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

ICES CIEM

ICES CM 2004/ACFM:12 Advisory Committee on Fishery Management

Report of the
Workshop on Sampling and Calculation Methodology for Fisheries Data (WKSCMFD)

26–30 January 2004
Nantes, France

International Council for the Exploration of the Sea

Conseil International pour l'Exploration de la Mer

Palægade 2–4 DK–1261 Copenhagen K Denmark
Telephone + 45 33 15 42 25 · Telefax +45 33 93 42 15
www.ices.dk · info@ices.dk

TABLE OF CONTENTS

1 GENERAL

1.1 Participants

Eero Aro, Finland	eero.aro@rktl.fi
Ulrich Berth, Germany	ulrich.berth@ior.bfa-fisch.de
Paolo Carpentieri, Italy (part time)	paolo.carpentieri@uniroma1.it
Liz Clarke, UK Scotland	e.d.clarke@marlab.ac.uk
Willem Dekker, Netherlands	willem.dekker@wur.nl
Wim Demaré, Belgium	wim.demare@dvz.be
Michele De Meo, Italy	demeo@irepa.org
Ole Folmer, Denmark	ofo@dfu.min.dk
Maria Hansson, Suede	maria.hansson@fiskeriverket.se
David Hirst, Norway	david.hirst@nr.no
Ernesto Jardim, Portugal	ernesto@ipimar.pt
Ciarran Kelly, Ireland	ciaran.kelly@marine.ie
Georgs Kornilovs, Latvia	georgs_k@latfri.lv
Stéphanie Mahévas, France	Stephanie.Mahevas@ifremer.fr
Cecilia Manzi, Italy	manzi@istat.it
David Maxwell, England	d.l.maxwell@cefas.co.uk
Philippe Moguedet, EU (part time)	Philippe.MOGUEDET@cec.eu.int
Iago Mosqueira, Spain	imosqueira@suk.azti.es
Hildrun Muller, Germany	hildrun.mueller@ior.bfa-fisch.de
Tiit Raid, Estonia	raid@sea.ee
Katja Ringdahl, Suede	katja.ringdahl@fiskeriverket.se
Paz Sanpedro, Spain	paz.sampedro@co.ieo.es
Sylvie Van Iseghem, france (part time)	Sylvie.Van.Iseghem@ifremer.fr
Pavlos Vidoris, Greece	fri@otenet.gr
Joël Vigneau, France (chair)	Joel.Vigneau@ifremer.fr

1.2 Background and terms of reference

In the report of their 2003 meeting, the ICES Planning Group on Commercial Catch, Discards and Biological Sampling (PGCCDBS) noted that:

"The Data Directive requires EU member countries to estimate precision levels for various types of data. Different methods can be implemented to determine precision of a sampling plan. Using Coefficient of Variation or confidence intervals will give different results."

"The standardisation of sampling methodology is linked to the notion of precision level. The beginning of this standardisation must be a complete statistical analysis of the different national programmes. A number of methods can be applied to raise samples to obtain statistical population estimates. The heterogeneity becomes a problem when data are merged."

As a result of the concerns, PGCCDBS proposed to ICES/ACFM that a workshop on sampling and calculation methodology for fisheries data, to be held in Nantes (Fr) in January 2004. At the 91th ICES Annual Science Conference, it was decided that the Workshop on Sampling and Calculation Methodology for Fisheries Data [WKSCMFD] will meet in Nantes, France, from 26-31 January with the following terms of reference :

a) produce guidelines for routine estimation of precision in connection with national sampling programmes;

b) identify data requirements and appropriate sampling strategies and methods (e.g. stratification, mandatory and optional variables, selection of vessels, gears, etc.) to collect fisheries data which fulfil the requirements related to stock assessment;

c) compile information on and review the statistical procedures implemented within the national sampling programs (length, age and other biological parameters);

d) propose methods to estimate precision and design sampling stratification schemes that will minimise bias and maximise precision.

1.3 Executive summary

The quality of scientific advice on management of fish stocks and fisheries depends on the adequacy and cost-effectiveness of the basic data collected. Regulation (EC) N° 1543/2000 establishes a Community framework for the collection and management of data needed to evaluate the situation of the fishery resources and the fisheries sector. A key issue in this Regulation is the balance between the precision acquired and the costs involved. Amongst others, this requires insight in the precision achieved by current programmes, as well as consideration of options for improvement.

In previous years, EC projects such as study contract 94/13, FIEFA, SAMFISH and EMAS have compiled information and worked on the subject of sampling precision. The WKSCMFD is the first workshop dealing with the problem of precision related to the numerous biological parameters collected within the Regulation (EC) N°1639/2001 at the scale of Europe and for any stocks where information is collected. In this context, and notwithstanding the terms of reference, a large number of issues were raised by the participants. We then have focussed our attention on the description of the methods usable to fulfil the Regulation and explained their domain of application. Besides this important chapter, summary information on the national programmes has been compiled with regards to stratification and statistical methods used. Important subjects have not been completely addressed such as sampling design and full description of what an exploratory analysis should be. The group advises the reader to look at the working documents in annex where some examples of studies are given.

25 participants from 18 countries attended this meeting. In all the report, the scope has been made to understand and find fully agreement on the implementation of the Regulation concerning the notion of precision. In that way, clear statements on the following points have been addressed:

- Stratification
- Parameters to estimate by module
- Level of precision to achieve defined in terms of coefficient of variation

The calculation of the precision of an estimate is not the only goal to achieve. A low CV does not guarantee an accurate estimate of the "true" parameter value. Further discussion on this issue can be found in this report.

There are a few steps to consider before coming to the calculation of the precision. These steps are described in the guidelines section but do not represent a "recipe book". Expert knowledge, statistical tools and feed-back from the users will always be necessary to build the optimal sampling scheme.

1.4 General considerations

STECF/SGRN in July 2003 has proposed an alternative approach to introduce the notion of precision level in the catch-at-length and catch-at-age sampling to introduce a scientific basis in the definition of sampling intensities. The group agrees with STECF/SGRN that the estimation of precision and the knowledge of the strength and weakness of a sampling scheme is highly desirable. The group has stressed the need of strong relation between the sampling specialists and the users for the definition of precision targets. The precision obtained given a sampling intensity can be highly variable between stocks and dependant on the disaggregation level requested (e.g. fleet-based data, …).

Ultimately, scientific advice on stock and fisheries management is based on (trends in) stock indicators such as spawning stock biomass, fishing mortality, recruitment, and exploitation patterns. The quality of the advice depends, amongst others, on the amount and quality of the underlying data, and the way they have been used in the analyses. Given the currently available data, the precision of the calculated stock indicators can be estimated (e.g. by bootstrap methods), and data collection programmes adapted where the achieved precision is inadequate. However, the number of stocks for which precision levels of stock indicators have been calculated is currently extremely limited, and does not warrant the derivation of rules-of-thumb on the relation between precision and the size of data collection programmes in general. Moreover, the relation between different sources of variation and ultimate stock indicators is not well understood. Analytical modelling of these relations is complex, due to the many practical complications in field sampling programmes, and has not covered the whole process from field sampling, through (national) aggregation, up to the final stock assessment.

Regulation (EC) No 1639/2001 addresses the problem of defining adequate precision levels from the other side. Instead of evaluating the acquired precision of the ongoing samplings, the required precision levels are defined, from which the intensity of field sampling programmes can be derived, in principle. To this end, three predefined levels of precision are listed (in Annex.B.4), which are discussed in more detail in Section 2.4.

Given the uncertain theoretical relation between primary field data and ultimate assessment results (SSB, F, R), the predefined precision requirements are applied to intermediate results, such as length and weight at age. Implicitly, it is assumed that imposing precision requirements on these intermediate results is a surrogate for the ultimate requirements, more or less achieving the same objective. Given that the relationship between these surrogates and the ultimate stock indicators is unknown, there is no way to test this (implicit) assumption. Using these surrogates raises two problems. First, it should be clear how the precision requirement applies to an inherently multidimensional variable, such as catch-at-age. It seems reasonable to assume that the requirement needs to apply to each age group separately, but mutual dependencies between estimates might complicate the matter. Secondly, most data collection programmes use stratified sampling techniques, primarily by country, but additionally by fleet, gear, quarter, area, etc. Precision levels can be calculated for each stratum (e.g. each country), but it is the precision of the aggregate of all strata in relation to the stock indicators that is of ultimate interest.

2 THE NOTION OF PRECISION IN THE REGULATION (EC) 1639/2001

2.1 Spirit of the regulation

Regulation (EC) No 1543/2000 establishes a Community framework for the collection and management of data needed to evaluate the status of fishery resources and the fisheries sector. This regulation stipulates that Member States set up national programmes for the collection and management of fisheries data in accordance with Regulation (EC) No 1639/2001. STECF SGRN noted in March 2002 that for a number of stocks the level of sampling (for age and length) specified would be completely inadequate to derive age or length distributions with acceptable levels of precision. STECF SGRN suggested (July 2003) that instead of trying to re-define sampling levels for all stocks in Appendix XV, to propose an alternative approach to length and age sampling. This approach would define targets based on precision levels with a fall-back option based on MP levels, if the target precision cannot be reached. The WKSCMFD agreed with SGRN that a statistical approach should be used to define sampling effort for age and length. WKSCMFD also notes that the Module G Surveys (EC Reg. 1639/2001) does not specify any parameters or precision to be estimated.

2.2 Pilot surveys

The first point of module B deals with pilot surveys. When it is not possible to define quantitative targets for sampling programmes, neither in terms of precision levels, nor in terms of sample size, pilot surveys in the statistical sense will be established. Pilot surveys are designed to evaluate the importance of the problem, i.e. quantify the variability of the parameter in order to design the sampling scheme to estimate this parameter without bias and quantify the sampling effort to reach a target precision. If no information at all is available on this parameter, a pilot survey will be designed with respect to time, space and if possible fishing gears to cover main sources of variations. If information is available on the parameter (time or space overall distribution or fishing gears), a pilot survey can use this information to focus the sampling effort within the known distribution. In any case, it should be possible at the end of the pilot survey to use statistical tools developed in chapter 6 to design the sampling strategies and it should be possible to estimate the sampling effort needed to achieve a precision goal.

2.3 Stratification in sampling

Presently the EU regulation states that the sampling strategies must be at least as efficient as Simple Random Sampling and that such sampling strategies must be described within the corresponding National Programmes. The necessary dis-aggregation levels are specified in regulation Appendix XV as well as the basic stratification and the sampling intensities.

Simple random sampling, as the name implies, is the simplest sampling design. However, the design and process to obtain simple random sample is not easy. In most cases we can improve and overcome these difficulties in simple random sampling design by using stratification.

It may be possible to divide a heterogeneous population into sub populations each of which is internally as homogenous as possible. The benefit of stratified sampling is that if each stratum is homogenous, a precise estimate of any stratum parameter can be obtained from a small sample in that stratum. These estimates can then be combined into a precise estimate for the whole population. The basis behind stratification is to avoid bias and increase precision.

The workshop noted that almost all established national sampling programmes are based on stratified sampling. This stratification is connected to all modules and to catch and landings monitoring (module B) within the market or sea sampling programmes and stratification is based on fishing technique, space and/or time as a minimum.

Stratification for estimation of biological parameters requires specific local knowledge of fish biology in the area and species of interest. Stratification for estimation of parameters related to commercial catches requires additional knowledge about the fisheries. With respect to stratification by gear, the nature of the fleet will often provide a natural stratification, and additionally, this provides information for the implementation of technical measures.

The regulation 1639/2001 gives general rules to stratify sampling for catch composition (by length, age and species), depending on the fish stock in question: stratification is by some combination of time (month, quarter or annual), gear (total or fleet), and space (rectangle, division or area).

In the sampling programmes, the bias and precision of the resulting catch composition estimates depends on things such as:

- number of strata
- sampling effort per stratum
- method of selecting samples
- variability in the data (within and between strata)
- estimator used

Strata should be chosen on the basis of scientific knowledge, and it may be that the situation requires many strata. However, when the number of strata is large with respect to the sampling effort, then the sampling scheme is said to be over-stratified. This leads to estimation problems, firstly because un-sampled strata need to be accounted for in some way and secondly because small sample sizes can result in inaccurate estimates of precision. An example of this is the long-running Scottish discard sampling scheme (Anon. 2003, Stratoudakis et al, 1999).

A well balanced sampling design will always give the best estimates. Hirst et al. 2003 (and WD1 this workshop) have shown that using modelling methodology in analysing catch-at-age data, can deal with overstratification at the cost of additional assumption. This Bayesian hierarchical model (Hirst et al. 2003) enabled them to obtain estimates of the catch-at-age with appropriate uncertainty, and also to provide advice on how best to sample data in the future.

2.4 Link between CI and CV

In the Regulation confidence levels are defined in relative terms with respect to the parameter concerned, and therefore effectively defines precision levels as an acceptable coefficient of variation CV, the standard deviation (of the estimated mean) divided by the mean. It is implicitly assumed that confidence intervals are symmetrical around the best estimate. This need not be true, as for instance in log-normally distributed statistics. Moreover, the use of CV's applies to any statistical distribution for which mean and standard deviation can be calculated, but in practice, is easily understood to imply a Normal distribution (e.g. a 5 % error rate conforming to a confidence interval of 1.96 times the standard deviation). Most estimates being the product of many unknown sources of variation, the Normal distribution will often fit reasonably well in practice, but there is no theoretical justification, and several observed statistics certainly do not fit. In addition to the issue of statistical variation, potential bias in parameter estimation procedures affects the overall precision too. Unrepresentative sampling can result from erroneous stratification, from undocumented landings etc. Unlike the statistical variation, bias in the sampling design cannot be detected from the data, and therefore inherently requires local knowledge of the field.

Let X be the parameter of interest in the whole population, let $\hat{X}$ be the estimate of X obtained from the data and let $\hat{\mathrm{Var}}(\hat{X})$ be the estimate of the variance of $\hat{X}$.

The estimate of the coefficient of variation (CV) of $\hat{X}$ is defined as: $\hat{\mathrm{CV}}(\hat{X}) = \frac{\sqrt{\hat{\mathrm{Var}}(\hat{X})}}{\hat{X}}$

If we suppose that $\hat{X}$ is normally distributed, then the 95% confidence interval for $\hat{X}$ is given by $\hat{X} \pm t(1-\frac{\alpha}{2};n-1)\sqrt{\hat{\mathrm{Var}}(\hat{X})}$.

If we suppose the sample size is not too low (>20), then $t(1-\frac{\alpha}{2};n-1) \approx 2$, and the 95% confidence interval is given by $\hat{X} \pm 2\sqrt{\hat{\mathrm{Var}}(\hat{X})}$

In the Regulation, required precision levels are defined in relative terms to the parameter of interest (5, 10 or 25 % respectively). Expressing the width of the last equation in relative terms (that is: dividing by $\hat{X}$ and subtracting 100 %) yields

95% confidence interval for $\hat{X}$ in relative terms is $\pm \frac{2\sqrt{\hat{\mathrm{Var}}(\hat{X})}}{\hat{X}} = \pm 2\,\hat{\mathrm{CV}}(\hat{X})$

The precision levels listed in the Regulation thus imply the following for a parameter with an approximate Normal distribution and a reasonable sample size (n>20).

a) level 1: level making it possible to estimate a parameter with precision of plus or minus 25% for a 95 % confidence interval, implies that

   the estimated CV of the parameter is (at most) 12.5%

b) level 2: level making it possible to estimate a parameter with precision of plus or minus 10% for a 95 % confidence interval, implies that

   the estimated CV of the parameter is (at most) 5%

c) level 3: level making it possible to estimate a parameter with precision of plus or minus 5% for a 95 % confidence interval, implies that

   the estimated CV of the parameter is (at most) 2.5%

Note that we are interested in the measuring the precision of the estimate $\hat{X}$, not of the original data. For example, suppose the parameter to be estimated is the mean length at age 2 in the catch, and let x_i denote the observed lengths at age 2 in a sample of size *n*. Then $\hat{X} = \bar{x}$, and $\hat{\mathrm{Var}}(\hat{X}) = \hat{\mathrm{Var}}(\bar{x}) = \frac{1}{n}\hat{\mathrm{Var}}(x)$ and

$$\hat{\mathrm{CV}}(\bar{x}) = \frac{\sqrt{\hat{\mathrm{Var}}(\bar{x})}}{\bar{x}} = \frac{\sqrt{\frac{1}{n}\hat{\mathrm{Var}}(x)}}{\bar{x}} = \frac{1}{\sqrt{n}}\frac{\sqrt{\hat{\mathrm{Var}}(x)}}{\bar{x}} = \frac{1}{\sqrt{n}}\hat{\mathrm{CV}}(x)$$

A numerical example of this is as follows: 2 year old herring in some region of the Baltic had an average length of 21.09 cm with a standard deviation of 1.92 cm.. Since this measurement was based on *n*=67, the mean length of this age group has a standard error of $1.92/\sqrt{67} = 0.24$, and the 95 % confidence interval for the mean thus reads: (21.09-2*0.24) to (21.09+2*0.24), which is between 20.61 and 21.57 cm. This corresponds to a CV for the mean of approximately 1% (0.24/21.09).

2.5 Parameters to estimate in Modules H and I

Various parameters are to be estimated in module H and I, each of them being very specific. The precision usually applies to a scalar type estimator. When the estimator is of a multi-elements (vector) type, the Regulation states that precision must be calculated for each element of the vector corresponding to specifically defined criteria. Table 2.1 gives an overview of all the information required in module H and I to fulfil the Regulation.

Table 2.1. Overview of the parameters to be estimated according to Modules H and I of the EU Reg. 1639/2001, their dimension and the level of precision that is required.

Parameter	Dimension	Required level of precision [1]
Module H		
Landings		
Length distribution	vector	not defined
Age distribution	vector	not defined
Discards		
Length distribution	vector	not defined
Age distribution	vector	not defined
Recreational fisheries		
Catch in weight	scalar	not defined
Catch in number	scalar	not defined
Module I		
Length at age	vector	3 for stocks that can be aged, all ages accounting at least 95% of landings 2 for the others, ages accounting at least 90% of landings
Weight at age	vector	3 for stocks that can be aged, all ages accounting at least 95% of landings 2 for the others, ages accounting at least 90% of landings
Maturity at length²	vector	3 within the 20% and 90% limits of mature fish
Maturity at age²	vector	3 within the 20% and 90% limits of mature fish
Fecundity at length²	vector	3 within the 20% and 90% limits of mature fish
Fecundity at age²	vector	3 within the 20% and 90% limits of mature fish
Species composition in the landings	scalar	1
Sex ratio at age²	vector	3 for all ages accounting at least 95% of landings
Sex ratio at length²	vector	3 for all lengths accounting at least 95% of landings

[1] As defined in the EU regulation 1639/2001

² For maturity, fecundity and sex ratios reference can be made to age **or** length (for more detailed information see Module I (b) (ii))

3 TOR A – GUIDELINES FOR ROUTINE ESTIMATION OF PRECISION

1. Which variable do you have to estimate? Is it a scalar or a vector?

2. What is the sampling design to achieve this purpose? (Simple random, stratified, stratified multi-stage...); What is your sampling unit?; How many samples do you have per cell?

3. Provide an exploratory analysis to perform a diagnostic of your sampling

- **Simple random design :**

   i. Did you meet any problem to collect your sample with regards to your sampling design? If yes, which impact would it have on the results?

   ii. **Bias :** Are the samples representative of the population sampled? If not, you must take care with the interpretation of your results.

   iii. **Precision :** are there samples outliers? Should you remove them from the analysis?

   iv. **Design :** could you point out some patterns correlated to specific variables (such as space, season, gear, fleet, ...)?

♦ **<u>Stratified sampling design :</u>**

i. Do you meet any problem collecting your sample with regards to your sampling design? If yes, which impact would it have on the results?

ii. Precision :

*Within each stratum :*are there samples outliers? Should you remove them from the analysis? Is the sampling effort appropriate relative to the population level within this stratum?

Between each stratum : is the variability of the observations similar or larger between stratum than within stratum? Could you merge your stratum?

iii. **Design :** could you pointed out some patterns correlated to specific variables (such as space, season, gear, fleet, ...) within the stratum?

4. Which methods are applicable to calculate the coefficient of variation of the selected variable? See Table 6.1 in Section 6 comparing the methods. Select one :

- If you have only one strata and you estimate a scalar variable, you can use the analytical method.
- If you have more than around ten samples in each cell, then you can use the bootstrap method.
- A modelling approach is usable in both of the above situations.

4 TOR B – APPROPRIATE SAMPLING METHODS

The important question of sampling strategy has not been studied here. The improvement of a sampling scheme can only be done after primary analysis of the data and the coefficients of variation. In the guidelines chapter the reader will find advice to analyse the data in the scope of calculating the precision and to investigate the appropriateness of the sampling design.

A review of exploratory analysis tools of sampling design needs to be done.

Based on information contained in the tables of chapter 5 and with appropriate exploratory tools, sampling data should be analysed. This analysis should point out the source of potential bias in the current sampling design and ways to improve the precision.

These important issues need to be addressed specifically to another workshop.

5 TOR C – COMPILE AND REVIEW STATISTICAL PROCEDURES IMPLEMENTED WITHIN THE NATIONAL SAMPLING PROGRAMS

For member states within the EU, regulation EC 1639/2001 states levels of sampling intensities and targets of precision within the sampling programmes. Statistical procedures and sampling programme design within the different countries is however, often stock-dependant making a detailed review a scope too big to cover within this meeting. Instead the meeting decided to produce summary tables with the intention of getting a simple overview of methods in use and different approaches to sampling. The tables will also allow us to analyse similarities and discrepancies between countries and regional areas in the future. To our knowledge this is the first attempt to summarise information in this way.

5.1 Overview of methods used for precision calculation

The precision targets of regulation EC 1639/2001 give rise to a huge shift in the way countries treat data and have thereby raised a large number of questions regarding the methods to use. Many member states within the EU are now in a process of changing the statistical treatment of data collected within their national programme, but this process is slow compared to the timetable within the regulation. Table 1 summarises where we, as a community, are in this process as well as the methods in use. These methods are described in detail in Chapter 6.

5.2 Overview of sampling strategies, stratification and location

The choice of method for calculating precision is to a certain degree dependant on the choice of sampling strategy. Furthermore, risk of introducing bias is heavily related to design of the sampling programme. As noted underneath ToR b improvement of a sampling scheme can only be done after primary analysis of the data and the coefficients of variation. To start and understand this work in a wider context, knowledge of sampling strategies of today is of importance. Tables 2a, 2b & 2c summarize the sampling strategies of today regarding principal methods of sampling (Table 2a), stratification (Table 2b) and location of the sampling (Table 2c).

Table 1 Overview of 2003 National Programs in respect to calculation of precision levels.

	Stocks included in Appendix XV (EC 1639/2001)	Be	De	UK En	Est	Fra	GFR	Gr	Ire	Ita	Lat	NL	Pt	Fin	UK Sc	Sp	Sw
1.Biological sampling of landings: composition by age and length (aged based assessment) / length (length based assessment)	1. No of stocks sampled (total)	23	39	60[a]	10	42[*]	43	14	44	174	13[1]	25	40	16	44	38	17
	1.1.1 No of stocks sampled (aged based assessment)	18	28	35	10	23	16	8	32	*	4	10	12	14	22	19	12
	1.1.2 No of stocks sampled (length based assessment)	5	11	18	0	19	0	6	12		1	1	7	2	2	19	3
	1.2.1 No of stocks for which precision is reported (aged based assessment)	0	0	12[b]	0	4	12	8	0*		0	7	6	0	0	0	1
	1.2.2 No of stocks for which precision is reported (length based assessment)	0	0	0	0	2	0	6	0		0	0	0	0	0	0	0
	1.3.1 Methods[†] for calculation of precision (aged based assessment)	-	-	A	-	A+B	A	-	-		-	B	A+B	-	-	A+B	B
	1.3.2 Methods[†] for calculation of precision (length based assessment)	-	-	-	-	A+B	-	-	-		-	-	-	-	-	-	-
2. Discard sampling	2.1 No of target stocks sampled	13	20	79	0	18	47	35	30		1[2]	0	17	2	30	16	8
	2.2 No of target stocks for which precision is reported	0	0	17	0	9	11	0	0		0	0	17	0	0	6	0
	2.3 Methods[†] for calculation of precision	-	-	A	-	A	A	0	-		-	-	A+B	0	B	GLM	-
3. Other biological parameters (SMALK)	3.1 Number of stocks sampled	2	24	45[c]	10	15	21	14	0[†]		13[3]	21	10	14	6	59	16
	3.2 Number of stocks for which precision is reported	0	0	[c]	0	15	11	0	0		0	0	10	0	0	0	0
	3.3 Methods[†] for calculation of precision	-	-	A	-	A	A	-	-		-	-	A+B	-	-	-	-

4. Comments	UK En :(a): assessment definition of stock used or definition by species and ICES division or subdivision if no assessment. (b): precision routinely supplied to assessments for 12 species but is available for all 35 age based stocks (c): Sex, weight and maturity. Precision reported for sex based ALKs, growth curves and maturity modelling. Fr : Tropical water stocks excluded GFR : No of Stocks indicated for which precision calculation is planned but not yet done for 2003 data Ire: * Precision work to be commenced in 2004 [†]No. of stocks to be determined following survey in Q1 2004 It Data not used in the purpose of assessment Sp : Analytical and bootstrap to be implemented in National database in 2004 Lat : [1] this includes 5 stocks (5 species) for which analytical assesssment is performed, 4 species from Appendix XIII, and 3 species which are not listed in EC Regulation; [2] discards of cod, besides seal damaged salmon is being recorded in recent years; [3] for 3 species only weight additionally to length and age is recorded.

[†]A: analytical, B: bootstrapping, BM: Bayesian heirachical modelling, O: other (explain in comments)

Table 2A Overview of 2003 sampling strategies for age/length composition of commercial landings.

	Stocks included in Appendix XV (EC 1639/2001)	Bel	De	UK En	Est	Fra	GFR	Gre	Ire	Ita	Lat	NL	Pt	Fin	UK Sc	Sp	Sw
1.	No of stocks sampled (total)	23	39	60	10	42	52	14	44	174	13	21	40	16	44	38	17
2.	No of stocks sampled for age composition	18	28	35	10	23	16	8	32	85	12	15	21	14	22	19	13
3.	No of stocks sampled with ALK method†	18	28	35	0	23	16	8	32	0	5	3	21	10	22	19	0
3.1	No of stocks with lengths sampled from unsorted landings	12	20	9	0	9	49	6	28	174	5	11	40	14	11	0	0
	No of stocks with lengths sampled from market categories	11	15	35	0	33	3	2	16*	0	0	3	0	0	11	19	0
3.2	No of stocks with otoliths sampled independently of length distribution	18	16	35	0	16	0	0	0	85	4	3	21	8	0	19	0
	No of stocks with otoliths sampled from length distribution	0	14	35	0	7	0	0	22	23	1	12	0	8	22	0	0
3.3	No of stocks with otoliths sampled from unsorted landings	7	20	9	0	2	16	6	6	85	5	9	0	10	11	0	0
	No of stocks with otoliths sampled from market categories	11	15	35	0	21	1	2	26	0	0	6	0	0	11	19	0
3.4	No of stocks with otolith samples stratified by length	7	28	35	0	19	16	8	32	23	1	3	21	8	22	19	0
	No of stocks with random otolith samples	11	0	0	10	3	0	0	0	85	4	12	0	8	0	0	0
4.	No of stocks sampled with direct method‡	0	0	0	10	0	0	8	0	174	7	12	0	0	0	0	13
4.1	No of stocks with fish sampled from unsorted landings	0	0	0	10	0	0	6	28	174	7	9	0	14	0	0	6
	No of stocks with fish sampled from market categories	0	0	0	0	0	0	2	16	0	0	12	0	0	0	19	7
5.	No of stocks sampled with other methods – described in Comments below	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4 *
6.	Comments	Ire : * Sometimes from market categories, sometimes from unsorted landings															

†ALK method: Ages and lengths are sampled for each stratum (either independently or at the same time) to obtain a length distribution and an ALK, which are combined to estimate the proportion by age in the landings.

‡Direct method: A random sample of fish is taken from the landings (either unsorted or by size category) for each stratum, and used to estimate the proportion by age in the landings directly.

Table 2B Overview of 2003 stratification level of sampling strategies for age/length compositions of commercial landings.

	Stocks included in Appendix XV (EC 1639/2001)	Bel	De	UK En	Est	Fra	GFR	Gr	Ire	Ita	Lat	NL	Pt	Fin	UK Sc	Sp	Sw
1.	No of stocks sampled (in total)	23	34	60	10	42	52	14	44	174	13	25	40	16	44	38	17
2.	No of stocks stratified in time (total)	23	34	60	10	33	49	13	44	0	13	25	40	16	30	38	17
	No of stocks stratified by quarter	22	34	10	10	29	49	0	35	0	8	25	0	16	0	38	16
	No of stocks stratified by time unit shorter than quarter	1	0	60[a]	0	4	0	0	9	0	3	3	40	0	30	0	1
	No of stocks stratified by time unit longer than quarter	0	0	0	0	0	0	13	0	174	2	0	0	0	0	0	0
3.	No of stocks stratified in space (total)	23	34	60	10	15	49	14	44	174	6	0	40	16	30	38	17
	No of stocks stratified by ICES/FAO division	23	12	0	0	0	46	0	5	174		21	0	0	0	0	0
	No of stocks stratified by ICES subdivision / GFCM geographical sub-area	23	22	60	10	0	3	14	39	174	4	21	24	16	0	31	17
	No of stocks stratified by space unit smaller than ICES subdivision / GFCM geographical sub-area	1	0	5	0	15	0	0	9	0	2	0	24	0	30	7	0
4.	No of stocks stratified by gear (total)	23	6	45	10	21	2	6	44	52	5	0	40	10	30	38	10
5	Overstratification* : No of stocks for which this has not been analysed	23	34	0[a]	10	0	44	8	0	0	0	0	40	16	0	-	17
	No of stocks for which this is experienced to be a problem	0	-	~6	0	9	0		7	0	-		-	-	0	-	-
	No of stocks for which this is experienced NOT to be a problem	23	-	~47	0	33	0	14	37	0	-	21	-	-	0	-	-
6.	Comments	UK En : (a): sample collection stratified by coastal region and monitored against monthly targets to ensure even coverage. Analysis stratified by quarter															

- *Overstratification= strata with landing for which there is no corresponding sufficient catch-sampling

Table 2C Overview of 2003 locations of sampling of commercial landings.

	Stocks included in Appendix XV (EC 1639/2001)	Bel	DK	UK En	Est	Fra	GFR	Gr	Ire	Ita	Lat	NL	Pt	Fin	UK Sc	Sp	Sw
1.Length distributions	No of stocks sampled by observer on board	13	25	0	5	1	47	11	31[1]	150	6	0	16	6		10	2
	No of stocks sampled at port by ship	0	15	60	10	41	6	0	44	174	12[1]	0	24	16	44	28	7
	No of stocks sampled at market/auction	15	16	60	0	37	0	14	22[2]	-	-	14	24	-	0	0	8
	No of stocks sampled by coastguards	0	16	0	0	0	0	0	0	-	-	0	0	-	0	0	4
	No of stocks sampled other (explain in comments)	0	0	0	0	0	0	0	22	-	-	7	0	-	0	0	-
2.Otolith origin	No of stocks sampled by observer on board	7	15	0	5	0	14	11	34[3]	85	6	0	11	6	0	4	0
	No of stocks sampled at port by ship	0	17	35	10	23	4	0	7	85	11[1]	0	10	16	22	14	5
	No of stocks sampled at market/auction	11	9	35	0	21	0	14	26	-	-	8	0	-	0	14	8
	No of stocks sampled on surveys	2	23	*	10	14	0	0	0[4]	-	7	0	10	4	0	14	0
	No of stocks sampled by coastguards	0	14	0	0	0	0	0	0	-	-	0	0	-	0	0	4
	No of stocks sampled other (explain in comments)	0	0	0	0	0	0	0	6[5]	-	-	7	0	-	0	0	-
3.Other Biological parameters	No of stocks sampled by observer on board	0	5	0	5	0	18	11	-	-	3	0	11	6	0	10	1
	No of stocks sampled at port by ship	0	21	45	10	9	3	0	-	-	10	0	10	16	6	49	7
	No of stocks sampled at market/auction	2	9	35	0	9	0	14	-	-	-	12	0	-	0	49	0
	No of stocks sampled on surveys	0	6	*	10	8	1	0	-	-	3	2	10	4	6	49	12
	No of stocks sampled by coastguards	0	0	0	0	0	0	0	0	-	-	0	0	-	0	-	0
	No of stocks sampled other (explain in comments)	0	0	0	0	0	0	0	0	-	-	7	0	-	0	-	0

Comments	UK En : * Research surveys are used but not for commercial landings data Ire : [1]Including smapling for discards for stocks only specified in the Data collection Regulation [2]Sampled at factory and at boat [3]Includes all stocks sampled for discards from fleet based programme, i.e. includes stocks not specified in the Data collection Regulation [4]Survey ALK's not used for commercial data [5]All pelagic stocks sampled at factory NL : Samples taken by crew onboard Lat : [1] for 5 stocks (4 species) besides prevailing sampling at port also additional sampling at sea is performed

6 TOR D – PROPOSE METHODS TO ESTIMATE PRECISION

6.1 Introduction

This section addresses the issue of how precision should be estimated. Three methods are suggested, the advantages and disadvantages discussed and the methods compared. The most important factor in deciding which method to use is the sampling design. No method of analysis can provide good estimates of variables if the sampling scheme is poor, or there are a large number of missing observations. Some methods are however better at analysing unbalanced schemes.

A second important issue is the use to which the precision estimate will be put. If precision is required for a vector such as catch-at-age the covariance between estimates at different ages is needed for almost every purpose, eg use in stock projection, or assessing the uncertainty in VPA.

6.2 Analytical approach

6.2.1 Description

In this section, we present the analytical formulation of the catches at age variance with application in a simple random design and a stratified sampling design.

Notation

We use the following notation:

$\hat{N}$: the total landings estimator (in number)

$\hat{N}_i$: the landings estimator at age *i* (in number)

$\hat{N}_j$: the landings estimator at length *j* (in number)

$\hat{p}_i$: the estimator of age *i* proportion of landings

$\hat{p}_{ji}$: the estimator of age *i* at length *j* proportion of landings

$\hat{N}_{ij}$: the estimator of the j^{th} length class landings in age *i*

The estimator of the landings at age *i* is

$$\hat{N}_i = \hat{N}\hat{p}_i$$

and its variance estimator decomposed into three elements is:

$$Var\left(\hat{N}_i\right) = Var\left(\hat{N}\hat{p}_i\right) = \hat{p}_i^2 Var\left(\hat{N}\right) + \hat{N}^2 Var\left(\hat{p}_i\right) + Var\left(\hat{N}\right)Var\left(\hat{p}_i\right) = V1 + V2 + V3 \qquad (1)$$

This expression shows the great importance of the precision of the age-length key compared to the precision in the estimation of the landings. The second element ***V2*** of the variance will often be the larger component of $Var\,\hat{N}_i$, since it is a function of the squared landings in number.

An analytical formulation of the aged landings variance estimator is available (e.g. in Quinn and Deriso 1999) for the case of a random sampling design (without any stratification).
First, we explicit the variance of the total landings and the variance of the proportion of landings at age *i*. The total landings estimator is the result of the length sampling, while the estimator of the proportion of landings at age i is provided by the age sampling. Thus the total landings estimator is

$$\hat{N} = \sum_j \hat{N}_j$$

and the estimator of the proportion of landings for age i is

$$\hat{p}_i = \frac{\sum_j \hat{N}_{ji}}{\hat{N}} = \frac{\sum_j \hat{N}_j \hat{p}_{ji}}{\sum_j \hat{N}_i}$$

These estimators are calculated using the estimator of the landings at length *j*, $\hat{N}_j$, and the estimator of the proportion of landings of age *i* in the length class *j*, $\hat{p}_{ji}$, estimated from the length sampling and the age sampling respectively.

Simple stratified sampling

For each defined stratum, a sample of the landings is collected and each individual of the sample is then assigned to a length group.

Some additional notation for the stratified sampling is as follows:

k : k^{th} stratum of the strata

K : the number of stratum of the strata

W_k : total landings of the k^{th} stratum in weight

n_k : samples number of the k^{th} stratum

v : the v^{th} sample

W_{kv} : the v^{th} sample weight of the k^{th} stratum

J : the number of length class

I : the number of age group

j : j^{th} length class

i : i^{th} age group

N_{jkv} : the number of fish belonging to the j^{th} length class in sample v

W_{jkv} : the weigth of fishes belonging to the j^{th} length class in sample v

M : the number of individual used to construct the age-length key

m_j : the number of individual of length j of the age-length key

pl_j : the proportion of individuals of length j of the age-length key

q_{ji} : the proportion of individuals of length j and age i of the age-length key

The variance estimator of the total landings is the following.

$$Var(\hat{N}) = \sum_j Var(\hat{N}_j) + \sum_{j \neq j'} Cov(\hat{N}_j, \hat{N}_{j'})$$

We assume that $\sum_{j \neq j'} Cov(\hat{N}_j, \hat{N}_{j'}) = 0$, for all *(j,j')*, to simplify the calculation of the variance estimate.

Variance of landings at length (in number)

From the sampling design of the landings at length, the estimator of the landings at length *j* can be decomposed as follows,

$$\hat{N}_j = \sum_{k=1}^{K} \frac{W_k}{\sum_{v=1}^{n_k} W_{kv}} \left(\sum_{v=1}^{n_k} N_{jkv} \right) = \sum_{k=1}^{K} W_k \frac{\sum_{v=1}^{n_k} N_{jkv}}{\sum_{v=1}^{n_k} W_{kv}}$$

and the estimator of the variance is

$$Var(\hat{N}_j) = \sum_{k=1}^{K} W_k^2 Var\left(\frac{\sum_{v=1}^{n_k} N_{jkv}}{\sum_{v=1}^{n_k} W_{kv}} \right) \qquad \textbf{(2)}$$

and from Cochran (1977),

$$Var\left(\frac{\sum_{v=1}^{n_k} N_{jkv}}{\sum_{v=1}^{n_k} W_{kv}}\right)=\frac{1-\frac{\sum_{v=1}^{n_k} W_{kv}}{W_k}}{\frac{1}{n_k}\left(\sum_{v=1}^{n_k} W_{kv}\right)^2}\frac{\sum_{v=1}^{n_k}\left(N_{jkv}-\frac{\sum_{v=1}^{n_k} N_{jkv}}{\sum_{v=1}^{n_k} W_{kv}}W_{kv}\right)^2}{n_k-1}$$

This last equation points out that variability of sample sizes, expressed in weight (W_{kv}) and in number (N_{jkv}), would penalized the variance of $\hat{N}_j$. To quantify this penality, we could introduce an average proportion at age.
The estimator of the landings at length *j* would thus be :

$$\hat{N}_j=\sum_{k=1}^{K}\frac{W_k}{\sum_{v=1}^{n_k} W_{kv}}\left(\sum_{v=1}^{n_k} N_{kv}\overline{p}_{jk}\right)$$

An optimal sample size would then be defined as the value of N_{kv} inducing a low variability in p_{jk} .

Variance of the proportion of landings at age
From the sampling design of the age-length key, the estimator of the proportion of landings at age *i* is calculated with the following equation,

$$\hat{p}_i=\sum_{j=1}^{J} q_{ij}\, pl_j$$

With an assumption of proportional allocation, the estimate of the variance of the proportion at age *i* (Kimura 1977, Lai 1987) is

$$Var(\hat{p}_i)=\sum_{j=1}^{J}\frac{q_{ij}\, pl_j^2(1-q_{ij})}{m_j}\frac{pl_j(q_{ij}-pl_j)^2}{M} \qquad \textbf{(3)}$$

Finally, the estimate of coefficient of variation of the landings at age *i* can be calculated putting results of equations **(2)** and **(3)** into equation **(1)**.

6.2.2 Assumptions

First of all, each sample must be representative of the underlying population of the stratum to which it belongs.
The analytical statistic of the variance is derived respecting exactly the sampling scheme.
The sampling design is supposed to be random within each stratum.
The stratification scheme is supposed to split the population into separate (no overlapping) subsets, constituting a partition of the population.
The above formulation can only be applied if the sampling design is either a simple random or a simple stratified one. In a more complex stratified sampling design (for instance a multi-stage one), an analytical formulation might still be possible to derive, but it would probably be very un-attractive and un-generalizable.

6.2.3 Implementation

No important constraints are identified.

6.2.4 Advantages

It is possible to identify and quantify the part of variance due to age and due to length sampling in the total variance.
Statistics can be derived to analyse sampling design.
It is a deterministic method.

6.2.5 Disadvantages

It might be complex to derive an analytical formulation, especially in a complex stratified design. An estimate of the total landings in number can be calculated that will not take into account the covariance between numbers at different ages. The estimator presented does not provide any estimation of the covariance between numbers at different ages.

6.3 Non-Parametric Bootstrap

6.3.1 Description

The aim of the non-parametric bootstrap (Efron 1979) is to provide information on the variability of a statistic, for example, catch numbers-at-age, by resampling from the observed data. New samples, with the same number of observations as the original one, are created by sampling with replacement from the original data. For stratified schemes resampling is carried out independently within each stratum. For each new sample the statistic of interest is then calculated. This gives a bootstrap distribution for the statistic from which its precision is estimated.

A standard reference on resampling methods is Efron and Tibshirani (1993), which gives detailed coverage of a range of topics. Manly (1997) gives an accessible account of the methods with references to biological applications. Patterson et al. (2001) provide an up-to-date review of methods of estimating uncertainty in fish stock assessment, including the bootstrap and jackknife.

6.3.2 Assumptions

In common with the analytical calculations and modelling approaches, the non-parametric bootstrap assumes that the observed data are representative of the underlying population. The non-parametric bootstrap does not require any modelling assumptions, such as normal error distributions to produce estimates of uncertainty.

6.3.3 Implementation

For market sampling data, the bootstrap is implemented by resampling age and length samples separately. These new samples are then processed and raised in the same manner as the original market sampling data to produce estimates of catch numbers-at-age. Therefore, data need to be available at the sample level of recording and not in aggregated form. It is possible to implement bootstrapping on most statistical or data processing systems.

If very few samples are present in the original data, few distinct choices of samples are possible when bootstrapping. Chan and Lee (2001) suggested a different algorithm for small sample bias reduction and based their work on sample sizes less than 10. Therefore, a guideline of at least 10 samples in each cell at the lowest level of stratification is suggested.

Some specific examples for estimating the precision of catch numbers-at-age estimates can be found in: Jardim et al. (2004), Vigneau and Mahevas (2004), the EMAS project (Anonymous 2001) and work from it (O'Brien et al 2001a, 2001b, Simmonds et al 2001), and the SAMFISH project (2000) and work related to it (Maxwell et al 2001). These documents give algorithms describing the procedures used. Two points are highlighted here: the fundamental issue of resampling the correct unit, and the more technical issue of combining bootstrapped age and length samples.

Unit to resample

Knowing which unit to resample comes from knowledge of the country's sampling and analysis scheme. The bootstrap sampling unit must be the same as the independent units in the sampling scheme. For example, when length samples are collected by vessel the unit to resample must be vessel too. If only individual length measurements are resampled then variation between boats will be missed, leading to an underestimate of the variance. For age samples, if the individual otoliths in an age-length key are considered independent observations then resampling otoliths within each length group is reasonable.

Combining age and length bootstrap samples.

Implementing the bootstrap for catch numbers-at-age differs from standard examples as two bootstrap samples, age and length, are generated not one. There is a choice of how to combine them, either within iterations or between all iterations.

Method 1, within iterations.

Iteration 1 age-length key 1 & length distribution 1 => age composition 1
Iteration 2 age-length key 2 & length distribution 2 => age composition 2
…
Iteration n age-length key n & length distribution n => age composition n

Method 2 between all iterations.

		length dist. 1	length dist 2	…	LD r
		⇓	⇓		⇓
Age-length key, 1	=>	age composition 1	age composition 2	…	
Age-length key, 2	=>	.			.
…		.			.
Age-length key, s	=>	.	…	…	age composition rs

Method 1 was used in the majority of the references given earlier. However, if the age and length samples are independent, this choice is not thought to create major differences in the results.

6.3.4 Advantages

Non-parametric bootstrapping can be used for any sampling scheme, it only requires a method of estimating the catch number-at-age that can be repeated. The method makes no assumptions about the statistical distribution of the data. The concept of resampling the original data is simple to explain (although the implementation requires care).

It is easy to combine the outputs of bootstrapping. National estimates resulting from different sampling schemes and analysis procedures can be added and then passed to stock assessment software to judge the effect of variation due to market sampling on assessment results. The bootstrap produces more than just a variance estimate, correlation estimates between different ages and distributions of the catch numbers-at-age estimates are produced. It is also possible to extent the method to calculate the variance of the variance estimate (jack-knife after bootstrap).

6.3.5 Disadvantages

The non-parametric bootstrap needs a suitable number of samples within each cell of the stratification scheme. It may become necessary to use a higher level of aggregation e.g. quarter instead of month. This will give more heterogeneous strata and increase the variance of the estimate, if the stratification at the lower level was effective. As there are no modelling assumptions, the bootstrap does not make use of correlation between levels of each stratum to increase the precision of estimates.

The method requires more computational resources and time than analytical calculations, but this should not be a major constraint nowadays.

6.4 Modelling Approach

6.4.1 Description

A model is able to describe explicitly both (1) the process by which the data is generated, i.e. statistical distributions for the variables and (2) mathematical relationships between variables. For example, in the context of proportions-at-age, one might assume that age samples are multinomial, and that proportion-at-age for some combination of gear and region can be estimated by a linear combination of gear and region effects.

Bayesian hierarchical Modelling
The distinction between the modelling and sampling approaches is more important than that between frequentist and Bayesian modelling. With the Bayesian approach it is easier to make a model that includes all sources of variation (eg age reading errors are harder to include in a frequentist model). It is also easier to combine different types of data (age

only, length only, age-given-length). Implementation can be harder though, and issues such as convergence of MCMC chains and choice of prior distributions arise. These become more important when the model is very complex and the data are sparse. One potential advantage of the Bayesian approach is that prior distributions developed from similar situations can help when data is poor (eg. at the start of a new sampling scheme). An example of a modelling approach to estimating catch at age is in Hirst et al (working document).

6.4.2 Assumptions

The assumptions are the distribution chosen for the data and the mathematical relationship between the variables. Thus the assumptions increase with the complexity of the model.

6.4.3 Implementation

Depends on complexity of model. It is possible to design software for easy use for specific models and sampling designs, but it is difficult to write generic programs that work in all possible situations. Speed of computation is not much of a problem now.

6.4.4 Advantages

The modelling approach can be used in any sampling situation, with missing or unbalanced data less of a problem than in sampling methods. Estimates of uncertainty are correct given model assumptions, parameter estimates can be interesting biologically, and it is possible to simulate from the model to optimise sampling design. A key advantage is that assumptions are explicit. This means that it is clear when assumptions have been made and what they are. With other methods, ad-hoc approaches may rely on hidden assumptions. Models can also give information about the process, and provide a useful exploratory tool. The output (at least from a Bayesian model) is a full distribution of catch at age statistics, including covariance.

6.4.5 Disadvantages

Assumptions are necessary. Some are testable, eg that length|age is log normal, but if the model is complex some are not. For example it is very difficult to test the assumption that logistically transformed proportion parameters are normal, or that residual variances are equal in all cells. Bias can result if the model is wrong. Implementation may require some sophisticated statistical knowledge, and requires model selection, testing and fitting for each application.

6.5 Comparison of modelling and sampling approaches

The most important difference between the two methods is in the effect of missing or undersampled cell. For example suppose we have 4 areas (A) and 2 gears (G), and we want to estimate the mean length of a fish in each cell. The table below gives possible sample sizes in each cell:

	A1	A2	A3	A4
G1	20	1	6	0
G2	0	10	6	6

Note that cells G2A1 and G1A4 have not been sampled. A sampling theory estimate would simply be the mean of the data in each cell independently, and without further assumptions it would not be possible to estimate the mean for G1A4 or G2A1. Also, the mean for G1A2 would simply be equal to the single data value and therefore be very uncertain. The usual approach would be to assume that ‘similar’ cells are in fact identical, eg that G1A4=G2A4, or to abandon one level of stratification eg gear. The ‘hidden’ assumption is that cells are either completely independent (if there is any data), or identical (if there is no data). A particular disadvantage is that the assumptions are completely driven by the distribution of samples. It is also impossible to improve the estimation of cells with little data (eg G1A2) by using information from other cells. The modelling approach would use that fact that the effects of G1 and A2 can be estimated from the rest of the data, with some uncertainty added to account for any interaction.

The main disadvantage of the modelling approach in this situation would be if there was a much larger interaction between G1 and A2 than between any of the other effects. In this case the modelling approach would be biased, and the bias may be a worse problem than the extra uncertainty in the sampling approach. This kind of issue (bias vs variance) may be more serious in more complex situations.

A second illustration:

The sampling approach to the ALK is to use a matrix of observed counts (which may be bootstrapped later). It may look something like this:

	A1	A2
L1	7	2
L2	0	1
L3	2	4
L4	0	6

This table says that 100% of fish of length L2 are age 2, although in reality if the length categories are quite narrow, the age distribution for L2 will be similar to that for L1 and L3. The modelling approach will take this into account, effectively getting an estimate of the age distribution for L2 which utilizes the information in the samples of L1 and L3. This means that the precision of the model based estimate will almost always be higher than that based on sampling theory, because the model uses more information. The model will however be biased if it is wrong, eg if for some (unmodelled) reason 100% of fish of length L2 really are age 2.

	Design-based		Model-based	
	Analitical	Non-parametric bootstrap	Frequentist	Bayesian
Assumptions	Sample representative of the population, sampling scheme unbiased Strata must be a partition of the space.	Resampling unit must be independent.	Distributions and relationships between variables.	
Advantages	Explicit, identify variance due to age and due to length, can derive statistics to analyse sampling design.	Non-parametric, can deal with complex processes, simple concept, estimates covariance.	Explicit, deal with complex situations, id var comps, estimations of uncertainty, parameters can have biological interest, can include expert knowledge.	Idem frequentist model, easier to deal with missing observations, include more complex expert knowledge and different sources of data.
Disadvantages	It becomes extremely complex to apply to more than 1 strata situation, no covariance between ages.	Sensitive to low number of samples in strata which can underestimate variance or produce biased estimates due to merging of strata.	Complex assumptions, requires model testing and fitting, different sampling schemes and stocks may require different models.	Idem frequentist model, more difficult to implement, MCMC convergence problems.
Implementation	Simple	Simple, uses simulations.	Complex.	More complex, uses simulations.
Example (ref)	WD 4, 5, 6, 7 & 8			WD1

Table 6.1 Comparison of methods - Summary

7 FAQ : FINAL ASKED QUESTIONS

At the end of the Workshop, participants were free to raise anonymous questions they felt were important to address. These questions are given below without classification. Some of these points are discussed in this report, some are not but all of them are relevant.

1) What is a good sampling scheme?

2) How can estimates of precision from various countries and fleets be combined?

3) How many samples are necessary to be taken by stratum?

4) What is the effect of changing the number of samples?

5) Can strata be combined?

6) What is the effect of age reading errors?

7) Can I construct only one routine for all stocks?

8) Why post stratification is a problem?

9) How can you stratify sampling if there are multiple purposes for your sampling?

10) Can results of different approaches for the same data be different?

11) Is it possible to create an expert group for support to countries that need help with statistical methods?

12) What does correlation between ages imply for estimating uncertainty?

8 RECOMMENDATIONS

- National programmes should be analysed in term of precision before going to another step.

- There is no recipe, no simple guideline to estimate precision for all stocks and everywhere.

- Precision should be estimated at a stock level.

- A tool need to be developed at the international level to produce estimates of precision.

- A workshop devoted exclusively on sampling design should be organised in the beginning of 2005.

  The terms of reference should be

  a) analyse the results of precision obtained by each country

  b) advise on sampling strategies including stratification and sampling effort

9 REFERENCES

Anonymous (2000). Final report of CFP Study Project 99/009. Improving Sampling of Western and Southern European Atlantic Fisheries, - SAMFISH.

Anonymous (2001). Final report of CFP Study Project 98/075, Evaluation of market sampling strategies for a number of commercially exploited stocks in the North Sea and development of procedures for consistent data storage and retrieval (software) - EMAS.

Anonymous. (2003). The Report of Workshop on Discard Sampling Methodology and Raising Procedures, Copenhagen Denmark 2-4 September 2003, 27 pp.

Chan, K. and S. Lee (2001). "An exact iterated bootstrap algorithm for small-sample bias reduction." Computational Statistic & Data Analysis **36**: p. 1-13.

Cochran, W.G. (1977). Sampling techniques. Wiley series in probability and mathematical statistics-applied.. J.W.a. sons. 413 pp

Efron, B. (1979). "Bootstrap methods: another look at the jackknife." Annals of Statistics 7: 1-26.

Efron, B. and R. J. Tibshirani (1993). An introduction to the bootstrap. New York, Chapman & Hall.

Hirst, D.J., Aanes, S., Storvik, G., Huseby, R.B. and Tvete, I.F.(2003). Estimating catch-at-age from market sampling data using a Bayesian hierarchical model. (Applied Statistics, in press).

Hirst, D., G. Storvik, et al. (2004). Working document. Estimating catch-at-age by combining data from different sources. WKSCMFD, Nantes

Jardim, E., V. Trujillo And P. Sampedro (2004 in press). Uncertainties in Sampling Procedures and Age Composition for Hake and Sardine in Atlantic Iberian Waters.

Kimura, D.K. (1977). Statistical Assessment of the Age-length Key. Journal of Fisheries Research on Board of Canada. 34:3. p

Lai, H.-L. (1987). Optimum allocation for estimating age composition using age-length key. Fishery Bulletin 85:

Manly, B. F. J. (1997). Randomization, bootstrap and Monte Carlo methods in biology. London, Chapman & Hall.

