Wageningen IMARES
Institute for Marine Resources & Ecosystem Studies

Location IJmuiden
P.O. Box 68
1970 AB IJmuiden
The Netherlands
Tel.: +31 255 564646
Fax: +31 255 564644

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

Location Den Helder
P.O. Box 57
1780 AB Den Helder
The Netherlands
Tel.: +31 22 363 88 00
Fax: +31 22 363 06 87

Location Texel
P.O. Box 167
1790 AD Den Burg Texel
The Netherlands
Tel.: +31 222 369700
Fax: +31 222 319235

Internet: www.wageningenimares.wur.nl
E-mail: imares@wur.nl

Report
C033/07

Bayesian analysis of research vessel surveys: trends in North Sea plaice abundance

Hans Bogaards, Lisa Borges, Marcel Machiels, Sarah Kraak

Commissioned by: Ministerie van Landbouw, Natuur en Voedselkwaliteit
Postbus 20401
2500 EK Den Haag

Table of Contents

Uitgebreide Nederlandse samenvatting

Plaats binnen het F-project

In het F1-werkpakket van het F-project houden we ons bezig met verbetering van de toestandsbeoordeling van schol en tong. Problemen rond de onzekerheid en bias in de toestandsbeoordeling en de gegevens die daarvoor worden gebruikt, worden onderzocht in een serie van kleinere deelstudies, die elk een probleem bestuderen. In vier deelstudies, producten A10, A11, A12 en A14, benaderen we het probleem van de onzekerheid in de toestandsbeoordeling. Dit rapport betreft product A11, onzekerheid in de toestandsbeoordeling van Noordzee schol aan de hand van een “survey-based assessment” model.

Onzekerheid

In bestandsschattingen zijn twee bronnen van onzekerheid te onderscheiden: onzekerheid ten gevolge van de gebruikte gegevens en onzekerheid ten gevolge van de aannames die gemaakt worden om het model te gebruiken. In een serie van vier deelstudies proberen we inzicht te krijgen in het aandeel van beide onzekerheidsbronnen in de totale onzekerheid van de bestandsschatting. In de eerste drie deelstudies onderzoeken we drie structureel zeer verschillende modellen voor de bestandsschatting. Deze modellen doen elk verschillende aannames over de populatiedynamica en de manier waarop je deze kunt reconstrueren vanuit de gegevens. Vergelijking van de modeluitkomsten (vierde deelstudie) geeft ons een beeld van de onzekerheid die veroorzaakt wordt door de modelkeuze. Binnen elk van de eerste drie deelstudies variëren we modelaannames, waardoor we inzicht krijgen in de onzekerheid veroorzaakt door deze aannames. Door het toepassen van de Bayesiaanse methode krijgen we informatie over de onzekerheid die door de gegevens veroorzaakt wordt. In deze deelstudie is het model dat we onderzoeken een zogenaamd “survey-based assessment” model. Het model wordt alleen geëvalueerd met behulp van schol gegevens.

Het is van belang te beseffen dat de methode die hier gepresenteerd wordt twee nieuwe elementen heeft: de Bayesiaanse benadering en het feit dat het een “survey-based assessment” model betreft. Deze twee nieuwigheden staan in principe los van elkaar.

De Bayesiaanse benadering

Bayesiaanse statistiek wordt de laatste jaren meer en meer gebruikt voor bestandschattingen binnen en buiten ICES. Twee redenen om Bayesiaanse statistiek te gebruiken zijn:

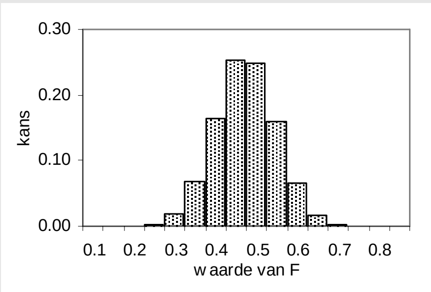
(1) dat onzekerheden in de schattingen expliciet gemaakt kunnen worden / wordt

(2) dat informatie uit andere bronnen meegenomen kan worden in een analyse.

Het feit dat de parameters met onzekerheid geschat worden maakt het mogelijk om deze onzekerheid ook door te vertalen naar de prognoses en te presenteren in de vorm van risicotabellen. Met zulke tabellen, waarin bijvoorbeeld de kans dat het bestand onder Blim komt wordt aangegeven voor verschillende vangstopties, hebben de beheerders een instrument om beter de risico's in te schatten van de verschillende opties. In de traditionele optietabellen wordt de onzekerheid niet getoond, en wordt ten onrechte de indruk gewekt dat als de beheerder een bepaalde TAC kiest we zeker weten welke SSB er in het jaar daarna zal zijn.

1. Onderzekerheden expliciet gemaakt.

In de Bayesiaanse statistiek wordt voor elke te schatten parameter (bijvoorbeeld visserijsterfte F) de waarschijnlijkheid berekend dat deze een bepaalde waarde heeft, gegeven de gebruikte gegevens. Er wordt dus niet een puntschatting van de parameter gegeven, maar een kansverdeling, zoals in de fictieve figuur hieronder te zien is.



Volgens deze fictieve figuur is kans dat F = 0.45 of F = 0.5 erg groot, maar er bestaat ook een kleine kans dat F = 0.3 of F = 0.65. Deze verdeling kan smal of breed zijn, wat overeenkomt met respectievelijk grote zekerheid of grote onzekerheid over de parameter. In de traditionele benadering maken we de mate van onzekerheid niet zichtbaar.

2. Kennis of informatie uit andere bronnen meenemen

Als men vooraf al een idee heeft rond welke waarde een bepaalde parameter zal zitten, bijvoorbeeld gebaseerd op andere studies, dan kan deze kennis meegenomen worden. Het is bijvoorbeeld mogelijk dat een analyse van een ei-survey aangeeft dat de paaibiomassa (SSB) in een bepaald jaar tussen de 280 000 en de 300 000 ton ligt. Het is essentieel dat die andere studie op andere gegevens gebaseerd is, zoals aantallen eieren uit een ei-survey; dezelfde gegevens, zoals vangstgegevens, mogen niet twee keer worden gebruiken. Met deze informatie wordt dan voor elke te schatten parameter de waarschijnlijkheid berekend dat deze een bepaalde waarde heeft, gegeven de gebruikte data èn gegeven deze extra informatie. Op deze manier kunnen resultaten van verschillende studies aan het model toegevoegd worden.

Survey-based assessment (SURBA)

Toestandsbeoordelingen die alleen op survey gegevens zijn gebaseerd noemt wel ook wel SURBA ("survey-based assessment") schattingen. Onder een survey wordt hier verstaan: een onderzoeksreis waarin relevante biologische informatie wordt verzameld die een beeld geeft van de omvang en verspreiding van die populaties waar specifiek op gevist wordt. Er zijn een aantal verschillende surveys, die elk hun eigen karakteristieken hebben voor wat betreft de tijd van het jaar waarin gevist wordt, het vistuig dat gebruikt wordt en het gebied van de zee dat bevist wordt. Vissen staat in deze context gelijk aan bemonsteren, oftewel aan het nemen van steekproeven waarmee een beeld kan worden verkregen van de gehele populatie.

In deze deelstudie wordt informatie van drie surveys gebruikt: (i) de "beam trawl survey" die jaarlijks in augustus plaats vindt met het onderzoeksschip ISIS en die het zuidelijke deel van de Noordzee bemonstert (afgekort BTS-I); (ii) de "beam trawl survey" die jaarlijks in augustus en september plaats vindt met het onderzoeksschip TRIDENS en die het centrale deel van de Noordzee bemonstert (afgekort BTS-T); (iii) de "sole net survey" die jaarlijks in september wordt uitgevoerd met het onderzoeksschip ISIS en zich voornamelijk op kustzones richt (afgekort SNS). Gezamenlijk dekken deze drie surveys een belangrijk gedeelte van het verspreidinggebied van schol in de Noordzee.

