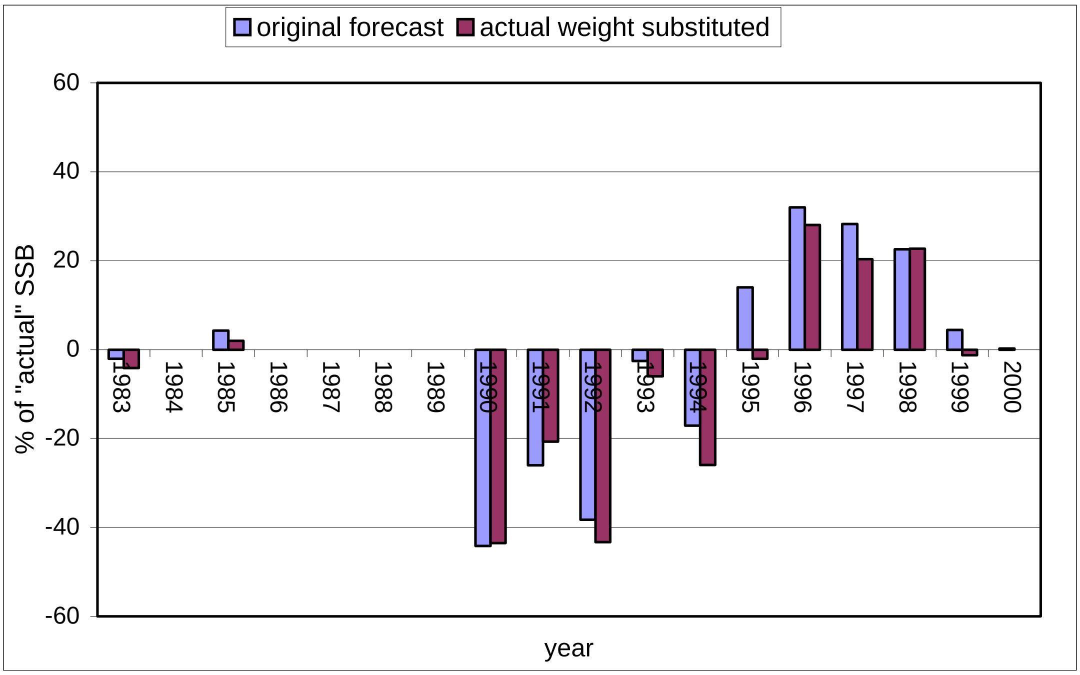


Figure 14. North Sea plaice. Weights-at-age in the stock.