Maxwell D.M., B.D.Rackham, S.Warnes, M.Armstrong, A.W.Newton, R.Officer. (2001). A study of market sampling data, precision of age compositions and the effect of mis-ageing: ICES Division VIIa cod. ICES CM 2001/P:11.

O'Brien C.M., C.D. Darby, D.L. Maxwell, B.D. Rackham, H. Degel, S. Flatman, M.A. Pastoors, E.J. Simmonds and M. Vinther (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., C.D. Darby, B.D. Rackham, D.L. Maxwell, H. Degel, S. Flatman, M. Mathewson, M.A. Pastoors, E.J. Simmonds and M. Vinther (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

Patterson, K. R., R. M. Cook, C. D. Darby, S. Gavaris, L. T. Kell, P. Lewy, B. Mesnil, A. E. Punt, V. R. Restrepo, D. Skagen and G. Stefansson (2001). "Estimating Uncertainty in Fish Stock Assessment and Forecasting." Fish and Fisheries

Quinn, T.J. and Deriso, R.B. (1999). Quantitative Fish Dynamics. Oxford University Press. 542 pp

Simmonds E.J., C. L. Needle, , H Degel, S Flatman, C. M. O'Brien, M. A. Pastoors, A. P. Robb and M. Vinther (2001).The precision of international market sampling for North Sea herring and its influence on assessment. ICES CM 2001/P:21.

Stratoudakis, Y, Fryer, R.J., Cook, R.M. and Pierce, G.J. (1999) Fish discarded from Scottish demersal vessels: Estimators of total discards and annual estimates for targetted gadoids. ICES. J. Mar. Sci., 56, 592-605.

Vigneau J. and S. Mahevas. (2004) Working document. Precision in catch at age data with regard to sampling design. WKSCMFD, Nantes.

10 WORKING DOCUMENTS

The following material was available to the Workshop in the form of Working Documents:

WD01	Hirst, D., G. Storvik, et al. (2004). Estimating catch-at-age by combining data from different sources. WKSCMFD, Nantes.
WD02	Oeberst, R. (2004). A universal cost function for the optimization of the number of age readings and length measurements for Age-Length-Key-Tables. WKSCMFD, Nantes.
WD03	Oeberst, R. (2004). Precision of length-at-age and weight-at-age based on data from commercial landings according to requirement of EU No. 1639/2001. WKSCMFD, Nantes.
WD04	Vigneau, J. and S. Mahevas (2004). Precision in catch at age data with regard to sampling design. WKSCMFD, Nantes.
WD05	O'Brien, C. M., C. D. Darby, et al. (2001). The precision of international market sampling for North Sea Plaice (Pleuronectes platessa L.) and its influence on stock assessment. ICES CM 2001/P:13.
WD06	O'Brien, C. M., C. D. Darby, et al. (2001). The precision of international market sampling for North Sea cod (Gadus morhua L.) and its influence on stock assessment. ICES CM 2001/P:14.
WD07	Simmonds, E. J., C. L. Needle, et al. (2001). The precision of international market sampling for North Sea herring and its influence on stock assessment. ICES CM 2001/P:21
WD08	Sampedro, P. and V. Trujillo (2004). Sampling uncertainties of the main fish caught on the Spanish Atlantic Coast. WKSCMFD, Nantes.

Estimating catch-at-age by combining data from different sources

David Hirst[1], Geir Storvik[1,3], Magne Aldrin[1], Sondre Aanes[2], Ragnar Bang Huseby[1]

1:Norwegian Computing Center, Oslo, Norway. 2:Institute for Marine Research, Bergen, Norway. 3. University of Oslo

Abstract

Estimating the catch-at-age of commercial fish species is an important part of the quota-setting process, for many different species and almost all countries with a fishing fleet. Current procedures are usually very time-consuming and somewhat ad hoc, and the estimates have no measure of uncertainty. Hirst et al (2003) developed a method for catch-at-age of Norwegian cod, but this only considered aged fish sampled randomly from random hauls. In most countries the sampling scheme is not so simple. There are usually a very large number of length-only samples, from which the age must be estimated using an age-length relationship, and often some or all of the age samples are collected from data which are first stratified by length. This adds considerably to the difficulties in the estimation. In this paper we model the three different kinds of data (random samples of age, length and weight, age and weight stratified by length, and length-only) simultaneously using a development of the Bayesian hierarchical model in

Hirst et al (2003). This enables us to obtain estimates of the catch-at-age with appropriate uncertainty, and also to provide advice on how best to sample data in the future.

Keywords: Bayesian hierarchical model, catch-at-age, fishery statistics, Markov Chain Monte Carlo.

1. Introduction

As part of the process of setting fishing quotas, every country in Europe with a fishing fleet reports the total annual catch-at-age of various species to the International Council for the Exploration of the Seas (ICES). Strictly speaking, catch-at-age means the total number of fish caught at each age. However, it is common to group less frequent ages together to form a number of age groups. In our case fish older than 12 are considered one group. Also, an unknown number of fish are caught but discarded at sea. We do not take into account these discards. This kind of data is sometimes known as 'market sampling', although in Norway a substantial part of the data is taken directly from the boat rather than from the market.

The weight of the total catch is usually considered to be known at a fairly fine resolution (in Norway season by gear by area by year) and the aim of the analysis is (a) to estimate proportion at age, and (b) to estimate the mean weight of fish in order to convert the total to numbers from weight. A variety of different sampling schemes have been established for this purpose, and the data are analysed in a range of different ways. A common feature of most of these methods is that there is no statistical model for the sampling process. Ad-hoc methods are used which are very time-consuming, and rely on individual judgement, which by definition is not repeatable. The Norwegian approach is outlined in Hirst *et al.* (2003). With such methods it is very difficult to get a measure of the uncertainty in the reported results. In order to address this problem, Hirst *et al.* (2003) developed a Bayesian hierarchical model for the Norwegian catch of Northeast Arctic cod (*Gadus morhua*). This model however only addressed the strategy of sampling fish at random from random boats, and estimating the age and measuring the weight of all the fish in the sample. There was no modelling of length in this paper, and weight was modelled directly in terms of age.

The Norwegian sampling scheme is probably unique in Europe. Elsewhere there is an emphasis on sampling large numbers of fish for which only length is measured, and weighing and estimating the ages of only a few of these. These aged fish are usually stratified by length, eg one fish from each 5cm length class in a sample might be aged. This kind of sampling in fact takes place to a lesser degree in Norway as well, and additional length-only or age-given-length data (often from independent sources such as the coastguard) are utilised in the estimation. In this paper we develop a model for all these kinds of data.

The difficulties in the analysis arise mostly because it is not possible to develop a proper sampling scheme for fishing vessels. In general they are sampled when and if they are available. There are important differences in the catch between different seasons, fishing gears and regions of the sea, and if we call each combination of these factors a cell, there are necessarily many cells with no samples. In addition, there is a large within haul correlation in the ages and sizes of the caught fish. Thus the ‘effective sample size’ is very much smaller than the number of fish sampled. See Aanes and Pennington (2003). This leads to a larger uncertainty than would be apparent from a naïve assumption of independence of fish.

The aim of this paper is to establish a proper statistical framework within which market-sampling data can be analysed. The Bayesian hierarchical framework is very appropriate for this kind of modelling, because it can easily accommodate the different sampling schemes, and because it provides a full measure of uncertainty. Bayesian approaches are now slowly emerging into the fisheries community, see eg Millar and Meyer (2000) and the references therein.

2. The data

There are 3 main sources of data available to the Norwegian Institute of Marine Research (IMR), which is responsible for estimating the catch-at-age of cod in Norway. They are

1) The Amigo, a research vessel hired by IMR: This sails from port to port along the north Norwegian coast over a period of about 6 weeks, 4 times a year (roughly

corresponding to the 4 seasons). At each port it takes a sample of about 80 fish from any boats available at the time. There is rarely more than one boat available. The fish are weighed, the length measured, and the otoliths extracted for estimating the age of the fish (Campana, 2001). Each year about 200 boats, and thus about 16,000 fish are sampled. Note that the program only samples landings. There are an unknown number of small fish discarded at sea, though we refer to catch-at-age in this paper.

2) The coastguard: One of the coastguard's tasks is to make sure that the Norwegian fishery laws and regulations are kept and they have the right to inspect any vessel and to sample the catch. In most cases the vessels sampled by the coastguard are a random sample of the vessels operating within an area, but in a few cases the inspections may be based on suspicion of illegal fishing. Thus, it might be expected that some of the samples would be biased or unrepresentative for the total catch, although this does not appear to be the case. In general these samples will only provide length measurements of the fish sampled, though occasionally there are some ages and weights as well. The coastguard samples more of the trawlers than the Amigo. The number of fish sampled in each haul is very variable, but averages about 100.

3) The Reference Fleet: This is a fleet of commercial fishing vessels that have agreed to provide IMR with data on their catch. The reference fleet was started in 2001 with 6 vessels, and consists currently of 8 vessels. The fleet targets several commercially important species including cod. This sampling program is developing and will expand in the years ahead. So far it has consisted mostly of

length-only data, but there are an increasing number of age samples. In 2002 this fleet sampled approximately 500 hauls of cod with around 90 fish sampled in each haul.

The 'cells' we consider in this paper are the individual combinations of the regions in figure 1, season (corresponding roughly to the quarters of the year), gear (bottom trawl, Danish seine, gillnet, longline and handline) and year (1995 to 2002, but the reference fleet only began in 2001). One cell therefore represents one gear, in one region, in one season of one year. Our sampling unit is the haul, and we do not consider the actual boat which was sampled to be of interest. For the Amigo and coastguard data it is very unlikely that the same boat would be sampled twice (at least in the same year), but clearly the reference fleet provides many samples from the same few boats. Any boat effect however is largely due to the particular gear being used, and the remaining effect will be very small compared to the differences between hauls. For the purposes of the analyses in this paper, we have formed 'super-regions' by grouping the regions in the map. In fact we have used the 8 'standard' IMR groups of {3, 2, 10, 11, 13, 14, 15, 16, 17, 24, 1}, {12}, {4}, {5, 37, 39}, {0}, {6}, {7, 28} and {20, 21, 22, 23, 25, 27}. It is necessary to do some grouping because most regions have little or no data, although other groupings are possible. We have grouped ages over 12 together, and there are no fish younger than two, giving us 12 age groups.

In the next sections we develop the various components of the model: The proportion at age, length-given-age, and weight-given-length. The components are brought together in the likelihood for the whole data set. We then explain how to obtain samples from the posterior distribution of the parameters given the data using Markov Chain Monte Carlo (MCMC) (Gilks et. al 1996). Finally we show some results, and illustrate how these change when different data sources are included in the analysis. This also enables us to provide some guidance on how best to sample in the future.

3. The model for proportion at age

The samples from a boat are assumed to be randomly drawn from the total population of fish in that haul, and the hauls are themselves assumed to be randomly sampled from all those within the appropriate cell. The numbers at age in a sample from haul h from cell c, $\boldsymbol{X}_{c,h}$, are therefore multinomial,

$$\boldsymbol{X}_{c,h} \sim \text{multinomial}(\boldsymbol{p}_{c,h.}, n_{c,h})$$

The number of fish sampled from the haul, $n_{c,h}$, is assumed not to depend in any way on $\boldsymbol{p}_{c,h}$.

The vector of proportions at age in the haul, $\boldsymbol{p}_{c,h}$ has A elements, one for each age group. Let $p_{c,h}(a)$ be the a^{th} element, where $0 \le p_{c,h}(a) \le 1$ and $\sum_{a'=1}^{A} p_{c,h}(a') = 1$. This is reparameterised as $p_{c,h}(a) = \dfrac{\exp(\alpha_{c,h}^{a})}{\sum_{a'=1}^{A} \exp(\alpha_{c,h}^{a'})}$.

We model $\alpha_{c,h}^{a}$ in terms of the various covariates as

$$\alpha_{c,h}^{a} = \alpha^{base,a} + \alpha_{y(c)}^{year,a} + \alpha_{s(c)}^{season,a} + \alpha_{g(c)}^{gear,a} + \varsigma_{r(c)}^{region,a} + \varsigma_{c}^{cell,a} + \varsigma_{c,h}^{haul,a}.$$

Here *y(c)* means the year, *s(c)* the season, *g(c)* the gear and *r(c)* the region corresponding to cell *c*. From now on for clarity we drop the *c* and just refer to $\alpha_{y}^{year,a}$ etc.

The α terms and $\varsigma_{r}^{region,a}$ are the main effects for year, season, gear and region. The α terms are fixed effects and $\varsigma_{r}^{region,a}$ is a spatially smoothed random effect. Some spatial smoothing is necessary in order to estimate the proportions for areas with no data. Here this is accomplished by assuming $\varsigma_{r}^{region,a}$ follows a Gaussian conditional autoregressive distribution (CAR) (e.g. Carlin & Louis, 1996). It is assumed that there will always be some data for all levels of the fixed effects that are of interest. The $\varsigma_{c}^{cell,a}$ terms are independent random effects modelling the interactions between the main effects. In other words the differences between the fit from the main-effects-only model and the true cell means are modelled by the $\varsigma_{c}^{cell,a}$ terms. The differences between hauls within a cell are modelled by the random effects $\varsigma_{c,h}^{haul,a}$. For more details of the parameters, including identifiability constraints and the prior distributions, see the appendix.

4. The models for length-given-age and weight-given-length

In figure 2a we plot log(length) against log(age) for all the hauls in one cell. Each symbol corresponds to one haul, and the lines are the estimated regression lines for each haul. It can be seen that the linear relationship is plausible, but there may be some problems at high and low ages. There appears to be some variation in the slope, but rather more in the intercept. The equivalent plot of log(weight) against log(length) is in figure 2b. Here the linear relationship is very clear, and there is rather more variation in the intercept than the slope. We therefore model both relationships as linear with a constant slope and variable intercept. Other models would of course be possible, and may in particular be necessary for other species with different growth patterns. Note that it would be possible to model weight given age directly, but that modelling it via length enables us to get a better estimate of the mean weight at age in cells with length but no age data.

We assume length-given-age and weight-given-length are log Normal, with constant variances, and means linear in log(age) and log(length) respectively in an individual haul. The slopes are constant, but the intercepts vary between cells and boats within a cell:

$$\log(length_{c,h,f}) = \beta_{0,c,h} + \beta_1 \log(age_{c,h,f}) + \varepsilon^{fish}_{c,h,f}$$
$$\log(weight_{c,h,f}) = \delta_{0,c,h} + \delta_1 \log(length_{c,h,f}) + \nu^{fish}_{c,h,f}$$

Here $length_{c,h,f}$ is the length of the f^{th} fish from haul h in cell c, $weight_{c,h,f}$ its weight and $age_{c,h,f}$ its age. $\varepsilon^{fish}_{c,h,f}$ and $\nu^{fish}_{c,h,f}$ are independent zero mean Gaussian random variables. The slopes β_1 and δ_1 are common to all cells and hauls, and the intercepts $\beta_{0,c,h}$ and $\delta_{0,c,h}$ are given by:

$$\beta_{0,c,h} = \beta^{base} + \beta^{year}_y + \beta^{season}_s + \beta^{gear}_g + \varepsilon^{region}_r + \varepsilon^{cell}_c + \varepsilon^{haul}_{c,h}$$

$$\delta_{0,c,h} = \delta^{base} + \delta_y^{year} + \delta_s^{season} + \delta_g^{gear} + \nu_r^{region} + \nu_c^{cell} + \nu_{c,h}^{haul}$$

ε_r^{region} and ν_r^{region} are CAR parameters with similar properties to $\varsigma_r^{region,a}$ in the age model (see appendix). ε_c^{cell} and ν_c^{cell} are random 'all interactions' effects equivalent to $\varsigma_c^{cell,a}$. ε_c^{haul} and ν_c^{haul} are between haul random terms. The β and δ terms are fixed effects similar to the α terms in the model for proportions at age. For more details see the appendix.

5. Inference on unknown parameters

Parts of the model are standard, and ordinary methods such as maximum likelihood could have been applied. This is certainly true for the length-given-age model and the weight-given-length model. With all ages known and with no random effects involved in the age model also the parameters in the multinomial model describing the age proportions could easily be found by maximum likelihood. Both the inclusion of random effect in the multinomial model and missing ages makes maximum likelihood estimation much more complicated. Further, a frequentistic approach to estimation makes it difficult to take the uncertainty in the parameters into account.

Our approach has been the Bayesian one. The full information about the parameters are described through the posterior distribution. This distribution is difficult to calculate, but approximations can be obtained through Monte Carlo sampling. The actual sampling is performed through an MCMC algorithm using a combination of Gibbs sampling and

Metropolis-Hastings steps. The details are left to a more technical paper, but in outline the approach is as follows:

1) If the ages and lengths of all the sampled fish were known, it would be simple to simulate the parameters of the length-given-age model (since this is just a linear model). It would also be relatively simple to simulate the parameters of the proportion-at-age model (although the inclusion of random effects complicates the simulations somewhat). Also parameters from the different submodels (age model, length given age model and weight given length model) are independent in this case.

2) If the parameters of the length-given-age model and the proportion-at-age model are known, it is simple to simulate the ages of the fish with only length data (since age-given-length is multinomial).

3) We therefore treat the missing ages as parameters and use Gibbs sampling to alternate between simulating the missing ages, and simulating the other model parameters. It is also possible to use block updating for most of the parameters apart from the precisions.

Using this approach it is possible to find the joint posterior distribution of all the parameters very quickly. On a reasonably powerful pc, one year's data can easily be analysed in less than 5 minutes. Obviously the time increases with the number of years of data, but even 8 years worth takes under an hour. Convergence of the MCMC chains is fast because of the block-updating. Research is currently underway to make this even more efficient.

6. Estimating catch-at-age

We need to estimate the total catch at age *a* in cell *c*, T_{ca}. We have

$T_{ca} = T_c \text{ x } mean_c(p(a))$, where T_c is the total catch in the cell in numbers of fish, and $mean_c(p(a))$ is the mean proportion at age over all hauls in the cell. We assume that there are a large number of hauls in the cell, so that the mean is equal to the expected value, giving us $T_{ca} = T_c E_c(p(a))$.

The total catch in a cell, W_c is given in weight, rather than numbers. We therefore need the mean weight of fish caught in the cell, $\overline{w}_c$, in order to calculate $T_c = W_c / \overline{w}_c$.

We have:

$$\log(weight_{c,h,f}) = \delta_{0,c,h} + \delta_1 \log(length_{c,h,f}) + \nu_{c,h,f}^{fish}$$

$$\log(length_{c,h,f}) = \beta_{0,c,h} + \beta_1 \log(age_{c,h,f}) + \varepsilon_{c,h,f}^{fish}$$

Thus

$$\begin{aligned}\log(weight_{c,h,f}) &= \delta_{0,c,h} + \delta_1\left(\beta_{0,c,h} + \beta_1 \log(a_{c,h,f}) + \varepsilon_{c,h,f}^{fish}\right) + \nu_{c,h,f}^{fish} \\ &= \delta_{0,c,h} + \delta_1\beta_{0,c,h} + \delta_1\beta_1 \log(a_{c,h,f}) + \left(\delta_1\varepsilon_{c,h,f}^{fish} + \nu_{c,h,f}^{fish}\right) \\ &= A_c + A_h + B_c \log(a_{c,h,f}) + C_f\end{aligned}$$

Here A_c and B_c are constant for all hauls in a cell, A_h is a random haul dependent intercept, and C_f is a random fish effect.

From the earlier equations:

$$A_c = \left(\delta^{base} + \delta_y^{year} + \delta_s^{season} + \delta_g^{gear} + \nu_r^{region} + \nu_c^{cell}\right) + \delta_1\left(\beta^{base} + \beta_y^{year} + \beta_s^{season} + \beta_g^{gear} + \varepsilon_r^{region} + \varepsilon_c^{cell}\right)$$

$$B_c = \delta_1 \beta_1$$

A_h is a random haul dependent intercept

$$A_h = \nu_{c,h}^{haul} + \delta_1 \varepsilon_{c,h}^{haul}.$$

This is constant in a haul, but a Gaussian random variable within a cell:

$$A_h \sim N\left(0, \frac{1}{\tau_{weight}^{haul}} + \frac{\delta_1^{\,2}}{\tau_{length}^{haul}}\right).$$

C_f is a random fish effect:

$$C_f = \left(\delta_1 \varepsilon_{c,h,f}^{fish} + \nu_{c,h,f}^{fish}\right)$$

This is random fish effect, with a constant distribution:

$$C_f \sim N\left(0, \frac{1}{\tau_{weight}^{fish}} + \frac{\delta_1^{\,2}}{\tau_{length}^{fish}}\right)$$

Hence the weight of a random fish f of age $a_{c,h,f}$ in haul h, cell c is lognormal:

$$\log(weight_{c,h,f} \mid a_{c,h,f}) \sim N\left(A_c + A_h + B_{cell} \log(a_{c,h,f}), \frac{1}{\tau_{weight}^{fish}} + \frac{\delta_1^{\,2}}{\tau_{length}^{fish}}\right).$$

Its expectation (over all fish in the haul) is

$$E_{c,h}(weight_{c,h,f} \mid a_{c,h,f}) = \exp(A_c + B_c \log(a_{c,h,f}))\exp(A_h)\exp\left(\tfrac{1}{2}\left(\frac{1}{\tau_{weight}^{fish}} + \frac{\delta_1^{\,2}}{\tau_{length}^{fish}}\right)\right)$$

Taken over all boats in a cell, this expectation is itself a random variable, also lognormal, since $\exp(A_h)$ is lognormal:

$$\log\left(E_{c,h}(weight_{c,h,f} \mid a_{c,h,f})\right) \sim N\left(A_c + B_c \log(a_{c,h,f}) + \tfrac{1}{2}\left(\frac{1}{\tau_{weight}^{fish}} + \frac{\delta_1^{\,2}}{\tau_{length}^{fish}}\right), \frac{1}{\tau_{weight}^{haul}} + \frac{\delta_1^{\,2}}{\tau_{length}^{haul}}\right)$$

The expected weight of a fish of age *a* in a cell is therefore

$$E_c(weight \mid a) = \exp\left(A_c + B_c \log(a) + \tfrac{1}{2}\left(\frac{1}{\tau_{weight}^{fish}} + \frac{\delta_1^{\ 2}}{\tau_{length}^{fish}} \right) + \tfrac{1}{2}\left(\frac{1}{\tau_{weight}^{haul}} + \frac{\delta_1^{\ 2}}{\tau_{length}^{haul}} \right) \right).$$

Again assuming a large number of hauls in a cell, so that the mean weight at age is equal to its expected value, the mean weight of a fish in the cell is given by

$$\overline{w}_c = \sum_a E_c(p(a)) E_c(weight \mid a)$$

There is no explicit formula for $E_c(p(a))$, but the estimator

$$\hat{E}_c(p(a)) = \frac{\exp\left(E\left(\alpha_{c,h}^{a}\right)\right)}{\sum_{a'=1}^{A} \exp\left(E\left(\alpha_{c,h}^{a'}\right)\right)}$$

is almost unbiased, so long as the within cell variance of $\alpha_{c,h}^{a}$ for each age is small compared to the difference between the $E\left(\alpha_{c,h}^{a}\right)$ for the different ages within the cell. This is certainly the case for our data.

These formulae enable us to calculate the catch at age *for a given set of parameters.* In order to take the uncertainty of the parameters into account, parameters are sampled from the joint posterior distribution (as described in the previous section). A Monte Carlo estimate of the catch at age is then obtained for each set of parameters, and uncertainty measures can be calculated simultaneously.

7. Results

The method described gives as output the joint posterior distribution of total catch for each age group for any combination of cells, as well as the (joint) posteriors of all the individual parameters. We expect that interest will usually be mostly in the catch-at-age results, along with the uncertainty in the estimates which is directly available from the posteriors. One example of these results is shown in figures 3a and b where the posterior means for each age, along with 95% credibility intervals, are plotted for 1995 and 2002, using data from 1995 to 2002. Obviously the 12 age groups are not independent, but the error bars give a good indication of the uncertainty. If for example the results were to be used for a virtual population analysis (VPA), samples form the full joint posteriors for all required years would be available. These could be used to obtain the posterior distributions of the parameters calculated in the VPA.

These plots also illustrate the effect of using the length-only data in addition to the length and age data from the Amigo. The numbers of Amigo and ‘extra’ (ie reference fleet and coastguard) hauls sampled per year are shown in figure 4. The ‘extra’ samples are virtually all length-only although it is expected that in the future there will be more age-given-length samples. In 1995 there were only about 50% as many extra samples as Amigo samples, and there is almost no difference in the results. By 2002 however there were about 4 times as many extra as Amigo samples, and there is a useful reduction in the size of the error bars. Clearly there is some information in the length-only data, but per sample age readings are far more informative.

The model parameters are also of interest. Some examples are shown in figures 5 and 6. The first shows the posterior distributions of standard deviations of the random effects in the age model, ie $\sqrt{1/\tau_{age}^{cell}}$, $\sqrt{1/\tau_{age}^{haul}}$ and $\sqrt{1/\tau_{age}^{region}}$. The region sd is the most uncertain, not surprisingly since we only use 8 regions in its estimation. The haul sd is the most precisely estimated. Note that the region sd should be scaled by the number of neighbours of a region in the distribution of $\varsigma_r^{region,a}$, and so for a region with several neighbours, the posterior mean of the sd would in fact be smaller than the posterior means of the other two sds.

The second plot shows the posterior means of the season effects in the age model, ie $\alpha_s^{season,a}$. Note that the values for age 6 and season 1 are defined to be zero. A high value for this parameter means a higher probability of catching fish from age group *a* in season *s*, so for example it can be seen that fish younger than 3 become more likely to be caught later in the year, presumably because they get bigger. This is also true to a lesser degree for ages 3 and 4, but the effect disappears, or maybe even reverses for older fish.

8. Conclusions

The model described is as far as we know the first comprehensive approach to analysing multiple sources of catch-at-age data, in a way that can include all types of sampling schemes we know of, and that properly accounts for the uncertainty in the estimation. It is very fast (at least compared to traditional methods), and in addition to the catch-at-age estimates, can also give information on the model parameters, which may be interesting biologically.

There are a number of additions and improvements that could be made. Perhaps the most interesting would be to include errors in the age-reading. This was done for the simpler sampling scheme of Hirst et al (2003), and with some development could be included in this model. This may be important because unpublished work from the Institute of Marine Research suggests that on average about 10% of the ages may be wrong by one year increasing up to 40% for older fish. It would also be possible to include different length-given-age or weight-given-length models, which may be appropriate for different fish species, or a different spatial model which may suit different fisheries.

Appendix

In all 3 models (proportion at age, length-given-age and weight-given-length) the fixed effect parameter values are relative to the baseline terms $\alpha^{base,a}$, β^{base} and δ^{base}, and it is necessary to set one level of each fixed effect to zero for identifiability:

$$\alpha_{y*}^{year,a} = \alpha_{s*}^{season,a} = \alpha_{g*}^{gear,a} = \beta_{y*}^{year} = \beta_{s*}^{season} = \beta_{g*}^{gear} = \delta_{y*}^{year} = \delta_{s*}^{season} = \delta_{g*}^{gear} = 0$$

For the proportion at age model the proportions must sum to 1, and so we have the additional restriction that all parameters for one age group $a*$ are set to zero:

$$\alpha^{base,a*} = \alpha_{y}^{year,a*} = \alpha_{s}^{season,a*} = \alpha_{g}^{gear,a*} = 0$$

We use a*=6 (usually the most common age group) and y*=s*=g*=1. The choice of y*, s* and g* is unimportant, but convergence is fastest if a* is one of the most common age groups. Setting the parameters to zero for some value of a* has the undesirable effect of giving the catch at age for this age group a smaller posterior variance than the other age groups. A better restriction might be to make the mean over all age groups constant.

We give all non-zero fixed effects non-informative prior distributions:

$$\alpha^{base,a} \sim N(0,1/0.001)\forall a \neq a*$$
$$\alpha_y^{year,a} \sim N(0,1/0.001)\forall y \neq y*, a \neq a*$$
$$\alpha_s^{season,a} \sim N(0,1/0.001)\forall s \neq s*, a \neq a*$$
$$\alpha_g^{gear,a} \sim N(0,1/0.001)\forall g \neq g*, a \neq a*$$

and so on for the β and δ terms.

The spatial terms $\varsigma_r^{region,a}$, ε_r^{region} and ν_r^{region} have Gaussian conditional autoregressive (CAR) prior distributions (see for example Carlin and Louis (1996)):

$$\varsigma_r^{region,a*} = 0$$
$$\varsigma_r^{region,a} \mid \varsigma_{j\neq r}^{region,a} \sim N\left(\bar{\varsigma}_r^{region,a}, \frac{1}{\tau_{age}^{region} n_r}\right), a \neq a*$$
$$\bar{\varsigma}_r^{region,a} = n_r^{-1} \sum_{j\in\partial(r)} \varsigma_j^{region,a}$$

$\partial(r)$ = set of neighbours of region r

n_r=number of neighbours of region r.

The priors for ε_r^{region} and ν_r^{region} are similar.

The ς, ε and ν terms are independent random effects, again set to zero for $a=a^*$, with the following priors:

$$\varsigma_c^{cell,a} \sim N\left(0,1/\tau_{age}^{cell}\right) \forall c, a \neq a^*$$
$$\varsigma_h^{haul,a} \sim N\left(0,1/\tau_{age}^{haul}\right) \forall h$$
$$\varsigma_c^{cell,a^*} = \varsigma_h^{haul,a^*} = 0 \; \forall c, h$$

The ε and ν priors are similar. All precision terms τ are given vague Gamma(0.01,0.01) priors.

References

Aanes,S. and Pennington,M. (2003) On estimating the age composition of the commercial catch of Northeast Arctic cod from a sample of clusters. ICES journal of marine science, 60: 297-303.

Carlin,B.P. and Louis,T.A. (1996) Bayes and empirical Bayes methods for data analysis. Chapman and Hall, London.

Campana,S.E.(2001) Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. Journal of fish biology 59, 197-242.

Gilks, W.R., Richardson, S. and Spiegelhalter, D. J.(1996), Markov chain Monte Carlo in practice. London: Chapman and Hall

Hirst,D.J., Aanes,S., Storvik,G., Huseby,R.B. and Tvete,I.F.(2003) Estimating catch-at-age from market sampling data using a Bayesian hierarchical model. Applied Statistics (in press)

Millar, R. B. and Meyer, R.(2000), Non-linear state space modelling of fisheries biomass dynamics by using Metropolis-Hastings within-Gibbs sampling. Appl. Stat. 49, 327-342.

Acknowledgements

This work was funded by the Norwegian Institute of Marine Research.

Figures

Figure 1: Map of sampling area showing ICES regions.

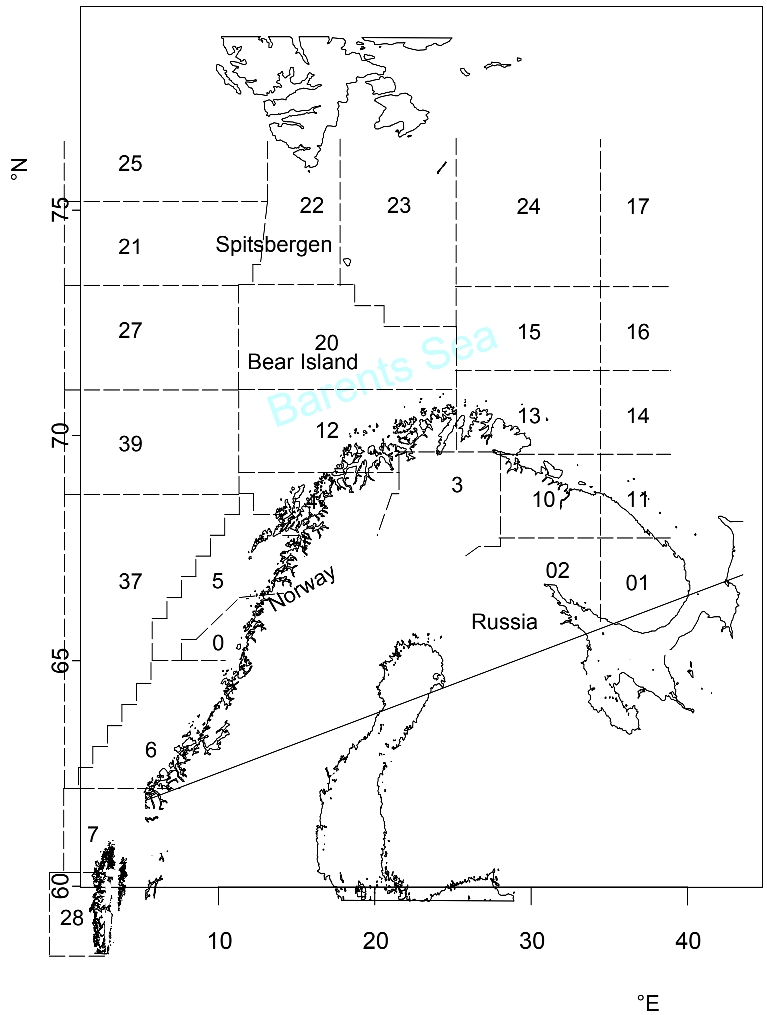
Figure 2: Plots showing relationship between (a) log(length) and log(age) and (b) log(weight) and log(length) all hauls within a cell.

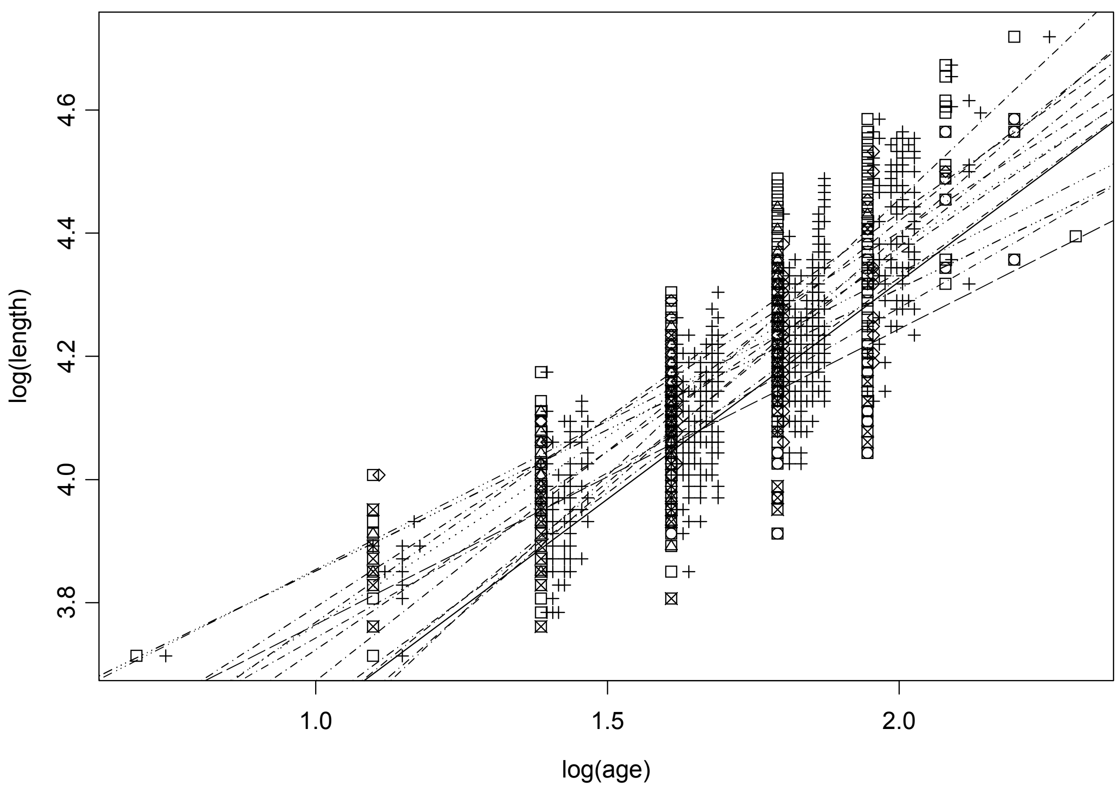
Figure 3: Estimates of catch at age in (a) 1995 and (b) 2002, using Amigo only data and all data. Bars are 95% credibility intervals.

Figure 4: Numbers of Amigo and extra hauls sampled in each year.

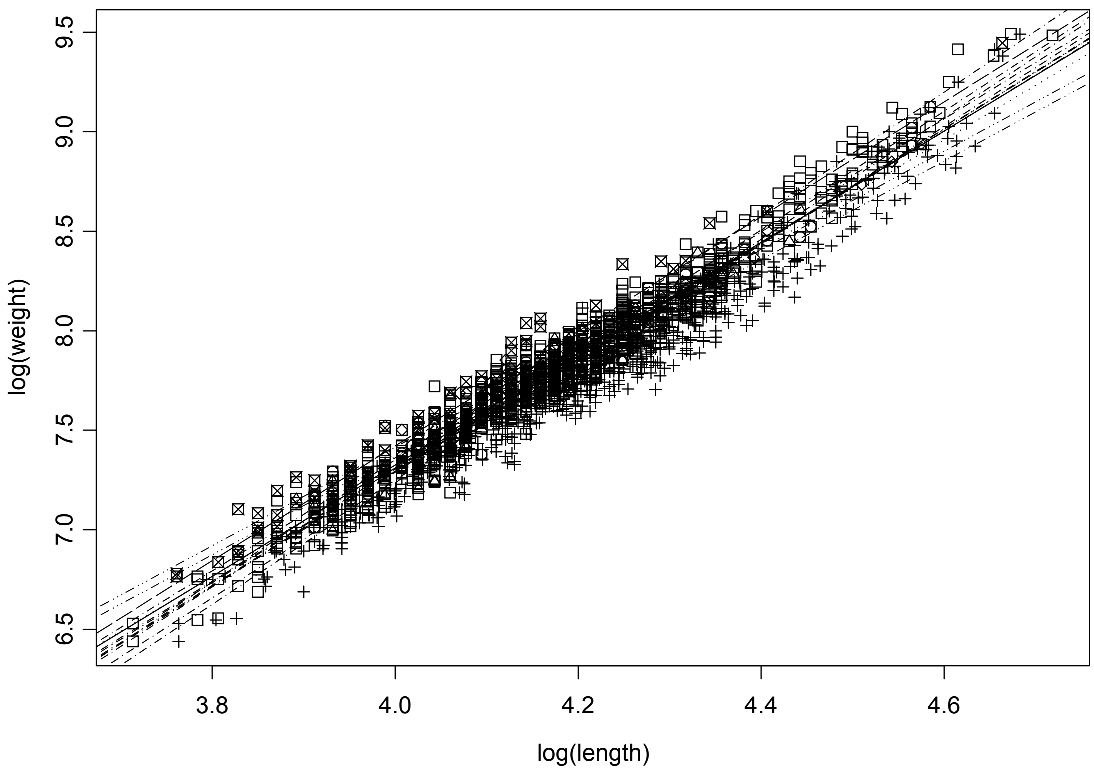
Figure 5: Posterior distributions of the standard deviations of the random effects in the proportion at age model.

Figure 6: Posterior distributions of the season effects in the proportion at age model.

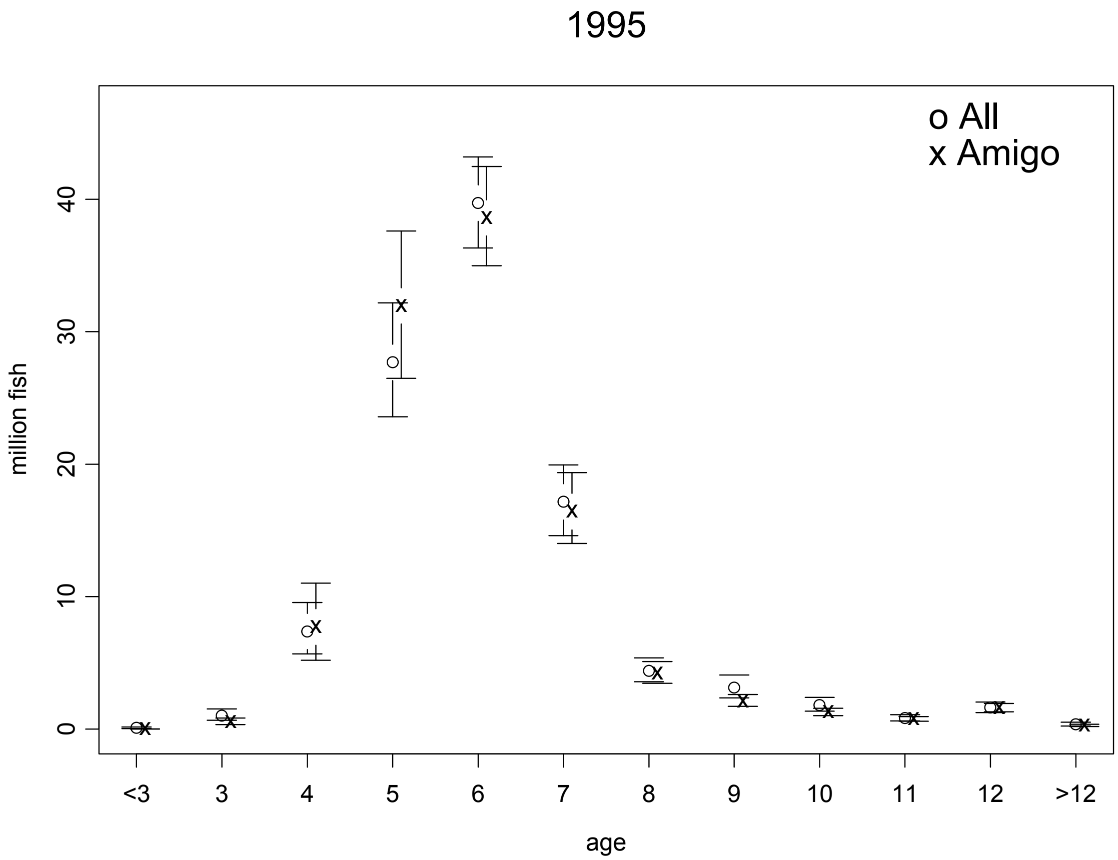

°N
75
70
65
60
10
20
30
40
°E
25
22
23
24
17
21
Spitsbergen
27
20
15
16
Bear Island
Barents Sea
13
14
39
12
3
10
11
37
5
Norway
02
01
0
Russia
6
7
28

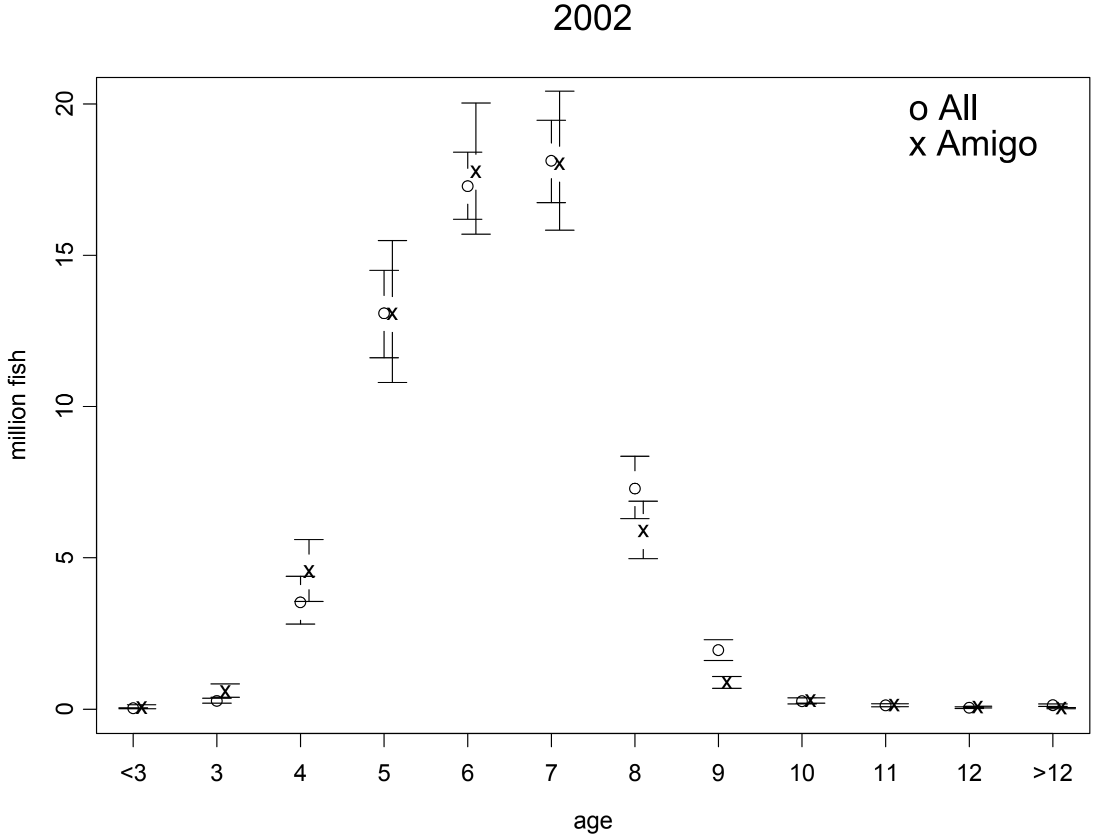


log(length)
4.6
4.4
4.2
4.0
3.8
1.0
1.5
2.0
log(age)

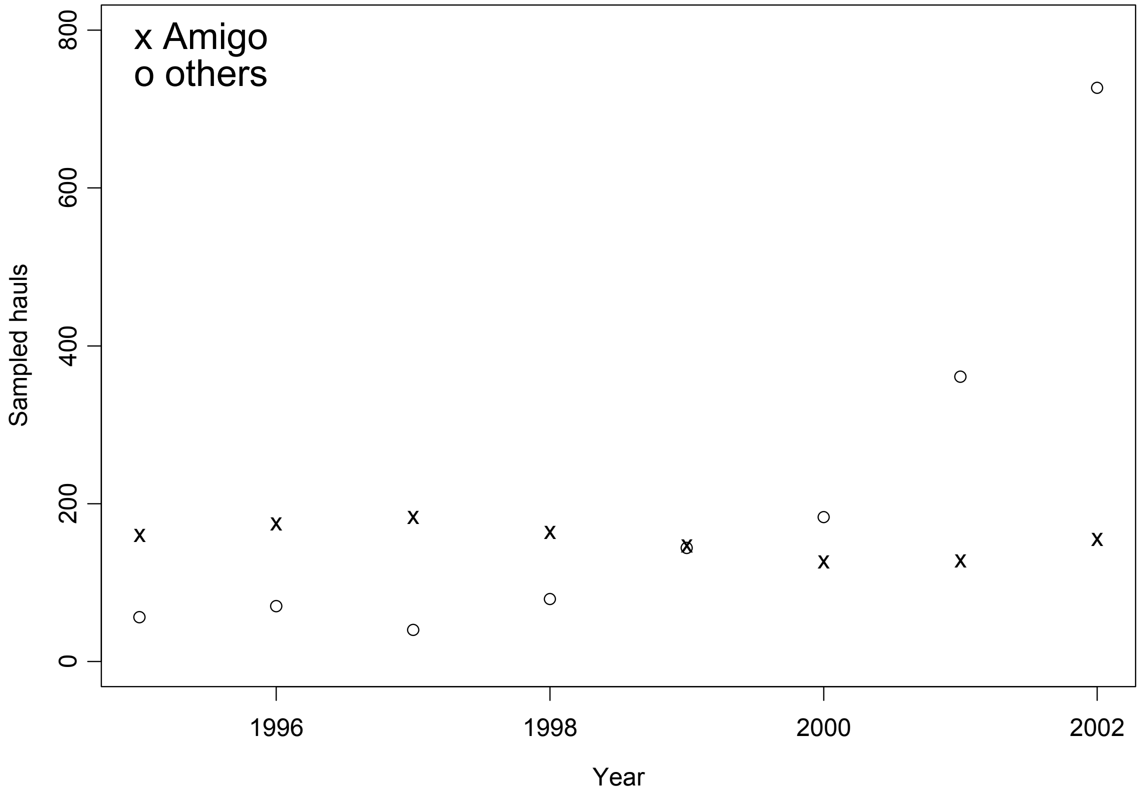


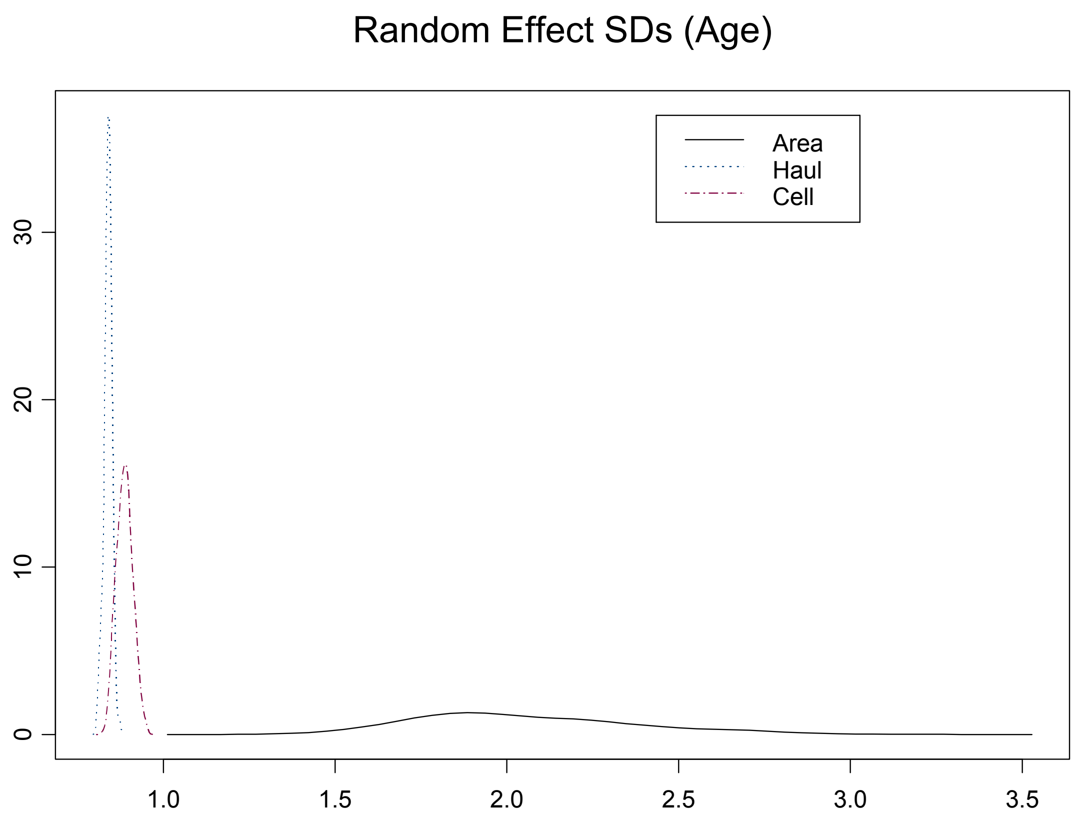
log(weight)
log(length)
6.5
7.0
7.5
8.0
8.5
9.0
9.5
3.8
4.0
4.2
4.4
4.6


1995
o All
x Amigo
million fish
0
10
20
30
40
<3
3
4
5
6
7
8
9
10
11
12
>12
age

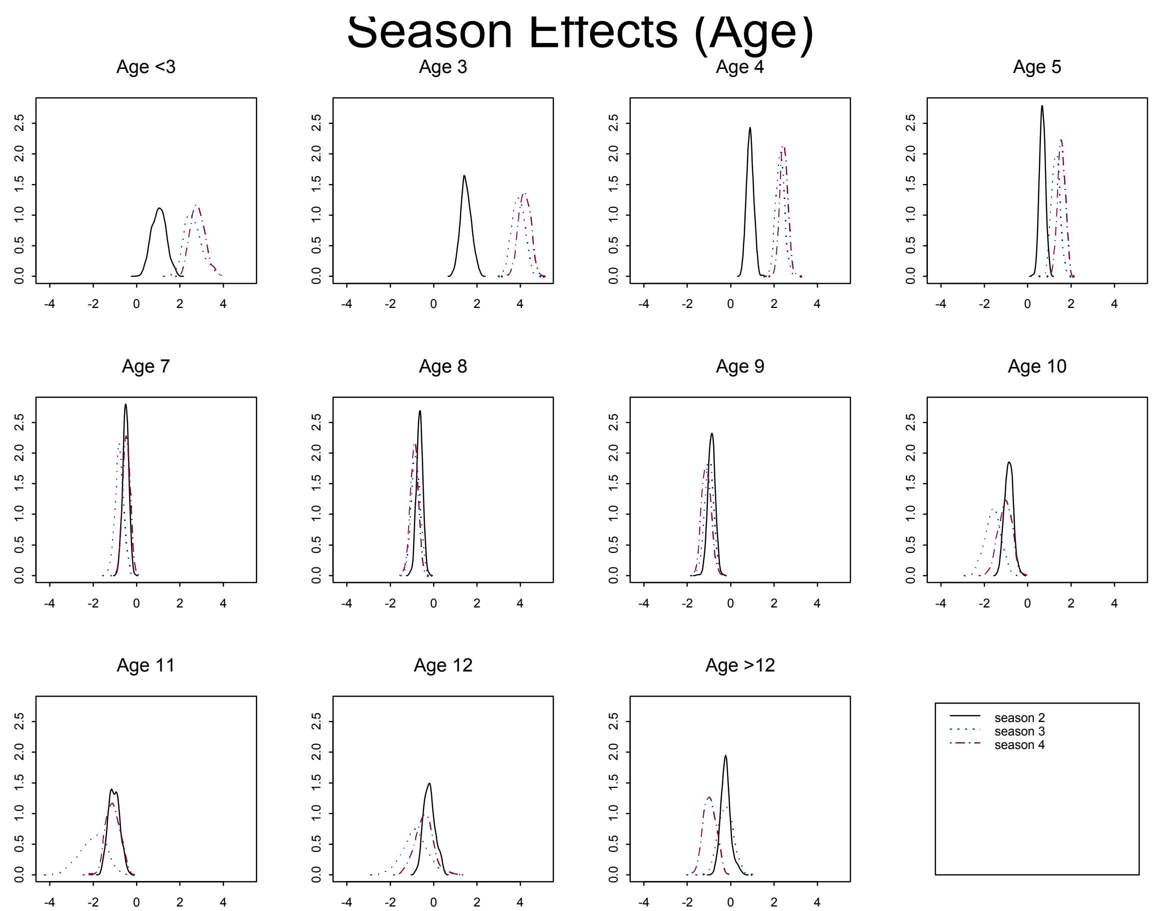


2002
o All
x Amigo
million fish
0
5
10
15
20
<3
3
4
5
6
7
8
9
10
11
12
>12
age


x Amigo
o others
Sampled hauls
0
200
400
600
800
1996
1998
2000
2002
Year



Random Effect SDs (Age)
Area
Haul
Cell
0
10
20
30
1.0
1.5
2.0
2.5
3.0
3.5


Season Effects (Age)
Age <3
Age 3
Age 4
Age 5
Age 7
Age 8
Age 9
Age 10
Age 11
Age 12
Age >12
season 2
season 3
season 4

A universal cost function for the optimization of the number of age readings and length measurements for Age-Length-Key-Tables

R. Oeberst
Institute of Baltic Sea Fishery, Rostock , Germany

Abstract

The accuracy of the estimated proportion of the age groups for surveys or commercial fisheries essentially influences the quality of the stock assessment. Using the method developed by Fridriksson (1934), several methods have been developed for estimating the necessary number of age readings and length measurements for Age-Length-Key-Tables (ALKT). Depending on the special situations different optimized cost functions were used for the optimization.