De tijdens de surveys waargenomen aantallen per leeftijd en per jaar dienen als input voor SURBA. Op basis van een statische analyse kunnen schattingen worden gemaakt betreffende trends in aanwas, sterfte en paaibestand. Voor alle duidelijkheid: SURBA geeft geen absolute, maar relatieve schattingen van aanwas en paaibestand. De waarde van SURBA heeft dus alleen betrekking op het beoordelen van trends. De sterfte schatting is in principe vergelijkbaar met die van conventionele ICES assessments, maar vereist wel een juiste aanname voor wat betreft de relatieve vangbaarheid van de verschillende leeftijdsklassen in een survey.

Het model dat in deze deelstudie wordt gebruikt is een modificatie van een reeds gepubliceerd model (Cook, 1997). In het model wordt de aanname gemaakt dat totale sterfte kan worden opgesplitst in natuurlijke sterfte M (die normaliter als een constante wordt beschouwd) en visserijsterfte F, en dat de laatste het product is van een jaareffect F_y en een leeftijdseffect F_a. Het leeftijdseffect wordt ook wel de selectiviteit van visserijsterfte genoemd en wordt voor (een deel van) de tijdreeks constant verondersteld. Deze eigenschap wordt "separability" genoemd en zorgt voor een reductie van het aantal te schatten parameters. Nochtans staat een dataset van Y jaar en A leeftijdsklassen niet meer schattingen toe dan Y–1 voor het jaareffect F_y en A–1 voor het leeftijdseffect F_a. Schatting van F_y voor het eerste jaar in de tijdreeks en schatting van F_a voor de jongste leeftijdsklasse zijn daarom niet mogelijk.

De nieuwigheid van onze analyse is dat we de parameters van het model op een Bayesiaanse manier hebben geschat. Deze benadering vergemakkelijkt de schatting van onzekerheid in de parameters en maakt het in principe mogelijk kennis uit onafhankelijke studies te verdisconteren. Een kracht van Bayesiaanse analyse is voorts de transparantie waarmee onderlinge relaties tussen parameters kunnen worden uitgedrukt. In deze deelstudie wordt daar een voorbeeld van gegeven door de ontwikkeling van visserijsterfte in de tijd te modeleren als een zogenoemde "random walk", zodat de schatting van F_y voor een bepaald jaar afhankelijk wordt van de schatting voor het jaar ervoor. Dit zorgt voor een veel "gladder" verloop van visserijsterfte en maakt mogelijk duidelijk wat de lange termijn ontwikkeling is.

Resultaten

De beste fit van het model op BTS-I data werd verkregen door het leeftijdseffect van visserijsterfte F_y vrij te laten variëren van jaar op jaar en onafhankelijke leeftijdseffecten te schatten voor de periode 1985-1999 en voor de periode 2000-2005. De keus voor het jaar 2000 als breekpunt werd ingegeven door resultaten van de meest recente ICES assessment, waarin de selectiviteit van visserijsterfte vrijgelaten wordt om per jaar te variëren. De bevindingen van ICES komen overeen met onze resultaten: sinds 2000 is de visserijsterfte vooral heel hoog op 2-jarige vis, en relatief minder hoog op oudere vis (Figuur 3).

Vanuit statistisch oogpunt mag de beste fit van dit model op BTS-I data geen verrassing zijn, aangezien dit ook het model is met de meeste vrije parameters. Een model met veel parameters zal altijd beter op een specifieke dataset passen dan een model met weinig parameters. De generaliseerbaarheid naar andere (hypothetische) datasets wordt echter kleiner naarmate een model op maat van een specifieke dataset wordt gemaakt. Wanneer de verbetering in model fit wordt gecorrigeerd voor de toename van het aantal vrije parameters, blijkt nog steeds dat het model significant verbeterd is door de introductie van onafhankelijke leeftijdseffecten voor de periodes 1985-1999 en 2000-2005. Toevoeging van een “random walk” voor de schatting van F_y voegt dan nog maar weinig toe (Figuur 5).

De geschatte trends in aanwas komen redelijk goed overeen met die van ICES. Sinds het midden van de jaren tachtig hebben 1997 en 2002 de hoogste aanwas gekend (Figuren 3,6,7). In de analyse van SNS data komen de begin jaren negentig ook als mogelijke piekjaren naar voren (Figuur 7), maar dit wordt niet bevestigd in overige analyses noch in ICES schattingen. Schattingen van het paaibestand op BTS-I data laten zien dat dit in de loop van de jaren negentig gedaald is tot ongeveer de helft van het niveau in de eind jaren tachtig (Figuur 3). Waar ICES schattingen een tijdelijke toename in de periode 2000-2003 laten zien, laat onze analyse op BTS-I data een tijdelijke toename in de eind jaren negentig zien (Figuur 3). De analyse op BTS-T data suggereert echter een geleidelijke toename van het paaibestand sinds de eind jaren negentig (Figuur 6).

De grootste discrepantie tussen de analyses op verschillende survey data onderling, alsook met ICES schattingen, zit in de schatting van visserijsterfte. Waar BTS-T schattingen sinds de eind jaren negentig een redelijk stabiele en lage F suggereren, laten BTS-I schattingen een sterk fluctuerende en veel hogere F zien. De analyse op basis van BTS-I schattingen suggereert een piek in visserijsterfte in de tweede helft van de jaren negentig (Figuur 4), een bevestiging van de ICES schattingen. Het niveau van de visserijsterfte in de ICES schattingen ligt tussen dat van onze schattingen op basis van BTS-I en BTS-T data, respectievelijk.

Discussie

De resultaten van onze analyse maken duidelijk wat de voordelen en de beperkingen zijn van “survey-based assessment”. De uiteenlopende signalen die in de verschillende surveys doorkomen geven mogelijk een beeld van de uiteenlopende ontwikkelingen die zich afspelen in de verschillende subpopulaties van schol. Hierbij dient in aanmerking te worden genomen dat geschatte trends in sterfte sterk beïnvloed worden door migratie tussen subpopulaties. Het is moeilijk om op basis van verschillende surveys tot een eenduidige conclusie te komen voor wat betreft het totale scholbestand in de Noordzee. Integratie van de verschillende surveys in een enkele analyse vergt noodzakelijke aanpassingen aan het model.

De resultaten van onze analyse maken tevens inzichtelijk wat de voordelen zijn van de Bayesiaanse benadering. De incorporatie van een “random walk” voor de ontwikkeling van visserijsterfte in de tijd demonstreert de flexibiliteit van SURBA in een Bayesiaans framework. Hoewel het resulterende “gladde” verloop wellicht onrealistisch is in de zin dat werkelijke fluctuaties bewust worden gladgestreken, geeft het mogelijk een helder beeld van de lange termijn ontwikkeling.

De afhankelijkheid van relatief kleine steekproeven wordt algemeen als grootste tekortkoming van SURBA beschouwd. Integratie van surveys is daarom des te belangrijker. Een juiste specificatie van de onderlinge afhankelijkheid van parameters speelt daarbij een cruciale rol, alsmede het gebruik van kennis uit onafhankelijke studies. De Bayesiaanse benadering levert hiervoor een krachtig conceptueel framework.

