RIVO-Netherlands Institute for Fisheries Research

P.O. Box 68
NL 1970 AB Ymuiden
The Netherlands
Phone: +31 255 564646
Fax: +31 255 564644
E-mail: fisheriesresearch.asg@wur.nl
Internet: www.rivo.wageningen-ur.nl

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

RIVO report

Number: C047/05

Surveys used in the stock assessment of North Sea plaice and sole

O.A. van Keeken, R.E. Grift and A.D. Rijnsdorp

Commissioned by: Ministerie van Landbouw, Natuur en Voedselkwaliteit
T.a.v. de directeur Visserij
De heer drs. G. de Peuter
Postbus 20401
2500 EK DEN HAAG

Project number: 324-12470-02

Approved by: Dr. E. Jagtman
Head Department Biology and Ecology

Signature: ________________________

Date: 10 August 2005

Number of copies: 17
Number of pages: 55
Number of tables: 7
Number of figures: 5
Number of annexes: 3

Table of Contents

Samenvatting

Een van de drie werkpakketten van het F-project, werkpakket “F1”, richt zich op de verbetering van de toestandsbeoordeling. Dit rapport “A5” vormt onderdeel van dit werkpakket F1 en behandeld de surveys (bemonstering van het visbestand met een onderzoeksschip, dat elke keer met hetzelfde vistuig volgens dezelfde procedure vist), die gebruikt worden in de bestandsschatting van schol en tong uit de Noordzee. Het rapport is opgedeeld in drie delen. Het eerste deel van het rapport beschrijft de surveys en hoe de indices berekend worden. Het tweede deel onderzoekt de surveyindices en beschrijft de effecten van verschillende bemonsteringsschema’s op de bestandsschatting. Het laatste deel onderzoekt in hoeverre de surveys nog geschikt zijn als indices voor jonge schol, omdat de laatste jaren een verschuiving heeft plaats gevonden in de ruimtelijke verspreiding van schol, terwijl de opzet van de surveys niet veranderd zijn.

Voor de bestandsschatting van schol en tong uit de Noordzee worden gegevens gebruikt afkomstig van drie surveys: Beam Trawl Survey (BTS), Sole Net Survey (SNS) en Demersal Fish Survey (DFS). De surveys verschillen in het gebied dat ze bestrijken, het vistuig, verspreiding over de dieptes, stratificatie etc. De BTS survey is begonnen in 1985 en vindt plaats in het najaar. In 1996 is de survey uitgebreid zodat ook de diepere gedeeltes in de Noordzee bevist werden. Deze survey richt zich op schol en tong tot ongeveer 10 jaar oud. Per ICES rectangle worden 1-4 trekken gedaan. De SNS survey is begonnen in 1969 en richt zich op jonge schol en tong in ondiepere kustzones. Jaarlijks worden 10 gebieden bevist, waarbij gevist wordt langs een rechte lijn (transect). Tot 1989 werd zowel in het voorjaar als in het najaar gevist, daarna alleen in het najaar met uitzondering van 2003 (voorjaar). De DFS survey is begonnen in 1969 en richt zich op 0 en 1 jarige vis in de kustzones als de waddenzee. Jaarlijks worden een aantal gebieden bevist waarbij gestratificeerd wordt over verschillende dieptezones. Tot 1987 vond deze survey plaats in zowel het voorjaar als het najaar, daarna enkel in het najaar.

De werking van de surveys werd in 2002-2003 onderzocht in het kader van EU project EVARES. Dit project bestond uit twee delen. Het eerste deel onderzocht de samenhang tussen indices (een indices is een waarde voor het aantal gevangen vissen van een bepaalde leeftijd in een jaar in de desbetreffende survey), het tweede deel onderzocht het effect van veranderingen in bemonsteringsopzet op de bestandsschatting van schol en tong.
Het eerste deel onderzocht statistische relaties tussen de indices van dezelfde survey (volgen van de samenhang van jaarklassen in een survey), tussen de indices van verschillende surveys (samenhang van dezelfde jaarklasse in een jaar tussen de surveys) en tussen de indices en de bestandsschatting. De correlatie tussen jaarklassen in een survey en tussen verschillende surveys was over het algemeen het hoogst voor 2-4 jarige vissen. De BTS liet hogere correlaties zien met de bestandsschatting voor oudere leeftijden terwijl de SNS dit liet zien voor jonge leeftijden.
Het tweede deel onderzocht het effect van verschillende bemonsteringsopzetten van surveys op de bestandsschatting. Verschillende bemonsteringsopzetten zijn bijvoorbeeld een survey niet meer doen, een survey elke twee jaar uitvoeren in plaats van elk jaar, verandering in het aantal trekken etc. Het effect werd onder andere onderzocht op het paaibestand en de visserijsterfte. Verwijdering van een survey gaf de grootste verandering in de bestandsschatting. Een survey elke twee jaar uitvoeren in plaats van elk jaar gaf minder verandering, maar deze was wel groter dan veranderingen in de jaarlijkse bemonstering zoals verandering van het aantal trekken.

Zoals vermeld verschillende de surveys in stratificatie, locatie, het vistuig en de verdeling over diepteklassen. De afgelopen jaren heeft een verschuiving plaats gevonden in de ruimtelijke verspreiding van schol naar diepere gebieden. De surveys die gebruikt worden zijn echter niet veranderd en de verschuiving van schol kan effect hebben op de effectiviteit van de surveys voor gebruik in de bestandsschatting. De vangsten in de surveys werden via een model voorspeld aan de hand van onder andere diepte. Vanaf 1997 liet de DFS een scherpe daling

zien in de hoeveelheid 1 jarige schol, tegelijkertijd dat de verschuiving in schol plaats vond. De BTS en SNS lieten deze daling niet zien. De DFS kan de hoeveelheid van 1 jarige schol onderschatten. Een model dat onder andere lengte en breedte van een positie en de afstand tot de kust meeneemt geeft een betere indicatie maar dit was op dit moment niet mogelijk deze uit te voeren. In een vervolg studie zal een ruimtelijk model worden opgesteld die hiermee rekening houden kan.

Summary

This report A5 of the F-project deals with surveys and is divided into three sections:

1. Description of the survey sampling and raising procedures
2. Investigation of survey strategy and effect on stock assessment
3. Suitability of surveys for recruitment indices

In the stock assessment of North Sea plaice and sole three surveys are used: Beam Trawl Survey (BTS), Sole Net Survey (SNS) and the Demersal Fish Survey (DFS). Each survey varies in spatial distribution, gear types, locations and distribution over depth classes. The BTS survey started in 1985 and was expanded in 1996. The survey takes place in the autumn. This survey targets plaice and sole up to age ten in the offshore areas with stratification over ICES rectangles. The SNS started in 1969. Up to 1989 this survey took place in the spring and autumn, afterwards only in the autumn with the exception of 2003 (spring). This survey targets juvenile plaice and sole and is stratified over ten transects in the near shore areas. The DFS targets ages 0-1 plaice and sole in the inshore areas and is stratified over different depth ranges. Up to 1987 this survey took place in the spring and autumn, afterwards only in the autumn.

The effects of changes in survey sampling strategy on the stock assessment of North Sea plaice and sole were investigated within EU project EVARES during 2002-2003. For this project variations within and correlations within and between surveys were investigated. The BTS survey showed higher correlation with XSA for older age groups, while the SNS showed higher correlations for younger age groups. Also the effects of different sampling strategies, for example conducting a survey every two years, were investigated on the outputs of stock assessment. The entire removal of surveys gave the greatest changes on the stock assessment parameters like SSB and F. The removal of even or odd years led to greater changes in stock assessment parameters than changes in sampling intensity during a year.

The beam trawl surveys that are used in the stock assessment of North Sea plaice and sole vary in spatial distribution, gear types, locations and distribution over depth classes. Because the spatial distribution of plaice has changed significantly while the survey design and data analysis has remained unaltered, the suitability of these surveys for plaice recruitment were investigated. Since 1997, the index for age-1 plaice from the DFS sharply declined, concurrent with the change in spatial distribution of plaice, whereas the index from the SNS and BTS did not. The DFS could underestimate the abundance of age 1 plaice. In this stage of the project we could not quantify the effect on the index estimated from each survey and this should be done during the rest of the F-project.

Introduction

The F-project is a 5-year research project with the objective to improve the mutual understanding between fishermen, scientists and fisheries managers, by stimulating communication and collaboration between fishermen, fisheries scientists and fisheries managers. One of the three working packages of the F-project is concerned with the improvement of stock assessment of plaice and sole (F1). 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 analyzing 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

Product A5 within working package F1 investigates the representative nature of the input data and deals with biological parameters. This product investigates the survey indices used in the assessment of North Sea plaice and sole.

Surveys are hauls conducted by a research vessel during a certain period, covering a certain area to obtain fisheries independent estimates on the status of a fish stock. A survey is usually conducted annually during the same period of the year with a constant rigging of the gear. Because of the fixed mesh size used and area covered, surveys are size and therefore age selective. Some surveys target juvenile ages, while other surveys target a wider range of age classes. Most surveys estimate the abundance per age class. Some surveys, like acoustic or egg surveys, estimate biomass.

In the stock assessment of North Sea plaice and sole, survey data are used from three surveys: Beam Trawl Survey, Sole Net Survey and Demersal Fish Survey. For each survey catches per haul are combined for each age and year to form indices. These indices are used as a catch at age matrix in a tuning fleet file, which forms one of the input datasets for the assessment models. Indices are used in the assessment to scale the catch at age matrix.

This report is divided into three sections:

1) Description of the surveys and raising procedures
2) Investigation of survey strategy and effect on stock assessment
3) Suitability of surveys for recruitment indices

In chapter two the rigging of the gears and the areas covered by each surveys are described. The second part of this chapter describes the raising procedures used to calculate the tuning indices, which are used by the ICES demersal working group.

During 2002-2003 the effects of changes in survey sampling strategy on the stock assessment of North Sea plaice and sole were investigated within EU project EVARES (Evaluation of research surveys in relation to management advice, EVARES – FISH/2001/02 – LOT 1). A common framework of analyses was defined to quantify the correlations and variation in survey data and the influence on statistics of interest obtained from stock assessments (e.g. Spawning Stock Biomass, Status quo TACs) of certain changes in the research survey sampling strategy. Results from this study are presented in chapter 3.

In chapter 4 we investigated if the three beam trawl surveys still give representative estimates for plaice recruitment, since the spatial distribution of plaice has changed significantly while the survey design and data analysis have remained unaltered.

2. Description surveys

Three surveys are used in the stock assessment of North Sea plaice and sole. Indices from the Beam Trawl Survey (BTS) and the Sole Net Survey (SNS) are used for tuning the VPA, while the Demersal Fish Survey (DFS) is used for estimating recruitment. The first part of this chapter gives a description of the surveys, while the second part gives a description of the raising procedures.

Table 2.1. Survey characteristics.

	BTS		SNS	DFS		
	south-eastern North Sea	**central & western North Sea**		**coast**	**Wadden Sea**	**Scheldt estuary**
period	Aug-Sep	Aug-Sep	Sep-Oct	Sep-Oct	Sep-Oct	Sep-Oct
first year	1985	1996	1969	1969	1969	1969
vessel	Isis	Tridens	Isis	Isis	Stern	Schollevaar
haul duration	30 min	30 min	15 min	15 min	15 min	15 min
trawling speed	4 knots	4 knots	3.5-4 knots	2-3 knots	2-3 knots	2-3 knots
beam trawl size	8 m	8 m	6 m	6 m	3 m	3 m
mesh size net	120 mm	120 mm	80 mm	35 mm	35 mm	35 mm
mesh size cod-end	40 mm	40 mm	40 mm	20 mm	20 mm	20 mm
tickler chains	8	8	4	1	1	1
other traits		flip-up rope	'sole-net'			

2.1 Sampling

2.1.1. Beam Trawl Survey

The Beam Trawl Survey (BTS) is directed at obtaining fisheries-independent indices of abundance for North Sea plaice and sole up to age 10. The survey is conducted in the North Sea and waters around the United Kingdom in cooperation with institutes from England, Germany and Belgium, and is coordinated by the ICES Working Group on Beam Trawl Surveys (WGBEAM). The Dutch contribution to the survey was initiated in 1985 with research vessel “ISIS”, covering the southeastern part of the North Sea (Figure 2.1, left). From 1995 the survey is also conducted with research vessel “TRIDENS”, covering the southwestern, western and central part of the North Sea (Figure 2.1, right). The BTS is conducted each year in August/September using an 8 m beam trawl with 40 mm mesh cod-end. The haul duration is 30 minutes. The research vessel “ISIS” executes 1-4 pseudo-random hauls per ICES rectangle, while “TRIDENS” only executes 1 haul per rectangle.

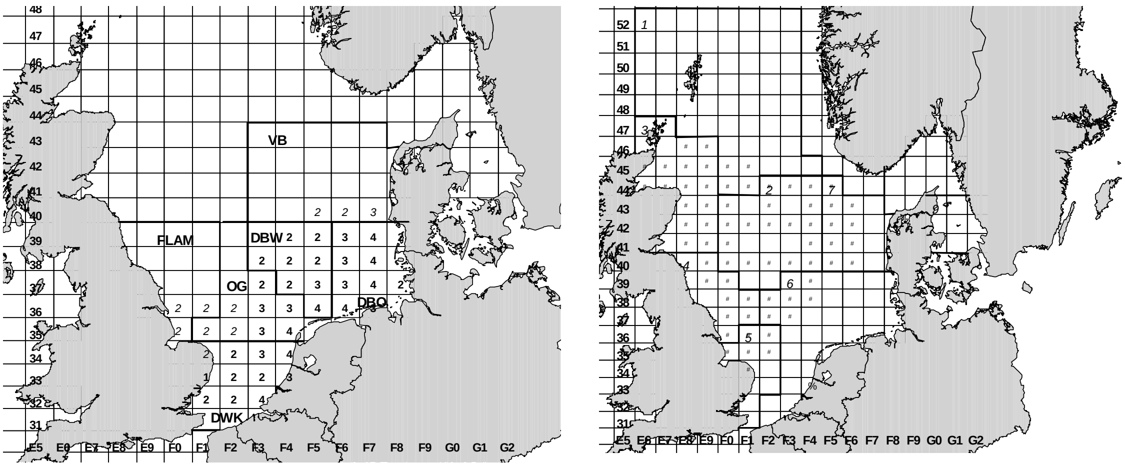


Figure 2.1. BTS. Sampling coverage of research vessel ISIS (left) and Tridens (right).

2.1.2. Sole Net Survey

The Sole Net Survey (SNS) is directed at obtaining fisheries-independent indices of abundance for age 1-4 North Sea plaice and sole. The survey was initiated in 1969 and is based on 10 fixed transects (parallel or perpendicular to the coastline) along the Dutch, German and Danish coast (Figure 2.2). The position of stations within a transect is chosen such that the entire depth-range of the transect is covered. During 1969-1989 the survey was conducted in spring and in autumn. Since 1990, the SNS is carried out once a year in September – October, with the exception of 2003 when the survey was carried out in spring instead of autumn. The SNS is using a 6 m beam trawl with 40 mm mesh cod end. The haul duration is 15 minutes.

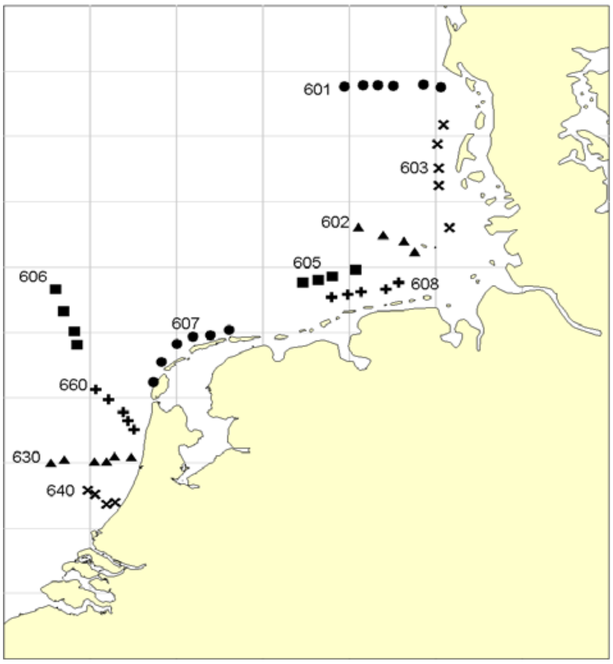


Figure 2.2. SNS. Sampling coverage of the 10 transects.

2.1.3. Demersal Fish Survey

The Demersal fish Survey (DFS) is directed at monitoring juvenile plaice and sole, shrimps and non-commercial demersal fish species. The survey was initiated in 1969 and covers the Wadden Sea, the Westerschelde, the Oosterschelde and the coastal zone (Figure 2.3). During 1969-1986 the survey was conducted in spring and autumn, but from 1987 onwards the survey is only conducted in the autumn. A 6 m trawl is used by research vessel “ISIS” in the coastal zone, while a 3 m trawl is used by research vessel “Stern” in the Wadden Sea and by “Schollevaar” in the Westerschelde and Oosterschelde. All trawls use a 20 mm mesh cod end, and the haul duration is 15 minutes. Other “young fish surveys” are executed by institutes from England, Germany and Belgium, and are combined to form international indices.

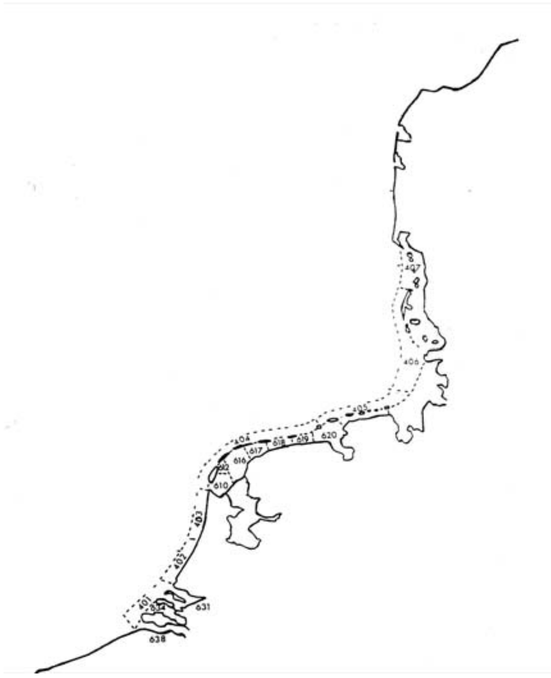


Figure 2.3. DFS. Sampling coverage along the coast.

2.2 Raising indices

For each of the surveys, the numbers at length and age by haul are derived from the numbers at length by applying age-length keys (proportion of each age within a length-class) to the data. These age-length keys differ between areas within each survey and will be described for each survey in the paragraphs below.

2.2.1. Beam Trawl Survey

For the BTS seven different age-length key areas are selected (Figure 2.4). Each of these areas comprised of a number of ICES rectangles. For both the ISIS and Tridens, separate age-length keys are used for the different areas. The numbers at length are summed for each age to obtain the numbers at age by haul. From the numbers at age by haul, the mean number at age by ICES rectangle is calculated. The indices are calculated as the mean number at age over all ICES rectangles.

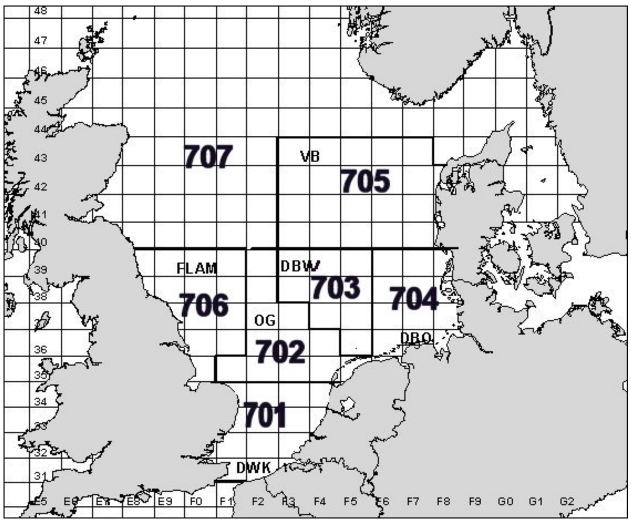


Figure 2.4. BTS. Age-length key areas based on combinations of ICES rectangles.

2.2.2. Sole Net Survey

The 10 transects presented in Figure 2.2 are selected as age-length keys areas for the SNS survey. For each haul the mean number at age is calculated. From the mean number at age per haul the mean number at age per transect is calculated. The indices are calculated as the mean number at age over all transects, expressed as numbers at age per 100 hour of fishing.

2.2.3. Demersal Fish Survey

The different sampling areas for the DFS presented in Figure 2.3 are combined to 7 age-length keys areas (Figure 2.5). The index calculation for the DFS is area and depth stratified (Table 2.2). For each haul the mean number at age is calculated. Each area is divided in depth zones and for each depth zone the mean number at age per hectare is calculated. The mean number at age per depth zone is multiplied with the surface area of this depth zone within the area. The sum of all numbers at age corrected for the surface area of a depth zone within an area are summed to obtain an index for each area. The indices used eventually are the sum of all area indices.