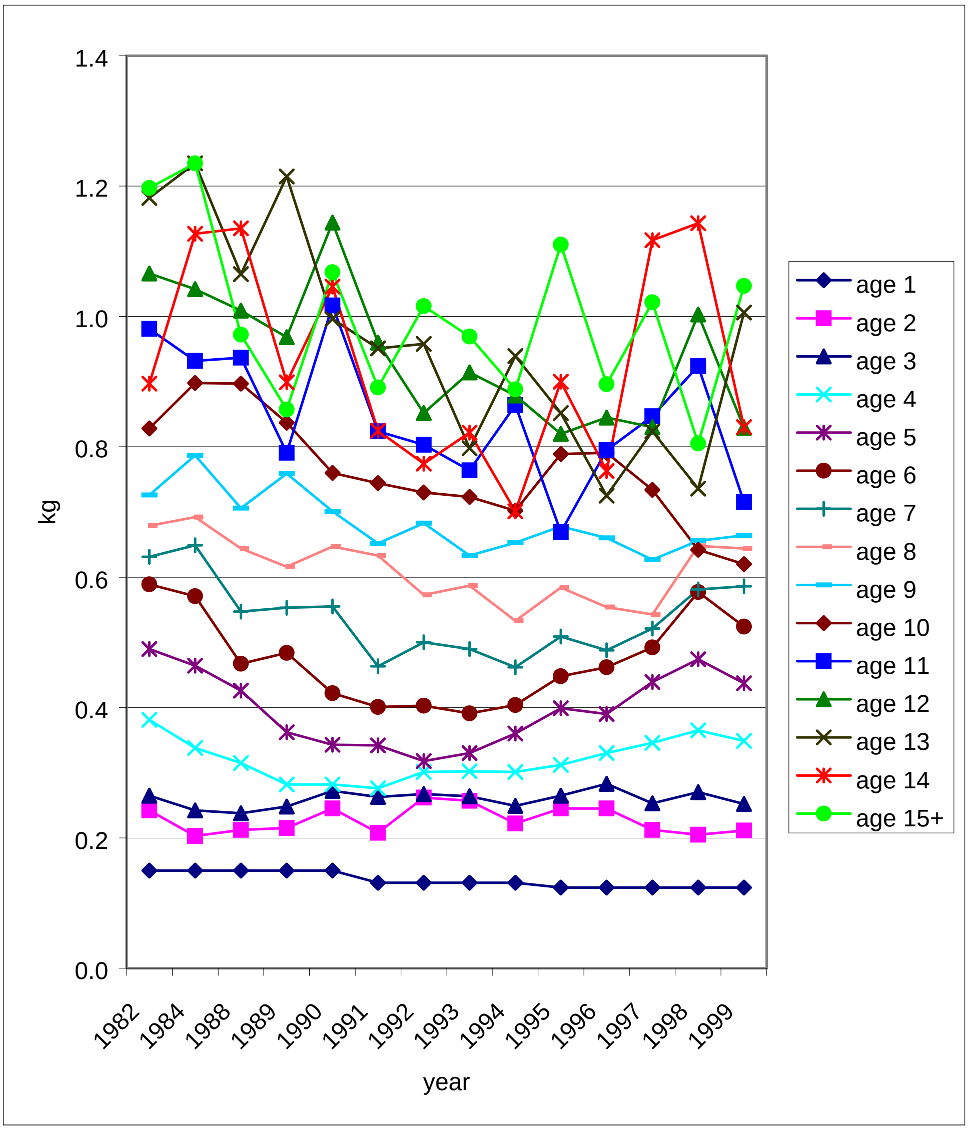


Figure 15a. North Sea plaice. Weights-at-age per year class.