1. Introduction

For a number of years, fishery management authorities have tried to limit fishing mortality on many fished stocks through a total allowable catch (TAC) regime. ICES provides annual advice on TACs, derived from stock assessments. While data from research vessel surveys are used to calibrate the margins of the age-structured population matrix, conventional ICES VPA assessment is dominated by the commercial catch-at-age data. Consequently, estimated stock trends may be misleading whenever official landings figures are not representative of the true catches – e.g. due to illegal landings, discards or by-catch in other fisheries – or whenever significant changes in fishing effort have not been taken into account. The proportion of the catch not included in the official landings figures is likely to increase with a restrictive TAC regime. This would not only obscure stock decline in ICES assessments, but could in turn give rise to an overly optimistic advice when restriction of fishing effort is warranted.

Research vessels tend to perform routine hauls at specified locations. Although not affected by misreporting or changes in fishing effort, they are sensitive to changes in spatial distribution of a fish stock from year to year. Also, because the fishing effort of surveys constitutes only a fraction of the commercial fishing effort, research vessel data are inherently less precise (in the sense that observed numbers-at-age are more affected by measurement error) as compared to commercial catch data. Even though there are usually multiple independent surveys that can serve as input for assessment, they often do not agree with each other. Each survey has distinct characteristics regarding geographic area, time of year and fishing gear used, which are all likely to affect the measurements.

Cook (1997) first presented an analytical model for survey-based stock assessment, which has formed the basis for SURBA, a Fortran-based package for the analysis of research vessel data (Needle, 2003). Although the method does not allow estimation of absolute population size, it can reveal fisheries-independent trends in fish stocks. Also, SURBA yields an estimate of fishing mortality which (provided that the catchability is specified correctly) should be comparable to that of conventional assessments. However, SURBA estimates of fishing mortality are well-known to be sensitive to noise in the data and adequate specification of catchability remains elusive. For this reason, Cook (1997) recommended the investigation of alternative parameterizations, but developments have been scarce in recent years.

A serious drawback of SURBA is its inability to provide a quantification of uncertainty for relevant parameters. For example, estimation of fishing mortality is facilitated by assuming separable temporal and age effects. Even though standard errors can be provided for the point estimates of these separable effects, the standard error of their product is not defined. Since the precautionary approach has become a key concept in fisheries management, uncertainties in assessment, be it survey-based or not, have to be taken into account. Numerous stochastic assessment methods have been proposed (see Lewy and Nielsen, 2003, plus references contained therein), of which those that fall within the Bayesian framework have the advantage that prior beliefs about parameters can be incorporated into the estimation procedure (Punt and Hilborn, 1997). Although Bayesian methods have been criticized for their potential to give too much weight to vested interests (Cotter et al., 2004), they have proved very insightful when applied to virtual population analysis (Virtala et al., 1998).

The purpose of this paper is to apply SURBA in a Bayesian framework. Specifically, we investigate the effect of an alternative parameterization for fishing mortality. Sensitivity to variability in natural mortality and catchability is also assessed. We use research vessel data on North Sea plaice (*Pleuronectes platessa*) from two beam trawl surveys (BTS-I and BTS-T, by the research vessels Isis and Tridens respectively) and from the sole net survey (SNS). The BTS-I supplies information on the southern North Sea, the BTS-T on the central North Sea and the SNS on the coastal zones. Together, the three surveys cover the distribution area of the North Sea plaice stock.

2. Methods

2.1 Data

The analysis is based on survey data used in the most recent assessment for North Sea plaice (ICES, 2006). The SNS provides data on the longest period, from 1970 up to and including 2005. However, this survey only provides data on the first three age classes (age range 1-3). The two beam trawl surveys (BTS) provide data on nine age classes (age range 1-9), but for a shorter period of years: the Isis survey (BTS-I) from 1985 onwards and the Tridens survey (BTS-T) from 1996 onwards. Missing values (SNS: 4.6%, BTS-I: 1.0%, BTS-T: 2.2%) were replaced with the mean of the non-missing values over the years for the age class concerned. Zero values were replaced with an abundance of 0.1 fish. In addition to survey data, we used observed data on weight-at-age and assumed maturity to be negligible at the age of recruitment, 50% in age classes two and three, and 100% at older ages.

2.2 Model

The basic equations of the model are described by Cook (1997). The model is a separable model with regard to fishing mortality: fishing mortality at age by year $F_{a,y}$ is the product of a year effect F_y and an age effect F_a. We made a few modifications. First, as the model only allowed estimation of A–1 age effects in fishing mortality, we let u_{ry} denote the abundance index at the age of recruitment in the month of the survey (as opposed to the start of the year). Second, as only Y–1 year effects in fishing mortality can be estimated, we assumed the cumulative mortality-at-age in the first year to be similar to that in the second year of survey data. These modifications ensured that the model is not under-specified with respect to the number of parameters that need to be estimated: abundance indices at recruitment u_{ry}, the temporal component of fishing mortality F_y, the age-effect of fishing mortality F_a, and overall measurement error (on a log-normal scale) ε. All these parameters were given vague uniform priors (see Table 1).

For comparability to Cook (1997), we initially set natural mortality $M_{a,y}$ to 0.1 and relative catchability q'_a to 1 for all age classes. Variability, both between years and between ages for $M_{a,y}$ and for all ages relative to recruitment for q'_a, was introduced by giving these parameters increasingly vague priors. It should be noted, however, that both $M_{a,y}$ and q'_a essentially act as a correction to fishing mortality and cannot be estimated independently.

The model was further refined by breaking down fishing mortality into two distinct periods and applying separate age-effect to both periods. Recent analyses (ICES, 2006) have demonstrated a marked shift in the age-specific pattern of fishing mortality, in that older fish seem to be less affected by fishing effort in recent years than they were before. As this shift seems to have taken place at the turn of the century, we set F_{a1} for the period up to and including 1999 and set F_{a2} for the period from 2000 onwards. In order to study an alternative parameterization for fishing mortality, we modeled the temporal component of fishing mortality F_y as a random walk, with $\log(F_y) = \log(F_{y-1}) + Z_y$, where Z_y denotes a continuous, purely random process. Giving this process an informative prior enables us to put sensible boundary conditions on the magnitude by which fishing mortality can change from one year to the next.

2.3 Simulations

Results were obtained by simulation using WinBUGS, version 1.4. Subsequent calculations and graphs were produced using R, version 2.3.1. In standard runs, 5 Markov chain Monte Carlo (MCMC) simulations using Gibbs sampling were used to construct posterior distributions. Each run was initialized by drawing random starting points from roughly overdispersed distributions for the estimable parameters. In standard runs, each MCMC chain ran for 2000 iterations of which the first half were discarded in order to diminish the effect of the starting distribution. From the second half of each MCMC simulation, one in five iterations was stored to remove the possible effect of autocorrelation. Thus, we stored a total of 200 times 5 realizations from which we obtained quantiles of the posterior distributions.

Convergence was monitored in various ways. First, the trajectories of each parameter were inspected visually in order to assess the extent to which the various chains mixed. Second, convergence of each parameter was assessed by calculating a scale reduction factor, which expresses the scale for which the posterior distribution might be reduced if iterative simulations were continued indefinitely (Gelman et al., 2004). At convergence, these factors should be one. Third, for selected simulations where convergence was achieved after the default number of simulations, we performed extended runs to confirm that results were indeed the same.

Model fit was evaluated by the deviance, which is defined as –2 times the log-likelihood. Model complexity was measured by estimating the effective number of parameters pD, which can be thought of as the number of unconstrained parameters in the model (with constraints depending both on the data and on the priors). The deviance information criterion (DIC) was used to judge which of the various models would have the lowest predictive deviance. Models are penalized to have larger DIC both by the deviance (the larger this is, the worse the fit) and by pD, thus favoring models with a smaller effective number of parameters.

After fitting the model, spawning stock biomass (SSB) was calculated as the sum of the product of the numbers-at-age, weight-at-age and maturity-at-age. In common with conventional ICES assessments, annual fishing mortality was averaged over the age range 2-6. For graphical output, we used median estimates together with 95% credible intervals (ranging from the 2.5th to 97.5th quantiles of the simulated posterior distributions).