Table 2.2. Depth zones used in the DFS index calculation.

Depth zone	Wadden Sea	Other areas
1	0-6	0-5
2	6-12	5-10
3	12-20	10-20
4	>20	>20

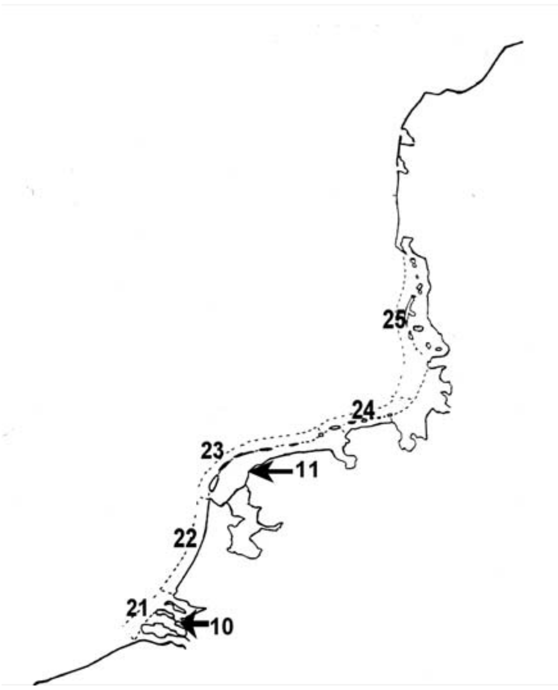


Figure 2.5. DFS. Age-length key areas based on combinations of different sampling areas

3. Sampling strategy

3.1 Introduction

During 2002-2003 the effects of changes in survey sampling strategy on the stock assessment of North Sea plaice and sole were investigated within EU project EVARES (Evaluation of research surveys in relation to management advice, EVARES – FISH/2001/02 – LOT 1). A common framework of analyses was defined to quantify the correlations and variation in survey data and the influence on statistics of interest obtained from stock assessments (e.g. Spawning Stock Biomass, Status quo TACs) of certain changes in the research survey sampling strategy.

3.2 Methods

Two different analyses were conducted on the survey data:
1) correlation and variation studies and
2) simulation of the variability of survey results and the effects of that variability on the outputs of stock assessment.

3.2.1 Correlation and variation

Three different kinds of correlations were investigated:
- Within survey consistency

Log transformed survey indices at age were correlated with log transformed indices from the same survey for the same year class one year later.

- Between survey consistency

Correlations between abundance indices from two surveys for the same age in the same year were investigated.

- Consistency between survey indices and stock abundance

For the consistency between survey indices and stock abundance, the correlation between the survey indices and the abundances derived from the numbers at age obtained by the ICES assessment were investigated.

3.2.2 Effects on stock assessment

The precision or variance of a survey estimate of numbers at age for use in an assessment may be estimated by applying the bootstrapping technique. The assumptions behind a bootstrap are that the samples may be considered as equivalent to a simple random sample from the population. Bootstrapping is carried out for either single stage or multiple stages by resampling the same number of observations from the original samples with replacement. Each trawl is used to obtain an estimate of numbers at age, estimates at age are taken as a set within the bootstrap, to ensure that correlation at age that occurs in the sampling is represented in the realizations used for the assessments.

The effects of the variability of survey indices on the outputs of stock assessments were investigated, using altered tuning indices in the stock assessment model. Different sampling strategies were simulated, constructing if possible 500 tuning indices datasets for each scenario. To create these new tuning indices, catch-at-age numbers from the BTS and SNS were resampled by applying the bootstrapping technique to the numbers-at-age per stratum, which was an ICES rectangle for BTS or a transect for SNS. From the resampled numbers-at-age per haul the mean number-at-age in each stratum was calculated. For BTS, the resampled indices per age group were obtained by calculating the mean numbers-at-age over all of the ICES rectangles in the index area. For SNS, the resampled indices per age group were obtained by calculating the mean numbers-at-age (expressed as numbers per 100h fishing) over all transects.

Different sampling strategies were explored:

- A complete stop of a survey. A survey (either BTS or SNS) was stopped over a period of 10 years (1991-2000). Data before this period were left unchanged.
- Conducting a survey every second year. A survey was conducted only once every two years (even or odd years) during the last 10 years. This was done for a single survey (BTS and SNS) as well as for both surveys simultaneously.
- Bootstrapping both surveys, keeping the same sampling intensity. For both surveys the same sampling intensity was kept within an ICES rectangle (BTS) or transect (SNS).
- Changing sampling intensity on one survey. Three different sampling intensity scenarios were considered:
    - reducing intensity by half
    - keeping the same intensity
    - doubling the sampling intensity within an ICES rectangle or transect.
- One survey and the commercial fleet taken out
- Both surveys stopped over the last 10 years.
- Sampling half of the rectangles or transects. For BTS half of the ICES rectangles were selected randomly, for SNS half of the transects. Odd numbers of rectangles/transects were randomly rounded up or down. Catch numbers per haul were kept unchanged.
- Reducing sampling intensity for one survey by half and doubling the sampling intensity on the other survey.

In order to evaluate the performance of surveys in the assessments, a number of key parameters were selected to provide a way of summarising the performance. Three concepts were considered:

- The current state of the stock evaluated by spawning stock biomass and mean F for adults in the fishery in the terminal year (SSB0 and Fbar0)
- Recent change in the stock evaluated by the change in SSB and mean F from 5 years before to the terminal assessment year (SSB0-4 and Fbar0-4)
- Projections of TAC and SSB to the TAC year with two different management options. In both cases the fishing mortality in the next year was set as F=Fstatus quo. Then for the TAC year F is set either as F=Fstatus quo or F=Fpa. (SSBFSQ, SSBFPA, TACFSQ and TACFPA)

These eight parameters, their sensitivity to changes in the surveys sampling strategy and, where bootstrap input data have been used, their spread are presented as a measure of the influence of the surveys in the assessments. All options used within the ICES Working Group have been kept in order to facilitate comparisons between the simulation outputs and results from the available assessments. The period within which changes in a survey corresponds to the last ten years.

3.3 Results

3.3.1. Correlation and variation

Correlations between datasets are expressed as correlation coefficient, which is an indication for how well two datasets agree: 0 means two datasets do not agree, 1 means they fully agree. For North Sea plaice in the BTS, correlation coefficients were highest for ages 2-3 and ages 3-4 (Table 3.1, Figure I.I in Appendix I). The SNS showed similar correlation coefficients, except for ages 5-6 (Table 3.1 Figure I.II). Correlation coefficients were the highest for ages 2-3 and ages 3-4.

Table 3.1. Plaice. Within-survey consistency, expressed as correlation coefficients (r).

Ages	BTS	SNS
0-1	0.40	0.39
1-2	0.55	0.58
2-3	0.78	0.79
3-4	0.70	0.80
4-5	0.68	0.71
5-6	0.47	0.70

Correlations between BTS and SNS were highest for ages 2 and 3 plaice (Table 3.2, Figure I.III) Correlations between survey indices and estimated XSA population numbers showed that the BTS (Figure I.IV) appears to provide a better tuning index for the older ages (3-6) and SNS (Figure I.V) for the youngest age (age 1) whereas both surveys perform about equal for age 2 (Table 3.2).

Table 3.2. Plaice. Between-survey consistency (BTS / SNS) and consistency between survey indices and stock abundance (survey / Assessment), expressed as correlation coefficients (r).

Ages	BTS / SNS	BTS / Assessment	SNS / Assessment
0	0.77	0.49	0.77
1	0.57	0.59	0.61
2	0.89	0.81	0.55
3	0.85	0.84	0.65
4	0.60	0.59	0.42
5	0.79	0.65	0.38
6	0.50		

For North Sea sole, low correlations were observed between ages 0-1 for both BTS and SNS (Figure 3.3, Figure I.VI, I.VII). Higher correlations were observed for sole up to age 5 for both SNS and BTS. For age 4 sole caught with SNS, CPUE data were only available for 2 years.

Table 3.3. Sole. Within-survey consistency expressed as correlation coefficients (r).

Ages	BTS	SNS
0-1	0.20	0.21
1-2	0.66	0.84
2-3	0.68	0.74
3-4	0.69	1
4-5	0.67	
5-6	0.37	

Correlation coefficients between BTS and SNS were high except for age 0 sole (Table 3.4, Figure I.VIII). Correlations between survey indices and estimated XSA population numbers showed that BTS (Figure I.IX) performs best for the older age-groups (ages 3-6) whereas SNS (Figure I.X) performs best for the youngest age-groups (ages 1-2) (Table 3.4)

Table 3.4. Sole. Between-survey consistency (BTS / SNS) and consistency between survey indices and stock abundance (survey / Assessment), expressed as correlation coefficients (r).

Ages	BTS / SNS	BTS / Assessment	SNS / Assessment
0	0.36	0.87	0.94
1	0.79	0.80	0.85
2	0.70	0.84	0.78
3	0.83	0.72	
4		0.76	
		0.73	

3.3.2. Effects on stock assessment

3.2.2.1. North Sea plaice

Survey taken out

The largest change in SSB in the year 2000 (Figure II.1) was created by taking out BTS. SSB ratio 2000/1996 (Figure II.2), F ratio (Figure II.3) and recruitment ratio (Figure II.4) were also affected most by taking out BTS. The lower SSB and recruitment in 2000 caused by the removal of BTS resulted in markedly lower predictions in 2002 of SSB (Figure II.5) and TAC (Figure II.6).

Conducting survey every second year

Differences were observed when odd or even years were taken out for BTS and SNS separately or BTS and SNS combined. Because of selecting only every second year of survey data, signals from a strong year class disappears for some years from the tuning dataset, altering the tuning. Tuning weights also change with changing data influencing the estimations. The SSB in 2000 was higher for even years taken out than for odd years taken out for all 3 scenarios (Figure II.1). For the F ratio SNS was more in agreement with the basic set when even years were taken out, while BTS, and BTS and SNS combined where more in agreement when odd years were taken out (Figure II.3). Large differences were observed in recruitment ratio between odd and even years taken out for SNS and for BTS and SNS combined (Figure II.4). In almost all cases the three scenarios gave a lower estimation of the predicted SSB and TAC in 2002 compared with the basic set (Figure II.5 and II.6).

Changing sampling intensity

Reducing the sampling intensity by half gave a lower SSB in 2000 compared to keeping the same intensity, but was in general more in agreement with the basic set then conducting a survey every second year. Doubling the sampling intensity, however, did not differ much from keeping the same sampling intensity. SNS showed higher fluctuations when changing the sampling intensity than BTS. For SNS the indices are calculated over transects, which each cover a considerable depth range. Indices for BTS are calculated over ICES rectangles, which in general cover a smaller depth range. Juvenile plaice prefer the shallow zone while adults prefer deeper water. Resampled catch per haul data for SNS might therefore differ more than for BTS, resulting in higher fluctuations of the various diagnostics.

One survey and commercial fleet taken out

SSB in 2000 (Figure II.1), SSB ratio (Figure II.2) and recruitment ratio (Figure II.4) were higher than the basic data set with both SNS and commercial fleet taken out and lower with both BTS and commercial fleet taken out. F ratio was much higher for both scenarios compared with the basic data set (Figure II.3). Predicted SSB (Figure II.5) and TAC (Figure II.6) for 2002 differed for both scenarios up to around 25% with the basic data set.

Both surveys taken out

With both surveys taken out, large differences in SSB in 2000 were observed with the basic set (Figure II.1), SSB ratio (Figure II.2) and recruitment ratio (Figure II.4). The difference in F ratio was larger with one survey and both the commercial fleet taken out than with both surveys taken out (Figure II.3).

Sampling half of the rectangles or transects

Taking out 50% of the rectangles for BTS gave a higher SSB in 2000 (Figure II.1) and a lower F ratio (Figure II.3) compared to taking out 50 % of the transects for SNS. In most cases, taking out 50% of the rectangles or transects gave a lower prediction of the predicted TAC for 2002 (Figure II.5 and II.6).

Reducing sampling intensity for one survey by half and doubling the sampling intensity on the other survey

Reducing sampling intensity by half for one survey and doubling sampling intensity on the other did not show large differences between both scenarios. Both scenarios showed lower SSB in 2000 (Figure II.1) compared to the basic set, which indicated that the reduction in sampling intensity could not be compensated by a doubling of the intensity of the other survey.

3.2.2.2. North Sea sole

Survey taken out

Taking out BTS gave a higher SSB in 2000 compared with the basic set while taking SNS out gave a lower value. Removal of the SNS gave slightly larger differences in F ratio and recruitment ratio than the removal of BTS (Figures II.9 and II10). Taking out either BTS or SNS gave a higher estimation of the predicted SSB (5-9%) and TAC (3-7%) for 2002 (Figure II.11 and II.12).

Conducting survey every second year

The removal of odd and even years of BTS separately showed an inverse effect on estimates of SSB, SSB ratio, F ratio and recruitment ratio compared with the removal of SNS separately. The removal of both surveys followed the pattern observed for the removal of BTS only. Only when the even years of both surveys were taken a markedly higher prediction was derived for both SSB and TAC in 2002 (Figures II.11 and II.12).

Changing sampling intensity

Changing sampling intensity for both BTS and SNS influenced the fluctuation of values of the different diagnostics, but no major differences were observed between the basic set and the data sets.

One survey and commercial fleet taken out

Taking out both commercial fleet and one survey gave a lower SSB in 2000 (Figure II.7) and SSB ratio (Figure II.8), a higher F ratio (Figure II.9) and a lower recruitment ratio (Figure II.10) compared with the basic set. For both scenarios predicted SSB (Figure II.11) and TACFpa (Figure II.12) in 2002 were much lower than the basic data set.

Both surveys taken out

Taking out both surveys resulted in a SSB in 2000 (Figure II.7) and F ratio (Figure II.9) almost similar to the basic set. The recruitment ratio was high compared to the basic set (Figure II.10). Consequently predicted SSB and TAC in 2002 (Figure II.11 and II.12) were much higher than for the basic data set.

Sampling half of the rectangles or transects

Sampling half of the rectangles or transects did not show large differences with the basic data. Small differences with the basic set were only observed for the predicted TAC (Figure II.12).

Reducing sampling intensity for one survey by half and doubling the sampling intensity on the other survey

Reducing sampling intensity by half for one survey and doubling sampling intensity for the other survey did not show large differences with the basic data set for both scenarios. The predicted SSB and TAC were slightly higher then the basic set (Figure II.11 and II.12).

3.4 Synthesis

A review of the corresponding correlation coefficients makes it possible to assess the consistency between surveys for each age. This can be of specific importance when abundance in research surveys and commercial C.P.U.E. show different trends. In such a case it is importance to make sure that the assessments will not be strongly influenced by commercial C.P.U.E indices that are not proportional to stock abundance. On the other hand the identification of a strong correlation pattern between different surveys could pave the way for tuning techniques that would recognize them.

There are limits to the interpretation of the correlation coefficients between survey indices and stock abundance. Good correlations for a specific survey can be due to the fact that the assessment has been “attracted” by that tuning series. Conversely, poor correlations can be due to attraction to other series, e.g. commercial fleets. The weights assigned to all tuning fleets during the assessment process will also affect correlations. The correlations nevertheless provide a valuable summary of the consistency of the various survey indices and stock abundances, which is especially useful when they are uses in conjunction with other results. If the correlation is poor while the corresponding sampling variance is low it is likely for instance that a moderate change in the sampling intensity will not strongly affect the final assessments results. On the other hand if correlations are poor for a set of research surveys, which on the other hand perform well in terms of within and between survey consistency, this suggests that the influence of other inputs to the assessment, first of all those coming from commercial fleets, would deserve further analysis.

Bootstrapping made it possible to simulate time series associated with various levels of sampling error for a survey (or the case being –see below refined scenarios), while taking into account the corresponding correlations. However the corresponding outputs should not be interpreted as predicting what would result from a real change in the number of hauls, because, among other reasons, the survey design would not stay unchanged. Basic simulations have anyway been associated with a/ the existing level of sampling errors, b/ doubling it and c/ halving it. It also appeared necessary to specify which baseline should be considered for comparisons. Such a baseline can correspond to what was obtained with the available complete data sets. This has the merit of simplicity. But it does not take into account the fact that the existing assessments are affected by (among others) the sampling variability. This is why it can also be useful, and in some cases more legitimate to refer the set of values that could have been obtained due to sampling errors in a set of surveys, conducted in each case with the present intensity. Whenever possible the two possible references have been provided.

4. Suitability of surveys for recruitment indices

4.1 Introduction

As described in Chapter 2, the three beam trawl surveys that are used in the stock assessments have different characteristics. They vary in spatial distribution, gear types, locations and distribution over depth classes (Figure III.1). Also, the procedure to estimate recruitment indices varies (Table 4.1).

Table 4.1. Overview of estimating recruitment indices from the three beam trawl surveys.

Survey	Age length key	Stratification	Starting year
BTS	Area	ICES rectangle	1985
SNS	Transect	Transect	1969
DFS	Area	Depth class	1969

Because the spatial distribution of plaice has changed significantly while the survey design and data analysis have remained unaltered, we investigated if the three beam trawl surveys still give representative estimates for plaice recruitment. All surveys have been executed since long time and during this period the sampling design and raising procedures have remained unchanged. Over this period, however, the spatial distribution of plaice has changed significantly (Grift et al., 2004; Van Keeken et al., 2004). Especially for under-sized plaice, for which the indices are most important in the stock assessments, the spatial distribution has changed (Figure III.2).

4.2 Methods

Changes in the spatial distribution of sole and plaice populations were investigated to assess if the three surveys still give representative indices for the abundance of sole and plaice. As described, the three surveys have different sampling designs and cover different areas. The DFS, for example, is limited to shallow depths and does not cover the entire distribution area of the North Sea plaice population. The BTS on the other hand, only covers deeper areas. The three surveys combined cover the entire distribution area of all age classes of plaice and sole. Therefore, we used all three of them to quantify changes in the spatial distribution of plaice.

In the current analysis, we use depth as a proxy for space because in previous studies we observed that the spatial distribution of plaice mainly changed along a depth gradient from onshore to offshore. Obviously, a more sophisticated analysis including e.g. latitude, longitude and distance offshore would be preferable but was not feasible in this phase of the project. Therefore, we chose to explore first possible changes in the spatial distribution along depth gradients and elaborate on these analyses in later stages of the project.

Possible changes in the distribution along depth gradients were quantified by using a statistical model that predicts the number of plaice and sole per age group. In this model, effects of water depth, survey, cohort, year and month were taken into account:

$$\mathrm{Log(CpUE)} \sim D + D^2 + C_i + S_j + M_k + (D \times Y) + (D \times M_k) + (D^2 \times C_i) + (D \times S_j)$$

where CpUE is the catch per unit of effort (Log (numbers per hectare+1)), *D* is depth (meters), *C* is cohort (year class), *S* is survey (DFS, BTS, SNS), *M* is month of the year, and *Y* is year.

The interaction between depth and year (D x Y) is the most important term in this analysis because it allows the effect of depth to change over years. If this term is significant, we can conclude that the distribution of fish over depths has changed in time. The model yields parameters with which densities of fish at certain depths in specific years can be estimated. These parameters can be used to illustrate the changed distribution of fish over depths.

Although the SNS and DFS have been executed since 1969, the BTS has been executed since 1985. Therefore, we only used data from 1985 on.

4.3 Results

The model performed well, especially for the age groups 0-5 of plaice and 0-4 of sole a large proportion of the variance was explained (R^2>30%).

For plaice of ages 0-6, there has been a highly significant (*P*<0.01) change in the abundance of fish at different depths (significance of D x Y). For ages 7-10, there was no significant change. For sole, the change in abundance at different depths was highly significant (*P*<0.01) for ages 0, 2 and 7 and significant for ages 1 and 10 (Figure III.2).

The change in the abundance of plaice and sole is clearly visible when the model parameters are used to describe the abundance at different depths at different years (Figures III.4-III.10). Figure III.3 shows the abundance of age 1 plaice at different depths in 1985 and 2002 (2002 is the last year in the analysis in which all surveys were executed in fall). The shift from age 1 plaice to deeper areas is clear; the average depth has increased from 21 to 32 m (average of predicted values, over all surveys).

A comparison of the model results with the indices that are estimated from these surveys shows that the trend in the DFS index rather faithfully mirrors the trend in the depth x cohort parameter (Figure III.11). This parameter indicates the distribution of plaice along the depth gradient. Obviously, the catches in the DFS strongly influence this parameter because the DFS has many hauls and also the shallowest depths. The trends also show discrepancies between indices from the different surveys, especially since 1997. The index for age-1 plaice from the DFS sharply declined since 1997 whereas those from the SNS and BTS did not. It was especially during those years that age-1 plaice moved to deeper areas.

4.4 Synthesis

The significant change in the spatial distribution of several age classes of plaice and sole could influence the accuracy of the indices retrieved from the three beam trawl surveys. In this stage of the project we could not quantify the effect on the index estimated from each survey and this should be done during the rest of the F-project. We think, however, that the consequence of this finding could be that the DFS, for example, could have become less representative for age-1 plaice. In 1985, the distribution of DFS hauls over depth covered the range in which the abundance of age-1 plaice peaked (Figure III.4). In 2002, however, this peak laid outside the range of the DFS programme and consequently, the DFS could underestimate abundance of age 1 plaice.

