

ICES Report to EU commision
on survey variance 2007

**CONFIDENCE LIMITS ESTIMATION OF
ABUNDANCE INDICES FROM
BOTTOM TRAWL SURVEY DATA:
IMPLEMENTATION IN DATRAS**

FEBRUARY 2007

ICES HEADQUARTERS, COPENHAGEN



International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

H. C. Andersens Boulevard 44-46

DK-1553 Copenhagen V

Denmark

Telephone (+45) 33 38 67 00

Telefax (+45) 33 93 42 15

www.ices.dk

info@ices.dk

Recommended format for purposes of citation:

ICES. 2012. Report of the Workshop on Implementation in DATRAS of Confidence Limits Estimation of, 10-12 May 2006, ICES Headquarters, Copenhagen. . 53 pp.

For permission to reproduce material from this publication, please apply to the General Secretary.

The document is a report of an Expert Group under the auspices of the International Council for the Exploration of the Sea and does not necessarily represent the views of the Council.

© 2012 International Council for the Exploration of the Sea.

Contents

1	Introduction	1
2	Description of surveys stored in DATRAS	2
2.1	Data stored.....	2
2.2	Survey design.....	4
2.2.1	BITS.....	4
2.2.2	NS-IBTS.....	5
2.2.3	BTS.....	7
2.2.4	EVHOE.....	8
2.2.5	Scottish Groundfish Survey.....	10
3	Calculation of indices	12
3.1	BITS.....	13
3.2	NS-IBTS indices calculation.....	14
3.3	BTS indices.....	14
3.4	EVHOE.....	14
3.5	Scottish Western IBTS.....	15
4	Choice of method for uncertainty estimation	15
5	How to bootstrap in DATRAS	16
5.1	Length distribution bootstrapping.....	17
5.2	Bootstrap age-length keys by length class.....	18
5.3	Number of bootstraps.....	19
5.4	The index calculations.....	19
5.5	Reference Uncertainty for Abundance Indices.....	19
5.6	Presentation of results.....	19
5.7	Conclusion.....	21
6	Effect of a change in sampling intensity	22
6.1	Effect of change in coverage – No of hauls.....	23
6.2	Effect of change in age sampling - No of aged fish per length class.....	23
7	Final conclusion	24
8	References	24
	Annex 1: Range/median figures for all surveys stored in DATRAS	26

1 Introduction

Basic data from several ICES coordinated trawl and beam trawl surveys in the Baltic Sea, the North Sea, the area west of Scotland and France are stored in ICES' **DATA**base of **TRAWL** Surveys, DATRAS, in the ICES Secretariat (Figure 1). Every year and for each survey, member countries report data, and indices of fish stock abundance are calculated and fed into the routine fish stock assessment work.

Procedures to calculate uncertainties of these indices were not developed for the first version of DATRAS released in 2003. Such estimates would be very useful for the population modelling used in the assessment work and for evaluating individual surveys, e.g. to determine whether they should be intensified to improve the precision of the fish stock assessment models.

Therefore, the EU Commission requested ICES to implement uncertainty estimation in DATRAS and supplement the routine abundance indices with uncertainty estimates.

The terms "variance estimation" and "uncertainty of the abundance estimator" are used loosely. In the context of DATRAS, these terms refer to the confidence interval, i.e. the interval where we have, $x\%$ falling below and $x\%$ falling above. There are therefore two decisions to be made:

- Decision on the percentage (x) corresponding to the tail of the distribution of the abundance index estimator;
- Decision on the procedure to calculate the distribution

Obviously, the first decision is arbitrary but it is required for consistency between years and between surveys. For convenience and to avoid excessive computer time it was decided that present the range for the abundance indices i.e. the difference between the quartiles: Q75%–Q25%.

Therefore, herein, focus is on the second question: how to calculate the distribution. This paper analyses statistical methods and their implementation.

We considered the technical advantages and disadvantages of a range of methods, including their suitability for implementation in DATRAS, making use of the work reported in:

- ICES Workshop on the Analysis of Trawl Survey Data (ICES 1992/D:6), which reviewed survey design and index definition;
- Nordic Council of Ministers Workshop on evaluations of fish stocks (Lassen, H. 1999. Ed.);
- The EVARES project (2003, EVARES - FISH/2001/02 - Lot 1) on evaluation of research surveys in relation to management advice;
- ICES (WKSCFMD) (2004), which reviewed analytical variance estimators, bootstrapping and modelling approaches;
- ICES Workshop on Survey Design and Data Analysis (WKSAD) (ICES 2005/B:07), which considered the effect of spatial structure of the population and provides information on geo-statistical models.

In addition many individual scientists have provided input to the issue of estimating survey index variance (e.g. Pennington (1983) on the use of the Delta distribution where zero values are treated separately and positive values are assumed to follow a log-normal distribution and Petitgas (1993) on a geostatistical approach).

Bootstrapping has been widely used within fisheries in recent years (see e.g. O'Brien et al 2001a, 2001b, Simmonds et al 2001). It is relatively easy to explain, and does not have many assumptions. The lack of assumptions on spatial distribution also suggests it will be robust to changes in spatial distribution from year to year.

For many surveys the sampling design operates with many strata and thus few observations within each stratum. This is creating complications when trying to estimate variance, because the bootstrap sampling will underestimate the within stratum variance - if the stratum includes only two samples the variance is underestimated by a factor of two. Ideally each stratum should include more than 20 samples. When there is only one sample in a stratum, analytical estimations will not be possible and bootstrapping will of course give a variance estimate of zero. Dealing with strata with few hauls is one of the main challenges in the estimating variance of survey indices.

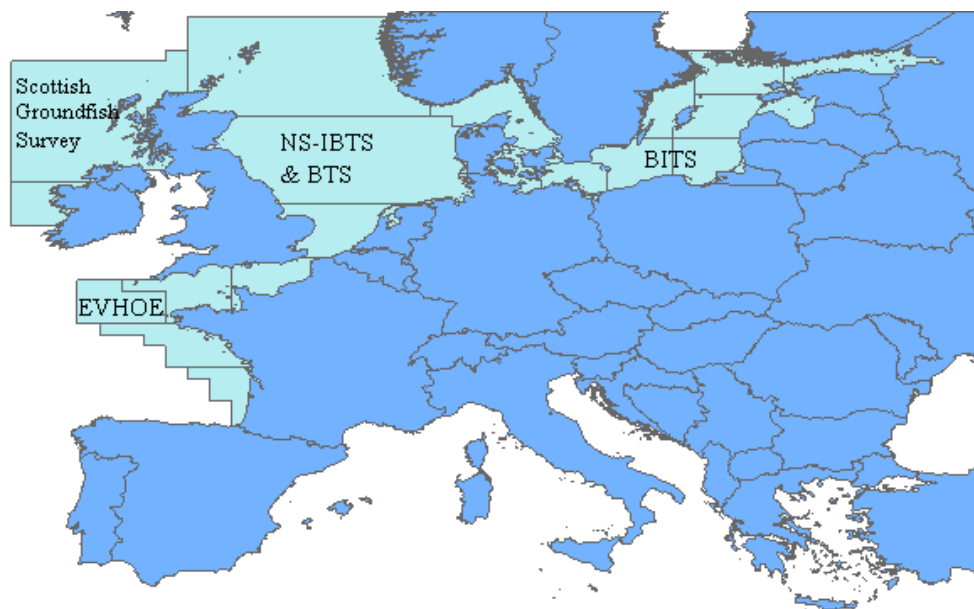


Figure 1. Surveys in ICES area for which data are stored in DATRAS

2 Description of surveys stored in DATRAS

2.1 Data stored

Table 3 shows which countries collect abundance trawl survey data and which submit the data to DATRAS. Spain, Portugal and Ireland will start submitting their data to ICES in 2007.

Table 1. Gear and area combinations against the countries that conduct the surveys. S indicates the country participates in the survey, D indicates the data are in the database.

COUNTRY	BOTTOM TRAWL				BEAM TRAWL
	BALTIC	NORTH SEA	WESTERN AREAS (SCOTTISH GROUND FISH SURVEYS)	SOUTHERN AREAS (EVHOE)	NORTH SEA
Belgium					S
Denmark	S,D	S,D			
England		S,D	S		S
Estonia	S,D				
Finland	S,D				
France		S,D	S,D	S,D	
Germany	S,D	S,D			S
Ireland			S		
Latvia	S,D				
Lithuania	S,D				
Netherlands		S,D			S,D
Norway		S,D			
Poland	S,D				
Portugal				S	
Russia	S,D				
Scotland		S,D	S,D		
Spain			S	S	
Sweden	S,D	S,D			

Table 2 shows the periods for which data are stored in DATRAS for the various surveys. Some of the surveys started as country-performed survey and were later standardized into a common international coordinated survey with e.g. one type of gear. It will therefore only make sense to make a survey performing analysis on the years where the survey have been standardized into one common survey with one gear type. The year ranges in which this has been the case can be seen in Table 3.

Table 2. Year ranges of data in DATRAS and to be used in a variance analysis.

SURVEY	QUARTER	YEAR RANGE IN	YEAR RANGE TO USE FOR ANALYSIS
		DATRAS DATABASE	
Scottish Groundfish Survey	1	1985–2006	1985–2006
	4	1990–2005	1990–2005
BITS	1	1991–2006	2001–2006
	4	1991–2005	2001–2005
BTS-Tridens	3	1985–2006	2003–2006
BTS-Isis	3	1985–2006	2003–2006
EVHOE	4	1998–2003	1998–2003
NS-IBTS	1	1965–2006	1983–2006
	2	1991–1997	1991–1997
	3	1991–2005	1991–2005

4	1991–1997	1991–1997
---	-----------	-----------

2.2 Survey design

The surveys stored in DATRAS have different designs, varying from fix design with 1–2 hauls at the same position in each e.g. ICES statistical rectangle to random selected hauls in a depth stratified design. The design of each survey is described in the following section.

2.2.1 BITS

Baltic cod stock has been monitored annually in first quarter since 1982, and in fourth quarter since 1991; and both bottom trawl surveys are carried out by most countries surrounding the Baltic (Figure 2). The national research vessels have each surveyed part of the area with some overlap in coverage and applied a depth stratified sampling design. However, different gears and design were applied and, in 1985, ICES established a Study Group on Young Fish Surveys in the Baltic in order to standardize the surveys and based on the Group's recommendations standard trawl gear and survey design were implemented in 2001.

The BITS design is based on an area split based on subdivisions and 20 m depth layers. Each year, the total number of planned stations is allocated to the subdivisions and the depth layers according to agreed procedure. Before the international trawl surveys start, the necessary stations of the depth layers are randomly selected from a list of clear haul data. These stations are a subsample of the possible trawl tracks. If the number of possible tracks is not large enough for a random selection in some strata, fixed stations can be used. However, this situation only occurs in shallow waters in subdivision 28.

The following table summarizes the number of planned stations by subdivision and depth layer for the BITS in spring 2006. In most cases, there are more than 7 stations planned. Only in SD 23 (Øresund) and SD 27 (West of Öland), the numbers of planned stations are less than 5 due to various reasons.

SUBDIVISION	22	23	24	25	26	27	28
Depth layer							
10 – 40	23	3	21	20	9	3	11
41 – 60			14	28	9	2	8
61 – 80			11	27	17	2	15
81 – 100				12	16	3	12
100 - 120					10		
Total	23	3	46	87	61	10	46

The cpue values (catch per hour) are corrected in units of the large standard gear TVL using conversion factors. Sub sample are taken from the different hauls for analysing age, sex, weight and maturity. The individuals for estimating the ALK are sampled from different hauls according to the requirement of the BITS manual. The table below gives the minimum number of otoliths from each length group, which must be cut per country, survey, ICES Subdivision and species based on the length distribution.

LENGTH-CLASS	MINIMUM NUMBER OF AGE READINGS
With probably only one age group (age group 0, 1)	2 to 5

With probably more than on age group	
Portion of the length class less than 5%	10
Portion of the length class more than 5%	20

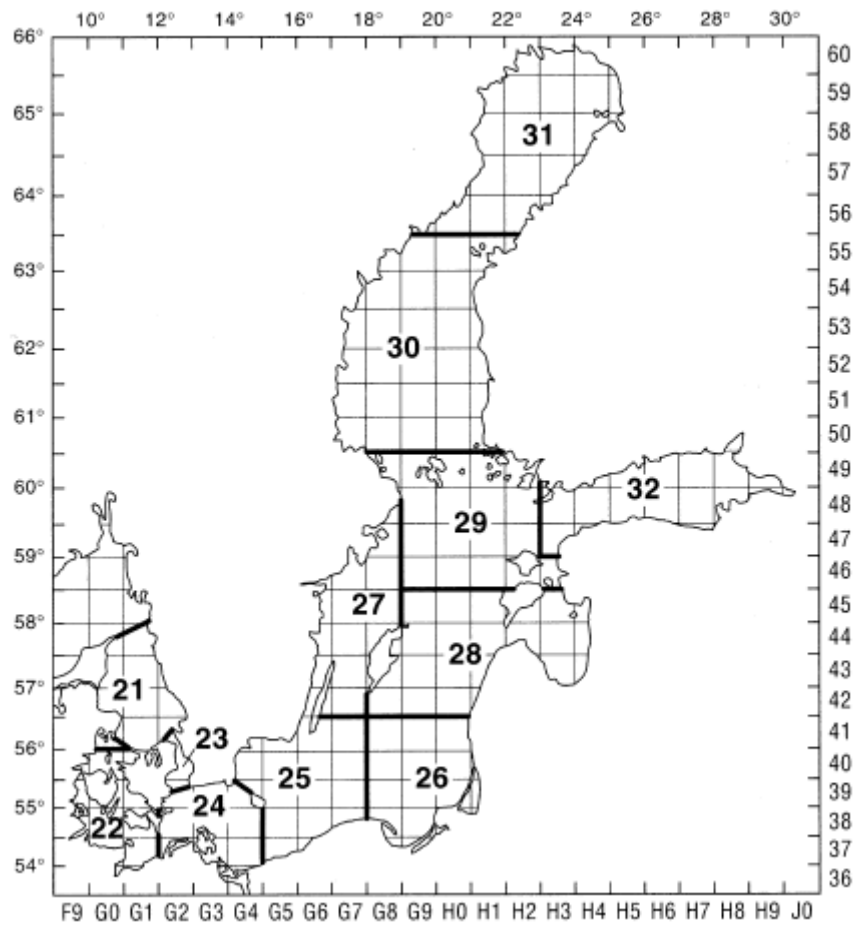


Figure 2. Map of ICES Subdivisions and ICES statistical squares within the Baltic Sea area covered by BITS and IBTS (Kattegat).

2.2.2 NS-IBTS

In the North Sea, the IBTS started in the 1960s as a survey that was directed at juvenile herring and was at that time called the International Young Herring Survey (IYHS).

As it was gradually realized that the survey also yielded valuable information for other fish species, such as cod and haddock, the objectives were broadened and the survey was renamed into the International Young Fish Survey (IYFS). Besides the IYFS, which was carried out in the first quarter, a number of national surveys developed in the 1970s and 1980s that were mainly carried out in the third quarter.

In 1990, ICES decided to combine the international and the national surveys into the IBTS. The IBTS has been carried out twice per year (first and third quarter) since 1997, and on a quarterly basis in the period 1991–1996.

The stratification of the survey grid has always been based on ICES statistical rectangles (one degree longitude x 0.5 degree latitude ~ 30 x 30 nautical miles). Each rectangle is usually fished by the ships of two different countries, so that at least two hauls are made per rectangle. The ICES-rectangles are used to evenly distribute the hauls over the whole North Sea (Figure 3).

Most statistical rectangles contain a number of possible tows that are deemed free of obstructions, and vessels are free to choose any of these positions in the rectangles that they are surveying. In some rectangles, sampling may be further stratified due to significant changes in seabed depth which may, in turn, cause variations in the fish population.

In rectangles or strata that are to be sampled more than once by the same vessel, is the IBTS manual recommends that valid hauls are separated by at least one day or by at least 10 miles wherever possible. Tows in adjacent rectangles should also be separated by at least 10 miles.

All countries except England (Q3) and Norway (Q1 and Q3) select the hauling position randomly from a list of clear haul positions. England and Norway use the same fixed hauls every year.

To obtain the length distribution, the catch is sorted into species or species/sex. Where the numbers of individuals are too large for all to be measured (due to time constraints etc), a representative subsample is selected of at least 75 fish (although sampling a very limited length range could be adequately achieved with less). In the event that a truly representative subsample cannot be selected, it will be necessary to further sort the species into two or more size grades or categories (Manual for the International Bottom Trawl Surveys, version VII).

Otolith samples are collected within 9 specified Roundfish areas as illustrated in figure 3. For all species, the same areas are used. For the target species, the following minimum sampling levels are tried to be obtained for each sampling area:

herring:	8 otoliths per 1/2 cm group
sprat:	16 otoliths per 1/2 cm group 8.0–11.0 cm
	12 otoliths per 1/2 cm group >11.0 cm
mackerel:	8 otoliths per 1 cm group
cod:	8 otoliths per 1 cm group
haddock:	8 otoliths per 1 cm group
whiting:	8 otoliths per 1 cm group
Norway pout:	8 otoliths per 1 cm group
saithe:	8 otoliths per 1 cm group

For the smallest size groups, that presumably contain only one age group, the number of otoliths per length class can be reduced. Inversely, more otoliths per length are required for the larger length classes.

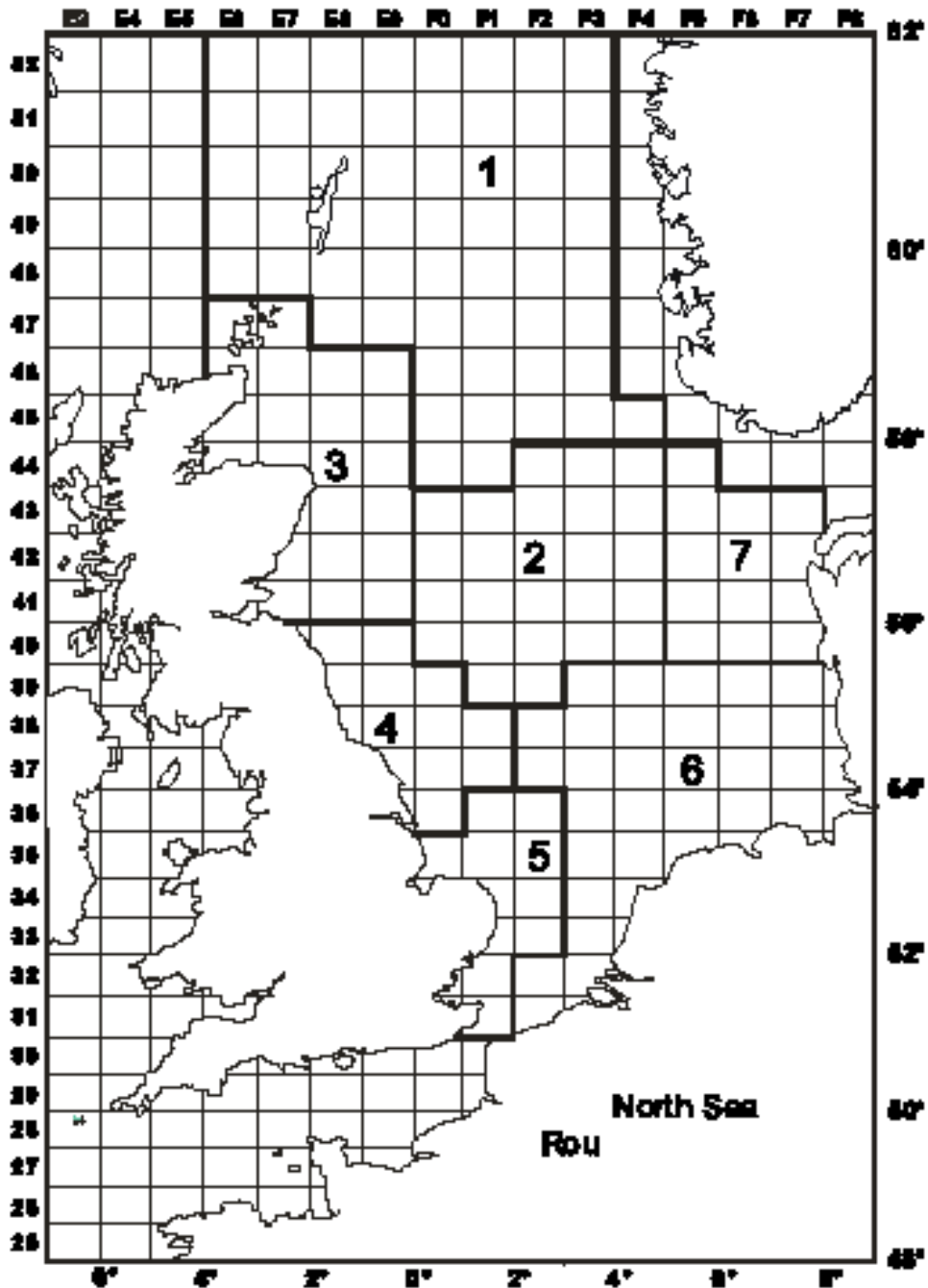


Figure 3. IBTS survey area with statistical rectangles and roundfish areas used for stratification of age-length keys

2.2.3 BTS

The Netherlands BTS was initiated in 1985 to estimate the abundance of the dominant age groups of plaice and sole including pre-recruits. Initially the survey was only carried out in the southeastern North Sea (ICES Area IV) using RV "Isis" equipped with a pair of 8 m beam trawls.

The survey was designed to take between one and three hauls per ICES rectangle. The stations are allocated over the fishable area of the rectangle on a "pseudo-random" basis to ensure that there is a reasonable spread within each rectangle. No attempt is made to return to the same tow positions each year.