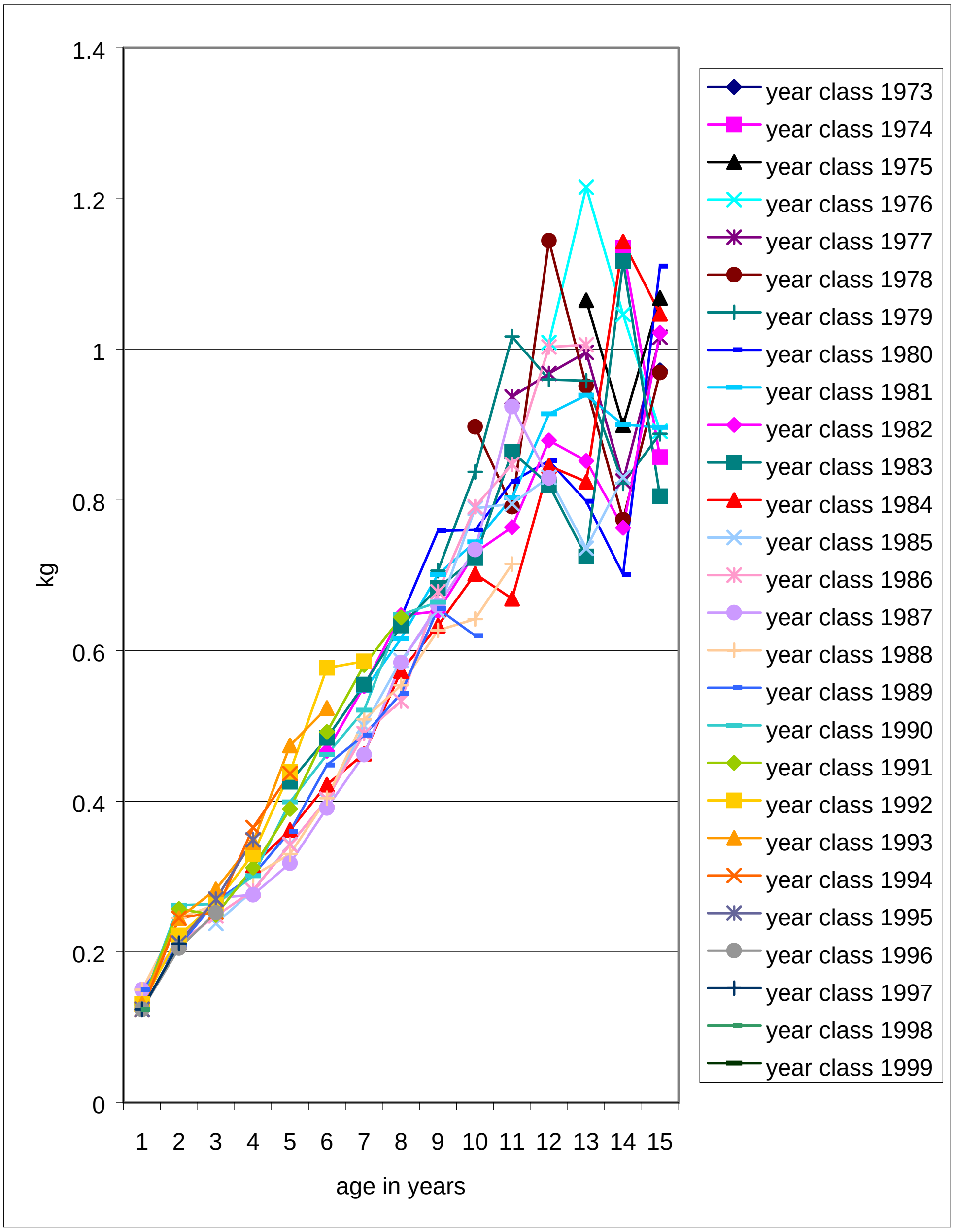


Figure 15b. North Sea plaice. Weights-at-age per year class, for the ages 10-15+.