In subsequent analyses, the change in spatial distribution should be described more accurately by including latitude, longitude and distance offshore. This will be done in study A5.1, which will look in closer detail to the models to be used. The results of these more accurate models should be used to quantify the effects of a change in spatial distribution on the indices and hence stock assessments.

References

Grift, R. E., I. Tulp, L. Clarke, U. Damm, A. McLay, S. Reeves, J. Vigneau and W. Weber. 2004. Assessment of the ecological effects of the Plaice Box. Report of the European Commission Expert Working Group to evaluate the Shetland and Plaice boxes. Brussels: 121 p.

Van Keeken, O. A., M. Van Hoppe, R. E. Grift and A. D. Rijnsdorp. 2004. The effect of changes in the spatial distribution of juvenile plaice (*Pleuronectes platessa*) in the North Sea on the management of its stocks. ICES C.M. 2004/K:25.

Appendix I

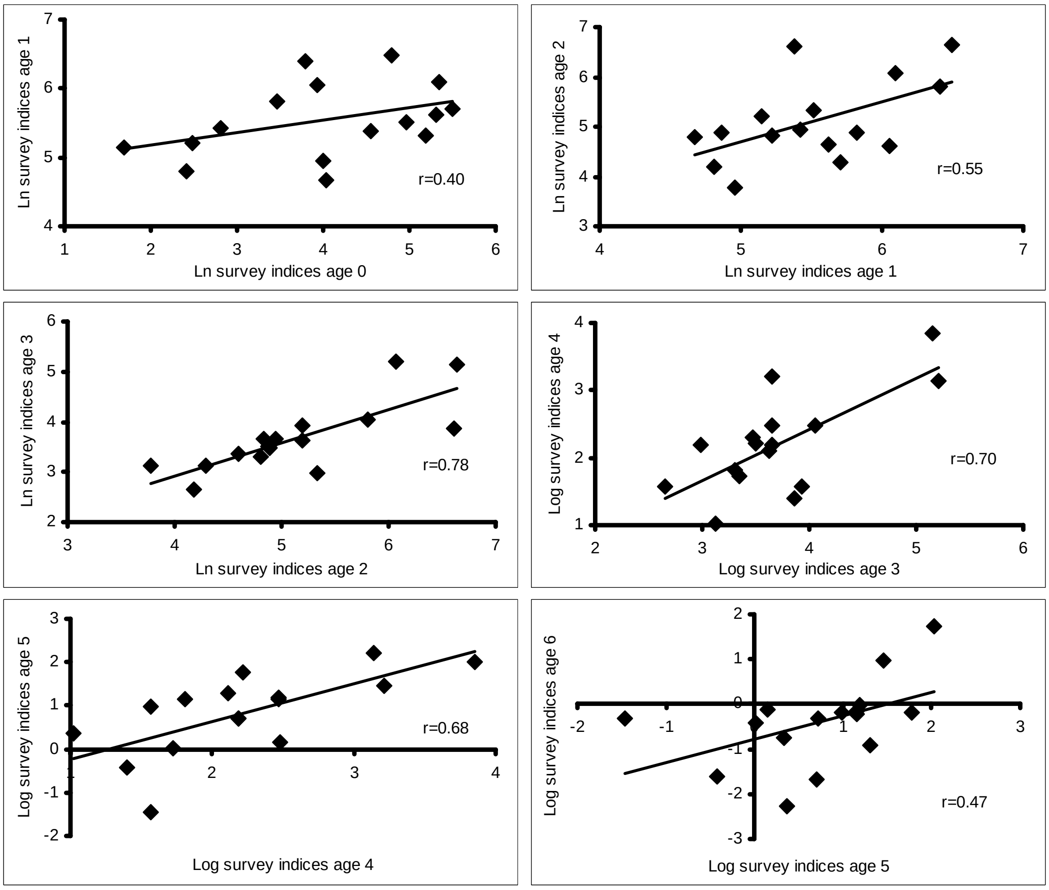


Figure I.1. Plaice. Correlations within logarithm of survey indices BTS for ages 1-6.

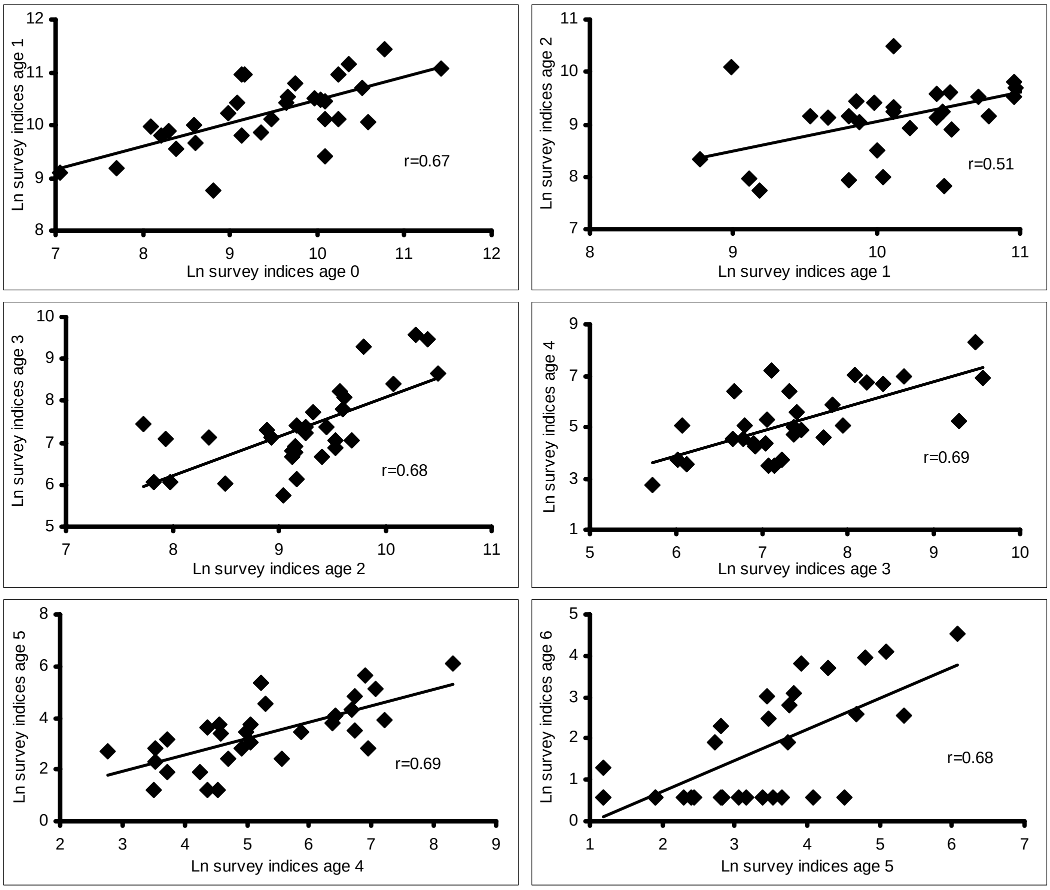


Figure I.2. Plaice. Correlations within logarithm of survey indices SNS for ages 1-6

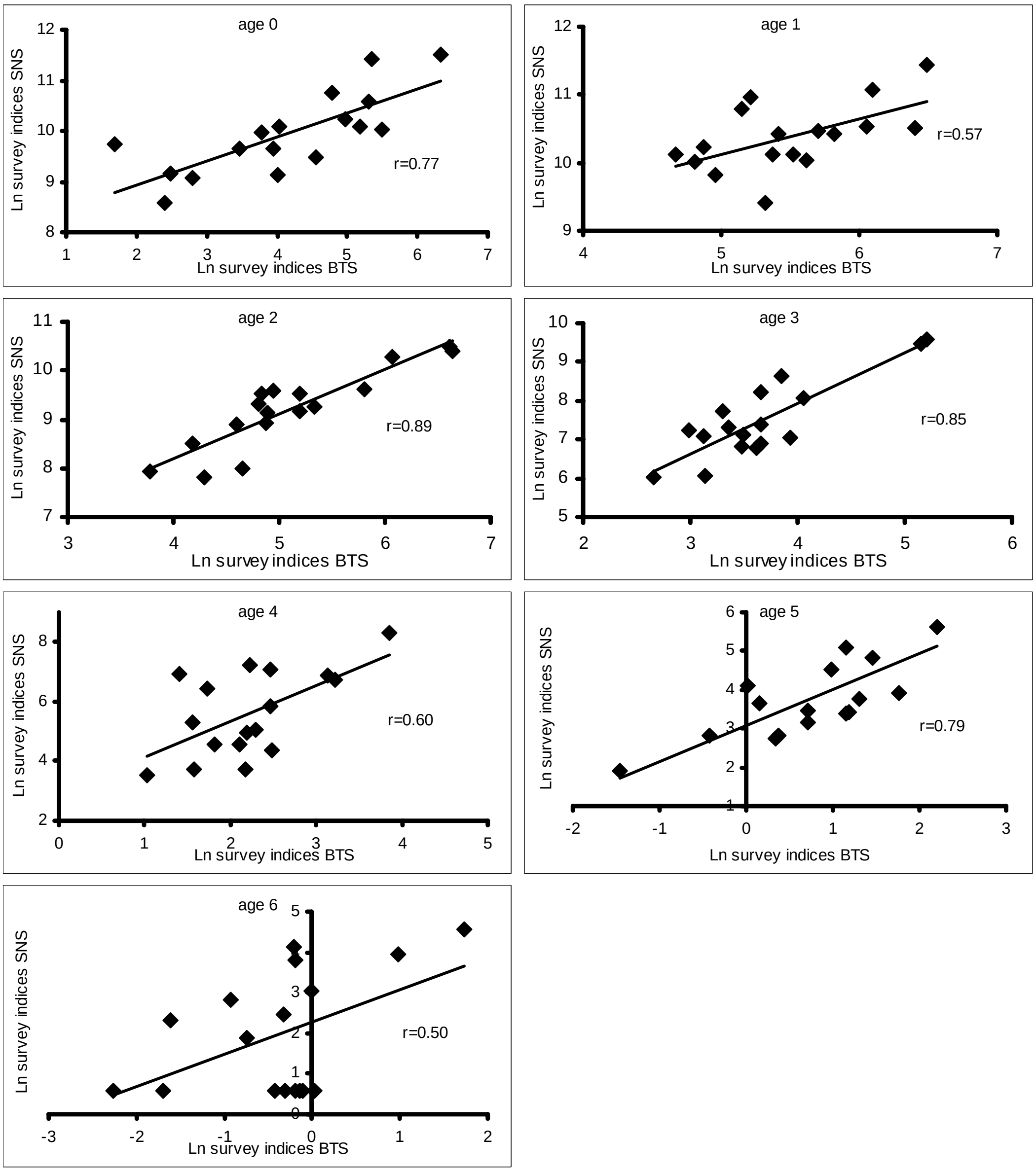


Figure I.3. Plaice. Correlations between logarithm of survey indices BTS and logarithm of survey indices SNS for ages 0-6.

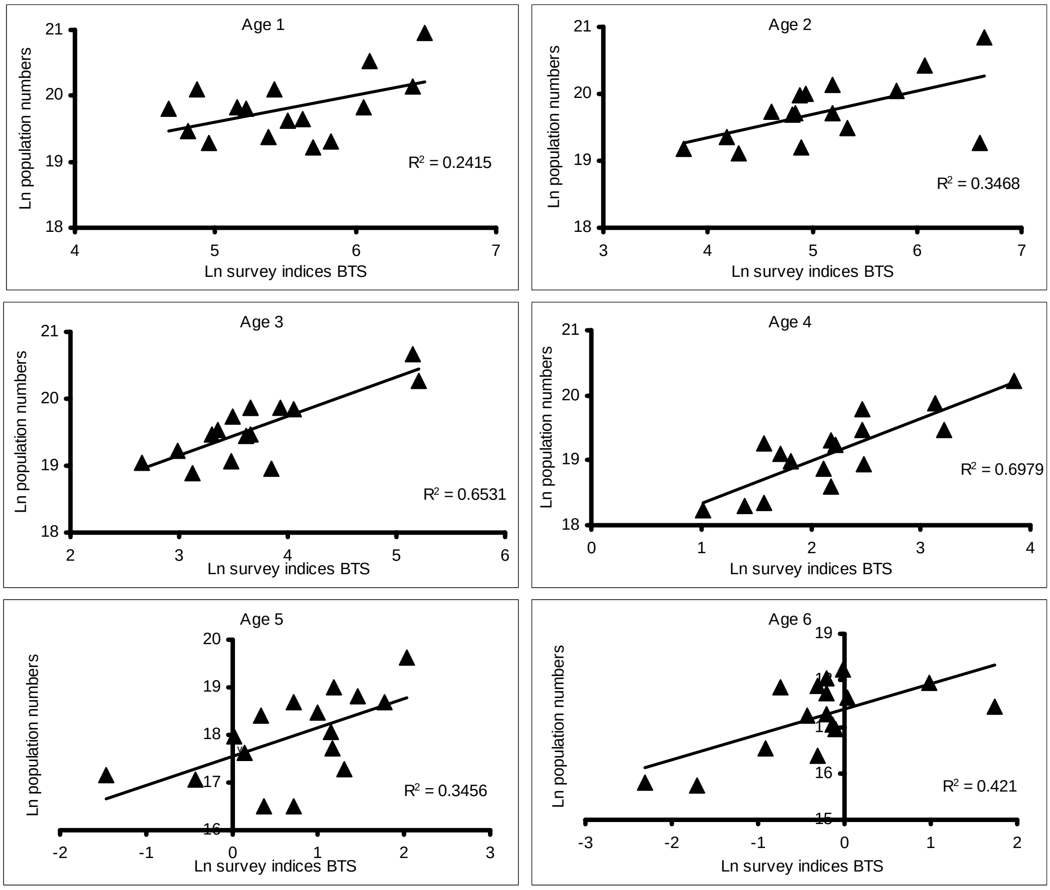


Figure I.4. Plaice. Correlations between logarithm of survey indices BTS and logarithm of estimated population numbers for ages 1-6.

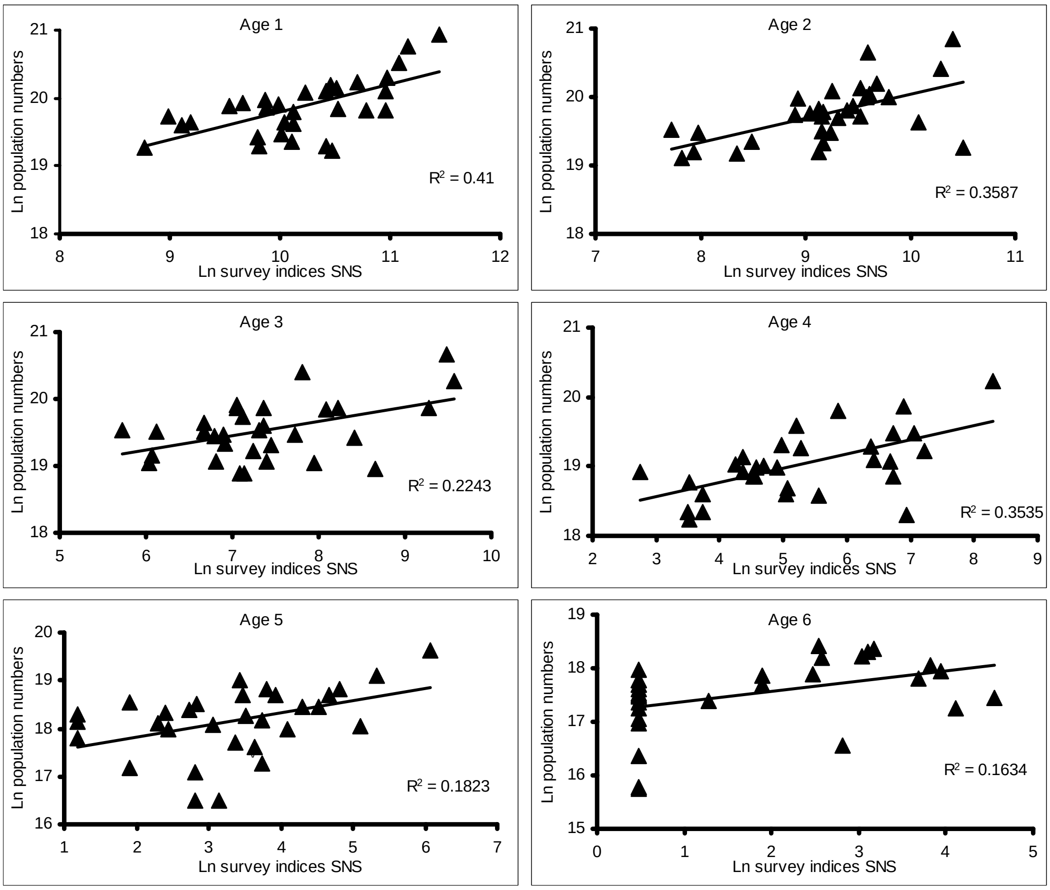


Figure I.5. Plaice. Correlations between logarithm of survey indices SNS and logarithm of estimated population numbers for ages 1-6.

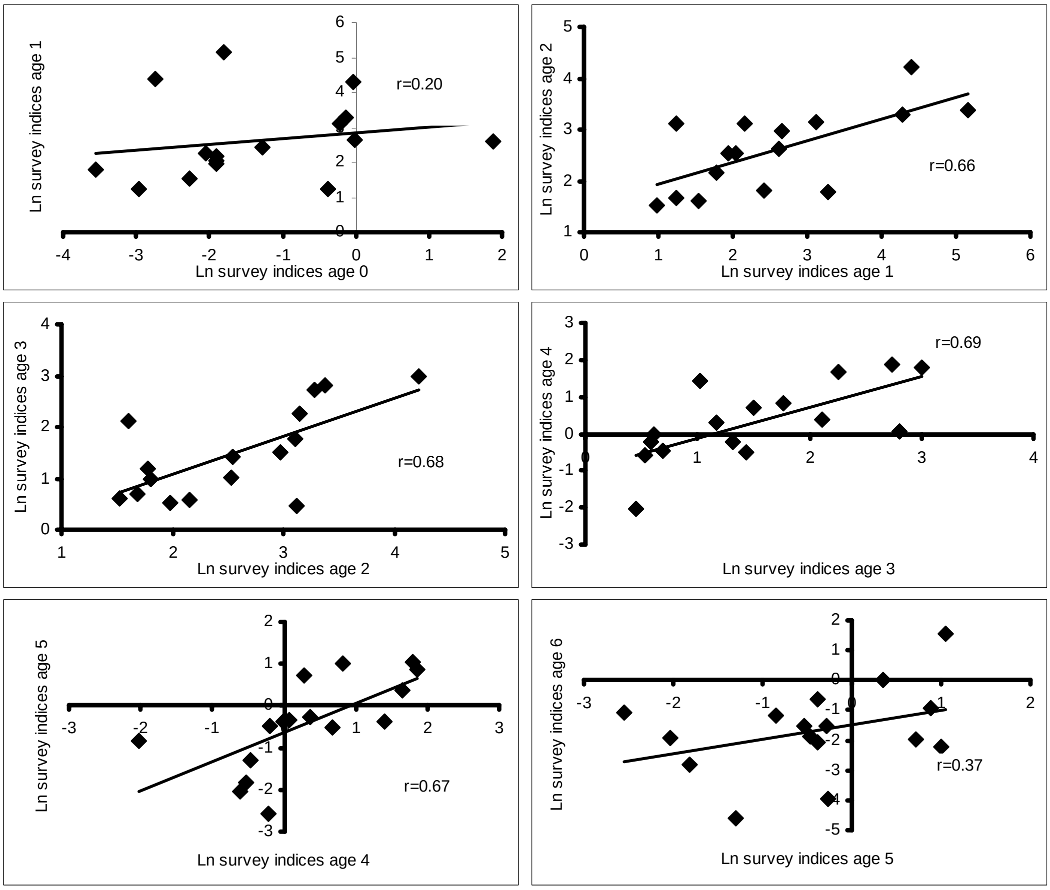


Figure I.6. Sole. Correlations within logarithm of survey indices BTS for ages 1-6.