In 1995, the survey was expanded into the central and northern part of the North Sea using RV "Tridens". Sampling strategy is similar, but only one haul per rectangle is taken, preferably close to the centre of the rectangle (figure 4).

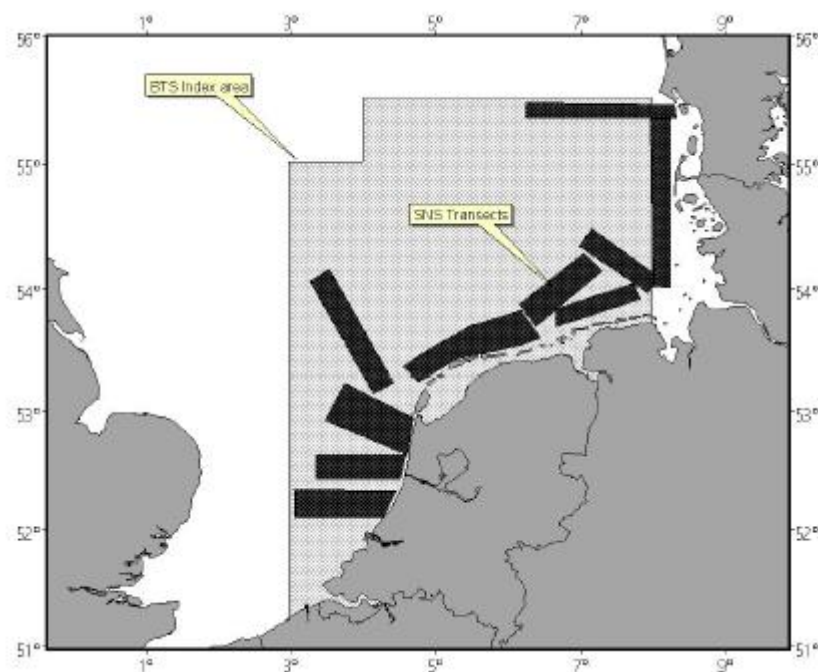


Figure 4. Locations of the BTS Index area and SNS transects in the North Sea. In this report only BTS is relevant.

2.2.4 EVHOE

Between 1987 and 1996, the EVHOE survey was conducted in the Bay of Biscay on an annual basis with the exception of 1993 and 1996. It was conducted in the third or fourth quarter except in 1991, when it took place in May. In 1988, two surveys were conducted: one in May the other in October.

The Celtic Sea was surveyed from 1990 to 1994, but the sampling was restricted to a small geographical area.

Since 1997, the survey has covered all the Celtic Sea and Bay of Biscay during the 4th quarter for 40 to 45 days depending on year and availability of ship. The survey has the following main objectives:

- construction of time-series of abundance indices for all the commercial species in the Bay of Biscay and the Celtic Sea with an emphasis on the yearly assessed species where abundance indices at age are computed.
- to describe the spatial distribution of the species and to study their interannual variations.
- to estimate and/or update biological parameters (growth, sexual maturity, sex ratio...)

The stratification scheme adopted defines 6 depth strata within a geographic stratification that separates the Bay of Biscay in 2 areas and the Celtic Sea in 3 areas (figure 5).

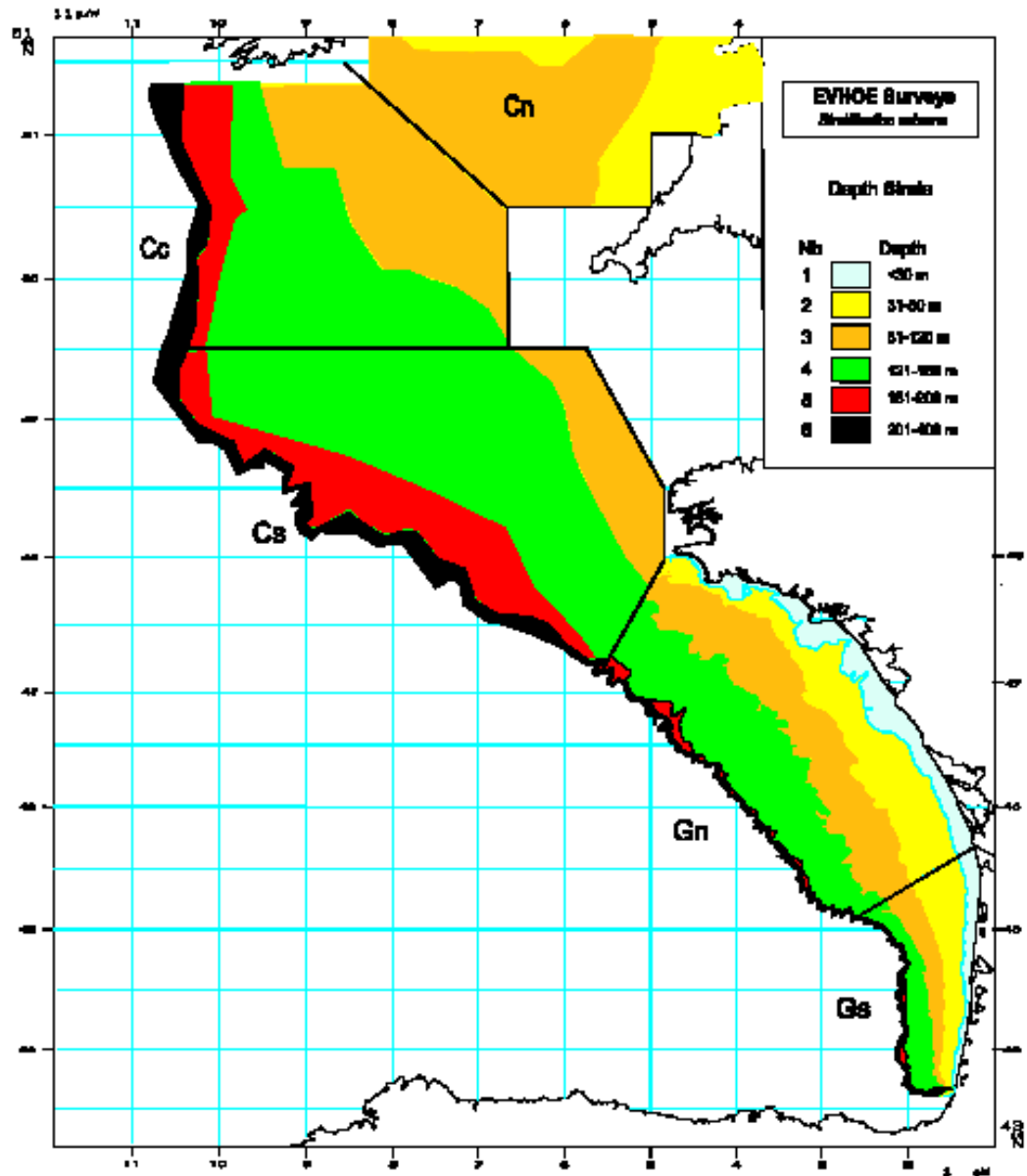


Figure 5. Area covered and stratification used in the EVHOE surveys.

The sampling strategy is a stratified random allocation: the number of sets per stratum optimized by a Neyman allocation on numbers variance averaged on the 4 most important commercial species (hake, monkfishes and megrim) leaving at least two stations per stratum. 140 sets are planned every year. This number of sets is adjusted according to the time at sea available.

The catch is sorted by species, counted and weighted. In the case of a huge catch of one dominant species, only a fraction of the catch is sorted.

Biological parameters (length, weight, status of maturity among others) and hard structures (otoliths and *illicia*) are collected. The specification of the sampling level of otoliths and *illicia* is described in table 3 (see also Manual for the International Bottom Trawl Surveys in the Western and Southern Areas, version 2).

Table 3. Specification of the sampling level of otoliths and *illicia* by country.

Species	Country	Otoliths, <i>illicia</i> or spines
<i>Gadus morhua</i>	FR,SC IR	All individuals Juveniles: 5/cm/ICES Div, Adults: 10/cm/ICES Div
<i>Lepidorhombus whiffiagonis</i>	FR,SC IR SP P	8/cm/sex/area Juveniles: 5/cm/ICES Div, Adults: 10/cm/ICES DIV 20/cm 3/cm/sex/area
<i>Lepidorhombus boscii</i>	FR,SC SP P	No 10/cm 3/cm/sex/area
<i>Lophius piscatorius</i>	All (IR from 2002) SC	All individuals (illicium) None
<i>Lophius budegassa</i>	FR,SP,P (IR from 2002)	All individuals (illicium & 2 nd fin ray) All individuals (Illicium only)
<i>Melanogrammus aeglefinus</i>	IR,SC	1/cm/ICES Rectangle
<i>Merlangius merlangus</i>	FR IR,SC	Proportional 1/10/cm/sex/area 1/cm/ICES DIV.
<i>Merluccius merluccius</i>	FR IR SP P	8/cm/sex/area 8/cm/undet./area Juveniles: 5/cm/ICES Div, Adults: 10/cm/ICES Div < 17 cm - 1 each 3 individuals > 17 cm - all individuals 3/cm/sex/area >40cm - all individuals 10/cm/undet./area
<i>Micromessistius poutassou</i>	SP P	10/haul (random) 10/cm/sex/area
<i>Microstomus kitt</i>	IR	5/cm/ICES Div
<i>Molva molva</i>	FR	All individuals
<i>Pleuronectes platessa</i>	IR	1/cm/ICES Rectangle
<i>Pollachius pollachius</i>	FR	All individuals
<i>Pollachius virens</i>	IR,SC	All individuals
<i>Scomber scombrus</i>	SP,SC P	10/cm/area 5/cm/sex/area
<i>Scomber japonicus</i>	P	5/cm/sex/area
<i>Solea vulgaris</i>	FR IR	All individuals Juveniles: 5/cm/ICES Div, Adults: 10/cm/ICES Div
<i>Squalus acanthias</i>	IR	Spines All individuals
<i>Trachurus trachurus</i>	SP P	15/cm 5/cm/sex/area 10/cm/undet./area

2.2.5 Scottish Groundfish Survey

The Quarter 1 (March) Scottish Groundfish survey started in 1981, and was initially intended to cover the fishing grounds on the continental shelf to the west of Scotland. In 1996, the survey area was extended to include the northern Irish Sea (Figure 6).

The survey covers Division VIa and extends into the northern part of the Irish Sea and NW of Ireland. The depth range covered has been 20 to 500 m since 2000. The target species are cod, haddock, whiting, saithe and herring and age frequencies are constructed for these species. All other fish species encountered are also sampled for at least length frequencies.

Indices of abundance-at-age are calculated for all the target species and these data are used at the Northern Shelf Assessment Working Group and also made available for the Herring Assessment Working Group.

The Scottish Mackerel recruit Quarter 4 survey began in 1990, and has a depth range of 20 - 500 m. The survey extended to the area west of the British Isles between 56 and 61 N, and bounded by the 200 m depth-contour and the coast. Initially the survey area did not include the area of the Minch and the north channel of the Irish Sea, but gradually the spatial coverage altered so that now it mimics the Quarter 1 survey.

The target species have now been extended to include cod, haddock, whiting, saithe and herring as well as the original target species mackerel.

The Scottish West Coast Surveys use a ICES rectangle based sampling strategy similar to that used in the North Sea. Trawl stations are selected at one tow per rectangle based on a library of clear tows. There is no explicit return to the same trawling position every year, although this is generally the case.

Since 1999, the potential for using a depth rather than rectangle based stratification has been under investigation. To this end, and where possible, those rectangles which display substantial internal depth variation have been sampled twice at different depths. The recent inclusion of samples collected between 200 and 500 m would suggest that depth stratification should be initiated as soon as possible.

The sampling scheme for length and age distribution is described under the EVHOE survey.

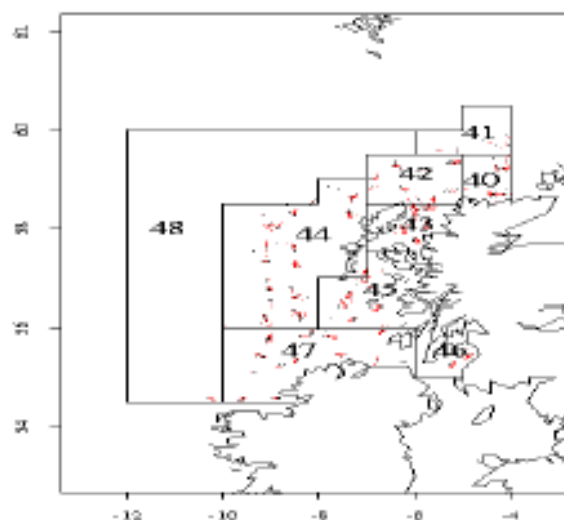


Figure 6. Scottish demersal sampling areas

3 Calculation of indices

DATRAS delivers fishery-independent abundance indices to be used for tuning in stock assessments. Table 4 gives an overview of the different species, where they are caught and which working groups use their indices.

Table 4: Overview of the species for which the surveys provide indices

SURVEY	AREA	SPECIES	ASSESSMENT WG
BTS	North Sea, Celtic Sea and Irish Sea	plaice	northern shelf WG (VIIa), southern shelf (VIIe, VIIe, VIIg,f), and North Sea and Skagerrak WG
		sole	northern shelf WG (VIIa), southern shelf (VIIe, VIIe, VIIg,f), and North Sea and Skagerrak WG
IBTS Q1	North Sea and Skagerrak	cod	North Sea and Skagerrak WG
		whiting	North Sea and Skagerrak WG
		saithe	North Sea and Skagerrak WG
		haddock	North Sea and Skagerrak WG
		Norway pout	North Sea and Skagerrak WG
		herring	Herring Assessment south of 56 WG
		sprat	Herring Assessment south of 56 WG
IBTS Q3	North Sea and Skagerrak	Cod	North Sea and Skagerrak WG (Scottish & English separately)
		Whiting	North Sea and Skagerrak WG
		Haddock	North Sea and Skagerrak WG (English data only)
		Norway pout	North Sea and Skagerrak WG (Scottish & English separately)
Scottish grundfish survey Q1	VIa	Cod	Northern Shelf WG
		Haddock	Northern Shelf WG
		Whiting	Northern Shelf WG
		Norway pout	Northern Shelf WG
		Saithe	Northern Shelf WG
Scottish grundfish survey Q4	Via	Cod	Northern Shelf WG
		Haddock	Northern Shelf WG
		Whiting	Northern Shelf WG
		Norway pout	Northern Shelf WG
		Saithe	Northern Shelf WG
EVHOE	VII	Whiting	WG on assessment of Southern Shelf demersal stock
		Cod	WG on assessment of Southern Shelf demersal stock
		Hake	WG on assessment of Southern stocks of hake, monk and megrim
		Monkfish	WG on assessment of Southern stocks of hake, monk and megrim
		Anglerfish	WG on assessment of Southern stocks of hake, monk and megrim
		Megrim	WG on assessment of Southern stocks of hake, monk and megrim
BITS	Baltic Sea	cod	Baltic WG
	SubDiv 24	Flounder	Baltic WG

For all the surveys, there is one standard routine in DATRAS to allocate the ALK key to the length based cpue data.

The cpue per age (a), length and haul (H) is calculated as the fraction of the age distribution:

$$CPUE_{H,a,l} = \frac{CPUE_{H,l} * ALK_{a,l}}{ALK_l}$$

There are three possibilities for obtaining age information for a length class if an age distribution is missing for that length class:

If length is less than a minimum length predefined in DATRAS, the age is set to age 1 in first quarter and 0 in all other quarters.

If length is between minimum length and maximum length (also predefined in DATRAS), then age is set to the nearest ALK either at a length class before or at a length class after the one which misses an ALK. If there is one below and one after the length class at equal distance in length, a mean is taken.

If the length is larger than max length, the age is set to the plus group.

After the age distribution is allocated to the length distribution, the age based indices are calculated. The indices are calculated differently for each survey. Below, the calculation for each survey is described.

3.1 BITS

The BITS survey provides abundance indices for cod and flounder. There are two cod indices: one for the eastern Baltic, subdivisions 25+26+28, and one for the western Baltic, subdivisions 22+24. At present, only indices for the eastern cod are calculated by DATRAS.

cpue for cod at length per haul from before 2001 is standardized to the standard TV3 trawl used by all vessels since 2001 by multiplying with a conversion factor (ISDBITS. 2001). Then the mean age at length per depth stratum and subdivision are calculated and weighted with the surface area (m²) of the stratum. From these means, the mean catch-at-age per subdivision and then the mean catch per index area are calculated.

For flounder, the index is only calculated for subdivision 25, and no conversion factor is applied to data before 2001. Otherwise the calculation is identical to the cod calculation.

Calculations:

- cpue per length (l) and haul (H) are multiplied with the conversion factor (conf) for the gear to give cpue adjusted for gear performances (conf)
- $confCPUE_{H,l} = CPUE_{H,l} * conf$
- Age is allocated to the length distribution as described above.
- Number per length (l) (1 cm group) per haul is summed by year, quarter, subdivision and depth stratum (DS) and divided with total hauls in the depth stratum.

$$mCPUE_{DS,a,l} = \frac{\sum_{SD} confCPUE_{H,a,l}}{\sum_{SD} H}$$

- Depth stratum '0-19' is deselected
- Mean cpue per length and stratum is multiplied by the area of the depth stratum (km²)
- The multiplied values are summed over depth stratum and index area (IA)
- The summed values are divided with sum of all stratum areas

$$mCPUE_{IA,a,l} = \frac{\sum_{IA} mCPUE_{DS,a,l} * A_{DS}}{\sum_{IA} A_{DS}}$$

- The final indices by age are calculated by taking the sum of the length classes for the given age within the index area:

$$mCPUE_{IA,a} = \sum_l mCPUE_{IA,a,l}$$

3.2 NS-IBTS indices calculation

In IBTS North Sea, the indices are calculated per index area, which are specific for each species. The indices are calculated as mean at age per statistical rectangle and then as a mean of the statistical rectangles over the index area. Some statistical rectangles are reduced in size due to land or very shallow water. For herring, sprat and saithe, the mean cpue at age are weighted with the percent covered with water depths between 10 m and 200 m for these statistical rectangles.

- Number per length (l) and age (a) (1 cm group and semi cm for herring and sprat) per haul is summed by year, quarter, statistical rectangle (ST) and divided with total hauls in the statistical rectangle.

$$mCPUE_{ST,a,l} = \frac{\sum_{ST} CPUE_{H,a,l}}{\sum_{ST} H}$$

- Mean number by index area, sum of mean cpue in all fished rectangles in index area (IA) divided by number of fished rectangles in index area:

$$mCPUE_{IA,a,l} = \frac{\sum_{IA} CPUE_{ST,a,l}}{\sum_{IA} ST}$$

- The final indices by age are calculated by taking the sum of the length classes for a given age within the index area:

$$mCPUE_{IA,a} = \sum_l mCPUE_{IA,a,l}$$

3.3 BTS indices

The BTS survey consists of several surveys that are not totally integrated and cover different areas. Therefore, indices are calculated separately for each survey/area combination.

The DATRAS database only includes the Dutch BTS surveys: BTS-Tridens and BTS-Isis. Both surveys produce abundance indices for plaice and sole. The indices are calculated the same way as for NS-IBTS.

3.4 EVHOE

The indices in the EVHOE survey are calculated similar to the BITS indices as both surveys are depth stratified; however, EVHOE does not use conversion factors.

Calculations:

- Age is allocated to the length distribution as described in the general introduction.
- Number per length (l) (1 cm group) per haul is summed by year, quarter, EVHOE area (EA) and depth stratum (DS) and divided with total hauls in the depth stratum.

$$mCPUE_{DS,a,l} = \frac{\sum_{DS} CPUE_{H,a,l}}{\sum_{DS} H}$$

- Mean cpue per length and stratum is multiplied by the area of the depth stratum (km^2)
- The multiplied values are summed over depth stratum and index area (IA)
- The summed values are divided with sum of all stratum areas

$$mCPUE_{IA,a,l} = \frac{\sum_{IA} mCPUE_{DS,a,l} * A_{DS}}{\sum_{IA} A_{DS}}$$

- The final indices by age are calculated by taking the sum of the length classes for the given age within the index area:

$$mCPUE_{IA,a} = \sum_l mCPUE_{IA,a,l}$$

3.5 Scottish Western IBTS

The indices for the Scottish Western IBTS are calculated slightly differently from the indices in other surveys in DATRAS. Whereas all the other surveys are weighted by either area size or number of strata, the Scottish indices are weighted with the number of hauls samples in each sampling area.

- For each age, length and subarea, sum across valid hauls and standardize to per 10 hours

$$CPUE_{SA,a,l} = \frac{\sum_{h=1}^{nhauls_{SA}} N_{a,l,h} \times 10}{nhauls_{SA}}$$

- Take a weighted sum over areas

$$Index_{a,l} = \frac{\sum_{SA=1}^{nareas} (CPUE_{SA,a,l} \times nhauls_{SA})}{\sum_{SA=1}^{nareas} nhauls_{SA}}$$

- Sum across length classes

$$Index_a = \sum_{l=1}^{nlength} Index_{SA,a,l}$$

4 Choice of method for uncertainty estimation

We considered the technical advantages and disadvantages of a range of methods, including their suitability for implementation in DATRAS, making use of work reported in ICES WKSCFMD (2004), which reviewed analytical variance estimators,

bootstrapping and modelling approaches and ICES WKSAD (2005), which considered the effect of spatial structure of the population and provides information on geo-statistical models.

Bootstrapping has technical advantages over analytical calculations because asymmetric distributions do not cause problems when calculating a confidence interval and covariance between ages is part of the output. A further issue with the usual analytical calculations is that they are based on the assumption of random (or stratified random) sampling, so they are not strictly valid for fixed stations designs.

The usefulness of geo-statistical methods has been demonstrated for individual analyses. But overall, the bootstrap approach was considered most appropriate for routine analysis in DATRAS. It is technically adequate, relatively straightforward to implement and easy to explain. It represents the calculations currently used to produce survey indices and a consistent definition of bootstrap sampling is possible across all the surveys in the system.