3. Results

Models fitted on the entire BTS-I dataset gave suspiciously high values for F_a in the highest age class, as well as wildly fluctuating estimates of fishing mortality in the last 10 years of observation (Figure 1, upper panels). Upon investigation of the dataset, we decided to leave out the highest age class in analyses. This gave more sensible estimates of fishing mortality, both over age and over time (Figure 1, lower panels). Moreover, by leaving out the highest age class, the estimate of overall measurement error decreased from 0.83 (s.d. 0.05) to 0.56 (s.d. 0.04).

Simulations with the baseline model (age range 1-8) yielded scale reduction factors close to 1 for all parameters (maximum 1.019), confirming our visual impression that the various MCMC simulations converged well. The estimate of overall measurement error corresponds to a c.v. of 61%, suggesting that the model does not fit the data very well (deviance 855; pD 43.4; DIC 898). There is no overall trend in fishing mortality considering the whole of the time series, but estimates still fluctuate rather drastically from one year to the next (see Figure 1d). Estimates of relative recruitment follow the same trends as observed numbers in the youngest age class, but the abundance index at recruitment is overestimated at the beginning of the time series (Figure 2a). This might indicate that the assumption of a constant cumulative mortality-at-age for the first couple of years does not hold. With the exception of 1987, and to a lesser extent 1995/1996, the SSB calculated from the model agrees well with the SSB calculated from the survey data. However, for the final years of observations, SSB tends to be slightly overestimated (Figure 2b).

Simulations with the model modified to include separate estimates of the age effect of fishing mortality also converged well (maximum scale reduction factor 1.023). Overall measurement error is estimated at 0.54 (s.d. 0.04), corresponding to a c.v. of 58%. The increase in model complexity is offset by the improvement in model fit (deviance 840; pD 46.3; DIC 886). The better model fit is apparent from the plots of relative recruitment (Figure 3a) and SSB (Figure 3b). The estimates of the age effect of fishing mortality are markedly different for the periods 1985-1999 and 2000-2005. Fishing mortality was primarily focused at the fourth age class until 2000 (Figure 3c), whereas it shifted to the younger age classes from 2000 onwards (Figure 3d). However, the temporal component of fishing mortality is only marginally stabilized and the trends in fishing mortality are similar to those from the baseline model.

Simulations with the temporal component of fishing mortality modeled as a random walk have trouble in achieving convergence, if fishing mortality in the first year is given a random starting point. If fishing mortality in the first year is fixed to a given value, convergence is reached after the default number of iterations. [Because the mean of the temporal component is scaled to unity, the precise value at which fishing mortality in the first year is fixed does not matter.] If fishing mortality is allowed to change drastically from one year to the next (prior c.v. for change in F_y 50%), results are comparable to the default simulations. Although model fit is less satisfactory, complexity is reduced and overall the model performs slightly better as compared to the baseline model (deviance 856; pD 40.5; DIC 897). If fishing mortality is allowed to change less drastically from one year to the next, the model fits worse yet complexity is further reduced. The loss in model fit is offset by the reduced complexity for prior c.v.'s of 25% (deviance 857; pD 37.7; DIC 895) and 10% (deviance 860; pD 33.9; DIC 894). Further restriction of the temporal change in fishing mortality leads to a loss of predictive power, as assessed for prior a c.v. of 5% (deviance 865; pD 32.0; DIC 897) and 2% (deviance 875; pD 31.8; DIC 907). Simulations with prior c.v.'s of less than 2% do not converge easily, as a significant number of scale reduction factors are above 1.02 even after extended runs. The effect of the use of various prior c.v.'s on the estimates of fishing mortality is shown in Figure 4.

Simulations with separate estimates of the age effect of fishing mortality as well as a restriction on the temporal change in fishing mortality take somewhat longer to achieve acceptable convergence (maximum scale reduction factor 1.018). Using a 10% prior c.v. for change in F_y, overall measurement error is estimated at 0.55 (s.d. 0.04). Compared to the model without restriction on the temporal change in fishing mortality, the model fits the data less well but the effective number of parameters is also reduced (deviance 846; pD 40.0; DIC 886). The less satisfactory fit is apparent from the plots of relative recruitment (Figure 5a) and SSB (Figure 5b). SSB calculated from the survey data still falls within 95% credible intervals for most of the time series, but agreement is poor for the last four years of observation. Estimates of the age effect in the period 2000-2005 (Figure 5c) are similar to those in the period 1985-1999 (Figure 5d).

Variability in $M_{a,y}$ (between years and ages) had no impact on the estimates of fishing mortality, as long as the prior was informative. With a prior c.v. of 25%, neither model fit nor complexity are affected and plots of relative recruitment and SSB look the same. Using increasingly vague priors resulted in significant negative correlations between annual fishing mortality and annual natural mortality, as averaged over the age range 2-6, but on average results remain the same.

Variability in q'_a similarly has a small impact in the estimates of fishing mortality, but the model-fitting algorithm encountered problems if the prior became too vague. Essentially, there is no information about this parameter and posterior distributions reflected uncertainty in the priors. There was a slight tendency for sampled values of q'_a to shift towards higher values for the highest age class in the analysis (this shift was larger with less informative priors). Close investigation revealed that age-specific estimates of relative catchability were positively correlated to estimates of fishing mortality over the previous year, but negatively correlated to those estimates for the following year. The absence of data on survival from the highest age class likely explains this tendency (Table 2).

Models fitted on the BTS-T data set were not affected by exclusion of the highest age class. In order to compare results to those obtained on the BTS-I data set, we nevertheless restricted the analysis to the age range 1-8. Estimates of relative recruitment are significantly higher than numbers observed in the youngest age class, suggesting that the assumption of a constant catchability for all age classes does not hold in the BTS-T data. Still, model results confirm the peak estimates of recruitment in the years 1997 and 2002 (Figure 6a). The SSB calculated from the model agrees well with the SSB calculated from the BTS-T data and shows an increasing trend over the entire period of observations (Figure 6b). This result sharply contrasts with those obtained through BTS-I observations (see Figure 2b). Estimates of fishing mortality are even more divergent. According to BTS-T data, fishing mortality was primarily focused at the age range 5-7 (Figure 6c) and has remained rather stable over the past 10 years (Figure 6d). Moreover, the model suggests a several-fold lower level of fishing mortality than obtained through BTS-I data (see Figure 1d).

Because SNS data were entirely missing for the year 2003, we restricted the analysis to the period from 1970 up to and including 2002. As the SNS only provides data on the first three age classes, we did not look at estimates of SSB or selectivity. Annual fishing mortality was calculated as the arithmetic mean over the age classes two and three only. The model estimates a peak recruitment in 1986, followed by lesser peaks in 1991/1992, 1997 and 2002 (Figure 7a). Estimates of fishing mortality are very unstable over time, and are even higher than those obtained through BTS-I data (Figure 7b). It should be noted, however, that estimates of fishing mortality on the basis of this survey are not very meaningful.

4. Discussion

In this paper, we present fisheries-independent estimates of trends in North Sea plaice abundance. Our estimates are based on survey data and have been obtained by Bayesian analysis. The Bayesian framework allows a natural estimation of uncertainty in the model parameters, and permits the use of prior information or expert opinion.