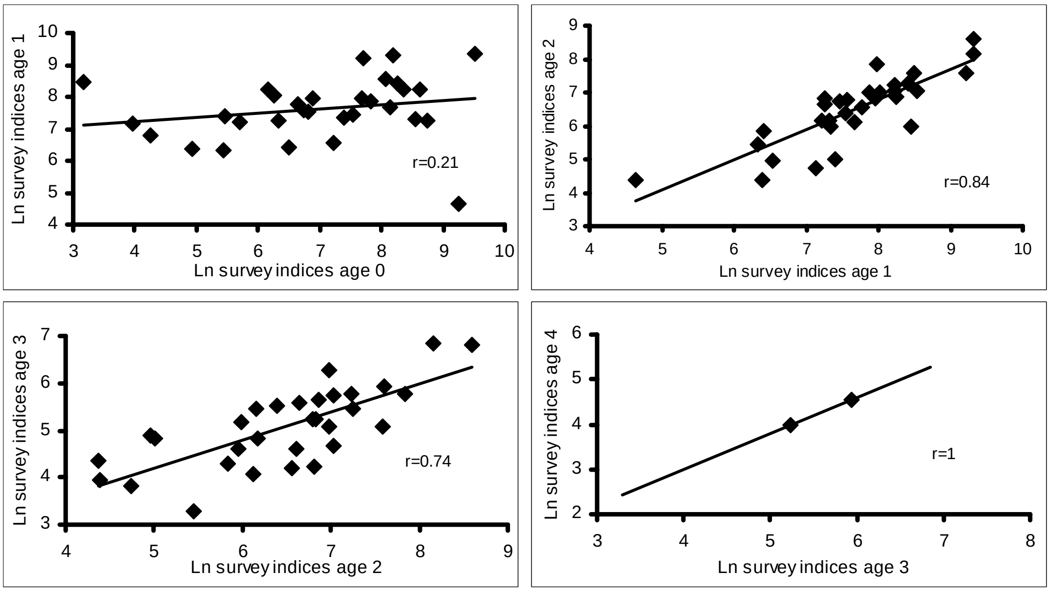


Figure I.7. Sole. Correlations within logarithm of survey indices SNS for ages 1-4.

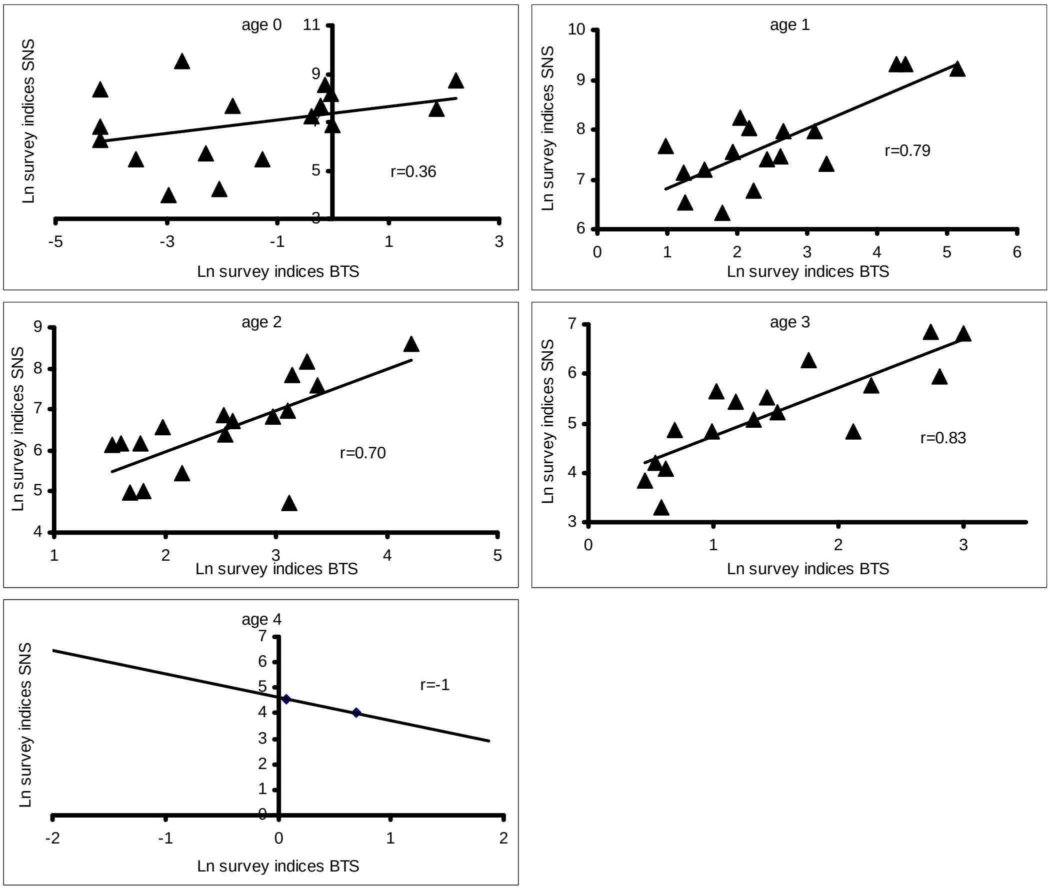


Figure I.8. Sole. Correlations between logarithm of survey indices BTS and logarithm of survey indices SNS for ages 0-4.

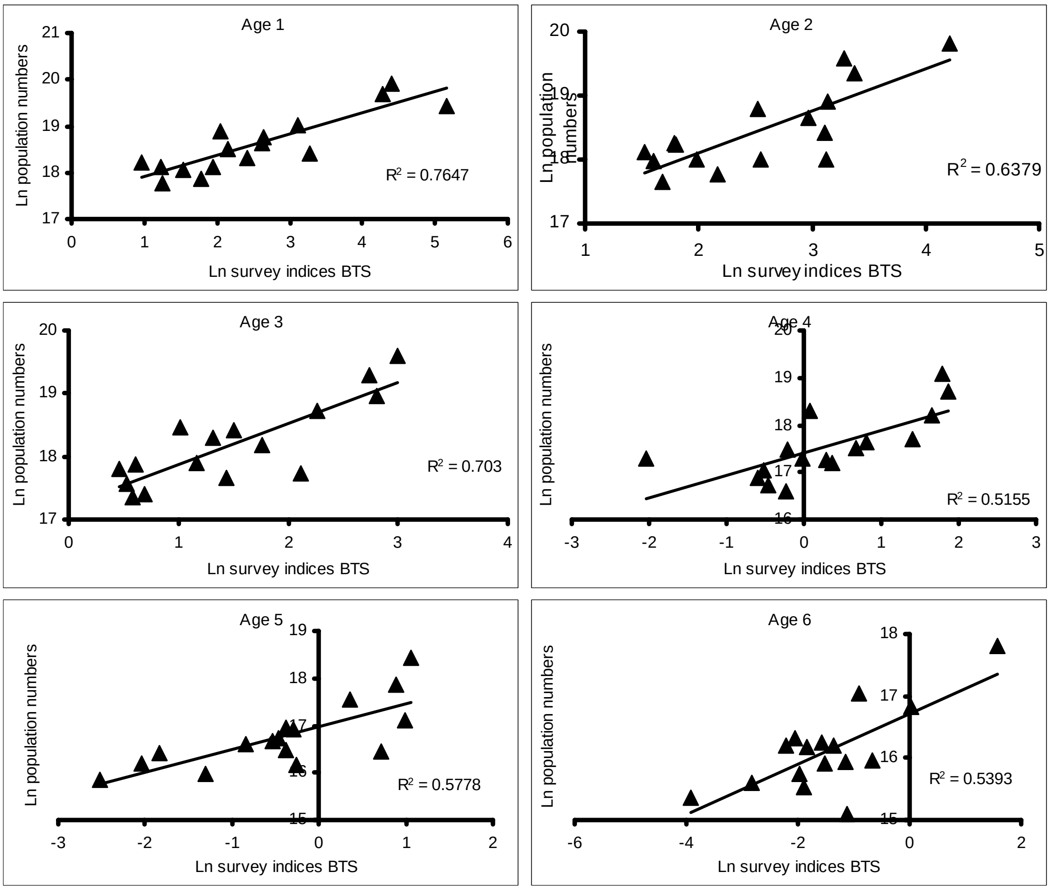


Figure I.9. Sole. Correlations between logarithm of survey indices BTS and logarithm of estimated population numbers for ages 1-6.

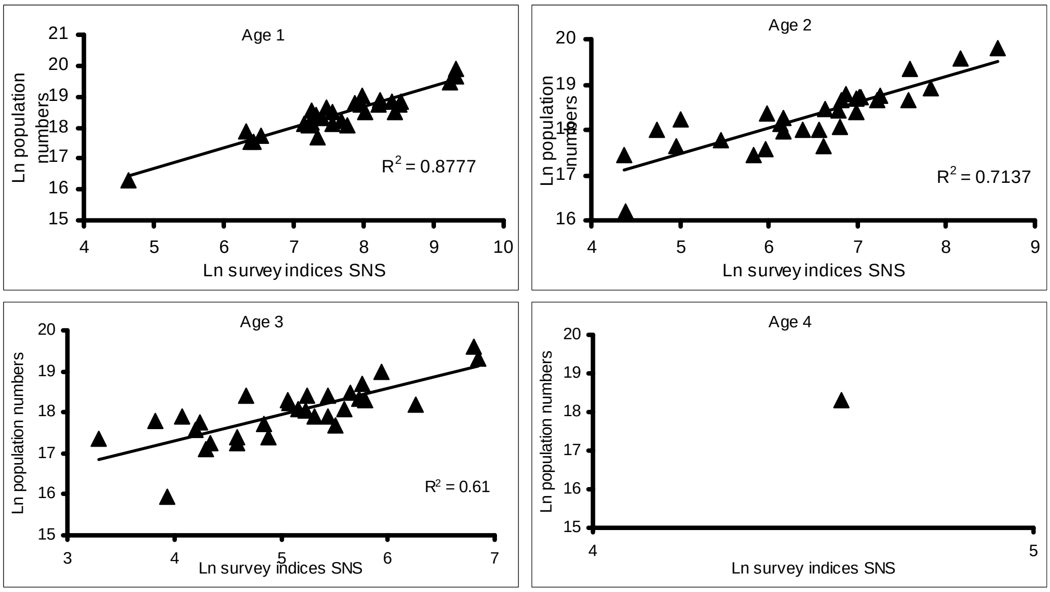
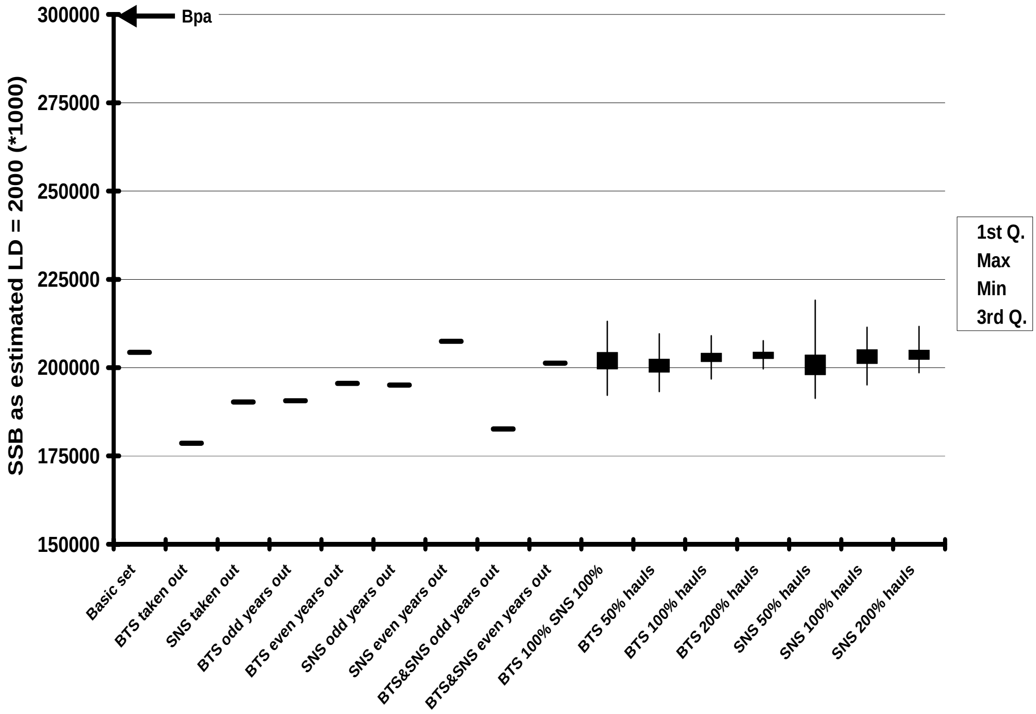


Figure I.10. Sole. Correlations between logarithm of survey indices SNS and logarithm of estimated population numbers for ages 1-4.

Appendix II



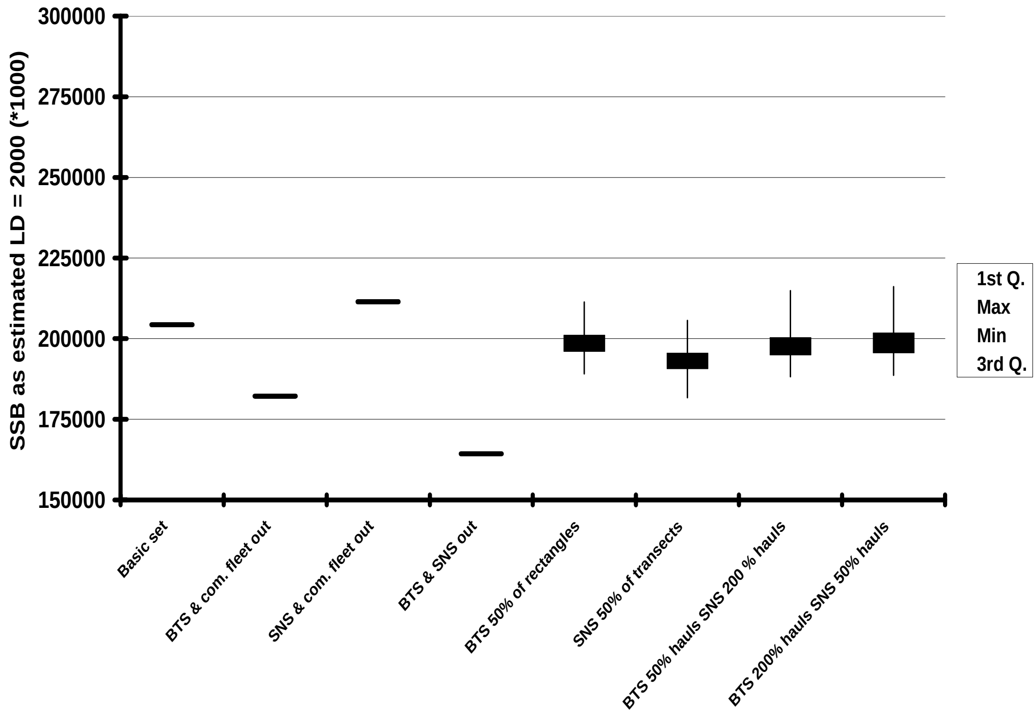
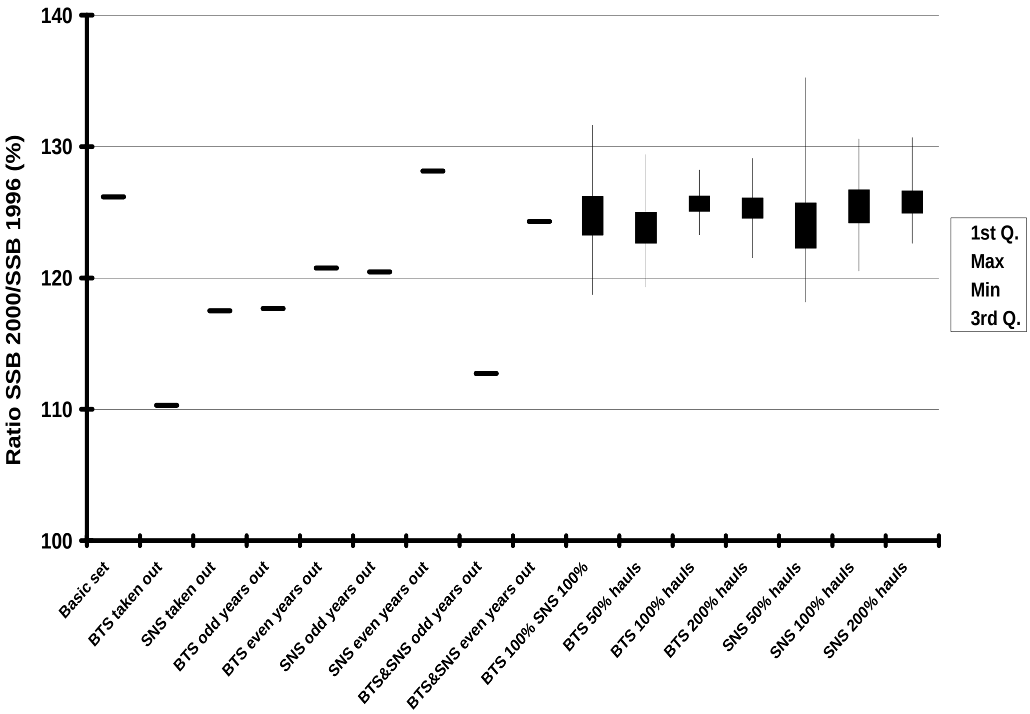


Figure II.1. Plaice. Spawning stock biomass in 2000.



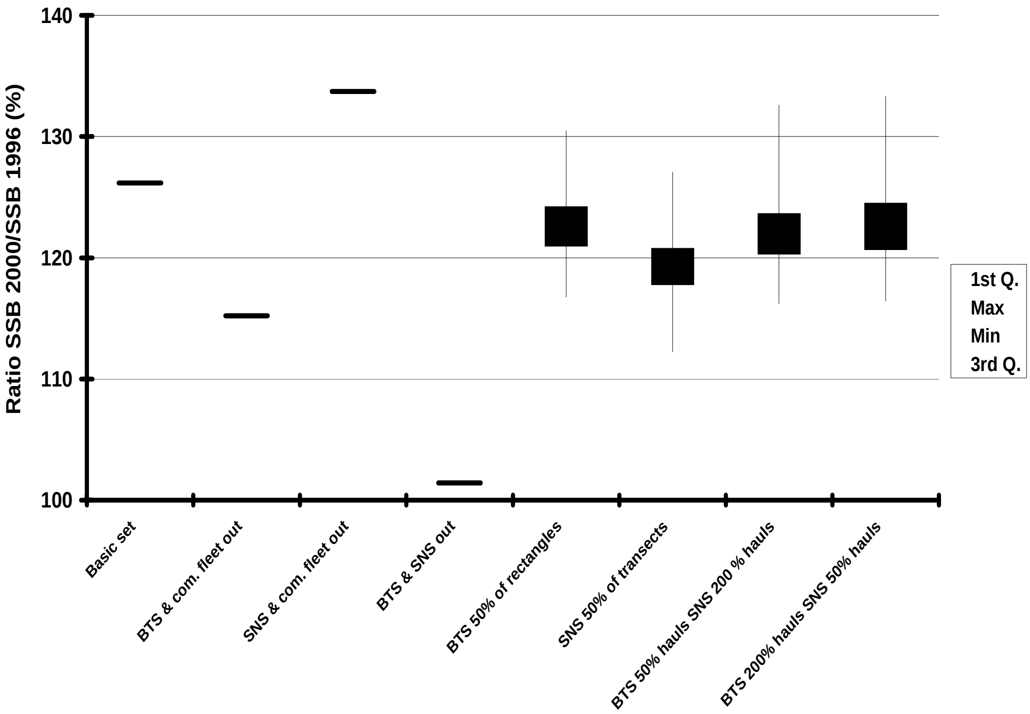
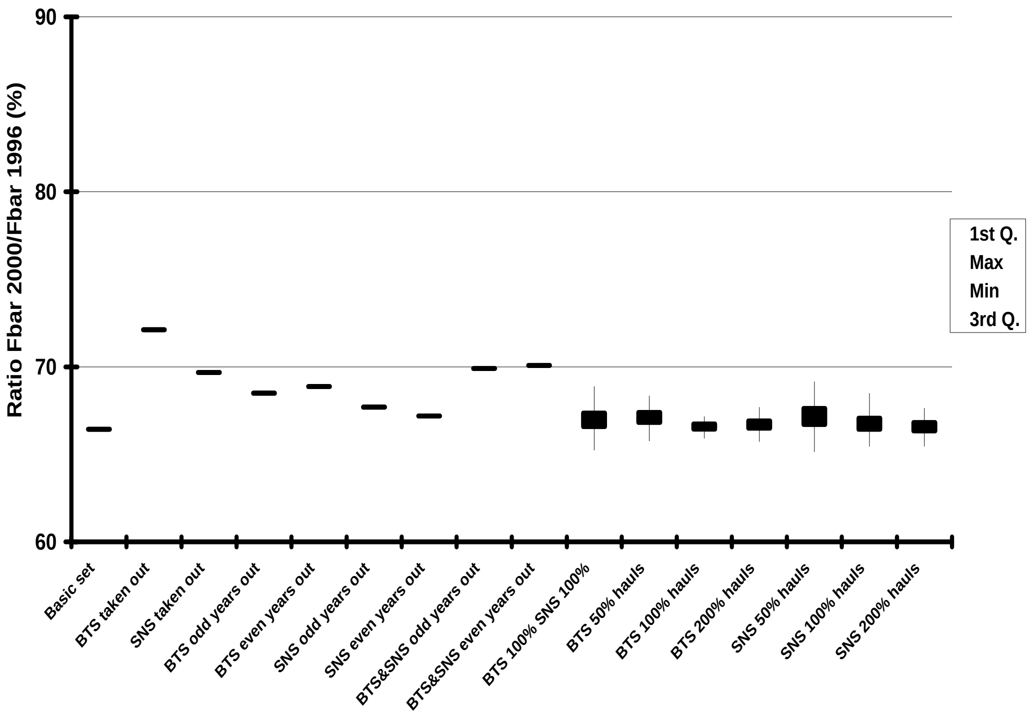


Figure II.2. Plaice. Ratio SSB 2000 and SSB 1996.