5 How to bootstrap in DATRAS

Within each survey for each species the calculation covers a so-called index area. This is the area within which hauls are considered and the average over the length compositions is done. In several cases the index area and the entire survey area are identical. The age-length keys are aggregated by a set of "otolith areas" without restricting the data to the index area.

The abundance estimator works in six steps: 1) First, the length composition by subareas (e.g. rectangles) and haul is calculated, 2) In parallel, ALK's are aggregated on a separate set of "otolith areas" (e.g. roundfish areas, the phrase "otolith area" is used by the North Sea beam trawl survey), 3) the length distributions calculated per subarea and haul in step 1 are raised to age compositions using ALK's found in step 2, 4) the mean number per age and length in a subarea are calculated and 5) these age/length compositions are averaged over an index area (e.g. the entire North Sea), i.e. ignoring length compositions that refer to subareas outside the index area. This later step means that some subareas that are fished will not be considered in the index calculation because the rectangles are not part of the index area. 6) Finally the age/length composition in the index area is summed to mean cpue per age.

This suggests that the haul information in a survey shall be considered as two components: the length frequency distribution of the catch and the age-length keys.

These components are bootstrapped independently. The bootstrapping approach chosen is the *naïve* approach (Lehtonen, R., and Pahkinen, E. 2004). The haul bootstrap unit are the entire length distribution of a haul (not bootstrapping the individual length groups) thereby maintaining covariance between the length groups within a haul. The ALKs are bootstrapped as individual aged fish length group by length group.

Bootstrapping the length frequency distribution of the catch by haul will consider how the hauls are distributed within an area. A certain amount of pooling among subareas is need in some of the surveys where there are too few observations per subareas. There are therefore two steps in this bootstrap procedure, 1) selecting a haul and 2) allocating this haul to a subarea.

Figure 7 below shows the bootstrapping data flow; the flow is described in the following section.

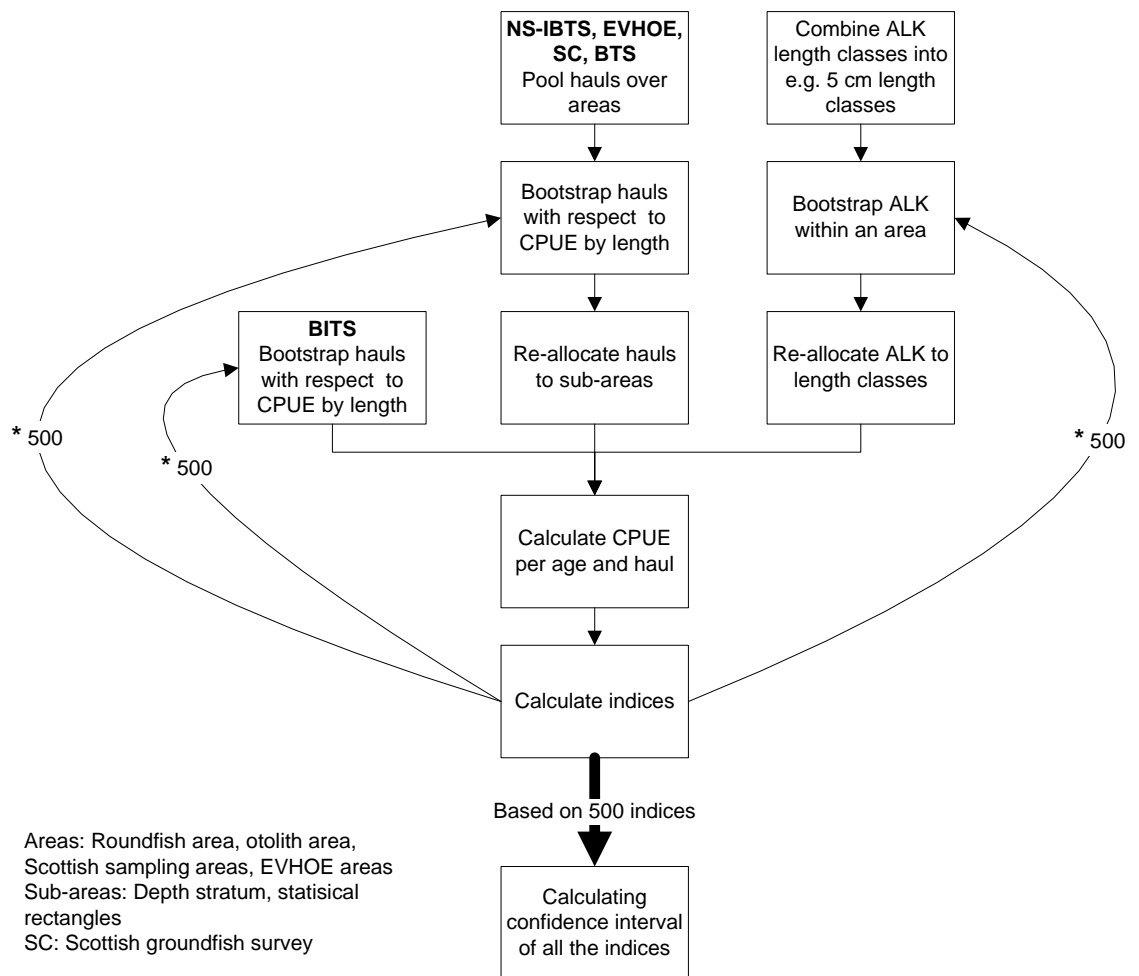


Figure 7. Bootstrapping data flow

5.1 Length distribution bootstrapping

Only hauls within the index area for a given species are used in the bootstrapping.

The number of hauls within each subarea was kept when bootstrapping. For those surveys where the index calculation is done by large strata with many hauls per stratum (generally 5 to 20) hauls, we draw the bootstrap from the hauls in that stratum and the approach is straightforward. For surveys like the North Sea IBTS and BTS, which operate with rectangles as strata (subareas) for index calculation, too few hauls was available within a stratum (rectangle) and the bootstrap sample was drawn from a larger pool of hauls i.e. by RoundFish Area (RF) (Table 5).

The abundance index calculation requires that each haul is assigned to a subarea (rectangle) and the bootstrap therefore needs a second step to allocate the selected haul to a subarea.

Table 5. Overview of sampling strata and pooling strata

SURVEY	SPECIES	SAMPLING STRATA	POOLING STRATA
NS-IBTS	All	Statistical rectangle	Roundfish area combined with index area
BITS	All	SubDiv + depth strata	SubDiv + depth strata
EVHOE	All	EVHOE area + depth strata	Combined depth strata of same depth within Celtic Sea or Bay of Biscay
BTS	All	Statistical rectangle	Otolith areas
Scottish groundfish survey	Cod, whiting, haddock, monkfish	Scottish demersal sampling areas	Scottish demersal sampling areas

5.2 Bootstrap age–length keys by length class

The Age-Length-Keys for all the surveys are an aggregation of individual samples from a haul combined over a larger area (Table 6).

It is assumed for most surveys that the individuals for the ALK are randomly taken from all parts of the area. Furthermore, this sampling regime for ALK data assumes that the age distribution of a length class does not significantly differ within the different parts of the sampling area.

Table 6. Overview of ALK sampling areas

Survey	Area of ALK
BITS	ICES subdivision
IBTS	Roundfish area
BTS	Otolith areas
EVHOE	EVHOE areas
Scottish groundfish survey	Demersal sampling area

The sampled numbers of fish per length class are used as basis for bootstrapping the ALK's. Analyses of the data which are available in DATRAS have shown that in many cases the number of aged fish per length class is significantly lower than the required number for bootstrapping.

We therefore investigated, if the age distribution could be pooled over larger length classes than the sample length class. The conclusion was, that for the smaller length groups i.e. where the growth curve is steep this pooling bias the variance measure (upwards) for the younger age groups. As the DATRAS surveys are focused on recruitment and young age groups estimates we decided avoid this. We were looking for an approach for only start pooling for larger fish (actually where the need for pooling is more pronounced) but then these groups only present small numbers of fish and is not central for the application in the assessment. A pooling scheme with 3–5 length groups would need to be species and most likely also survey dependent. Simulations indicate that the advantage of pooling is offset by the added bias for the young age groups and that we do not really gain much precision in the variance for the older age groups. The sampled numbers of fish per length class was therefore bootstrap though number of samples is not sufficient in all length classes.

5.3 Number of bootstraps

The length cpue and the ALK data were bootstrapped 500 times.

5.4 The index calculations

For each of the 500 bootstrap samples, the normal procedure for applying ALK to length data and the usual index calculations were carried out.

5.5 Reference Uncertainty for Abundance Indices

In order to compare results between years within a survey or between surveys the graphs are provided with a reference level. This level is chosen to correspond approximately to a 20% CV level. Using the normal distribution $1.35 * CV = \text{Range/Median}$. We show a reference level of 0.25 for the Range/50% percentile uncertainty indicator.

5.6 Presentation of results

The bootstrap results are presented as the Range ($=[Q25\%;Q75\%]$) divided with the median,

The range in the bootstrapped indices is larger for the small age groups compare to the old age groups because the index is a skewed distribution (approx log-normal). The range increases with the mode. In this case the smaller ages have a larger mode and thereby also a larger range than the larger ages. By dividing with the median, the range is weighted with the median and the effect of the mode is removed. In conformity with using the coefficient of variation ($CV = \text{std. Dev./mean} * 100\%$). DATRAS uses the range/50% percentile.

As examples the Figures below (Figure 8) shows the Range/median for Cod, haddock and herring in NS-IBTS survey in quarter 1. The relative variance is smaller for the younger age groups and larger for the older age groups.

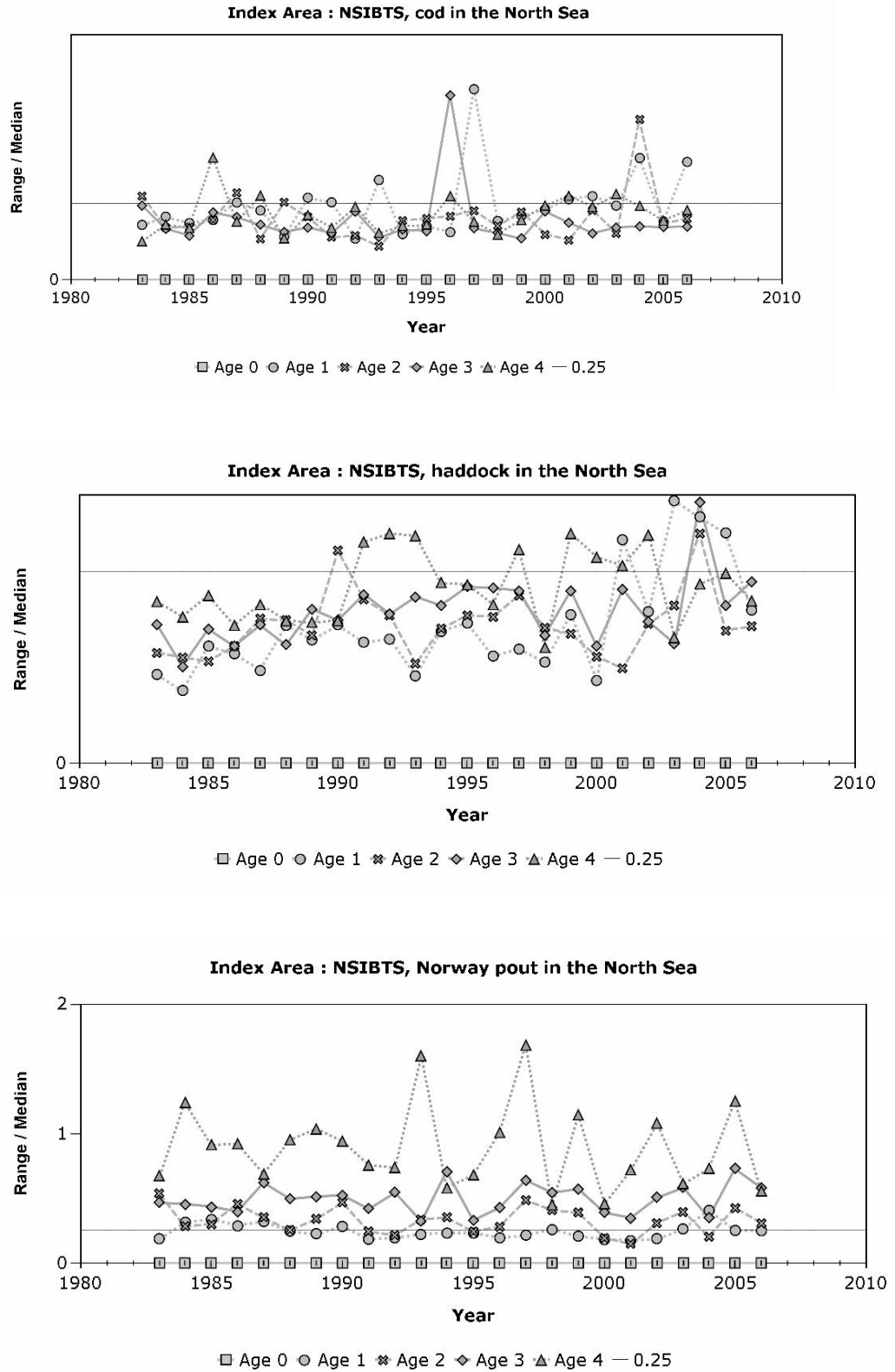


Figure 8. Range/median for Cod, haddock and herring in NS-IBTS survey in quarter 1

5.7 Conclusion

Annex I shows the graphs (Range/Median) for all the indices produced in DATRAS, including the IBTS quarter 2 and 4 survey. The surveys in these two quarters are no longer carried out but are included in this report to give a complete overview of the data in the system.

Table 7 summarizes all the graphs and shows roughly which uncertainty indicator are above, below or on the 0.25 reference line. The uncertainty indicators for NS-IBTS are generally above the reference line, the only exceptions are cod, haddock and whiting. In the BITS survey, cod are varying around the reference line, while the flounder uncertainty indicators are above. For the indices in the two BTS surveys only the BTS-Tridens plaice uncertainty indicators are on the line the other uncertainty indicators are above.

All the uncertainty indicators for the two surveys, French EVHOE and Scottish groundfish survey are either on the reference line or below.

Table 7. The table shows whatever the uncertainty indicator is above, below or on the reference line.

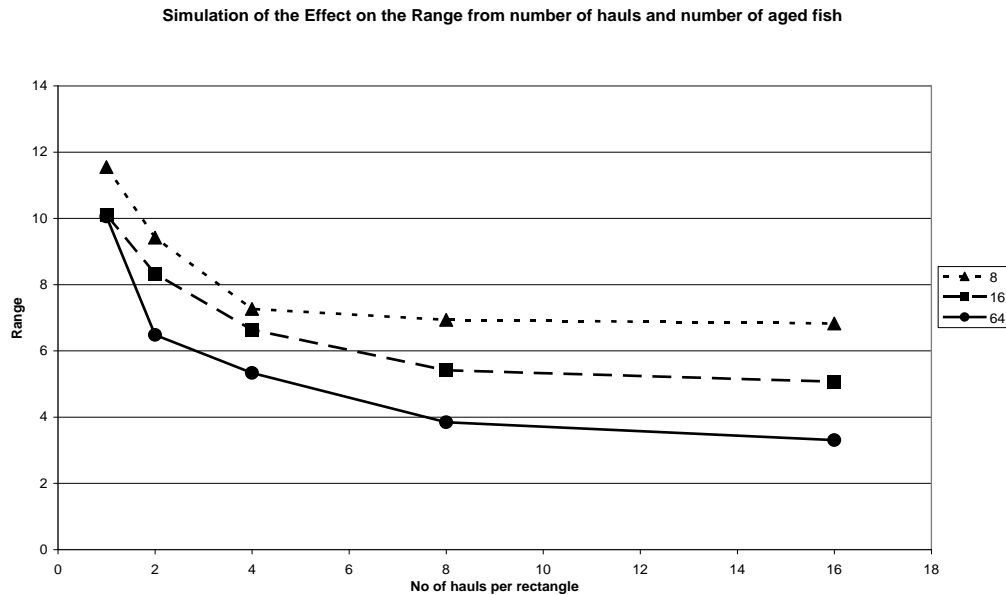
INDICES	ABOVE THE REFERENCE LEVEL	AROUND THE REFERENCE LEVEL	BELOW REFERENCE LEVEL
BITS Quarter 1			
COD_EastBaltic		X	
Flounder_SubDiv24	X		
BITS Quarter 4			
COD_EastBaltic		X	
Flounder_SubDiv24	X		
BTS_Isis Quarter 3			
Plaice_NorthSea	X		
Sole_NorthSea	X		
BTS_Tridentis Quarter 3			
Plaice_NorthSea		X	
Sole_NorthSea	X		
NS-IBTS Quarter 1			
Cod_Kattegat	X		
Cod_NortSea			X
Haddock_NorthSea			X
Herring_NorthSea	X		
Norway Pout_NorthSea	X		
Mackerel_NorthSea	X		
Plaice_IIIa	X		
Sprat_NorthSea	X		
Spart_IIIa	X		
Whiting_NorthSea		X	
NS-IBTS Quarter 3			
Cod_Kattegat	X		
Cod_NortSea		X	
Haddock_NorthSea			X
Herring_NorthSea	X		

INDICES	ABOVE THE REFERENCE LEVEL	AROUND THE REFERENCE LEVEL	BELOW REFERENCE LEVEL
Norway Pout_NorthSea	X		
Mackerel_NorthSea	X		
Plaice_IIIa	X		
Sprat_NorthSea	X		
Spart_IIIa	X		
Whiting_NorthSea		X	
EVHOE Quarter 4			
Hake_BayBiscay_CelticSea			X
Cod_CelticSea			X
Monk fish_BayBiscay_CelticSea			X
Megrim_BayBiscay_CelticSea			X
Anglerfish_BayBiscay_CelticSea			X
Whiting_BayBiscay		X	
Scottish groundfish survey Quarter 1			
Saithe_VIa			X
Norway pout_VIa			X
Haddock_VIa			X
Whiting_VIa			X
Saithe_VIa		X	
Scottish groundfish survey Quarter 4			
Cod_VIa			X
Norway Pout_VIa		X	
Haddoc_VIa			X
Saithe_VIa		X	
Whiting_VIa			X

6 Effect of a change in sampling intensity

There are many examples that the quality of the survey results can be significantly improved by standardizing gears and other procedures and by choosing an appropriate design (e.g. depth stratification). This present study is restricted to investigate the effect of the number of hauls and the number of aged fished. These are key parameters determining the accuracy obtained with a survey. The number of fish in the length distribution is a third parameter. This parameter is ignored in this analysis as the length composition is usually sampled at a high intensity (it is fairly cheap to measure the length of fish) and that for several surveys all fish on deck are length measured. There are also examples where all fish on deck of a species are aged.

The graph below shows the general effects calculated from a theoretical survey. The main feature is that the increase in accuracy is levelling out for constant age sampling with increasing number of hauls. The same is the result is the number of hauls is kept constant and the number of aged fish is increased.



In order to investigate the effect of changing the sampling intensity the bootstraps were repeated

- 1) Hauls were included in the basic data set by random selection where the probability of including each haul was 80%
- 2) Aged fish were included in the data set by random selection where the probability of including each fish was 80%

6.1 Effect of change in coverage - No of hauls

Table 8 shows the ratio of the range obtained by the 80% haul reduced data set and the full data set.

Table 8. The ratio between the 80% haul reduced data set and the full data set.

YEAR	AGE1	AGE2	AGE3	AGE4
Cod in North Sea (NS_Cod)	0.93	1.09	1.27	1.11
Cod in Kattegat (NS_CodCat)	1.12	1.08	1.13	1.07
Haddock in North Sea (NS_Had)	1.26	1.13	1.10	1.17
Herring in North Sea (NS_Her)	1.09	1.08	1.04	1.16
Norway pout in North Sea (NS_NorPout)	1.17	1.13	1.13	1.08
Sprat in IIIa (NS_SpratIIIa)	1.26	1.06	0.97	1.01
Sprat in IV (NS_SpratIV)	1.22	1.03	1.07	1.05
Whiting in North Sea (NS_Whit)	1.08	1.13	1.05	1.12

The conclusion from this analysis is that a change in the number of hauls of 20% in general will change the accuracy with 10–15%. This applies in both directions both an increase in the number of hauls which increases the accuracy or a decrease which will decrease the accuracy. The examples where the reduction of number of hauls suggest an increase in accuracy – where the ratio given above is less than 1 – are related to cases where the number of fish are small or a single/few large hauls are dominating.

6.2 Effect of change in age sampling - No of aged fish per length class

The table below shows the ratio of the range obtained by the 80% aged fish reduced data set and the full data set

Table 9. The ratio between the 80% age fish reduced data set and the full data set.

YEAR	AGE1	AGE2	AGE3	AGE4
Cod in North Sea (NS_Cod)	0.97	1.00	1.13	1.08
Cod in Kattegat (NS_CodCat)	1.04	1.02	1.11	1.09
Haddock in North Sea (NS_Had)	1.02	1.04	1.12	1.38
Herring in North Sea (NS_Her)	1.00	1.10	0.99	1.19
Norway pout in North Sea (NS_NorPout)	0.97	1.09	1.20	1.22
Sprat in IIIa (NS_SpratIIIa)	1.03	1.01	1.33	3.24
Sprat in IV (NS_SpratIV)	1.02	1.14	1.37	1.49
Whiting in North Sea (NS_Whit)	0.99	1.08	1.10	1.50

The conclusion from this analysis is that a change in the number of aged fish of 20% in general will change the accuracy with a few percent for the younger but that the effect increases with age to 20–40%. However, for several of the examples it is not possible to increase the number of aged fish without increasing the number of hauls as all fish on deck are aged.