Due to the difficulty of adaptating the model used for different types of cost such as the time taken or money required to carry out age readings and length measurements, a universal cost function has been developed.

The factor used, referred to as Z, can be interpreted as the quotient from the expenditure of one age determination and one length measurement. Depending on the factor Z chosen, the necessary number of age readings, X' and length measurements, F' can be estimated for the required level of accuracy.

A second step reduces the necessary number of readings using an optimum distribution of the number of age readings per length group.

Since the estimated optimum essentially depends on the quality of the preliminary experiment, an objective criterion has been developed judging its value.

Kurzfassung

Eine verallgemeinerte Kostenfunktion zur Optimierung der Anzahl von Altersbestimmungen und Längenmessungen für Längen - Alters - Schlüssel - Tabellen

Die Genauigkeit der geschätzten Anteile der Altersgruppen bei Surveys und in kommerziellen Fängen beeinflußt wesentlich die Genauigkeit der Bestandsschätzung. Auf der Grundlage des von Fridriksson (1934) entwickelten Verfahrens wurden verschiedene Methoden für eine Optimierung des Probenumfanges entwickelt, um die notwendige Anzahl der Altersbestimmungen und Längenmessungen für Längen – Alters - Schlüssel - Tabellen zu schätzen. In Abhängigkeit von den speziellen Bedingungen wurden unterschiedliche Kostenfunktionen für die Optimierung genutzt.

Weil eine Anpassung der genutzten Modelle an verschiedene Kostenstrukturen, wie die notwendige Bearbeitungszeit oder das Geld für die Altersbestimmung und die Längenmessungen schwierig ist, wurde eine verallgemeinerte Kostenfunktion entwickelt.

Der genutzte Faktor Z kann als Quotient aus dem Aufwand für eine Altersbestimmung und eine Längenmessung interpretiert werden. In Abhängigkeit von dem gewählten Faktor Z kann dann für eine geforderte Genauigkeit die notwendige Anzahl von Altersbestimmungen X' und Längenmessungen F' berechnet werden.

In einem zweiten Iterationsschritt wird der notwendige Probenumfang weiter reduziert. Dafür wird eine optimale Aufteilung der Anzahl der Altersbestimmungen auf die Längenklassen genutzt.

Weil die Schätzung des optimalen Probenaufwandes wesentlich von der Qualität des Vorversuches abhängig ist, wurde ein objektives Kriterium entwickelt, um die Brauchbarkeit des Vorversuches einschätzen zu können.

Introduction

Since 1934, when Fridriksson (1934), first developed a method for using length measurements as a criterion of stratification in order to estimate the proportion of age groups several, authors have discussed the possibilities of optimum data sampling.

Equations for the variance, $V(p_j)$ of proportion, p_j of age group, j were described by Tanaka (1953) using these length stratification criteria. Further, Tanaka (1953) discussed the principle problems of this strategy for data sampling. These are:

a) How many age determinations and length measurements are necessary to guarantee a given level of accuracy of p_j ?

b) Is it useful to implement a cost function in order to take different costs for length measurements and age determinations into account?

c) Is it useful to apply different strategies for sampling age determinations within the length intervals?

Kuthun (1963) compared the estimation of proportions of age groups for unstratified and length stratified age estimations. He also discussed the use of a cost function. Kimura (1977) compared different strategies for age sampling within length classes in cases where the:

a) number of age determinations is proportional to the distribution of length measurements

b) number of age determinations is constant in all length classes.

Kimura (1997) recommended the first strategy as optimal. Pope and Knights (1975) and Chester and Waters (1983) discussed an optimal sampling strategy for commercial fish where the data for the estimation came from several harbours and were collected at different times (monthly basis). Lai (1987) used a special cost function to estimate the number of age determinations and length measurements for several fish species.

All these investigations show that the estimation of an optimal number of age determinations and length measurements depends on the specific situation of the fish species considered. This is the reason why the workshop on sampling strategies for age and maturity (Anon. 1994) stated:

"Optimal allocations are seldom startlingly better than the suboptimal solutions that pragmatic schemes generate. In any case allocations of otoliths which would have been optimal for last year's age distribution may not be for this year's. Furthermore, an optimal allocation means that some specific goal function has been optimized and it follows that other goal functions will not be optimized for that particular strategy. A strategy should therefore be tested for its performance in various settings."

This study describes a method for estimating the size and structure of the sample required to give the portion of each age group to a certain level of accuracy. A universal cost function is also included. Furthermore, a strategy is shown for distributing the number of age determinations over the different length classes. An example of an ALKT is used to demonstrate the equations and strategies.

Calculation of the proportion of an age group, p_j

In most cases, especially for surveys and commercial sampling, the proportion of each age group is calculated by a combination of age readings and length samples.
Table A1 shows an example of how a typical ALKT is calculated by the Institute of Baltic Sea Fishery, Rostock. The data come from a real young fish survey carried out in November 1992 in ICES subdivision 24.
The age-length data are combined with the length distribution to calculate the proportion, p_j of age group, j in parts per thousands (o/oo).

Table 1 shows the formal notation of Table A1 and some of the equations for the calculation of the ALKT.

Table 1: Equations for the "AGE-LENGTH-KEY Table"

Length classes	Age groups $j = 1..K$		
$I = 1 . . M$	X_{11} X_{12} X_{1K} X_{ij} X_{M1} X_{M2} X_{MK}	$\boldsymbol{X_{i.}} = \Sigma_j\ X_{ij}$	F_i
	$P_j = \Sigma_i (H_i * Pij)$	$\boldsymbol{X} = \Sigma_i X_{i.}$	$\boldsymbol{F} = \Sigma_i F_i$

The following notations and equations were used.

M	number of length classes
K	number of age groups
i	index of length class
j	index of age group
X_{ij}	number of fish of age, j and length class, i in the age sample
F_i	number of fish of length class, i in the length sample
p_j	proportion of age group, j

F	$= \Sigma_i F_i$	number of fish in the length sample	(1.1)
H_i	$= F_i / F$	proportion of length class, i in the length sample	(1.2)
$X_{i.}$	$= \Sigma_j\ X_{ij}$	number of fish of length class, i in the age sample (marginal distribution)	(1.3)
X	$= \Sigma_i X_{i.}$	number of age readings	(1.4)
P_{ij}	$= (X_{ij} / X_{i.})$	portion of the age group, j in the length class, i	(1.5)

Then the proportion, p_j of age group, j in the catch can be calculated by

$$p_j = \Sigma_i (H_i * Pij) \qquad (1.6)$$

In order to get an impression of the quality of data and estimations, their variabilities can be considered in the form of variances and confidence intervals. This is done in the following section.

Calculation of the variance of the proportion of an age group p_j

The estimation of the variance of p_j is necessary for assessing the accuracy of the estimated p_j. The variance, $V(p_j)$ (Tanaka, 1953) of the proportion, p_j of age group, j is the sum of two components and is affected by different parameters:

$$V(p_j) = VA(j) + VZ(j) \qquad (2.1)$$

$$VA(j) = \Sigma_i VA(ij) = \Sigma_i \{H_i^2 * P_{ij} * [1 - P_{ij}] / (X_{i.} - 1)\} \qquad (2.2)$$

and

$$VZ(j) = \Sigma_i VZ(ij) = \Sigma_i \{ H_i * (P_{ij} - P_j)^2 \} / F \qquad (2.3)$$

VA(j) describes the variation within the length strata and VZ(j) the variation between length strata.

Revision of the equation (2.3) into the form (2.3’) is used to separate the variance component, VA(j) in such a way that the calculation of the variation of the age readings, X is made possible. For the variance component, VZ(ij) the non-randomised value, F is already separated in the denominator. There is no change to this equation.

For the variance components, VA(j) the following equation can be used.

$$VA(j) = \Sigma_i VA(ij) = 1 / X * \{ \Sigma_i \{H_i^2 * P_{ij} * [1 - P_{ij}] / (X_{i.}/X - 1/X) \}\} \qquad (2.3')$$

In this equation the non-randomised value, X is nearly excluded in the sum equation. $X_{i.}/X$ is the relative frequency of age readings in the length class, i. The factor 1/X is still included in the sum equation after this transformation. But if X is large, a small change in X does not essentially change the denominator.

The Variance, $V(p_j)$ can now be written as

$$V(p_j) = 1 / X * \{ \Sigma_i \{H_i^2 * P_{ij} * [1 - P_{ij}] / (X_{i.}/X - 1/X) \}\} + 1 / F * \{\Sigma_i \{ H_i * (P_{ij} - P_j)^2 \}\} \qquad (2.4)$$

and $V(p_j)$ can be assessed by

$$V(p_j) = 1 / X * VA°(j) + 1/F * VZ°(j) \qquad (2.5)$$

with

$$VA°(j) = \{ \Sigma_i \{H_i^2 * P_{ij} * [1 - P_{ij}] / (X_{i.}/X - 1/X) \}$$ and

$VZ°(j) \quad = \{\Sigma_i \{ H_i * (P_{ij} - P_j)^2 \}\}.$

The component VZ°(j) is only dependent on the relative proportions of age-length combinations within the length intervals. The influence of X on VA(j) is very low because $X_{i.}/X$ is a relative frequency and the influence of 1/X is low if X is large.

These equations illustrate the following:

a) It is not useful to reduce only one of the additive components VA(j) or VZ(j) because the limit of $V(p_j)$ is VZ(j) for $X \Rightarrow \infty$ and F is fixed as well as the limit of $V(p_j)$ is VA(j) for $F \Rightarrow \infty$ and X is fixed.

b) The component VA(j) can be influenced by the total number of age readings, X and the distribution of the age readings, $X_{i.}$ within length class i.

c) The component VZ(j) can be influenced only by the number of length measurements, F.

The non-randomised variables $X_{i.}$, X and F are indicated in bold and italics in Table 1.

Calculation of the confidence interval of p_j

In order to be able to estimate the accuracy of the estimated p_j it is important to consider confidence intervals. Half of the confidence interval, $D(p_j)$ of p_j can be estimated by

$$D(p_j) \quad = (V(p_j))^{1/2} * 1{,}96 \qquad (3.1)$$

if the type I error is fixed at $\alpha = 0.05$ (COCHRAN 1972).

With equation (2.5), the variance, $V(p_j)$ can be estimated and the confidence interval (3.1) can be calculated for all age groups, j. In Table A1 the components VA(j) and VZ(j) are shown.

In order to design a sampling programme and estimate the necessary number of age readings and length measurements, the following information is necessary:

- an unbiased estimate of the variance, $V(p_j)$,
- the chosen type I error α, and
- definition of the required level of accuracy of the p_j.

The following equations demonstrate two possible ways to fix / choose the accuracy of half the confidence interval

a) $D(p_j) \quad \leq d$ for example, $d = 0.05$

b) $D(p_j) \leq k * p_j$ with k constant, for example, k = 0.1:

Method a) is demonstrated in Figure 1. The width of the confidence interval is independent of p_j .

Method b) is illustrated in Figure 2. The width of the confidence interval is dependent on p_j

Figure 1 Method a)

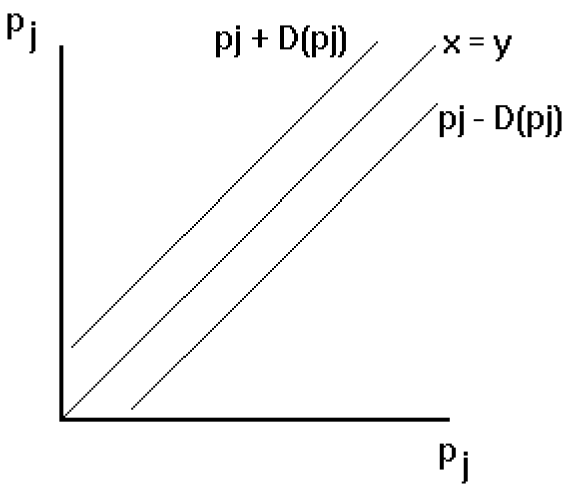
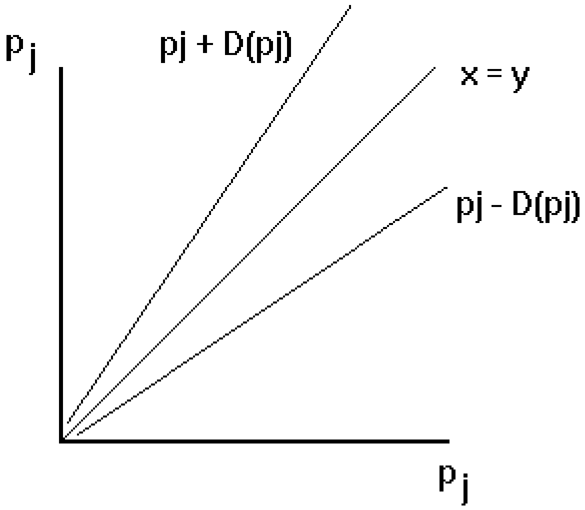


Figure 2 Method b)



The estimated sample size is dependent on the method chosen. Investigations show that method b) has significant consequences. The necessary number of age readings and length measurements is very high if p_j is lower than 10%. In this case, the sample sizes are very high for age groups which constitute only a small proportion of the catch. Calculations made by the Institute of Baltic Sea Fishery, Rostock use the first method where:

d = 0.05

α = 0.05 (type I error)

Experiments showed that the greatest sample size for any age groupin an ALKT were for the age group representing the highest proportion of the catch because $V(p_j)$ is the maximum.

The estimation of the necessary sample sizes, F' and X', depending on the chosen demand and the $V(p_j)$, can be calculated with

$$(VA°(j) / X' + VZ°(j) / F') * 1.96^2 \leq d^2. \quad (3.2)$$

In this case the VA°(j) and VZ°(j) are the estimates from the preliminary experiment.

The goal of this strategy is to make it possible to vary the number of age readings and length measurements, assuming that the pattern of variance in the next experiment is almost the same.

If the number of age readings, X' is chosen with

$$X' > 1.96^2 / d^2 * VA°(j) \quad (3.3)$$

the necessary number of length measurements, F' could be calculated by

$$F' \geq VZ°(j) / \{ d^2 / 1.96^2 - VA°(j) / X' \}. \quad (3.4)$$

If we choose the number of length measurements, F' with

$$F' > 1.96^2 / d^2 * VZ°(j) \quad (3.5)$$

the necessary number of age readings, X' could be calculated by

$$X' \geq VA°(j) / \{ d^2 / 1.96^2 - VZ°(j) / F' \}. \quad (3.6)$$

The interpretation of these equations shows that different numbers of age readings (X') and length measurements (F') can be calculated to produce the same accuracy. However, in practice age readings are more difficult and time consuming than length measurements.

In order to take this into account a variable, Z is introduced in the next step. Z can be interpreted as representing the ratio of the effort of one age determination and one length measurement. This factor can be interpreted as a possible price for the fish, the necessary work in the laboratory, etc.

Now the sample sizes F' and X' can be calculated, depending on Z, using the following equations

$$F' = [Z * VA°(j) + VZ°(j)] / (d^2 / 1.96^2) \quad (3.7)$$

and

$$X' = F' * Z \quad (3.8)$$

Table 2 shows the estimated number of age readings (X') and length measurements (F') necessary, depending on the variable Z for the ALKT of Table A1.

Table 2: The necessary size of age samples, X' and length samples, F' using VA(j), VZ(j) and $X_{i.}$ for the AGE-LENGTH-KEY data of Table A1.

Restriction (A) with d = 0.05 and $\alpha = 0.05$

Age group

	0		1		2		3		4		5		6		8	
Z	F'	X'	F'	X'	F'	X'	F'	X'	F'	X'	F'	X'	F'	X'	F'	X'
1	239	239	206	206	126	126	128	128	72	72	20	20	6	6	3	3
2	284	142	299	150	213	106	218	109	133	56	35	18	11	6	7	3
3	329	110	382	131	300	100	308	103	154	51	50	17	16	5	10	3
4	374	94	484	121	388	97	397	99	195	49	66	16	21	5	13	3
5	420	84	577	115	475	95	487	97	235	47	81	16	26	5	16	3
6	465	77	670	110	563	94	577	96	276	46	97	16	31	5	19	3
7	510	73	762	109	650	93	666	95	317	45	112	16	36	5	22	3
8	555	69	855	107	738	92	756	94	358	45	128	16	41	5	26	3
9	601	67	948	105	825	92	846	94	399	44	143	16	46	5	29	3
10	646	65	1040	104	913	91	935	94	439	44	159	16	51	5	32	3

Estimation of the sample sizes of F' and X' for different years show that the chosen value of Z should not be too high. If Z is higher than five, the number of length measurements increases while the number of age readings decreases very slowly.

The calculated sample size will be different for each age group. This is caused by the different proportions of each age group in the catch and the need for the same level of accuracy for all age groups.

In practice the maximum sample size should be used based on the chosen Z. If this strategy is used, the accuracy for age groups with a lower estimated sample size will actually be higher than the required level of accuracy.

Distribution of age readings per length class, Xi.

As explained in the previous step two of the three non-random variables (F', X', $X'_{i.}$) can be chosen from the ALKT. The following describes how, $X_{i.}$ can be changed in order to minimize the sample size required for age readings, X' and length measurements, F' still further.

Table A2 shows the values of VA(ij) calculated from the ALKT in Table A1.

There are great differences between the VA(ij). The aim of the optimisation here is the reduction of the VA(ij) without changing the total number of age readings (X).

From equations and tables the following can be concluded:

By definition the VA(ij) are zero if

a) only one age group is in the length class, i (P_{ij} is zero or one for all j) or

b) the proportion, H_i of the length distribution of the length class (i) is zero.

The VA(ij) are high if

a) there are two age groups both present in the same proportions ie. 50%, in the length class, i because $max_j(P_{ij} * (1 - P_{ij})) = 0.5 * 0.5$ and

b) the proportion, H_i of the length class, i in the length distribution is high.

In order to optimize the $X_{i.}$ within each length class the following options are possible:

A) Choose the $X_{i.}$ proportional to max_j VA(ij) for all i.

B) Choose the $X_{i.}$ proportional to sum_j VA(ij) for all i.

C) Or any other criterion.

Interpretation of these results is essential for the development of an age reading strategy. If missing values or bias are included in the data of the preliminary experiment, the resulting algorithm will produce biased estimates. For this reason it is necessary to evaluate the optimized sample structure.

The calculations are based on the variant A). The estimated number of age readings per length class, $X'_{i.}$ were chosen proportional to the maximum VA(ij) throughout all age groups, j.

$$X'i. = X' * max_j (VA(ij)) / \Sigma_i max_j (VA(ij)) \quad (4.1)$$

$$X''i. = X' * \Sigma_j (VA(ij)) / \Sigma_i \Sigma_j (VA(ij)) \quad (4.2)$$

The following step is a further optimization procedure. After the calculation of $X'_{i.}$, they can be used to estimate a new variance of p_j with chosen $X'_{i.}$. For this procedure the length distribution, H_i and the proportions, p_{ij} of all age groups, j and all length classes i are the same as in the ALKT of Table A1. Only $X_{i.}$ has been replaced by $X'_{i.}$. Based on this new estimation of $V(p_j)$, as well as the new structure of VA(ij), new sample sizes for the total number of age readings, X'' and the length measurements, F'' necessary can be calculated.
The results of this calculation are shown in Table 3.

Table 3 : The necessary size for age samples, X" and length samples, F" using the VA(ij), VZ(ij) and optimum X'i. from Table 4. Restriction (A) with d = 0.05 and α = 0.05

Age group

	0		1		2		3		4		5		6		8	
Z	F"	X"	F"	X"	F"	X"	F"	X"	F"	X"	F"	X"	F"	X"	F"	X"
1	211	211	152	152	101	101	134	134	98	98	21	21	10	10	2	2
2	228	114	190	95	163	82	229	115	164	82	39	19	18	9	3	1
3	245	82	228	73	226	75	325	108	231	77	56	19	27	9	4	1
4	262	66	266	66	288	72	420	105	297	74	73	18	36	9	5	1
5	280	56	303	61	351	70	516	103	364	73	91	18	44	9	6	1
6	297	49	341	57	413	69	611	102	430	72	108	18	53	9	8	1
7	314	45	379	54	476	68	707	101	497	71	125	18	62	9	9	1
8	331	41	417	52	538	67	802	100	563	70	142	18	70	9	10	1
9	349	39	455	51	601	67	898	100	630	70	160	18	79	9	11	1
10	366	37	493	49	663	66	993	99	696	70	177	18	88	9	13	1

As expected, the comparison of results from Table 2 and Table 3 shows that the necessary sample sizes for age readings (X") and length measurements (F") are lower if optimized $X'_{i.}$ are used instead of non-optimized X_i. Further investigations illustrated that the two optimization criterion variants, A and B, produce nearly the same $X'_{i.}$. Table 4 shows the optimal number of age readings per length class in the ALKT. $X'_{i.}$ were calculated using criterion A) and $X''_{i.}$ using criterion B). While there are big differences between $X_{i.}$ and $X'_{i.}$, the differences between $X'_{i.}$ and $X''_{i.}$ are very small.

Table 4: Optimum distribution of age samples using data from Table A1

Xi. - number of fish in the age sample in length class i

$X'_{i.}$ - optimum corresponding to the max_j VA(ij)

$X''_{i.}$ - optimum corresponding to the sum_j VA(ij)

X°i. - estimated optimum number of age readings per length class (Z = 3, X" = 82, Table 3) with corrections

Length in mm	$X_{i.}$	$X'_{i.}$	$X''_{i.}$	$X°_{i.}$
100	0	0	0	2
105	0	0	0	2
110	5	0	0	2
115	8	0	0	2
120	13	0	0	2
125	13	0	0	2
130	16	0	0	2
135	13	0	0	2
140	13	0	0	2
145	13	0	0	2
150	13	0	0	2
155	26	17	20	2
160	18	42	49	4
165	26	60	71	6
170	26	146	153	13
175	26	0	0	5
180	26	85	95	8
185	26	71	76	6
190	29	32	38	5
195	24	14	10	5
200	39	16	19	5
205	39	25	21	5
210	42	57	51	5

Length in mm	$X_{i.}$	$X'_{i.}$	$X''_{i.}$	$X°_{i.}$
215	39	44	47	5
220	39	97	91	8
225	39	58	52	5
230	39	93	75	6
235	39	54	46	5
240	47	37	30	5
245	42	16	14	5
250	55	5	6	5
255	39	12	14	5
260	47	5	5	5
265	31	2	1	5
270	39	1	1	5
275	13	0	0	5
280	5	0	0	5
285	8	0	0	5
290	10	0	0	5
295	5	0	0	5
300	0	0	0	5
305	5	0	0	5
310	0	0	0	5
sum	1000	989	985	194

Correction of the estimated number of age readings per length class, $X'_{i.}$

Since an estimate of the variance is necessary for the estimation of the sample size, a preliminary experiment must be carried out. In most areas of fishery research (surveys, commercial samples over short time periods) it is not possible to repeat the investigations again to use the results as a preliminary experiment because ageing is a time consuming procedure and the results are normally useable later.

Therefore, it is advantageous if the results from the previous year can be used in place of a preliminary experiment. The suitability of data from the preliminary experiment depends on the following conditions:

- the data do not include any bias (ie. age determinations exist for all length classes, the influence of errors in ageing can be negated, the influence of migration processes is neither low, the length distribution, nor the growth of the age groups vary greatly from year to year, etc.).

Only if the above conditions are true, can the results from the previous year used instead of a preliminary experiment without any reservations.

Examples where data from the previous year could be used include young fish surveys and commercial sampling of stocks with low migration and relatively constant recruitment.

Since there are variations in growth and year-class strength, it is necessary to select the optimal sampling strategy using previous experience.

The recommendation of X'i. implies some problems.

1. The results are dependent on historic data from previous years.
2. If a length class, i contains only one age group, the number of age readings $X'_{i.}$ necessary will be zero.
3. For length classes i with $F_i = 0$ the recommended number of age readings, $X'_{i.}$ will also be zero.

If are used the $X'_{i.}$ the following year without corrections it may be impossible to calculate the age of many length classes. Also, a change in the growth of fish results in a change in the length distribution of the age groups. This is the reason why experience recommends a correction of $X'_{i.}$. For this the following aspects should be considered:

Especially in young fish surveys, a range of length classes may be observed within only one age group (age group 1 or 2). Acoustic surveys in the western Baltic Sea have found herring (*Clupea harengus*) with a total length of less than 10 cm in age class 1. However, many individuals are caught each year with a length range of between 5 and 10 cm. The recommendation for $X'_{i.}$ for this length classes is zero. Furthermore it is not possible to detect all changes in the length distribution of the age groups 1 and 2 using optimal $X'_{i.}$. This is the reason why two age readings should be the minimum number in these length classes.

Since the recommendation for $X'_{i.}$ is zero if $X_{i.}$ or H_i is zero too, it should be ensured that where necessary age readings are carried out in the next experiment for length classes which normally cover more than one age group (no catch, no age readings). These length intervals normally include two or more age groups. More than two otoliths should be read in these length classes and a minimum of 5 age readings should be carried out, if enough individuals can be caught. This problem can occur in length intervals where there are very low catches. The recommended correction can be interpreted also as „as many age readings as possible, but not more than 5“. Experience is necessary to judge the minimum number.

Due to this, the number of all age readings will be higher than the recommended X“. It therefore, follows that the accuracy of the current experiment will be higher than the required level of accuracy.

In Table 4 the corrected $X°_{i.}$ are shown in the last column.

The following standards should be used for surveys and commercial sampling:

a) <u>Surveys</u>

The ALKT for statistical units should be analysed for each survey. The recommended optimal sample sizes for X', F', $X'_{i.}$ and $X°_{i.}$ should be recorded in a special table. These data can then be used the following year for the same survey and for comparing the recommendations of successive years..

b) Commercial fishery

The ALKT for quarters and ICES subdivisions should be analysed for all commercial species . The recommended values of X', F', $X'_{i.}$ and $X°_{i.}$ can then be used next season.

Tables 5.1 to 5.3 and A3 illustrate the results of this special analysis for the autumn acoustic survey in ICES subdivision 24 for different years.

Table 5.1: Optimum number of length measurements (F') and age readings (X') based on the factor Z for the using method A) using $X_{i.}$

	year	1992		1993		1994
Z	F'	X'	F'	X'	F'	X'
1	250	250	286	286	142	142
3	330	110	767	256	285	95
5	410	82	1247	249	427	85

Table 5.2: Optimum number of length measurements (F'') and age readings (X'') based on the factor Z for the using method A) using $X'_{i.}$

	year	1992		1993		1994
Z	F''	X''	F''	X''	F''	X''
1	222	222	208	208	128	128
3	245	82	532	177	132	44
5	269	54	855	171	136	27

The large differences between the X', F' and X'', F'' of the different years were caused by a strong change in the proportion of older herring (*Clupea harengus*) (age group 2+) in subdivision 24.

In 1992 and 1994 the majority of herring (*Clupea harengus*) in the catch were either age group zero or one. However, in 1993 a higher proportion of older herring (*Clupea harengus*) had already migrated into subdivision 24.

Table 5.3: Estimated mean length (cm) per age group and standard deviation (Std) (N, number of age readings per age group, j)

age	year	1992			1993			1994	
group	N	Mean	Std	N	Mean	Std	N	Mean	Std
0	42	13,2	1,2	47	12,6	1,2	24	11,6	1,3

1	127	17,6	1,2	50	18,0	1,0	26	17,5	0,7
2	89	21,5	1,4	46	19,0	1,2	57	19,0	1,0
3	62	229	1,7	91	21,0	2,1	38	21,1	2,2
4	41	24,2	2,1	90	21,0	2,7	39	22,4	3,1
5	12	25,5	2,4	55	23,6	3,0	20	24,2	2,9
6	4	24,8	2,2	34	25,9	2,1	8	28,0	0,9
7				10	22,9	3,9	5	29,8	0,6
8				8	26,5	1,9			
9				2	28,1	9,4			
10				1					
11+									

The following conclusions can be drawn from this sample:

- there is no dramatic change in growth of the different year classes exists and
- the age readings seem to be unbiased.

The procedure described can assume an accuracy of $D(p_j) <= 0.05$ with, a minimum sample size for all estimated age proportions.

As the estimated optimum sampling strategy is essentially dependent on the model assumptions used, it would apear necessary to investigate those factors that can produce bias or high variances.
Possible causes could be:

- errors in age reading,
- a low number of length measurements,
- biased data sampling.

Since it is difficult to judge the accuracy of the preliminary experiment, a method has been developed which indicates any possible inaccuracies and defines an objective criterion.

Use of ALKT for indicating problems in age readings

With the AKLT data it is possible to estimate the mean length per age group and the variance.

If one assumes that the length is normally distributed for each age group, a theoretical total length distribution, F_{ti} can be calculated for the ALKT.

Investigations showed that the use of a log - normal distribution model for length data did not essentially produce any changes in the results.

If we use the notations

l_j mean length of the age group j

s^2_j variance of the length of age group j

we can estimate the mean length for age groups and the variance by

$$lj = \Sigma_j \{l_i * p_{ij} / \Sigma_j p_{ij} \tag{6.1}$$

$$s^2_j = \Sigma_j \{l_i^2 * p_{ij} / \Sigma_j p_{ij} \} - lj^2 \tag{6.2}$$

The cumulative length distribution of an age group, j can be described by

$$\Phi_j(l_{i+1}) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{l_{i+1}} e^{-\frac{(t-l_j)^2}{2s_j^2}} dt \tag{6.3}$$

If we denote the proportion of the length interval (l_i,l_{i+1}] of the age group, j by

$$\Phi_{ji} = \Phi_j(l_{i+1}) - \Phi_j(l_i) \tag{6.4}$$

we can further calculate the theoretical distribution of the length sample, F_{ti} with

$$F_{ti} = F * \Sigma_j \{ p_j * (\Phi_{ji}) \} . \tag{2.3.5}$$

Investigations over recent years have shown that the assumption that the length is normally distributed for all age groups holds true for all commercial fish in the Baltic Sea.

Comparing the observed (F_i) with the theoretical (F_{ti}) length distribution can indicate any possible problems due to age reading.

Figures B.1 to B.4 (see Appendix B) show some of the comparisons of the observed (line) and the theoretical (dots) length distributions for commercial species. Large differences in length classes with low frequencies can be an indication of possible problems arising from the age readings.

Extreme age-length combinations were found in all cases, when the difference between the observed (F_i) and the theoretical (F_{ti}) length distribution were large.

A more precise analysis of the influence of possible sources of error such as the selectivity of trawls, lake of complete cover of age groups within the trawl, inaccurate length samples and errors in age reading would be necessary in order to judge the accuracy of the theoretical length distribution.

Conclusions

The algorithm described can be used to estimate the accuracy of the proportion of age groups from ALKT. The number of age readings (X') and length measurements (F') necessary and the optimum distribution of the age readings per length class ($X'_{i.}$) (depending on the chosen criterion) can be determined according to the required level of accuracy. Where changes in length and age distribution occur from year to year, professional judgement is required to decide on the relevant correction.

It is necessary to point out that the results of the algorithm depend on the quality of the AKLT data . Bias in the sampling of otoliths or length measurements produces biased estimations of X',F' and $X'_{i.}$.

Comparison of observed (F_i) and computed theoretical (F_{ti}) length distributions can be used to routinely analyze the quality of the age readings.

References

Anon., 1994: Report of the Workshop on Sampling Strategies for Age and Maturity. ICES C.M. 1994 /D:1

Armstrong, D.W.; Iladia, S., 1986: On bias in estimates of population parameters obtained by means of length-stratified sampling for age/length keys. ICES, Doc. C.M. 1986 /G:10

Chester, A.J.; Waters,J.R., 1983: Two Stage Sampling for Age Distribution in the Atlantic Menhaden Fishery with comments on optimal Survey design. ICES C.M. 1983 /D:19

Cochran, W.G., 1972: Stichprobenverfahren. Walter de Gryter Berlin New York

Fridriksson, A., 1934: On the calculation of age-distribution within a stock of cod by means of relatively few age-determinations as a key to measurements on a large scale. Rapp. Proc.-Verb. Reun. Vol 86:1-5

Gudmunsdottir, A.; Steinarsson, B.A.; Stefánsson, G., 1988: A simulation procedure to evaluate the efficiency of some otolith and length sampling schemes. ICES, DOC. C.M. 1988 /D:14

Gulland, J.A., 1955: Estimation of growth and mortality in commercial fish populations. Min. Agric. and Fish Invest., er. 2, Vol. 18, No. 9. 46 pp

Gulland, J.A., 1957: Sampling problems and methods in fisheries research. FAO Fish Bull., Vol. 10: 157-181

Kimura, D.K., 1977: Statistical Assessment of Age-Length-Key. J. Fish. Res. Board Can., Vol. 34, 1977

Kutkuhn, J.H., 1963: Estimating absolute composition of California Salmon landings. Fish Bulletin No. 120

Lai, H.-L., 1987: Optimum Allocation for Estimating Age composition used AGE-LENGHT KEY. Fishery Bulletin; Vol. 85, No. 2, 1987

Nicholson, M.D.; Armstrong, D.W., 1974: A note on a possible source of bias in the estimation of mean length at age. ICES, C.M. 1974 / F:35

Pope, J.A., Knights, B., 1975: Sources of variations in catch at age data and the optimal use of age reading effort. ICES, C.M. 1975 /F:20

Rasch, D.,1968: Elementare Einführung in die mathematische Statistik. VEB Deutscher Verlag der Wissenschaften Berlin

Rasch, D.; Herrendörfer, G.; Bock, J.; Busch, K., 1978: Verfahrensbibliothek Versuchsplanung und -auswertung. VEB Deutscher Landwirtschaftsverlag Berlin

Schweigert, J.F.; Sibert, J.R., 1983: Optimizing survey design for determining age structure of fish stocks: an example from British Columbia Pacific herring (*Clupea harengus pallasi*). Can. J. Fish. Aquat. Sci. 40: 588-597

Tanaka, S., 1953: Precision of age-composition of fish estimated by double sampling method using the length for stratification. Bull. Japan. Soc. Sci. Fish., Vol. 19, No. 5: 657-670

Author's address: Rainer Oeberst, Institut für Ostseefischerei Rostock, An der Jägerbäk 2, 18069 Rostock, Germany, Fax: +49 381 810445

Appendix A: Large Tables

Table A1 : Example of an AGE-LENGTH-KEY Table as ASCII file of the IOR

FISH :	Herring	(029)	RACE / TYPE :	
AREA of CATCH:	Arkonasee	124-124	TIME :	221192-281192
FISHING :	Ground 10-49		STORAGE :	fresh / / / /
NUMB. of SAM.:			DEPTH of CATCH :	ALL
SPAWNING TYPE:	ALL		SEX :	ALL

LENGTH CLASS (mm)	AGE GROUPS 0	1	2	3	4	5	6	7	8	SUM	LENGTH DISTR. F	0/00	AGE unkn..	Mean W.(g)
85.00 -	-	-	-	-	-	-	-	-	-	-	2	0	-	-
90.00 -	-	-	-	-	-	-	-	-	-	-	10	0	-	-
95.00 -	-	-	-	-	-	-	-	-	-	-	32	2	2	-
100.00 -	-	-	-	-	-	-	-	-	-	-	65	3	3	-
105.00 -	-	-	-	-	-	-	-	-	-	-	195	9	9	-
110.00 -	2	-	-	-	-	-	-	-	-	2	405	19	-	-
115.00 -	3	-	-	-	-	-	-	-	-	3	728	35	-	-
120.00 -	5	-	-	-	-	-	-	-	-	5	998	48	-	-
125.00 -	5	-	-	-	-	-	-	-	-	5	1120	54	-	-
130.00 -	6	-	-	-	-	-	-	-	-	6	1247	60	-	-
135.00 -	5	-	-	-	-	-	-	-	-	5	1240	59	-	-
140.00 -	5	-	-	-	-	-	-	-	-	5	1466	70	-	-
145.00 -	5	-	-	-	-	-	-	-	-	5	808	39	-	4.0
150.00 -	5	-	-	-	-	-	-	-	-	5	765	37	-	-
155.00 -	8	2	-	-	-	-	-	-	-	10	376	18	-	-
160.00 -	3	4	-	-	-	-	-	-	-	7	382	18	-	-
165.00 -	4	6	-	-	-	-	-	-	-	10	565	27	-	2.6
170.00 -	3	6	-	1	-	-	-	-	-	10	831	40	-	-
175.00 -	-	10	-	-	-	-	-	-	-	10	959	46	-	-
180.00 -	-	8	1	-	-	1	-	-	-	10	802	38	-	-
185.00 -	-	7	2	1	-	-	-	-	-	10	628	30	-	-
190.00 -	-	6	5	-	-	-	-	-	-	11	431	21	-	4.2
195.00 -	-	3	3	2	1	-	-	-	-	9	213	10	-	4.1
200.00 -	-	-	10	5	-	-	-	-	-	15	387	18	-	-
205.00 -	1	1	7	4	2	-	-	-	-	15	384	18	-	-
210.00 -	-	1	7	6	1	-	1	-	-	16	611	29	-	3.4
215.00 -	-	-	9	5	1	-	-	-	-	15	577	28	-	4.5
220.00 -	-	1	4	8	2	-	-	-	-	15	787	38	-	6.3
225.00 -	-	-	3	7	4	-	1	-	-	15	592	28	-	4.8
230.00 -	-	1	6	5	2	-	-	-	1	15	728	35	-	8.0
235.00 -	-	-	3	7	2	3	-	-	-	15	562	27	-	13.1
240.00 -	-	-	4	7	6	1	-	-	-	18	514	25	-	16.4
245.00 -	-	-	3	4	8	-	1	-	-	16	324	15	-	8.4
250.00 -	-	-	-	15	6	-	-	-	-	21	273	13	-	18.1
255.00 -	-	-	-	4	11	-	-	-	-	15	349	17	-	36.7
260.00 -	-	-	-	4	11	3	-	-	-	18	218	10	-	40.1
265.00 -	-	-	-	5	3	2	2	-	-	12	97	5	-	37.9
270.00 -	-	-	-	1	6	5	3	-	-	15	104	5	-	32.3
275.00 -	-	-	-	1	3	-	1	-	-	5	32	2	-	-
280.00 -	-	-	-	-	2	-	-	-	-	2	60	3	-	-
285.00 -	-	-	-	-	-	3	-	-	-	3	11	1	-	60.0
290.00 -	-	-	-	-	-	1	2	-	1	4	8	0	-	-
295.00 -	-	-	-	-	-	1	1	-	-	2	2	0	-	219.5
300.00 -	-	-	-	-	-	-	-	-	-	-	-	-	-	-
305.00 -	-	-	-	-	-	1	1	-	-	2	2	0	-	125.0
310.00 -	-	-	-	-	-	-	-	-	-	-	38	2	2	-
SUM	60	56	67	92	71	21	13	-	2	382	20928	1002	16	10.2
p_j(0/00)	466	174	116	130	73	15	7	-	2	983			16	
VA(j)	77.1	157.8	149.0	152.7	69.5	26.4	8.5	-	5.4					
VZ(j)	6.0	3.5	1.2	1.2	1.0	0.1	0.0	-	0.0					
MEAN LENGTH	137.2	180.0	214.9	227.1	244.0	234.8	242.1	-	234.9	175.5				
MEAN WEIGHT	3.7	3.8	6.5	11.1	18.7	25.5	15.5	-	8.0	10.2				
CATCH in N	9.742	3.642	2.430	2.718	1.532	0.323	0.147	-	0.051	20.586	(* 10 ** 3)		0.335	

MEAN WEIGHT IN THE CATCH(g) INPUT : 39.700 CALCULATED : 10.185

CATCH(t) : 0.830844 CATCH in NUMBER 20.928 (* 10 ** 3)

Table A2 : The VA(ij) from the AGE-LENGTH-KEY of Table A1

Length	VA(ij) * 1.0E+9 Age								
(mm)	0	1	2	3	4	5	6	7	8	Sum
85.0	0	0	0	0	0	0	0	0	0	0
90.0	0	0	0	0	0	0	0	0	0	0
95.0	0	0	0	0	0	0	0	0	0	0
100.0	0	0	0	0	0	0	0	0	0	0
105.0	0	0	0	0	0	0	0	0	0	0
110.0	0	0	0	0	0	0	0	0	0	0
115.0	0	0	0	0	0	0	0	0	0	0
120.0	0	0	0	0	0	0	0	0	0	0
125.0	0	0	0	0	0	0	0	0	0	0
130.0	0	0	0	0	0	0	0	0	0	0
135.0	0	0	0	0	0	0	0	0	0	0
140.0	0	0	0	0	0	0	0	0	0	0
145.0	0	0	0	0	0	0	0	0	0	0
150.0	0	0	0	0	0	0	0	0	0	0
155.0	5738	5738	0	0	0	0	0	0	0	11477
160.0	13599	13599	0	0	0	0	0	0	0	27198
165.0	19436	19436	0	0	0	0	0	0	0	38872
170.0	36789	42045	0	15767	0	0	0	0	0	94601
175.0	0	0	0	0	0	0	0	0	0	0
180.0	0	26108	14686	0	0	14686	0	0	0	55479
185.0	0	21011	16008	9005	0	0	0	0	0	46023
190.0	0	10516	10516	0	0	0	0	0	0	21031
195.0	0	2877	2877	2238	1279	0	0	0	0	9272
200.0	0	0	5428	5428	0	0	0	0	0	10856
205.0	1496	1496	5985	4703	2779	0	0	0	0	16460
210.0	0	3330	13984	13318	3330	0	3330	0	0	37291
215.0	0	0	13031	12066	3378	0	0	0	0	28475
220.0	0	6285	19753	25140	11672	0	0	0	0	62851
225.0	0	0	9145	14225	11177	0	3556	0	0	38104
230.0	0	5378	20744	19207	9988	0	0	0	5378	60695
235.0	0	0	8242	12820	5952	8242	0	0	0	35256
240.0	0	0	6133	8433	7885	1862	0	0	0	24313
245.0	0	0	2434	2996	3995	0	936	0	0	10361
250.0	0	0	0	1736	1736	0	0	0	0	3473
255.0	0	0	0	3885	3885	0	0	0	0	7769
260.0	0	0	0	1103	1517	886	0	0	0	3507
265.0	0	0	0	475	366	271	271	0	0	1383
270.0	0	0	0	110	423	392	282	0	0	1207
275.0	0	0	0	94	140	0	94	0	0	327
280.0	0	0	0	0	0	0	0	0	0	0
285.0	0	0	0	0	0	0	0	0	0	0
290.0	0	0	0	0	0	9	12	0	9	30
295.0	0	0	0	0	0	2	2	0	0	5
300.0	0	0	0	0	0	0	0	0	0	0
305.0	0	0	0	0	0	2	2	0	0	5
310.0	0	0	0	0	0	0	0	0	0	0
Sum	77059	157819	148966	152748	69503	26352	8486	0	5387	646321

Table A3: Comparison of the $X_{i.}$, $X'_{i.}$ and $X°_{i.}$ for different years from the autumn acoustic survey in ICES subdivision 24

Length		1992			1993			1994	
(mm)	$X_{i.}$	$X'_{i.}$	$X°_{i.}$	$X_{i.}$	$X'_{i.}$	$X°_{i.}$	$X_{i.}$	$X'_{i.}$	$X°_{i.}$
80	0	0	2	2	0	2	0	0	2
85	0	0	2	2	0	2	0	0	2
90	0	0	2	5	0	2	9	0	2
95	3	0	2	5	0	2	9	0	2
100	0	0	2	5	0	2	5	0	2
105	0	0	2	5	0	2	5	0	2
110	0	0	2	7	0	2	9	0	2
115	5	0	2	7	0	2	9	0	2
120	8	0	2	9	0	2	9	0	2
125	5	0	2	9	0	2	9	0	2
130	5	0	2	9	0	2	9	0	2
135	5	0	2	9	0	2	9	0	2
140	16	0	2	7	0	2	9	0	2
145	11	0	2	12	0	2	9	0	2
150	27	22	2	16	0	4	18	15	2
155	27	45	4	18	0	4	0	0	2
160	34	178	15	21	6	4	9	0	2
165	48	160	13	39	6	4	14	0	5
170	48	181	15	46	26	5	23	187	5
175	45	0	5	25	83	15	65	64	5
180	56	22	5	46	116	21	92	126	5
185	50	18	5	35	114	20	92	77	5
190	45	20	5	46	88	16	74	113	5
195	29	25	5	46	42	8	37	52	5
200	19	45	5	30	117	21	65	43	5
205	27	48	5	30	53	10	46	37	5
210	50	25	5	30	58	10	41	34	5
215	42	47	5	39	21	5	18	46	5
220	61	41	5	35	47	8	28	49	5
225	64	31	5	35	15	5	14	14	5
230	48	26	5	21	41	7	18	19	5
235	37	14	5	37	15	5	5	0	5
240	29	14	5	35	34	6	9	28	5
245	21	5	5	25	25	5	14	11	5
250	29	3	5	37	21	5	23	29	5
255	27	1	5	53	9	5	23	10	5
260	29	4	5	32	11	5	23	11	5
265	11	5	5	25	1	5	23	9	5
270	21	2	5	30	3	5	23	4	5
275	11	0	5	14	17	5	23	2	5
280	5	3	5	30	9	5	23	9	5
285	0	0	5	12	2	5	5	0	5
290	0	0	5	14	7	5	23	1	5
295	0	0	5	5	0	5	18	0	5
300	3	0	5	0	0	5	5	0	5
305	0	0	5	0	0	5	5	0	5
310	0	0	5	2	0	5			5
									5
sum	1000	985	217	1000	987	276	1000	990	181
X	377			434			217		

Appendix B: Figures

Figure B 1: Comparison of the observed length distribution, F_i (line) and the theoretical length distribution, F_{ti} (dots) for samples of cod (*Gadus morhua morhua*) and pike-perch (*Stizostedion lucioperca*)

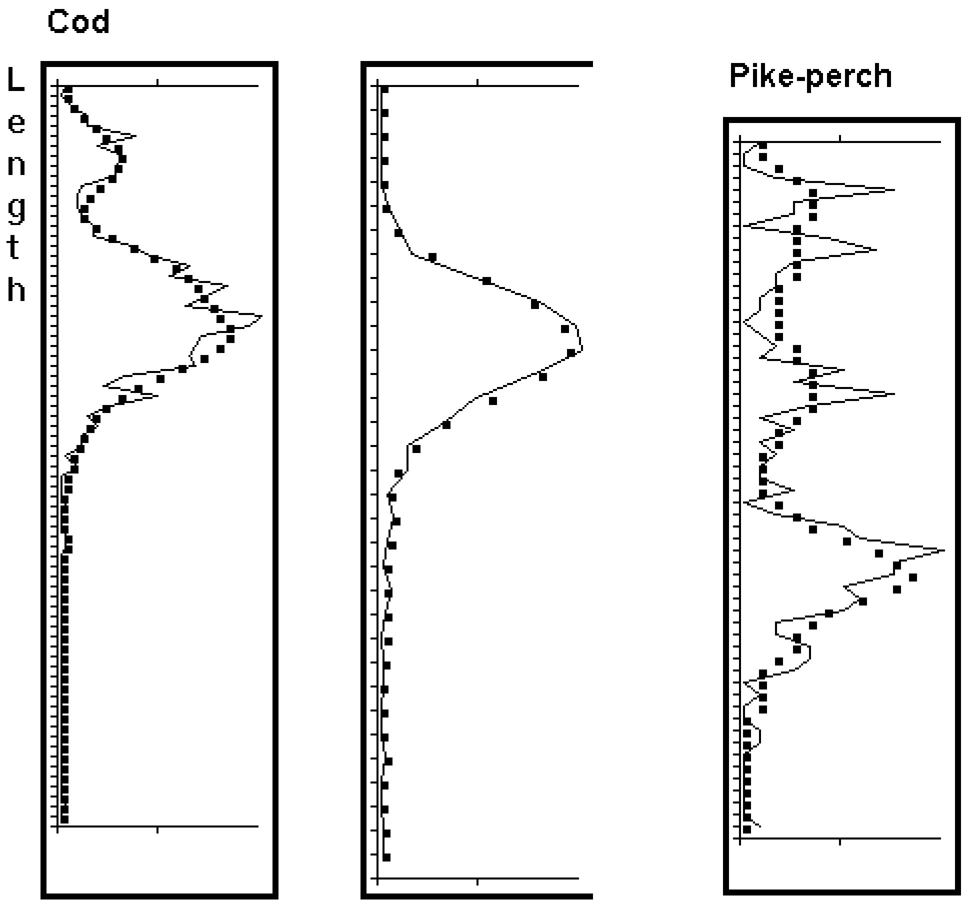


Figure B 2: Comparison of the observed length distribution, F_i (line) and the theoretical length distribution, F_{ti} (dots) for samples of herring (*Clupea harengus*)

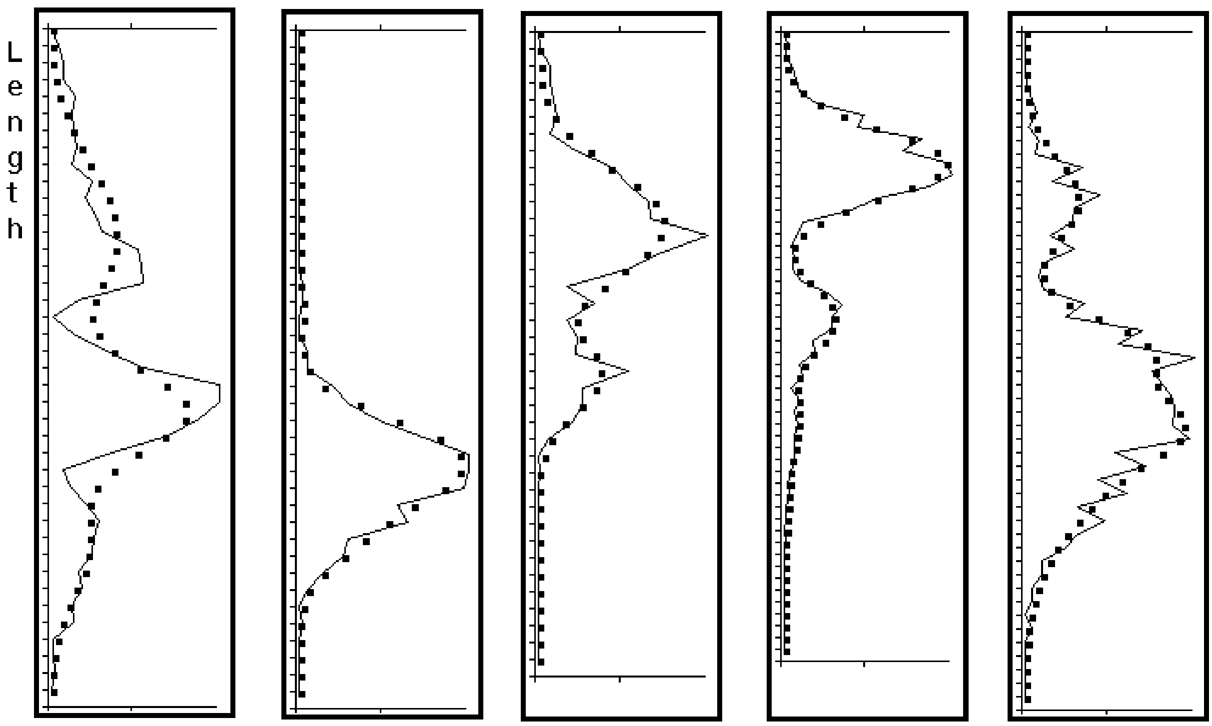


Figure B 3: Comparison of the observed length distribution, F_i (line) and the theoretical length distribution, F_{ti} (dots) for samples of sprat (*Sprattus sprattus*)

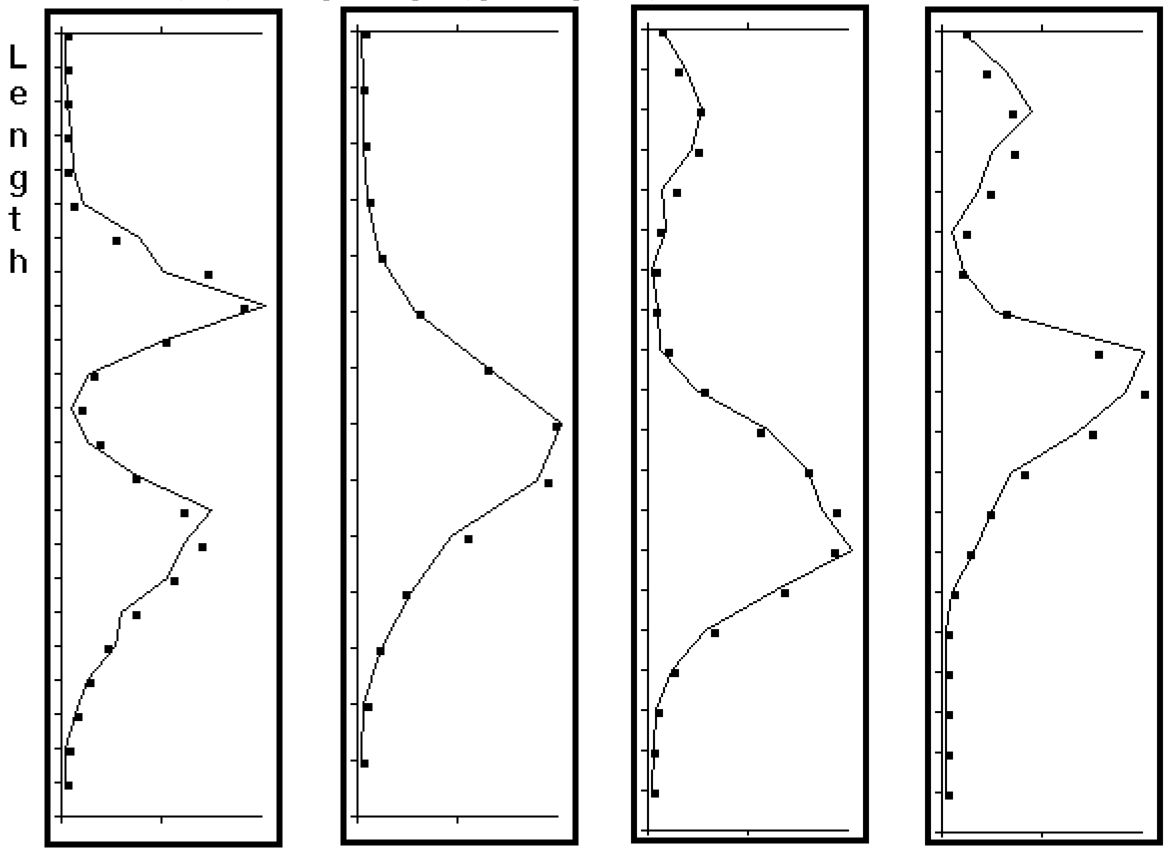
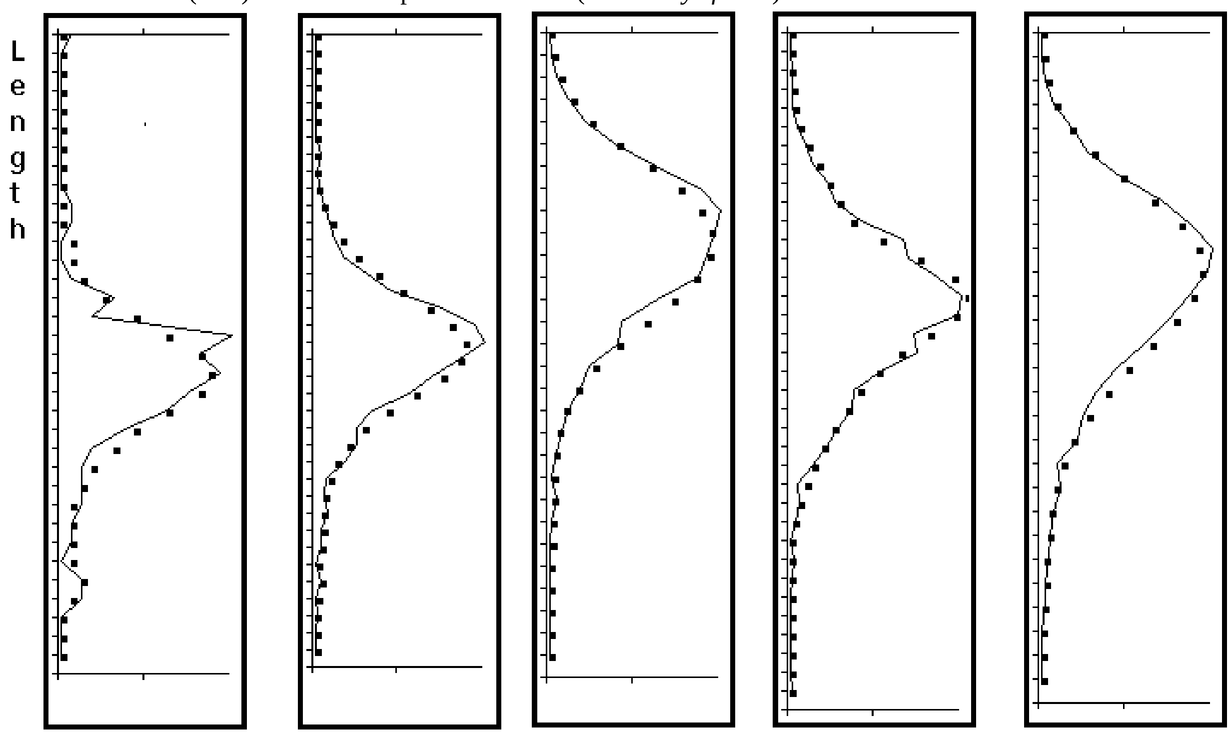


Figure B 4: Comparison of the observed length distribution F_i (line) and the theoretical length distribution F_{ti} (dots) for some samples of flounder (*Platichthys flesus*)



Working Paper | | **Workshop on Sampling and Calculation Methodology for Fisheries Data Nates 27 – 30 January 20043**

Precision of length-at-age and weight-at-age based on data from commercial landings - according to requirement of EU No. 1639/2001 -

by

R. Oeberst
Bundesforschungsanstalt für Fischerei Hamburg
Institut für Ostseefischerei Rostock
An der Jägerbäk 2
D - 18069 Rostock, Germany

Abstract

Simulation of weight data based on length data in combination with weight-length relation showed that weight of age groups is skewed distributed. Depending on the skewness of length data the skweness of weight data change. Statistical test showed that weight of age groups can be lognormal distributed, but, it is also possible that the hypothesis of lognormal distribution must be reject. The simulations and study of data from commercial landings showed that the precision of length-at-age better than the required precision if the required precision is reached for weight-at-age data due to the significant larger coefficient of variation of weight-at-age in comparison to length-at-age data.
Based on the data from commercial landings in 2002 the precision of length-at-age and weight-at-age were estimated and the necessary number of measurements were estimated which is necessary to reach the precision required by EU. The necessary number was estimated for both cases that weight of age groups is normally or lognormally distributed. The studies showed that in most cases the necessary number measurements is lower if it is assumed that weight is lognormally distributed due to the reduced effect of extreme large single weight data.
The studies show that the precision required by EU No. 1639/2001 is unrealistic because the necessary number of weight measurements can no be realised during the regular observation process.
An alternative definition of required precision is presented.