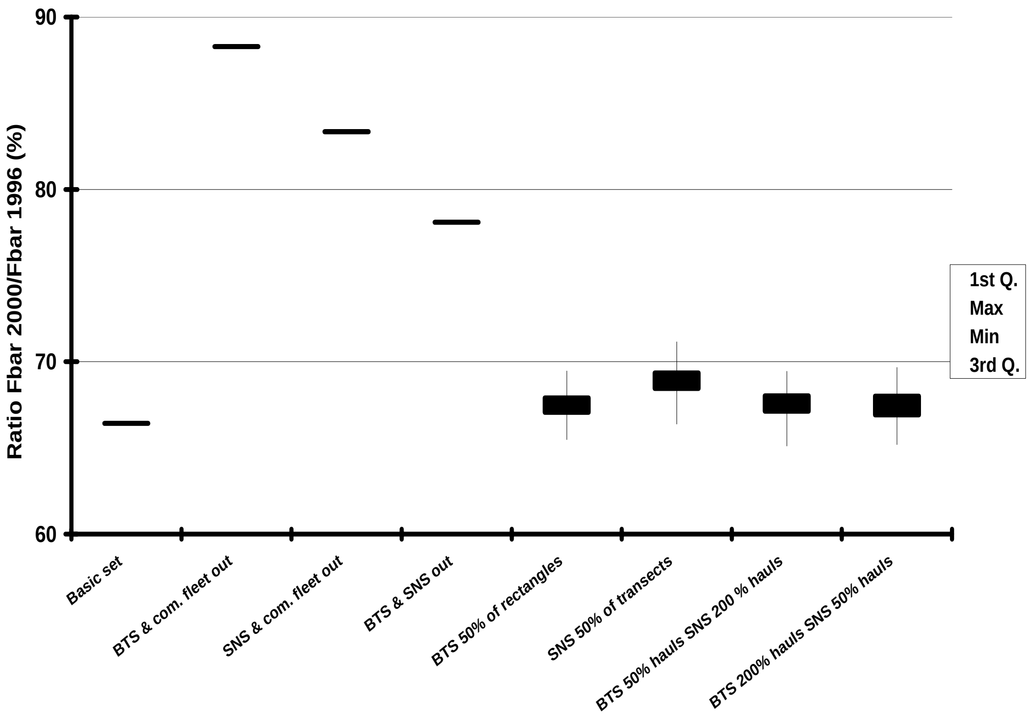


Figure II.3. Plaice. Ratio Fbar 2000 and Fbar 1996.

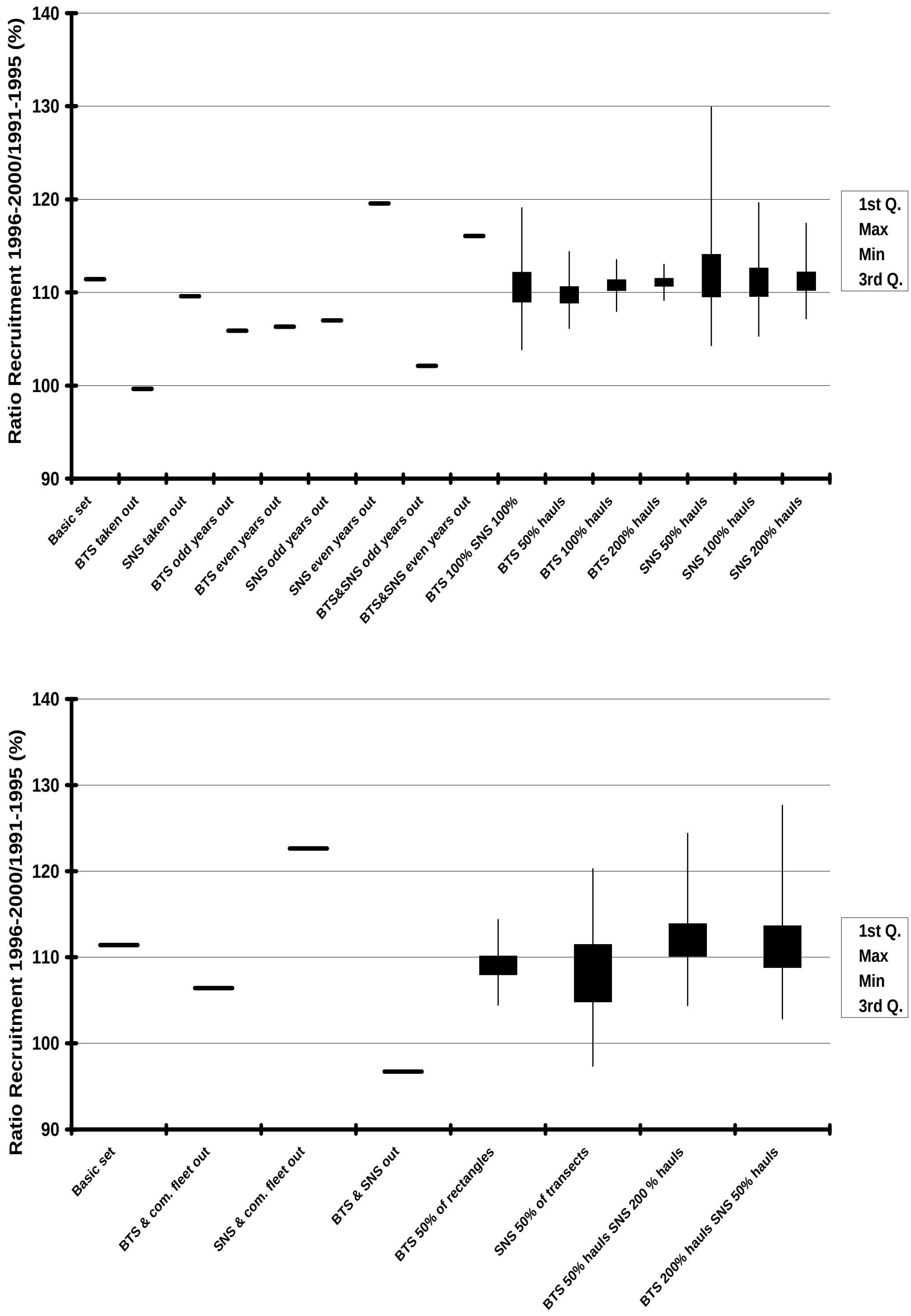
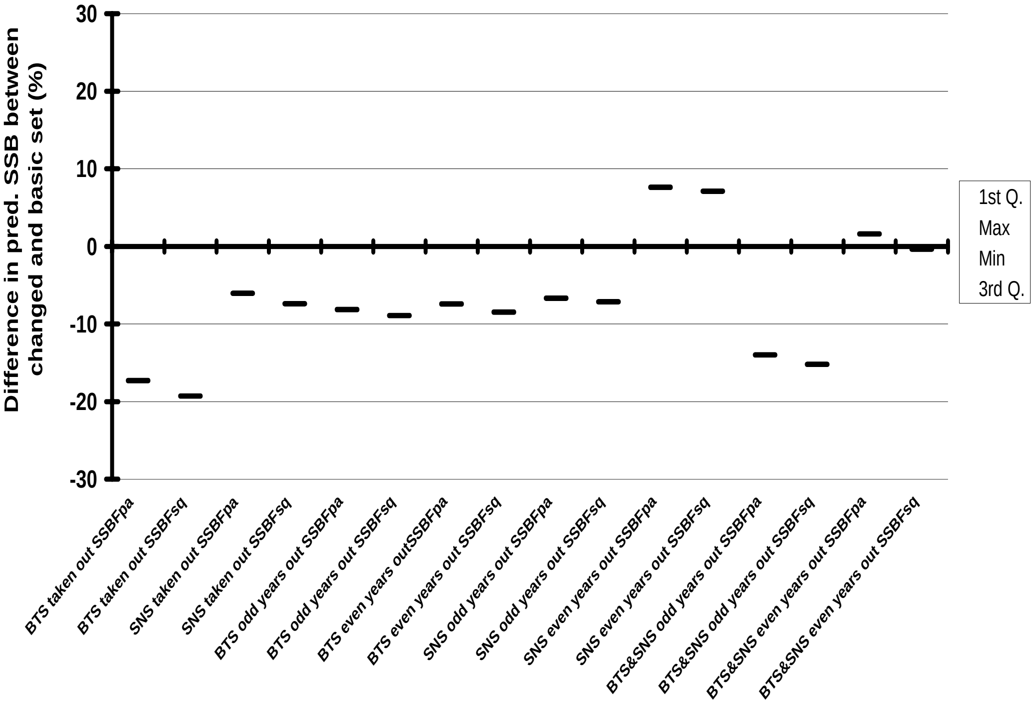


Figure II.4. Plaice. Ratio recruitment 1996-2000 / 1995-1991.



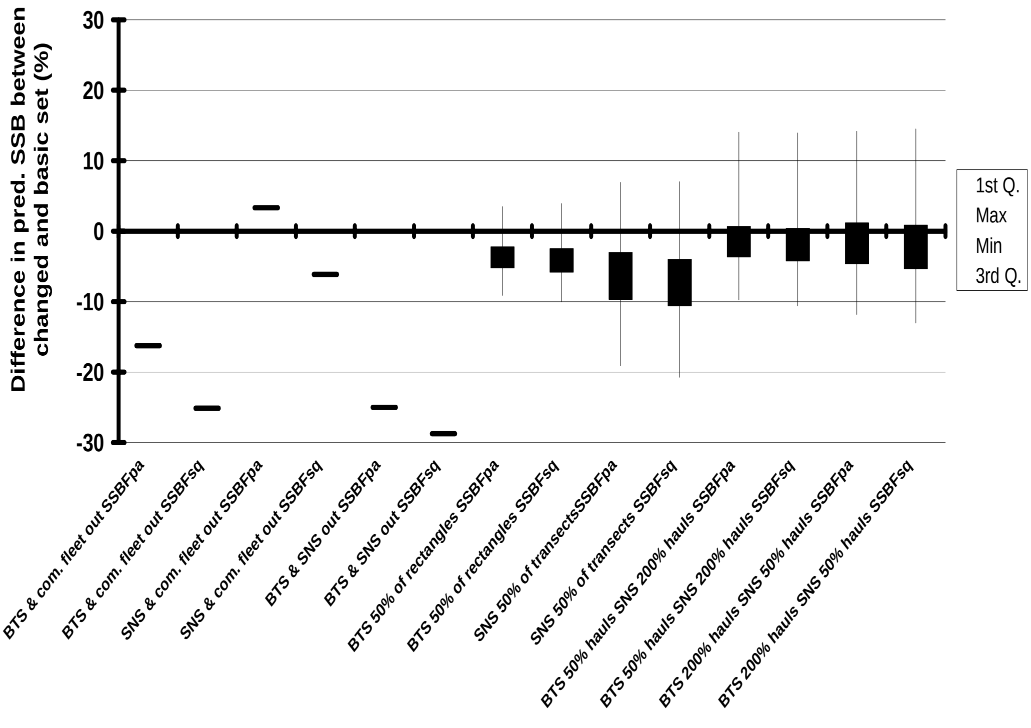


Figure II.5. Plaice. Difference in SSBFsq and SSBFpa for 2002 between changed and basic set.

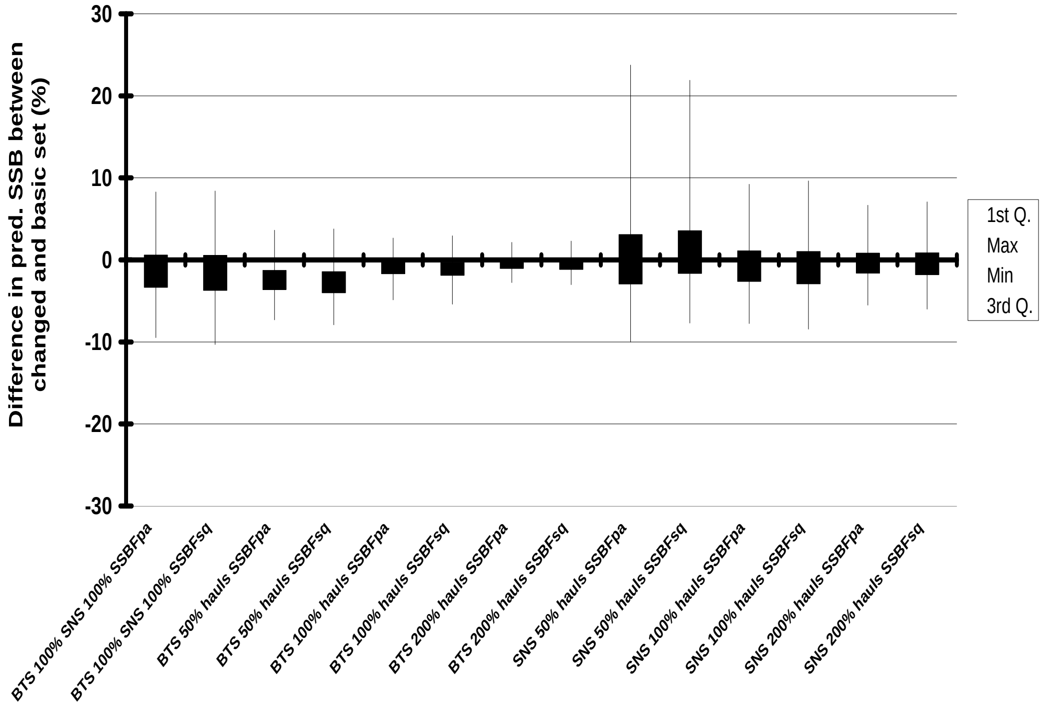
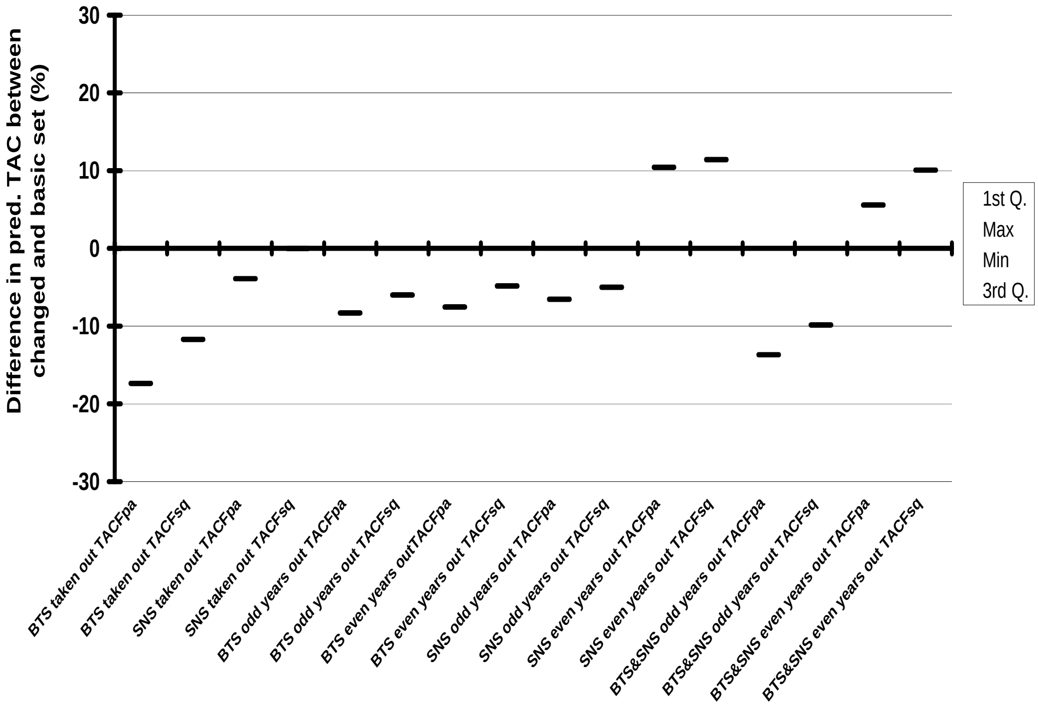


Figure II.5 Continued. Plaice. Difference in SSBFsq and SSBFpa for 2002 between changed and basic set.



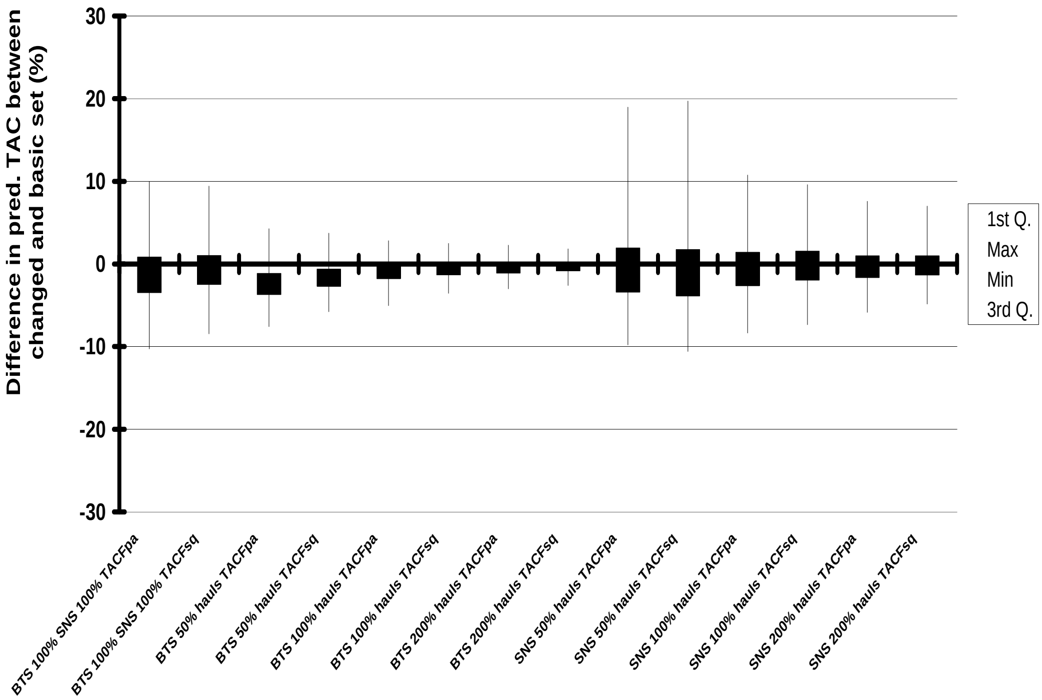


Figure II.6. Plaice. Difference in TACFsq and TACFpa for 2002 between changed and basic set.

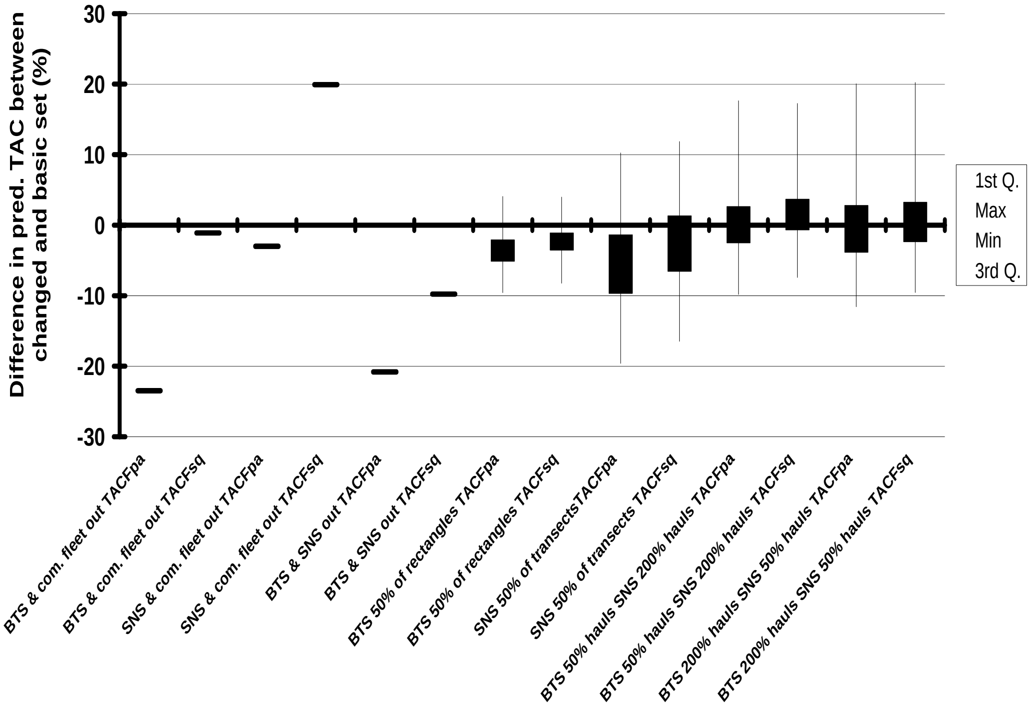
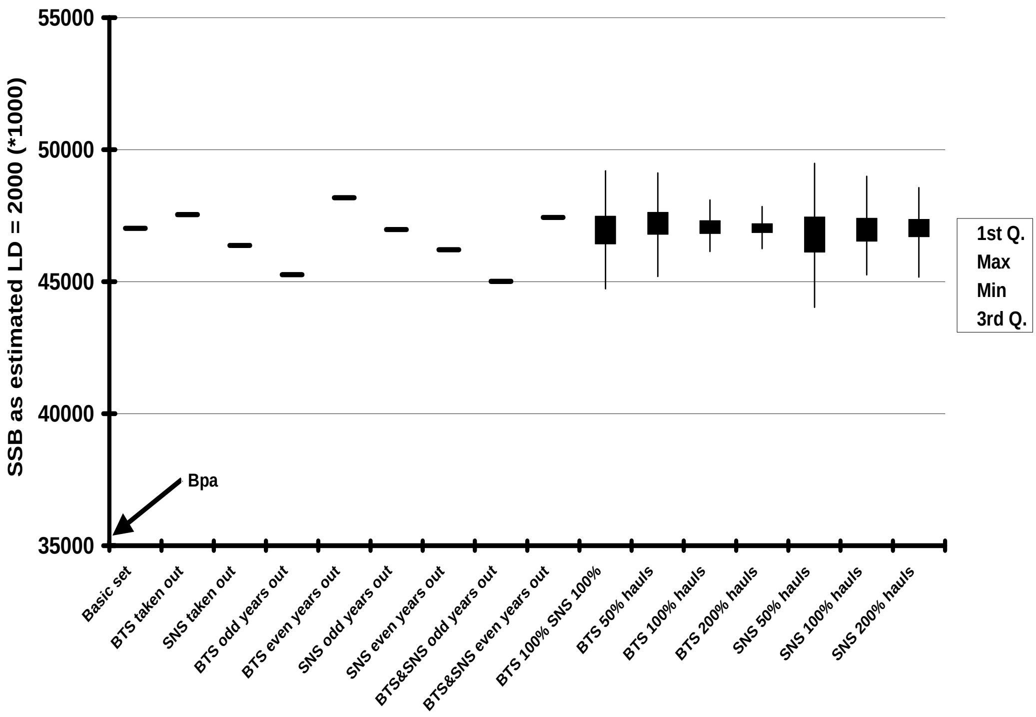


Figure II.6. Continued. Plaice. Difference in TACFsq and TACFpa for 2002 between changed and basic set.