7 Final conclusion

In general, most uncertainty indicators are either above or on the reference line. Only the French EVHOE survey and the Scottish groundfish survey show a major part of their uncertainty indicators for the abundance indices below the references line. A reduction of either number of hauls or number of age samples with 20% will increase the uncertainty of the indices with up to 20%.

The effect of a reduction in samples will move uncertainty indicators even further above the references line and the quality of the survey abundance indices will be reduced.

8 References

- Efron, B., and Tibshirani, R. J., 1993. *An Introduction to the Bootstrap*. CRC Press, Boca Raton.
- ICES, 2004. Report of the Workshop on Sampling and Calculation Methodology for Fisheries Data (WKSCMFD), 26–30 January 2004, Nantes, France. ICES CM 2004/ACFM:12
- ICES, 2005. Report of the Workshop on Survey Design and Data Analysis (WKSAD), 9–13 May 2005, Sète, France. ICES CM 2005/B:07.
- Lehtonen, R., and Pahkinen, E. 2004. "Practical Methods for design and Analysis of Complex Surveys". John Wiley & Sons Ltd. ISBN 0–470–84769–7.
- ISDBITS, 2001. Improvement of Stock Assessment and Data Collection by Continuation, Standardization and Design Improvement of the Baltic International Bottom Trawl Survey for Fishery Resource Assessment, Final Report. EU Project No. 98/099
- Pennington M. R., and M. D. Grosslein, 1978. Accuracy of abundance indices based on stratified random trawl surveys. ICNAF Res. Doc. 78/IV/77 : 42 p.
- O'Brien C. M., Darby, C. D., Maxwell, D. L., Rackham, B. D., Degel, H., Flatman, S., Pastoors, M. A., Simmonds, E. J., and Vinther, M. 2001a. The precision of international market sampling for North Sea plaice (*Pleuronectes platessa* L.) and its influence on stock assessment. ICES CM 2001/P:13.
- O'Brien C. M., Darby, C. D., Rackham, B. D., Maxwell, D. L., Degel, H., Flatman, S., Mathewson, M., Pastoors, M. A., Simmonds, E. J., and Vinther, M. 2001b. The precision of

international market sampling for North Sea cod (*Gadus morhua* L.) and its influence on stock assessment. ICES CM 2001/P:14.

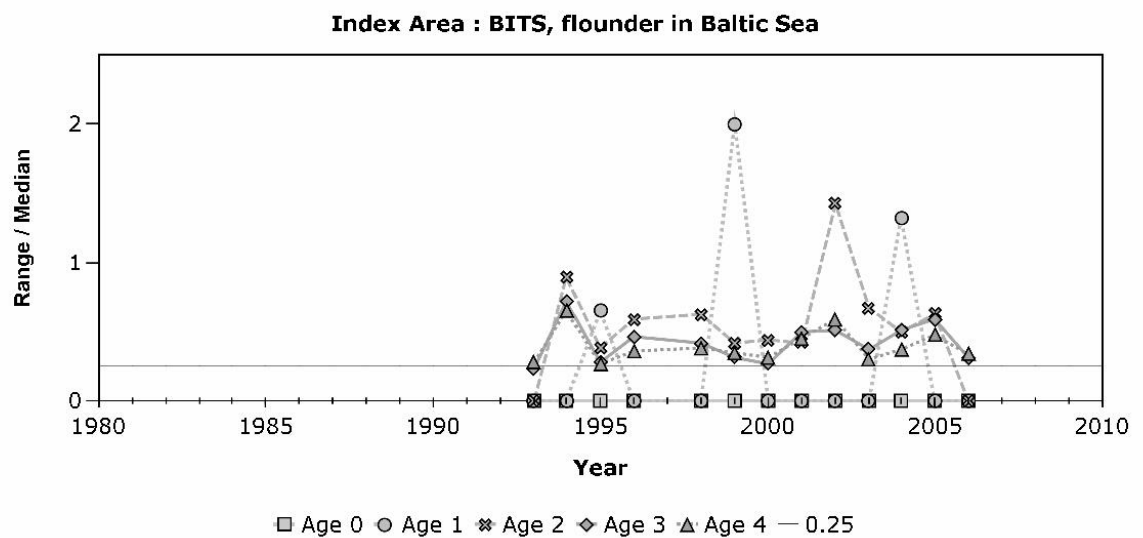
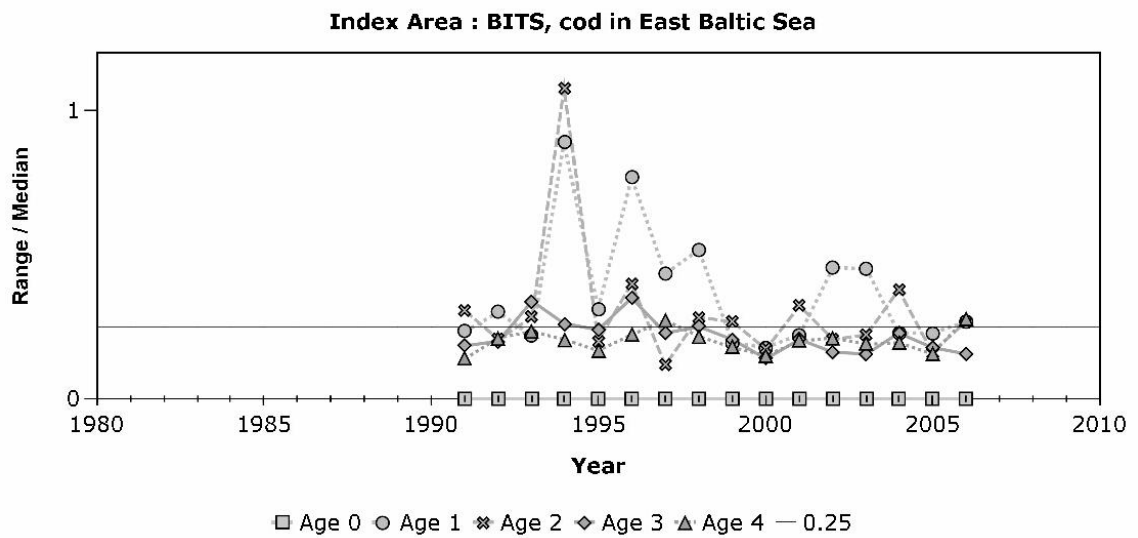
Pennington, M. 1983. Efficient estimators of abundance, for fish and plankton surveys. *Biometrics*, 39: 281–286.

Petitgas, P. 1993. Geostatistics for fish stock assessments: a review and an acoustic application. *ICES Journal of Marine Science*, 50: 285–298.

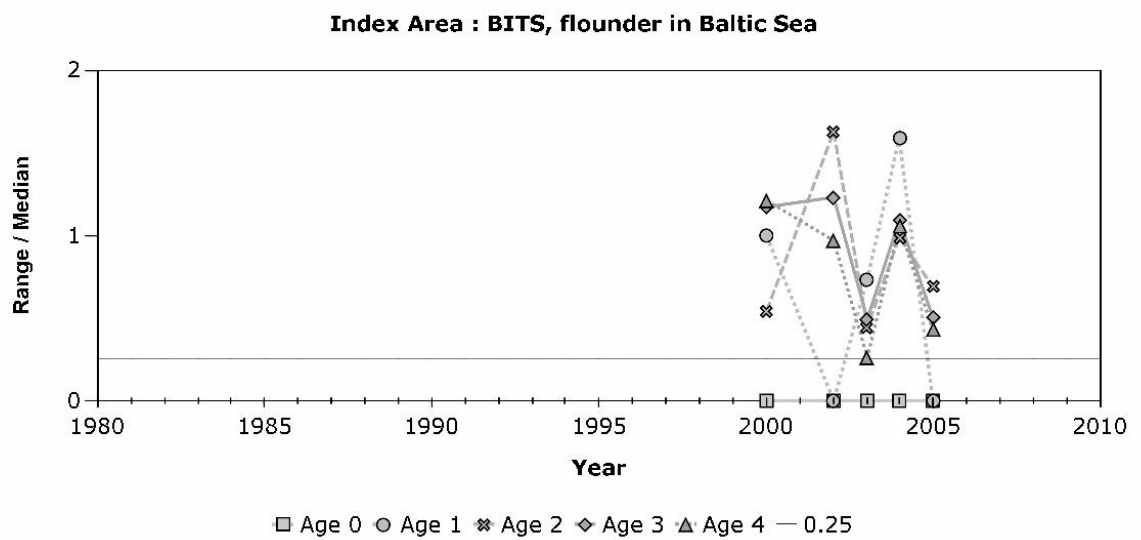
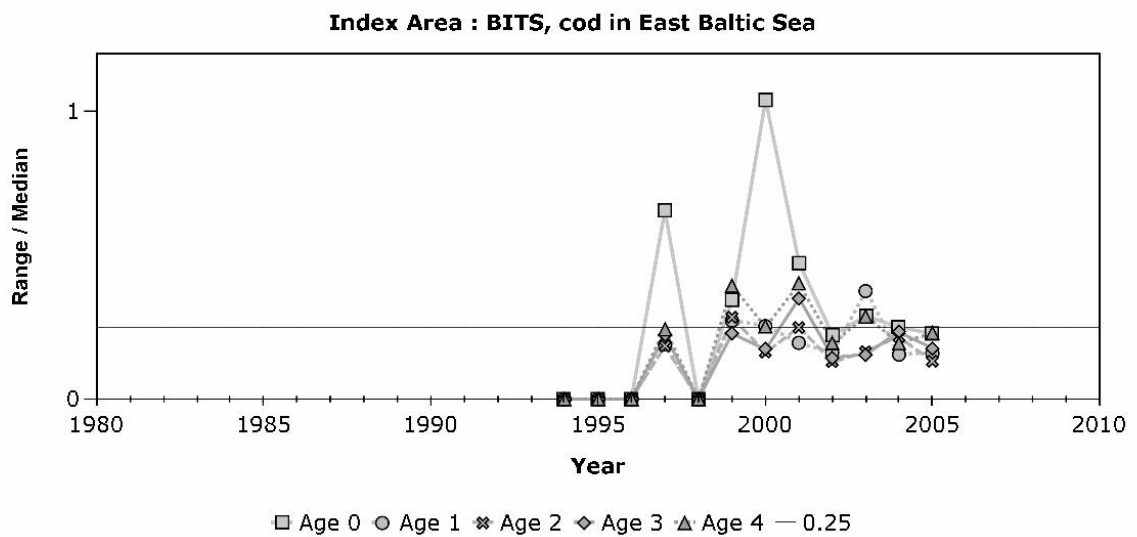
Simmonds, E. J., Needle, C. L., Degel, H., Flatman, S., O'Brian, C. M., Pastoors, M., Robb, A. P., and Vinther, M. (2001) The precision of international market sampling for North Sea herring and its influence on assessment. *Theme Session: Quality and Precision Data Underlying Fish Stock Assessment 2001/P:21*, 1–22 Copenhagen: ICES.

Annex 1: Range/median figures for all surveys stored in DATRAS

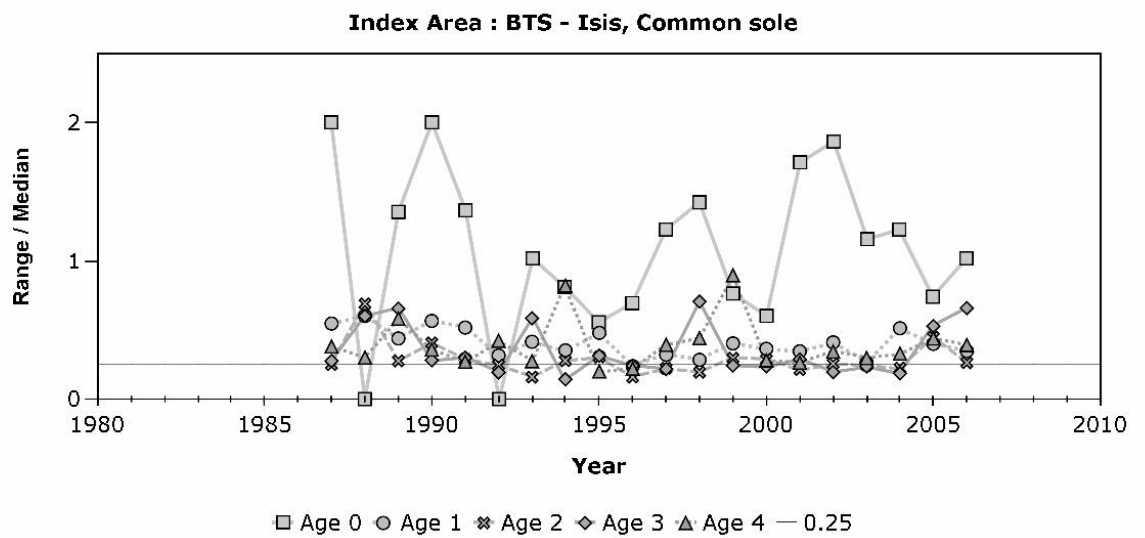
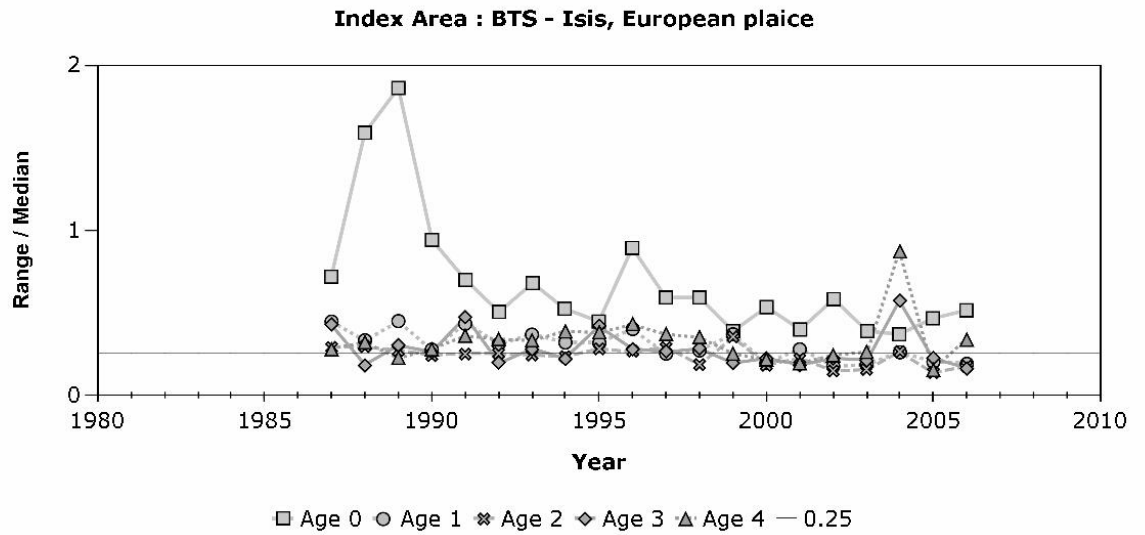
Survey: Baltic International Trawl Survey (BITS)—Quarter: 1



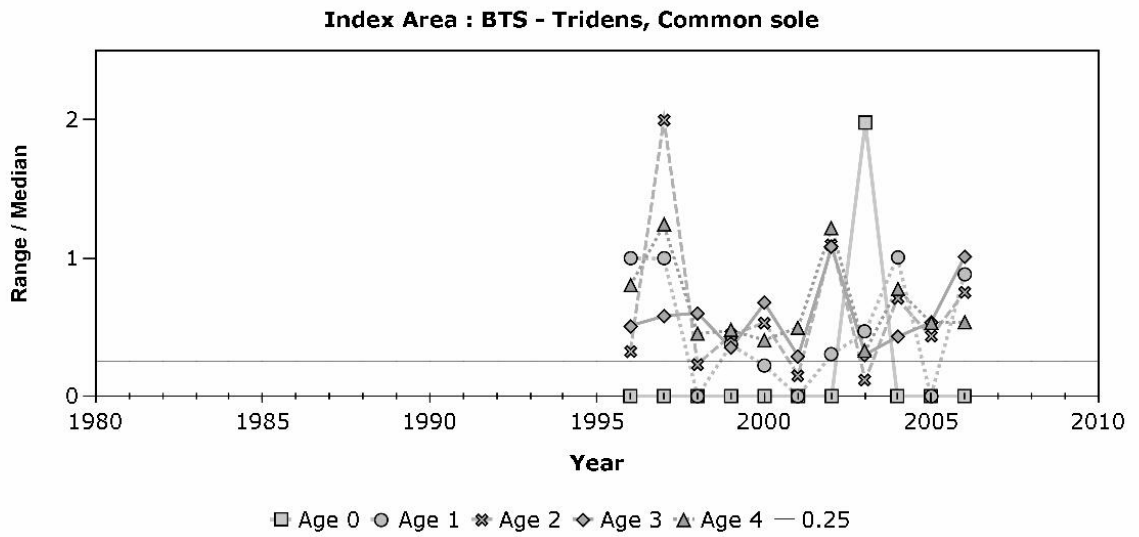
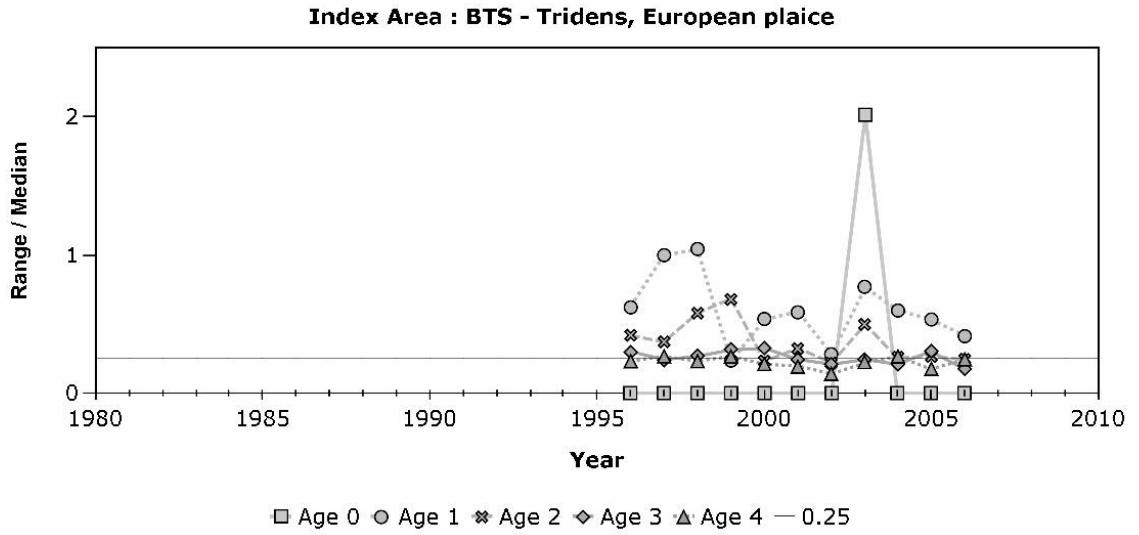
Survey: Baltic International Trawl Survey (BITS)—Quarter: 4