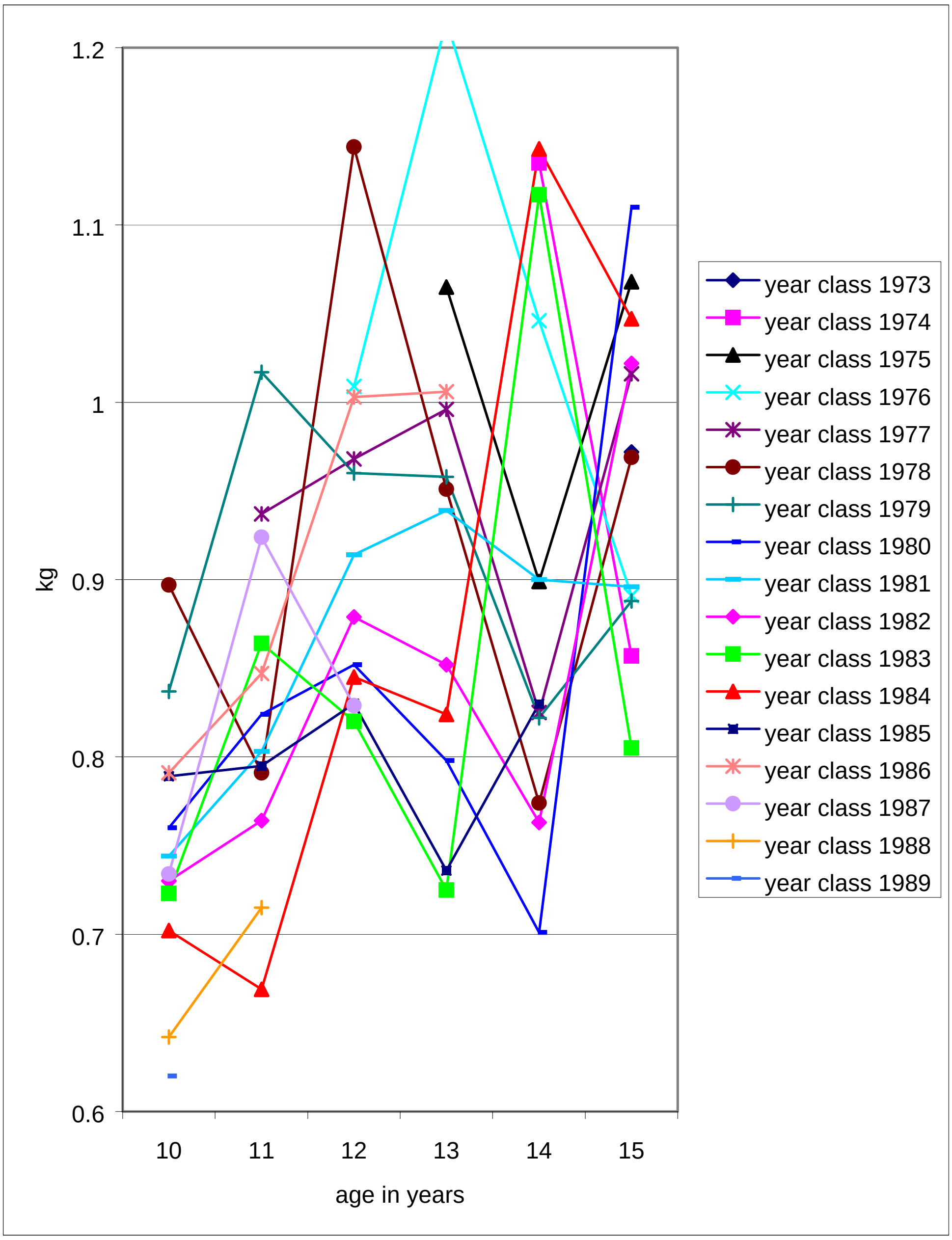


Figure 16. North Sea plaice. Prediction errors in SSB surviving the current year in (1) the conventional forecasts with *status quo* F assumption, (2) when the 3-year-averages of the "smoothed" weight values are substituted, and (3) when the "smoothed" values of actual weight are substituted. Weight in the stock data as well as weight in the catch data are changed.