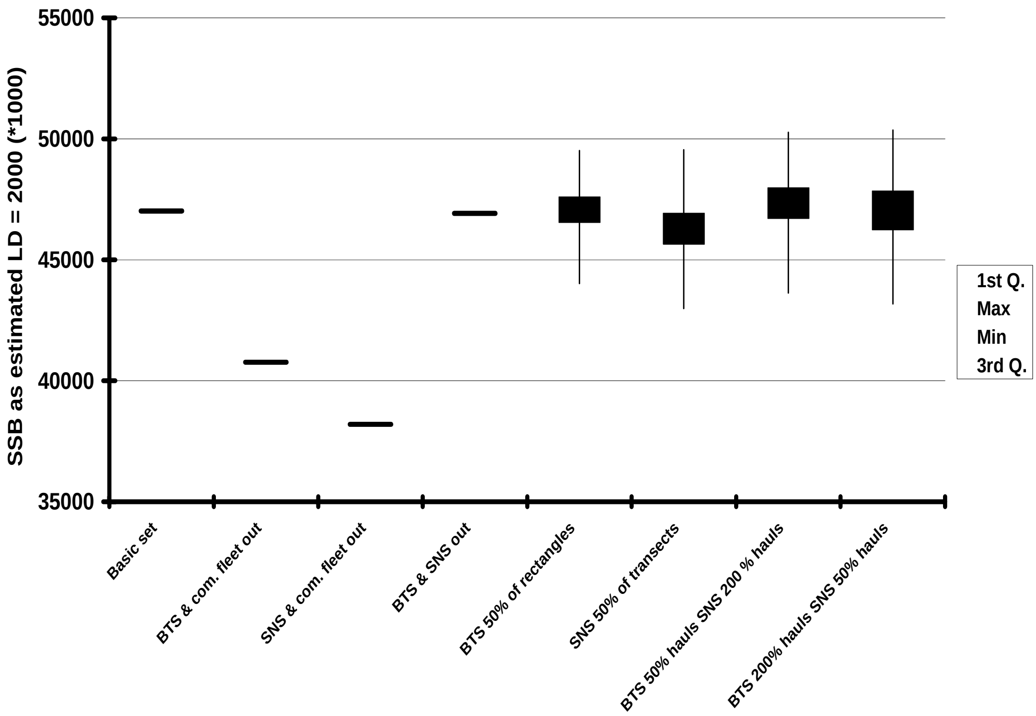


Figure II.7. Sole. Spawning stock biomass in 2000.

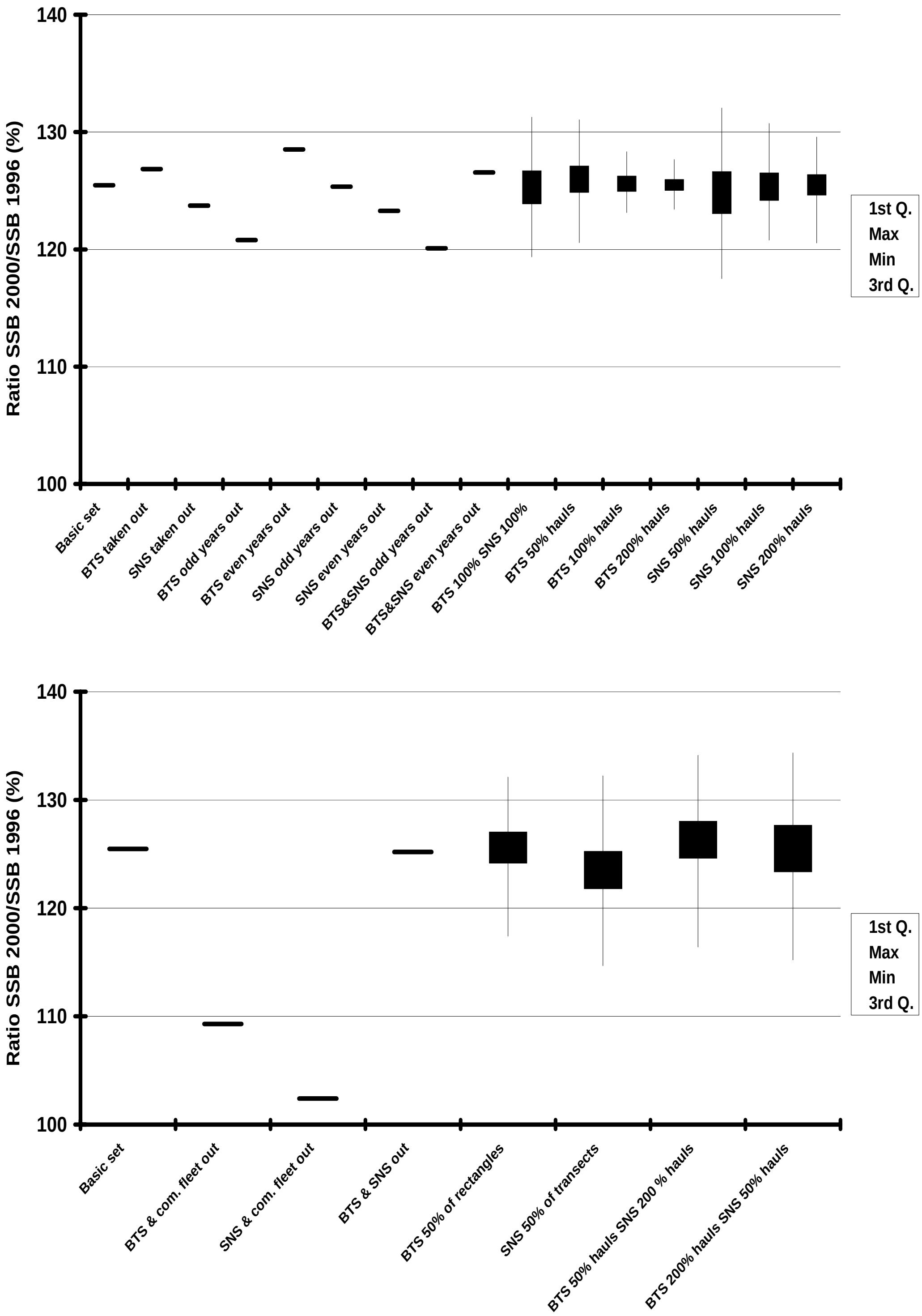
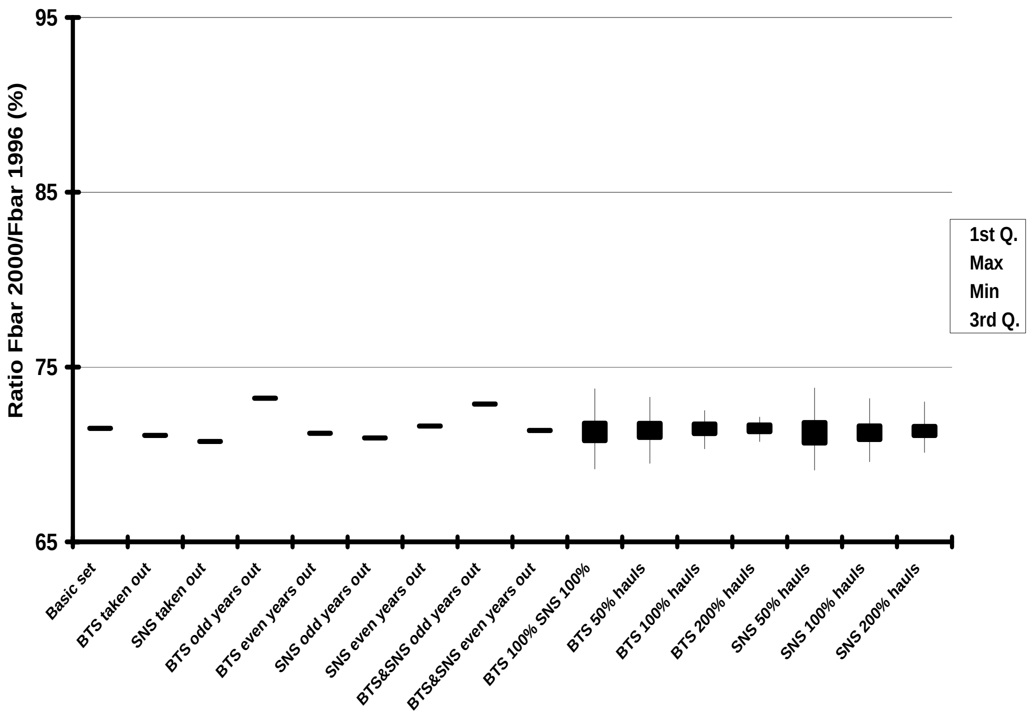


Figure II.8. Sole. Ratio SSB 2000 and SSB 1996.



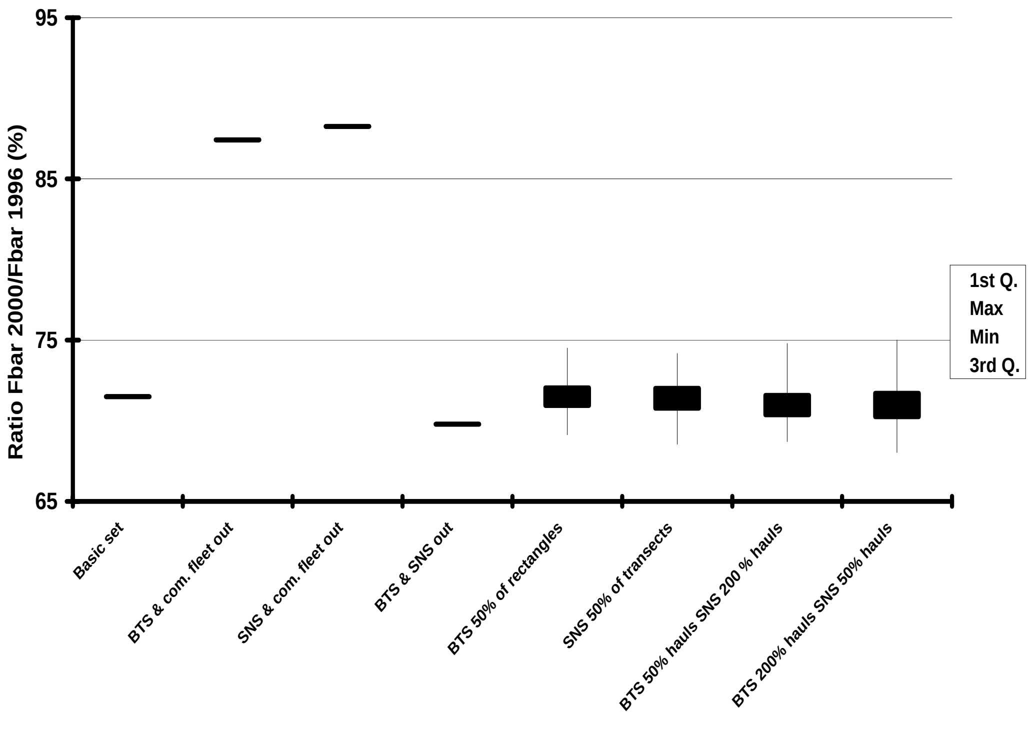
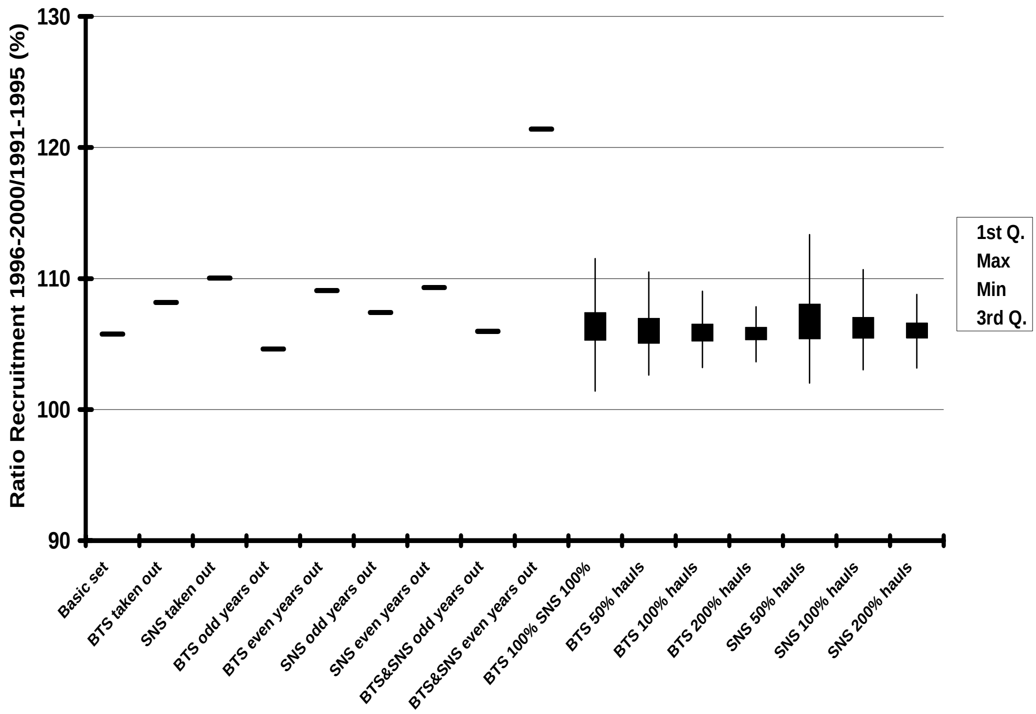


Figure II.9. Sole. Ratio Fbar 2000 and Fbar 1996.



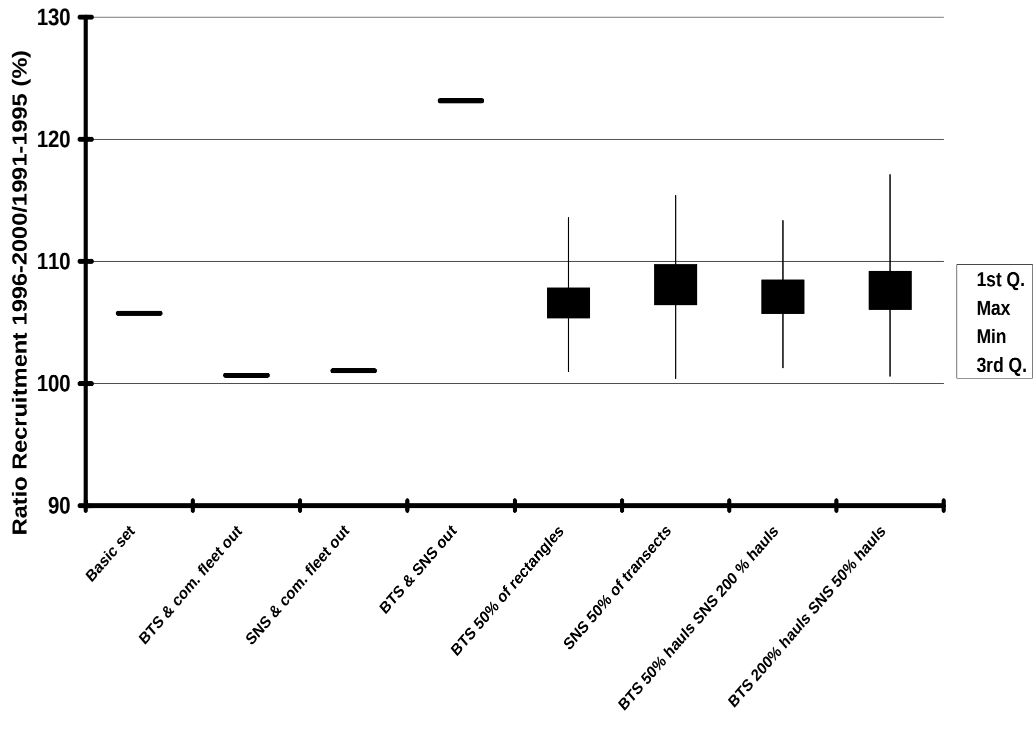


Figure II.10. Sole. Ratio recruitment 1996-2000 / 1995-1991.

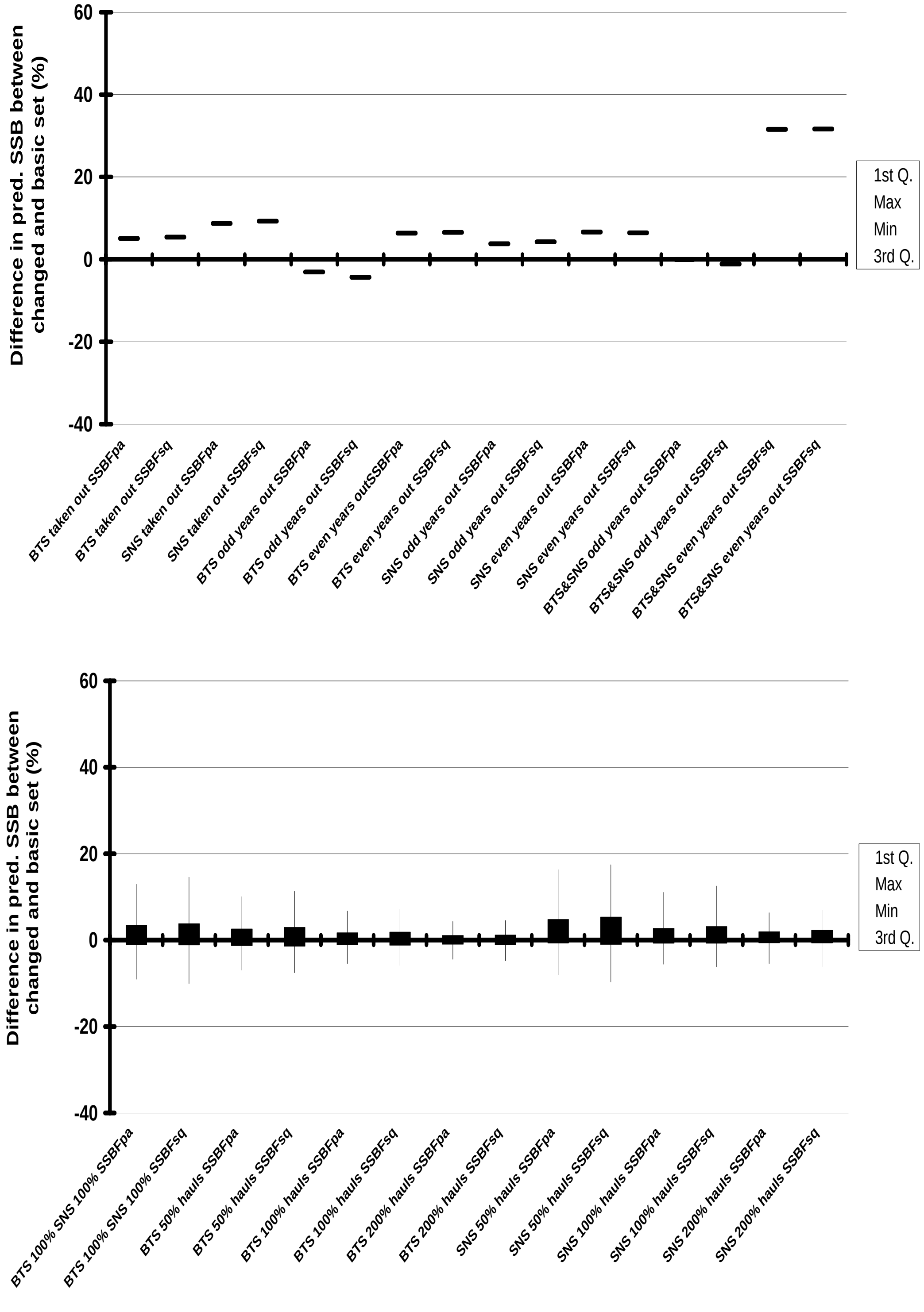


Figure II.11. Sole. Difference in SSBFsq and SSBFpa for 2002 between changed and basic set.