Survey: Beam Trawl Survey -Isis (BTS-Isis)—Quarter: 3

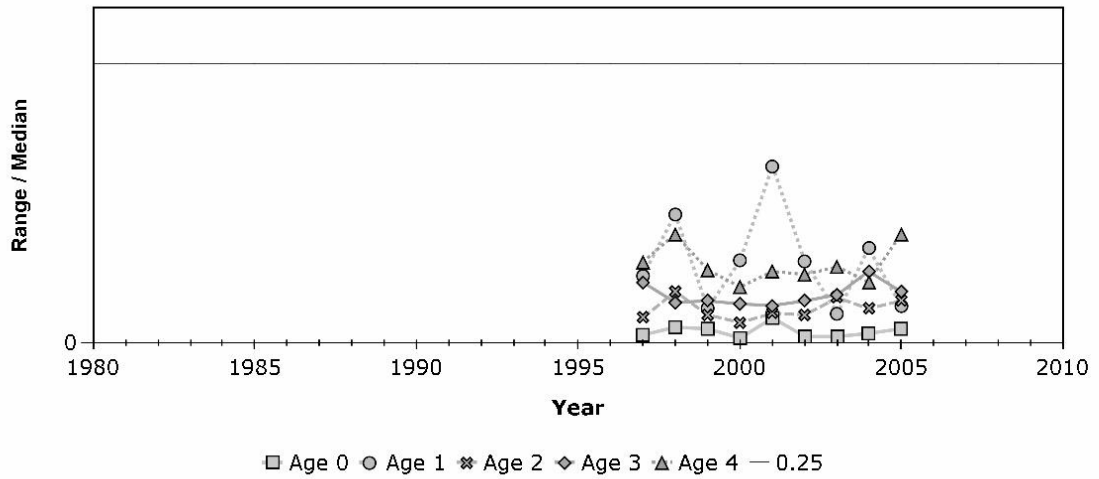


Survey: Beam Trawl Survey -Tridens (BTS-Tridens)—Quarter: 3

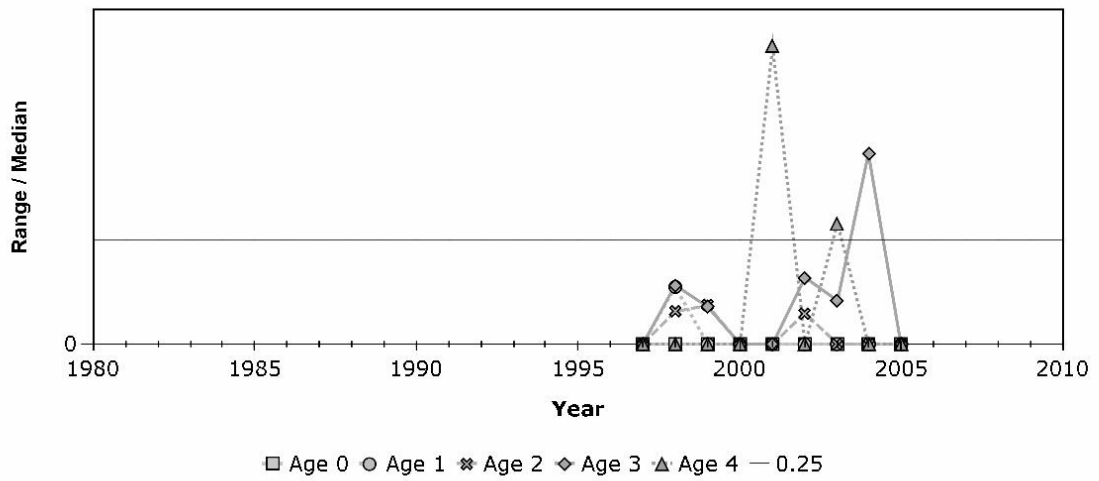


Survey: French Groundfish Survey (EVHOE)—Quarter: 4

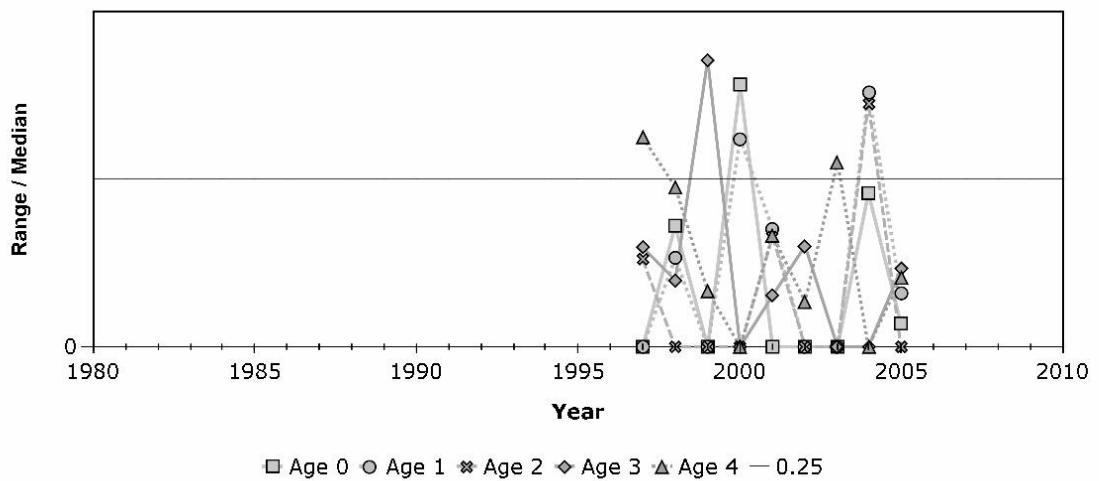
Index Area : EVHOE, hake in Bay of Biscay and Celtic Sea



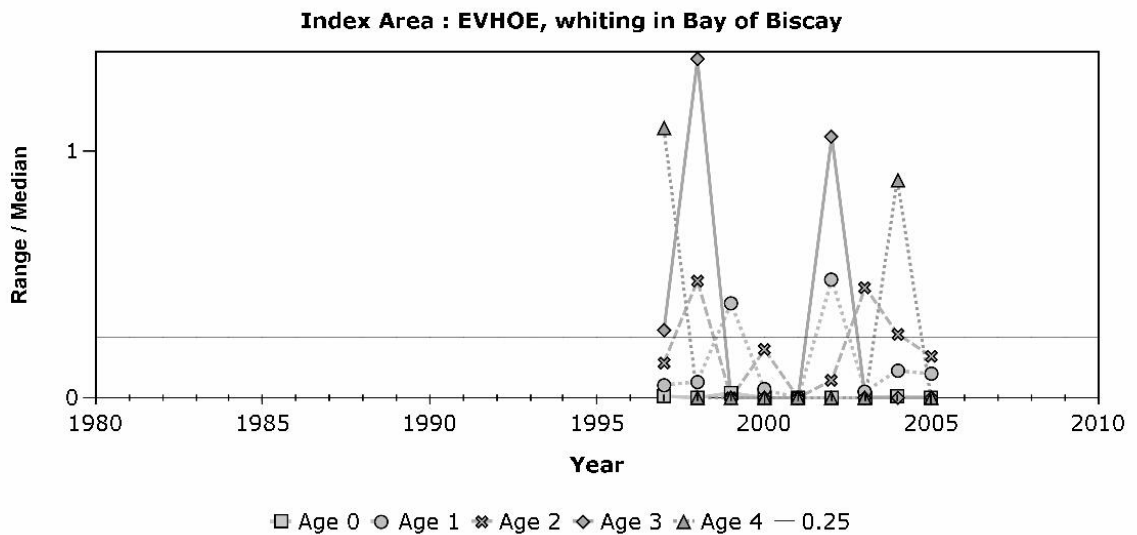
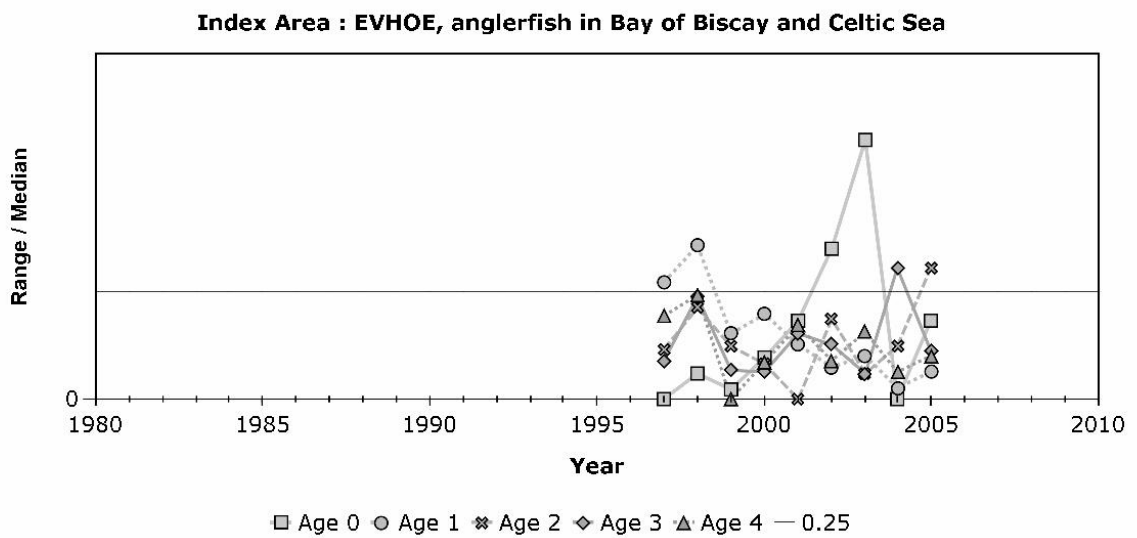
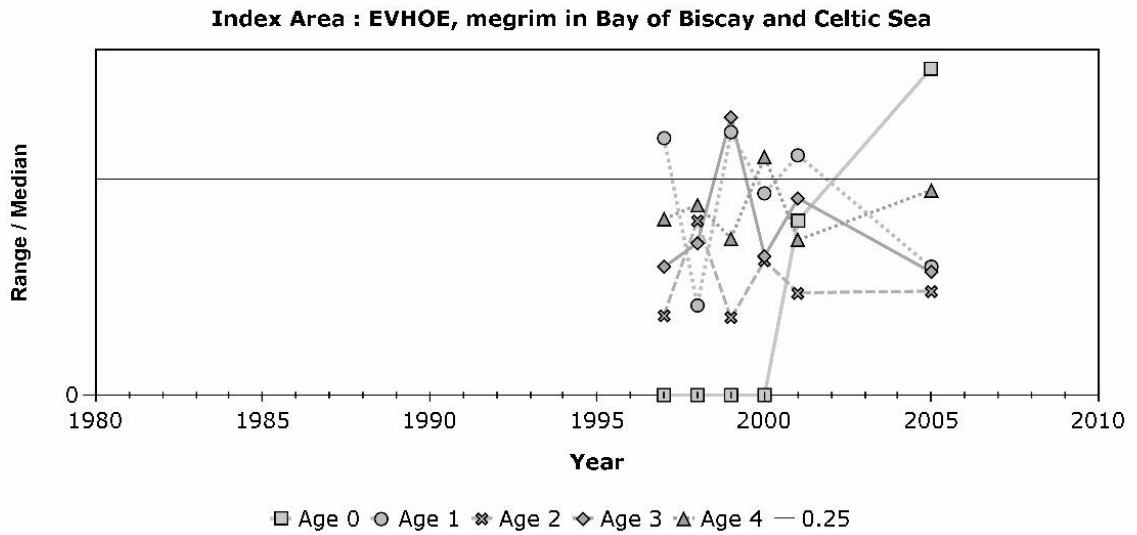
Index Area : EVHOE, cod in Celtic Sea



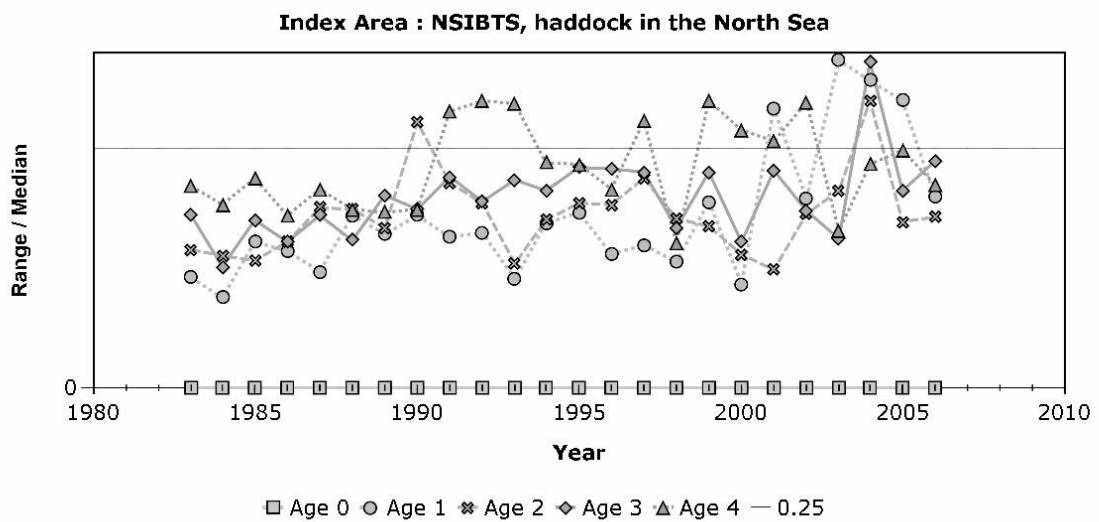
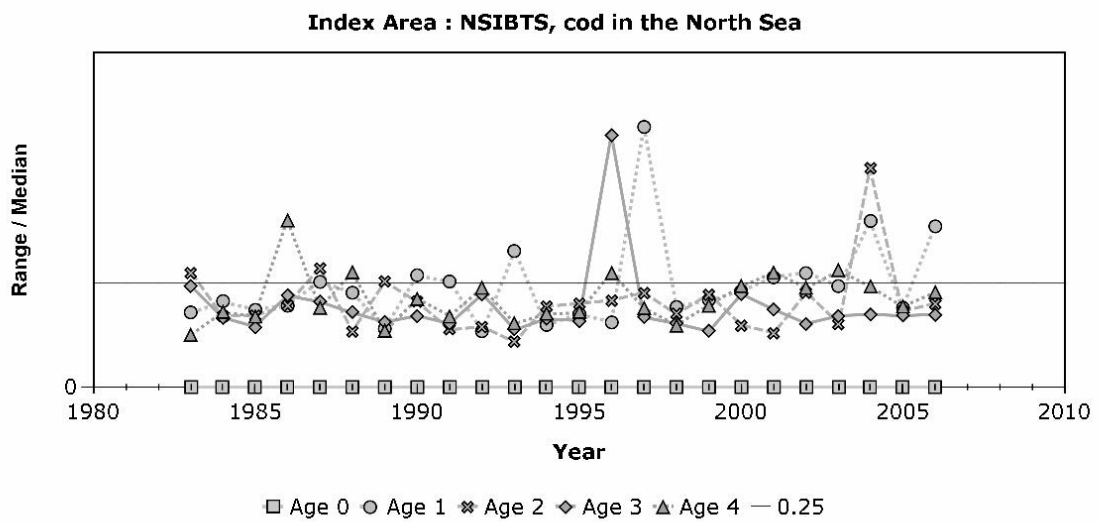
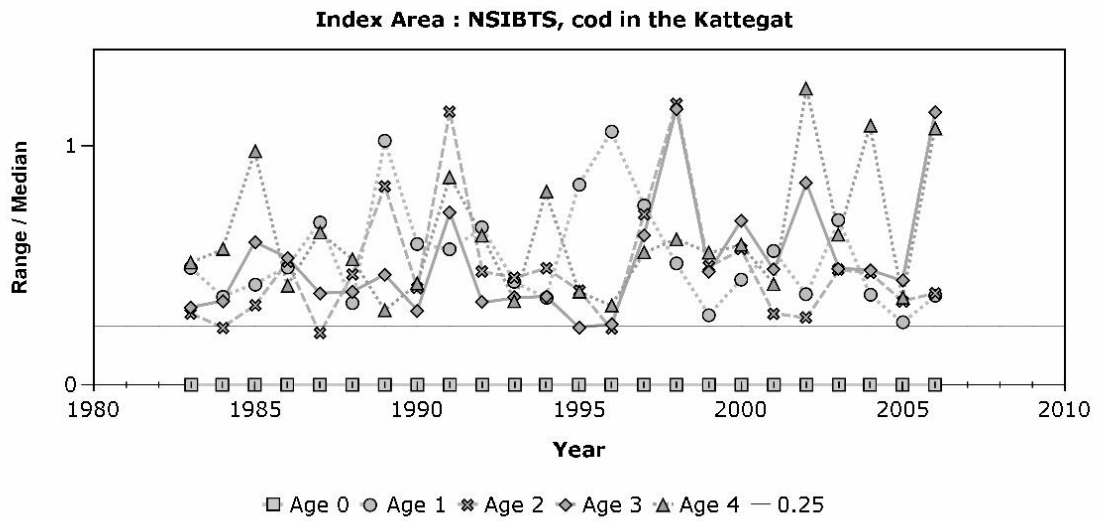
Index Area : EVHOE, monkfish in Bay of Biscay and Celtic Sea

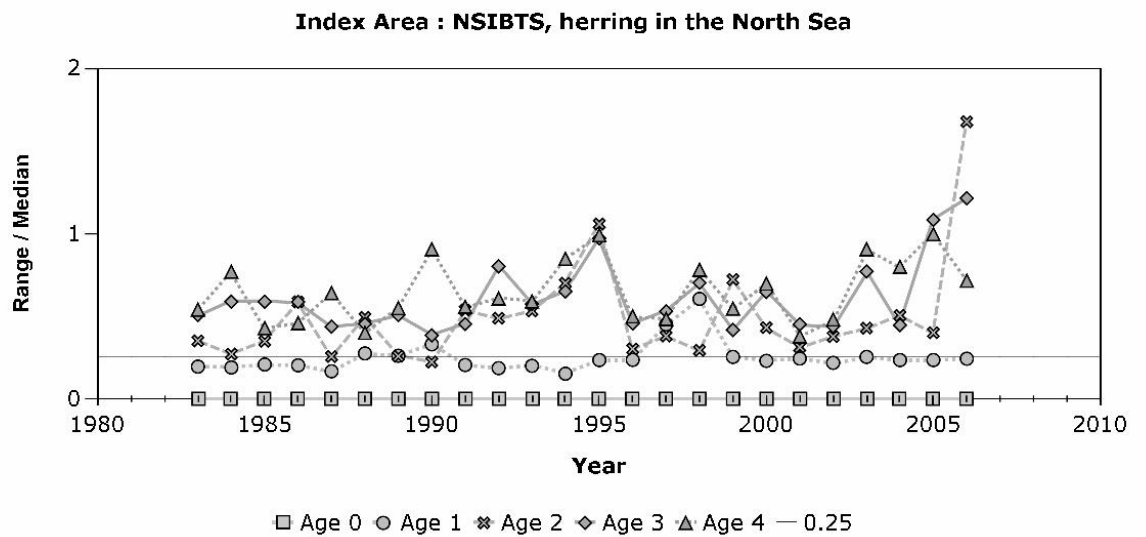
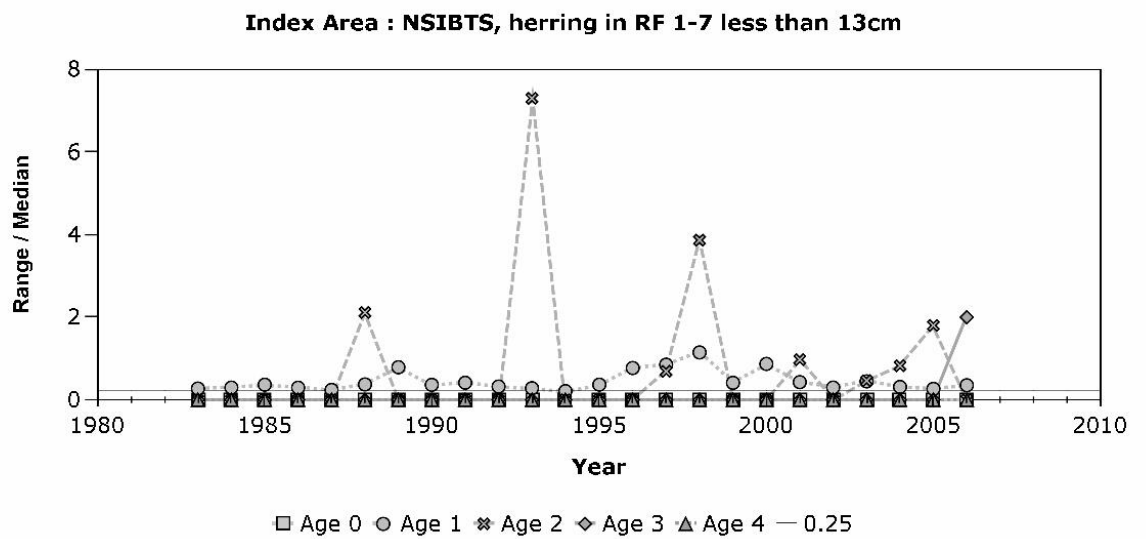
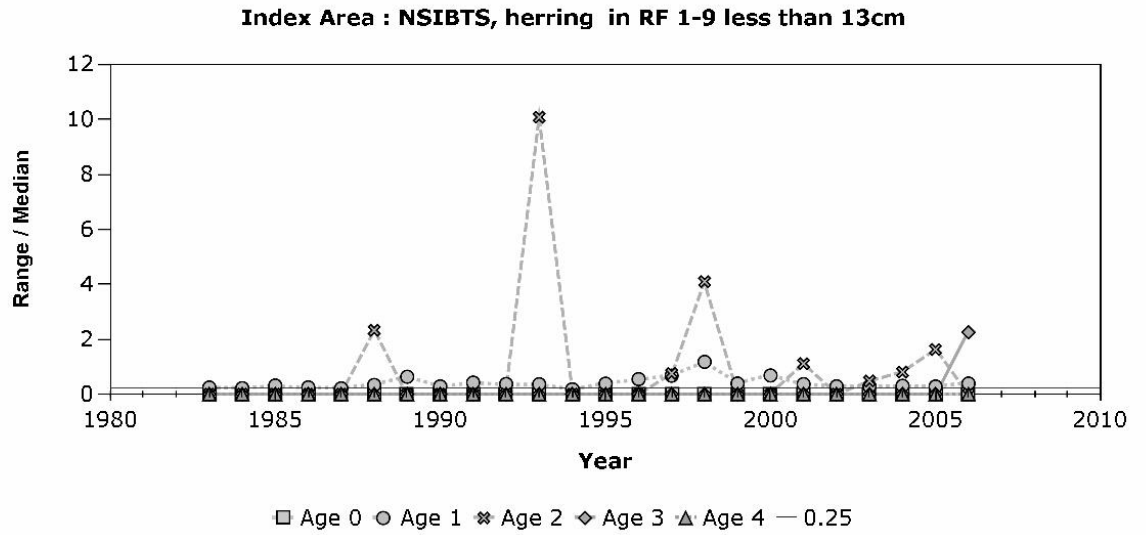