To our best knowledge, this is the first reported Bayesian survey-based stock assessment. Cook (1997) and Needle (2003) described a methodology for estimating stock trends and fishing mortality without the need for commercial catch-at-age data, and applied their analyses to a number of species. Their estimates regarding plaice are comparable to ours, with a few notable differences. First, in Cook's results, the drop in fishing mortality as of 1988 appeared as a single outlier for the period up to 1994. Our results suggest that fishing mortality declined from 1986 onwards, only to stabilize as of 1990. Second, Needle's analysis suggested large fluctuations in SSB, with a particular high estimate as of 1998. Although our analysis also suggests SSB peaks towards the late nineties, we do not confirm the unusually high estimate of 1998. In our analyses, highest SSB estimates were found in the late eighties. Thereafter, SSB steadily decreased until the late nineties whereas fishing mortality steadily increased during that period.

Our results are partly in agreement with the most recent ICES assessment results on North Sea plaice, where SSB is estimated to have increased in the late eighties, followed by a subsequent decline in the nineties (ICES, 2006). However, our analysis suggests a temporary recovery of SSB in the late nineties, whereas ICES estimates show a temporary increase in the early 2000s (ICES, 2006). Our peak estimates of recruitment in 1986, 1997 and 2002 are also in line with ICES estimates, with the difference that their number of recruits in 1986 is more than twofold that of 1997 and 2002. The latter finding is only confirmed in our analyses based on SNS data. In all our other analyses, these cohorts were approximately of the same size. The greatest disparity arises in the estimate of fishing mortality. As in Cook's results, our analysis on BTS-I data suggests mortality rates around 1. Results obtained through BTS-T data however suggest stable mortality rates below 0.5, whereas results based on SNS data show mortality rates generally between 1 and 2 for the period 1985-2005. For comparison, ICES estimates have mortality rates around 0.6 for the period 1985-2005. Although our estimates of a gradual increase in fishing mortality until the late nineties are confirmed by ICES (2006), the trends from 1998 onwards are unclear.

A possible explanation for the different estimates of fishing mortality, from models based on different survey data, is that each model relates to a different subpopulation. As these subpopulations clearly interact with each other, it is best to think of the North Sea plaice stock as a metapopulation. This illustrates the need to combine various surveys within one assessment, and would involve the accommodation of Cook's model for the purpose of metapopulation analysis. Clearly, analyzing trends in subpopulations is bound to be of limited value if rates of exchange between subpopulations are high. Nevertheless, our results do provide an interesting picture when fishing mortality is interpreted as the sum of both mortality and net migration. For example, the disparity between our results based on BTS-I and BTS-T data could be taken as a difference in fishing mortality between the southern and the central North Sea or as the result of migration of mature plaice from the southern to the central North Sea.

The specification of a random walk for the temporal component may have led to an overly smooth development of fishing mortality in our analysis. In fact, our choice of a random walk was motivated by the desire to achieve a smoothed representation of this development (cf. Cook, 1997). Although the 10% prior c.v. for change in F_y is supported by an analysis of the magnitude by which fishing mortality changed in ICES assessments, it might have been more realistic to incorporate autoregressive terms as well. An advantage of performing SURBA in a

Bayesian framework is the flexibility with which such processes can be incorporated in the estimation procedure. Moreover, the parameters of such a process have a clear interpretation, which is not the case for the parameters of a penalty function (Needle, 2003).

In this paper, we made minimal use of the possibility to construct informative priors. This makes our results all the more comparable to those of Cook (1997), Needle (2003) and ICES (2006), but neglects a main benefit of parameter estimation within the Bayesian framework. Use of external data to formulate a useful stochastic process for the development of fishing mortality naturally would lead to construction of informative priors. The same applies to selectivity and catchability, but there is very little information on these quantities. Moreover, as our analysis demonstrates, these parameters are correlated and independence of priors is another consideration (Cotter et al., 2004).

Finally, our analysis estimates an overall measurement error in the order of 60% c.v.. Although this might indicate the poor performance of our model, it also might indicate the inherent sensitivity to noise in survey data. As Cook (1997) pointed out, the potential gain in accuracy and detail when performing stock assessment on such type of data, which are generally of a small sample size, should be weighed against the loss in precision and their power to detect changes accordingly. In a sense, survey-based assessment aims to trade precision for a potential higher accuracy. Whether this is indeed the case remains to be investigated, but our analysis demonstrates its usefulness as a flexible tool for extracting signals from noisy data.

References

Cook R.M. 1997. Stock trends in six North Sea stocks as revealed by an analysis of research vessel surveys. ICES Journal of Marine Science, 54: 924-933.

Cotter A.J.R., Burt L., Paxton C.G.M., Fernandez C., Buckland S.T., Pan J.X. 2004. Are stock assessment methods too complicated? Fish and Fisheries, 5: 235-254.

Gelman A., Carlin J.B., Stern H.S., Rubin D.B. 2004. Bayesian data analysis, second edition. Chapman & Hall/CRC, Boca Raton, Florida.

ICES. 2006. Report of the Working Group on the assessment of demersal stocks in the North Sea and Skagerrak. 5-14 September, Copenhagen. ICES CM 2007/ACFM:09, 971pp.

Lewy P., Nielsen A. 2003. Modelling stochastic fish stock dynamics using Markov Chain Monte Carlo. ICES Journal of Marine Science, 60: 743-752.

Needle C.L. 2003. Survey-based assessments with SURBA. Working document to the ICES Working Group on Methods of Fish Stock Assessment, Copenhagen, February 2003.

Punt A.E., Hilborn R. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries, 7: 35-63.

Virtala M., Kuikka S., Arjas E. 1998. Stochastic virtual population analysis. ICES Journal of Marine Science, 55: 892-904.

Tables

Table 1. Prior distributions used in the baseline Bayesian SURBA model.

Parameter	Description	Prior distribution	Note
ε	Measurement error (s.d. on log scale)	uniform[0,1]	s.d. 1 on log scale equals c.v. 131%
u_{ry}	Abundance index of recruitment at age 1 by year	uniform[0,max]	max is set to twice maximum observed
F_y	Year effect of fishing mortality	uniform[0,3]	mean of the year effects is scaled to 1
F_a	Age effect (selectivity) of fishing mortality	uniform[0,3]	selectivity at lowest age is not estimable
M	Natural mortality (at age by year)	lognormal(log(0.1), s.d.)	s.d. is varied in sensitivity runs
$q_a^{'}$	Relative catchability (scaled to lowest age)	lognormal(log(1), s.d.)	s.d. is varied in sensitivity runs

Table 2. Correlations between the age effect of fishing mortality in the period 1985-1999 and relative catchability.

Parameters	q_2'	q_3'	q_4'	q_5'	q_6'	q_7'	q_8'
F_2^1	**0.74**	0.05	–0.05	0.02	–0.05	–0.01	0.02
F_3^1	**–0.54**	**0.59**	0.11	0.02	0.03	–0.02	–0.04
F_4^1	–0.01	**–0.57**	**0.55**	0.07	–0.03	–0.01	0.03
F_5^1	0.01	–0.02	**–0.52**	**0.58**	0.03	0.01	–0.02
F_6^1	–0.03	–0.00	–0.05	**–0.62**	**0.50**	0.06	0.00
F_7^1	0.00	–0.03	–0.04	–0.02	**–0.51**	**0.48**	0.10
F_8^1	0.07	0.02	0.04	–0.01	–0.10	**–0.49**	**0.52**

Figures

Figure 1. Selectivity (A, C) and trends in fishing mortality (B, D) obtained through BTS-I data. The upper panel shows results from an analysis on the age range 1-9, whereas the lower panel shows results from an analysis restricted to the age range 1-8. The heavy line is the median estimate, the area between the thin lines is the 95% credible interval.