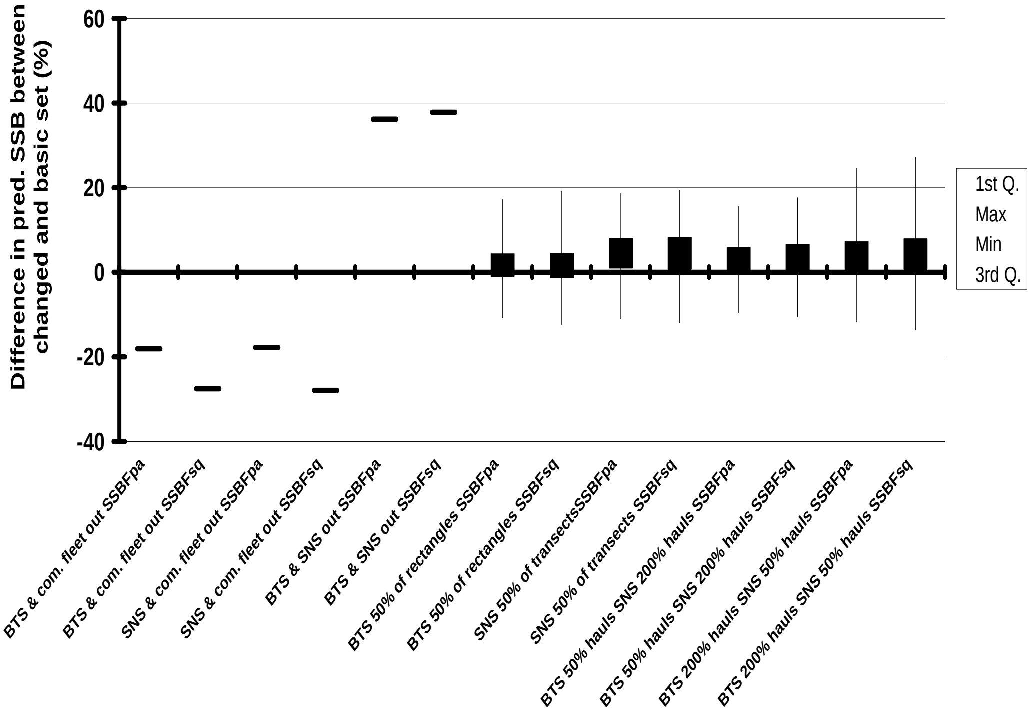


Figure II.11. Continued. Sole. Difference in SSBFsq and SSBFpa for 2002 between changed and basic set.

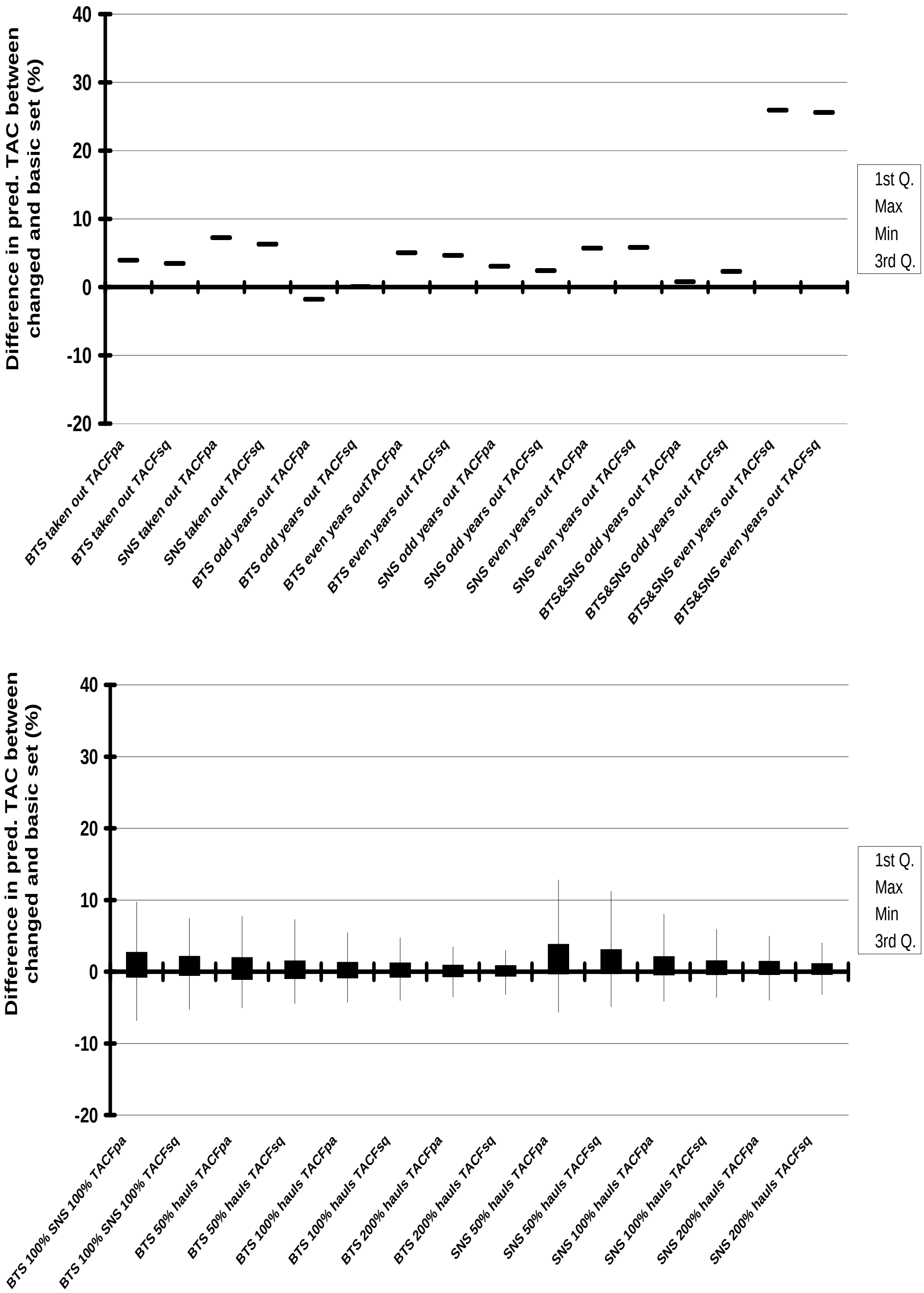


Figure II.12. Sole. Difference in TACFsq and TACFpa for 2002 between changed and basic set.

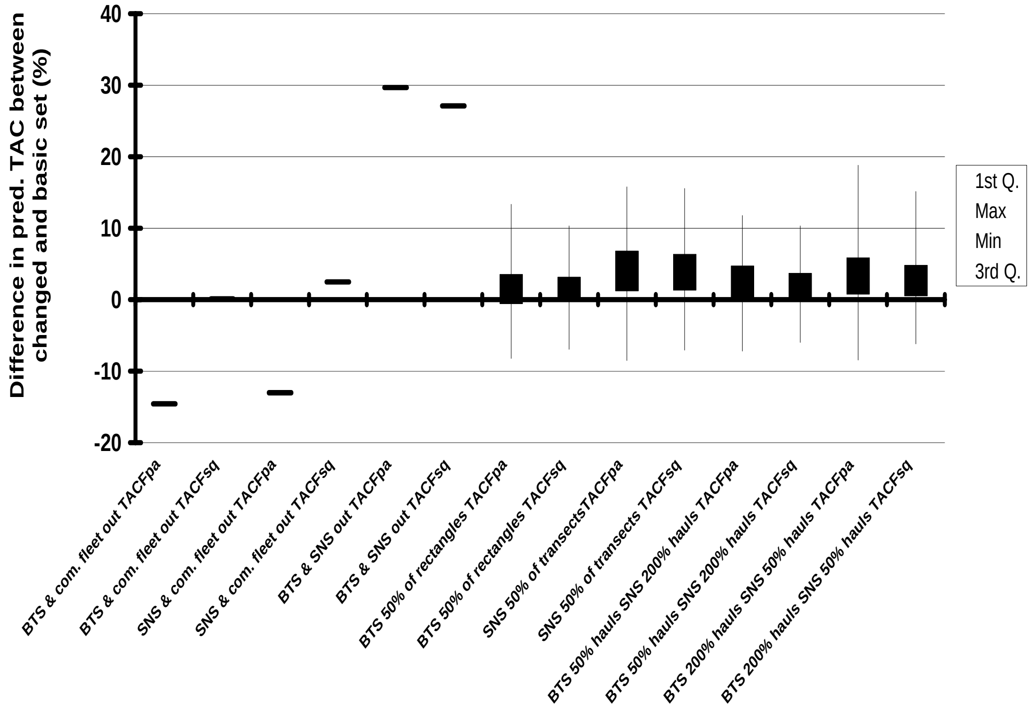


Figure II.12. Continued. Sole. Difference in TACFsq and TACFpa for 2002 between changed and basic set.

Appendix III

Table III.1. Results of the analysis of variance for 10 age classes of plaice and sole. Symbols indicate significance of each term in the model (D is depth (meters), C is cohort (year class), S is survey (DFS, BTS, SNS), M is month of the year, and Y is year, R^2 is the percentage of variance explained by the model.

Species	Age	D	D^2	C	S	M	D x Y	D x M	D^2 x C	D x S	R^2
Plaice	0	+++	+++	+++	+++	+++	+++	+++	+++	+++	49%
	1	+++	+++	+++	+++	+++	+++	+++	+++	+++	39%
	2	+++	+++	+++	+++	+++	+++	+++	+++	+++	52%
	3	+++	+++	+++	+++	+++	+++	n.s.	+++	+++	51%
	4	+++	n.s.	+++	++	n.s.	+++	n.s.	+++	+++	39%
	5	+++	+++	+++	n.s.	+++	+++	n.s.	+++	++	32%
	6	+++	+++	++	+++	+	+++	+++	+++	+++	29%
	7	n.s.	+++	+++	++	++	n.s.	n.s.	+++	n.s.	14%
	8	n.s.	n.s.	n.s.	n.s.	++	n.s.	++	+++	n.s.	13%
	9	n.s.	+++	n.s.	+	n.s.	n.s.	n.s.	+++	n.s.	8%
	10	n.s.	+	n.s.	n.s.	n.s.	n.s.	n.s.	+++	n.s.	5%
Sole	0	+++	n.s.	+++	+++	n.s.	+++	n.s.	+++	+++	54%
	1	+	+	+++	+++	n.s.	+	n.s.	+	+++	34%
	2	+++	n.s.	+++	+++	+++	+++	+++	+++	+++	41%
	3	n.s.	+++	+++	+++	+++	n.s.	+++	+++	+++	41%
	4	n.s.	n.s.	+++	+++	++	n.s.	+++	+++	+++	30%
	5	n.s.	++	++	+++	n.s.	n.s.	n.s.	+++	+++	25%
	6	n.s.	n.s.	++	+++	n.s.	n.s.	++	+++	+++	20%
	7	+++	n.s.	n.s.	+	n.s.	+++	n.s.	+++	+++	14%
	8	n.s.	n.s.	+++	+++	n.s.	n.s.	n.s.	+++	+++	14%
	9	n.s.	n.s.	+++	+++	n.s.	n.s.	n.s.	++	++	3%
	10	+	n.s.	n.s.	+	n.s.	+	n.s.	+++	++	5%

Significance: "n.s.": $P>0.10$
'+': $0.05<P<0.10$
'++': $0.01<P<0.05$
'+++': $P<0.01$

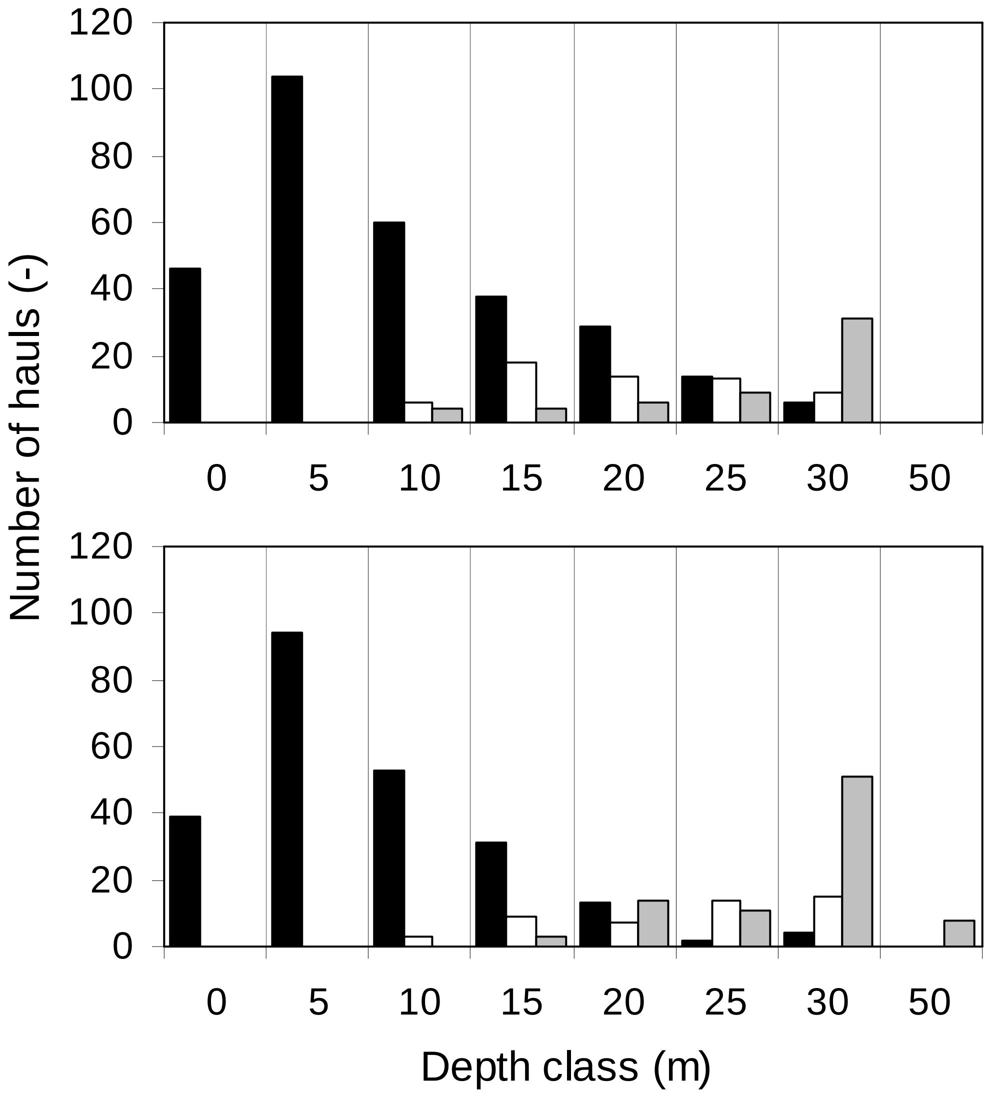


Figure III.1. Distribution of hauls from the three surveys over depth classes in 1985 (top panel) and 2002 (bottom panel). Black bars: DFS; white bars: SNS; grey bars BTS.

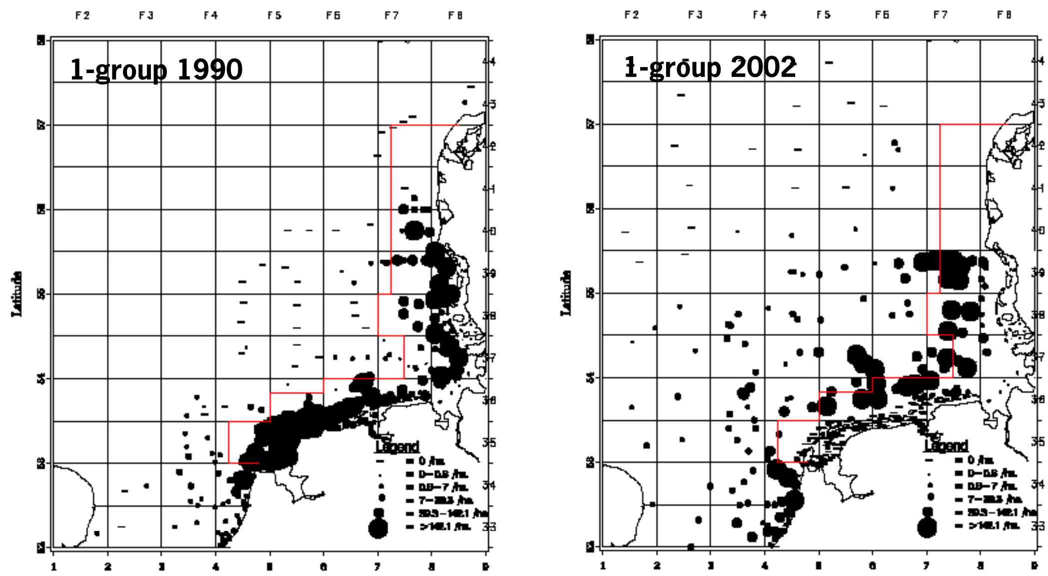


Figure III.2. Spatial distribution of age-1 plaice in 1990 and 2002. There is a clear movement of age-1 plaice to deeper areas (From Grift et al 2004).

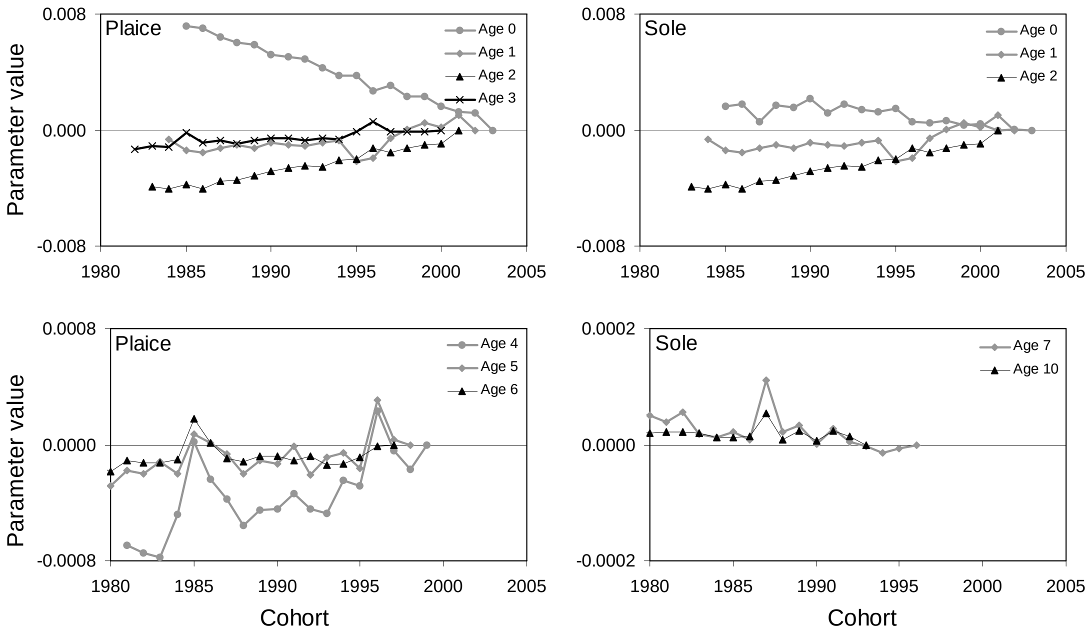


Figure III.3. Trends in parameter values of the interaction term (depth x cohort) for plaice and sole per age group. Only significant trends are shown (Table III.2). A higher value indicates of the parameter indicates that abundance increases more with depth.