Survey: French Groundfish Survey (EVHOE)—Quarter: 4 Continued

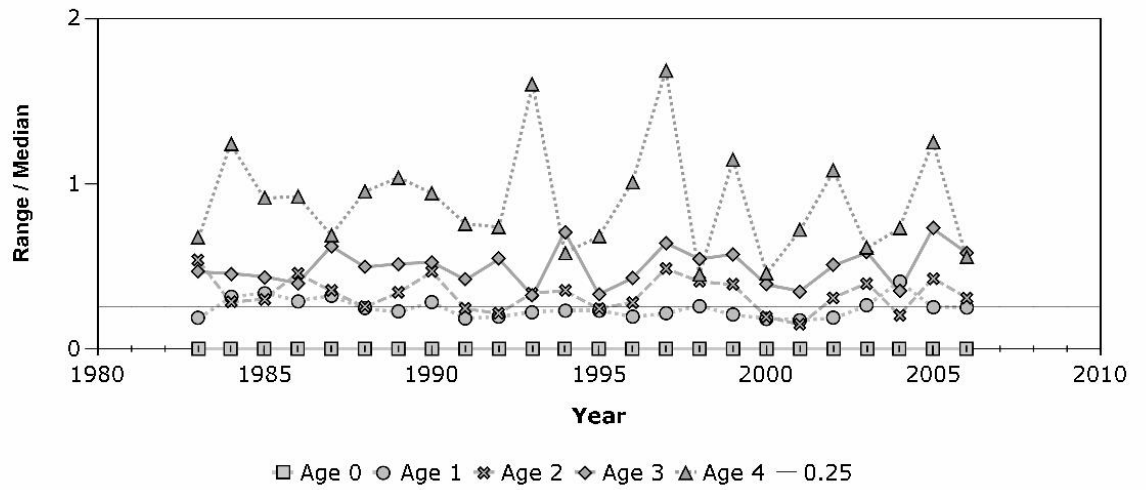


Survey: North Sea International Bottom Trawl Survey (NSIBTS)—Quarter: 1

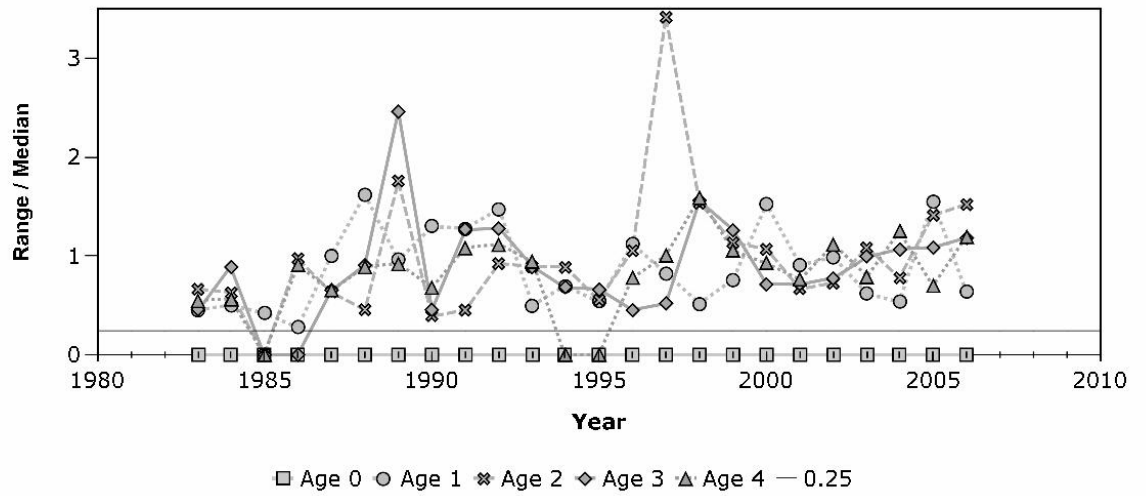




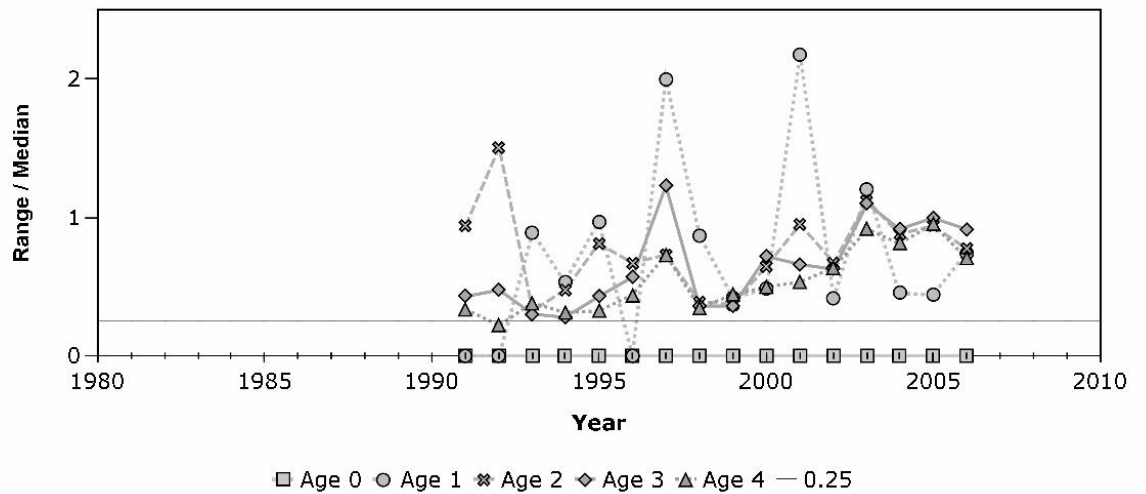
Index Area : NSIBTS, Norway pout in the North Sea

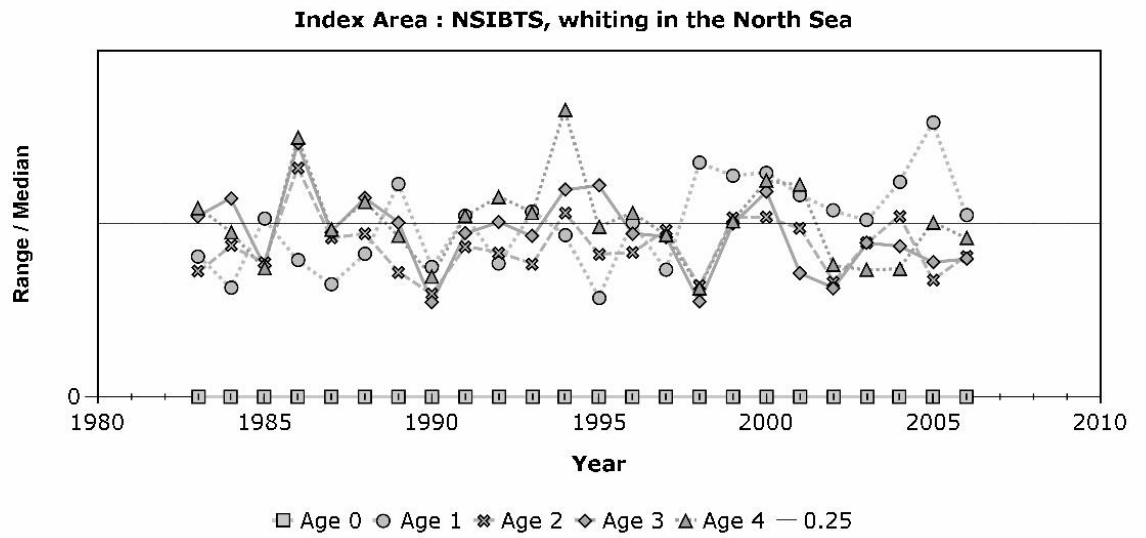
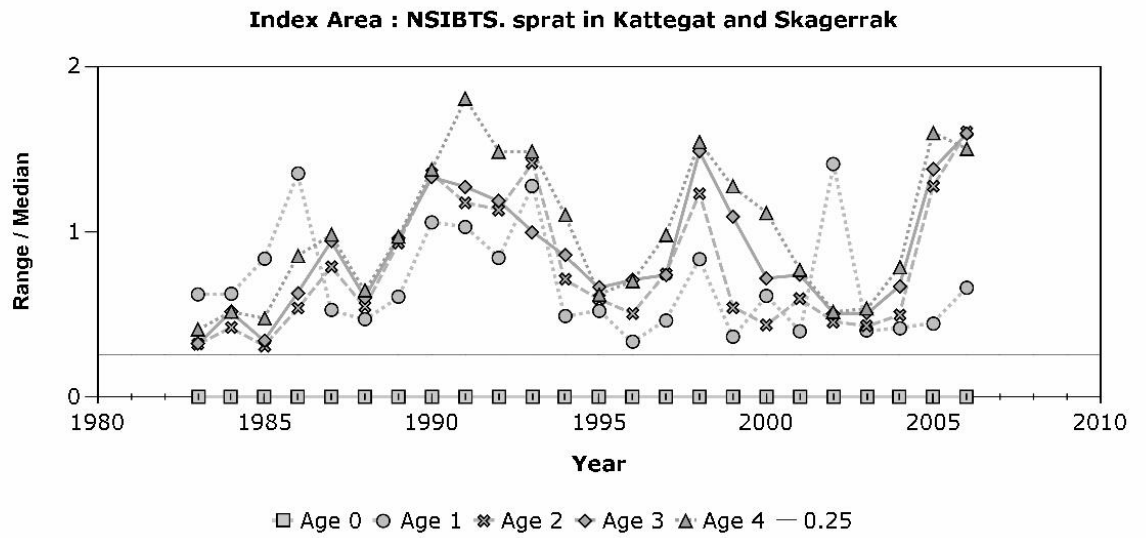
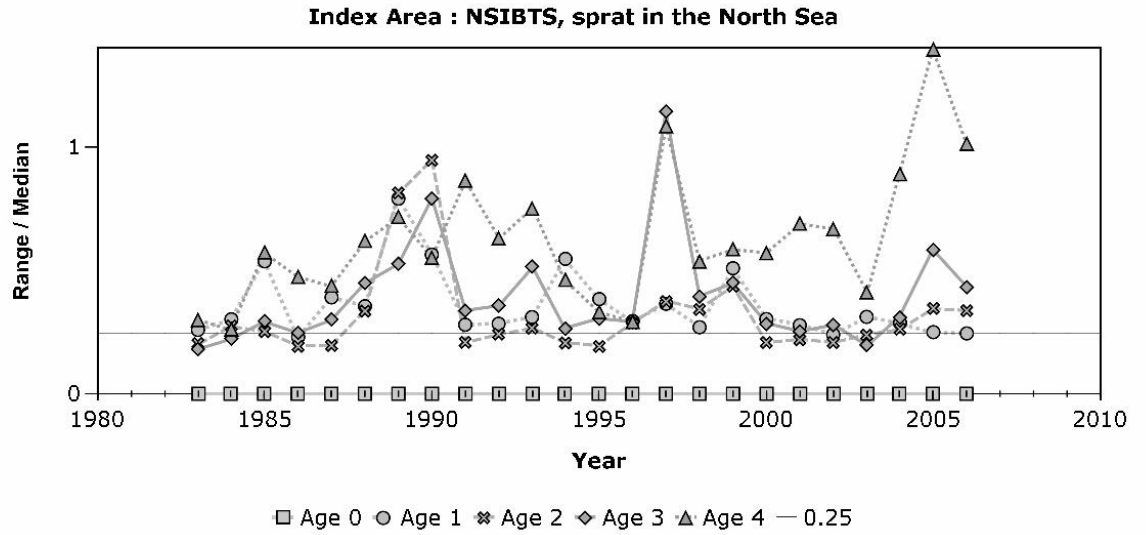


Index Area : NSIBTS, mackerel in the North Sea

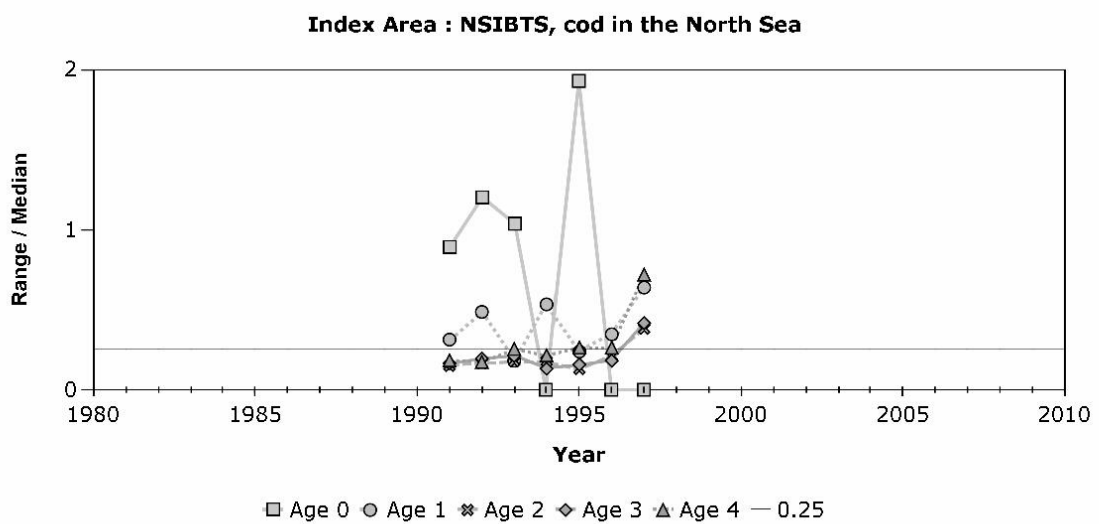
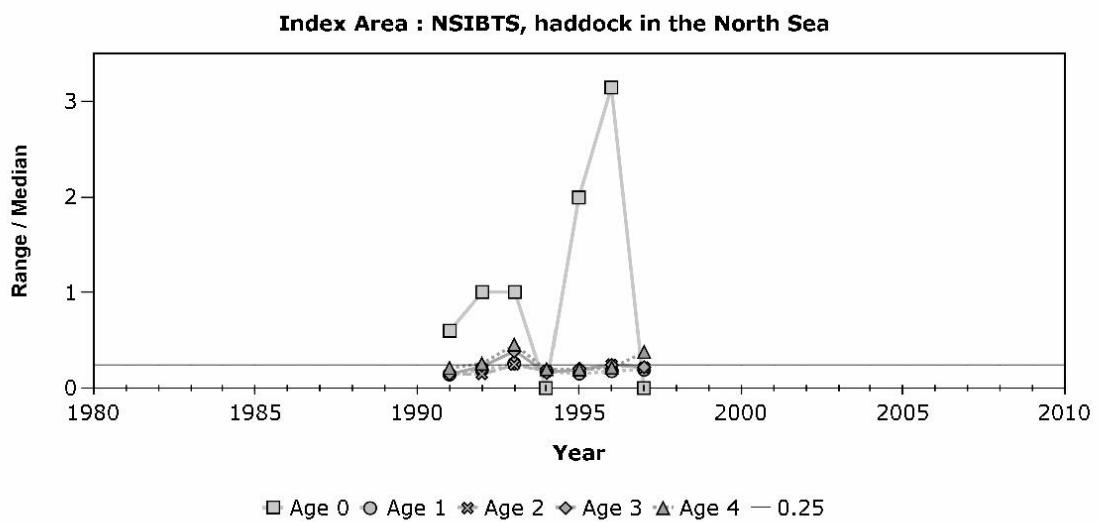


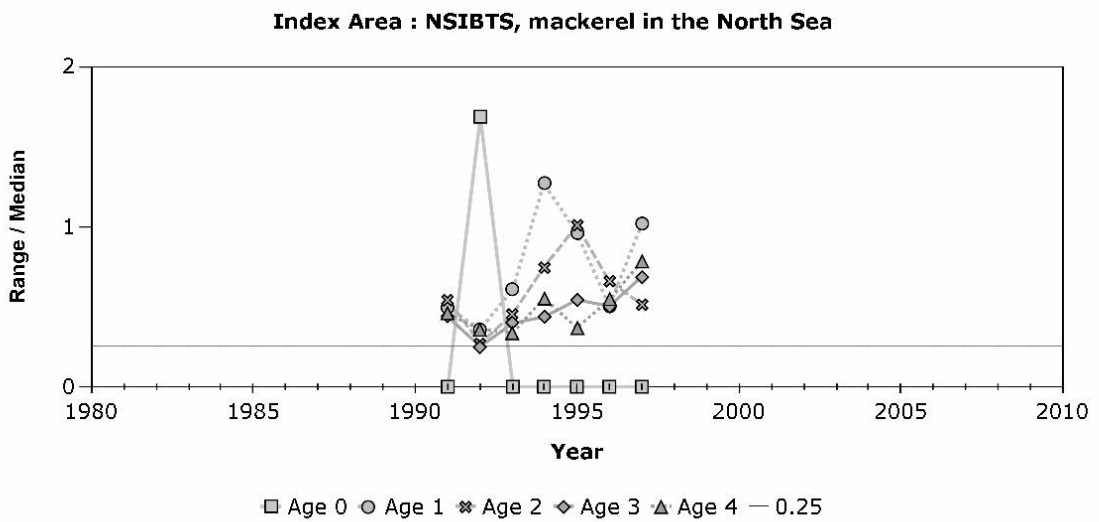
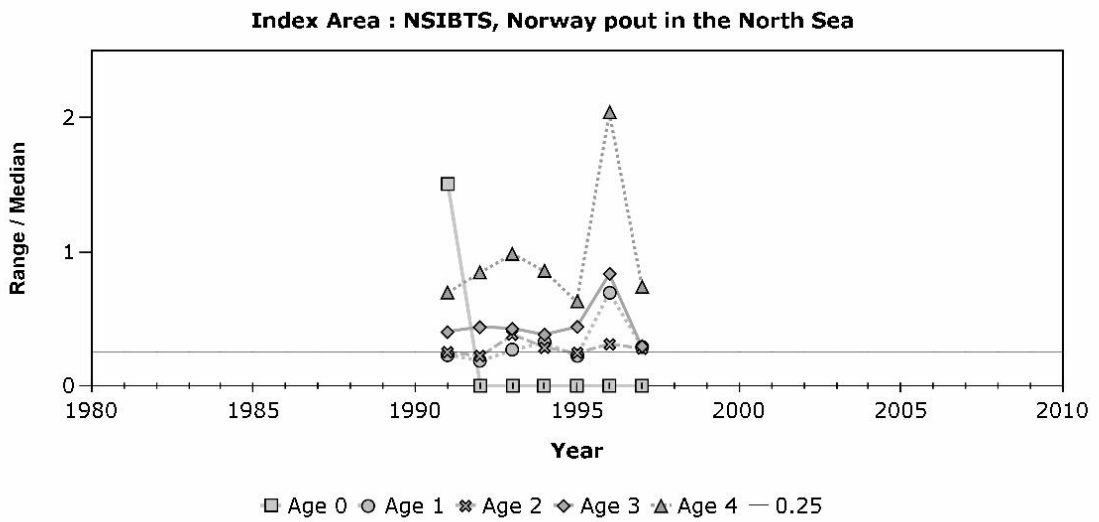
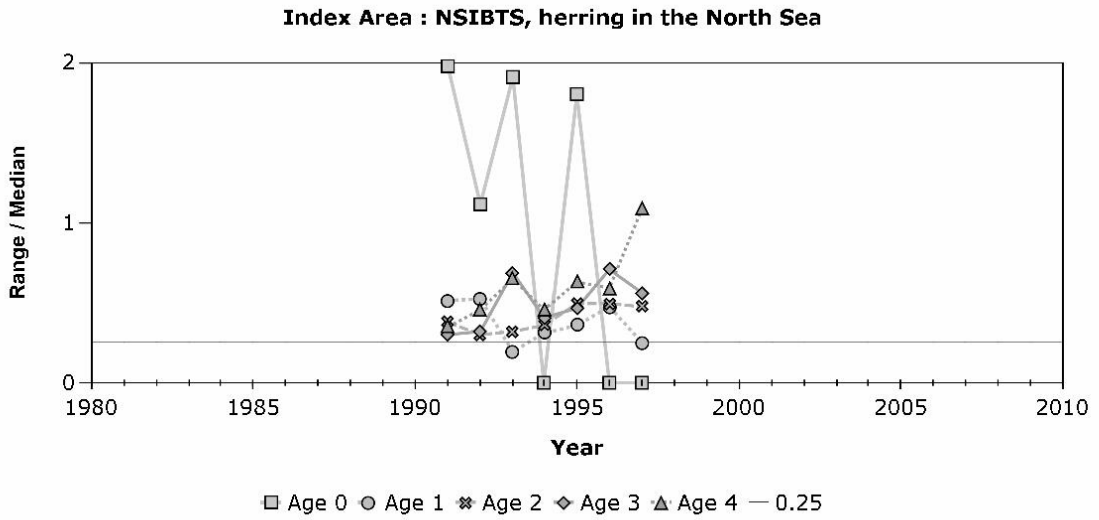
Index Area : NSIBTS, plaice in Kattegat and Skagerrak

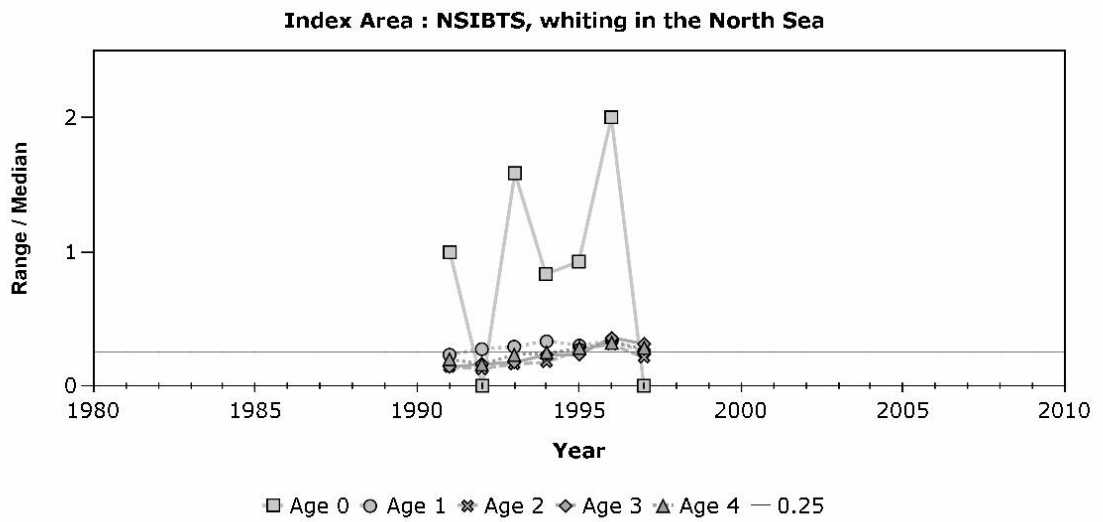
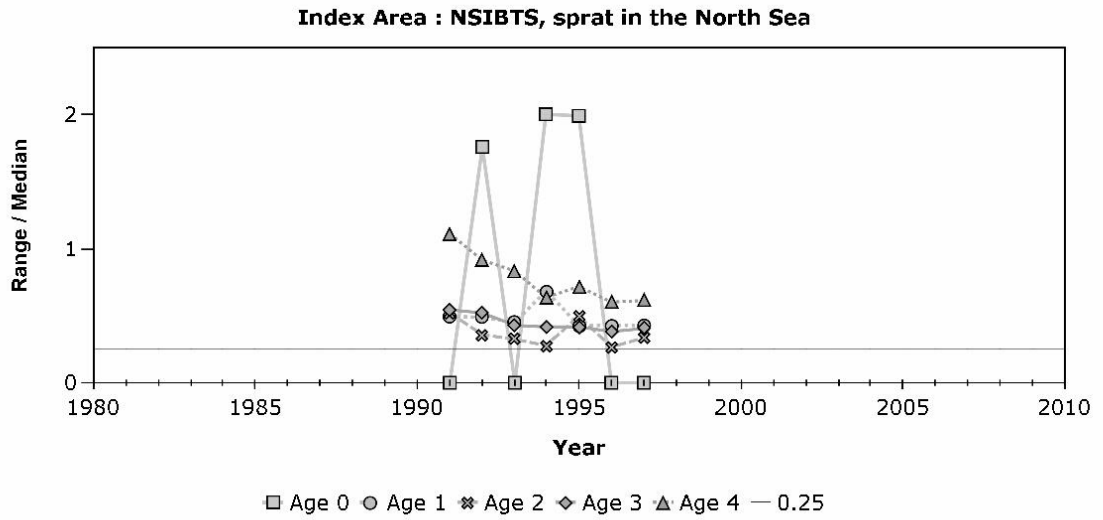