The decree (EG) No. 1639/2001 of the commission from 25.7.2001 Point I : Other biological parameters (Modul I) requires “the mean weight and total length of each age group is to assess with Degree 3 for all age groups which landings together are more than 95 % of the landings of the country of fish stock (Appendix XVI)”. The Degree 3 of precision is defined as the confidence interval of mean is supposed to be smaller than 5 % of the mean using first kind of error α=0.05. It is necessary to show every third year that these requirements are fulfilled.
Unfortunately necessary additional requirements were not defined.
Which landings should be used for defining the age groups for which the precision must be fulfilled, landings in number or landings in weight. Depending on the choice of one of the options the oldest age group change for that the precision should be fulfilled.

Furthermore, the time interval is not defined that should be used to estimate the precision. It is known that fish species growing some cm within a quarter with the effect that the variability of total length and weight increases with increasing time interval. Consequently the necessary number of measurements increases if the same level of precision is required. On the other hand the total number of necessary samplings probably also increases if the time interval is too small chosen due to the high variability of weight of age group in comparison to the increase of the mean weight from month to month.
It is well documented that the development weight of male and female individuals can significantly differ in the different periods of the year. Therefore, it is necessary to lay down whether the precision must be separately reached for both sexes
In 2002 an intensive sampling program of the commercial landings was carried out by the Federal Research Centre for Fisheries in Germany. The data of Baltic cod are used to check the precision of the length-at-age and weight-at-age. Furthermore, the necessary total number of measurements are estimated depending the required precision and proposal is presented to optimise the possible available number of measurements for obtaining the best accuracy.

Material and Methods

Report of the working group "Baltic Fish Assessment" (ICES 2003) in 2003 was used to check for which age groups the required precision must be reached for Baltic cod.
Simulations were use to investigate the relations between standard deviation of total length and weight based on weight length relation and to study the distribution function of weight dependent on the distribution of length.
Sampling data of commercial cod landings in 2002 were used for statistical analyses. Means and standard deviation of total length and weight were estimated by ICES subdivision, month or quarter, sex and age.
Chi-Square goodness-of-fit statistics, Shapiro-Wilks W statistics, Z score skewness and Z score kurtosis test were use to check whether length and weight of individuals of the same age group is normally or lognormally distributed:
T-test was used to compare the mean weight of the same age group by sex. Furthermore, it was tested whether significant change of mean weight occur from month to month within a quarter.
Following notations were used:

d	index of ICES subdivision
q	index of quarter
s	index of sex
a	index of age
$\alpha = 0.05$	first kind of error

$N_{d,q,s,a}$	number of measurements by subdivision d, quarter q, sex s, and age a
$\overline{W}_{d,q,s,a}$	mean weight
$V(W_{d,q,s,a})$	variance of weight
$T(\alpha, N_{d,q,s,a} - 1)$	quantil of T-distribution

According to the required precision of degree 3 the symmetric confidence interval is given by
$$\overline{W}_{d,q,s,a} \pm T(\alpha/2, N_{d,q,s,a} - 1)\sqrt{V(W_{d,q,s,a}} / \sqrt{N_{d,q,s,a}})$$
with half confidence interval
$$D_{d.q.s.a} = T(\alpha/2, N_{d,q,s,a} - 1)\sqrt{V(W_{d,q,s,a})} / \sqrt{N_{d,q,s,a}}$$

The necessary number of measurements $N*_{d,q,s,a}$ is given by

$$2*T(\alpha/2, N*_{d,q,s,a}-1)\sqrt{V(W_{d,q,s,a})}/\sqrt{N*_{d,q,s,a}} \leq 0.05*\overline{W}_{d,q,s,a}$$

or

$$[2*T(\alpha/2, N*_{d,q,s,a}-1)\sqrt{V(W_{d,q,s,a})}/0.05/\overline{W}_{d,q,s,a}]^2 \leq N*_{d,q,s,a}$$

Because the age of individuals, the current distribution of age and the current distribution of sex in landings is unknown during the sampling process and because the required precision is required for all significant age groups two options can be used for estimating the total number necessary measurements $N*_{total,d,q,s}$ by sex, subdivision and quarter.

As many as necessary individuals must be measured to be save with high probability that required precision is reached for all significant age groups using the maximum of $N*_{1,total,d,q,s}$ of all significant age groups for sex, quarter and subdivision based on:

$$N*_{1,total,d,q,s,a} = N*_{d,q,s,a}/N_{d,q,s,a}\sum_a N_{d,q,s,a}$$

estimates the necessary number of measurements by age groups of given subdivision, quarter and sex and

$$N*_{1,total,d,q,s} = Max_a N*_{1,total,d,q,s,a},$$

estimates the number of necessary measurements using those age groups for that the precision is required.

This method uses the proportions of age groups in landings $N_{d,q,s,a}/\sum_a N_{d,q,s,a}$ because it is necessary that the required precision is reached also for the age groups with low landings. Furthermore, it is hypothised that the age distribution of landings is nearly the same in the subsequent year.

Alternative method uses a two step algorithm. During the first step total number of individuals is measures as given by

$$N*_{2,total,d,q,s} = \sum_a N*_{d,q,s,a}$$

Using $N*_{2,total,d,q,s}$ the accuracy of age groups which are dominant in the catch is higher than the required level. Otherwise the accuracy of the oldest age group is to low due to the low proportion of these individuals in landings. Therefore, additional weighting of the large individuals is necessary. However, these measurement can not be used for estimating the proportion of age groups in the catch.

Results

Simulations

Following procedure was used for simulating the relation between the length and weight distribution of given age group based on 1001 simulation for each case. Length-weight-relation

$W_j = a * L_j^b * \varepsilon_j$

with a =0.00397 and b = 3.26 as regression parameter was used based on female cod captured in first quarter in ICES subdivision 22 (**Fig. 1**). ε_j is normally distributed with mean 1 and variance $\sigma^2_R = 0.122$, the residual variance of the regression.

Furthermore it is assumed that length of individuals Lj Є NV(E(L),σ^2_L) is normally distributed with mean E(L) and standard deviation $\sigma^2_{L.}$

The distribution of length was simulated using the following data

E(L)	given mean length
CV(L)	given coefficient of variation of length
S(L) = CV(L) * E(L)	resulting standard deviation of length
E(Lj)	mean length of simulated data
S(Lj)	standard deviation of the length of simulated data
CV(Lj)	coefficient of variation of simulated length distribution
E(Wj)	mean weight of simulated data
S(Wj)	standard deviation of simulated weight
CV(Wj)	coefficient of determination of simulated weight.

Using constant CV(L) = 0.15 mean, standard deviation and coefficient of variation of length and weight were estimated for different values of E(L) assuming the length is normally distributed. The simulations show that CV(Wj) is significantly higher than CV(Lj) (**Tab. 1**). That means that it is only necessary to estimate the number of measurements $N*_{d,q,s,a}$ depending on variability of weight since this number of measurements also saves the precision of length. Furthermore, the data show that CV(Wj) is independent of E(Lj) if CV(Lj) is constant. A constant standard deviation of length S(Lj) = 3.7 was chosen during second simulation. That means that CV(Lj) and CV(Wj) decrease with increasing mean length (**Tab. 2**) and that CV(Wj) is dependent on E(Lj). Furthermore, distribution of Wj was studied. The used test (Chi-Square goodness-of-fit statistics, Shapiro-Wilks W statistics, Z score skewness and Z score kurtosis) showed the simulated Wj are neither normally nor lognormally distributed. **Figure 2** presents the density traces of length distribution for E(L) = 25 and 45 with S(L) = 3.7. **Figure 3** shows the density traces of the corresponding weight based on the described length-weight relation. These simulation showed that in both cases the skewness and the kurtosis of Wj were larger than the values that is expected for normal distribution (**Tab. 3**). That means that the assumption of an symmetric confidence interval of Wj is not true. Further simulations showed that Wj is lognormally distributed if the length of age groups Lj is lognomally distributed (skewed distributed) (**Tab. 4**).
Summing up the simulations follows that the non linear relation between length and weight of fish results in a skewed distribution of the weight of age groups. Skewness of weight increases with increasing skewness of length. Depending on the skewness of length distribution of age group it is possible that weight of individuals is lognormally distributed. Therefore, it seems to be useful to compare the estimated total number of necessary measurements for the assumptions that weight is normally or lognormally distributed. Especially, if it must be taken into account that length of age groups is skewed distributed due to problems in ageing.

<u>Age groups for that the precision is required</u>

Based on the estimates of Assessment working group in 2003 (ICES 2003) the landings of cod in length and weight and sum of landing by subdivision, quarter and age group are given in **Table 5 to 7**. According to the requirement of EU No. 1639/2001 significant age groups are marked for which the precision is required. In all cases the precision is required for age groups 2 and 3. Whether the precision is also required for age group 4+ is depending on whether the landings in number or weight are used. The tables also show that age groups changes from quarter to quarter for which the precision must be reached.

<u>Studies of individuals captured in 2002</u>

Statistical test showed that total length of age groups combined for month/quarter, sex and age group are neither normally nor lognormally distributed in the most cases probably due to

inaccuracies in the ageing process. Consequently the distribution structure of weight is also different. In some cases the hypothesis can not be rejected that weight is lognormally distributed. In other cases weight is neither normally nor lognormally distributed. These results are the reason why it is very difficult to estimate the necessary sampling size with one procedure. Especially if some extreme large individuals were observed in age groups weight of these individuals are lognormally distributed. In these cases CV(Lj) is large.
Statistical test (T-test) also showed that mean weight of male and female cod significantly differed in some cases (e.g. subdivision 22, age group 2, month 2 and 3). Further study showed that significant increase of weight is possible within quarter (e.g. subdivision22, age group 3, males and month 1 and 3). These results suggest that the mean weight by subdivision, quarter and age group should be separately estimated by male and female. The significant increase of weight by age groups within quarter shows that the period for which the precision is to estimate is not simple to define. However, it must be taken into account that the increase and weight within a quarter is relative low in relation to the high variability of weight. That means that the necessary number of measurements increases by the power of three if period of month is used as basis for estimating the precision since the variance of weight of different month is nearly the same. Therefore, **it is proposed that period of quarter is used in the future**.

Estimation of necessary number of measurements based on the mean weight of cod in 2002

Landings of cod in subdivision 22 were dominated by age group 2 and 3 (**Tab. 8 to 10**, Number). Furthermore, landings of age group 1 are large in subdivision 24 in fourth quarter. Standard deviation of length strongly varied from age group to age group in the same subdivision and quarter probably in some cases influenced by ageing process. The data suggest that standard deviation of length is close to 3.5 for all age groups. Standard deviation larger than 3.5 are mostly observed for age groups with low sampling number. In other cases length is skewed distributed (**Fig. 4**). Data also show that CV(Lj) and CV(Wj) is correlated with R = 0.87 based on 82 data. Standard deviation of weight also strongly varied, but, it can be assumed that standard deviation of weight can be expected in the range from 250 to 450. As it can be expected from the simulations weight of age groups is skewed distributed.
Using standard deviation of weight necessary number of weight measurement $N*_{d,q,s,a}$ by subdivision, quarter, sex and age, and the total number of measurements $N*_{1,total,d,q,s}$ and $N*_{2,total,d,q,s}$ based on the assumption that weight is normally distributed is given in **Table 11** to **13**. Additionally the proportion of age groups of landings and of necessary number of weight measurements $N*_{d,q,s,a}$ is presented. The same estimates are added based on the assumption that weight is lognormally distributed. The necessary number of measurements are marked for each estimation using the assumption the those age groups are significant which together present more than 95 % of landings. The estimates showed that in all cases (subdivision, quarter, sex and age) the required number of measurements is higher than the realized number. Consequently follows that the required precision of weight-at-age was not reached for Baltic cod. The quotient CV(Wj) / E(Wj) varied between 0.07 and 0.37 for the significant age groups. These values are larger than the required 0.05.
The estimated necessary number of weight by age group $N*_{d,q,s,a}$ varied between 270 and 2400 if it is assumed that the weight of age groups is normally distributed. Comparable values are estimated if it is assumed that the weight of age groups is lognormally distributed. In most cases $N*_{d,q,s,a}$ is larger based on the assumption of normally distributed weight, and in some cases the difference is large (e. g. $N*_{24,3,M,4}$) probably due to the large skewness of the

distribution. Therefore, data based on lognormally distribution should be used for checking the precision.
In all cases $N*_{1,total,d,q,s}$ are larger or equal than $N*_{2,total,d,q,s}$ due to the large difference of proportion of age groups in landings and the estimated necessary number of measurements by age groups $N*_{1,d,q,s,a}$. $N*_{1,total,d,q,s}$ varies between 1616 (d=24, q=3, s=F) and more than 200 000 (d=22, q=4, s=M). It is easy to see that such number of measurements can not be realized. $N*_{2,total,d,q,s}$ varies between 1540 (d=25,q=1,s=M) and 4685 (s=24, q=4, s=M). That means that least 2000 males and 2000 females by subdivision and quarter must be weighted if it is taken into account that additional weights of large individuals are necessary. Using these estimates about 12 000 individuals must be weighted by quarter assuming the proportion of male is about 50 % for Baltic cod landings. But, besides cod samples additional samples from other commercial landings must be taken.

These data illustrate that the defined precision required by EU No. 1639/2001 is unrealistic and must be change as fast as possible.

Proposal for defined precision

The studies showed that V(Wj) and CV(Wj) strongly varied by subdivision, quarter, sex and age group and that consequently $N*_{d,q,s,a}$ also strongly varied. That means that for each unit (subdivision, quarter, sex and age group) different sampling sizes must be realized. Furthermore, it is known that E(Wj) and V(Wj) are not only influenced by the variability of length and weight of age group, these parameters are although influences by subjective factors as small measurement and typing errors if the data are stored in the computer. Furthermore, during the process of ageing subjective errors are possible which can significantly influence E(Wj) and V(Wj). For excluding these possible effects concerning the estimated necessary number of measurement the following procedure is proposed:

Precision is required for subdivision, quarter, sex and the tree age groups with the highest proportions in landings assuming that weight of age groups is lognormally distributed.
It is required that confidence interval of weight of age groups. CI(Wj), is smaller or equal than 10% of mean weight, E(Wj).
If furthermore it is assumed that the coefficient of variation CV(Wj) is nearly constant for the significant age groups between 0.30 and .035. Then the necessary number can be estimated by

$$2*T(\alpha/2, N*-1)\sqrt{V(W)}/\sqrt{N*} \leq 0.1\overline{W}$$

Using the constant CV(Wj) follows

$$\sqrt{V(W)} = CV(W_j)\overline{W}_j,$$

$$2*T(\alpha/2, N*-1)CV(W_j)\overline{W}/\sqrt{N*} \leq 0.1\overline{W}$$

and

$$2*T(\alpha/2, N*-1)CV(W_j)\overline{W}/(0.1\overline{W}) \leq \sqrt{N*} \ = \ 2*T(\alpha/2, N*-1)CV(W_j)/0.1 \leq \sqrt{N*}$$

Using the equation follows that the necessary number of measurements are dependent of $CV(W_j)$ and the required width of confidence interval of weight.
N* is about 180 for $CV(W_j) = 0.3$ and about 250 for $CV(W_j) = 0.35$.
Based on this assumption the necessary number of weight data should be 250 individuals multiplied by the number of significant age groups. That mean that 1 500 individuals of Baltic cod should be measured by subdivision, quarter. That is a constant value which can be planned. These requirements mean that the precision of age groups with the highest

proportion in landings is better than the required level and that precision of age groups with small proportion in landings is lower than the required level.

Addendum: precision of sex by subdivision, quarter and age group

Notations:

$P_{d,q,a}$ proportion of female in subdivision d, quarter q and age group a

$P_{d,q,a} * (1- P_{d,q,a})$ variance of $P_{d,q,a}$

Required precision is

$$2T(\alpha/2, N*_{d,q,a} - 1)\sqrt{V(P_{d,q,a}}\sqrt{N*_{d,q,a}} \leq 0.05P_{d,q,a}$$

Using this equation $N*_{d,q,a}$ is larger than 1000. That means that the required precision for estimating the proportion of sex ration is also unrealistic.

The use of 250 individuals save a precision of less than $0.1*P_{d,q,a}$.

Tables

Table 1: Relation between CV(Lj) and CV(Wj) based on simulated data using constant coefficients of determination, CV(Lj), for different mean length, E(Lj)

E(L)	E(Lj)	S(Lj)	CV(Lj)	E(Wj)	S(Wj)	CV(Wj)
25	25.1	3.7	0.1486	158.7	79.5	0.501
25	24.9	3.8	0.152	157.5	79.3	0.503
25	24.9	3.7	0.149	154.8	74.6	0.482
25	24.8	3.8	0.151	154.8	80.2	0.518
25	24.8	3.7	0.148	153.3	72.6	0.474
35	35.0	5.3	0.151	472.3	227.3	0.481
35	35.1	5.4	0.153	474.3	237.9	0.502
35	34.8	5.2	0.149	464.4	230.9	0.497
35	35.0	5.4	0.155	472.9	236.8	0.501
35	35.3	5.4	0.154	483.6	239.3	0.495
45	45.0	6.7	0.150	1072.4	546.1	0.509
45	45.1	6.9	0.153	1076.2	536.2	0.498
45	45.0	6.8	0.151	1070.0	550.4	0.154
45	45.4	6.6	0.145	1099.2	540.3	0.492
45	44.7	6.7	0.151	1046.2	519.0	0.496
55	54.8	8.1	0.148	2031.3	1018.3	0.501
55	54.9	8.0	0.147	2049.6	1006.1	0.491
55	55.1	8.7	0.158	2099.8	1094.5	0.521
55	55.1	8.1	0.148	2051.5	1001.6	0.488
55	55.3	8.2	0.148	2096.1	1024.0	0.489

Table 2: Relation between CV(Lj) and CV(Wj) based on simulated data using constant standard deviation of length, S(Lj), for different mean length, E(Lj)

E(L)	E(Lj)	S(Lj)	CV(Lj)	E(Wj)	S(Wj)	CV(Wj)
25	24.9	3.7	0.149	155.7	74.5	0.478
25	25	3.7	0.148	157.6	78.2	0.496
25	24.9	3.6	0.145	155.2	73.2	0.472
25	24.9	3.6	0.145	155.6	75	0.482
25	25.1	3.7	0.147	159.5	76.5	0.480
35	35	3.7	0.106	452.3	158.1	0.350
35	35	3.7	0.106	445.2	161.9	0.364
35	34.8	3.7	0.106	443	160	0.361
35	35.2	3.7	0.105	458.1	165.5	0.361
35	34.9	3.8	0.109	450.9	170.7	0.379
45	44.9	3.6	0.080	1007.6	298.8	0.297
45	44.8	3.9	0.087	1002.3	310.8	0.310
45	45	3.7	0.082	1017.1	298.9	0.294
45	44.9	3.7	0.082	999.9	299.2	0.299
45	45.1	3.6	0.080	1020.4	298.2	0.292
55	55	3.7	0.067	1927.9	481.1	0.250
55	54.8	3.6	0.066	1903.8	460.6	0.242
55	54.8	3.8	0.069	1888.3	474.2	0.251
55	55.1	3.7	0.067	1938.8	482.6	0.249
55	55	3.8	0.069	1942.1	504.6	0.260

Table 3: Distribution parameter of simulated length and weight data related by $W_j = a * L_j^b * \varepsilon_j$ using normally distributed length with S(L) = 3.7 for different mean length E(L)

E(L)	25	45
Number of simulations	2002	2002
E(Lj)	24.99	44.97
S(Lj)	3.76	3.67
Skewness(Lj)	0.79	0.42
Kurtosis(Lj)	-1.11	0.74
E(Wj)	158.01	1007.02
S(Wj)	79.91	293.54
Skewness(Wj)	22.24	12.29
Kurtosis(Wj)	24.25	6.50
E(ln(Wj))	4.94	6.87
S(ln(Wj))	0.52	0.29
Skewness(ln(Wj))	-6.04	-3.80
Kurtosi(ln(Wj))	0.84	1.07

Table 4: Distribution parameter of simulated length and weight data related by $W_j = a * L_j^b * \varepsilon_j$ using lognormally distributed length

E(L)	34	55
Number of simulations	3003	3003
E(Lj)	33.60	55.05
S(Lj)	5.01	8.38
Skewness(Lj)	10.67	10.23
Kurtosis(Lj)	4.85	3.07
E(ln(Lj))	3.50	4.00
S(ln(Lj))	0.15	0.15
Skewness(ln(Lj))	0.24	0.50
Kurtosis(ln(Lj))	1.02	-1.40
E(Wj)	415.89	2075.34
S(Wj)	222.40	1126.43
Skewness(Wj)	39.41	38.96
Kurtosis(Wj)	61.74	64.47
E(ln(Wj))	5.91	7.51
S(ln(Wj))	0.50	0.51
Skewness(ln(Wj))	-0.10	0.69
Kurtosis(ln(Wj))	1.37	-1.73

Table 5: Landings in number and weight in subdivision 22 in 2002 based on Report of WGBFAS 2003 and sum of landings

	Catch in number				Mean weight in catch				Sum of catch in number				Sum of catch in weight			
Age	1. Quarter	2. Quarter	3. Quarter	4. Quarter	1. Quarter	2. Quarter	3. Quarter	4. Quarter	1. Quarter	2. Quarter	3. Quarter	4. Quarter	1. Quarter	2. Quarter	3. Quarter	4. Quarter
1				40704				550	0	0		3.4	0	0.0		1.8
2	1120364	442916		1070278	595	601		1023	48.8	56.1		92.8	30.6	38.3		88.6
3	1007868	272381		81408	1130	1055		1545	92.7	90.6		**99.6**	83.0	79.7		**98.6**
4	94129	41944		2394	1700	1479		2299	**96.8**	**95.9**		99.8	90.4	88.6		99.0
5	27550	26843		2394	2494	2139		5193	98.0	99.3		100	93.5	**96.9**		100.0
6	41335	5527			2756	3894			99.8	100		100	**98.8**	100		100.0
7	4592				5894				100	100		100	100.0	100.0		100.0
Total	2295838	789611		1197178												

Table 6: Landings in number and weight in subdivision 24 in 2002 based on Report of WGBFAS 2003 and sum of landings

	Catch in number				Mean weight in catch				Sum of catch in number				Sum of catch in weight			
Age	1. Quarter	2. Quarter	3. Quarter	4. Quarter	1. Quarter	2. Quarter	3. Quarter	4. Quarter	1. Quarter	2. Quarter	3. Quarter	4. Quarter	1. Quarter	2. Quarter	3. Quarter	4. Quarter
1	2076	22258	22423	874769	456	511	528	635	0.5	5.2	7.5	53.6	0.2	2.8	3.5	44.3
2	323409	136542	122877	497770	724	667	769	808	78.4	37.1	48.6	84.1	62.1	25.4	31.2	76.4
3	80956	199891	118393	226853	1317	988	1328	1102	**97.9**	83.8	88.2	**98.0**	90.2	74.3	77.4	**96.3**
4	7059	58212	34083	31009	3098	1195	2173	1385	99.6	**97.4**	**99.6**	99.9	**96.0**	91.5	**99.2**	99.7
5		8133	1196	1632		2443	2241	2182	99.6	99.3	100	100	96.0	**96.5**	100.0	100.0
6	1661	2140			9083	4268			100	99.8	100	100	100.0	98.7	100.0	100.0
7		856				5998			100	100	100	100	100.0	100.0	100.0	100.0
Total	415161	428032	298972	1632033												

Table 7: Landings in number and weight in subdivision 25 in 2002 based on Report of WGBFAS 2003 and sum of landings

	Catch in number				Mean weight in catch				Sum of catch in number				Sum of catch in weight			
Age	1. Quarter	2. Quarter	3. Quarter	4. Quarter	1. Quarter	2. Quarter	3. Quarter	4. Quarter	1. Quarter	2. Quarter	3. Quarter	4. Quarter	1. Quarter	2. Quarter	3. Quarter	4. Quarter
1				1593				455	0			0.3	0.0			0.1
2	87151			109149	592			541	20.3			19.7	16.9			11.3
3	247285			368872	674			877	77.9			85.2	71.4			72.5
4	93591			26523	909			1198	**99.7**			**89.9**	**99.3**			78.6
5	1288			49396	1721			1803	100			98.7	100.0			**95.4**
6				7170				3386	100			100	100.0			100.0
Total	429315			562703												

Table 8: Distribution parameter of length and weight in subdivision 22 by quarter, sex and age

SD	Quarter	Sex	Age	Number	E(Lj)	S(Lj)	CV(Lj)	E(Wj)	S(Wj)	CV(Wj)
22	1	F	2	316	37.6	3.6	9.5	546.7	173.3	31.7
22	1	F	3	295	47.4	4.8	10.2	1241.4	411.4	33.1
22	1	F	4	37	53.6	3.7	6.8	1796.1	462.0	25.7
22	1	F	5	7	60.9	3.4	5.5	2643.1	481.1	18.2
22	1	F	6	4	63.5	3.4	5.3	2922.3	537.5	18.4
22	1	M	1	1	32.0	0.0	0.0	379.0	0.0	0.0
22	1	M	2	342	37.3	3.6	9.7	549.5	162.8	29.6
22	1	M	3	311	46.3	4.0	8.6	1051.8	284.1	27.0
22	1	M	4	22	53.0	5.1	9.7	1558.5	492.1	31.6
22	1	M	5	3	57.0	0.8	1.4	2077.3	98.3	4.7
22	2	F	1	1	30.0	0.0	0.0	286.0	0.0	0.0
22	2	F	2	173	39.5	3.2	8.1	625.5	168.8	27.0
22	2	F	3	70	47.4	4.8	10.1	1088.0	349.5	32.1
22	2	F	4	4	52.5	4.4	8.4	1450.0	421.3	29.1
22	2	F	5	3	52.0	4.3	8.3	1586.0	298.8	18.8
22	2	M	1	1	29.0	0.0	0.0	221.0	0.0	0.0
22	2	M	2	200	38.0	3.3	8.6	554.5	141.1	25.5
22	2	M	3	128	46.6	4.2	9.0	991.3	273.2	27.6
22	2	M	4	17	49.5	5.8	11.7	1219.5	433.9	35.6
22	2	M	5	1	50.0	0.0	0.0	1013.0	0.0	0.0
22	4	F	1	58	35.3	3.8	10.9	441.1	148.1	33.6
22	4	F	2	357	46.6	3.8	8.2	1027.8	254.7	24.8
22	4	F	3	44	56.9	6.0	10.6	1941.6	657.6	33.9
22	4	F	4	1	61.0	0.0	0.0	2528.0	0.0	0.0
22	4	F	5	4	71.3	7.1	10.0	4072.3	1817.5	44.6
22	4	M	1	92	35.1	3.6	10.3	433.3	146.9	33.9
22	4	M	2	408	46.0	3.5	7.5	986.0	221.2	22.4
22	4	M	3	41	53.4	5.1	9.5	1570.8	497.0	31.6
22	4	M	4	2	59.5	4.5	7.6	2555.5	868.5	34.0
22	4	M	5	1	68.0	0.0	0.0	3864.0	0.0	0.0

Table 9: Distribution parameter of length and weight in subdivision 24 by quarter, sex and age

SD	Quarter	Sex	Age	Number	E(Lj)	S(Lj)	CV(Lj)	E(Wj)	S(Wj)	CV(Wj)
24	1	F	1	8	28.5	2.6	8.9	253.0	67.6	26.7
24	1	F	2	92	39.0	4.9	12.7	678.8	271.0	39.9
24	1	F	3	28	49.6	3.7	7.5	1318.0	306.8	23.3
24	1	F	4	4	64.3	2.5	3.9	3080.8	539.6	17.5
24	1	F	6	1	103.0	0.0	0.0	9683.0	0.0	0.0
24	1	M	1	15	29.3	2.9	10.0	285.5	94.3	33.0
24	1	M	2	105	37.7	4.8	12.6	607.2	229.2	37.8
24	1	M	3	14	49.4	3.1	6.3	1361.1	323.2	23.7
24	2	F	1	17	32.5	3.4	10.4	382.6	136.6	35.7
24	2	F	2	22	38.4	4.4	11.4	674.3	240.9	35.7
24	2	F	3	34	45.4	4.0	8.9	1113.2	282.6	25.4
24	2	F	4	10	47.6	5.5	11.5	1283.9	485.3	37.8
24	2	F	5	4	62.8	3.1	5.0	3046.5	655.3	21.5
24	2	F	7	1	83.0	0.0	0.0	5329.0	0.0	0.0
24	2	M	1	39	31.8	3.4	10.8	340.9	112.2	32.9
24	2	M	2	61	38.0	3.1	8.1	599.6	169.3	28.2
24	2	M	3	63	44.0	4.0	9.1	926.8	288.5	31.1
24	2	M	4	16	46.4	5.4	11.6	1136.8	420.3	37.0
24	2	M	5	2	52.0	2.0	3.8	1456.0	194.0	13.3
24	3	F	1	44	32.9	4.5	13.5	387.9	154.2	39.8
24	3	F	2	110	41.6	3.7	8.9	742.2	203.3	27.4
24	3	F	3	110	51.3	4.0	7.8	1359.6	344.4	25.3
24	3	F	4	28	59.9	6.6	11.0	2146.6	728.9	34.0
24	3	F	5	1	70.0	0.0	0.0	3128.0	0.0	0.0
24	3	M	1	47	32.8	3.2	9.8	367.6	109.8	29.9
24	3	M	2	72	42.6	4.2	9.8	804.0	237.1	29.5
24	3	M	3	62	49.9	4.3	8.7	1251.1	384.9	30.8
24	3	M	4	19	60.1	5.5	9.1	2160.8	495.6	22.9
24	3	M	5	1	53.0	0.0	0.0	1513.0	0.0	0.0
24	4	F	0	3	26.0	0.8	3.2	191.3	11.8	6.2
24	4	F	1	489	38.4	4.9	12.8	632.4	251.2	39.7
24	4	F	2	382	43.3	5.4	12.5	856.7	357.7	41.8
24	4	F	3	148	49.5	7.3	14.8	1280.8	864.2	67.5
24	4	F	4	32	54.9	7.0	12.7	1576.9	797.0	50.5
24	4	F	5	3	67.7	2.1	3.0	2686.3	746.5	27.8
24	4	M	0	7	26.7	3.4	12.8	212.7	66.3	31.2
24	4	M	1	322	36.8	4.4	12.0	545.3	208.6	38.2
24	4	M	2	210	42.3	5.9	13.9	797.8	374.7	47.0
24	4	M	3	119	47.5	6.4	13.5	1100.1	603.4	54.8
24	4	M	4	12	47.7	2.9	6.1	960.6	246.7	25.7
24	4	M	5	2	84.0	10.0	11.9	8283.0	2994.0	36.1
24	4	M	0	2	25.5	2.5	9.8	193.0	61.0	31.6

Table 10: Distribution parameter of length and weight in subdivision 25 by quarter, sex and age

SD	Quarter	Sex	Age	Number	E(Lj)	S(Lj)	CV(Lj)	E(Wj)	S(Wj)	CV(Wj)
25	1	F	1	2	26.0	1.0	3.8	161.5	5.5	3.4
25	1	F	2	28	34.5	5.5	15.9	452.2	233.5	51.6
25	1	F	3	57	41.2	3.6	8.7	733.8	193.8	26.4
25	1	F	4	21	46.3	3.8	8.1	1023.6	215.6	21.1
25	1	F	5	2	53.0	6.0	11.3	1417.0	304.0	21.5
25	1	M	1	4	27.3	1.1	4.0	196.0	38.4	19.6
25	1	M	2	48	35.4	4.0	11.2	476.9	173.3	36.3
25	1	M	3	52	39.1	2.9	7.3	629.5	140.5	22.3
25	1	M	4	33	44.5	2.7	6.1	932.2	161.7	17.3
25	4	F	1	2	31.5	0.5	1.6	315.5	34.5	10.9
25	4	F	2	77	38.8	4.7	12.2	665.6	305.9	46.0
25	4	F	3	215	44.1	3.9	8.8	981.7	289.2	29.5
25	4	F	4	162	48.5	3.5	7.2	1312.5	292.6	22.3
25	4	F	5	23	55.0	3.9	7.2	1836.4	390.3	21.3
25	4	F	6	5	69.2	5.7	8.3	3425.6	557.3	16.3
25	4	M	1	12	32.4	2.1	6.5	355.1	86.2	24.3
25	4	M	2	132	36.5	3.6	9.9	526.8	171.2	32.5
25	4	M	3	322	41.3	3.5	8.5	762.7	208.9	27.4
25	4	M	4	179	46.7	3.2	6.9	1095.4	251.7	23.0
25	4	M	5	33	53.8	4.2	7.8	1589.5	436.6	27.5

Table 11: Number of necessary measurements based on the assumption of normally and lognormally distributed weight for cod captured in subdivision 22 by quarter, sex and age

SD	Quarter1	Sex	Age	$N^*_{d,q,s,a}$ normally distributed weight	$N^*_{1,total,d,q,s}$ normally distributed weight	$N^*_{2,total.d,q,s}$ normally distributed weight	Proportion of age group in landings	Proportion of age group in $N^*_{d,q,s,a}$ normally	$N^*_{d,q,s,a}$ lognormally distributed weight	$N^*_{1,total,d,q,s}$ lognormally distributed weight	$N^*_{2,total,d,q,s}$ lognormally distributed weight	Proportion of age group in $N^*_{d,q,s,a}$ lognormally
22	1	F	2	811	1693	811	52	48	639	1335	639	45
22	1	F	3	886	1982	1697	48	52	767	1716	1406	55
22	1	F	4	535	**9540**	**2231**	6	19	535	**9543**	**1941**	22
22	1	F	5	269	25391	2501	1	10	208	19611	2149	9
22	1	F	6	275	45360	2776	1	10	255	42075	2404	11
22	1	M	1									
22	1	M	2	708	1406	708	50	33	661	1312	661	34
22	1	M	3	590	1287	1298	46	28	540	1179	1201	28
22	1	M	4	804	**24826**	**2102**	3	38	714	**22037**	**1915**	37
22	1	M	5	21	4721	2123	0	1	21	4753	1936	1
22	2	F	1									
22	2	F	2	588	853	588	69	25	500	725	500	24
22	2	F	3	833	2985	1421	28	35	680	2438	1180	33
22	2	F	4	682	**42768**	**2102**	2	29	557	**34952**	**1737**	27
22	2	F	5	288	24119	2391	1	12	325	27192	2062	16
22	2	M	1									
22	2	M	2	524	1208	909	43	22	470	1083	755	23
22	2	M	3	613	2209	1406	28	26	610	2197	1220	30
22	2	M	4	1020	**27673**	**2331**	4	44	910	**24677**	**2005**	45
22	2	M	5									
22	4	F	1	909	7272	909	13	23	755	6040	755	25
22	4	F	2	497	645	1406	77	13	465	604	1220	15
22	4	F	3	925	**9755**	**2331**	9	24	785	**8278**	**2005**	26
22	4	F	4									
22	4	F	5	1604	186079	3935	1	41	1060	122960	3065	35
22	4	M	1	927	5482	927	17	30	740	4376	740	27
22	4	M	2	407	543	1334	75	13	412	549	1152	15
22	4	M	3	808	10717	2142	8	26	647	8585	1799	24
22	4	M	4	931	**253351**	**3074**	0	30	895	243440	2694	33
22	4	M	5									

Table 12: Number of necessary measurements based on the assumption of normally and lognormally distributed weight for cod captured in subdivision 24 by quarter, sex and age

SD	Quarter1	Sex	Age	$N^*_{d,q,s,a}$ normally distributed weight	$N^*_{1,total,d,q,s}$ normally distributed weight	$N^*_{2,total.d,q,s}$ normally distributed weight	Proportion of age group in landings	Proportion of age group in $N^*_{d,q,s,a}$ normally	$N^*_{d,q,s,a}$ lognormally distributed weight	$N^*_{1,total,d,q,s}$ lognormally distributed weight	$N^*_{2,total,d,q,s}$ lognormally distributed weight	Proportion of age group in $N^*_{d,q,s,a}$ lognormally
24	1	F	1	578	9602	578	6	23	620	10308	620	26
24	1	F	2	1284	1856	1862	69	50	1140	1648	1760	48
24	1	F	3	439	2083	2300	21	17	347	1648	2107	15
24	1	F	4	250	**8300**	**2550**	3	10	270	**8978**	**2377**	11
24	1	F	6									
24	1	M	1	879	7856	879	11	35	920	8219	920	39
24	1	M	2	1149	1466	2028	78	46	1060	1353	1980	45
24	1	M	3	456	**4367**	**2484**	10	18	382	**3656**	**2362**	16
24	2	F	1	1028	5323	1028	19	25	990	5125	990	26
24	2	F	2	1029	4115	2057	25	25	980	3920	1970	26
24	2	F	3	521	1348	2578	39	13	505	1307	2475	13
24	2	F	4	1151	**10133**	**3729**	11	28	980	**8624**	**3455**	26
24	2	F	5	375	8247	4104	5	9	301	6622	3756	8
24	2	F	7									
24	2	M	1	873	873	873	22	25	1050	1050	1050	36
24	2	M	2	644	412	1517	34	18	465	297	1515	16
24	2	M	3	782	484	2299	35	22	610	378	2125	21
24	2	M	4	1102	**2685**	**3401**	9	31	663	**1616**	**2788**	23
24	2	M	5	146	2839	3546	1	4	145	2828	2933	5
24	3	F	1	1273	8480	1273	15	38	1030	6859	1030	33
24	3	F	2	606	1614	1879	38	18	585	1558	1615	19
24	3	F	3	519	**1382**	**2398**	38	16	505	**1345**	**2120**	16
24	3	F	4	930	9729	3328	10	28	968	10129	3088	31
24	3	F	5									
24	3	M	1	721	3081	721	23	28	710	3036	710	29
24	3	M	2	702	1960	1423	36	27	633	1767	1343	26
24	3	M	3	764	**2477**	**2186**	31	29	612	**1984**	**1955**	25
24	3	M	4	426	4504	2612	9	16	452	4782	2407	19
24	3	M	5									
24	4	F	0	34	11841	34	0	0	34	11979	34	1

24	4	F	1	1271	2748	1305	46	14	1100	2378	1134	20
24	4	F	2	1404	3886	2709	36	16	1075	2975	2209	20
24	4	F	3	3663	**26159**	**6372**	14	40	1355	**9677**	**3564**	25
24	4	F	4	2056	67923	8428	3	23	1133.3	37434	4697	21
24	4	F	5	624	219775	9052	0	7	727	256146	5424	13
24	4	M	0	784	75457	784	1	9	800	77029	800	14
24	4	M	1	1179	2467	1962	48	14	960	2009	1760	17
24	4	M	2	1776	5702	3739	31	21	1310	4204	3070	23
24	4	M	3	2421	13712	6160	18	28	1185	6712	4255	21
24	4	M	4	533	**29940**	**6693**	2	6	430	**24152**	**4685**	8
24	4	M	5	1053	354943	7746	0	12	1005	338685	5690	18
24	4	M	0	806	271614	8552	0	9	0	0	5690	0

Table 13: Number of necessary measurements based on the assumption of normally and lognormally distributed weight for cod captured in subdivision 2 by quarter, sex and age

SD	Quarter1	Sex	Age	$N^*_{d,q,s,a}$ normally distributed weight	$N^*_{1,total,d,q,s}$ normally distributed weight	$N^*_{2,total.d,q,s}$ normally distributed weight	Proportion of age group in landings	Proportion of age group in $N^*_{d,q,s,a}$ normally	$N^*_{d,q,s,a}$ lognormally distributed weight	$N^*_{1,total,d,q,s}$ lognormally distributed weight	$N^*_{2,total,d,q,s}$ lognormally distributed weight	Proportion of age group in $N^*_{d,q,s,a}$ lognormally
25	1	F	1	12	655	12	2	0	12	660	12	0
25	1	F	2	2146	**8431**	2158	25	62	1665	**6541**	1677	56
25	1	F	3	563	1087	2721	52	16	550	1061	2227	19
25	1	F	4	360	1883	**3081**	19	10	366	1917	**2593**	12
25	1	F	5	373	20514	3454	2	11	367	20185	2960	12
25	1	M	1	312	10692	312	3	15	290	9933	290	16
25	1	M	2	1065	**3039**	1377	35	53	865	**2469**	1155	48
25	1	M	3	403	1063	**1780**	38	20	385	1014	**1540**	21
25	1	M	4	245	1016	2025	24	12	270	1121	1810	15
25	2	F	1	99	23977	99	0	3	99	23958	99	4
25	2	F	2	1701	**10691**	1800	16	49	985	**6191**	1084	39
25	2	F	3	701	1577	2501	44	20	550	1238	1634	22
25	2	F	4	403	1203	**2903**	33	12	387	1156	**2021**	15
25	2	F	5	366	7704	3269	5	11	285	5997	2306	11
25	2	F	6	216	20880	3485	1	6	220	21296	2526	9
25	2	M	1	477	26929	477	2	16	485	27403	485	19
25	2	M	2	852	**4378**	1329	19	29	735	**3775**	1220	29
25	2	M	3	606	1276	1935	47	20	500	1053	1720	20
25	2	M	4	428	1619	**2363**	26	14	366	1386	**2086**	14
25	2	M	5	609	12520	2972	5	21	448	9204	2534	18

Figures

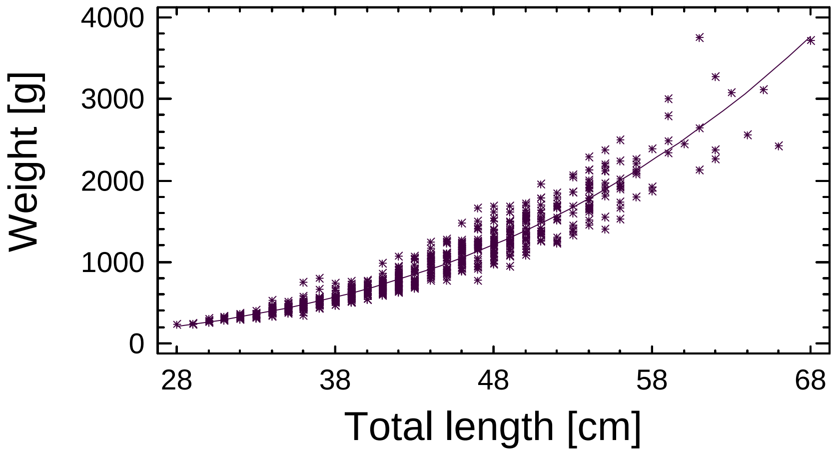


Figure 1: Weight-length relation of female cod captured in Subdivision 22 in first quarter 2002 – W = a =0.00397 * $L^{3.26}$, N = 660, R = 0.98

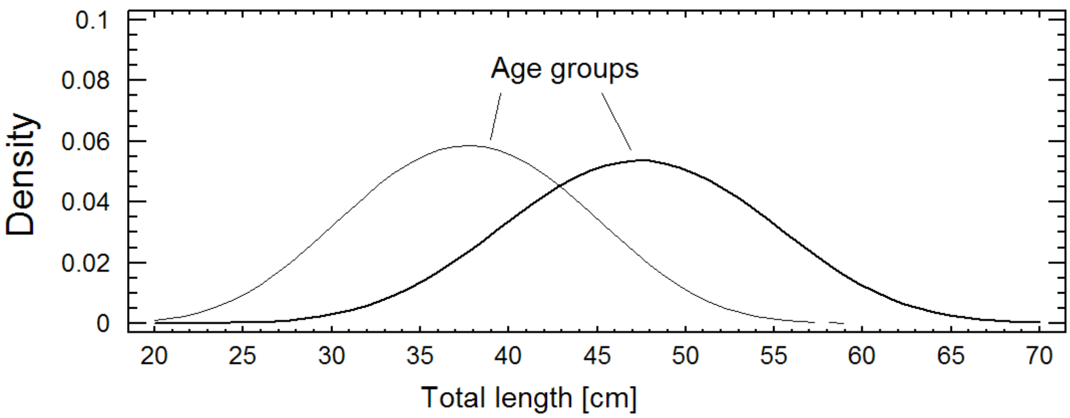


Figure 2: Density traces of length distribution for simulated age groups with mean length E(L) = 25 and 45 and standard deviation with S(L) = 3.7.