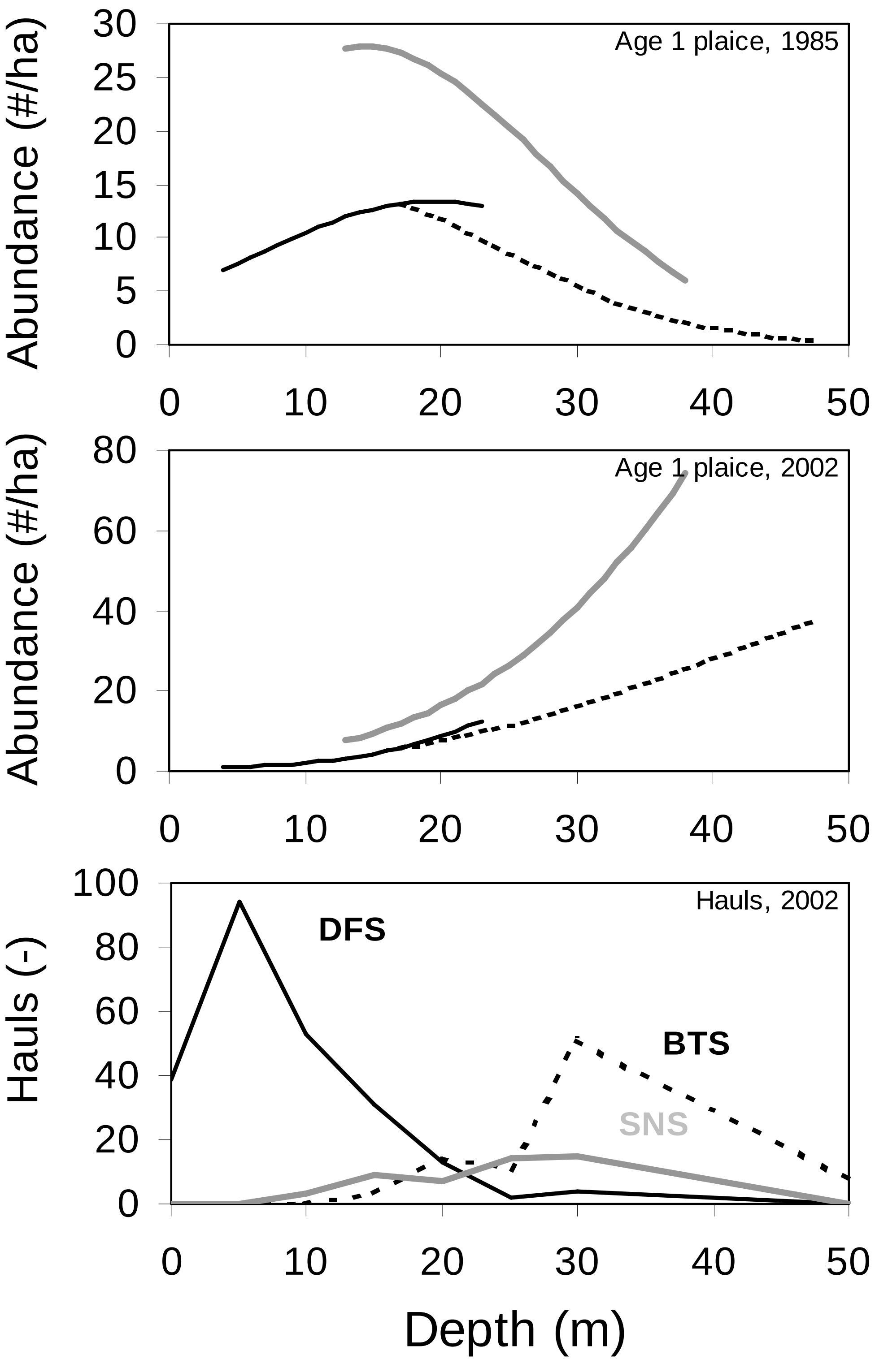


Figure 4.4. Changes in the abundance of Age-1 plaice at different depths between 1985 (upper panel) and 2002 (middle panel). The lower panel shows the number of hauls per depth class in each survey. Predicted values for limited depth classes (the 5-95 percentiles of depths sampled for each survey) for the month September. Solid black line: DFS; dashed black line: BTS; grey solid line: SNS.

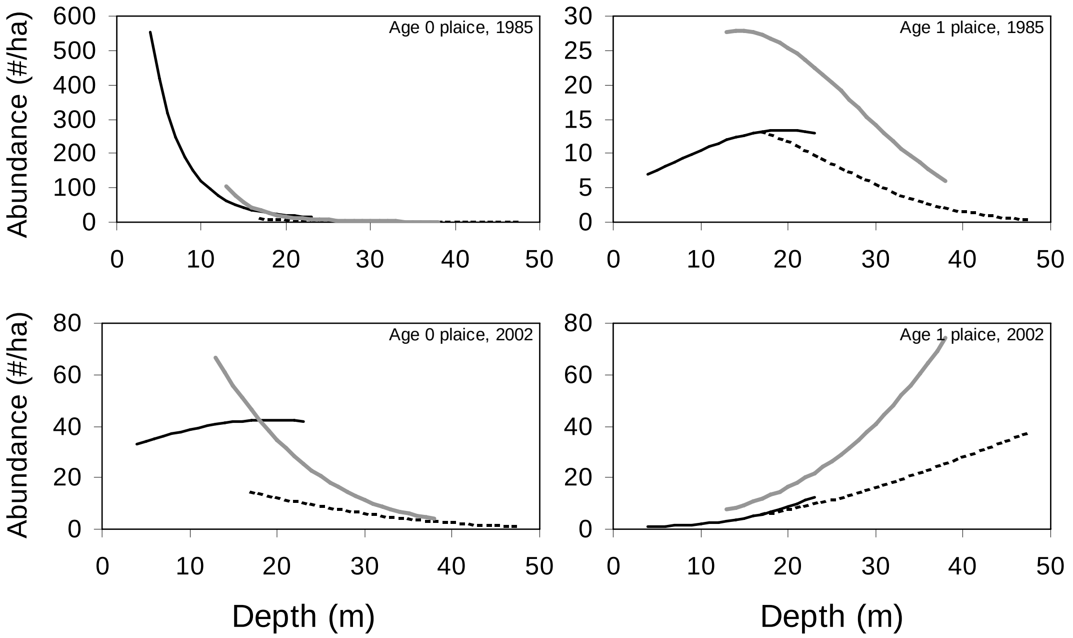


Figure 4.5. Abundance of plaice age 0 and 1 at different depths in 1985 and 2002. Predicted values for limited depth classes (the 5-95 percentiles of depths sampled for each survey) for the month September. Solid black line: DFS; dashed black line: BTS; grey solid line: SNS.

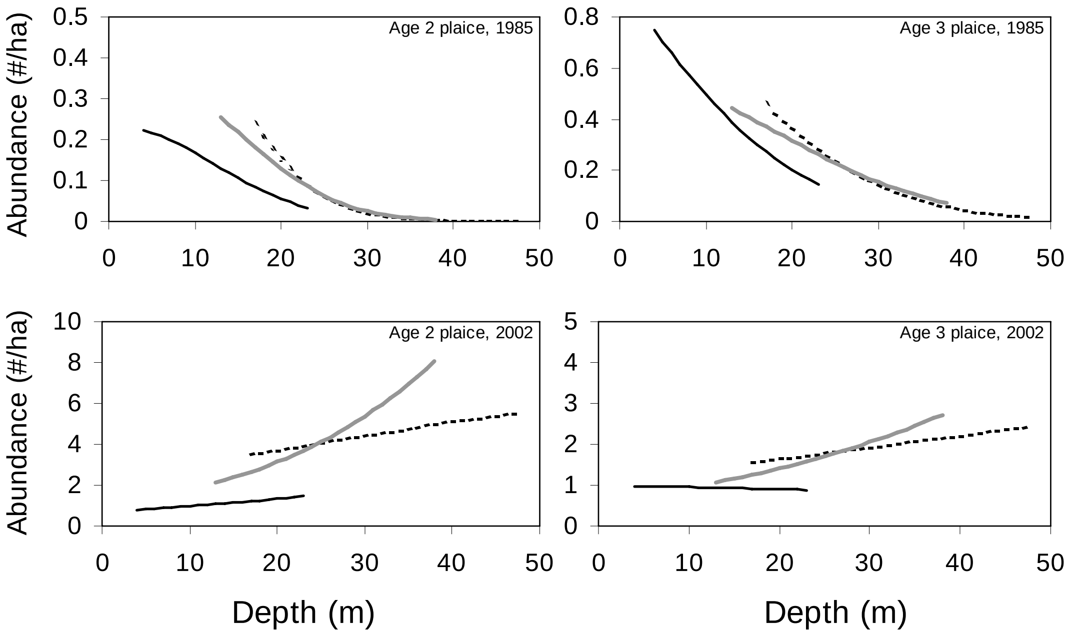


Figure 4.6. Abundance of plaice age 2 and 3 at different depths in 1985 and 2002. Predicted values for limited depth classes (the 5-95 percentiles of depths sampled for each survey) for the month September. Solid black line: DFS; dashed black line: BTS; grey solid line: SNS.

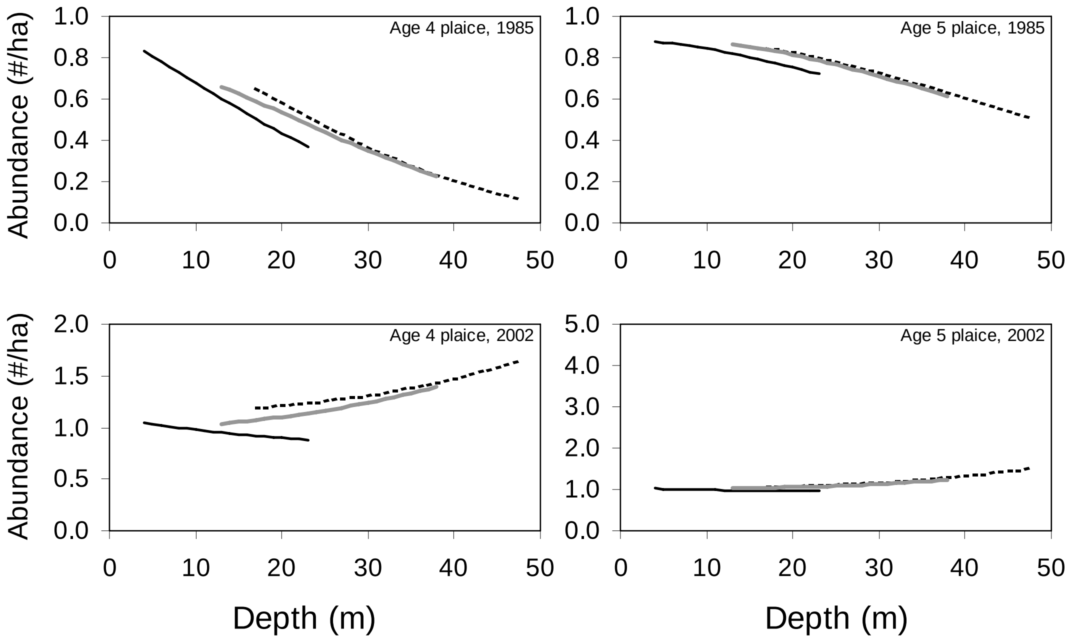


Figure 4.7. Abundance of plaice age 4 and 5 at different depths in 1985 and 2002. Predicted values for limited depth classes (the 5-95 percentiles of depths sampled for each survey) for the month September. Solid black line: DFS; dashed black line: BTS; grey solid line: SNS.

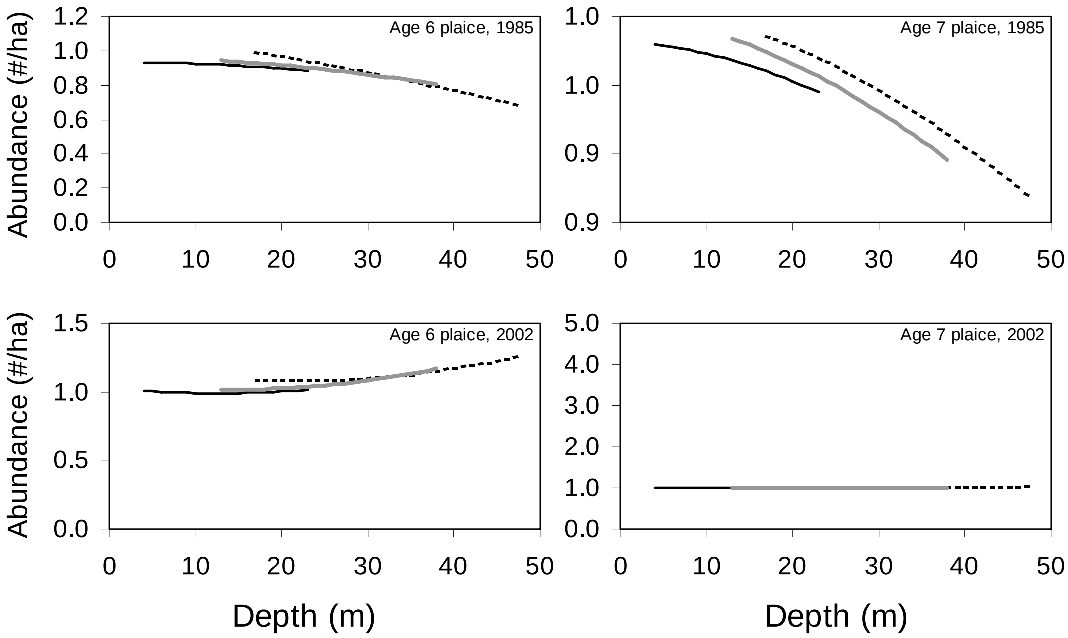


Figure 4.8. Abundance of plaice age 6 and 7 at different depths in 1985 and 2002. Predicted values for limited depth classes (the 5-95 percentiles of depths sampled for each survey) for the month September. Solid black line: DFS; dashed black line: BTS; grey solid line: SNS.

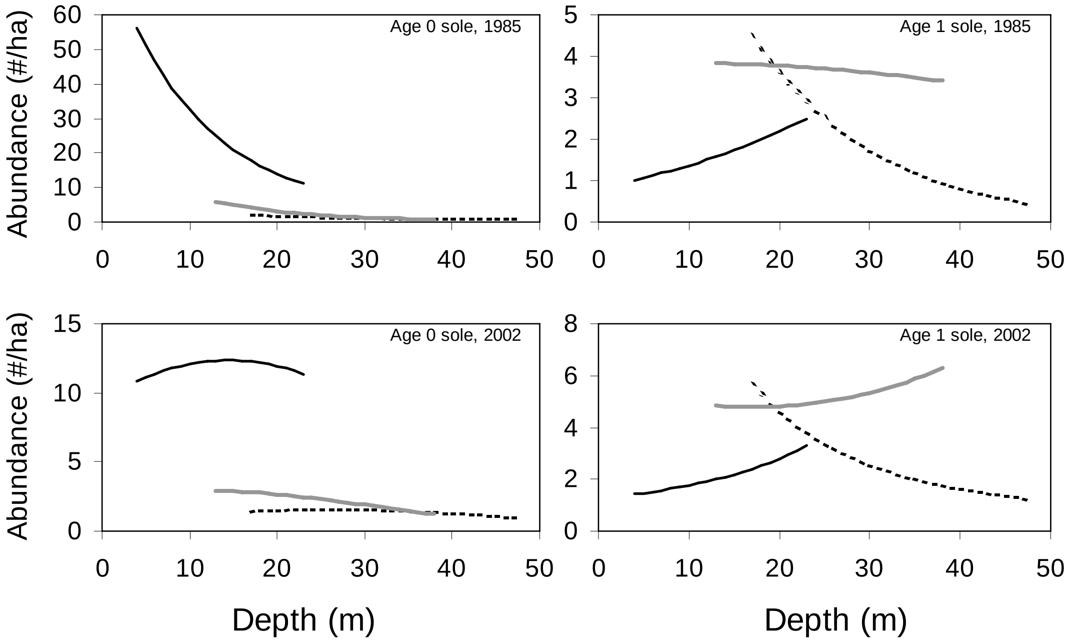


Figure 4.9. Abundance of sole age 0 and 1 at different depths in 1985 and 2002. Predicted values for limited depth classes (the 5-95 percentiles of depths sampled for each survey) for the month September. Solid black line: DFS; dashed black line: BTS; grey solid line: SNS.

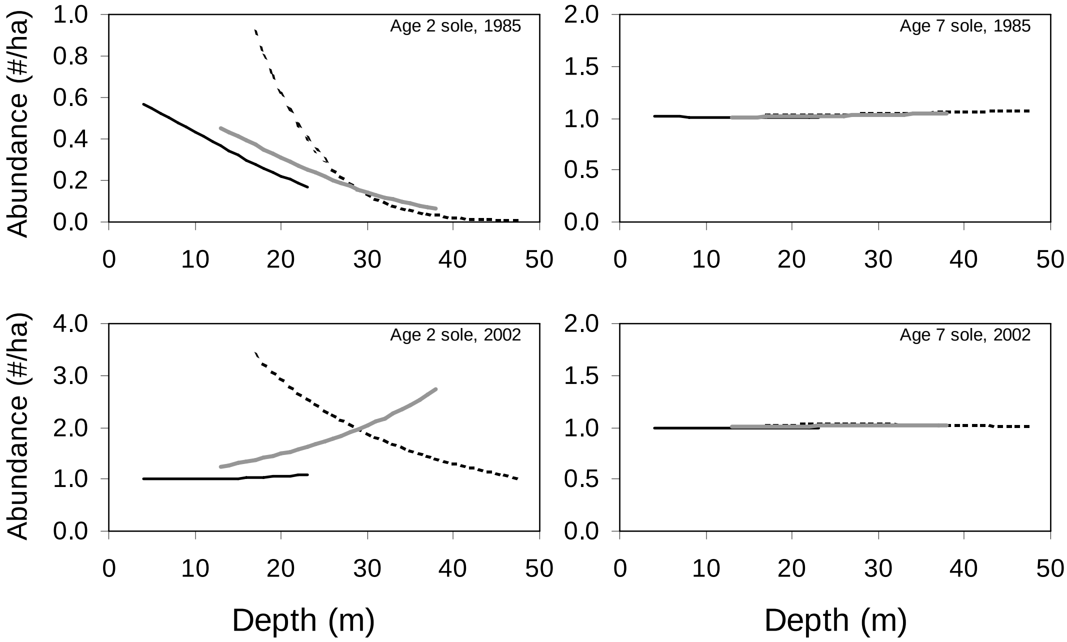


Figure 4.10. Abundance of sole age 2 and 7 at different depths in 1985 and 2002. Predicted values for limited depth classes (the 5-95 percentiles of depths sampled for each survey) for the month September. Solid black line: DFS; dashed black line: BTS; grey solid line: SNS.

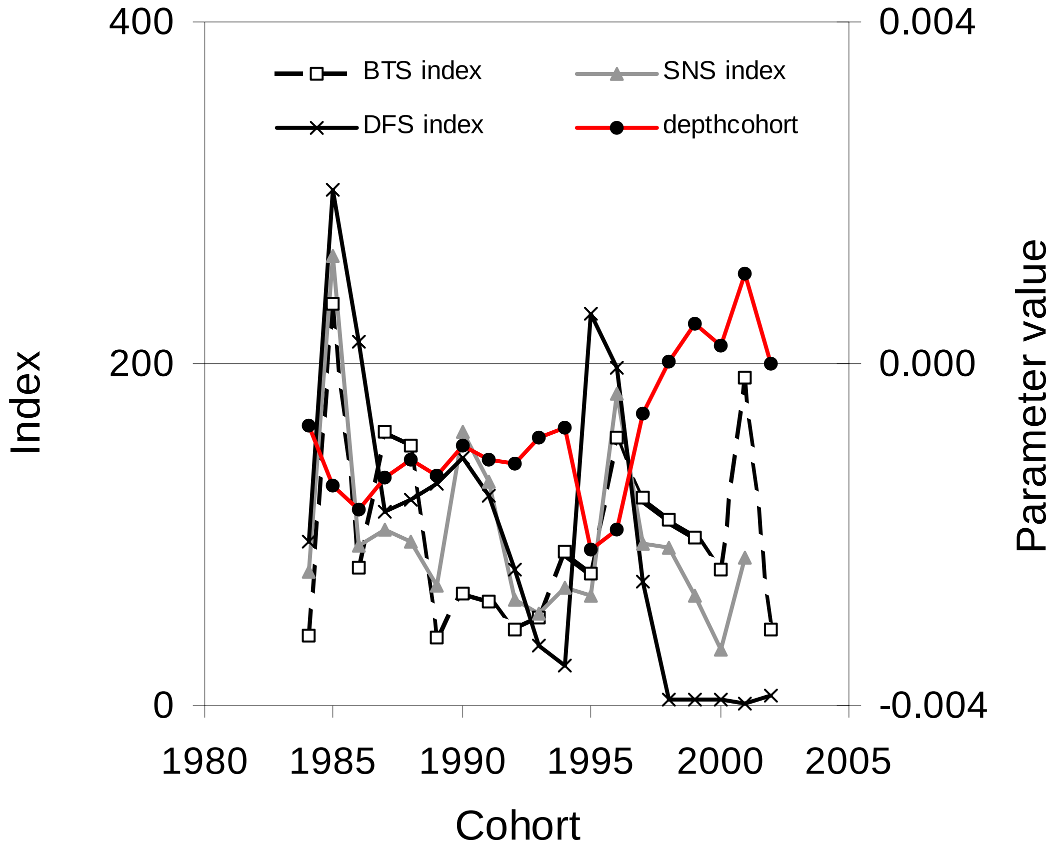


Figure 4.11. Trends in the index for Age-1 plaice from the BTS, SNS and DFS (left y-axis) and the parameter value for the interaction term between depth and cohort (depthcohort; red line; right y-axis). The parameter value indicates the abundance of age-1 plaice at different depths; when it is positive, abundance *increases* with depth whereas it *decreases* with depth when its value is negative.