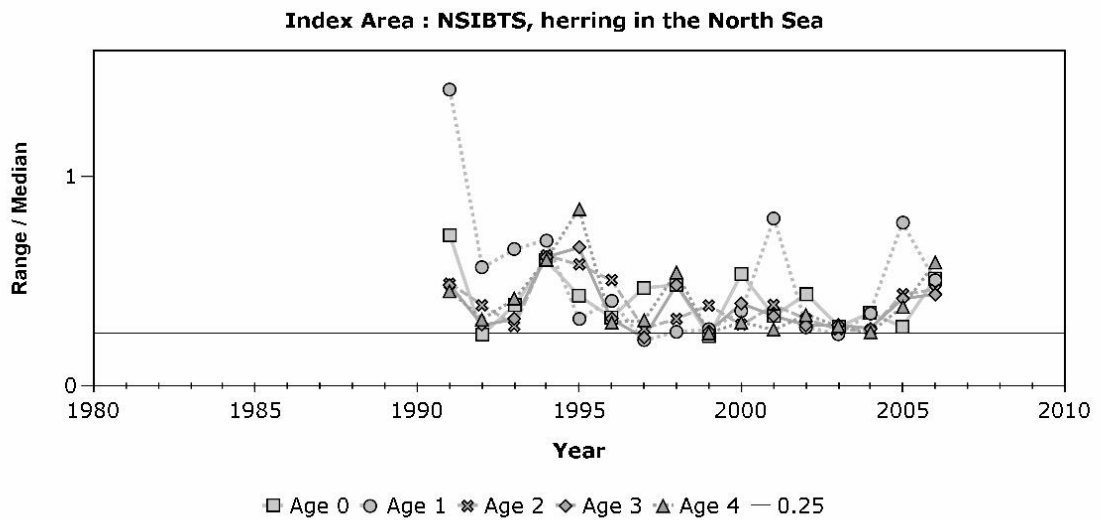
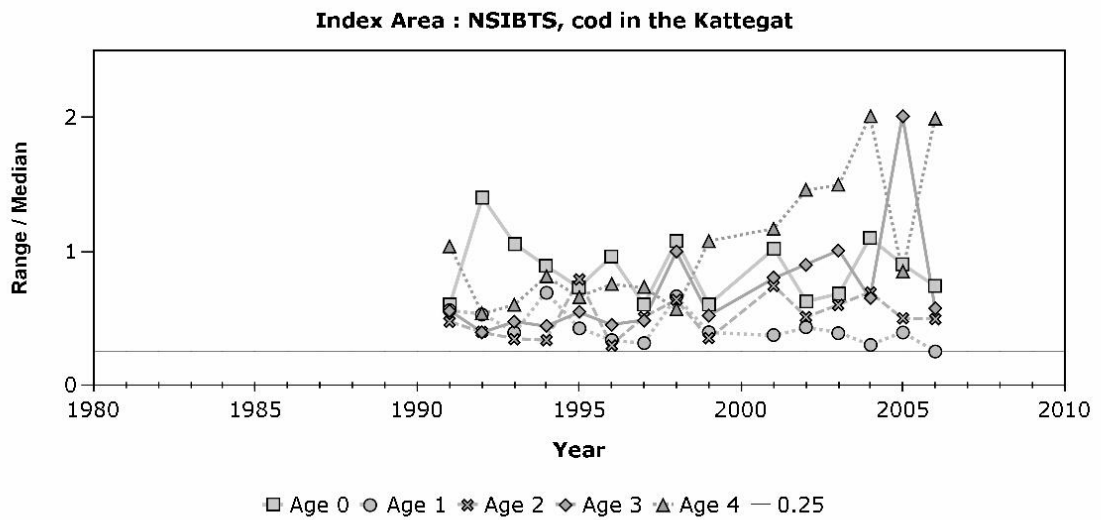
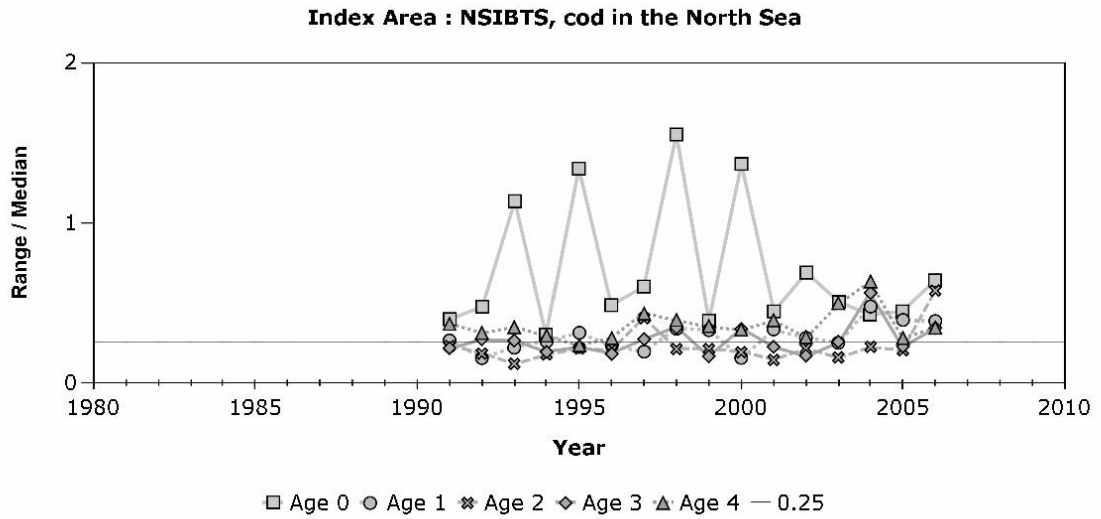
Survey: North Sea International Bottom Trawl Survey (NSIBTS)—Quarter: 2



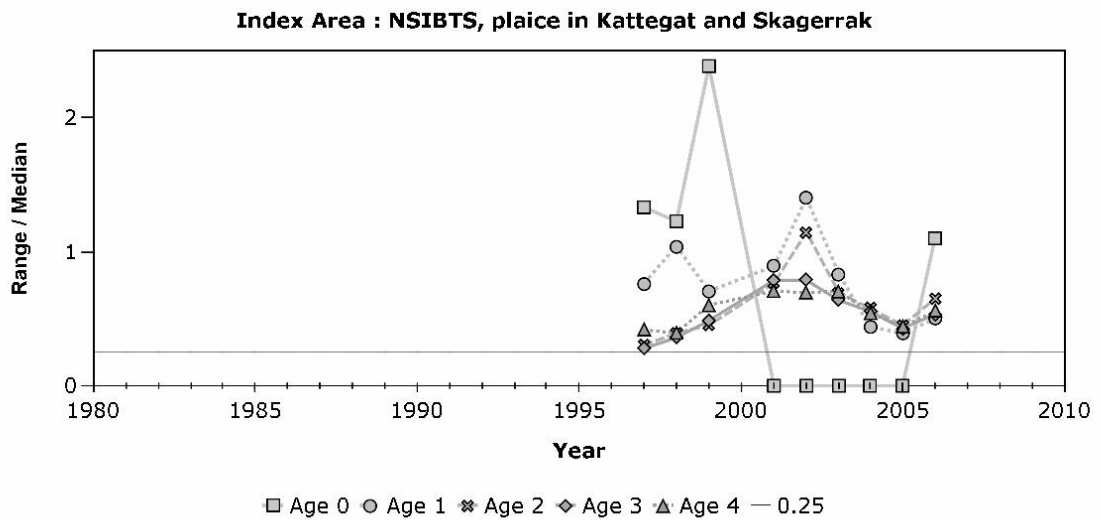
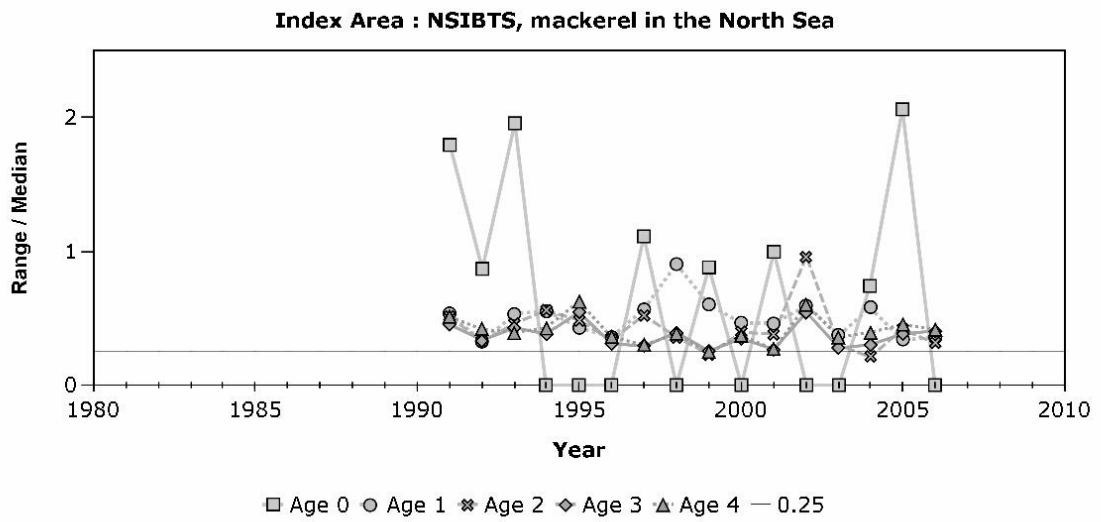
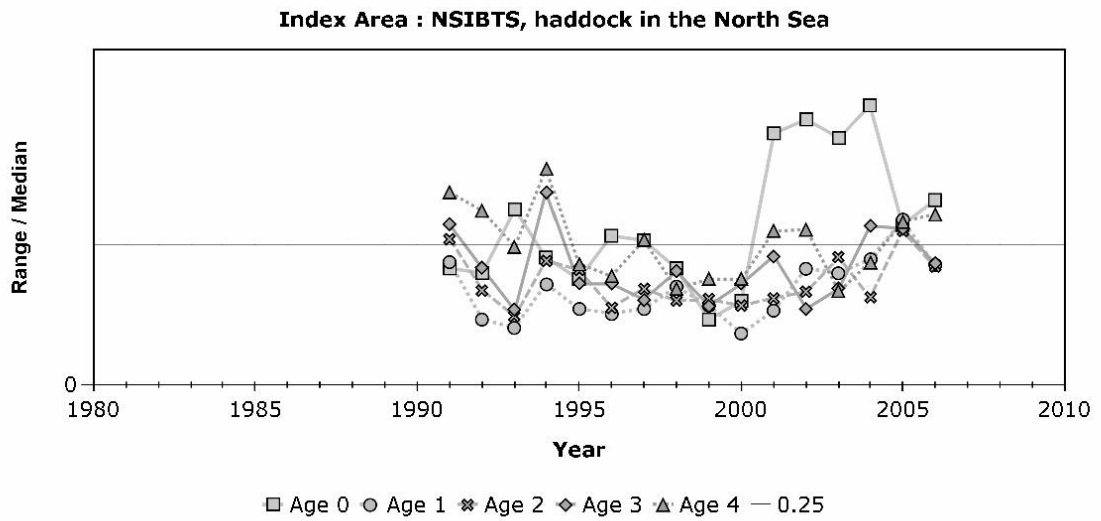




Survey: North Sea International Bottom Trawl Survey (NSIBTS)—Quarter: 3

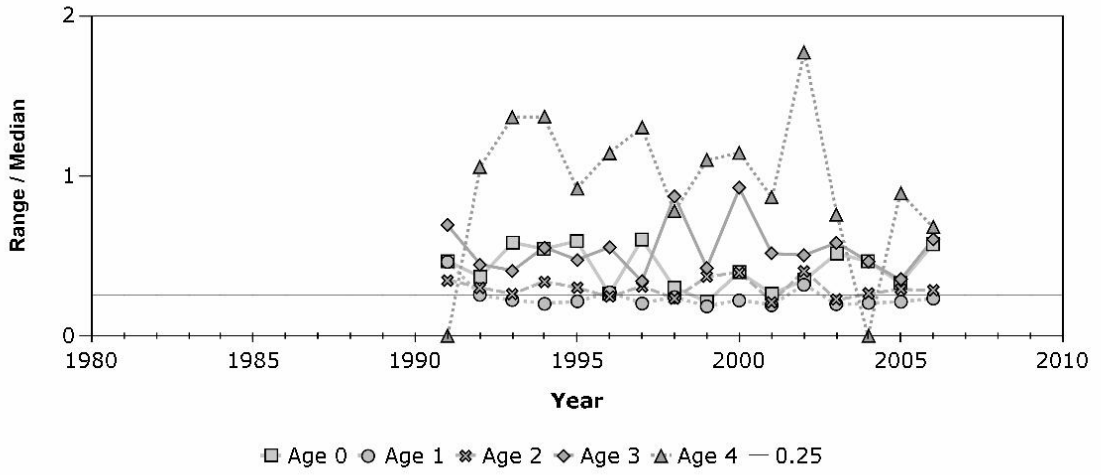


Survey: North Sea International Bottom Trawl Survey (NSIBTS)—Quarter: 3 Continued

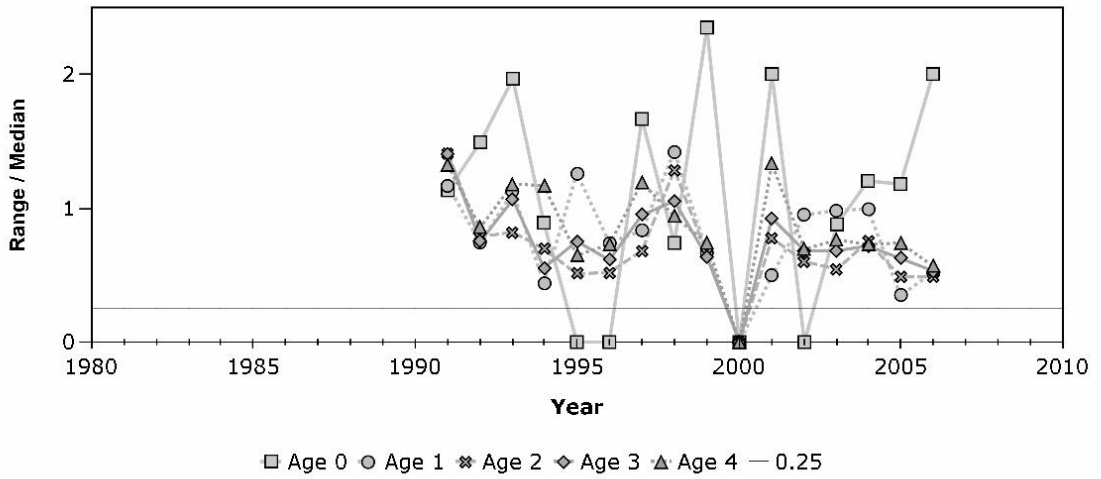


Survey: North Sea International Bottom Trawl Survey (NSIBTS)—Quarter: 3 Continued

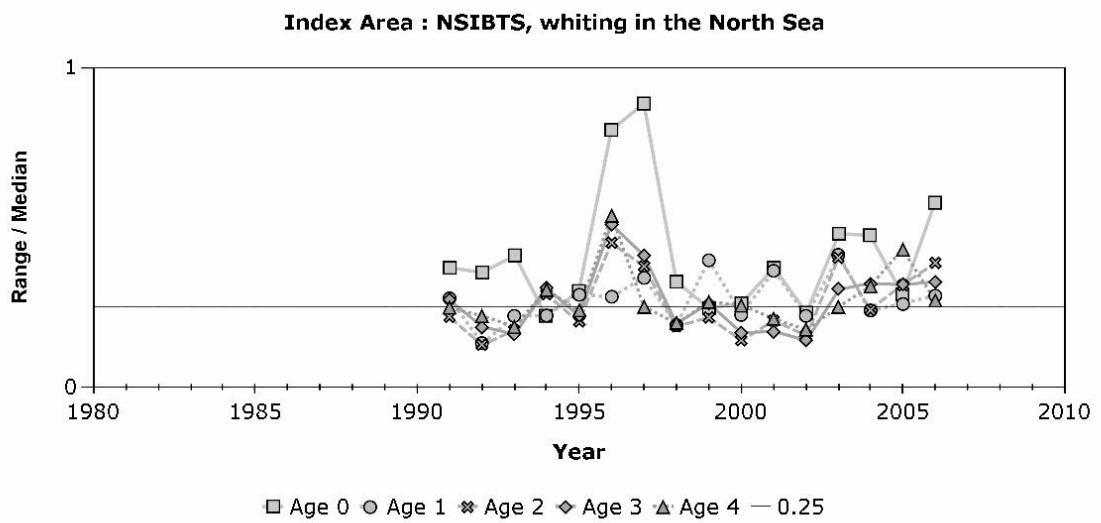
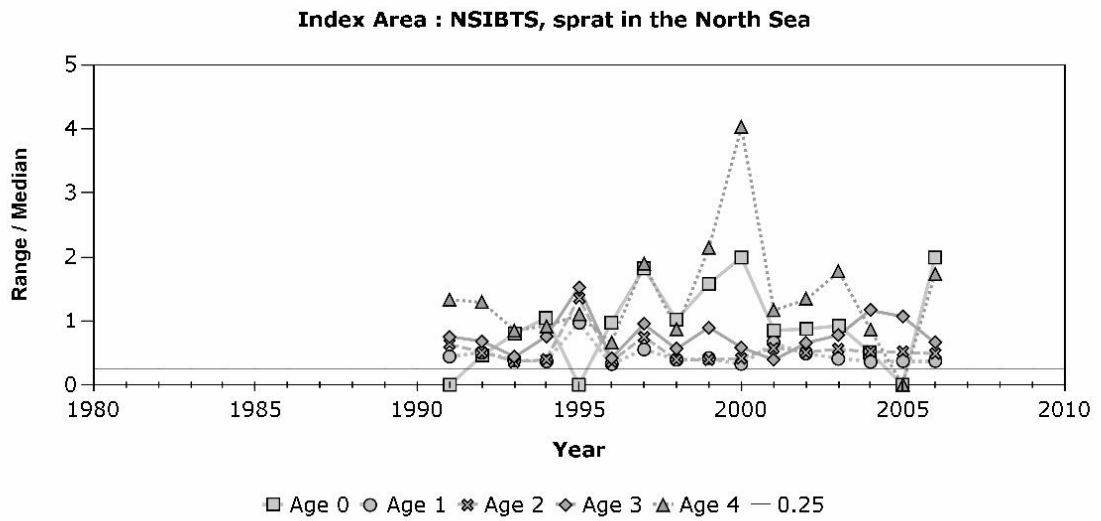
Index Area : NSIBTS, Norway pout in the North Sea