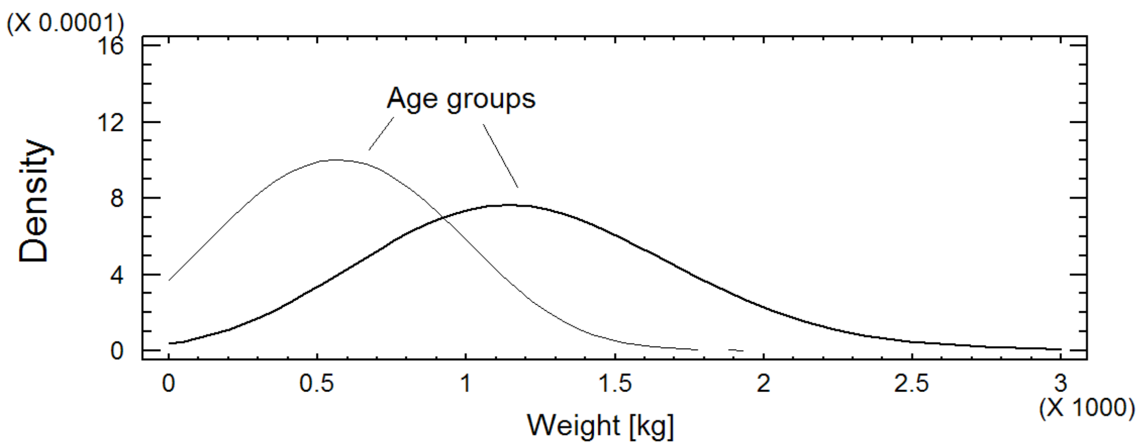


Figure 3: Density traces of the corresponding weight of simulated age groups based on the described length-weight relation

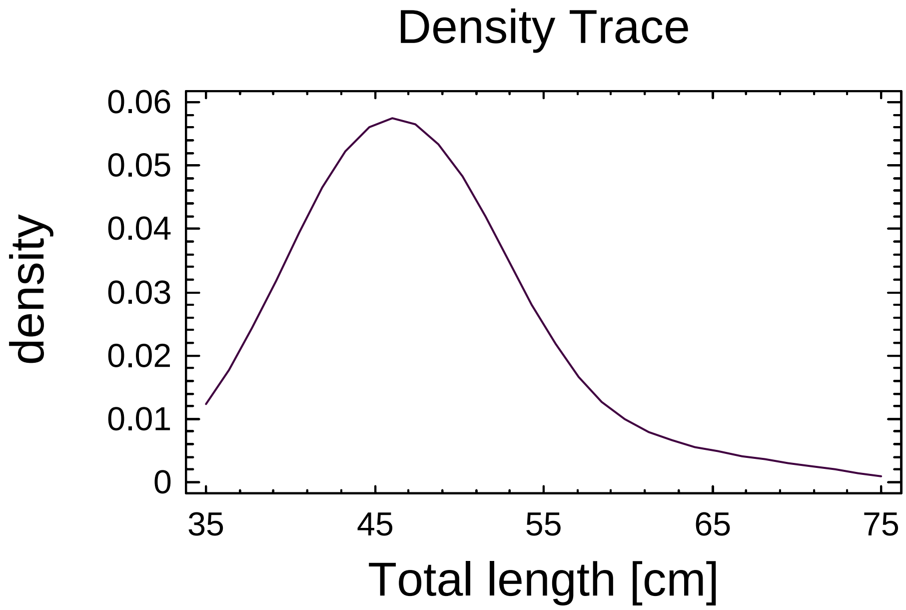


Figure 4: Density trace of female cod with age 3 captured in subdivision 24 in fourth quarter with standard deviation of length of 7.3

Precision in catch at age data with regard to sampling design

Joël Vigneau and Stéphanie Mahévas

Abstract

This working document is a contribution to the Term of Reference d) of the WKSCMFD : propose methods to estimate precision and design sampling stratification schemes that will minimise bias and maximise precision. Large sampling effort is needed all year long to provide working groups on stock assessment with essential data, the catch-at-length or catch-at-age data for each species combined at the International level. It is therefore essential to estimate the precision associated with these data with regards to the disagregated levels requested by the working groups. Each country is responsible for their own sampling which makes every national sampling design nearly unique. This document presents the analytical statistics and resampling techniques developed to analyse sampling shemes. Previous works have been largely used to develop this approach. A simulation algorithm is proposed to find graphically the optimum sampling effort for a target precision level. This paper also focus on exploratory analysis of sampling design and one statistic is proposed to quantify heterogeneities within and between strata. At the sample level, this statistic enables to point out possible outliers. At the strata level, the heterogeneities can be used to define *ad hoc* stratification by combining strata showing the same pattern or by adding another strata if patches are visible. Since the French sampling covers different fisheries, from the Mediterranean to the North Sea, different sampling designs are required to take into account the large disparities, mainly a market commercial category-based sampling and a fleet-based sampling. This analysis is applied to the two main french sampling designs using the 2002 Eastern Channel sole and Atlantic hake sampling database.

1 Introduction

The length and age distribution are used as input data to assess the state of the most important fish stocks all over the world. Inherent to any sampling procedure, the estimation of the length and age distribution may contain bias and uncertainties. The dissemination of input data errors or uncertainties in the assessment models have been studied by Kimura (1989), Restrepo and Powers (1990), Pelletier (1991) and more recently by Reeves (2003) and are given to be significant on stock evaluations. In the European fisheries sector, the sampling of the length and age distribution of species in the landings has always been under the Member States responsibility. As the national species/stock length and age distribution are combined to obtain the international distribution, there has been some regional coordination in the latest years (SAMFISH, FIEFA, North Sea sampling program) to avoid too much heterogeneity in the sampling process. To expand this coordination to every Member States, EU Regulation 1639/2001 hereafter named Data Directive has given a framework for the collection of data in the European fisheries sector since 2002. This Data Directive does not require a precision level for the numbers-at-length and numbers-at-age but the PGCCDBS (ICES, 2003) assumed that this information would be helpful to provide objective means for comparing national programmes, assist people involved in sampling and provide quality information on the input data for assessment models. Moreover, the Data Directive gives common rules for dealing with precision levels and sampling intensities. There are three case studies :

- when it is not possible to define quantitative targets for sampling programmes, neither in terms of precision levels, nor in terms of sample size, pilot surveys in the statistical sense will be established [. . .].

- When quantitative targets can be defined, they can be specified either directly by sample sizes or sampling rates, or by the definition of the levels of precision and of confidence to be achieved.

- When reference is made to a sample size or to a sampling rate in a population defined in statistical terms, the sampling strategies must be at least as efficient as simple random sampling. [. . .].

With respect to these common rules and to achieve PGCCDBS goal, the attention will be given on precision levels in number-at-length and number-at-age as a target and as a tool for sampling strategy analysis.
The length composition of the landings is assumed to be representative of the length composition in the stock to be used in the assessment models. Sampling the landings to estimate the overall species length structure introduces the notion of bias and precision with respect to sampling strategy and sampling effort. The first step of the study will be the research of heterogeneities within the length distribution related to factors like commercial market categories, time, fleet, harbours or areas. The heterogeneities evidences will permit to define stratas in the sampling procedure and try to minimize the variance of the numbers-at-length estimation by adopting a stratified sample design (Cochran, 1977). At a second stage, simple random sampling CV calculations will be proposed with respect to the Data Directive using analytical and resampling statistics. Only the resampling will be presented to estimate the precision of the stratified sampling CV's.

The age distribution sample is designed to be representative of the age distribution at length of the stock. The calculation of precision levels in numbers-at-age based on age and length sampling have been treated in numerous reports and the latest approach by Kimura (1977) was the starting point of new investigations. Kimura (1977) derivated the variance of an unbiased proportion at age estimator depending on a fixed, random and mixed allocation. He also introduced the VarTot function as an error index in the age length key (ALK). This function has been used to calculate coefficient of variation in the age distribution of NAFO cod (Baird, 1983) and to design an optimal sampling strategy for cod (Gavaris and Gavaris, 1983). Finally, Kimura (1989) proposed an analytical formula of the variance of the catch at age estimator, derived from the variance of the estimate of catch at weight and the variance of the age-weight relationship using the delta Method. This author also quantified the impact of this variability on the estimates produced by the cohort analysis for stock evaluation and discussed the choice of the allocation for age sampling, prefering a proportional allocation and showing the low impact the sampling level would have on the variance for older classes. Lai (1987) introduced a cost function to calculate the optimal sampling effort between length sample and age sample. This work was completed by Quinn et Deriso (1999) with the description of three optimisation methods in a stratified sampling.

Two sampling designs are considered in France to estimate landings length distribution (fig. 1). Both are first stratified by quarter, by harbours set and differed by their third strata which are either the commercial category or the métier, also called "captain sorting". For the stratified sampling design per métier, the selected units in the last strata are subsampled by commercial category. Samples are collected with a proportionnal allocation in each strata. For ageing, two approaches are practised whether the otholiths can be extracted without damaging the fish or not. In the former situation, a subsample of the length sample is taken, in the latter an independant age sample is collected.

One method will be proposed here corresponding to length sampling where sample weight is known and quarterly ALK with a fixed or proportional allocation. In accordance with the Data Directive, comparison will be made between simple random and stratified sampling to assess which method gives the best result in term of precision with a given cost. To perform this analysis, we will consider a single year age-length key agregating the traditional quater age-length keys. Examples of stock where sampling is stratified by fleet (Atlantic hake) and stocks where sampling is stratified by commercial categories (Eastern Channel sole) will be considered and different levels of sampling numbers will be given according to different levels of precision goal to be achieved.

2 Method

In this section, we first present 1/ the analytical formulation of the catches at age variance with regard to each of these sampling designs without the quater and harbours stages recommended by the CE and 2/ a bootstrap estimation of the catches at age variance for each completed stratified sampling design (i.e. with the quater and habour stages)as actually conducted in France. Then, a cost function is introduced to test the accuracy of the estimation according to sampling design and the number of samples required for a fixed level of precision.

2.1 Notations

We use the following notations:

- $\hat{D}$: the total landings estimator (in number)
- $\hat{D}_i$: the landings estimator at age i (in number)
- $\hat{p}_i$: the estimator of age i proportion of landings
- $\hat{D}_{ji}$: the estimator the j th class landings within age i landings

The sampling design considered in this analysis is stratified according to the following stratas :

Strata 1 : quater

Strata 2 : harbours set

Strata 3 : commercial category or mtier

The estimator of the landings at age i is

$$\hat{D}_i = \hat{D}\hat{pa}_i$$

and its variance estimator decomposed into three elements is:

$$\begin{aligned} Var(\hat{D}_i) &= Var(\hat{D}\hat{pa}_i) \\ &= \hat{pa}_i^2 Var(\hat{D}) + \hat{D}^2 Var(\hat{pa}_i) + Var(\hat{D}) Var(\hat{pa}_i) \\ &= V_1 + V_2 + V_3 \end{aligned} \quad (1)$$

This expression shows the great importance of the precision in estimation of the age-length key compared to the precision in the estimation of the landings. The second elementt (V_2) of the variance will indeed be the most determinant component of $Var(\hat{D}_i)$ since it is a function of the squared landings.

2.2 Analytical variance formulation

An analytical formulation of the aged landings variance estimator is available (e.g. in Deriso and Queen 1999) in the case of a random sampling design (without any stratification). The operational strata (commercial category or métier) of the french sampling bans from applying this analytical results on french sampling. Consequently, we delopped a more complex but still analytical formula of the variance. First, we explicit the variance of the total landings and the variance of the proportion of landings at age i. The total landings estimator is the result of the length sampling, while the estimator of the proportion of landings at age i is provided by the age sampling. Indeed, the total landings estimator is

$$\hat{D} = \sum_j \hat{D}_j$$

and the estimator of age i proportion of the landings is:

$$\hat{pa}_i = \frac{\sum_j \hat{D_{ji}}}{\hat{D}} = \frac{\sum_j \hat{D}_j \hat{pa}_{ji}}{\sum_j \hat{D}_j}$$

These estimators are calculated using the estimator of the landings at length j,$\hat{D}_j$, and the estimator of the proportion of landings of age i in the length class j, $\hat{pa}_{ji}$, estimated respectively from the length sampling and the age sampling.

2.2.1 Stratified sampling by commercial category

In each commercial category, a sample of the landings is collected and each individuals is then attributed to its length group (Fig. 0).

Some additionnal notations for the stratified sampling by commercial category :

- k : k th commercial category
- K : the number of commercial category
- W_k : total landings of the k th commercial category in weight
- n_k : samples number of the k th commercial category
- v : the v th sample
- w_{kv} : the v th sample weight of the k th commercial category
- J : the number of length class
- I : the number of age group
- j : j th length class
- i : i th age group

- d_{jkv} : the number of fish belonging to the j th length class of sample v
- w_{jkv} : the weigth of fishes belonging to the j th length class of sample v
- M : the number of individual used to construct the age-length key
- m_j : the number of individual of length j of the age-length key
- pl_j : the proportion of individuals of length j of the age-length key
- q_{ji} : the proportion of individuals of length j and age i of the age-length key

The variance estimator of the total landings is the following.

$$Var(\hat{D}) = \sum_j Var(\hat{D}_j) + \sum_{j \neq j\prime} Cov(\hat{D}_j, \hat{D}_{j\prime})$$

We assume that $Cov(\hat{D}_j, \hat{D}_{j\prime}) = 0$, for all (j, j'), to simplify the calculation of the variance estimate.

Variance of landings at length

From the sampling design of the landings at length, the estimator of the landings at length j can be decomposed as follows,

$$\hat{D}_j = \sum_{k=1}^{K} \frac{W_k}{\sum_{v=1}^{n_k} w_{kv}} (\sum_{v=1}^{n_k} d_{jkv}) = \sum_k W_k \frac{\sum_v d_{jkv}}{\sum_v w_{kv}}$$

The estimator of the variance is

$$var(\hat{D}_j) = \sum_k W_k^2 var(\frac{\sum_v d_{jkv}}{\sum_v w_{jkv}})$$

and from Cochran (1977),

$$var(\frac{\sum_v d_{jkv}}{\sum_v w_{kv}}) = \frac{1 - \frac{\sum_v w_{kv}}{W_k}}{\frac{1}{n_k}(\sum_v w_{kv})^2} \frac{\sum_v (d_{jkv} - \frac{\sum_v d_{jkv}}{\sum_v w_{kv}} w_{kv})^2}{n_k - 1} \quad (2)$$

This last equation pointed out that the variability of sample sizes expressed in weight (w_{kv}) and in number(d_{jkv}) would penalized the variance of $\hat{D}_j$. To quantify this penality, we could introduce an average proportion at age. The variance estimator would be :

$$\hat{D}_j = \sum_{k=1}^{K} \frac{W_k}{\sum_{v=1}^{n_k} w_{kv}} (\sum_{v=1}^{n_k} d_{kv} \bar{p_{jk}}$$

An optimal sample size would be defined as the value of d_{kv} inducing a low variability in p_{jk}.

Variance of the proportion of landings at age i

From the sampling design of the age-length key, the estimator of the proportion of landings at age i is calculated by the following equation,

$$\hat{pa}_i = \sum_{j=1}^{J} q_{ji} pl_j$$

With an assumption of proportional allocation, the estimate of the variance of the proportion at age i (Kimura 1977, lai 1987)

$$var(\hat{pa}_i) = \sum_{j=1}^{J} (\frac{pl_j^2 q_{ji}(1-q_{ji})}{m_j} + \frac{pl_j(q_{ji}-pl_j)^2}{M})$$

2.2.2 Stratified sampling by mtier

In each métier, some vessels are randomly selected. For each vessel, landings are splitted into commercial categories which are all proportionnaly sampled (Fig. 0). It should be noticed that the weight by commercial category of each strata métier is either unknown or unreliable : only landings total weight is available. Consequently, number at lengthis estimated for each vessel, summed over all the sampled vessels of the métier and finaly rased at the métier strata.

Let us introduce some additionnal notations :

- l : l th métier
- L : the number of métier
- n_l : the métier l number of sample (i.e. the number of vessel sampled)
- w_{lv} : the v th sample weight of the l th métier
- $\hat{D}_{jl}$: the estimator of the j th length class landings of the l th métier
- d_{jlv} : the number of fish belonging to the j th length class of sample v
- w_{jlv} : the weigth of fishes belonging to the j th length class of sample v

For each vessel we need to define the commercial category level since a sample is collected by commercial category.

- $dech_{lvk}$: the number of individual in the sample collected into the k th commercial category of the v th sample Ech_{lv}
- $wech_{lvk}$: the sample weight collected within the k th commercial category of the v th sample Ech_{lv}
- d_{lvk} : the number of individual of the k th commercial category of the v th sample Ech_{lv}
- w_{lvk} : weight of the k th commercial category of the v th sample Ech_{lv}

Variance of landings at length The variance of the total landings is identical to the commercial category case :

$$Var(\hat{D}) = \sum_j Var(\hat{D}_j) + \sum_{j \neq j\prime} Cov(\hat{D}_j, \hat{D}_{j\prime})$$

The estimate of the landings of length j is expressed as

$$\hat{D}_j = \sum_{l=1}^{L} \hat{D}_{jl} \quad (3)$$

$$= \sum_{l=1}^{L} \sum_{v=1}^{n_l} w_l \frac{d_{jlv}}{w_{lv}} \quad (4)$$

$$= \sum_{l=1}^{L} w_l \sum_{v=1}^{n_l} \frac{\sum_{k=1}^{K} d_{jlvk}}{\sum_{k=1}^{K} w_{lvk}} \quad (5)$$

$$Var(\hat{D}_j) = \sum_l w_l^2 var(\sum_{v=1}^{n_l} \frac{\sum_{k=1}^{K} d_{jlvk}}{\sum_{k=1}^{K} w_{lvk}}) = \sum_l w_l^2 \sum_{k=1}^{K} var(\frac{\sum_{k=1}^{K} d_{jlvk}}{\sum_{k=1}^{K} w_{lvk}})$$

Using theorem 2.5 from Cochran (1977), we write :

$$var(\frac{\sum_{k=1}^{K} d_{jlvk}}{\sum_{k=1}^{K} w_{lvk}}) = \frac{1 - \frac{\sum_k w_{lvk}}{W_{lv}}}{\frac{1}{K}(\sum_k w_{lvk})^2} \frac{\sum_k (d_{jlvk} - \frac{\sum_k d_{jlvk}}{\sum_k w_{lvk}} w_{lvk})^2}{K-1}$$

Often d_{jlvk} is estimated by sampling the commercial category and an estimator of this number is given by:

$$\hat{d_{jlvk}} = w_{lkv} \frac{dech_{jlvk}}{wech_{lvk}}$$

Variance of the proportion of landings at age i

This variance is calculated using the same formula as for the stratified sampling by commercial category (Kimura 1977) :

$$var(\hat{pa_i}) = \sum_{j=1}^{J} (\frac{pl_j^2 q_{ji}(1-q_{ji})}{m_j} + \frac{pl_j(q_{ji} - pl_j)^2}{M})$$

2.3 Analytical tool for samples exploratory analysis

In a stratified sampling, each strata is supposed to split the population into homogeneous subpopulation regarding the estimated statistic. It is important to detect outlier samples and quantify their influence in the estimation of the statistic variance. We define an indice, called Δ, which quantified the discrepancy between the number at length in the sample and the adjusted mean number at length to the sample weight. This indice of discrepancy can be used 1/ to explore the samples of a strata (either commercial category or métier) regarding the length class (a)) or over all length classes (b)) and 2/ to qunatify the heterogeneity within a strata. With regards to these differents cases, the formula of Δ is the following :

1. given a strata k,

(a) given a length class j

$$\Delta_{jv} = d_{jkv} - \frac{\sum_v d_{jkv}}{\sum_v w_{kv}} w_{kv} \tag{6}$$

(b) over all length classes

$$\Delta_v = \sum_j d_{jkv} - \frac{\sum_v d_{jkv}}{\sum_v w_{kv}} w_{kv} \tag{7}$$

2. over all strata k

(a) given a length class j

$$\Delta_{jv} = d_{jkv} - \frac{\sum_{k,v} d_{jkv}}{\sum_{k,v} w_{kv}} w_{kv} \tag{8}$$

(b) over all length classes

$$\Delta_v = \sum_j d_{jkv} - \frac{\sum_{k,v} d_{jkv}}{\sum_{k,v} w_{kv}} w_{kv} \tag{9}$$

Note that Δ_{jv} matchs with the last part of equation 2.

2.4 Resampling method

Resampling techniques such as jackknife and bootstrap are often used for estimating confidence intervals or standard errors for any statistics. The principal advantages of these techniques are the easiness of implementation and the non reliance on normal therory. The fundamental assumption of bootstrapping developped by Efron (1979) is that the observed data are representative of the underlying population. In our case, bootstrapping has the advantage of proving the exactness of the analytical calculation and giving a CV value for a multiple stage subsampling. The multi stage subsampling can lead to a number of small samples, which would cause bias in the calculation of an estimator by the bootstrap technique. Chan and Lee (2001) recommend another algorithm for small sample bias reduction and base their work on less than 10 sample sizes. It is therefore recommended to have more than 10 samples in each stage to ascertain the convergence of the bootstrapped CV calculation.

The bootstrap method consists of drawing with replacement a number of new samples (usually 1000) from the observed data, each of the same size as the observed data. The statistic is calculated for each new set of data, yielding a bootstrap distribution for the statistic. It is possible to simulate stratifying sample by resampling independantly within each stratum. The final statistic is therefore a linear combination of each independant subsample statistics.

2.4.1 Bootstrap of one-stage sampling catch-at-length estimator

To compare with analytical result it can be of interest to bootstrap a simple one-stage catch-at-length estimator. One-stage sampling means that there is no stratification at all, neither time nor space or fleet stratification. All the samples are combined into one group like asked by the Data Directive.
The algorithm for such a bootstrap is as follows :

set-up :

- create a list of unique identifiers for sampling units (sample number)
- calculate values that will not change during the bootstrap process : total landing weight and number of samples

Bootstrap loop repeated at each iteration :

- create a table with all the information contained in the randomly selectioned samples
- calculate the length structure by summing all the samples with their respective raising factor
- calculate the catch-at-length by raising the length structure to the total catch
- append the estimates from this iteration to the output file

final calculation

- calculate variance, mean, CV, 5th and 95th percentile from the bootstrap distribution of each length class
- calculate a weighted CV for the length range that corresponds to 90% of the stock

2.4.2 Bootstrap of stratified sampling catch-at-length estimator

The software S+ allows the setting of one grouping variable. It is therefore possible to simulate a stratified sampling by creating a variable that contains the complete label of the lower level strata. For example, stratification by quarter (Q), harbours (H) and fleet (F) would generate a variable where modalities would be ”Q1 H1 F1”, ”Q1 H1 F2”, ”Q1 H2 F1”, etc...

The algorithm for such a bootstrap is as follows :

set-up :

- create a list of unique identifiers for sampling units (sample number + stratification label variable)
- calculate values that will not change during the bootstrap process : total landing weights and number of samples by strata

Bootstrap loop repeated at each iteration :

- create a table with all the information contained in the randomly selectioned samples
- calculate the length structure by summing all the samples with their respective raising factor
- calculate the catch-at-length by raising the length structure to the total catch
- append the estimates from this iteration to the output file

final calculation

- calculate variance, mean, CV, 5th and 95th percentile from the bootstrap distribution of each length class
- calculate a weighted CV for the length range that corresponds to 90% of the stock

2.4.3 Bootstrap catch-at-age estimator

Bootstrap catch-at-age precision estimator can be obtained by two methods :
1 - using the bootstrap variance of the catch-at-length combined with the Kimura variance of the ALK with the formulas given in chapter 2.1
2 - bootstrapping the Age-Length-Key under the assumption that a sample is one otolith with its fixed parameters (quarter, zone, length, estimation of age, etc...) and the grouping variable is the length, i.e. the number of otoliths read by length class is constant during the bootstrap process.

The algorithm for such a bootstrap is as follows :

set-up :

- create a list of unique identifiers for length sampling units (sample number + stratification label variable)
- create a list of unique identifiers for age sampling units
- calculate values that will not change during the bootstrap process : total landing weights and number of length samples and number of otoliths by strata

Bootstrap loop repeated at each iteration :

- create a table with all the information contained in the randomly selectioned length samples
- create an Age-Length-Key with the otoliths randomly selectioned by length classes
- calculate the length structure by summing all the samples with their respective raising factor
- calculate the catch-at-length by raising the length structure to the total catch
- combine the catch-at-length with the ALK to obtain catch-at-age
- append the estimates from this iteration to the output file

final calculation

- calculate variance, mean, CV, 5th and 95th percentile from the bootstrap distribution of each age class
- calculate a weighted CV for the age range that corresponds to 90% of the stock

2.5 Simulation

For reasons of computer time consuming, only the analytical simple random stratification is used to implement the simulation algorithm. The simulation is therefore usable at the strata disaggregated level or simulates a non stratified scheme if used at the final aggregated level. The following algorithm is based on the resampling technique assumption, i.e. the observed data are representative of the underlying population which allows bootstrap like multiplication or division of the sample numbers.

set-up :

- create a list of unique identifiers for length sampling units
- create a list of unique identifiers for age sampling units

Simulation double loop

- First loop with a vector of length number of samples multipliers $nlmult$ (from n/10 to 3n)
- Second loop with a vector of age number of individuals sampled multipliers $namult$ (from n/10 to 3n)
- if $nlmult <= n$ Selection nlmult number of samples without replacement among the n length samples
- if $nlmult > n$ Selection nlmult number of samples with replacement among the n length samples
- if $namult <= n$ Selection namult individuals without replacement in the Age-Length-Kry
- if $namult > n$ Selection namult individuals with replacement in the Age-Length-Key
- combine the catch-at-length raised from nlmult length samples with the namult individuals ALK to obtain catch-at-age
- Calculate the precision
- append the estimates from this iteration to the output file

final graph

- draw a contour plot of the double loop precision matrix

3 Materials

Market sampling in France is done either by commercial categories, either by fleet or métier. Usually, sampling by commercial categories needs stability in the fish sorting process as a prerequisite. To represent both methodologies, Atlantic hake and Eastern Channel sole sampling scheme will be described and analysed.

3.1 Atlantic hake

On the French atlantic coast a large number of fleets lands hake using different fishing means : gillnets, trawls and lines. ICES working groups have defined Fishery Units (FU) to coordinate international sampling process. French sampling is based on these fishery units and landings are sampled within 5 of them : FU05 (inshore fish trawler in ICES area VII), FU09 (nephrops trawlers in ICES area VIII) , FU10 (trawlers in ICES area VIII), FU12 (longliners in ICES area VIII) and FU13 (gillnetters in ICES area VIII). Sampling scheme is distributed among 6 harbours from south of Biscaye to south Britanny.
The length sampling objective is based on a number of fish to sample per FU and per quarter except for FU09 and FU10 based on a number of trips per month to sample.

FU05	2000 fish
FU09	10 trips per month
FU10	10 trips per month
FU12	1000 fish
FU13	500 fish

Age sampling scheme combines different methodologies

- fish purchase : 10 fishes per length classes per quarter
- Direct otoliths removal from market length sampling : 5 fishes per length classes per quarter with special attention to the largest fish
- Supplementary otoliths are provided by surveys (RESSGASC, EVHOE)

3.2 Eastern Channel sole

Sole landings are shared mostly between trawlers and gillnetters. Landings occur out of scallop season, i.e. from march to november. Regional landing distribution is about 30% for harbours between Cherbourg and Fecamp and 70% for harbours between Dieppe and Dunkerque.
The length sampling objective is based on commercial categories distributed among the principal harbours and quarter. At each sampling day, at least 3 samples from each commercial categories are sampled. One sample consists of measuring around 50 individuals , that is to say boxes with large number of small fish are splitted in two or three equal parts to avoid too much differences in the within category number of individuals sampled. On the other hand, boxes with very few number of individuals are skipped.
The age sampling is based on quarterly fish purchase. Once a quarter (usually in the middle), a fixed weight of each commercial categories is bought in order to have all the length classes range.

4 Results

4.1 exploratory analysis of the samples

The precision level of a multistage sampling estimator depends on the adequation between sampling effort and within strata variance. The first analysis is therefore to investigate on the internal variabilities within strata. To do so, the formula developped in chapter 2.2 is very informative,

more precisely the last part called distance to the mean distribution and calculated for each sample. This statistic is also the principal component of the variance calculation and very high values points out possible outliers or sample that takes the larger part of the mean distribution information. This statistic applied to Eastern Channel sole (Fig. 1) shows the importance of stratifying by commercial categories which are well discriminated one from each other. This figure shows also that special sampling effort has to be made on small fish category as it has the most variability. Focus on only one category subsampled by harbour or quarter (fig. 2) shows the same range of variability and the same symmetry around 0 for each strata. This means that these strata require the same sampling effort and stratification can not be designed in the purpose of reducing the total variance but rather for having sufficient information at a disaggregated level.
The same statistic applied to Atlantic hake (fig. 3) shows very weak differences in the range of sample distances to the mean distribution within strata. This means that sampling effort has to be proportional to the level of landings as internal variabilities in the different strata are in the same magnitude.

To continue with this exploratory analysis, precision can be improved by optimizing the relative importance of sampling effort against relative importance of landings per commercial category (Fig. 4a), quarter (Fig. 4b and 5b) and Fishery unit (Fig. 5a). This can only be a post-analysis as it is difficult to know these informations before going sampling, but it can detect some discrepancies to correct for the next years. For example, we can see that a special effort has to be made in the second quarter for Easter Channel Sole (Fig. 4b) and it would be useful to sample a little more Fishery unit 13 for Atlantic hake maybe instead of sampling so much Fishery unit 12 ((Fig 5a). These are the kind of issues to be discussed in local workshops to optimize the precision at a given sampling effort.

4.2 catch-at-length precision

Precision on length distribution for Easter Channel sole (Fig. 6) and Atlantic hake (Fig.7) put in evidence two main issues. $Primo$, at a given sample number (Eastern Channel sole and Atlantic hake have respectively 6365 and 23086 individuals sampled) stratifying by commercial category will largely improve the precision. $Secundo$, analytical and bootstrap methods give exactly the same picture, i.e. good CVs on well represented length classes in the length distribution and poor precision on scarce length classes. Moreover, CV estimations are very close with both methodologies thus validating each other result. The overall CV is the weighted mean on length range representing 90% of the stock and calculated on an annual basis as required by the Data Directive.

4.3 Age length keys precision

Precision in ALKs (Fig. 8 for Eastern Channel Sole and Fig. 9 for Atlantic hake) shows the same pattern as precision in length distribution, i.e. poor precision in scarce ages and good precision in well represented ages. The CV range is narrower because special effort is made to get information from all the length classes with a fixed or proportional allocation. Eastern Channel sole and Atlantic hake have respectively 1102 and 1420 otoliths read and overall CV is the weighted mean on age range representing 90% of the stock. Quartely CVs for both species are found to be around 10% - 12% but represent only the sampling precision not the age reading errors which are meant to be very important, especially for hake.

4.4 catch-at-age precision

The age distribution, which is a combination between the length distribution and the ALK, is very sensitive to the precision of the latter. The analytical 3 terms variance formula (chapter 2.1) enables to discriminate the relative contributions of the precision in the length structure and the precision in the ALK to the overall precision (Table 1). For Eastern Channel sole in 2002, the weighted mean relative contributions are respectively 26.1%, 73.8% and 0.1% for the terms associated to catch-at-length variance, ALK internal variance and the product of both. The preponderance of the age information over the length information has already been underlined in the EMAS project (Anon, 2001) and is important to know for optimizing the sampling scheme (Fig. 11). The definition of a cost function combined to this analytical approach enables to quantify the precision and cost of different arrangements of stratification in the sampling (quarter/harbour/CommercialCategory, quarter/port, quarter/CommercialCategory, harbour/CommercialCategory, . . .).

5 Results

5.1 exploratory analysis of the samples

The precision level of a stratified sampling estimator depends on the adequation between sampling effort and within strata variance. Furthemore, a good estimation of variance can be performed if sampling effort devoted to a strata is proportional to the relative part of the sampled population within this strata. An explanatory analysis of the collected samples is necessary before any precision estimation to detect outliers and to start an analysis of the sampling design adequacy. The first analysis is therefore to investigate on the internal variabilities within strata. To do so, the indice, denoted Δ, developped in section 2.3 is very informative for each sample. This statistic is also the principal component of the variance calculation and very high values points out possible outliers or sample that takes the larger part of the mean distribution information.

- **Heterogeneity over all strata and all length class:** Δ_v This statistic applied to Eastern Channel sole (Fig. 2) shows the importance of stratifying by commercial categories which are well discriminated one from each other. The same statistic applied to Atlantic hake (fig. 3) shows very weak differences in the range of sample distances to the mean distribution within strata. This means that sampling effort has to be proportional to the level of landings as internal variabilities in the different strata are in the same magnitude.

- **Heterogoneity within each strata :**$\Delta_k v$ Focus on only one category subsampled by harbour or quarter (fig. 4) shows the same range of variability and the same symmetry around 0 for each strata except for the right part of the graph which represents the smallest fish category. This means that special sampling effort has to be made on small fish category as it has the most variability. The other strata require the same sampling effort and stratification can not be designed in the purpose of reducing the total variance but rather for having sufficient information at a disaggregated level.

Assumption of sample representativity of the underlying population can be distorted by samples with too few individuals measured. On the other hand, there is an asymptotic upper limit to the number of individuals to measure in one sample. Limit where continuing measuring fish does not bring more information. This limit depends on the number of length classes and is different from one species to the other. Figure 5 shows the heterogeneity in the sample number and sample weight within each strata. This graphs enables to point out possible discrepancies in the database.

To continue with this exploratory analysis, precision can be improved by optimizing the relative importance of sampling effort against relative importance of landings per commercial category (Fig. 6a), quarter (Fig. 6b and 6b) and Fishery unit (Fig. 6a). This can only be a post-analysis as it is difficult to know these informations before going sampling, but it can detect some discrepancies to correct for the next years. For example, we can see that a special effort has to be made in the second quarter for Eastern Channel Sole (Fig. 6b) and it would be useful to sample a little more Fishery unit 13 for Atlantic hake maybe instead of sampling so much Fishery unit 12 (Fig 7a). These are the kind of issues to be discussed in local workshops to optimize the precision at a given sampling effort.

5.2 catch-at-length precision

Precision on length distribution for Easter Channel sole (Fig. 8) and Atlantic hake (Fig.9) put in evidence two main issues. $Primo$, at a given sample number (Eastern Channel sole and Atlantic hake have respectively 6365 and 23086 individuals sampled) stratifying by commercial category will largely improve the precision. $Secundo$, analytical and bootstrap methods give exactly the same picture, i.e. good CVs on well represented length classes in the length distribution and poor precision on scarce length classes. Moreover, CV estimations are very close with both methodologies thus validating each other result. The overall CV is the weighted mean on length range representing 90% of the stock and calculated on an annual basis as required by the Data Directive. Figure 8 and 9 represents the sampling as requested by the Data Directive, i.e. annual CV without stratification. The important question to raise is to know if stratifying would give a better precision with the same effort. The stratified bootstrap allows to calculate CV with any sets of strata and it can be very informative to try each arrangement. Here, we have only calculated a stratification by quarter and commercial category and sets of harbours and commercial categories for Eastern Channel sole (fig. 10). It is evident that there is no need to split into sets of harbours (fig. 10b) but stratifying by quarter (fig. 10a) increases the precision.

5.3 Age length keys precision

Precision in ALKs (Fig. 11 for Eastern Channel Sole and Fig. 12 for Atlantic hake) shows the same pattern as precision in length distribution, i.e. poor precision in scarce ages and good precision in well represented ages. The CV range is narrower because special effort is made to get information from all the length classes with a fixed or proportional allocation. Eastern Channel sole and Atlantic hake have respectively 1102 and 1420 otoliths read and overall CV is the weighted mean on age range representing 90% of the stock. Quartely CVs for both species are found to be around 10% - 12% but represent only the sampling precision not the age reading errors which are meant to be very important, especially for hake. Annual CVs obtained by adding the quarterly ALK matrix gave respectively 5% and 3.2% for sole and for hake, which represents a very high increase of precision. ALK is the example where concatening quarterly information will lead to bias estimates as for most of the fisheries recruitment occurs during the year modifying at a great extent quarterly ALK information.

5.4 catch-at-age precision

Figure 13 shows the slight difference between analytical and bootstrap estimated precision. The reason is that this calculation has been done with the bootstrapped catch-at-length combined to the analytical estimation of variance of the annual ALK. The difference is only due to the difference between analytical and bootstrapped precision estimation of the catch-at-length. To complete this work, it remains to estimate a bootstrapped CV on the quarterly ALKs and estimate the real quarterly stratified catch-at-age used for the assessment. The age distribution, which is a combination between the length distribution and the ALK, is very sensitive to the precision of the latter. The analytical 3 terms variance formula (chapter 2.1) enables to discriminate the relative contributions of the precision in the length structure and the precision in the ALK to the overall precision (Table 1). For Eastern Channel sole in 2002, the weighted mean relative contributions are respectively 26.1%, 73.8% and 0.1% for the terms associated to catch-at-length variance, ALK internal variance and the product of both. The preponderance of the age information over the length information has already been underlined in the EMAS project (Anon, 2001) and is important to quantify for optimizing the sampling scheme (Fig. 14). The definition of a cost function combined to this analytical approach would enable to quantify the precision and cost of different arrangements of stratification in the sampling (quarter/harbour/CommercialCategory, quarter/port, quarter/CommercialCategory, harbour/CommercialCategory, . . .).

6 discussion

It is trivial to say that precision level in biological sampling for fisheries data depends on the sampling design and on the different variables that are either measured or estimated. The main purpose of this paper is to propose two complementary methods to estimate the precision in the biological sampling and build a logical reasonning for optimizing the sampling design.
The most usual demand comes from ICES Stock Assessment Working Groups that need an age structure of the landings for some species. The age composition of the landings is one of the most important parameters used in stock assessment modeling. Apart from gear and catchability parameters, the model assumption are that the age structure of the landings is first representative of the age structure of catches and is proportional to the age structure of the stock. It is therefore important, not only to quantify the precision of the age structure estimation but also to avoid source of bias. Among sources of bias, the most important lead to mis-estimation of discards and not adequate sampling design because of a bad spatial, temporal and selectivity coverage. In the future, the demand will come to the providing of fleet-based disaggregated age structure of the landings. The need of precision at disaggregated level is also an important issue to address.

In this document, we focus on the precision of the age structure estimation regarding different sampling designs and on analysis of the adequacy of the sampling design to reduce bias of this estimation.

The first point investigated is that sampling design splits correctly the landings heterogeneity. To ensure that sampling is representative of the national landings, the sampling design must cover all the fleets during all the fishing period with a sampling effort proportional to the landings. This is a basic rule-of-thumb to avoid bias risk when within and between strata variances are of the same magnitude. To avoid bias, it is important to see if landings age structure is linked to an external factor like quarter, geographic area, harbour, fleet, etc. . . . The indice Δ is a candidate statistic for this kind of exploration. In the case of commercial category-based sampling, we have seen that stratifying by quarter increased the precision but stratifying by sets of harbours reduced it. Such investigations are very time consumming and tools to analyse this in a comfortable way would be much appreciated. This kind of investigation has not been carried out yet on Atlantic hake but first results on Eastern Channel sole shows the importance to address such issues. What is the interest of stratifying by metier, by harbours ? The Δ statistic associated with quantification of different strata arrangements would give elements to answer. Some questions remain open like the need or not to have market commercial categories stable sorting to use this information as a raising factor. This paper has shown the efficiency of such a sampling design to gain precision. It could be sufficient to have a relative qualification of the fish length like small, medium, big usable as strata and raising factor to have a non-biased precise and cheap estimation of catch-at-length.

A second important point is the problem of number of individuals to measure in a sample. A balance has to be found between the number of samples and the number of individuals in each sample for optimizing the fixed sampling effort.

To settle the important issue of opimizing sampling scheme, we have on one hand a fixed sampling effort and on the other hand information from previous years sampling acting like a pilot study. The first step of the reasonning is the exploratory analysis where it is possible to detect possible outliers usually very influent in the final raised numbers. Distance to mean distribution has been introduced to quantify unit sample influence on the final raised number. When ordered by an external variable, this statistic also enables to detect possible stratification for the purpose of reducing the variance or inversely combining strata that show the same sampling pattern.
As the final age structure is the combination of two estimations in a double sampling procedure, the second step of the reasonning should be the search of a balance in the sampling effort between them. The problem comes from the fact that the easily measured variable have a low contribution to the final estimation unlike the costly hard-to-measure variable. EMAS (1991) considered the possibility of picking otoliths at random from catches without regard to size of the fish in a multi-stage sampling. It is obvious that a large effort is brought to collect catch-at-length data at every possible disaggregated level at the expense of the effort on otolith reading.

Concerning the CV calculation, the analytical method is difficult to write precisely because it depends on the sampling design and depends on which variables are measured and which variables are estimated. Once analytical writing done, it is easy to implement and allows exploratory analysis (Δ statistic), decomposition of variance and simulation as it is not computer time consumming. The analytical writing becomes very fastidious at a stratify sampling level. The bootstrap method is easy to implement, allows to stratify and test different strata arrangements but is computer time consumming. Both methods are thus complementaries and gives a complete analysis of a sampling design.
Like a few previous papers, we have shown that precision in the age structure was mostly driven by the precision in the age-length-key (ALK). This particularity raises a few comments : (i) the maximum of attention must be given to the collection of ALK, in particular the one-time-a-quarter fish purchase is probably not a random sample from the total population of fish landed in the quarter by all fleets from the all geographical area. (ii) reading errors are not mentionned at this stage and represent a non negligible component of the ALK precision, hake beeing probably the most concerned species for this problem. The improvement in the sampling methodology to increase precision and avoir bias must be accompanied by an improvement in data quality. Mis-reporting and/or under-reporting errors affect the raising procedure at an unknown extent. Like otolith reading errors this problem is probably non-negligible to the contribution of overall precision in the estimation of catch-at-length and catch-at-length.

At last, at the stock scale, the combination of different national fleet-based disaggregated data require a specific international coordination work for defining the same fleet definition and the same sampling methodology.

Aknowledgement The authors would like to thank Dr Michel Bertignac (Ifremer Lorient, France) for having placed Atlantic hake sampling database to their disposal.

Analytical variance at age decomposition (Absolute and relative)

	Absolute value (millions)			Relative value (%)			Total var	Total	
Age	V1	V2	V3	V1	V2	V3	(millions)	Number	CV
1	6.7	589.9	0.7	0.011	0.988	0.001	597	81054	0.302
2	4196.5	9003.1	10.8	0.318	0.682	0.001	13210	3207625	0.036
3	11149.6	16788.8	20.0	0.399	0.600	0.001	27958	4233195	0.039
4	1510.7	8725.3	10.4	0.147	0.852	0.001	10246	788669	0.128
5	5031.7	13972.2	16.7	0.265	0.735	0.001	19021	1201299	0.115
6	241.7	3743.5	4.5	0.061	0.938	0.001	3990	152493	0.414
7	65.6	2046.2	2.4	0.031	0.968	0.001	2114	88970	0.517
8	37.9	1553.7	1.9	0.024	0.975	0.001	1593	35870	1.113
9	8.5	737.0	0.9	0.011	0.987	0.001	746	22124	1.235
10	17.7	1008.1	1.2	0.017	0.982	0.001	1027	23146	1.385
11	34.0	1302.8	1.6	0.025	0.973	0.001	1338	40781	0.897
12	55.5	1633.2	2.0	0.033	0.966	0.001	1691	34470	1.193
13	2.6	403.2	0.5	0.006	0.992	0.001	406	5651	3.567
14	1.7	261.4	0.3	0.006	0.992	0.001	263	6075	2.671
15	60.4	1670.7	2.0	0.035	0.964	0.001	1733	23093	1.803

Table 1: *Sole VIID 2002 - Analytical variance decomposition*

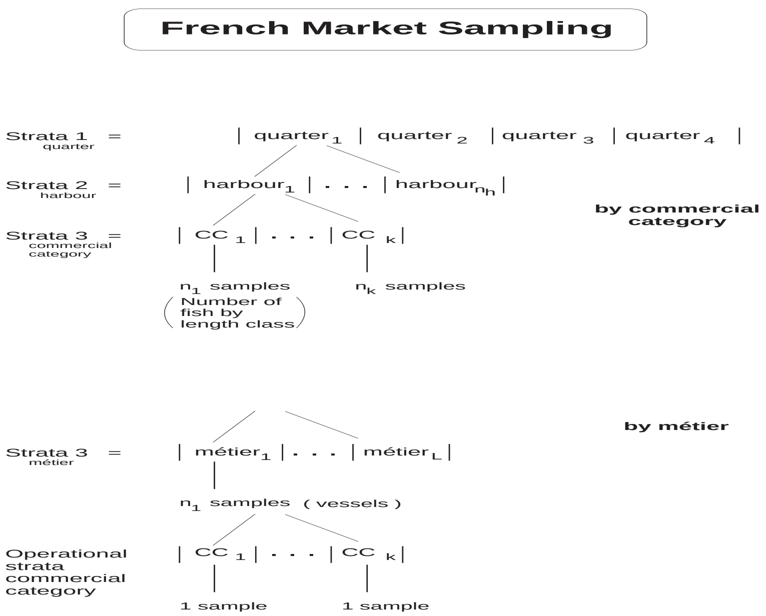


Figure 1: *French market sampling*

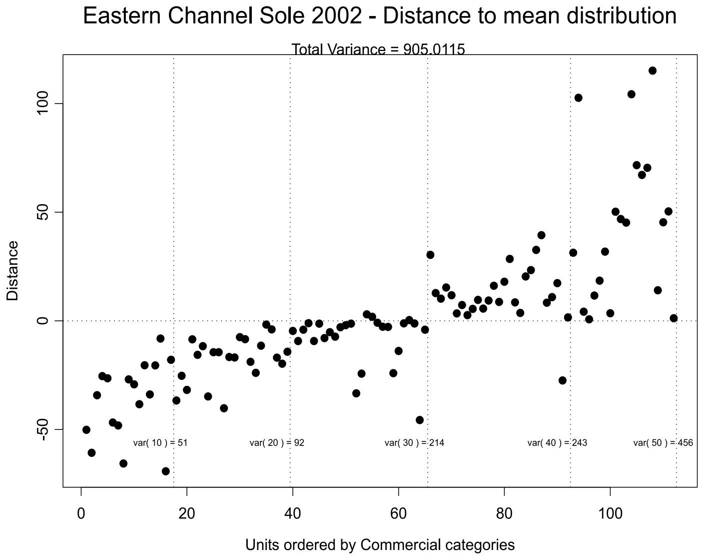


Figure 2: *Sole VIID 2002 - Distance to mean distribution. Samples ordered by commercial category*

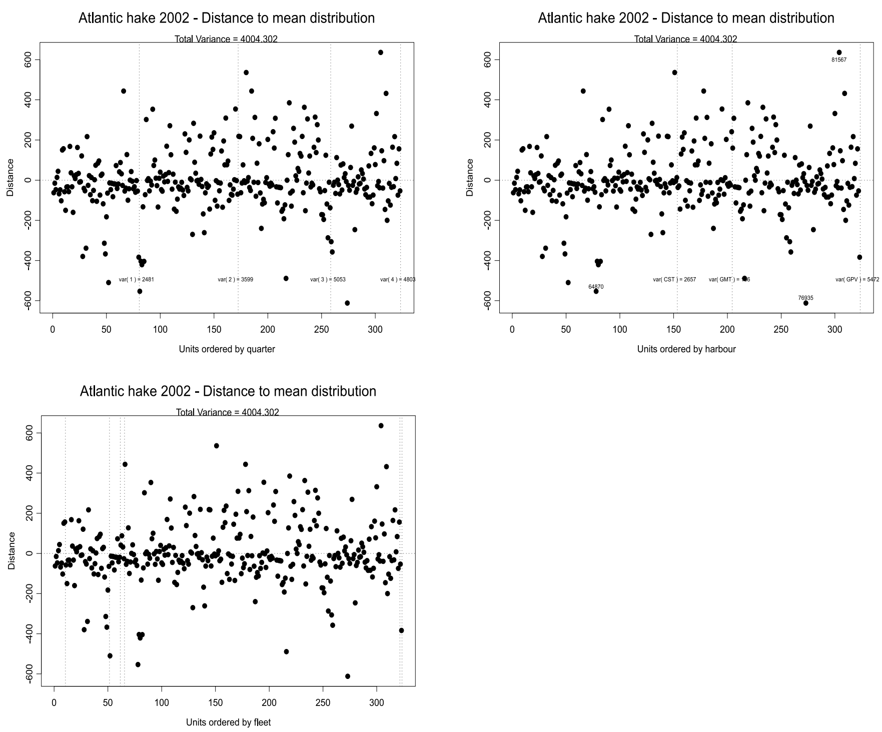


Figure 3: *Atlantic hake 2002 - Distance to mean distribution*

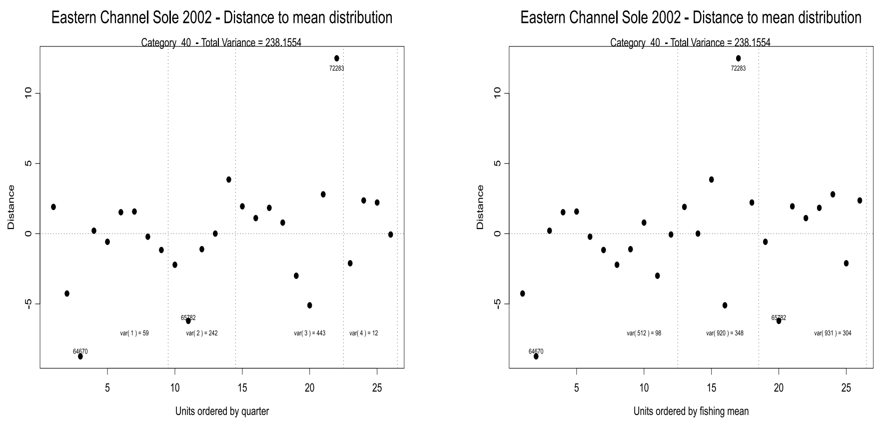


Figure 4: *Sole VIID 2002 - Distance to mean distribution for one commercial category. Samples ordered by a) quarter b) fishing mean*

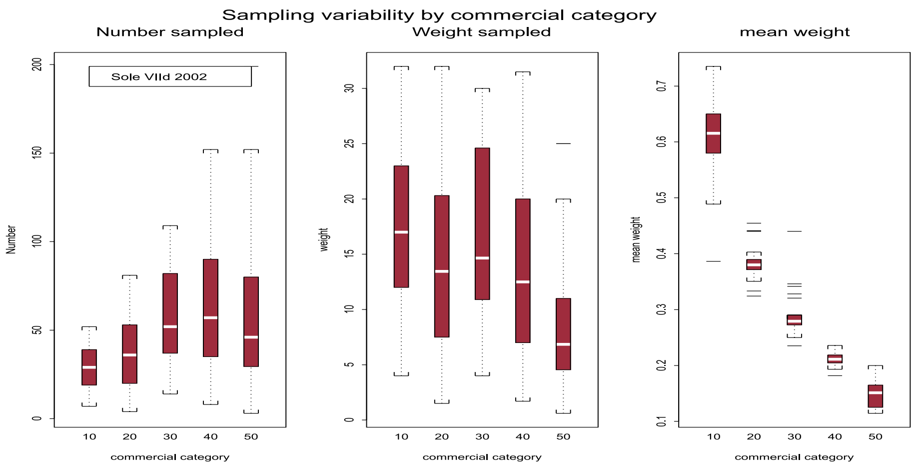


Figure 5: *Sole VIID 2002 - Sampling variability a) in number of fish measured by sample, b) sample weight and c) mean weight of a fish by sample*

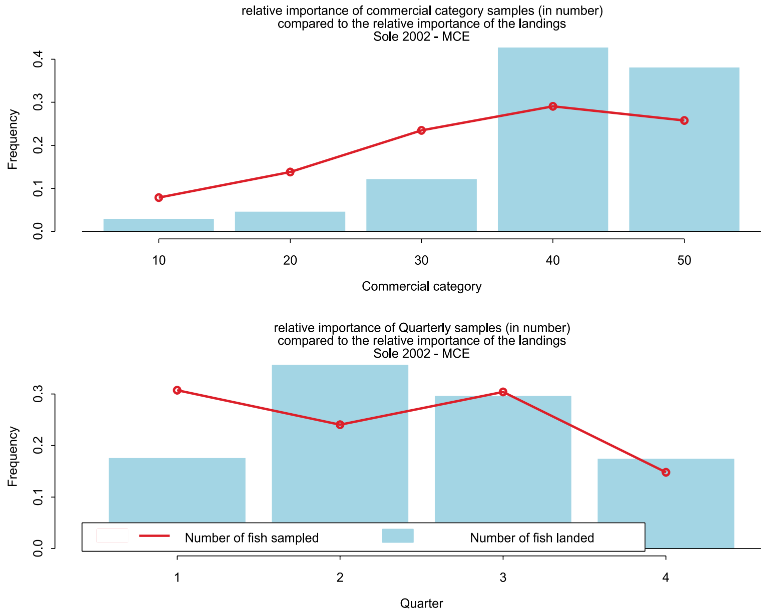


Figure 6: *Sole VIID 2002 - Relative importance of sampling effort against relative importance of landings by a) commercial category b)quarter*

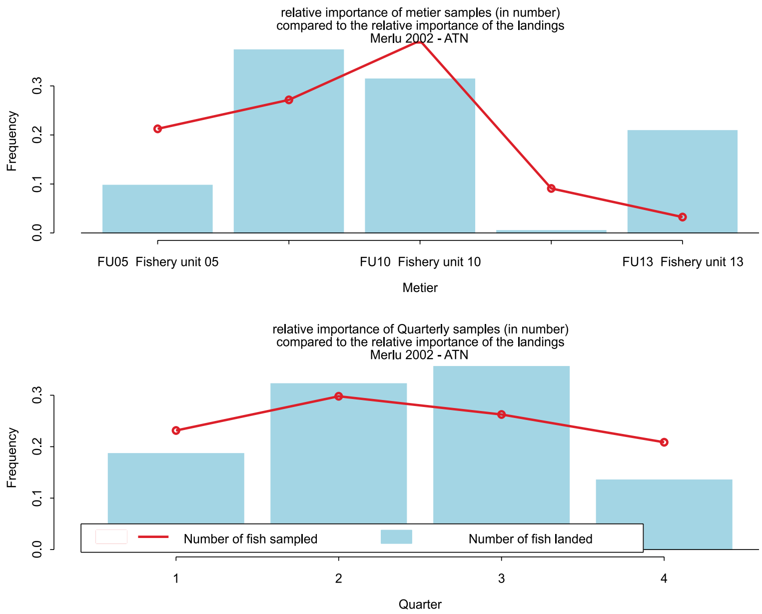


Figure 7: *Atlantic hake 2002 - Relative importance of sampling effort against relative importance of landings by a) fishery unit b) quarter*

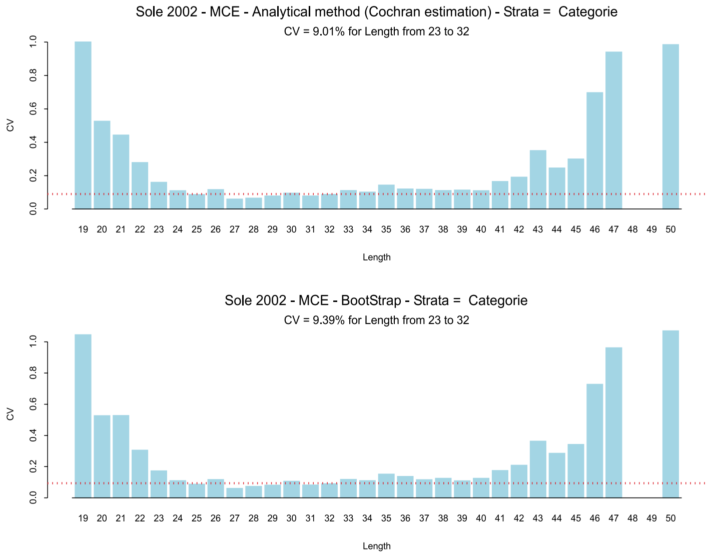


Figure 8: *Sole VIId 2002 - Precision in length distribution a) analytical b) bootstrap*