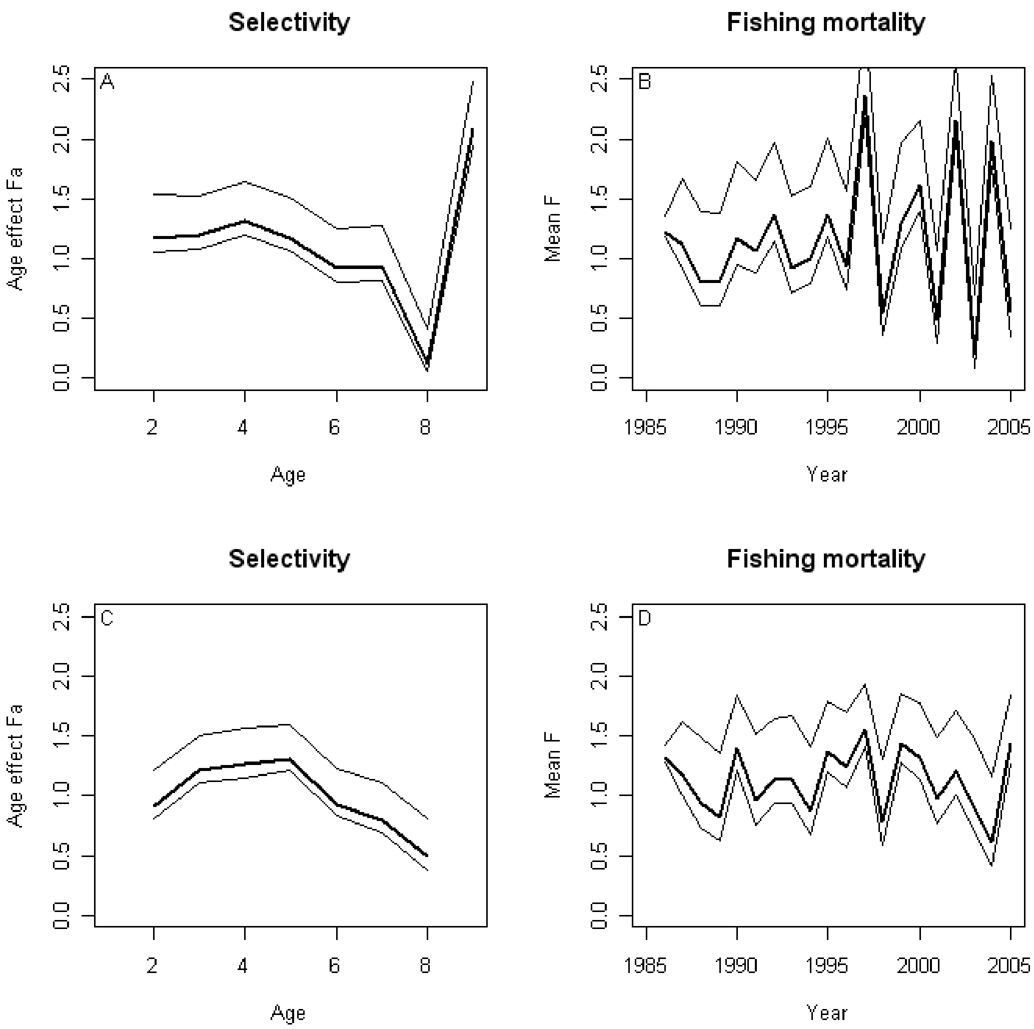


Figure 2. Trends in recruitment at age 1 (A) and spawning stock biomass (B) obtained through BTS-I data (age range 1-8). Circles represent observed number in the youngest age class (A) and SSB calculated from observed abundance (B), both scaled to match model estimates. The heavy line is the median estimate, the area between the thin lines is the 95% credible interval.

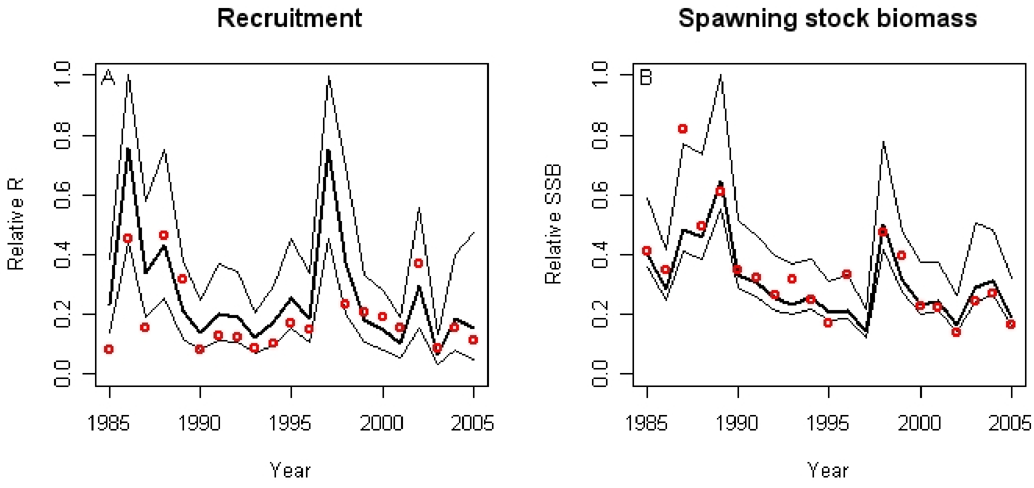


Figure 3. Results from the model with separable age effects for two separate periods, obtained through BTS-I data (age range 1-8). The upper panel shows trends in recruitment at age 1 (A) and spawning stock biomass (B). Circles represent observed number in the youngest age class (A) and SSB calculated from observed abundance (B), both scaled to match model estimates. The lower panel shows the age effects for the periods 1985-1999 (C) and 2000-2005 (D). The heavy line is the median estimate, the area between the thin lines is the 95% credible interval.