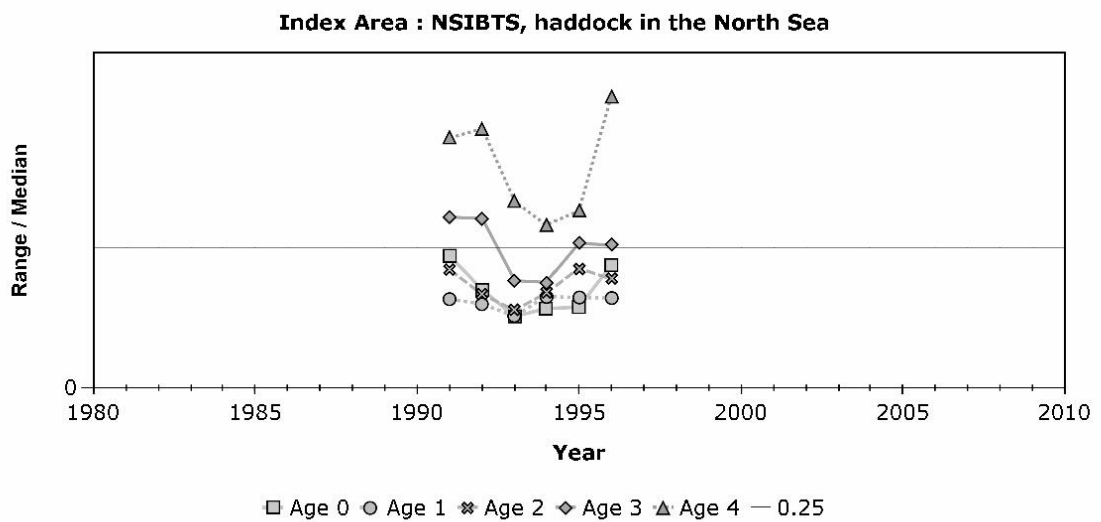
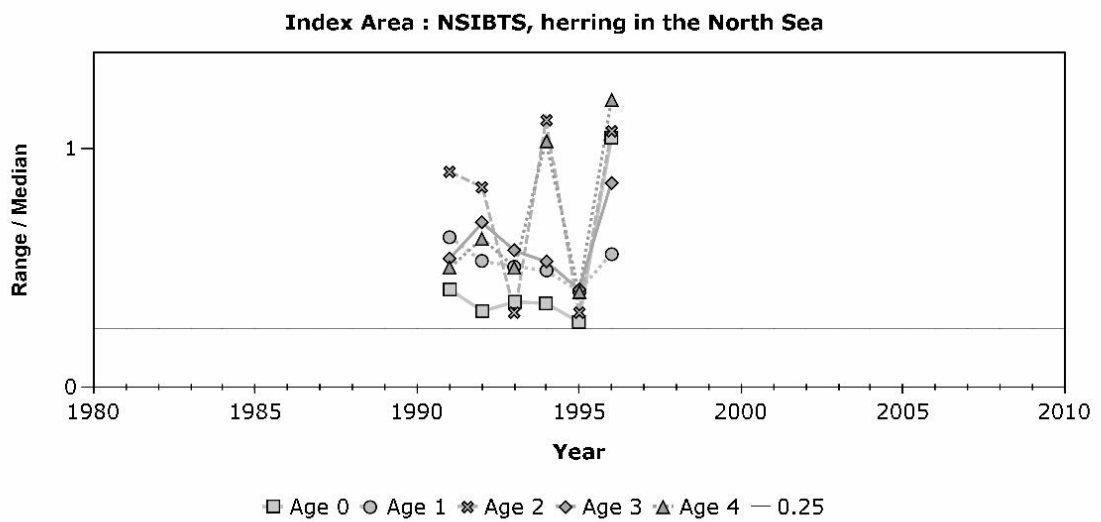
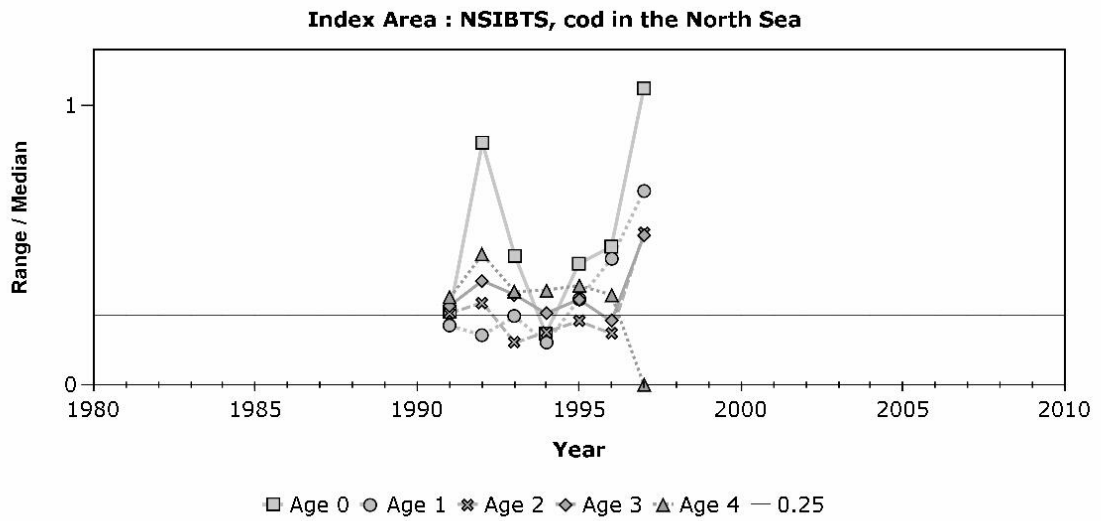
Index Area : NSIBTS. sprat in Kattegat and Skagerrak



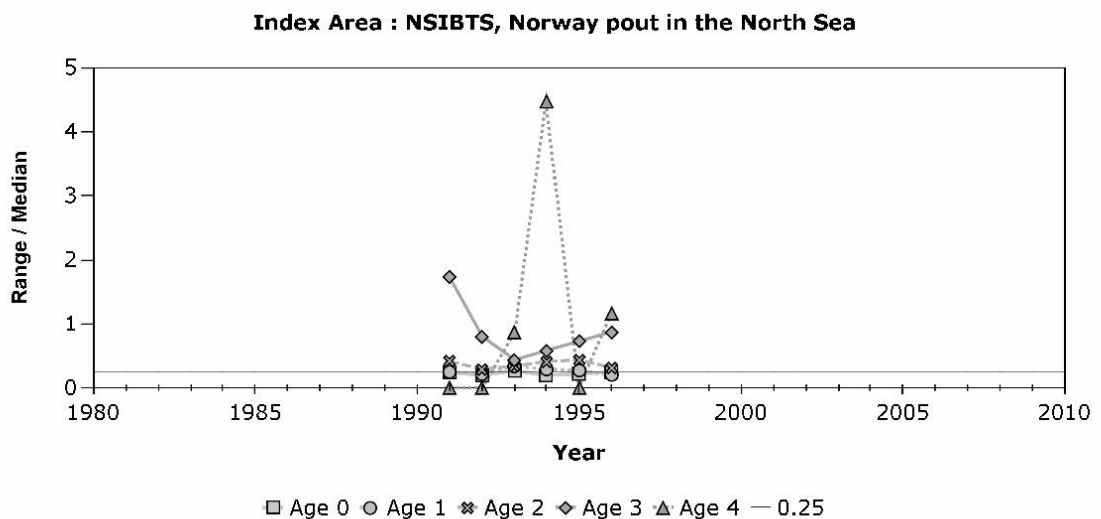
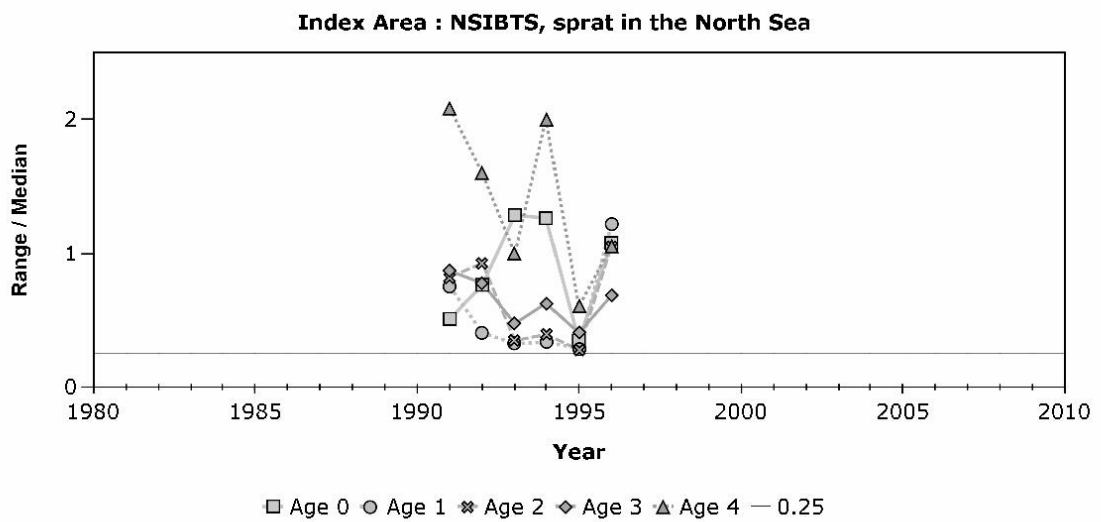
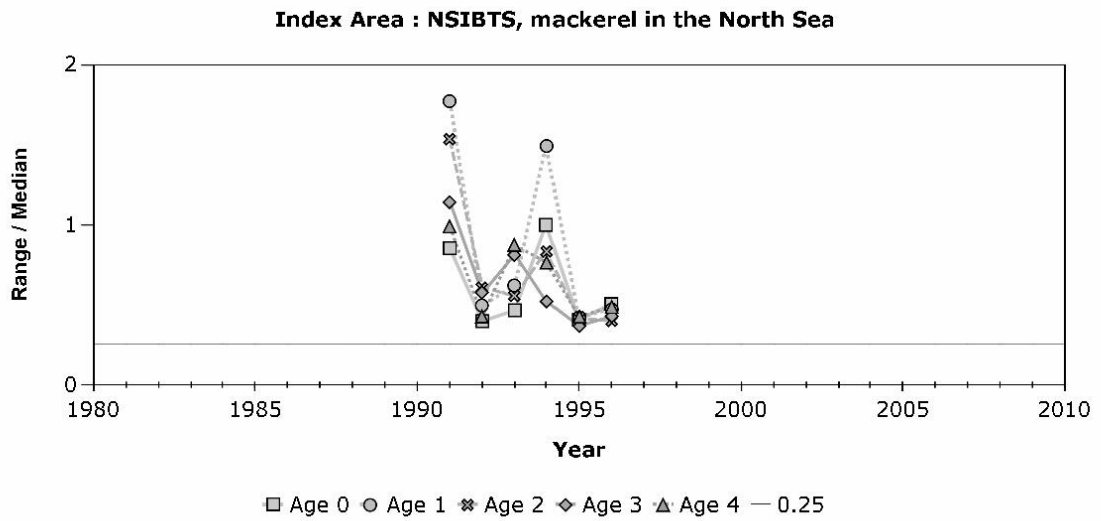
Survey: North Sea International Bottom Trawl Survey (NSIBTS)—Quarter: 3 Continued



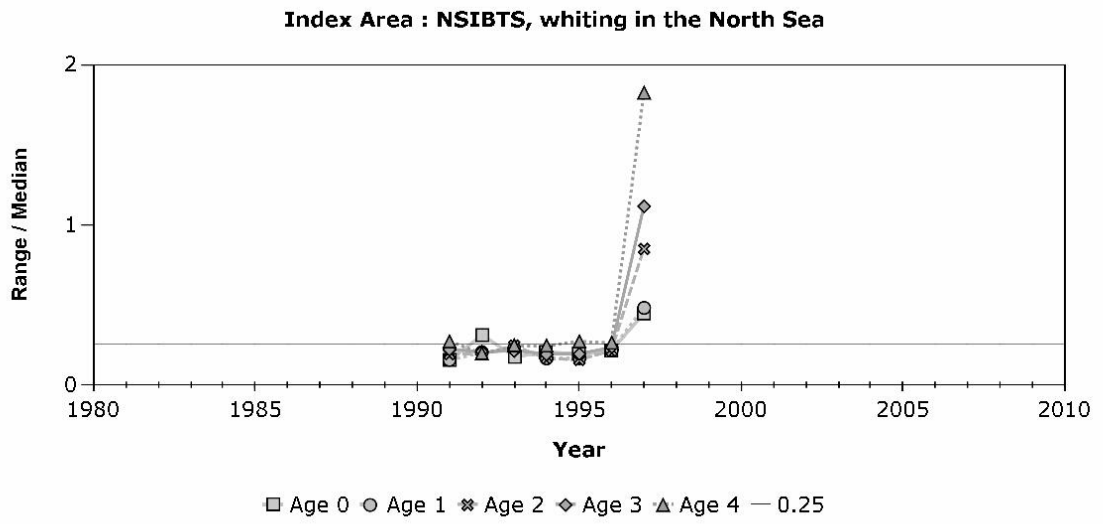
Survey: North Sea International Bottom Trawl Survey (NSIBTS)—Quarter: 4



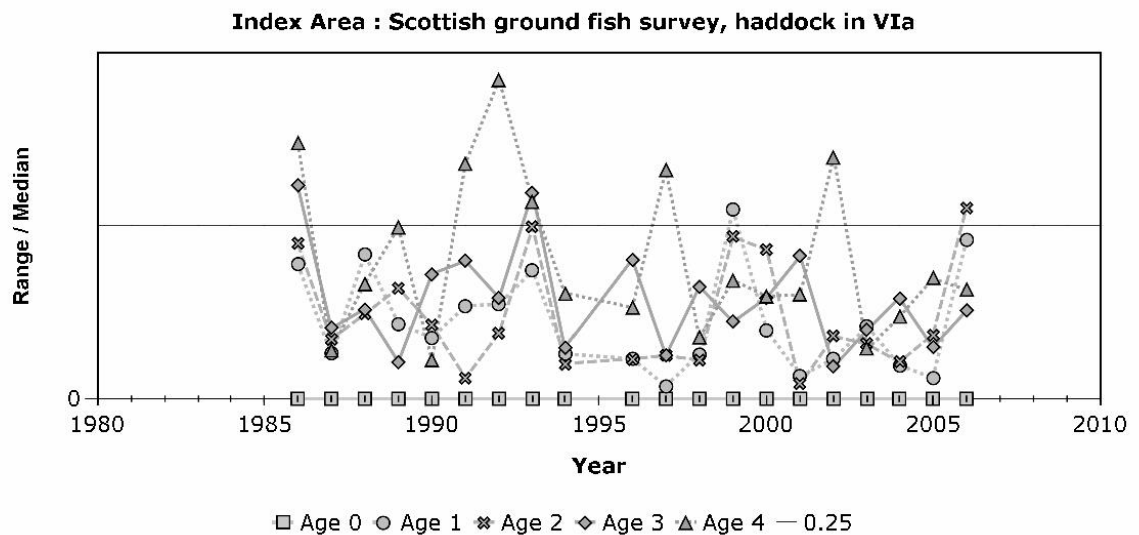
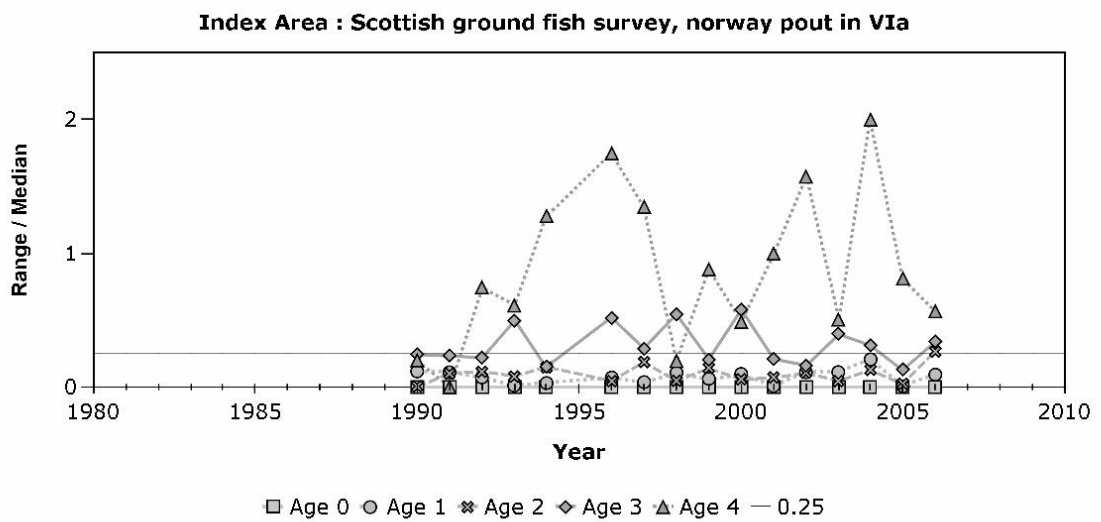
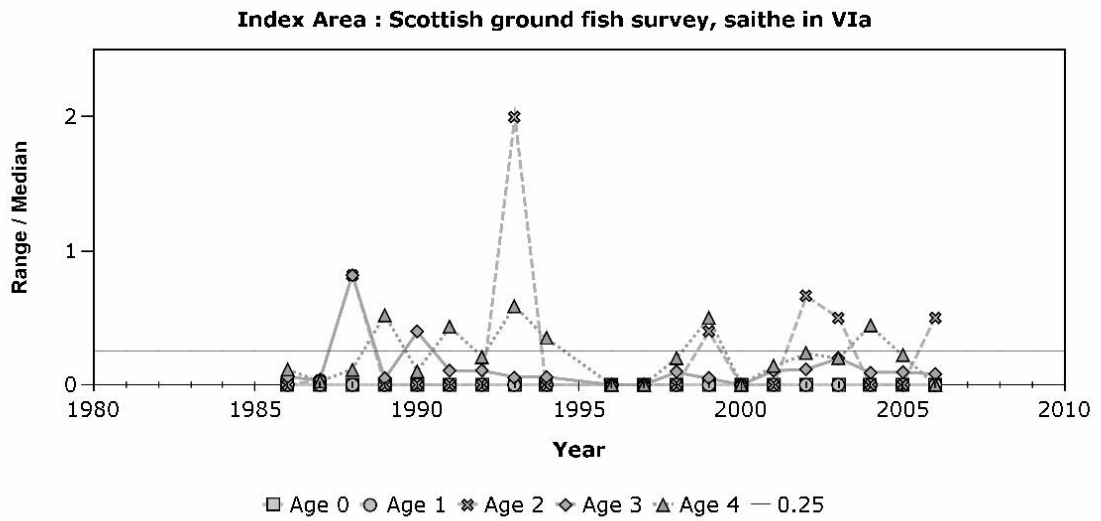
Survey: North Sea International Bottom Trawl Survey (NSIBTS)—Quarter: 4 Continued



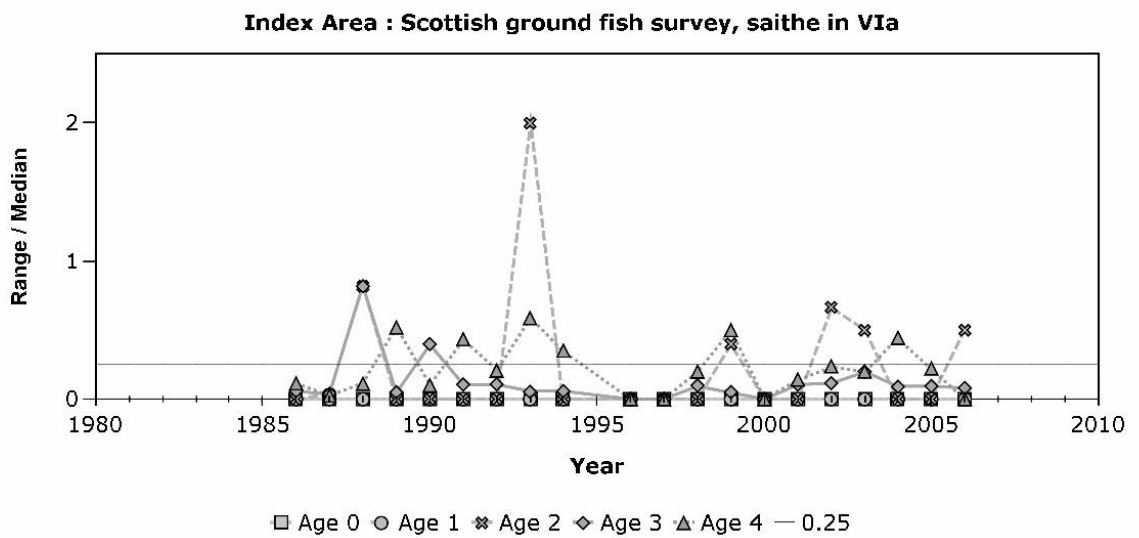
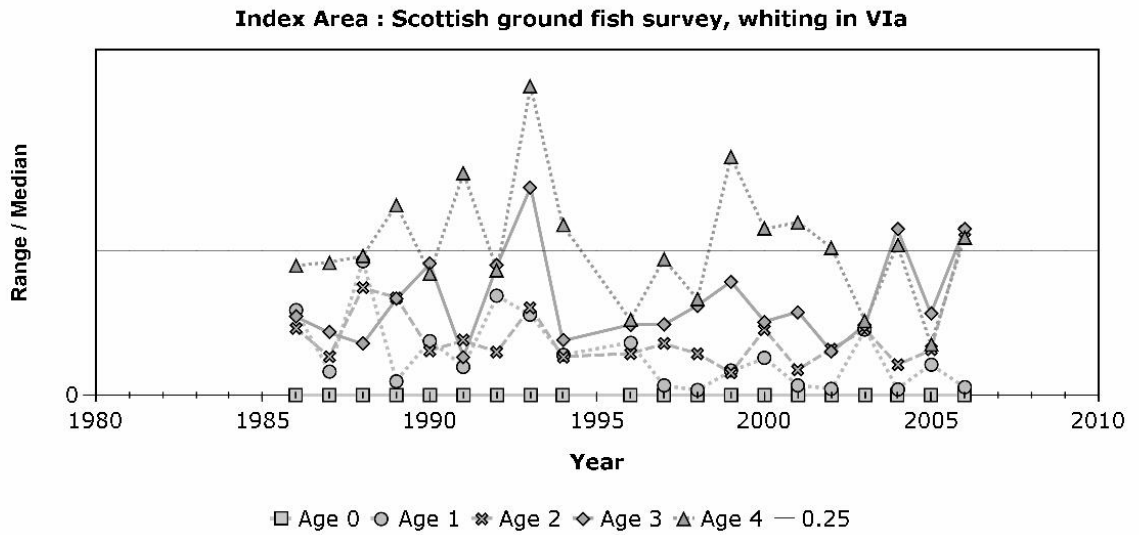
Survey: North Sea International Bottom Trawl Survey (NSIBTS)—Quarter: 4 Continued



Survey: Scottish Groundfish—Quarter: 1



Survey: Scottish Groundfish—Quarter: 1 Continued



Survey: Scottish Groundfish—Quarter: 4