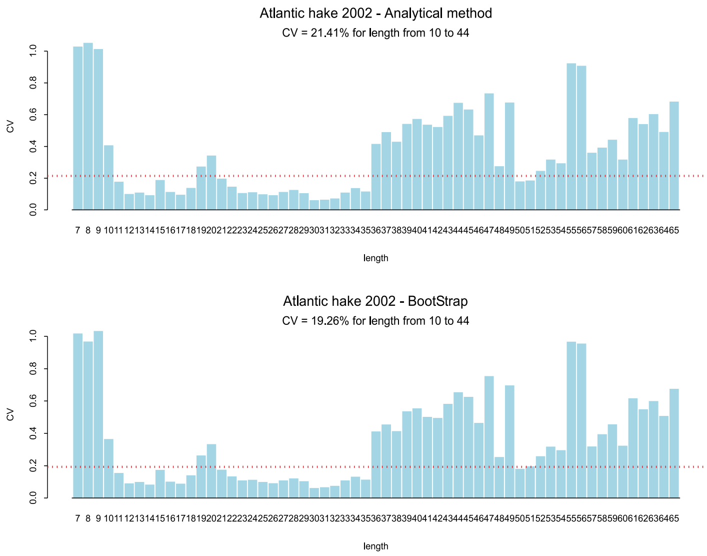


Figure 9: *Atlantic hake 2002 - Precision in length distribution a) analytical b) bootstrap*

Figure 10: *Sole VIID 2002 - Distance to mean distribution for one commercial category. Samples ordered by a) quarter b) fishing mean*

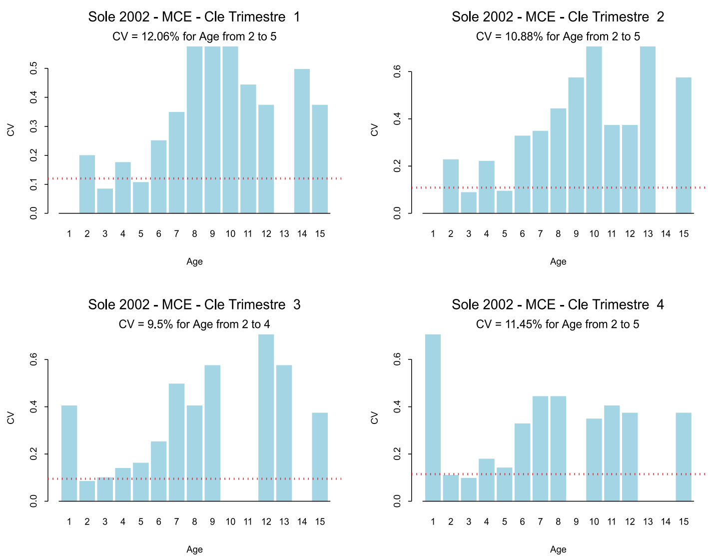


Figure 11: *Sole VIId 2002 - Precision in quarterly age length keys*

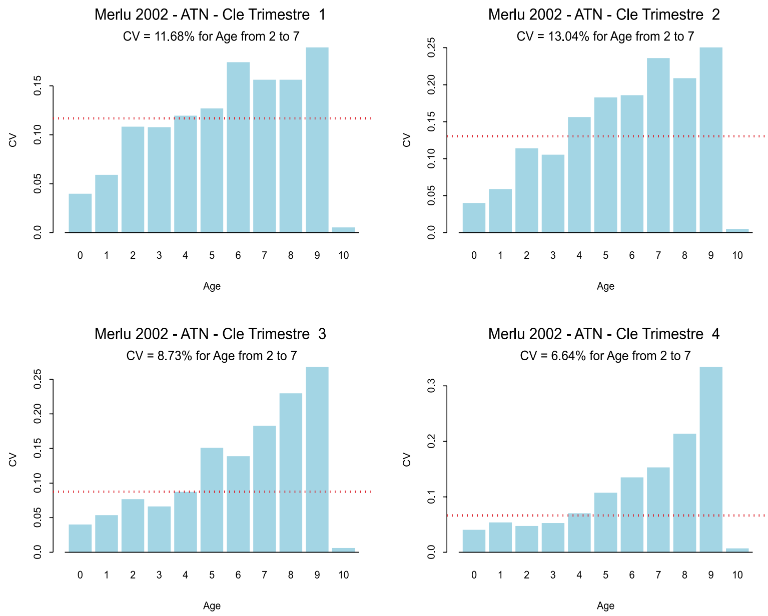


Figure 12: *Atlantic hake 2002 - CV at age in quarterly age length keys*

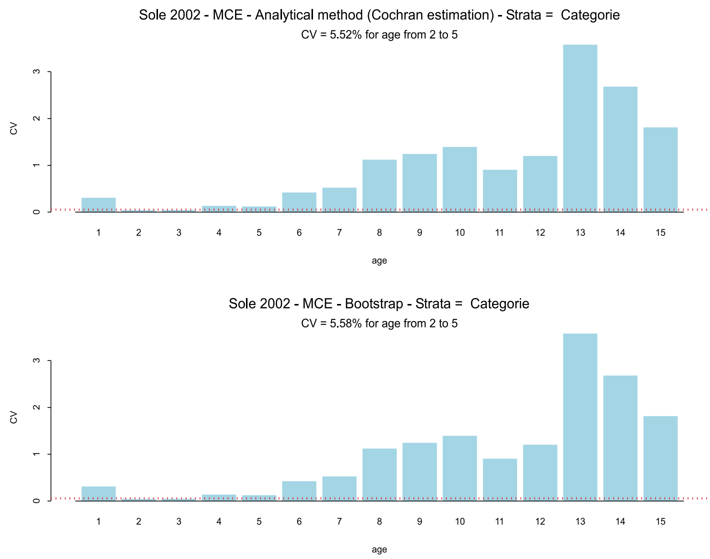


Figure 13: *Sole VIId 2002 - Precision in age distribution*

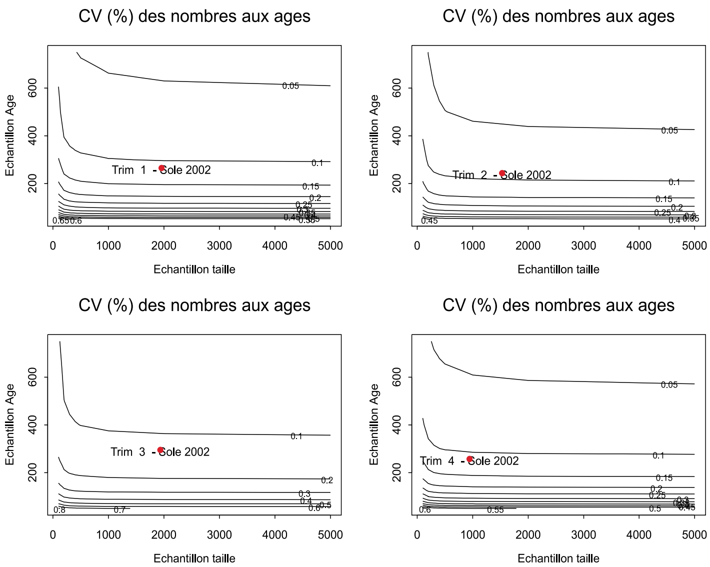


Figure 14: *Sole VIId 2002* - *simulation of precision from different sampling effort*

Precision in catch at age data with regard to sampling design

Joël Vigneau and Stéphanie Mahévas

Abstract

This working document is a contribution to the Term of Reference d) of the WKSCMFD : propose methods to estimate precision and design sampling stratification schemes that will minimise bias and maximise precision. Large sampling effort is needed all year long to provide working groups on stock assessment with essential data, the catch-at-length or catch-at-age data for each species combined at the International level. It is therefore essential to estimate the precision associated with these data with regards to the disagregated levels requested by the working groups. Each country is responsible for their own sampling which makes every national sampling design nearly unique. This document presents the analytical statistics and resampling techniques developed to analyse sampling shemes. Previous works have been largely used to develop this approach. A simulation algorithm is proposed to find graphically the optimum sampling effort for a target precision level. This paper also focus on exploratory analysis of sampling design and one statistic is proposed to quantify heterogeneities within and between strata. At the sample level, this statistic enables to point out possible outliers. At the strata level, the heterogeneities can be used to define *ad hoc* stratification by combining strata showing the same pattern or by adding another strata if patches are visible. Since the French sampling covers different fisheries, from the Mediterranean to the North Sea, different sampling designs are required to take into account the large disparities, mainly a market commercial category-based sampling and a fleet-based sampling. This analysis is applied to the two main french sampling designs using the 2002 Eastern Channel sole and Atlantic hake sampling database.

1 Introduction

The length and age distribution are used as input data to assess the state of the most important fish stocks all over the world. Inherent to any sampling procedure, the estimation of the length and age distribution may contain bias and uncertainties. The dissemination of input data errors or uncertainties in the assessment models have been studied by Kimura (1989), Restrepo and Powers (1990), Pelletier (1991) and more recently by Reeves (2003) and are given to be significant on stock evaluations. In the European fisheries sector, the sampling of the length and age distribution of species in the landings has always been under the Member States responsibility. As the national species/stock length and age distribution are combined to obtain the international distribution, there has been some regional coordination in the latest years (SAMFISH, FIEFA, North Sea sampling program) to avoid too much heterogeneity in the sampling process. To expand this coordination to every Member States, EU Regulation 1639/2001 hereafter named Data Directive has given a framework for the collection of data in the European fisheries sector since 2002. This Data Directive does not require a precision level for the numbers-at-length and numbers-at-age but the PGCCDBS (ICES, 2003) assumed that this information would be helpful to provide objective means for comparing national programmes, assist people involved in sampling and provide quality information on the input data for assessment models. Moreover, the Data Directive gives common rules for dealing with precision levels and sampling intensities. There are three case studies :

- when it is not possible to define quantitative targets for sampling programmes, neither in terms of precision levels, nor in terms of sample size, pilot surveys in the statistical sense will be established [. . .].
- When quantitative targets can be defined, they can be specified either directly by sample sizes or sampling rates, or by the definition of the levels of precision and of confidence to be achieved.
- When reference is made to a sample size or to a sampling rate in a population defined in statistical terms, the sampling strategies must be at least as efficient as simple random sampling. [. . .].

With respect to these common rules and to achieve PGCCDBS goal, the attention will be given on precision levels in number-at-length and number-at-age as a target and as a tool for sampling strategy analysis.
The length composition of the landings is assumed to be representative of the length composition in the stock to be used in the assessment models. Sampling the landings to estimate the overall species length structure introduces the notion of bias and precision with respect to sampling strategy and sampling effort. The first step of the study will be the research of heterogeneities within the length distribution related to factors like commercial market categories, time, fleet, harbours or areas. The heterogeneities evidences will permit to define stratas in the sampling procedure and try to minimize the variance of the numbers-at-length estimation by adopting a stratified sample design (Cochran, 1977). At a second stage, simple random sampling CV calculations will be proposed with respect to the Data Directive using analytical and resampling statistics. Only the resampling will be presented to estimate the precision of the stratified sampling CV's.

The age distribution sample is designed to be representative of the age distribution at length of the stock. The calculation of precision levels in numbers-at-age based on age and length sampling have been treated in numerous reports and the latest approach by Kimura (1977) was the starting point of new investigations. Kimura (1977) derivated the variance of an unbiased proportion at age estimator depending on a fixed, random and mixed allocation. He also introduced the VarTot function as an error index in the age length key (ALK). This function has been used to calculate coefficient of variation in the age distribution of NAFO cod (Baird, 1983) and to design an optimal sampling strategy for cod (Gavaris and Gavaris, 1983). Finally, Kimura (1989) proposed an analytical formula of the variance of the catch at age estimator, derived from the variance of the estimate of catch at weight and the variance of the age-weight relationship using the delta Method. This author also quantified the impact of this variability on the estimates produced by the cohort analysis for stock evaluation and discussed the choice of the allocation for age sampling, prefering a proportional allocation and showing the low impact the sampling level would have on the variance for older classes. Lai (1987) introduced a cost function to calculate the optimal sampling effort between length sample and age sample. This work was completed by Quinn et Deriso (1999) with the description of three optimisation methods in a stratified sampling.

Two sampling designs are considered in France to estimate landings length distribution (fig. 1). Both are first stratified by quarter, by harbours set and differed by their third strata which are either the commercial category or the métier, also called "captain sorting". For the stratified sampling design per métier, the selected units in the last strata are subsampled by commercial category. Samples are collected with a proportionnal allocation in each strata. For ageing, two approaches are practised whether the otholiths can be extracted without damaging the fish or not. In the former situation, a subsample of the length sample is taken, in the latter an independant age sample is collected.

One method will be proposed here corresponding to length sampling where sample weight is known and quarterly ALK with a fixed or proportional allocation. In accordance with the Data Directive, comparison will be made between simple random and stratified sampling to assess which method gives the best result in term of precision with a given cost. To perform this analysis, we will consider a single year age-length key agregating the traditional quater age-length keys. Examples of stock where sampling is stratified by fleet (Atlantic hake) and stocks where sampling is stratified by commercial categories (Eastern Channel sole) will be considered and different levels of sampling numbers will be given according to different levels of precision goal to be achieved.

2 Method

In this section, we first present 1/ the analytical formulation of the catches at age variance with regard to each of these sampling designs without the quater and harbours stages recommended by the CE and 2/ a bootstrap estimation of the catches at age variance for each completed stratified sampling design (i.e. with the quater and habour stages)as actually conducted in France. Then, a cost function is introduced to test the accuracy of the estimation according to sampling design and the number of samples required for a fixed level of precision.

2.1 Notations

We use the following notations:

- $\hat{D}$: the total landings estimator (in number)
- $\hat{D}_i$: the landings estimator at age i (in number)
- $\hat{p}_i$: the estimator of age i proportion of landings
- $\hat{D}_{ji}$: the estimator the j th class landings within age i landings

The sampling design considered in this analysis is stratified according to the following stratas :

Strata 1 : quater

Strata 2 : harbours set

Strata 3 : commercial category or mtier

The estimator of the landings at age i is

$$\hat{D}_i = \hat{D}p\hat{a}_i$$

and its variance estimator decomposed into three elements is:

$$\begin{aligned} Var(\hat{D}_i) &= Var(\hat{D}p\hat{a}_i) \\ &= \hat{pa_i^2}Var(\hat{D}) + \hat{D}^2Var(p\hat{a}_i) + Var(\hat{D})Var(p\hat{a}_i) \\ &= V_1 + V_2 + V_3 \end{aligned} \quad (1)$$

This expression shows the great importance of the precision in estimation of the age-length key compared to the precision in the estimation of the landings. The second elementt (V_2) of the variance will indeed be the most determinant component of $Var(\hat{D}_i)$ since it is a function of the squared landings.

2.2 Analytical variance formulation

An analytical formulation of the aged landings variance estimator is available (e.g. in Deriso and Queen 1999) in the case of a random sampling design (without any stratification). The operational strata (commercial category or métier) of the french sampling bans from applying this analytical results on french sampling. Consequently, we delopped a more complex but still analytical formula of the variance. First, we explicit the variance of the total landings and the variance of the proportion of landings at age i. The total landings estimator is the result of the length sampling, while the estimator of the proportion of landings at age i is provided by the age sampling. Indeed, the total landings estimator is

$$\hat{D} = \sum_j \hat{D}_j$$

and the estimator of age i proportion of the landings is:

$$\hat{pa}_i = \frac{\sum_j \hat{D_{ji}}}{\hat{D}} = \frac{\sum_j \hat{D}_j \hat{pa}_{ji}}{\sum_j \hat{D}_j}$$

These estimators are calculated using the estimator of the landings at length j,$\hat{D}_j$, and the estimator of the proportion of landings of age i in the length class j, $\hat{pa}_{ji}$, estimated respectively from the length sampling and the age sampling.

2.2.1 Stratified sampling by commercial category

In each commercial category, a sample of the landings is collected and each individuals is then attributed to its length group (Fig. 0).

Some additionnal notations for the stratified sampling by commercial category :

- k : k th commercial category
- K : the number of commercial category
- W_k : total landings of the k th commercial category in weight
- n_k : samples number of the k th commercial category
- v : the v th sample
- w_{kv} : the v th sample weight of the k th commercial category
- J : the number of length class
- I : the number of age group
- j : j th length class
- i : i th age group

- d_{jkv} : the number of fish belonging to the j th length class of sample v
- w_{jkv} : the weigth of fishes belonging to the j th length class of sample v
- M : the number of individual used to construct the age-length key
- m_j : the number of individual of length j of the age-length key
- pl_j : the proportion of individuals of length j of the age-length key
- q_{ji} : the proportion of individuals of length j and age i of the age-length key

The variance estimator of the total landings is the following.

$$Var(\hat{D}) = \sum_j Var(\hat{D}_j) + \sum_{j \neq j\prime} Cov(\hat{D}_j, \hat{D}_{j\prime})$$

We assume that $Cov(\hat{D}_j, \hat{D}_{j\prime}) = 0$, for all (j, j'), to simplify the calculation of the variance estimate.

Variance of landings at length

From the sampling design of the landings at length, the estimator of the landings at length j can be decomposed as follows,

$$\hat{D}_j = \sum_{k=1}^{K} \frac{W_k}{\sum_{v=1}^{n_k} w_{kv}} (\sum_{v=1}^{n_k} d_{jkv}) = \sum_k W_k \frac{\sum_v d_{jkv}}{\sum_v w_{kv}}$$

The estimator of the variance is

$$var(\hat{D}_j) = \sum_k W_k^2 var(\frac{\sum_v d_{jkv}}{\sum_v w_{jkv}})$$

and from Cochran (1977),

$$var(\frac{\sum_v d_{jkv}}{\sum_v w_{kv}}) = \frac{1 - \frac{\sum_v w_{kv}}{W_k}}{\frac{1}{n_k}(\sum_v w_{kv})^2} \frac{\sum_v (d_{jkv} - \frac{\sum_v d_{jkv}}{\sum_v w_{kv}} w_{kv})^2}{n_k - 1} \tag{2}$$

This last equation pointed out that the variability of sample sizes expressed in weight (w_{kv}) and in number(d_{jkv}) would penalized the variance of $\hat{D}_j$. To quantify this penality, we could introduce an average proportion at age. The variance estimator would be :

$$\hat{D}_j = \sum_{k=1}^{K} \frac{W_k}{\sum_{v=1}^{n_k} w_{kv}} (\sum_{v=1}^{n_k} d_{kv} \bar{p_{jk}}$$

An optimal sample size would be defined as the value of d_{kv} inducing a low variability in p_{jk}.

Variance of the proportion of landings at age i

From the sampling design of the age-length key, the estimator of the proportion of landings at age i is calculated by the following equation,

$$\hat{pa}_i = \sum_{j=1}^{J} q_{ji} pl_j$$

With an assumption of proportional allocation, the estimate of the variance of the proportion at age i (Kimura 1977, lai 1987)

$$var(\hat{pa}_i) = \sum_{j=1}^{J} (\frac{pl_j^2 q_{ji}(1-q_{ji})}{m_j} + \frac{pl_j(q_{ji}-pl_j)^2}{M})$$

2.2.2 Stratified sampling by mtier

In each métier, some vessels are randomly selected. For each vessel, landings are splitted into commercial categories which are all proportionnaly sampled (Fig. 0). It should be noticed that the weight by commercial category of each strata métier is either unknown or unreliable : only landings total weight is available. Consequently, number at lengthis estimated for each vessel, summed over all the sampled vessels of the métier and finaly rased at the métier strata.

Let us introduce some additionnal notations :

- l : l th métier
- L : the number of métier
- n_l : the métier l number of sample (i.e. the number of vessel sampled)
- w_{lv} : the v th sample weight of the l th métier
- $\hat{D}_{jl}$: the estimator of the j th length class landings of the l th métier
- d_{jlv} : the number of fish belonging to the j th length class of sample v
- w_{jlv} : the weigth of fishes belonging to the j th length class of sample v

For each vessel we need to define the commercial category level since a sample is collected by commercial category.

- $dech_{lvk}$: the number of individual in the sample collected into the k th commercial category of the v th sample Ech_{lv}
- $wech_{lvk}$: the sample weight collected within the k th commercial category of the v th sample Ech_{lv}
- d_{lvk} : the number of individual of the k th commercial category of the v th sample Ech_{lv}
- w_{lvk} : weight of the k th commercial category of the v th sample Ech_{lv}

Variance of landings at length The variance of the total landings is identical to the commercial category case :

$$Var(\hat{D}) = \sum_j Var(\hat{D}_j) + \sum_{j \neq j\prime} Cov(\hat{D}_j, \hat{D}_{j\prime})$$

The estimate of the landings of length j is expressed as

$$\hat{D}_j = \sum_{l=1}^{L} \hat{D}_{jl} \quad (3)$$

$$= \sum_{l=1}^{L} \sum_{v=1}^{n_l} w_l \frac{d_{jlv}}{w_{lv}} \quad (4)$$

$$= \sum_{l=1}^{L} w_l \sum_{v=1}^{n_l} \frac{\sum_{k=1}^{K} d_{jlvk}}{\sum_{k=1}^{K} w_{lvk}} \quad (5)$$

$$Var(\hat{D}_j) = \sum_l w_l^2 var(\sum_{v=1}^{n_l} \frac{\sum_{k=1}^{K} d_{jlvk}}{\sum_{k=1}^{K} w_{lvk}}) = \sum_l w_l^2 \sum_{k=1}^{K} var(\frac{\sum_{k=1}^{K} d_{jlvk}}{\sum_{k=1}^{K} w_{lvk}})$$

Using theorem 2.5 from Cochran (1977), we write :

$$var(\frac{\sum_{k=1}^{K} d_{jlvk}}{\sum_{k=1}^{K} w_{lvk}}) = \frac{1 - \frac{\sum_k w_{lvk}}{W_{lv}}}{\frac{1}{K}(\sum_k w_{lvk})^2} \frac{\sum_k (d_{jlvk} - \frac{\sum_k d_{jlvk}}{\sum_k w_{lvk}} w_{lvk})^2}{K-1}$$

Often d_{jlvk} is estimated by sampling the commercial category and an estimator of this number is given by:

$$\hat{d_{jlvk}} = w_{lkv} \frac{dech_{jlvk}}{wech_{lvk}}$$

Variance of the proportion of landings at age i

This variance is calculated using the same formula as for the stratified sampling by commercial category (Kimura 1977) :

$$var(p\hat{a}_i) = \sum_{j=1}^{J} (\frac{pl_j^2 q_{ji}(1-q_{ji})}{m_j} + \frac{pl_j(q_{ji} - pl_j)^2}{M})$$

2.3 Analytical tool for samples exploratory analysis

In a stratified sampling, each strata is supposed to split the population into homogeneous subpopulation regarding the estimated statistic. It is important to detect outlier samples and quantify their influence in the estimation of the statistic variance. We define an indice, called Δ, which quantified the discrepancy between the number at length in the sample and the adjusted mean number at length to the sample weight. This indice of discrepancy can be used 1/ to explore the samples of a strata (either commercial category or métier) regarding the length class (a)) or over all length classes (b)) and 2/ to qunatify the heterogeneity within a strata. With regards to these differents cases, the formula of Δ is the following :

1. given a strata k,

(a) given a length class j

$$\Delta_{jv} = d_{jkv} - \frac{\sum_v d_{jkv}}{\sum_v w_{kv}} w_{kv} \qquad (6)$$

(b) over all length classes

$$\Delta_v = \sum_j d_{jkv} - \frac{\sum_v d_{jkv}}{\sum_v w_{kv}} w_{kv} \qquad (7)$$

2. over all strata k

(a) given a length class j

$$\Delta_{jv} = d_{jkv} - \frac{\sum_{k,v} d_{jkv}}{\sum_{k,v} w_{kv}} w_{kv} \qquad (8)$$

(b) over all length classes

$$\Delta_v = \sum_j d_{jkv} - \frac{\sum_{k,v} d_{jkv}}{\sum_{k,v} w_{kv}} w_{kv} \qquad (9)$$

Note that Δ_{jv} matchs with the last part of equation 2.

2.4 Resampling method

Resampling techniques such as jackknife and bootstrap are often used for estimating confidence intervals or standard errors for any statistics. The principal advantages of these techniques are the easiness of implementation and the non reliance on normal therory. The fundamental assumption of bootstrapping developped by Efron (1979) is that the observed data are representative of the underlying population. In our case, bootstrapping has the advantage of proving the exactness of the analytical calculation and giving a CV value for a multiple stage subsampling. The multi stage subsampling can lead to a number of small samples, which would cause bias in the calculation of an estimator by the bootstrap technique. Chan and Lee (2001) recommend another algorithm for small sample bias reduction and base their work on less than 10 sample sizes. It is therefore recommended to have more than 10 samples in each stage to ascertain the convergence of the bootstrapped CV calculation.

The bootstrap method consists of drawing with replacement a number of new samples (usually 1000) from the observed data, each of the same size as the observed data. The statistic is calculated for each new set of data, yielding a bootstrap distribution for the statistic. It is possible to simulate stratifying sample by resampling independantly within each stratum. The final statistic is therefore a linear combination of each independant subsample statistics.

2.4.1 Bootstrap of one-stage sampling catch-at-length estimator

To compare with analytical result it can be of interest to bootstrap a simple one-stage catch-at-length estimator. One-stage sampling means that there is no stratification at all, neither time nor space or fleet stratification. All the samples are combined into one group like asked by the Data Directive.
The algorithm for such a bootstrap is as follows :

set-up :

- create a list of unique identifiers for sampling units (sample number)
- calculate values that will not change during the bootstrap process : total landing weight and number of samples

Bootstrap loop repeated at each iteration :

- create a table with all the information contained in the randomly selectioned samples
- calculate the length structure by summing all the samples with their respective raising factor
- calculate the catch-at-length by raising the length structure to the total catch
- append the estimates from this iteration to the output file

final calculation

- calculate variance, mean, CV, 5th and 95th percentile from the bootstrap distribution of each length class
- calculate a weighted CV for the length range that corresponds to 90% of the stock

2.4.2 Bootstrap of stratified sampling catch-at-length estimator

The software S+ allows the setting of one grouping variable. It is therefore possible to simulate a stratified sampling by creating a variable that contains the complete label of the lower level strata. For example, stratification by quarter (Q), harbours (H) and fleet (F) would generate a variable where modalities would be ”Q1 H1 F1”, ”Q1 H1 F2”, ”Q1 H2 F1”, etc...

The algorithm for such a bootstrap is as follows :

set-up :

- create a list of unique identifiers for sampling units (sample number + stratification label variable)
- calculate values that will not change during the bootstrap process : total landing weights and number of samples by strata

Bootstrap loop repeated at each iteration :

- create a table with all the information contained in the randomly selectioned samples
- calculate the length structure by summing all the samples with their respective raising factor
- calculate the catch-at-length by raising the length structure to the total catch
- append the estimates from this iteration to the output file

final calculation

- calculate variance, mean, CV, 5th and 95th percentile from the bootstrap distribution of each length class
- calculate a weighted CV for the length range that corresponds to 90% of the stock

2.4.3 Bootstrap catch-at-age estimator

Bootstrap catch-at-age precision estimator can be obtained by two methods :
1 - using the bootstrap variance of the catch-at-length combined with the Kimura variance of the ALK with the formulas given in chapter 2.1
2 - bootstrapping the Age-Length-Key under the assumption that a sample is one otolith with its fixed parameters (quarter, zone, length, estimation of age, etc...) and the grouping variable is the length, i.e. the number of otoliths read by length class is constant during the bootstrap process.

The algorithm for such a bootstrap is as follows :

set-up :

- create a list of unique identifiers for length sampling units (sample number + stratification label variable)
- create a list of unique identifiers for age sampling units
- calculate values that will not change during the bootstrap process : total landing weights and number of length samples and number of otoliths by strata

Bootstrap loop repeated at each iteration :

- create a table with all the information contained in the randomly selectioned length samples
- create an Age-Length-Key with the otoliths randomly selectioned by length classes
- calculate the length structure by summing all the samples with their respective raising factor
- calculate the catch-at-length by raising the length structure to the total catch
- combine the catch-at-length with the ALK to obtain catch-at-age
- append the estimates from this iteration to the output file

final calculation

- calculate variance, mean, CV, 5th and 95th percentile from the bootstrap distribution of each age class
- calculate a weighted CV for the age range that corresponds to 90% of the stock

2.5 Simulation

For reasons of computer time consuming, only the analytical simple random stratification is used to implement the simulation algorithm. The simulation is therefore usable at the strata disaggregated level or simulates a non stratified scheme if used at the final aggregated level. The following algorithm is based on the resampling technique assumption, i.e. the observed data are representative of the underlying population which allows bootstrap like multiplication or division of the sample numbers.

set-up :

- create a list of unique identifiers for length sampling units
- create a list of unique identifiers for age sampling units

Simulation double loop

- First loop with a vector of length number of samples multipliers $nlmult$ (from n/10 to 3n)
- Second loop with a vector of age number of individuals sampled multipliers $namult$ (from n/10 to 3n)
- if $nlmult <= n$ Selection nlmult number of samples without replacement among the n length samples
- if $nlmult > n$ Selection nlmult number of samples with replacement among the n length samples
- if $namult <= n$ Selection namult individuals without replacement in the Age-Length-Kry
- if $namult > n$ Selection namult individuals with replacement in the Age-Length-Key
- combine the catch-at-length raised from nlmult length samples with the namult individuals ALK to obtain catch-at-age
- Calculate the precision
- append the estimates from this iteration to the output file

final graph

- draw a contour plot of the double loop precision matrix

3 Materials

Market sampling in France is done either by commercial categories, either by fleet or métier. Usually, sampling by commercial categories needs stability in the fish sorting process as a prerequisite. To represent both methodologies, Atlantic hake and Eastern Channel sole sampling scheme will be described and analysed.

3.1 Atlantic hake

On the French atlantic coast a large number of fleets lands hake using different fishing means : gillnets, trawls and lines. ICES working groups have defined Fishery Units (FU) to coordinate international sampling process. French sampling is based on these fishery units and landings are sampled within 5 of them : FU05 (inshore fish trawler in ICES area VII), FU09 (nephrops trawlers in ICES area VIII) , FU10 (trawlers in ICES area VIII), FU12 (longliners in ICES area VIII) and FU13 (gillnetters in ICES area VIII). Sampling scheme is distributed among 6 harbours from south of Biscaye to south Britanny.
The length sampling objective is based on a number of fish to sample per FU and per quarter except for FU09 and FU10 based on a number of trips per month to sample.

FU05	2000 fish
FU09	10 trips per month
FU10	10 trips per month
FU12	1000 fish
FU13	500 fish

Age sampling scheme combines different methodologies

- fish purchase : 10 fishes per length classes per quarter
- Direct otoliths removal from market length sampling : 5 fishes per length classes per quarter with special attention to the largest fish
- Supplementary otoliths are provided by surveys (RESSGASC, EVHOE)

3.2 Eastern Channel sole

Sole landings are shared mostly between trawlers and gillnetters. Landings occur out of scallop season, i.e. from march to november. Regional landing distribution is about 30% for harbours between Cherbourg and Fecamp and 70% for harbours between Dieppe and Dunkerque.
The length sampling objective is based on commercial categories distributed among the principal harbours and quarter. At each sampling day, at least 3 samples from each commercial categories are sampled. One sample consists of measuring around 50 individuals , that is to say boxes with large number of small fish are splitted in two or three equal parts to avoid too much differences in the within category number of individuals sampled. On the other hand, boxes with very few number of individuals are skipped.
The age sampling is based on quarterly fish purchase. Once a quarter (usually in the middle), a fixed weight of each commercial categories is bought in order to have all the length classes range.

4 Results

4.1 exploratory analysis of the samples

The precision level of a multistage sampling estimator depends on the adequation between sampling effort and within strata variance. The first analysis is therefore to investigate on the internal variabilities within strata. To do so, the formula developped in chapter 2.2 is very informative,

more precisely the last part called distance to the mean distribution and calculated for each sample. This statistic is also the principal component of the variance calculation and very high values points out possible outliers or sample that takes the larger part of the mean distribution information. This statistic applied to Eastern Channel sole (Fig. 1) shows the importance of stratifying by commercial categories which are well discriminated one from each other. This figure shows also that special sampling effort has to be made on small fish category as it has the most variability. Focus on only one category subsampled by harbour or quarter (fig. 2) shows the same range of variability and the same symmetry around 0 for each strata. This means that these strata require the same sampling effort and stratification can not be designed in the purpose of reducing the total variance but rather for having sufficient information at a disaggregated level.
The same statistic applied to Atlantic hake (fig. 3) shows very weak differences in the range of sample distances to the mean distribution within strata. This means that sampling effort has to be proportional to the level of landings as internal variabilities in the different strata are in the same magnitude.

To continue with this exploratory analysis, precision can be improved by optimizing the relative importance of sampling effort against relative importance of landings per commercial category (Fig. 4a), quarter (Fig. 4b and 5b) and Fishery unit (Fig. 5a). This can only be a post-analysis as it is difficult to know these informations before going sampling, but it can detect some discrepancies to correct for the next years. For example, we can see that a special effort has to be made in the second quarter for Easter Channel Sole (Fig. 4b) and it would be useful to sample a little more Fishery unit 13 for Atlantic hake maybe instead of sampling so much Fishery unit 12 ((Fig 5a). These are the kind of issues to be discussed in local workshops to optimize the precision at a given sampling effort.

4.2 catch-at-length precision

Precision on length distribution for Easter Channel sole (Fig. 6) and Atlantic hake (Fig.7) put in evidence two main issues. $Primo$, at a given sample number (Eastern Channel sole and Atlantic hake have respectively 6365 and 23086 individuals sampled) stratifying by commercial category will largely improve the precision. $Secundo$, analytical and bootstrap methods give exactly the same picture, i.e. good CVs on well represented length classes in the length distribution and poor precision on scarce length classes. Moreover, CV estimations are very close with both methodologies thus validating each other result. The overall CV is the weighted mean on length range representing 90% of the stock and calculated on an annual basis as required by the Data Directive.

4.3 Age length keys precision

Precision in ALKs (Fig. 8 for Eastern Channel Sole and Fig. 9 for Atlantic hake) shows the same pattern as precision in length distribution, i.e. poor precision in scarce ages and good precision in well represented ages. The CV range is narrower because special effort is made to get information from all the length classes with a fixed or proportional allocation. Eastern Channel sole and Atlantic hake have respectively 1102 and 1420 otoliths read and overall CV is the weighted mean on age range representing 90% of the stock. Quartely CVs for both species are found to be around 10% - 12% but represent only the sampling precision not the age reading errors which are meant to be very important, especially for hake.

4.4 catch-at-age precision

The age distribution, which is a combination between the length distribution and the ALK, is very sensitive to the precision of the latter. The analytical 3 terms variance formula (chapter 2.1) enables to discriminate the relative contributions of the precision in the length structure and the precision in the ALK to the overall precision (Table 1). For Eastern Channel sole in 2002, the weighted mean relative contributions are respectively 26.1%, 73.8% and 0.1% for the terms associated to catch-at-length variance, ALK internal variance and the product of both. The preponderance of the age information over the length information has already been underlined in the EMAS project (Anon, 2001) and is important to know for optimizing the sampling scheme (Fig. 11). The definition of a cost function combined to this analytical approach enables to quantify the precision and cost of different arrangements of stratification in the sampling (quarter/harbour/CommercialCategory, quarter/port, quarter/CommercialCategory, harbour/CommercialCategory, . . .).

5 Results

5.1 exploratory analysis of the samples

The precision level of a stratified sampling estimator depends on the adequation between sampling effort and within strata variance. Furthemore, a good estimation of variance can be performed if sampling effort devoted to a strata is proportional to the relative part of the sampled population within this strata. An explanatory analysis of the collected samples is necessary before any precision estimation to detect outliers and to start an analysis of the sampling design adequacy. The first analysis is therefore to investigate on the internal variabilities within strata. To do so, the indice, denoted Δ, developped in section 2.3 is very informative for each sample. This statistic is also the principal component of the variance calculation and very high values points out possible outliers or sample that takes the larger part of the mean distribution information.

- **Heterogeneity over all strata and all length class:** Δ_v This statistic applied to Eastern Channel sole (Fig. 2) shows the importance of stratifying by commercial categories which are well discriminated one from each other. The same statistic applied to Atlantic hake (fig. 3) shows very weak differences in the range of sample distances to the mean distribution within strata. This means that sampling effort has to be proportional to the level of landings as internal variabilities in the different strata are in the same magnitude.

- **Heterogoneity within each strata :**$\Delta_k v$ Focus on only one category subsampled by harbour or quarter (fig. 4) shows the same range of variability and the same symmetry around 0 for each strata except for the right part of the graph which represents the smallest fish category. This means that special sampling effort has to be made on small fish category as it has the most variability. The other strata require the same sampling effort and stratification can not be designed in the purpose of reducing the total variance but rather for having sufficient information at a disaggregated level.

Assumption of sample representativity of the underlying population can be distorted by samples with too few individuals measured. On the other hand, there is an asymptotic upper limit to the number of individuals to measure in one sample. Limit where continuing measuring fish does not bring more information. This limit depends on the number of length classes and is different from one species to the other. Figure 5 shows the heterogeneity in the sample number and sample weight within each strata. This graphs enables to point out possible discrepancies in the database.

To continue with this exploratory analysis, precision can be improved by optimizing the relative importance of sampling effort against relative importance of landings per commercial category (Fig. 6a), quarter (Fig. 6b and 6b) and Fishery unit (Fig. 6a). This can only be a post-analysis as it is difficult to know these informations before going sampling, but it can detect some discrepancies to correct for the next years. For example, we can see that a special effort has to be made in the second quarter for Eastern Channel Sole (Fig. 6b) and it would be useful to sample a little more Fishery unit 13 for Atlantic hake maybe instead of sampling so much Fishery unit 12 (Fig 7a). These are the kind of issues to be discussed in local workshops to optimize the precision at a given sampling effort.

5.2 catch-at-length precision

Precision on length distribution for Easter Channel sole (Fig. 8) and Atlantic hake (Fig.9) put in evidence two main issues. $Primo$, at a given sample number (Eastern Channel sole and Atlantic hake have respectively 6365 and 23086 individuals sampled) stratifying by commercial category will largely improve the precision. $Secundo$, analytical and bootstrap methods give exactly the same picture, i.e. good CVs on well represented length classes in the length distribution and poor precision on scarce length classes. Moreover, CV estimations are very close with both methodologies thus validating each other result. The overall CV is the weighted mean on length range representing 90% of the stock and calculated on an annual basis as required by the Data Directive. Figure 8 and 9 represents the sampling as requested by the Data Directive, i.e. annual CV without stratification. The important question to raise is to know if stratifying would give a better precision with the same effort. The stratified bootstrap allows to calculate CV with any sets of strata and it can be very informative to try each arrangement. Here, we have only calculated a stratification by quarter and commercial category and sets of harbours and commercial categories for Eastern Channel sole (fig. 10). It is evident that there is no need to split into sets of harbours (fig. 10b) but stratifying by quarter (fig. 10a) increases the precision.

5.3 Age length keys precision

Precision in ALKs (Fig. 11 for Eastern Channel Sole and Fig. 12 for Atlantic hake) shows the same pattern as precision in length distribution, i.e. poor precision in scarce ages and good precision in well represented ages. The CV range is narrower because special effort is made to get information from all the length classes with a fixed or proportional allocation. Eastern Channel sole and Atlantic hake have respectively 1102 and 1420 otoliths read and overall CV is the weighted mean on age range representing 90% of the stock. Quartely CVs for both species are found to be around 10% - 12% but represent only the sampling precision not the age reading errors which are meant to be very important, especially for hake. Annual CVs obtained by adding the quarterly ALK matrix gave respectively 5% and 3.2% for sole and for hake, which represents a very high increase of precision. ALK is the example where concatening quarterly information will lead to bias estimates as for most of the fisheries recruitment occurs during the year modifying at a great extent quarterly ALK information.

5.4 catch-at-age precision

Figure 13 shows the slight difference between analytical and bootstrap estimated precision. The reason is that this calculation has been done with the bootstrapped catch-at-length combined to the analytical estimation of variance of the annual ALK. The difference is only due to the difference between analytical and bootstrapped precision estimation of the catch-at-length. To complete this work, it remains to estimate a bootstrapped CV on the quarterly ALKs and estimate the real quarterly stratified catch-at-age used for the assessment. The age distribution, which is a combination between the length distribution and the ALK, is very sensitive to the precision of the latter. The analytical 3 terms variance formula (chapter 2.1) enables to discriminate the relative contributions of the precision in the length structure and the precision in the ALK to the overall precision (Table 1). For Eastern Channel sole in 2002, the weighted mean relative contributions are respectively 26.1%, 73.8% and 0.1% for the terms associated to catch-at-length variance, ALK internal variance and the product of both. The preponderance of the age information over the length information has already been underlined in the EMAS project (Anon, 2001) and is important to quantify for optimizing the sampling scheme (Fig. 14). The definition of a cost function combined to this analytical approach would enable to quantify the precision and cost of different arrangements of stratification in the sampling (quarter/harbour/CommercialCategory, quarter/port, quarter/CommercialCategory, harbour/CommercialCategory, . . .).

6 discussion

It is trivial to say that precision level in biological sampling for fisheries data depends on the sampling design and on the different variables that are either measured or estimated. The main purpose of this paper is to propose two complementary methods to estimate the precision in the biological sampling and build a logical reasonning for optimizing the sampling design.
The most usual demand comes from ICES Stock Assessment Working Groups that need an age structure of the landings for some species. The age composition of the landings is one of the most important parameters used in stock assessment modeling. Apart from gear and catchability parameters, the model assumption are that the age structure of the landings is first representative of the age structure of catches and is proportional to the age structure of the stock. It is therefore important, not only to quantify the precision of the age structure estimation but also to avoid source of bias. Among sources of bias, the most important lead to mis-estimation of discards and not adequate sampling design because of a bad spatial, temporal and selectivity coverage. In the future, the demand will come to the providing of fleet-based disaggregated age structure of the landings. The need of precision at disaggregated level is also an important issue to address.

In this document, we focus on the precision of the age structure estimation regarding different sampling designs and on analysis of the adequacy of the sampling design to reduce bias of this estimation.

The first point investigated is that sampling design splits correctly the landings heterogeneity. To ensure that sampling is representative of the national landings, the sampling design must cover all the fleets during all the fishing period with a sampling effort proportional to the landings. This is a basic rule-of-thumb to avoid bias risk when within and between strata variances are of the same magnitude. To avoid bias, it is important to see if landings age structure is linked to an external factor like quarter, geographic area, harbour, fleet, etc. . . . The indice Δ is a candidate statistic for this kind of exploration. In the case of commercial category-based sampling, we have seen that stratifying by quarter increased the precision but stratifying by sets of harbours reduced it. Such investigations are very time consumming and tools to analyse this in a comfortable way would be much appreciated. This kind of investigation has not been carried out yet on Atlantic hake but first results on Eastern Channel sole shows the importance to address such issues. What is the interest of stratifying by metier, by harbours ? The Δ statistic associated with quantification of different strata arrangements would give elements to answer. Some questions remain open like the need or not to have market commercial categories stable sorting to use this information as a raising factor. This paper has shown the efficiency of such a sampling design to gain precision. It could be sufficient to have a relative qualification of the fish length like small, medium, big usable as strata and raising factor to have a non-biased precise and cheap estimation of catch-at-length.

A second important point is the problem of number of individuals to measure in a sample. A balance has to be found between the number of samples and the number of individuals in each sample for optimizing the fixed sampling effort.

To settle the important issue of opimizing sampling scheme, we have on one hand a fixed sampling effort and on the other hand information from previous years sampling acting like a pilot study. The first step of the reasonning is the exploratory analysis where it is possible to detect possible outliers usually very influent in the final raised numbers. Distance to mean distribution has been introduced to quantify unit sample influence on the final raised number. When ordered by an external variable, this statistic also enables to detect possible stratification for the purpose of reducing the variance or inversely combining strata that show the same sampling pattern.
As the final age structure is the combination of two estimations in a double sampling procedure, the second step of the reasonning should be the search of a balance in the sampling effort between them. The problem comes from the fact that the easily measured variable have a low contribution to the final estimation unlike the costly hard-to-measure variable. EMAS (1991) considered the possibility of picking otoliths at random from catches without regard to size of the fish in a multi-stage sampling. It is obvious that a large effort is brought to collect catch-at-length data at every possible disaggregated level at the expense of the effort on otolith reading.

Concerning the CV calculation, the analytical method is difficult to write precisely because it depends on the sampling design and depends on which variables are measured and which variables are estimated. Once analytical writing done, it is easy to implement and allows exploratory analysis (Δ statistic), decomposition of variance and simulation as it is not computer time consumming. The analytical writing becomes very fastidious at a stratify sampling level. The bootstrap method is easy to implement, allows to stratify and test different strata arrangements but is computer time consumming. Both methods are thus complementaries and gives a complete analysis of a sampling design.
Like a few previous papers, we have shown that precision in the age structure was mostly driven by the precision in the age-length-key (ALK). This particularity raises a few comments : (i) the maximum of attention must be given to the collection of ALK, in particular the one-time-a-quarter fish purchase is probably not a random sample from the total population of fish landed in the quarter by all fleets from the all geographical area. (ii) reading errors are not mentionned at this stage and represent a non negligible component of the ALK precision, hake beeing probably the most concerned species for this problem. The improvement in the sampling methodology to increase precision and avoir bias must be accompanied by an improvement in data quality. Mis-reporting and/or under-reporting errors affect the raising procedure at an unknown extent. Like otolith reading errors this problem is probably non-negligible to the contribution of overall precision in the estimation of catch-at-length and catch-at-length.

At last, at the stock scale, the combination of different national fleet-based disaggregated data require a specific international coordination work for defining the same fleet definition and the same sampling methodology.

Aknowledgement The authors would like to thank Dr Michel Bertignac (Ifremer Lorient, France) for having placed Atlantic hake sampling database to their disposal.