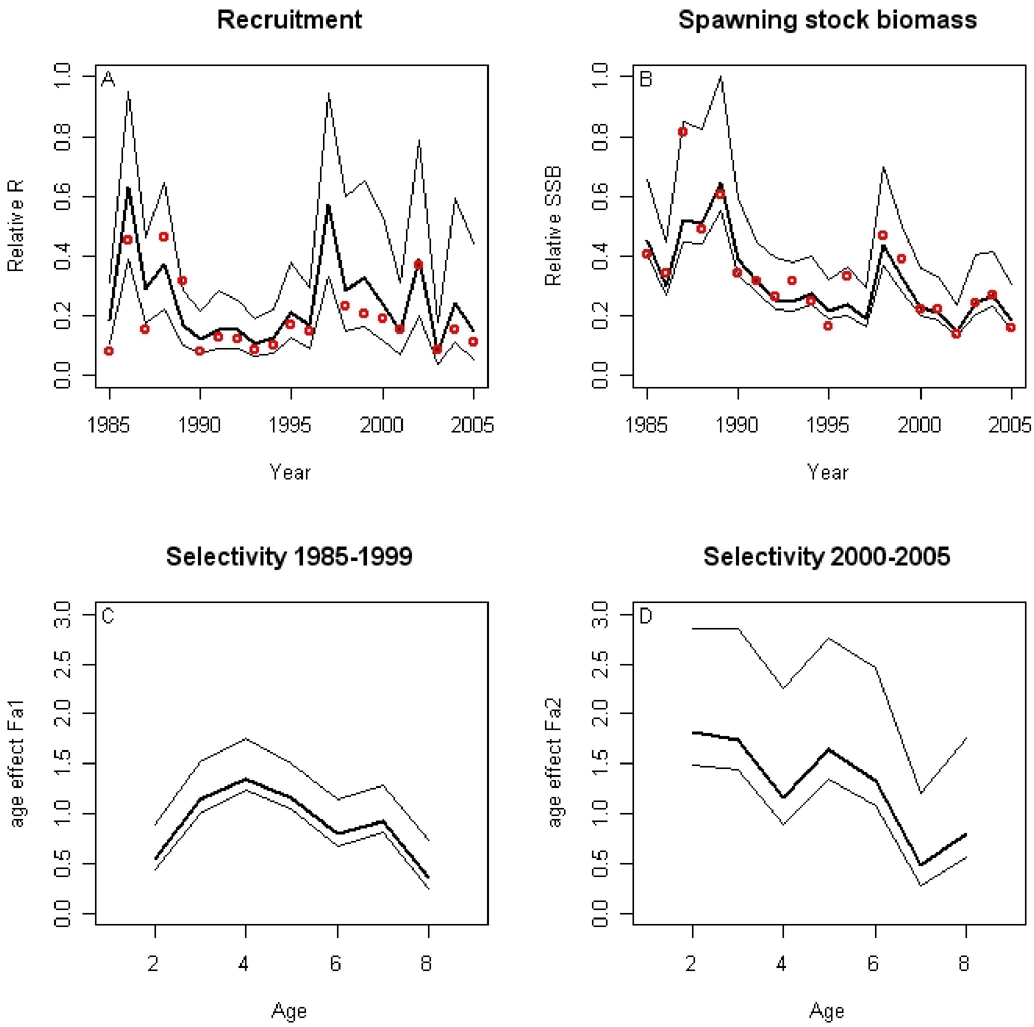


Figure 4. Estimates of fishing mortality, obtained through BTS-I data (age range 1-8), from analyses with the temporal component of fishing mortality modeled as a random walk. The prior c.v. for change in F_y varies from 50% (A), 25% (B), 10% (C) to 5% (D). The heavy line is the median estimate, the area between the thin lines is the 95% credible interval.

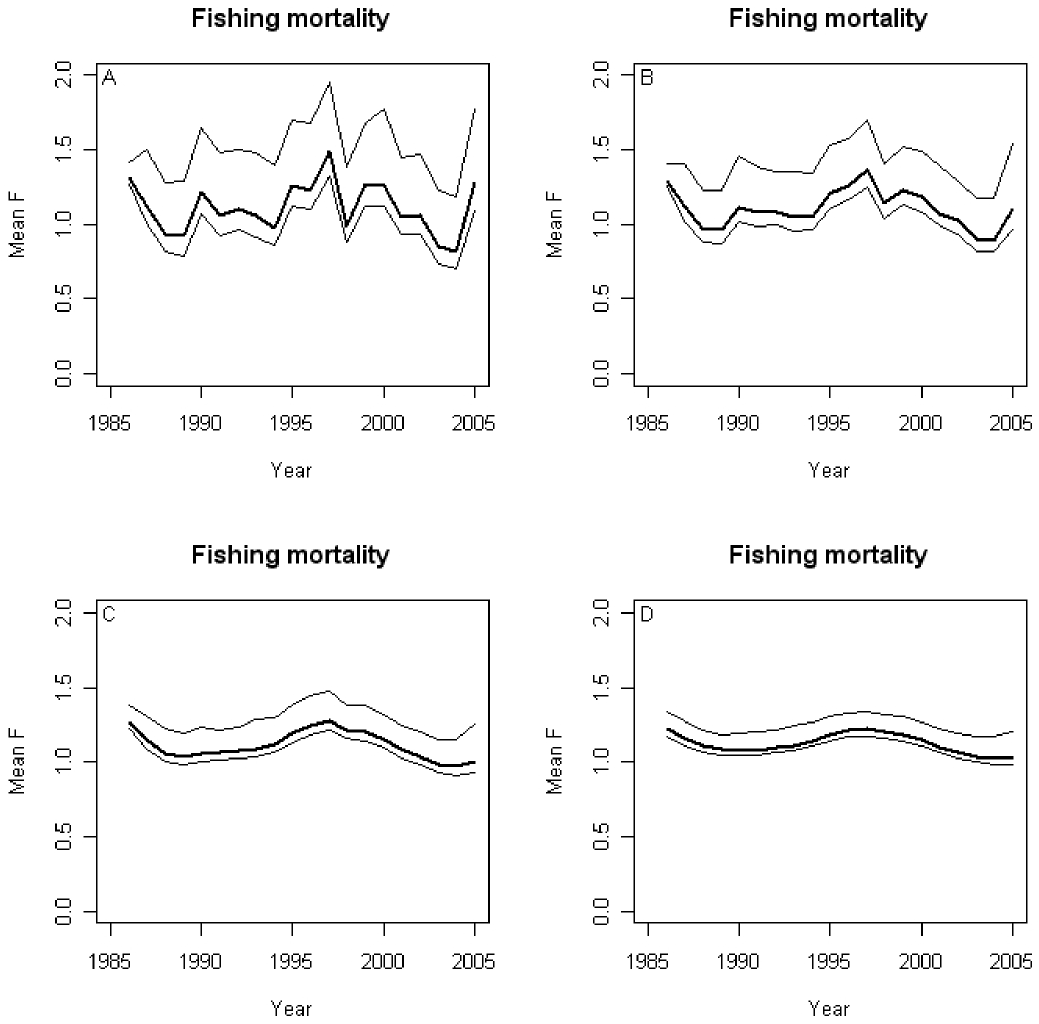


Figure 5. Results from the model with separable age effects for the periods 1985-1999 and 2000-2005, obtained through BTS-I data (age range 1-8), with a prior c.v. for change in F_y of 25%. The upper panel shows trends in recruitment at age 1 (A) and spawning stock biomass (B). Circles represent observed number in the youngest age class (A) and SSB calculated from observed abundance (B), both scaled to match model estimates. The lower panel shows the age effects for the periods 1985-1999 (C) and 2000-2005 (D). The heavy line is the median estimate, the area between the thin lines is the 95% credible interval.