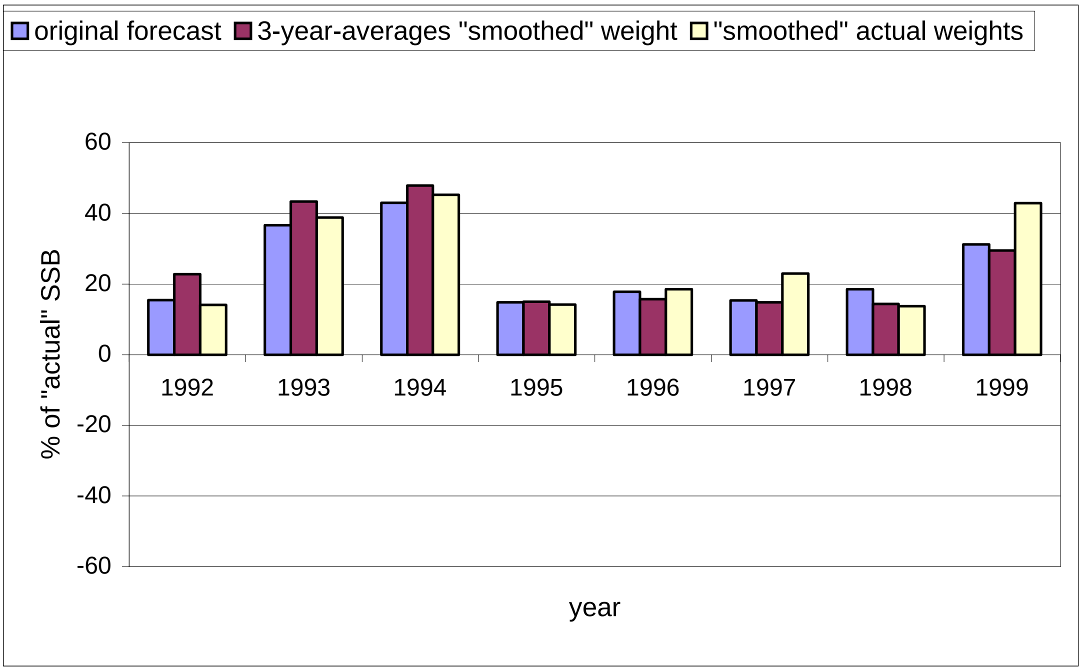


Figure 17. North Sea plaice. Prediction errors in SSB surviving the current year in the original forecasts with *status quo* F assumption, and when the “actual” relative exploitation-at-age is substituted.

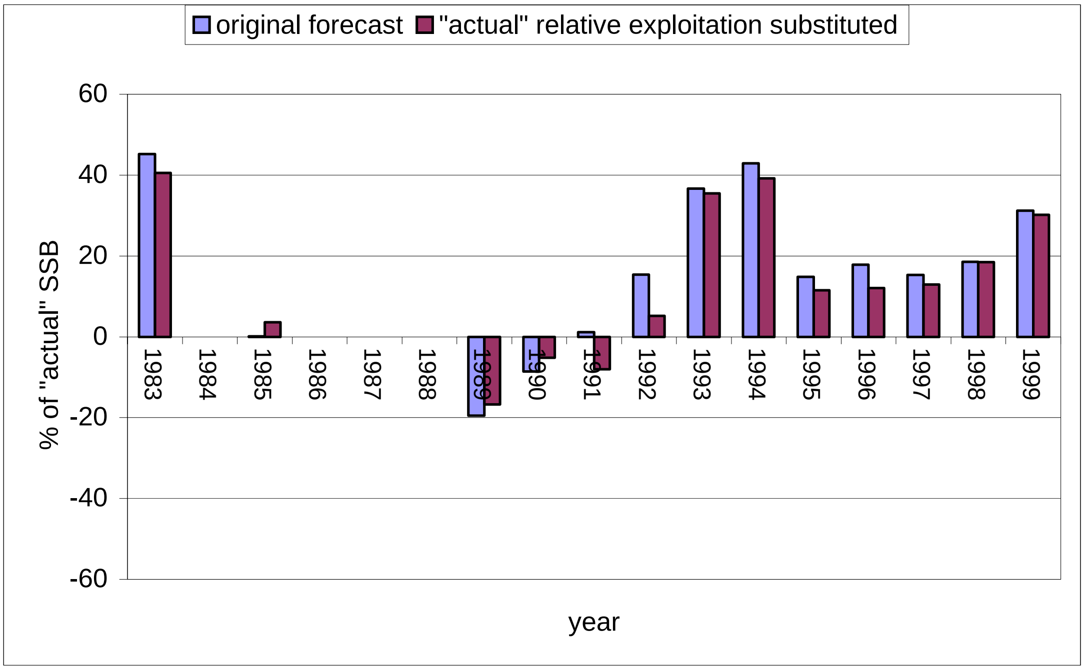


Figure 18. North Sea sole. Prediction errors in SSB surviving the current year in the original forecasts with *status quo* F assumption, and when the “actual” relative exploitation-at-age is substituted.