Analytical variance at age decomposition (Absolute and relative)

	Absolute value (millions)			Relative value (%)			Total var	Total	
Age	V1	V2	V3	V1	V2	V3	(millions)	Number	CV
1	6.7	589.9	0.7	0.011	0.988	0.001	597	81054	0.302
2	4196.5	9003.1	10.8	0.318	0.682	0.001	13210	3207625	0.036
3	11149.6	16788.8	20.0	0.399	0.600	0.001	27958	4233195	0.039
4	1510.7	8725.3	10.4	0.147	0.852	0.001	10246	788669	0.128
5	5031.7	13972.2	16.7	0.265	0.735	0.001	19021	1201299	0.115
6	241.7	3743.5	4.5	0.061	0.938	0.001	3990	152493	0.414
7	65.6	2046.2	2.4	0.031	0.968	0.001	2114	88970	0.517
8	37.9	1553.7	1.9	0.024	0.975	0.001	1593	35870	1.113
9	8.5	737.0	0.9	0.011	0.987	0.001	746	22124	1.235
10	17.7	1008.1	1.2	0.017	0.982	0.001	1027	23146	1.385
11	34.0	1302.8	1.6	0.025	0.973	0.001	1338	40781	0.897
12	55.5	1633.2	2.0	0.033	0.966	0.001	1691	34470	1.193
13	2.6	403.2	0.5	0.006	0.992	0.001	406	5651	3.567
14	1.7	261.4	0.3	0.006	0.992	0.001	263	6075	2.671
15	60.4	1670.7	2.0	0.035	0.964	0.001	1733	23093	1.803

Table 1: *Sole VIID 2002 - Analytical variance decomposition*

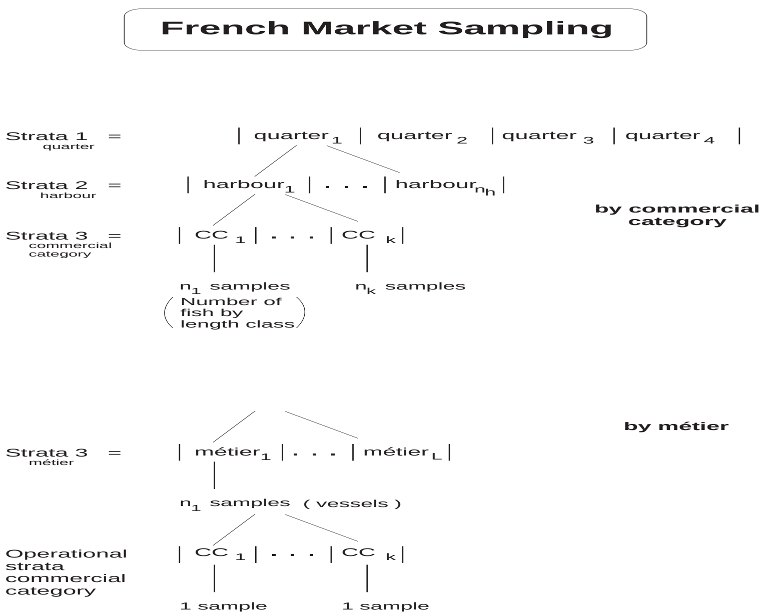


Figure 1: *French market sampling*

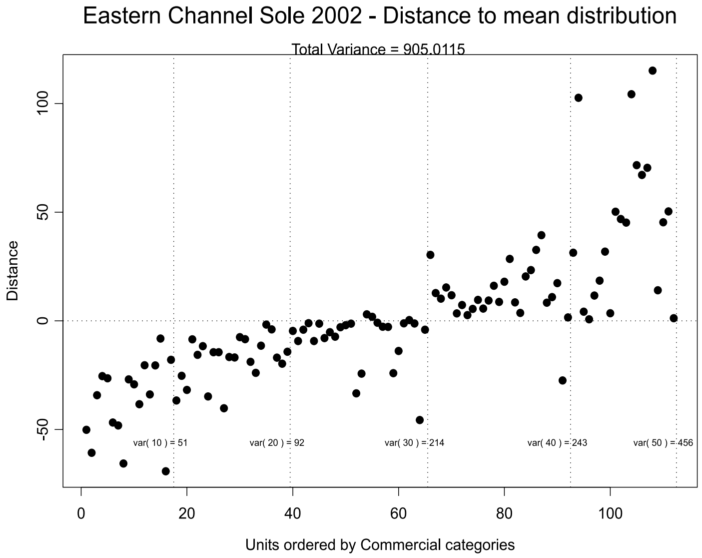


Figure 2: *Sole VIID 2002 - Distance to mean distribution. Samples ordered by commercial category*

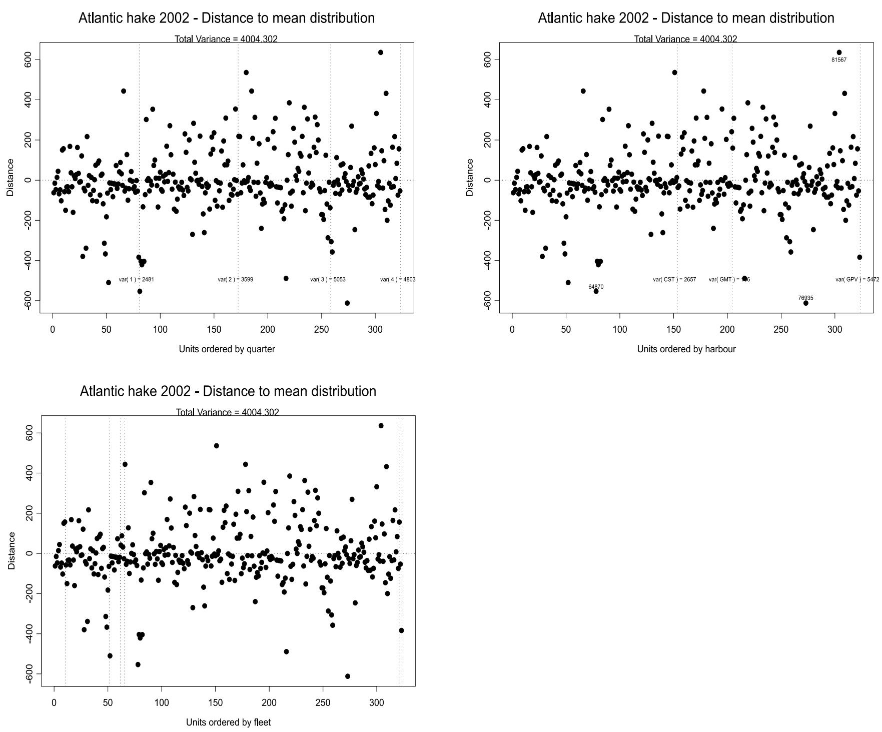


Figure 3: *Atlantic hake 2002 - Distance to mean distribution*

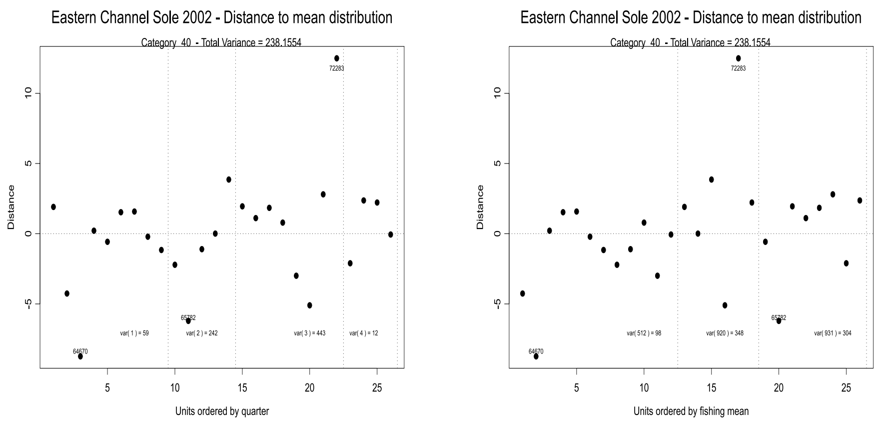


Figure 4: *Sole VIID 2002 - Distance to mean distribution for one commercial category. Samples ordered by a) quarter b) fishing mean*

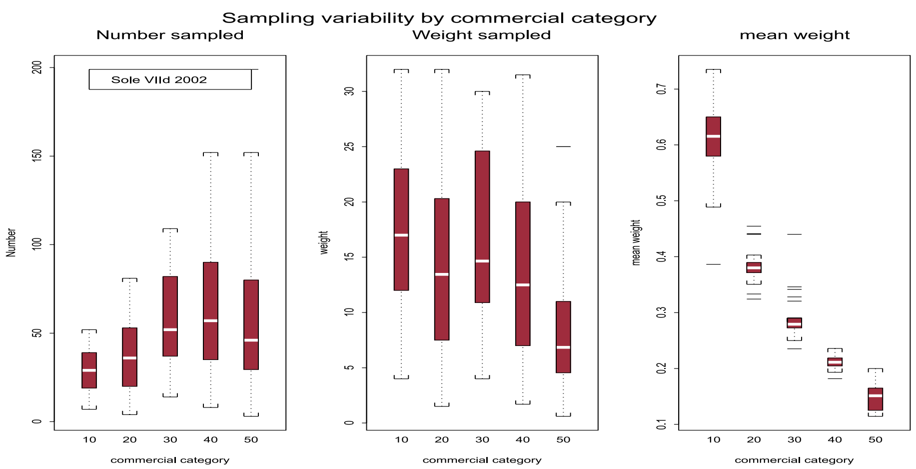


Figure 5: *Sole VIID 2002 - Sampling variability a) in number of fish measured by sample, b) sample weight and c) mean weight of a fish by sample*

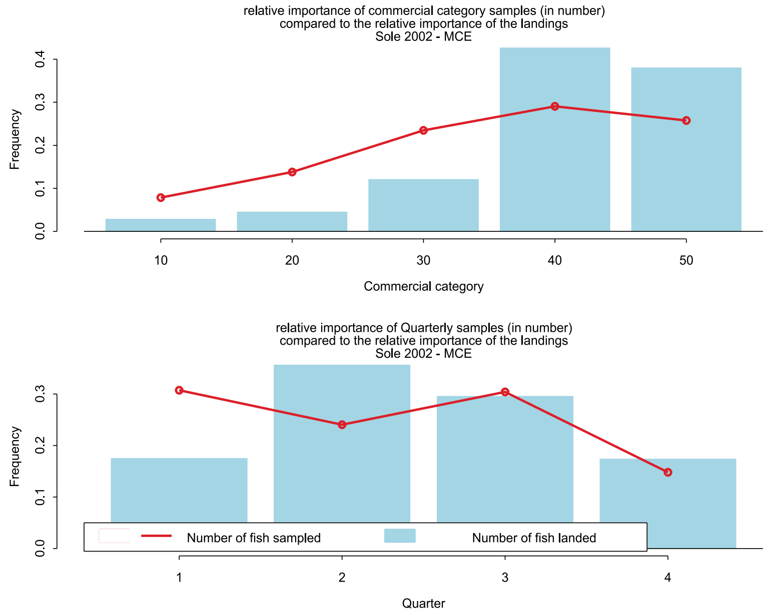


Figure 6: *Sole VIID 2002 - Relative importance of sampling effort against relative importance of landings by a) commercial category b)quarter*

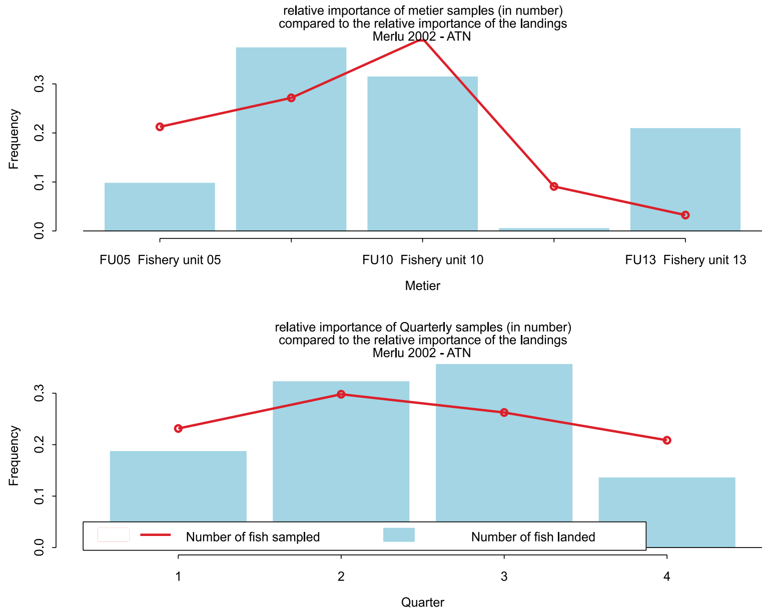


Figure 7: *Atlantic hake 2002 - Relative importance of sampling effort against relative importance of landings by a) fishery unit b) quarter*

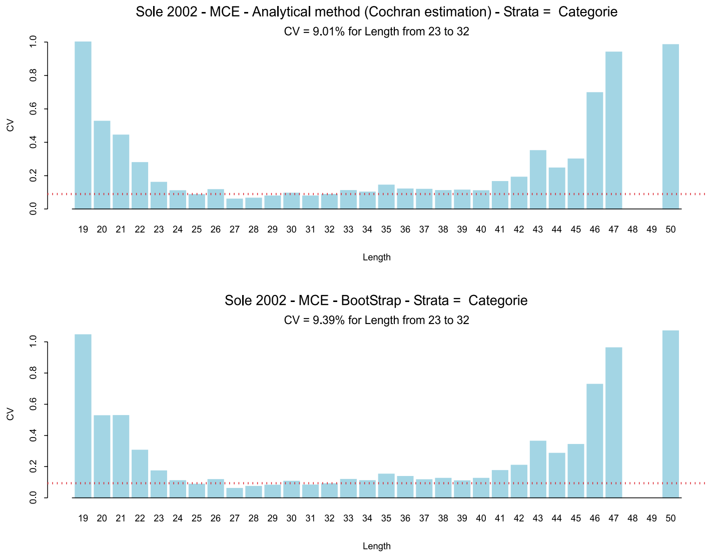


Figure 8: *Sole VIId 2002 - Precision in length distribution a) analytical b) bootstrap*

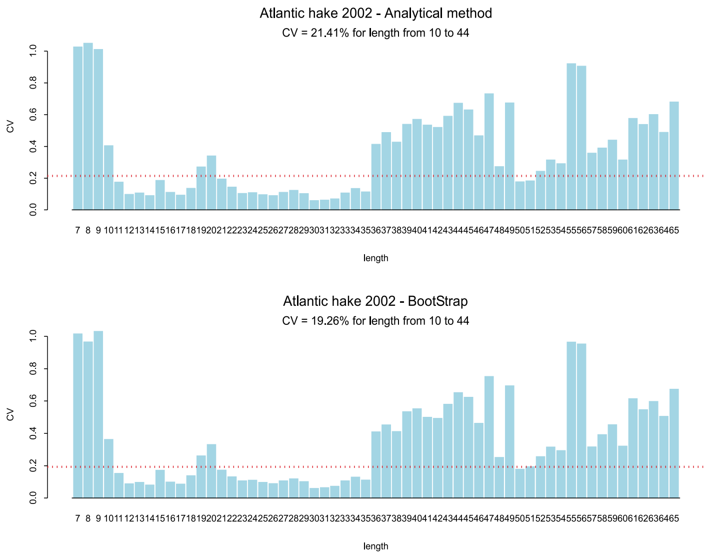


Figure 9: *Atlantic hake 2002 - Precision in length distribution a) analytical b) bootstrap*

Figure 10: *Sole VIID 2002 - Distance to mean distribution for one commercial category. Samples ordered by a) quarter b) fishing mean*

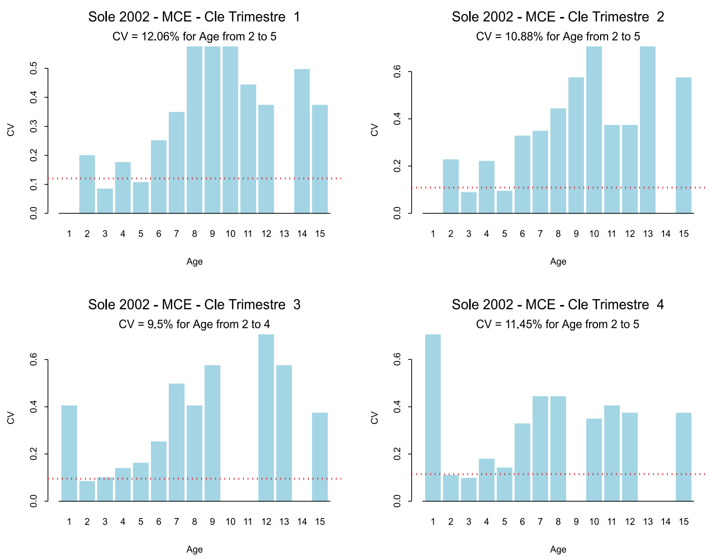


Figure 11: *Sole VIId 2002 - Precision in quarterly age length keys*

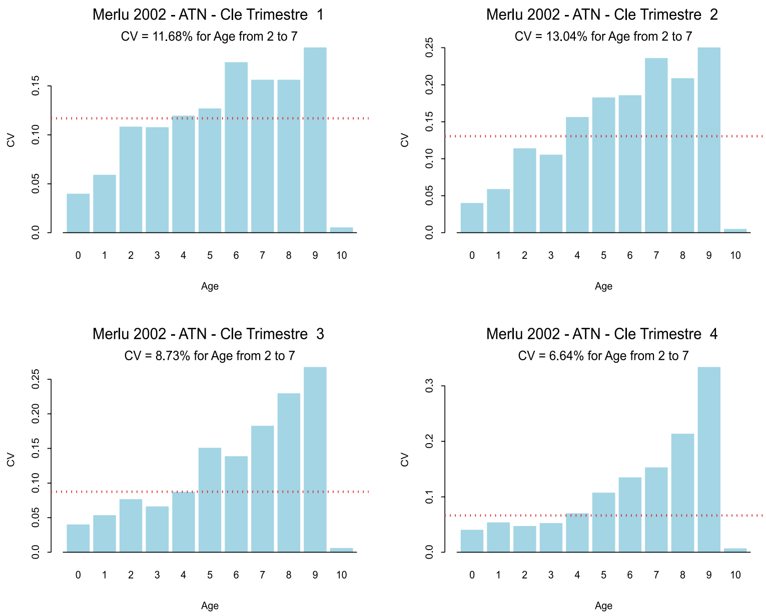


Figure 12: *Atlantic hake 2002 - CV at age in quarterly age length keys*

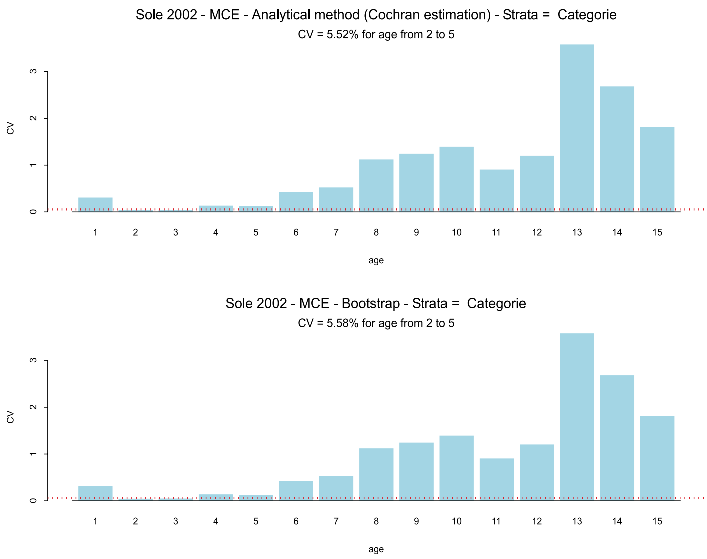


Figure 13: *Sole VIId 2002 - Precision in age distribution*

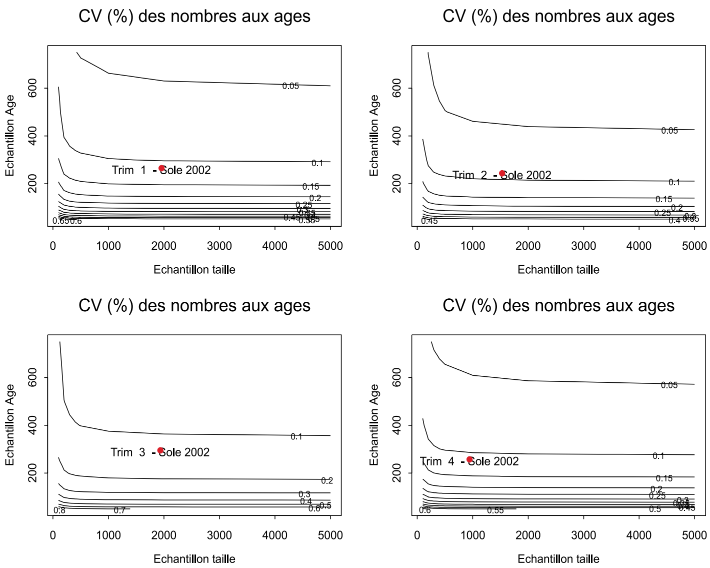


Figure 14: *Sole VIId 2002* - *simulation of precision from different sampling effort*

This paper is not to be cited without prior reference to the author(s)

International Council for the Exploration of the Sea

Quality and Precision of Basic Data Underlying Fish Stock Assessment and Implications for Fisheries Management Advice
CM 2001/P:13

The precision of international market sampling for North Sea plaice (*Pleuronectes platessa* L.) and its influence on stock assessment

C.M. O'Brien, C.D. Darby, D.L. Maxwell, B.D. Rackham, H. Degel, S. Flatman, M.A. Pastoors, E.J. Simmonds and M. Vinther

ABSTRACT

Market sampling is an essential source of data for age-based finfish stock assessment. Previously, little has been documented about the influence of potential error in these data on the precision of stock assessments and the management information they produce. This paper presents the results of an EU-funded study (CFP Study Project 98/075) of the precision of North Sea plaice (*Pleuronectes platessa* L.) fish market sampling carried out by the United Kingdom (England & Wales), Denmark and the Netherlands. Data from eight years of market sampling conducted for the period 1991-1998 were analysed to obtain the precision of estimated numbers-at-age in the landed catch. The annual market sample data were then used to generate 1000 realisations of the international numbers-at-age and mean weights-at-age in the landed catch for the eight-year period selected. Matrices of the catch numbers-at-age were computed for the generated realisations and these were used to produce 1000 stock assessments conditional on the XSA (Extended Survivors Analysis) model as implemented in 1999 at the ICES Working Group on the Assessment of Demersal Stocks in the North Sea. From the outcome of these assessments the influence of the market sampling programmes on the perception of the dynamics of the stock are graphically presented as simulation envelopes on the main management parameters.

C.M. O'Brien: CEFAS Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK [tel: +44 (0) 1502 524256, fax: +44 (0) 1502 513865, e-mail: c.m.obrien@cefas.co.uk]. C.D. Darby: CEFAS Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK [tel: +44 (0) 1502 524329, fax: +44 (0) 1502 513865, e-mail: c.d.darby@cefas.co.uk]. D.L. Maxwell: CEFAS Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK [tel: +44 (0) 1502 562244, fax: +44 (0) 1502 513865, e-mail: d.l.maxwell@cefas.co.uk]. B.D. Rackham: CEFAS Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK [tel: +44 (0) 1502 562244, fax: +44 (0) 1502 513865, e-mail: b.d.rackham@cefas.co.uk]. H. Degel: DIFRES, Charlottenlund Slot, 2920 Charlottenlund, Denmark [tel: +45 33 96 3386, fax: +45 33 96 3333, e-mail: hd@dfu.min.dk]. S. Flatman: CEFAS Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK [tel: +44 (0) 1502 562244, fax: +44 (0) 1502 513865, e-mail: s.flatman@cefas.co.uk]. M.A. Pastoors: RIVO, P.O. Box 67, 1970 AB IJmuiden, The Netherlands [tel: +31 255 564690, fax: +31 255 564 644, e-

mail:m.a.pastoors@rivo.wag-ur.nl]. E.J. Simmonds: FRS Marine Laboratory, P.O. Box 101, Aberdeen AB11 9DB, Scotland [tel: +44 1224 295366, fax: +44 1224 295511, e-mail: simmondsej@marlab.ac.uk]. M. Vinther: DIFRES, Charlottenlund Slot, 2920 Charlottenlund, Denmark [tel: +45 33 96 3300, fax: +45 33 96 3333, e-mail: mv@dfu.min.dk].

Keywords: bootstrap, mean-variance modelling, mortality, non-parametric, numbers-at-age, weights-at-age

INTRODUCTION

The catch-at-age matrix forms a major component of the assessment of many fish stocks. For stocks with international division of the catch, the sampling schemes are often diverse and the task of assembling the catch-at-age matrix is a complex and time-consuming process. The influence of this data on the assessment is rather poorly studied. To the knowledge of the authors, there are no comprehensive studies of the precision of international market sampling programmes and their implications for fisheries management advice. The papers published on this issue to date deal either with the potential effects of theoretical uncertainty in basic data on the advice provided (Pope and Gray 1983; Pope 1988; Pelletier and Gros 1994; Coggins and Quinn 1998) or on the estimation techniques and results of the analysis of uncertainty in the basic data itself (Tanaka 1953; Kimura 1977; Sparre et al. 1977; Gavaris and Gavaris 1983; Smith and Maguire 1983; ICES 1994; Reinert and Lewy 1998). In this paper, we present an attempt to combine these two approaches. This paper describes the analysis of catch-at-age data on North Sea plaice for the years from 1991 to 1998 inclusive, and their use in an age-based stock assessment.

To combine results from different sampling programmes to arrive at total international estimates of catch numbers-at-age and their associated variances, we could have followed two routes.

> Route 1: Attempt to combine the raw sampling data, calculate appropriate age-length keys (ALK) and raise the sampling data to the total international landings. In this way the variances of the procedure could be directly calculated (Gavaris and Gavaris 1983; Smith and Maguire 1983) or obtained from bootstrap analysis. A pre-requisite for this approach is that the sampling procedures (strata) are harmonised so that samples can be freely exchanged. This harmonisation is difficult to obtain from data already collected independently by different countries with different sampling and data storage methods.

Route 2: Use bootstrap techniques to generate a certain number of realisations of national age compositions and weights-at-age. Then combine these national realisations as a stock assessment working group would have done, delivering a number of realisations of the international age composition. These are then input into a stock assessment program to arrive at bootstrapped stock estimates.

It is this later approach that has been followed in this study.

We present first the results from studies of national market sampling programmes for estimating the catch numbers-at-age of North Sea plaice for the period 1991 to 1998, inclusive. Market sample data from the major fishing countries have been collated at the minimum aggregation level and used to generate 1000 national and then international replicates for use in bootstrapped assessments. The assessment procedure was that used at the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak ICES(2000).

MATERIAL AND METHODS

National data from the Netherlands, Denmark and the United Kingdom - England and Wales [subsequently denoted UK(E&W)], were used to provide 1000 replicates of national market sampling length, weight- and catch numbers-at-age. The sampling and raising methods are different for each nation, and consequently the methods for deriving the 1000 replicates were also different. These are described by nation below.

Bootstrapping the UK(E&W) data
Bootstrapping the catch-at-age data was carried out at the vessel level using code to replicate the raising calculations carried out by the market sampling system as closely as possible. The market samples were stratified by quarter, by fleet (beam trawl, otter trawl, seine) and by sex. The original data were extracted from the database that holds length and biological sample data, along with combined and raised processed data.

The algorithm used for the 1000 bootstraps was as follows:

Set-up:

- read in original length, age and weight data
- create a list of unique identifiers for sampling units – here used vessel code and sample number
- calculate values that will not change with each bootstrap sample – commercial weight totals and numbers of samples

Bootstrap loop that is repeated for 1000 iterations:

- set seed for random number generator
- form a bootstrap length sample by re-sampling length data

- calculate length distributions (LD) and analytical variance due to length sampling for bootstrap length sample using appropriate stratification and length groups (2cm for plaice)
- set seed for random number generator
- form bootstrap age sample by resampling age data
- calculate age-length key (ALK) and analytical variance due to ageing for the bootstrap age sample using appropriate stratification and length groups
- calculate age-length distribution (ALD) and analytical variance from LD and ALK
- calculate numbers-at-age using the LD and ALK
- calculate mean length within each length group and parameter for length-weight relationship
- calculate mean weight-at-age from ALD and length-weight relationship
- append the estimates from this iteration to the output file

Bootstrapping the Netherlands data

The bootstrap analysis of the Dutch data followed the same approach as UK(E&W), with the raising procedure adjusted to be specific to the Dutch case (Anon., 2001). Bootstrapping the catch-at-age data was carried out at the vessel level using code to replicate the raising calculations, with raising stratified by quarter, sex and market category.

Bootstrapping the Danish data

The Danish raising procedure for plaice is stratified by quarter of the year, landing region and market size class. Approximately 50% of the strata used have just one sample, which makes bootstrapping of just samples pointless. Therefore, the basic bootstrapping approach was extended by the bootstrapping of individual fish within a resampled market sample.

Internationally combined market data

The 1000 bootstrap replicates of mean weight-at-age and catch numbers-at-age for the Netherlands, Denmark and UK(E&W) were combined into 1000 replicates of international catch data sets. This fully sampled component constitutes around 75% of the North Sea plaice landings as given in ICES (2000) - refer to Table 1 in this manuscript. The major missing components compared to the Working Group values are landings' data from France and Belgium. In the assessments generated from the bootstrap data, the catch numbers-at-age have been scaled to the Working Group catch numbers-at-age in each year.

Calculating summary statistics

Once a 1000 catch numbers-at-age and weight-at-age replicates had been produced, mean values, variances, coefficients of variation (CVs) and correlations were calculated. The underlying relationship between mean and variance for catch numbers-at-age was investigated by fitting a linear regression of log(variance) on log(mean).

Assessment of North Sea plaice

The Extended Survivors Analysis (XSA) algorithm (Darby and Flatman 1994; Shepherd 1999) was modified to enable repeated fits of the model following replacement of the catch numbers-at-age and tuning fleet data for a user-specified range of years and ages. The estimates of the parameters of interest; namely, the recruitment, spawning stock biomass (SSB) and average fishing mortality calculated over a user-defined age range, were output during each iteration.

The XSA model was specified with the catchability and shrinkage constraints described in the report of the ICES Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (ICES 2000). Catchability was fitted as independent of population abundance for all ages; the catchability at each age greater than 10 was constrained to be equal to that estimated at age 10. The terminal populations were shrunk to the arithmetic mean of the fishing mortality estimated for the penultimate 5 oldest ages and the years 1993 – 1997. The coefficient of variation of the means used in the shrinkage was set at 0.5 and the minimum permitted value for the standard error of log-catchability set at 0.3. The assessment was applied to the catch numbers-at-age data for the years 1957 – 1998 as recorded by the ICES Working Group (ICES 2000). The catch per unit effort (CPUE) data for the tuning series was also extracted from that source. Two commercial and two survey CPUE series for the years 1989 – 1998 were used, no time series taper weighting was applied.

Two bootstrap assessments were run using the 1000 data sets from the bootstrap sampling of the international catch numbers-at-age and tuning fleet data. These were an assessment in which the catch at age data for the years 1991 – 1998 was replaced with a new sample at each iteration (series 1), and secondly replacement of the catch numbers-at-age data and the commercial fleet tuning data during each bootstrap iteration (series 2).

VPA-based assessment models make the assumption that the dominant portion of the uncertainty in the estimated population abundance and exploitation rate results from the process and measurement errors associated with the CPUE tuning series; the catch numbers-at-age data is assumed to be exact. In order to compare the potential magnitude of the errors generated from the tuning data series, a second non-parametric (NPB) bootstrap algorithm of the XSA model was developed. A base XSA model was fitted to the Working Group catch and tuning data sets. The base model log catchability residuals were sampled with replacement (independently by series and age) and used with the base model catchability parameters and VPA population abundance, to calculate new CPUE tuning data for all of the series. The bootstrapped CPUE values were then used with the base model catchability parameters to derive terminal population values for initiating a bootstrapped VPA. The weights assigned to terminal population estimates, calculated during the bootstrap process, were those estimated within the base XSA model. This non-parametric bootstrap algorithm assumes independence of the residuals by series and age and of the two commercial fleets' CPUE data from the catch number-at-age matrix.

The non-parametric bootstrap XSA model was used to generate 1000 replicates of the Working Group XSA model for North Sea plaice in order to examine the uncertainty associated with the tuning series information (series 3).

The final bootstrap model incorporated all of the bootstrap uncertainties, resulting from the sampling process and introduced by the tuning series. The bootstrap combined the algorithms described previously and performed 60 non-parametric bootstraps of the tuning series residual matrix for each of the 1000 catch and commercial CPUE data samples (series 4).

RESULTS

The bootstrapped catch numbers-at-age and weight-at-age of North Sea plaice from combining Dutch, Danish and UK(E&W) estimates are shown plotted with WG estimates in Figures 1 and 2, respectively. In all years the WG estimate of catch numbers-at-age is on the upper side of the bootstrap distribution as the WG estimates include additional age compositions from France and Belgium. The difference is particularly large for age 1 fish. Weights-at-age estimated from the bootstrap realisations seem to be in agreement with the WG estimates.

Catch numbers-at-age distributions
Histograms of the bootstrap estimates of catch numbers-at-age, scaled to the total international landings, are provided in Figures 3 and 4 for two arbitrarily chosen years (1991 and 1998). Superimposed on each histogram is a Gaussian kernel density estimate (solid line) (Sheather & Jones, 1991, MathSoft, 1998) and a normal distribution with the same mean and variance as the data (dashed line). These allow a visual inspection of how well normal distributions fit the data.

For North Sea plaice, the normal distribution closely matches the density estimate in all cases. This may be a result of the many stages of combining data involved in producing the international estimates.

Uncertainty and precision
The CV of catch numbers-at-age for the combined and national data sets are presented in Tables 2 through to 5. The CVs appear to be consistent across years and are similar for the three countries studied. CV of the national and international catch numbers-at-age follow the same pattern, with relatively higher CV on the very young and older age groups. As expected, the international CVs are lower than the national CVs. The international CVs of the age groups comprising the dominant proportion of the catch numbers-at-age are less than 5%.

The CV of the combined mean weight-at-age were generally less than 5% for most age groups and about 2% for the dominant age groups, further details are in the final report of CFP project 98/075 (Anon., 2001).

The underlying correlation of catch numbers-at-age was estimated using the numbers-at-age obtained from the resampling of the market sampling data. The patterns of positive and negative correlation were similar across the years so the mean correlation coefficients between estimates of catch numbers-at-age are given in the Table 6. Ages 2 and 3 show a negative correlation with older ages whilst the ages 5 through to 15 show a weak positive correlation.

Table 7 gives the coefficients by year from fitting a linear regression of log-variance against log-mean for international catch numbers-at-age. The coefficients are consistent across years with a slope of 1.42 (s.e. 0.05) for all years combined. Figure 5 shows the fit with all years combined. There is a series of seven outlying points above the line. These are the age 1 estimates from 1991-1997. Each is derived from only 2 quarters of data from one country so they have higher variances relative to their mean than the other estimates.

Evaluation of uncertainty in the stock assessment and management parameters

The time series of estimates of recruitment, SSB and average fishing mortality for ages 2 – 10 derived from the assessment in which the catch numbers-at-age data alone were bootstrapped (series 1) are presented in Figure 6. The results are consistent with those of the assessment carried out at the ICES Working Group; the Working Group estimates for each parameter lie on the median of the percentile distributions.

Re-sampling of the catch numbers-at-age data for the cohorts present in the years 1991 – 1998 has resulted in relatively negligible uncertainty in the estimates of average F, recruitment and SSB series; all CVs are less than 5%. The coefficients of variation of fishing mortality estimated in the final year are larger than those of the combined metrics, especially at age 1 where the expected value is close to zero. This result is not unexpected, as fishing mortality in the final year is a function of the ratio of two bootstrap replicates from the cohort, whereas SSB, recruitment and average F are derived from a weighted combination of the transformed replicates. Assessment models that are based on an underlying population structure reconstructed by VPA make the assumption that the catch numbers-at-age data are exact or, at least, that the effects of measurement errors in the catch numbers-at-age data can usually be ignored. The low CV values illustrated in Figure 6 indicate that, for the relatively well-sampled North Sea plaice assessment, this assumption appears reasonable.

Figure 7 illustrates the results of re-sampling the catch numbers-at-age data and the commercial fleet CPUE data for the years 1991 – 1998. Bootstrap re-sampling of the commercial fleet CPUE has increased the uncertainty associated with each of the estimated stock metrics. Uncertainty as measured by the CV decreases historically due to the convergence properties of the VPA equations. The results indicate increased uncertainty in the fishing mortality at the oldest ages of the assessment where sampling error increases due to reduced catch numbers.

The time series of percentiles of parameter values derived from 1000 non-parametric bootstrap assessments of the tuning series residuals with the catch numbers-at-age and tuning series catchability constant are illustrated in Figure 8. The bootstrap procedure re-constructs the CPUE series used to estimate the terminal population numbers for each cohort. Therefore, the main region of uncertainty is confined to the most recent years of the assessment time series and the oldest ages. Convergence of the VPA, conditional on the constant catch numbers-at-age matrix, produces the convergence of the percentiles of the historic estimates. The uncertainty in the estimates of fishing mortality in the final year is higher than the equivalent estimates for the sample bootstraps. The relative increase is higher at the youngest ages, which have lower CV as a result of less sampling noise in the calculation of the numbers-at-age, but relatively higher CV in the process and measurement error associated with the survey series.

The results for the fourth comparative series in which the two assessment bootstrap processes were combined are presented in Figure 9. It is seen that the two components of uncertainty have been additive. The historical variation induced by the re-sampling of catch numbers-at-age and CPUE data has been added to the uncertainty in the final year and oldest age population and exploitation rate estimates associated with the non-parametric bootstrapping of the tuning residuals. Apart from the first age, for this well-sampled stock, the CVs are all less than 30% with the majority of the uncertainty contributed by the errors in the tuning process.

DISCUSSION

The international sampling programmes appear to be delivering estimates of catch numbers-at-age that are rather precise, with CVs of 3.5% for plaice for the best estimated ages rising to about 15% at the older ages. While the precision of the best estimated ages is good, the current scheme is delivering poorer CVs on older ages. Care must be taken to ensure that the importance of estimating both old and young year classes is fully understood.

Based on the analysis of the histograms of the catch numbers-at-age, the normal distribution appears to be a reasonable description of the catch numbers-at-age data. This may be a result of the many stages of combining data involved in producing the international estimates and assumes independence among the national programmes.

The relationship between the mean and variance of the catch numbers-at-age is fundamental to any future statistical modelling; as is the assumption of independence between catch numbers-at-age. The underlying relationship between mean-variance of catch numbers-at-age was investigated by considering the mean and variance of the numbers-at-age obtained from the re-sampling of the market sampling data and compared to the power relationship:

$$\text{variance}\{\text{bootstrapped numbers-at-age}\} = e^{a} \,.\, \text{mean}\{\text{bootstrapped numbers-at-age}\}^{b}$$

Relationships between mean and variance were observed with slopes on the log variance-mean relationships of 1.42 (s.e. 0.05). Assessment models generally do not take this into account; changes to models or to weighting practices that would include these mean-variance relationships would be helpful. The apparent proportionality for the variance-mean relationship will facilitate the development of appropriate statistical models of catch-at-age that do not assume a log-normal distribution for catch numbers-at-age.

Negative correlations were observed between estimates of younger age classes, and weaker positive correlations were found between estimates of older ages. The correlations are considered to be a property of the population distributions, the fisheries and possible ageing errors. Future identification of the underlying processes will provide useful insights into their influence on assessment results and could lead to modification of sampling procedures.

These studies suggest that for the data sets examined the current levels of market sampling cause only small amounts of variability in assessment outputs for North Sea plaice, for the ages that are predominant in the catch data. The highest CVs, 15 - 18%, were estimated at the oldest ages and resulted primarily from sample noise in the catch numbers-at-age data.

As would be expected from a VPA method that assumes exact catch numbers-at-age, variance in the sampled catch numbers-at-age were transferred directly into the fishing mortality estimated for the final assessment year (c.f. 1998 in Table 2 and Figure 6). The introduction of sampling error to the commercial fleet CPUE series inflated the uncertainty in the management parameters but the increase in variance was less than that resulting from the catch numbers-at-age data. Again, this should be expected given that the commercial fleet CPUE data is a constituent part of the catch numbers-at-age data set and the effects of changes in abundance and sample error would be highly correlated. This raises the question as to the statistical validity in using commercial fleet tuning data twice when the expected correlations are high, firstly as part of the catch numbers-at-age and secondly as part of the tuning data.

On average the non-parametric bootstrap estimates of the uncertainty of the management parameters are higher than those derived from the re-sampling bootstrap. This results from the lower level of sampling associated with the smaller data sets used to obtain the CPUE data for tuning series. The CVs exhibit similar trends to the re-sampling bootstrap estimates, high values at the youngest and oldest ages. However, the method does not achieve the extreme values at the youngest age estimated from the sample data. As would be expected, combining the two bootstraps results in the highest estimates of uncertainty for the re-sampled time period. Figure 10 plots the CV of the terminal year fishing mortalities at age from the catch and fleet re-sampling bootstrap model, the non-parametric bootstrap model and the combined model. The figure illustrates that whereas the non-parametric bootstrap approach estimates similar values of uncertainty to the combined method when the re-sampling errors are low, as the magnitude of the sample errors increases at the older and youngest ages, uncertainty in the parameters is underestimated using this approach. The figure also shows that the errors are not additive and the combined method is required to allow for correlation.

ACKNOWLEDGEMENTS

This paper was prepared with funding support provided by the Department for Food, Environment and Rural Affairs (under contract MF0316) to the first four authors (CMO'B, CDD, DLM, BDR); together with funding support provided to all authors by the Commission of the European Communities Directorate General for Fisheries (DG XIV) under CFP Study Project 98/075 (Evaluation of market sampling strategies for a number of commercially exploited stocks in the North Sea and development of procedures for consistent data storage and retrieval (software) - EMAS).

REFERENCES

Anonymous (2001). Final report of CFP Study Project 98/075, Evaluation of market sampling strategies for a number of commercially exploited stocks in the North Sea and development of procedures for consistent data storage and retrieval (software) - EMAS.

Coggins, L. G. and T. J. Quinn (1998). A simulation study of the effects of aging error and sample size on sustained yield estimates. Fishery stock assessment models. F. Funk, T. J. Quinn, J. Heifetz et al**:** 955-975.

Darby, C.D. and Flatman, S. (1994). Virtual Population Analysis: Version 3.1 (Windows/DOS) User Guide. Information Technology Series, MAFF Directorate of Fisheries Research, Lowestoft, (1):85pp.

Gavaris, S. and C. A. Gavaris (1983). Estimation of catch at age and its variance for groundfish stocks in the Newfoundland region. Sampling commercial catches of marine fish and invertebrates. W. G. Doubleday and D. Rivard, Canadian Special Publication of Fisheries and Aquatic Sciences. **66:** 178-182.

ICES (1994). Report of the workshop on sampling strategies for age and maturity, Copenhagen, 3-9 February 1994. ICES C.M. 1994/D:1 Ref.: G,H,J.

ICES (2000). Report of the working group on the assessment of demersal stocks in the North Sea and Skagerrak. Copenhagen, 11-20 October 1999. ICES C.M. 2000/ACFM: 7.

ICES (2001). Report of the working group on the assessment of demersal stocks in the North Sea and Skagerrak. Copenhagen, 3-12 October 2000. ICES C.M. 2001/ACFM: 7.

Kimura, D. K. (1977). Statistical assessment of age-length-key. Journal of the Fisheries Research Board of Canada **34**: 317-324.

MathSoft. (1998). S-PLUS for windows, version 4.5. MathSoft, Inc., Seattle.

Pelletier, D. and P. Gros (1994). Assessing the impact of sampling errors on model-based management advice: an application to the Northwest Atlantic. Canadian Journal of Fisheries and Aquatic Sciences **48**: 2129-2139.

Pope, J. G. (1988). Collecting fisheries assessment data. Fish population dynamics**:** 63-82.

Pope, J. G. and D. Gray (1983). An investigation of the relationship between precision of assessment data and the precision of Total Allowable Catches. Sampling commercial catches of marine fish and invertebrates. W. G. Doubleday and D. Rivard, Canadian Special Publication of Fisheries and Aquatic Sciences. **66:** 151-157.

Reinert, R. and P. Lewy (1998). Sampling errors in estimated catch-at-age data in the Danish human consumption sampling program for the North Sea. Unpublished manuscript.

Sheather, S. J. and M. C. Jones (1991). A reliable data-based bandwidth selection method for kernel density estimation. Journal of the Royal Statistical Society series B **53**: 683-690.

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

Smith, S. J. and J. J. Maguire (1983). Estimating the variance of length composition samples. Sampling commercial catches of marine fish and invertebrates. W. G. Doubleday and D. Rivard, Canadian Special Publication of Fisheries and Aquatic Sciences. **66:** 165-170.

Sparre, P., H. Knudsen and S. Munch-Petersen (1977). Optimization of sampling programs for estimation of age distribution and fishing mortality.

Tanaka, S. (1953). Precision of age-composition of fish estimated by double sampling method using the length for stratification. Bulletin of the Japanese Society of Scientific Fisheries **19**: 657-670.

Table 1. Plaice comparison between mean bootstrap SOP and total landings estimated by the WGNSSK (ICES 2001).

Year	SOP bootstrapped countries	WG landings	%bootstrapped
1991	113048	148003	76%
1992	94383	125190	75%
1993	87612	117113	75%
1994	86098	110392	78%
1995	73789	98356	75%
1996	62292	81673	76%
1997	63425	83048	76%
1998	52949	71534	74%

Table 2. CV (%) of estimated catch numbers-at-age (combined data).

	1991	1992	1993	1994	1995	1996	1997	1998	Average
1	42.3	76.2	40.2	44.9	34.6	43.0	53.3		47.8
2	11.9	10.7	10.8	7.5	8.2	5.7	6.5	5.4	8.3
3	5.0	4.2	4.7	3.4	3.9	3.1	2.7	3.0	3.7
4	3.8	3.9	3.2	3.6	3.1	3.7	3.3	2.8	3.4
5	3.5	3.6	3.8	3.9	3.8	4.1	4.7	3.7	3.9
6	3.3	4.3	4.1	4.4	5.3	4.4	3.7	5.8	4.4
7	6.5	4.1	5.4	5.5	6.1	4.8	4.8	6.3	5.4
8	7.5	7.1	4.5	6.8	6.8	6.0	5.6	6.6	6.4
9	7.7	7.5	6.3	6.0	7.8	6.0	7.0	7.4	7.0
10	7.1	9.8	8.8	8.1	7.4	9.0	9.8	8.3	8.5
11	10.4	9.7	9.3	9.7	12.7	7.8	12.6	10.0	10.3
12	12.0	12.5	10.0	12.3	14.4	10.3	11.1	11.0	11.7
13	14.5	13.9	14.8	12.2	13.8	15.0	13.4	10.3	13.5
14	21.0	16.3	16.9	16.3	14.7	16.0	15.5	14.8	16.4
15+	8.4	10.1	8.0	9.5	10.5	9.9	9.0	9.2	9.3

Table 3.
CV of numbers at age

country	EW
species	PLE
area	IV

Average of cv_num	year							
age	1991	1992	1993	1994	1995	1996	1997	1998	Grand Total
1		80		90					85
2	17	16	19	17	17	18	22	14	17
3	9	7	7	7	9	10	8	6	8
4	5	6	4	5	5	7	5	4	5
5	5	5	5	6	7	5	7	6	6
6	3	6	5	7	7	6	6	7	6
7	7	5	8	8	9	7	7	7	7
8	10	11	6	11	10	9	8	9	9
9	12	12	9	9	11	9	9	9	10
10	10	16	12	12	11	11	14	10	12
11	16	13	12	14	16	10	16	12	14
12	15	20	12	14	17	13	13	13	15
13	18	17	19	15	15	17	14	11	16
14	27	20	19	17	16	18	17	16	19
15	9	11	8	10	11	10	9	9	10

Table 4.
CV of numbers at age

country	NL
species	PLE
area	IV

Average of cv_num	year								
age	1991	1992	1993	1994	1995	1996	1997	1998	Grand Total
1	42	72	40	44	35	43	53		47
2	13	13	12	8	9	7	8	8	10
3	6	5	7	5	5	4	4	4	5
4	6	6	5	6	5	5	5	4	5
5	6	7	7	7	6	7	8	6	7
6	5	7	8	8	10	8	7	10	8
7	11	9	9	10	11	10	10	16	10
8	13	13	9	10	10	12	13	13	12
9	14	13	12	11	13	11	15	16	13
10	13	15	16	14	10	18	16	16	15
11	15	15	13	16	19	12	19	17	16
12	23	22	24	21	22	21	18	18	21
13	29	30	28	21	38	26	27	24	28
14	28	27	30	56	33	41	32	30	35
15	21	19	33	28	31	33	25	40	29

Table 5.

CV of numbers at age

country	DK
species	PLE
area	IV

Average of cv_num	year								
age	1991	1992	1993	1994	1995	1996	1997	1998	Grand Total
1					57				57
2	7	6	9	16	9	5	7	3	8
3	8	5	3	3	3	3	3	3	4
4	4	5	4	4	4	4	4	4	4
5	5	3	3	6	5	4	7	5	5
6	5	6	4	6	9	6	7	10	6
7	7	6	7	8	11	6	8	9	8
8	12	11	7	13	16	10	10	13	11
9	11	13	10	12	19	14	20	16	15
10	14	13	17	18	19	18	27	30	20
11	26	20	21	26	25	25	55	48	31
12	32	25	27	46	49	29	66	57	41
13	43	39	45	36	41	54	68	52	47
14	59	53	67	46		61	49	50	55
15	42	53	48	52	50	57	61	51	52

Table 6. Mean correlation coefficient for catch numbers-at-age from 1991-1998, for combined Danish, UK(E&W) and Dutch bootstrapped estimates.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15+
1	1.00	0.23	-0.28	-0.19	-0.13	-0.11	-0.03	0.00	-0.01	-0.05	-0.03	0.01	-0.01	-0.01	-0.02
2		1.00	-0.30	-0.48	-0.39	-0.23	-0.15	-0.07	-0.07	-0.07	-0.07	-0.05	-0.03	-0.04	-0.02
3			1.00	-0.16	-0.26	-0.22	-0.23	-0.16	-0.12	-0.10	-0.06	-0.08	-0.05	-0.05	-0.06
4				1.00	0.11	0.01	-0.04	-0.08	-0.02	-0.01	-0.01	0.02	0.00	0.02	-0.04
5					1.00	0.24	0.18	0.06	0.05	0.03	0.04	0.00	0.00	0.00	-0.05
6						1.00	0.12	0.14	0.01	0.04	0.01	0.02	0.00	0.02	-0.02
7							1.00	0.17	0.14	0.05	0.02	0.03	0.05	0.03	0.03
8								1.00	0.13	0.05	0.05	0.09	0.02	0.07	0.07
9									1.00	0.10	0.04	0.09	0.05	0.10	0.05
10										1.00	0.09	0.05	0.06	0.08	0.08
11											1.00	0.09	0.03	0.04	0.10
12												1.00	0.04	0.08	0.14
13													1.00	0.10	0.16
14														1.00	0.19
15+															1.00

Table 7. Intercept (a) and slope (b) estimates from a linear regression of log(variance) on log(mean) for combined Danish, UK(E&W) and Dutch bootstrapped estimates (ages 1-14).

Year	a	b	st error a	st error b
1991	-0.48	1.47	0.94	0.11
1992	0.23	1.40	1.61	0.18
1993	-0.59	1.48	1.38	0.16
1994	0.26	1.37	1.04	0.12
1995	-0.67	1.48	1.26	0.15
1996	0.71	1.29	1.04	0.12
1997	1.00	1.25	1.00	0.12
1998	-0.97	1.44	0.39	0.05
All years	-0.22	1.42	0.40	0.05

Figure 1. Estimated catch numbers-at-age for 1991 to 1998, showing WG catch (line) and 1000 bootstrap estimates (points) from combination of Danish, Dutch and UK(E&W) bootstrap estimates.

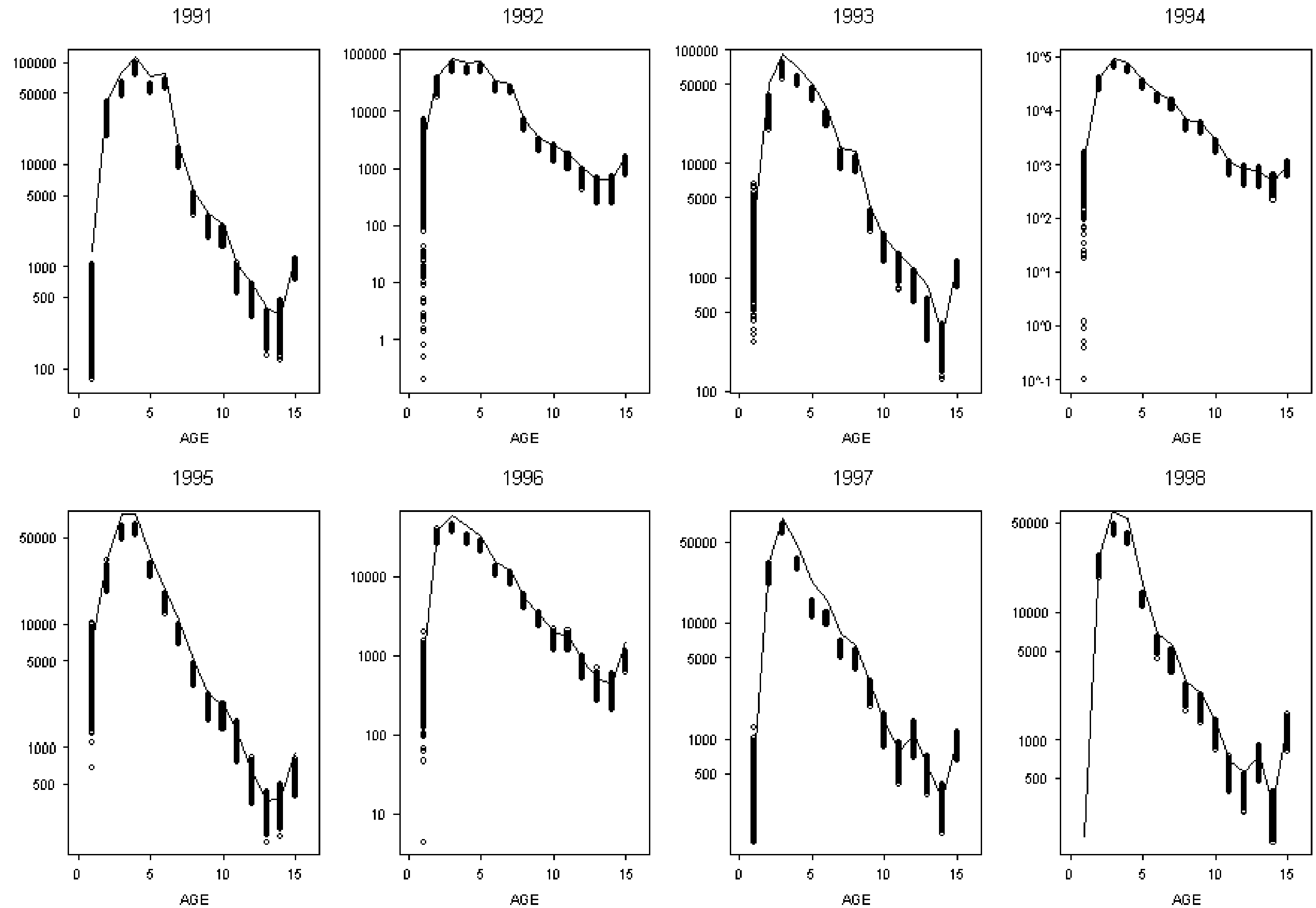


Figure 2. Estimated mean catch weights-at-age for 1991 to 1998, showing WG weights (line) and 1000 bootstrap estimates (points) from combination of Danish, Dutch and UK(E&W) bootstrap estimates.

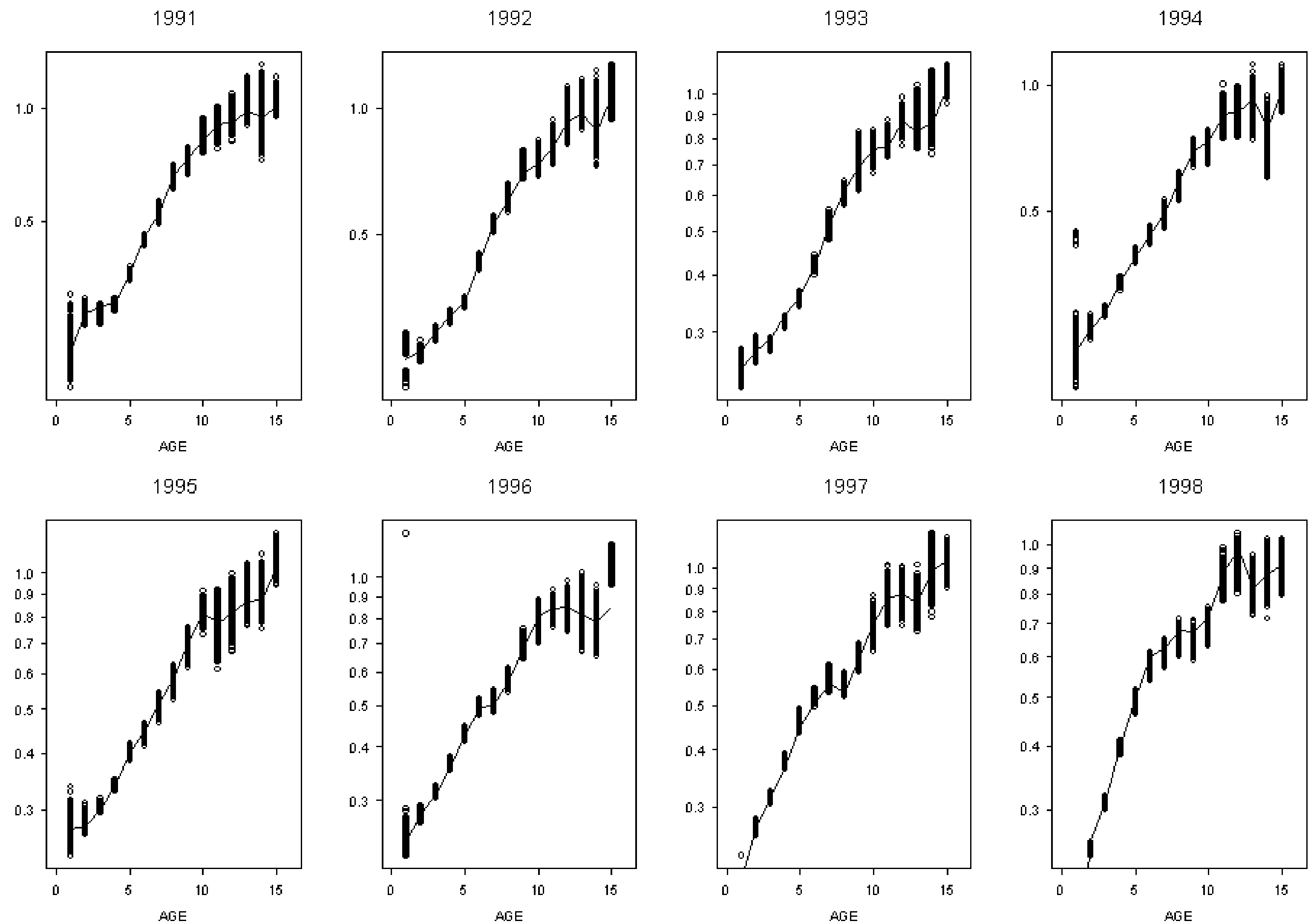


Figure 3. Histograms of bootstrap estimates of raised international catch numbers-at-age (for 1991).
Note that not all ages are shown in this figure (only ages 1-9 shown).