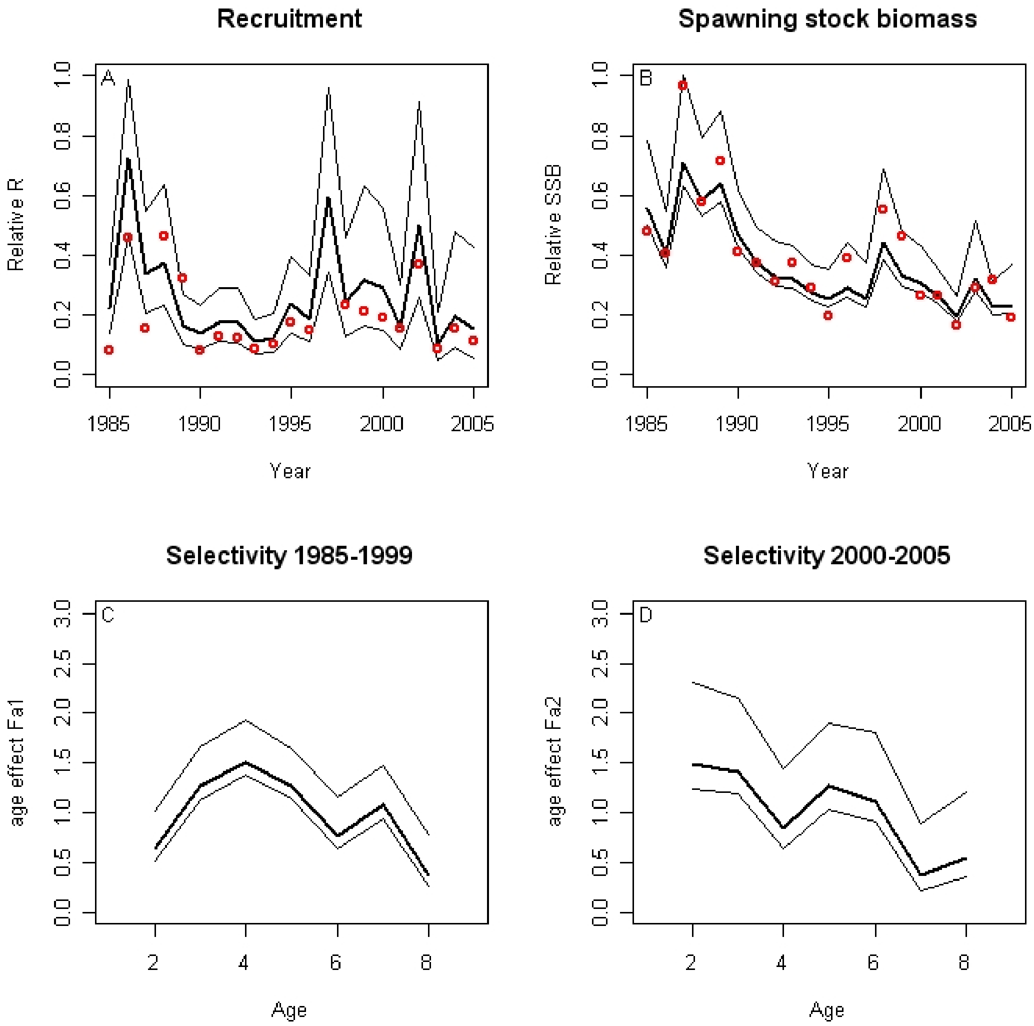


Figure 6. Results from the baseline model, obtained through BTS-T data (age range 1-8). The upper panel shows trends in recruitment at age 1 (A) and spawning stock biomass (B). Circles represent observed number in the youngest age class (A) and SSB calculated from observed abundance (B), both scaled to match model estimates. The lower panel shows the age effect (C) and the trend in fishing mortality (D). The heavy line is the median estimate, the area between the thin lines is the 95% credible interval.

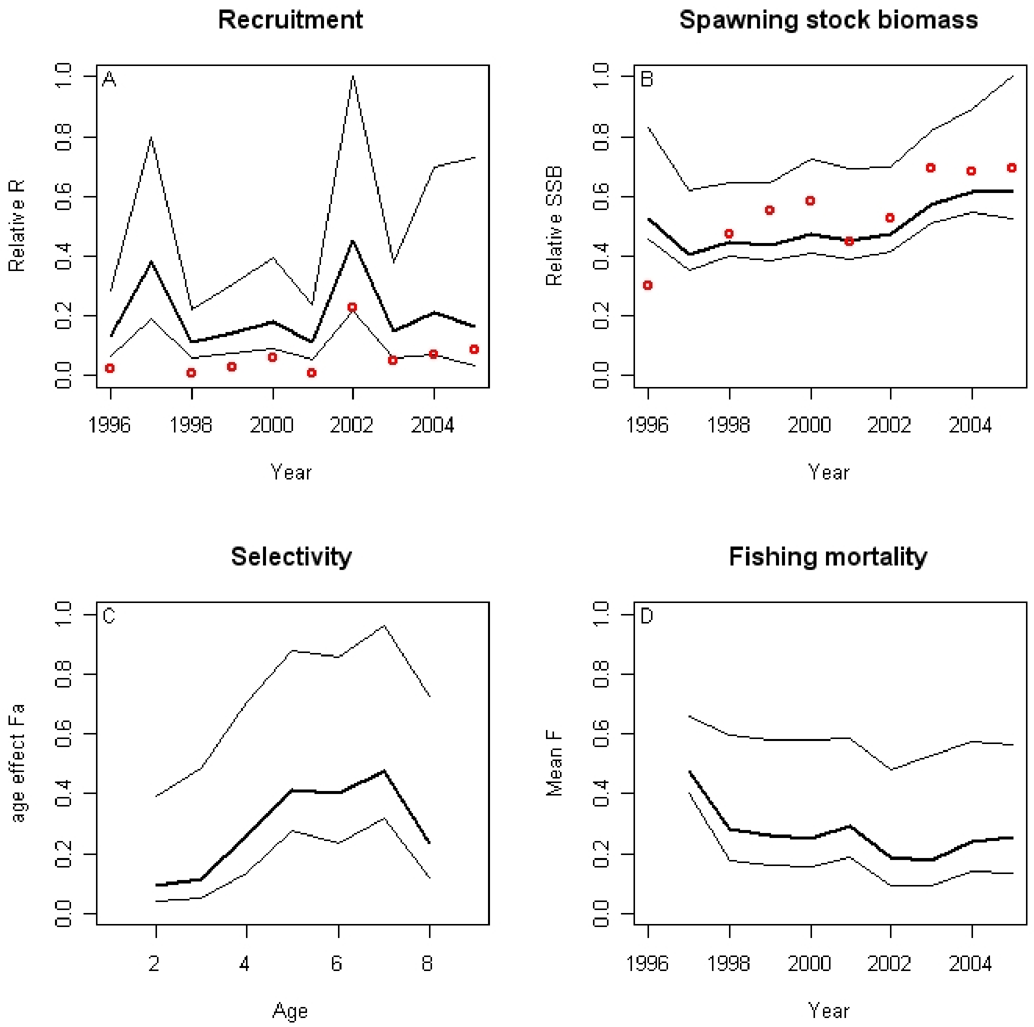
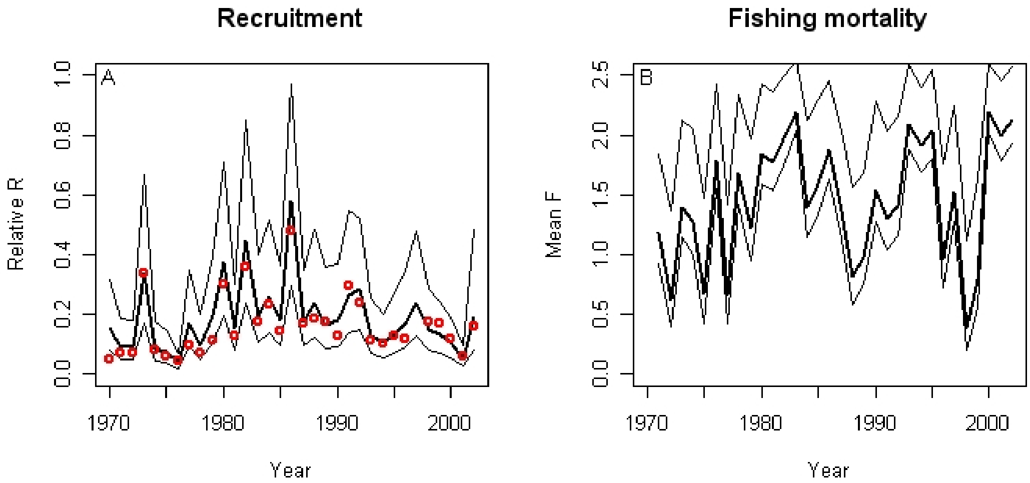


Figure 7. Trends in recruitment at age 1 (A) and fishing mortality (B) obtained through SNS data (age range 1-3). Circles represent observed number in the youngest age class (A), scaled to match model estimates. Annual fishing mortality is calculated as the arithmetic mean over the age classes two and three only. The heavy line is the median estimate, the area between the thin lines is the 95% credible interval.



Justification

This report

commissioned by:	Ministerie van Landbouw, Natuur en Voedsselkwaliteit Postbus 20401 2500 EK ‘s-Gravenhage
project number:	3.24.12470.02

has been produced with great care. The scientific quality has been peer-reviewed and assessed by or on behalf of the Scientific Board of Wageningen IMARES.

Drs. E. Jagtman
Head Fisheries Dept.

Signature: ____________________________

Date: 22 March 2007

Dr. A.D. Rijnsdorp
Scientific Board

Signature: ____________________________

Date: 22 March 2007

Number of copies:	20
Number of pages:	25
Number of tables:	2
Number of figures:	7
Number of annexes:	-