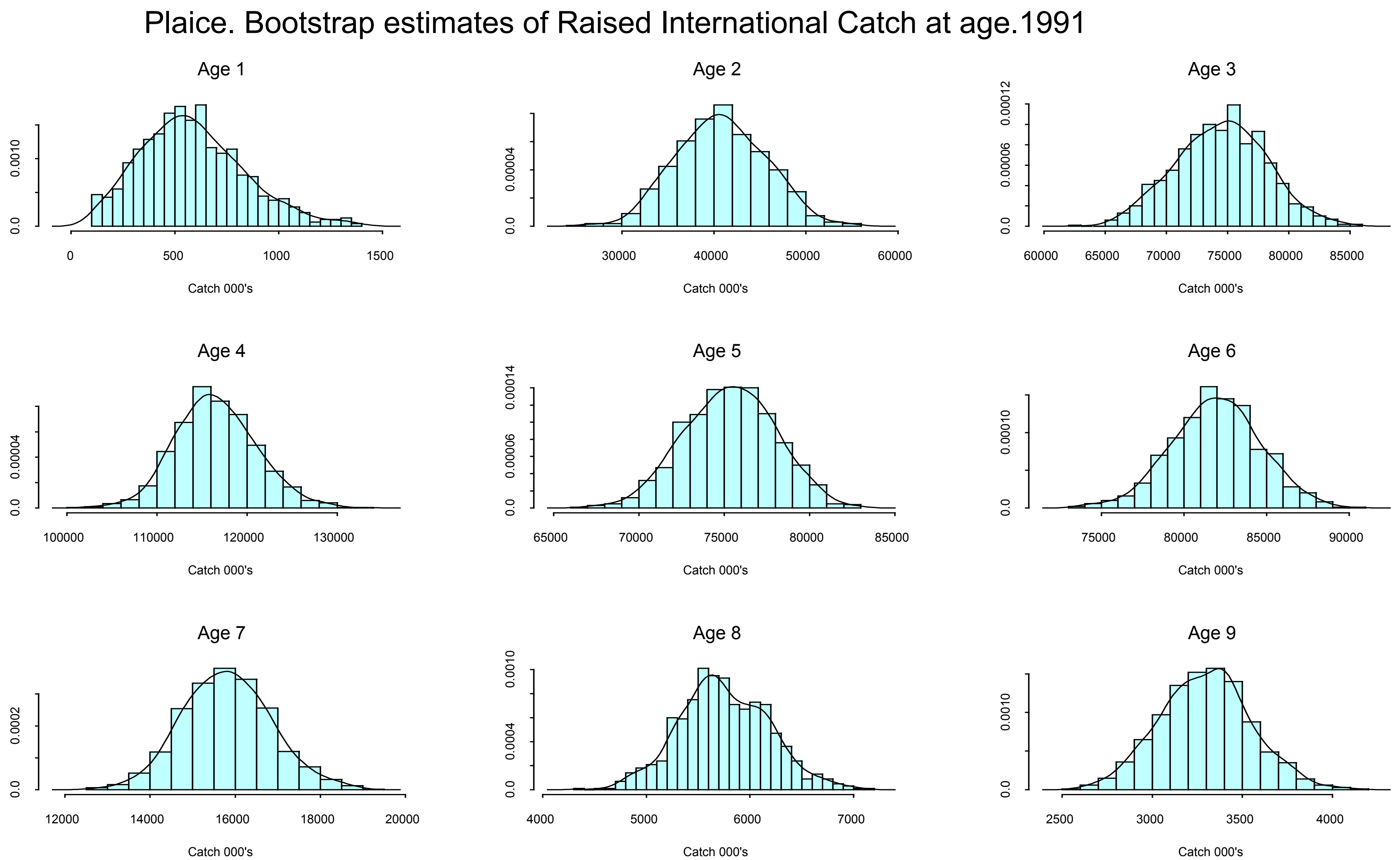


Figure 4. Histograms of bootstrap estimates of raised international catch numbers-at-age (for 1998).

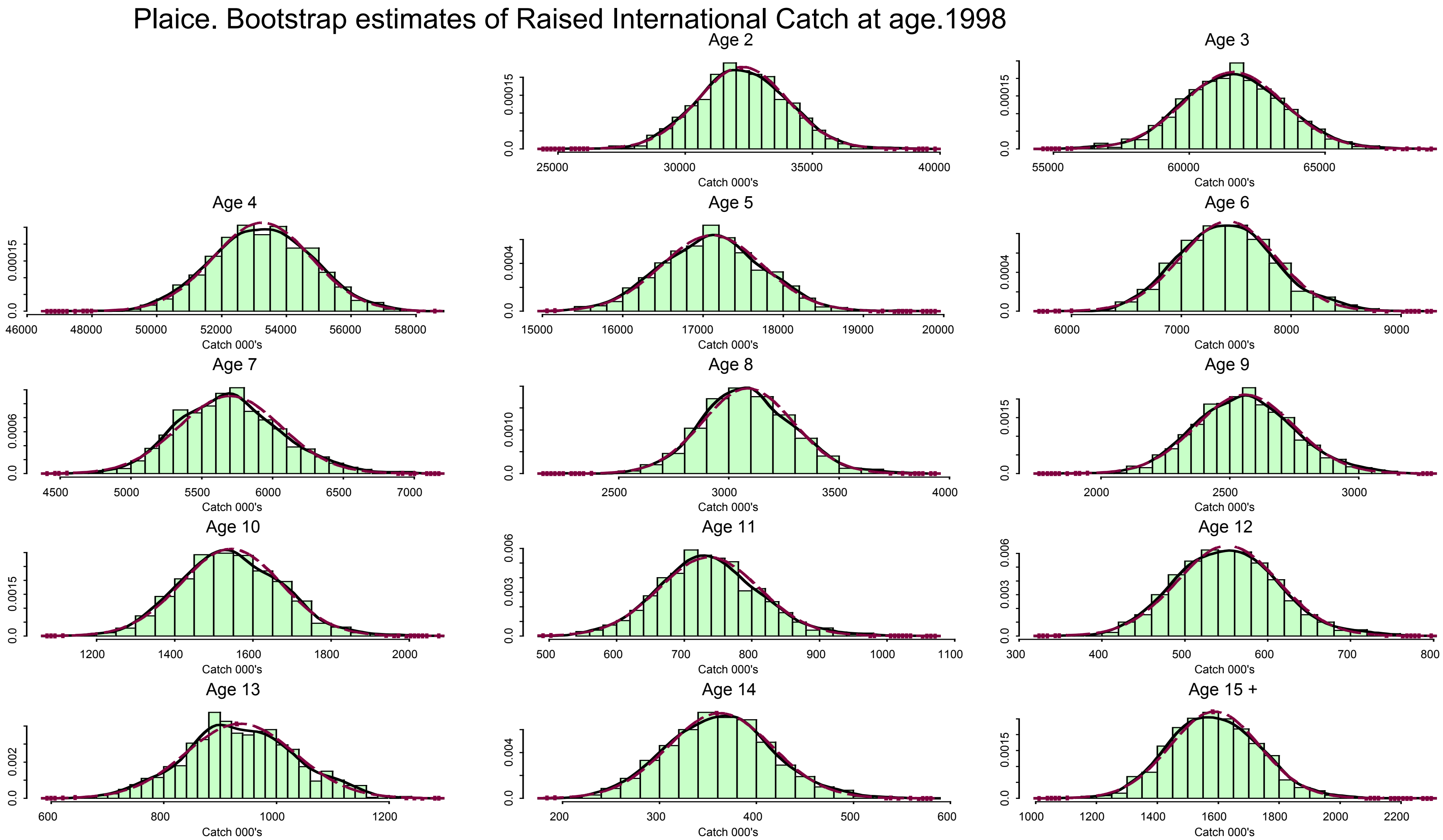
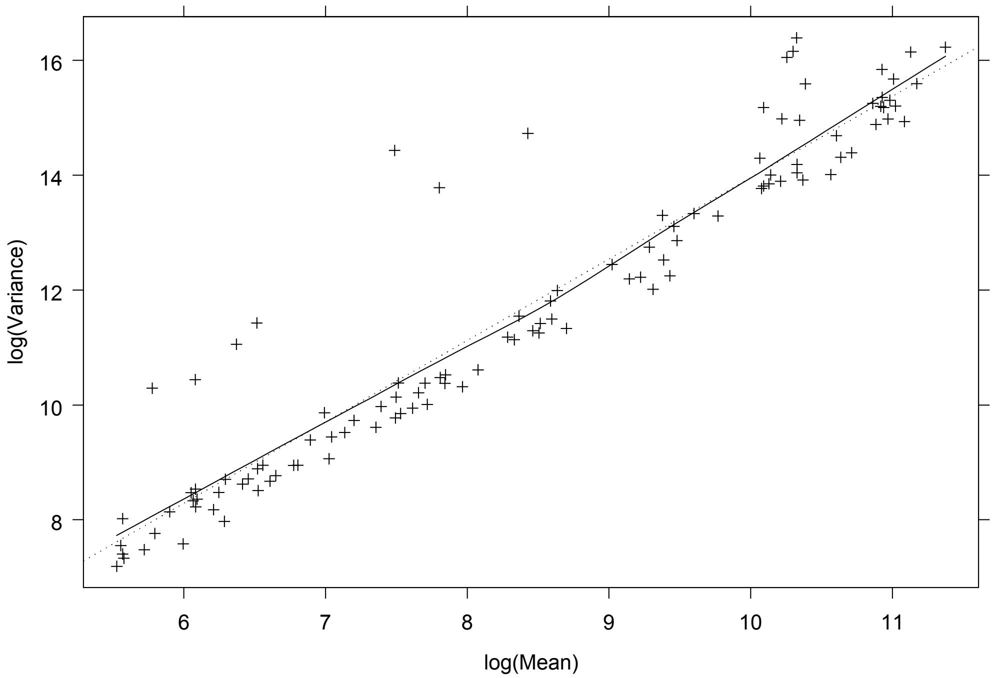


Figure 5. Log(variance)-log(mean) plots of combined Danish, UK(E&W) and Dutch bootstrapped estimates. Ages 1-14, years 1991 to 1998.



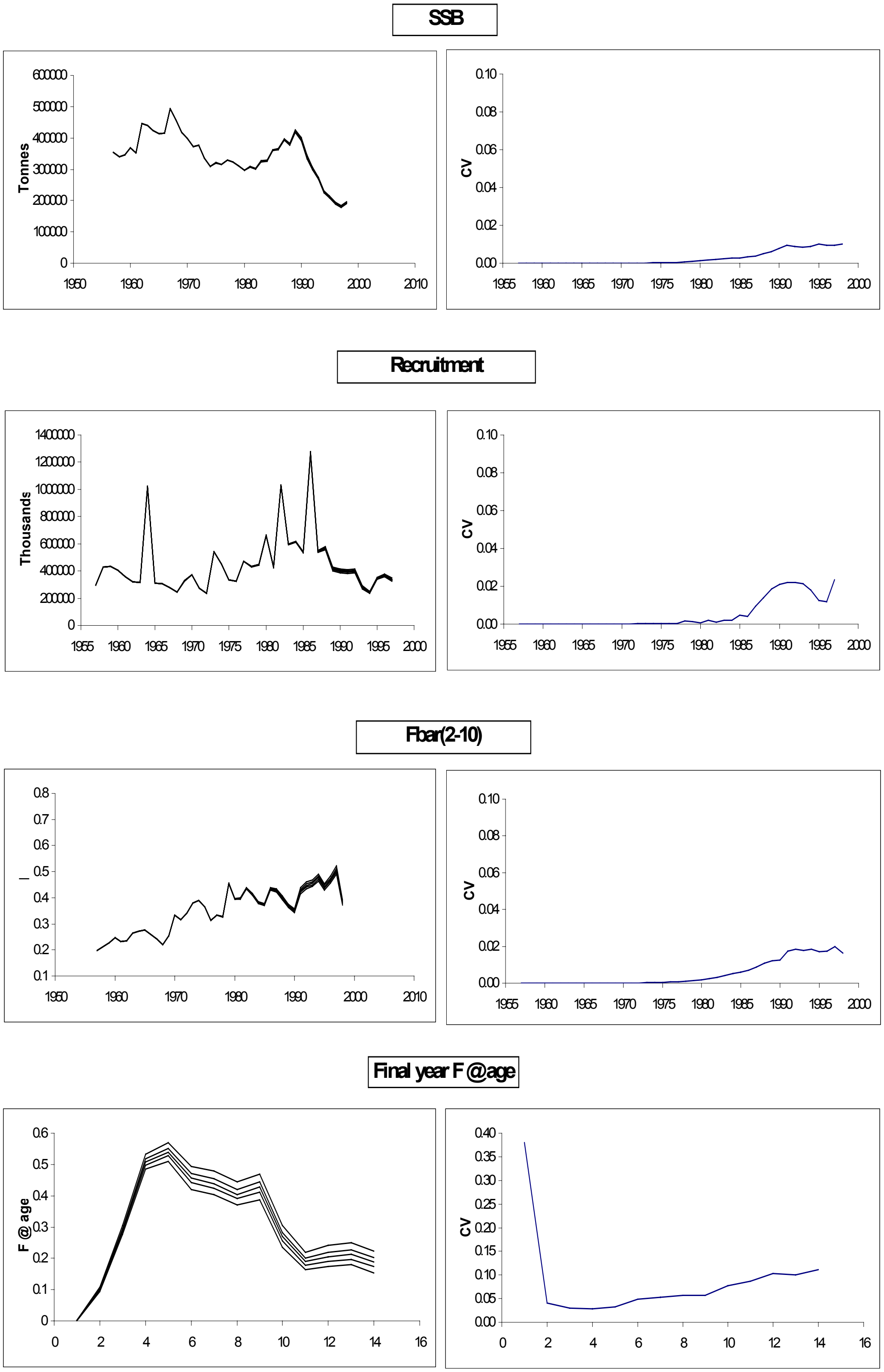


Figure 6. The 5,25,50,75,95th percentiles of Fbar (2-10), recruitment at age 1, SSB and F at age in the 1998 resulting from fitting the 1999 ICES WG XSA model to 1000 bootstraps of the North Sea plaice catch at age data for the years 1991 - 1998. Tuning fleet CPUE data constant.

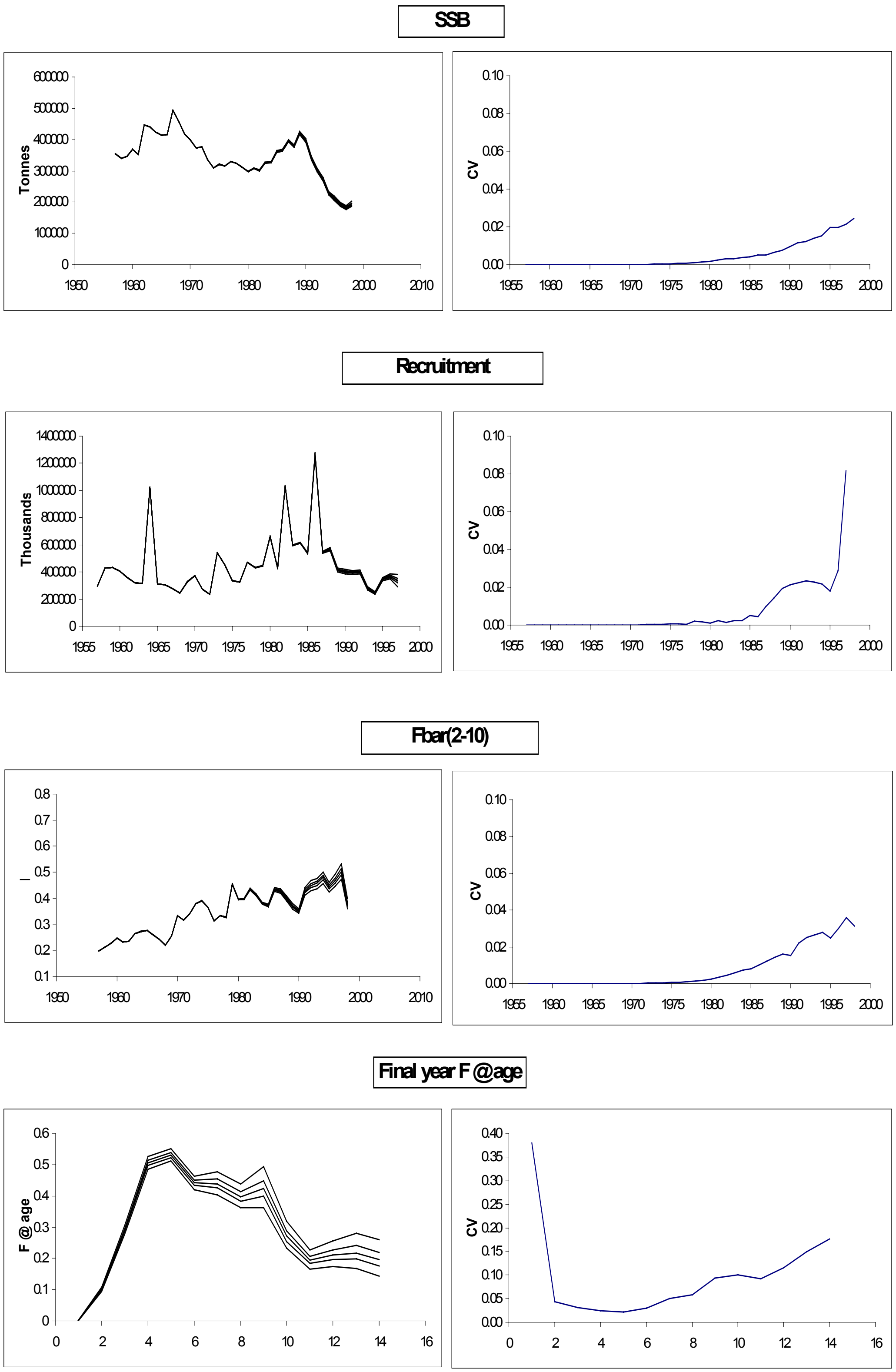


Figure 7. The 5,25,50,75,95th percentiles of Fbar (2-10), recruitment at age 1, SSB and F at age in the 1998 resulting from fitting the 1999 ICES WG XSA model to 1000 bootstraps of the North Sea plaice catch at age and commercial fleet tuning data for the years 1991 - 1998. Survey CPUE data constant.

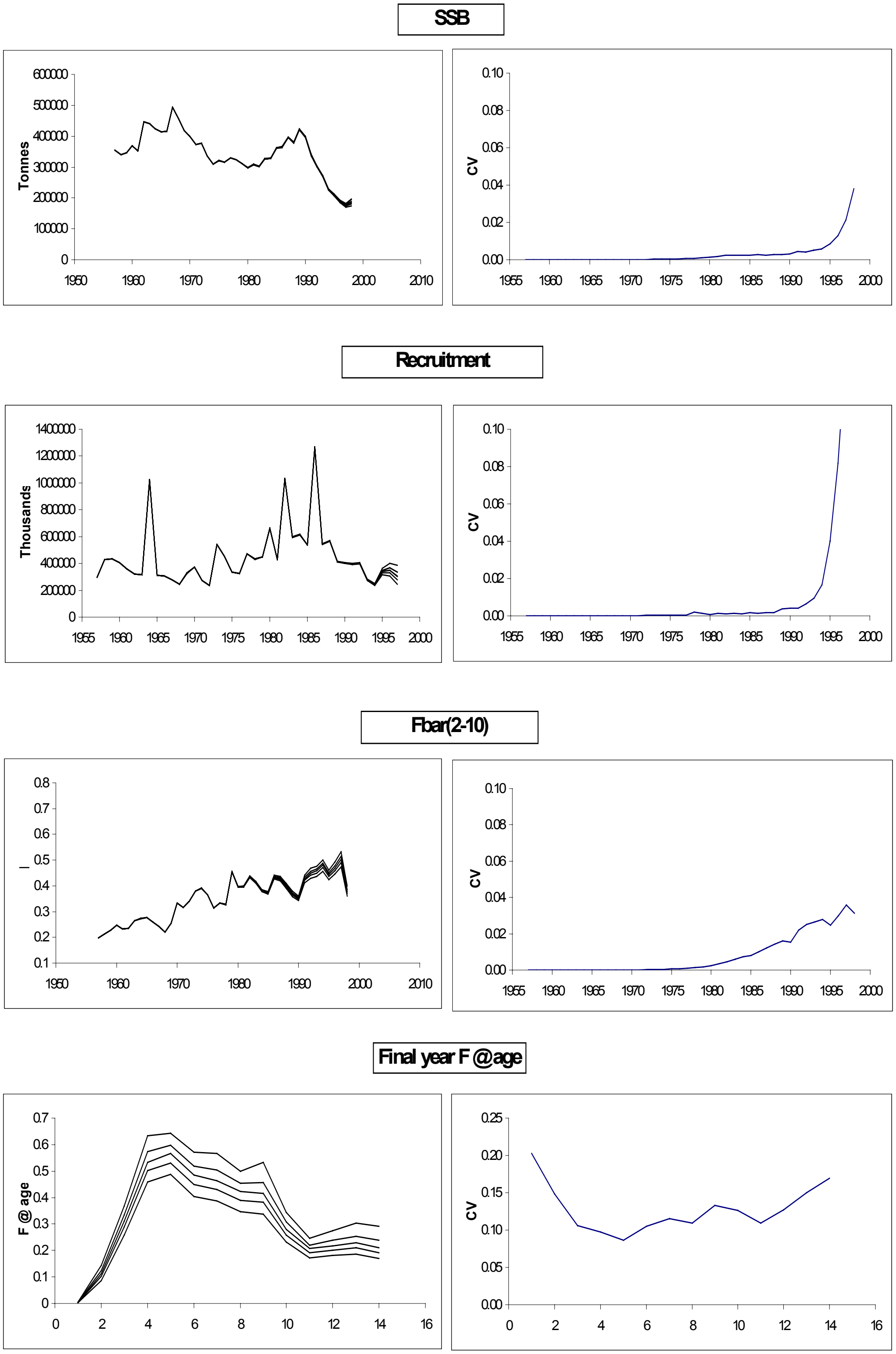


Figure 8. The 5,25,50,75,95th percentiles of Fbar (2-10), recruitment at age 1, SSB and F at age in 1998 resulting from fitting the 1999 ICES WG XSA model to 1000 non-parametric bootstraps samples of the North Sea plaice survey and commercial fleet catch per unit effort data. Catch at age data constant.

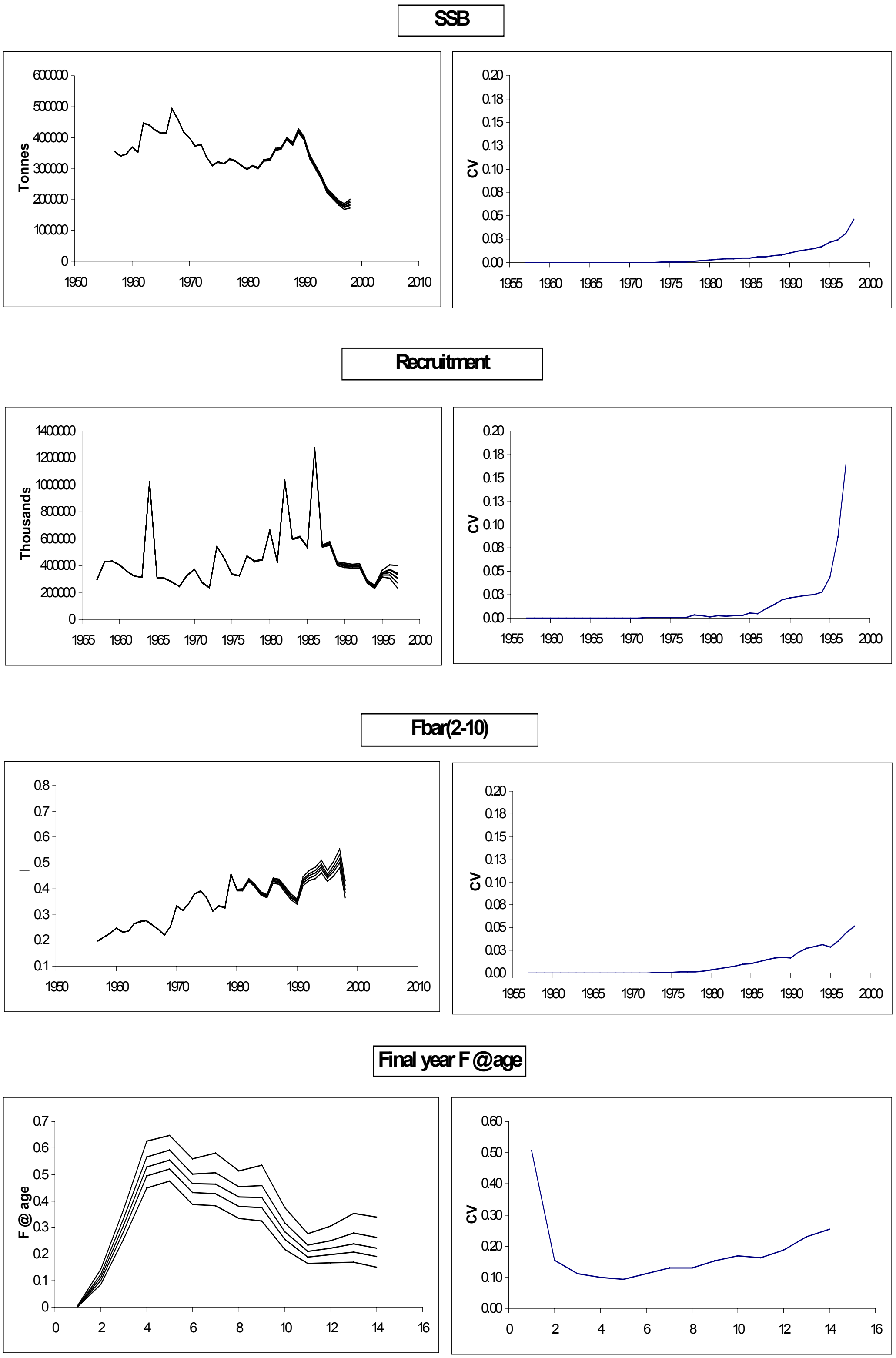


Figure 9. The 5,25,50,75,95th percentiles of Fbar (2-10), recruitment at age 1, SSB and F at age in 1998 estimated by fitting the 1999 ICES WG XSA model to 60 non-parametric bootstraps of each of 1000 assessments derived from the bootstrap samples of the North Sea plaice catch at age and fleet tuning data for the years 1991 - 1998.

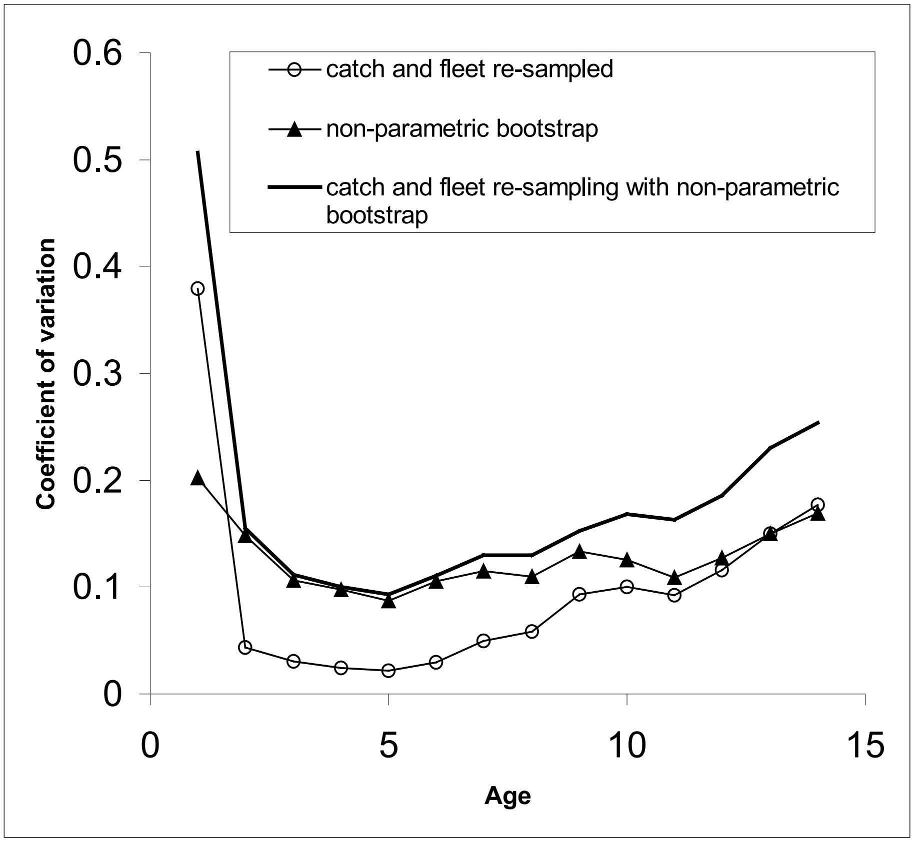


Figure 10 Estimated coefficients of variation of the North Sea plaice fishing mortality at age in the final assessment year derived from three bootstrap procedures.

International Council for the Exploration of the Sea

Quality and Precision of Basic Data Underlying Fish Stock Assessment and Implications for Fisheries Management Advice
CM 2001/P:14

The precision of international market sampling for North Sea cod (*Gadus morhua* L.) and its influence on stock assessment

C.M. O'Brien, C.D. Darby, B.D. Rackham, D.L. Maxwell, H. Degel, S. Flatman, M. Mathewson, M.A. Pastoors, E.J. Simmonds and M. Vinther

ABSTRACT

Market sampling is an essential source of data for age-based finfish stock assessment. Previously, little has been documented about the influence of potential error in these data on the precision of stock assessments and the management information they produce. This paper presents the results of an EU-funded study (CFP Study Project 98/075) of the precision of North Sea cod (*Gadus morhua* L.) fish market sampling carried out by the United Kingdom (England & Wales, Scotland), Denmark and the Netherlands. Data from eight years of market sampling conducted for the period 1991-1998 were analysed to obtain the precision of estimated numbers-at-age in the landed catch. The annual market sample data were then used to generate 1000 realisations of the international numbers-at-age and mean weights-at-age in the landed catch for the eight-year period selected. Matrices of the catch numbers-at-age were computed for the generated realisations and these were used to produce 1000 stock assessments conditional on the XSA (Extended Survivors Analysis) model as implemented in 1999 at the ICES Working Group on the Assessment of Demersal Stocks in the North Sea. From the outcome of these assessments the influence of the market sampling programmes on the perception of the dynamics of the stock are graphically presented as simulation envelopes on the main management parameters.

C.M. O'Brien: CEFAS Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK [tel: +44 (0) 1502 524256, fax: +44 (0) 1502 513865, e-mail: c.m.obrien@cefas.co.uk]. C.D. Darby: CEFAS Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK [tel: +44 (0) 1502 524329, fax: +44 (0) 1502 513865, e-mail: c.d.darby@cefas.co.uk]. B.D. Rackham: CEFAS Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK [tel: +44 (0) 1502 562244, fax: +44 (0) 1502 513865, e-mail: b.d.rackham@cefas.co.uk]. D.L. Maxwell: CEFAS Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK [tel: +44 (0) 1502 562244, fax: +44 (0) 1502 513865, e-mail: d.l.maxwell@cefas.co.uk]. H. Degel: DIFRES, Charlottenlund Slot, 2920 Charlottenlund, Denmark [tel: +45 33 96 3386, fax: +45 33 96 3333, e-mail: hd@dfu.min.dk]. S. Flatman: CEFAS Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK [tel: +44 (0) 1502 562244, fax: +44 (0) 1502 513865, e-mail: s.flatman@cefas.co.uk]. M. Mathewson: FRS Marine Laboratory, P.O. Box 101, Aberdeen AB11 9DB, Scotland [tel: +44 1224 295366, fax:

+44 1224 295511]. M.A. Pastoors: RIVO, P.O. Box 67, 1970 AB IJmuiden, The Netherlands [tel: +31 255 564690, fax: +31 255 564 644, e-mail:m.a.pastoors@rivo.wag-ur.nl]. E.J. Simmonds: FRS Marine Laboratory, P.O. Box 101, Aberdeen AB11 9DB, Scotland [tel: +44 1224 295366, fax: +44 1224 295511, e-mail: simmondsej@marlab.ac.uk]. M. Vinther: DIFRES, Charlottenlund Slot, 2920 Charlottenlund, Denmark [tel: +45 33 96 3300, fax: +45 33 96 3333, e-mail: mv@dfu.min.dk].

Keywords: bootstrap, mean-variance modelling, mortality, non-parametric, numbers-at-age, weights-at-age

INTRODUCTION

The catch-at-age matrix forms a major component of the assessment of many fish stocks. For stocks with international division of the catch, the sampling schemes are often diverse and the task of assembling the catch-at-age matrix is a complex and time-consuming process. The influence of this data on the assessment is rather poorly studied. To the knowledge of the authors, there are no comprehensive studies of the precision of international market sampling programmes and their implications for fisheries management advice. The papers published on this issue either deal either with the potential effects of theoretical uncertainty in basic data on the advice provided (Pope and Gray 1983; Pelletier and Gros 1994; Coggins and Quinn 1998), or on the estimation techniques and results of the analysis of uncertainty in the basic data itself (Tanaka 1953; Kimura 1977; Sparre et al. 1977; Gavaris and Gavaris 1983; Smith and Maguire 1983; ICES 1994; Reinert and Lewy 1998). In this paper, we present an attempt to combine these two approaches. This paper describes the analysis of catch-at-age data on North Sea cod from 1991 to 1998 inclusive, and their use in an age-based stock assessment.

To combine results from different sampling programmes to arrive at total international estimates of catch numbers-at-age and their associated variances, we could have followed two routes.

> Route 1: Attempt to combine the raw sampling data, calculate appropriate age-length keys (ALK) and raise the sampling data to the total international landings. In this way the variances of the procedure could be directly calculated (Gavaris and Gavaris 1983; Smith and Maguire 1983) or obtained from bootstrap analysis. A pre-requisite for this approach is that the sampling procedures (strata) are harmonised so that samples can be freely exchanged. This harmonisation is difficult to obtain from data already collected independently by different countries with different sampling and data storage methods.

Route 2: Use bootstrap techniques to generate a certain number of realisations of national age compositions and weights-at-age. Then combine these national realisations as a stock assessment working group would have done, delivering a number of realisations of the international age composition. These are then input into a stock assessment program to arrive at bootstrapped stock estimates.

It is this later approach that has been followed in this study.

We present first the results from studies of national market sampling programmes for estimating the catch numbers-at-age of North Sea cod for the period 1991 to 1998, inclusive. Market sample data from the major fishing countries have been collated at the minimum aggregation level and used to generate 1000 national and then international replicates for use in bootstrapped assessments. The assessment procedure was that used at the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak ICES(2000).

MATERIAL AND METHODS

National data from the Netherlands, Denmark, the United Kingdom - England and Wales [subsequently denoted UK(E&W)] and the United Kingdom – Scotland [subsequenly denoted UK(Scotland)], were used to provide 1000 replicates of national market sampling length, weight- and catch numbers-at-age. The sampling and raising methods are different for each nation, and consequently the methods for deriving the 1000 replicates were also different. These are described by nation below.

Bootstrapping the UK(E&W) data

Bootstrapping the catch-at age data was carried out at the vessel level using code to replicate the raising calculations carried out by the market sampling system as closely as possible. The market samples were stratified by quarter, fleet and area. To give enough samples for resampling the fleets were beam trawl, gillnet, long lines and 'bottom trawl' defined as a combination of seine, otter trawl and pair trawl gears. For cod, two areas were defined: North Sea roundfish areas 1,2,3,7 and North Sea roundfish areas 4,5,6 (Figure 1). The original data were extracted from the database that holds length and biological sample data, along with combined and raised processed data. Firstly, using the sample number and vessel codes in the database, two lists were formed: the boat-trips from which age samples and length samples had been taken. Each list was sampled with replacement to form a new list – the *bootstrap*. The *bootstrap* length and age samples were then comprised of the data from the boat-trips included in the new lists. Catch-at-age and weight-at-age estimates were calculated from the bootstrap length and age samples. This bootstrap procedure was repeated 1000 times for each period of interest.

The algorithm used for the 1000 bootstraps was as follows:

Set-up:

- read in original length, age and weight data

- create a list of unique identifiers for sampling units – here used vessel code and sample number
- calculate values that will not change with each bootstrap sample – commercial weight totals and numbers of samples

Bootstrap loop that is repeated for 1000 iterations:

- set seed for random number generator
- form a bootstrap length sample by re-sampling length data
- calculate length distributions (LD) and analytical variance due to length sampling for bootstrap length sample using appropriate stratification and length groups
- set seed for random number generator
- form bootstrap age sample by resampling age data
- calculate age-length key (ALK) and analytical variance due to ageing for the bootstrap age sample using appropriate stratification and length groups
- calculate age-length distribution (ALD) and analytical variance from LD and ALK
- calculate numbers-at-age using the LD and ALK
- calculate mean length within each length group and parameter for length-weight relationship
- calculate mean weight-at-age from ALD and length-weight relationship
- append the estimates from this iteration to the output file

Bootstrapping the Netherlands data
The bootstrap analysis of the Dutch data followed the same approach as UK(E&W), with the raising procedure adjusted to be specific to the Dutch case (Anon., 2001). Bootstrapping the catch-at-age data was carried out at the vessel level using code to replicate the raising calculations, with raising stratified by quarter, gear (1991-1993 only) and market category.

Bootstrapping the Danish data
The Danish raising procedure for cod is stratified by quarter of the year, landing region and market size class. Approximately 50% of the strata used have just one sample, which makes bootstrapping of just samples pointless. Therefore, the basic bootstrapping approach was extended by bootstrapping of individual fish within a resampled market sample.

Jack-knifing Scottish data
Market samples are aggregated to monthly-based region and gear length distributions with age length keys. These data are collected from multiple samples, however as data are combined before entry into the database, it is no longer possible to separate the individual samples. Approximately 560 datacells are used for North Sea stocks. In the resampling algorithm (simple jack-knife, Efron and Tibshirani 1993; ICES 2000b) these monthly and gear categories were resampled and combined to give a number of samples by region and quarter subsequently used in the raising procedure.

The following initial set-up was carried out:

1. Obtain total catch for commercial catch per data cell L
2. Obtain a length frequency distribution per data cell LF
3. Obtain an age length key for those cells for which it was available ALK
4. Obtain a list of links between length keys to age length keys, for all data cells, using fill-in rules as required
5. Select data cells for removal randomly without replacement; for simple jackknife select with equal probability;
6. Create new data set with selected samples removed
7. Find new fill-ins for data cells without length or age length keys
8. Calculate the mean weight W of fish for each data cell using standard monthly, region dependent length weight relationship and length frequency LF.
9. Calculate the total number of fish N for each data cell from the total catch L and the mean weight W
10. Calculate the number at age N_a for each data cell using the total N, the length frequency LF and the age length key ALK
11. Calculate the mean weight at age for each data cell using the length frequency LF, weight at length W_l and the age length key ALK
12. Calculate the total numbers at age by summing the numbers at age per data cell
13. Calculate the total biomass at age by summing the numbers * mean weight at age per data cell

Following 1000 replications:

1. Check that 1000 values have similar mean to original data
2. Calculate CV from mean and variance of 1000 replicates.
3. Correct the jack-knife estimates of CV for number of data cells and removed samples (Efron and Tibshirani, 1993)
4. For jack-knife inflate catch number at age of each replicate by scaling the replicates about the mean and setting the small number of negative observations (<1%) to zero.

The sampling procedure attempts to estimate data cells organised by month, area and gear. As the sampling is only partial, inevitably it is not possible to fill all of the cells where landings are reported in the year. In some cases no data is available at all, in others length keys only are available and age length key data must be supplied. The current method used is to assign length or more usually age/length key data from another cell. This process is in effect a step-wise spatial temporal based model, estimated by nearest neighbour method. The nearest neighbour is selected with a sequence of assignments from previous or following month based on the most similar areas. The presence of data in the same area gear cell is checked in previous and subsequent months, then in sequence (1st, 2nd,3rd etc.), until an alternative is found (Table 1 and Figure 2).

Internationally combined market data

The 1000 bootstrap replicates of mean weight and catch at age UK(Scotland), Netherlands, Denmark and UK(E&W) were combined into 1000 replicates of international catch data. This fully sampled component constitutes around 70% of the North Sea cod landings as given in ICES (2000) - refer to Table 2 in this manuscript. The major missing components compared to the Working Group are data from France, Germany and Belgium. In the assessments generated from the bootstrap data, the catch numbers at age were thenn scaled to the Working Group catch numbers-at-age in each year.

Calculating summary statistics

Once a 1000 catch numbers-at-age and weight-at-age replicates had been produced, mean values, variances, coefficients of variation (CVs) and correlations were calculated. The underlying relationship between mean and variance for catch numbers-at-age was investigated by fitting a linear regression of log(variance) on log(mean).

Assessment of North Sea cod

The Extended Survivors Analysis (XSA) algorithm (Darby and Flatman 1994; Shepherd 1999) was modified to enable repeated fits of the model following replacement of the catch numbers-at-age and tuning fleet data for a user-specified range of years and ages. The estimates of the parameters of interest; namely, the recruitment, spawning stock biomass (SSB) and average fishing mortality calculated over a user-defined age range, were output during each iteration.

The XSA model was specified with the catchability and shrinkage constraints described in the report of the ICES Working Group on the Assessment of Demersal Stocks in the North Sea (ICES 2000a). Catchability was fitted as proportional to population size at age 1, independent of population abundance for all older ages; the catchability at each age greater than 6 was constrained to be equal to that estimated at age 6. The terminal populations were shrunk to the arithmetic mean of the fishing mortality estimated for the penultimate 5 oldest ages and the years 1993 – 1997. The coefficient of variation of the means used in the shrinkage was set at 0.5 and the minimum permitted value for the standard error of log-catchability set at 0.3. The assessment was applied to the catch at age data for the years 1963 – 1998 as recorded by the ICES Working Group (ICES 2000a). The catch per unit effort (CPUE) data for the tuning series was also extracted from that source. Five commercial and three survey CPUE series for the years 1983 – 1998 were used, no time series taper weighting was applied.

Two bootstrap assessments were run using the 1000 data sets from the bootstrap sampling of the international catch at age and commercial tuning fleet data. These were an assessment in which the catch at age data for the years 1991 – 1998 was replaced with a new sample at each iteration (series 1), and secondly replacement of the catch at age data and four of the five commercial fleet tuning data series, over the same time period, during each bootstrap iteration (series 2).

VPA based assessment models make the assumption that the dominant portion of the uncertainty in the estimated population abundance and exploitation rate results from the process and measurement errors associated with the CPUE tuning series; the catch at age data is assumed to be exact. In order to compare the potential magnitude of the errors generated from the tuning data series, a second non-parametric (NPB) bootstrap algorithm of the XSA model was developed. A base XSA model was fitted to the Working Group catch and tuning data sets. The base model log catchability residuals were sampled with replacement (independently by series and age) and used with the base model catchability parameters and VPA population abundance, to calculate new CPUE tuning data for all of the series. The bootstrapped CPUE values were then used with the base model catchability parameters to derive terminal population values for initiating a bootstrapped VPA. The weights assigned to terminal population estimates, calculated during the bootstrap process, were those estimated within the base XSA model. This non-parametric bootstrap algorithm assumes independence of the residuals by series and age and of the two commercial fleets' CPUE data from the catch at age matrix.

The non-parametric bootstrap XSA model was used to generate 1000 replicates of the Working Group XSA model for North Sea cod in order to examine the uncertainty associated with the tuning series information (series 3).

The final bootstrap model incorporated all of the bootstrap uncertainties, resulting from the sampling process and introduced by the tuning series. The bootstrap combined the algorithms described previously and performed 60 non-parametric bootstraps of the tuning series residual matrix for each of the 1000 catch and commercial CPUE data samples (series 4).

RESULTS

The bootstrapped catch numbers-at-age and weight-at-age of North Sea cod from combining UK(Scotland), Dutch, Danish and UK(E&W) estimates are shown plotted with WG estimates in Figures 3 and 4, respectively. In all years the WG estimate of catch number at age is on the upper side of the bootstrap distribution as the WG estimates include additional age compositions from France, Germany and Belgium. Weights at age estimated from the bootstrap realisations seem to be well in line with the WG estimates.

Catch numbers-at-age distributions

Histograms of the bootstrap estimates of catch numbers-at-age, scaled to the total international landings, are provided in Figures 5 and 6 for two arbitrarily chosen example years (1991 and 1998**.** Superimposed on each histogram is a Gaussian kernel density estimate (solid line) (Sheather and Jones, 1991, MathSoft, 1998) and a normal distribution with the same mean and variance as the data (dashed line). These allow a visual inspection of how well normal distributions fit the data.

For age 1 cod, the mode of the density estimate is slightly greater than for the normal distribution and for the oldest ages the data are slightly skewed in the opposite direction.

The last histogram in 1998 reflects that few vessels caught fish aged 11+. Overall, the normal distribution appears to be reasonable description of the data. This may be a result of the many stages of combining data involved in producing the international estimates.

Uncertainty and precision

The CV of catch numbers at age for the combined and national data sets are presented in Tables 3 to 7. The CVs appear to be consistent across years and are similar for the three countries studied. CV of the national and international catch numbers follow the same pattern, with relatively higher CV on the very young and older age groups. As expected, the international CVs are lower than the national CVs. The international CVs of the most fished age groups are less than 5%.

The CV of the combined mean weight at age were generally less than 5% for most age groups and about 2% for the dominant age groups, further details are in the final report of CFP project 98/075 (Anon. 2001).

The underlying correlation of catch numbers-at-age was estimated using the numbers-at-age obtained from the resampling of the market sampling data. The patterns of positive and negative correlation were similar across the years so the mean correlation coefficients between estimates of catch numbers-at-age are given in the Table 8. There is some negative correlation amongst ages 1 to 6 and values are close to zero in the rest of the matrix.

Table 9 gives the coefficients by year from fitting a linear regression of log variance against log mean for international catch numbers at age. The coefficients are consistent across years with a slope of 1.4 for all years combined. Figure 7 shows the fit with all years combined.

Evaluation of uncertainty in the stock assessment and management parameters

The time series of estimates of recruitment, SSB and average fishing mortality for ages 2 – 8 derived from the assessment in which the catch at age data alone were bootstrapped (series 1) are presented in Figure 8. The results are consistent with those of the assessment carried out at the ICES Working Group; the Working Group estimates for each parameter lie on the median of the percentile distributions.

Re-sampling of the catch at age data for the cohorts present in the years 1991 – 1998 has resulted in relatively minor uncertainty in the estimates of average F, recruitment and SSB series; all CVs are less than 5%. The coefficients of variation of fishing mortality estimated in the final year are larger than those of the combined metrics, especially at ages greater than 6 where the CV exceed 10%. The higher CVs are not unexpected, fishing mortality in the final year is a function of the ratio of two bootstrap replicates from the cohort, whereas SSB, recruitment and average F are derived from a weighted combination of the transformed replicates. Assessment models that are based on an underlying population structure reconstructed by VPA make the assumption that the catch at age data are exact or, at least, that the effects of measurement errors in the catch at age data can be usually ignored. The high CV values illustrated in Figure 8 indicate that, for the relatively well-sampled ages of the North Sea cod this assumption would hold, but at the older less abundant ages significant levels of the uncertainty in estimated parameter values is being omitted from the assessment.

Figure 9 illustrates the results of re-sampling the catch at age data and the commercial fleet CPUE data for the years 1991 – 1998. Bootstrap re-sampling of the commercial fleet CPUE has increased the uncertainty associated with each of the estimated stock metrics. Uncertainty as measured by the CV decreases historically due to the convergence properties of the VPA equations and the lack of bootstrap re-sampling prior to 1981. The results indicate increased uncertainty in the fishing mortality at the oldest ages of the assessment where sampling error increases due to reduced catch numbers. However, the magnitude of the increase in the CV associated with the re-sampling of the commercial fleet data is substantially less than that linked to re-sampling the catch at age data indicating strong correlation in the bootstrapped data values.

The time series of percentiles of parameter values derived from 1000 non-parametric bootstrap assessments of the tuning series residuals with the catch numbers and tuning series catchability constant are illustrated in Figure 10. The bootstrap procedure re-constructs the CPUE series used to estimate the terminal population numbers for each cohort. Therefore, the main region of uncertainty is confined to the most recent years of the assessment time series and the oldest ages. Convergence of the VPA, conditional on the constant catch at age matrix, produces the convergence of the percentiles of the historic estimates. At the youngest ages the uncertainty in the estimates of fishing mortality in the final year is generally higher than the equivalent estimates for the re-sample bootstraps. It is 50% lower at the oldest ages.

The results for the fourth comparative series in which the two assessment bootstrap processes were combined are presented in Figure 11. It is seen that the two components of uncertainty have been approximately additive. The historical variation induced by the re-sampling of catch at age and CPUE data has been added to the uncertainty in the final year and oldest age population and exploitation rate estimates associated with the non-parametric bootstrapping of the tuning residuals. At ages greater than age 6 there is substantial uncertainty in the estimates of fishing mortality in the final year. The uncertainty is contributed equally by sampling error and by the errors associated with the tuning process.

DISCUSSION

The international sampling programmes appear to be delivering estimates of catch at age that are rather precise, with CV's of 2.5% for cod for the best estimated ages rising to about 40% at the older ages. While the precision of the best estimated ages is good, the current scheme is delivering poorer CVs on older ages. Care must be taken to ensure that the importance of estimating both old and young year classes is fully understood.

Based on the analysis of the histograms of the numbers at age, the normal distribution appears to be reasonable description of the catch at age data. However, this may be a result of the many stages of combining data involved in producing the international estimates and assumes independence among the national programmes.

The relationship between the mean and variance of the numbers-at-age is fundamental to any future statistical modelling of catch numbers-at-age; as is the assumption of independence between numbers-at-age. The underlying relationship between mean-variance of catch numbers-at-age was investigated by considering the mean and variance of the numbers-at-age obtained from the resampling of the market sampling data and compared to the power relationship:

$$\text{variance}\{\text{bootstrapped numbers-at-age}\} = e^{a} \, . \, \text{mean}\{\text{bootstrapped numbers-at-age}\}^{b}$$

Relationships between mean and variance are observed with slopes on the log variance-mean relationships of 1.37 (s.e. 0.02). Assessment models generally do not take this into account; changes to models or to weighting practices that would include these mean-variance relationships would be helpful. The apparent proportionality for the variance-mean relationship will facilitate the development of appropriate statistical models of catch-at-age that do not assume a log-normal distribution for catch-at-age.

These studies suggest that for the data sets examined the current levels of market sampling cause only small amounts of variability in assessment outputs for North Sea cod, for the ages that are predominant in the catch data. The highest CVs, 20 - 30%, were estimated at the oldest ages and resulted primarily from sample noise in the catch numbers-at-age data and this result would suggest that these ages should be combined into a plus-group.

As would be expected from a VPA method that assumes exact catch numbers-at-age, variance in the sampled catch numbers-at-age were transferred directly into the fishing mortality estimated for the final assessment year (c.f. 1998 in Table 3 and Figure 8). The introduction of sampling error to the commercial fleet CPUE series inflated the uncertainty in the management parameters but the increase in variance was less than that resulting from the catch numbers-at-age data. Again, this should be expected given that the commercial fleet CPUE data is a constituent part of the catch numbers-at-age data set and the effects of changes in abundance and sample error would be highly correlated. This raises the question as to the statistical validity in using commercial fleet tuning data twice when the expected correlations are high, firstly as part of the catch numbers-at-age and secondly as part of the tuning data.

On average the non-parametric bootstrap estimates of the uncertainty of the management parameters are higher than those derived from the re-sampling bootstrap. This results from the lower level of sampling associated with the smaller data sets used to obtain the CPUE data for tuning series. The CVs exhibit similar trends to the re-sampling bootstrap estimates, high values at the youngest and oldest ages. As would be expected, combining the two bootstraps results in the highest estimates of uncertainty for the re-sampled time period. Figure 12 plots the CV of the terminal year fishing mortalities at age from the catch and fleet re-sampling bootstrap model, the non-parametric bootstrap model and the combined model. The figure shows that the errors are not additive and the combined method is required to allow for correlation.

ACKNOWLEDGEMENTS

This paper was prepared with funding support provided by the Department for Food, Environment and Rural Affairs (under contract MF0316) to the first four authors (CMO'B, CDD, BDR, DLM); together with funding support provided to all authors by the Commission of the European Communities Directorate General for Fisheries (DG XIV) under CFP Study Project contract 98/075 (Evaluation of market sampling strategies for a number of commercially exploited stocks in the North Sea and development of procedures for consistent data storage and retrieval (software) - EMAS).

REFERENCES

Anonymous (2001). Final report of CFP Study Project 98/075, Evaluation of Market Sampling strategies for a number of commercially exploited stocks in the North Sea and development of procedures for consistent data storage and retrieva (software) - EMAS.

Coggins, L. G. and T. J. Quinn (1998). A simulation study of the effects of aging error and sample size on sustained yield estimates. In Fishery stock assessment models. F. Funk, T. J. Quinn, J. Heifetz et al, pp.955-975.

Darby, C.D. and Flatman, S. (1994). Virtual Population Analysis: Version 3.1 (Windows/DOS) User Guide. Information Technology Series, MAFF Directorate of Fisheries Research, Lowestoft, (1):85pp.

Gavaris, S. and C. A. Gavaris (1983). Estimation of catch at age and its variance for groundfish stocks in the Newfoundland region. Sampling commercial catches of marine fish and invertebrates. W. G. Doubleday and D. Rivard, Canadian Special Publication of Fisheries and Aquatic Sciences. **66:** 178-182.

ICES (1994). Report of the Workshop on Sampling Strategies for Age and Maturity. ICES Headquarters, Copenhagen, 3-9 February 1994. ICES C.M. 1994/D:1 ref: G,H,J.

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

ICES (2000b). Report of the Study Group on Market Sampling Methodology. Marine Laboratory, Aberdeen, 24-25 January 2000. ICES C.M. 2000/D:01.

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

Kimura, D. K. (1977). Statistical assessment of age-length-key. Journal of the Fisheries Research Board of Canada **34**: 317-324.

MathSoft. (1998). S-PLUS for windows, version 4.5. MathSoft, Inc., Seattle.

Pelletier, D. and P. Gros (1994). Assessing the impact of sampling errors on model-based management advice: an application to the Northwest Atlantic. Canadian Journal of Fisheries and Aquatic Sciences **48**: 2129-2139.

Pope, J. G. and D. Gray (1983). An investigation of the relationship between precision of assessment data and the precision of Total Allowable Catches. Sampling commercial catches of marine fish and invertebrates. W. G. Doubleday and D. Rivard, Canadian Special Publication of Fisheries and Aquatic Sciences. **66:** 151-157.

Reinert, R. and P. Lewy (1998). Sampling errors in estimated catch-at-age data in the Danish human consumption sampling program for the North Sea. Unpublished manuscript.

Sheather, S. J. and M. C. Jones (1991). A reliable data-based bandwidth selection method for kernel density estimation. Journal of the Royal Statistical Society series B **53**: 683-690.

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

Smith, S. J. and J. J. Maguire (1983). Estimating the variance of length composition samples. Sampling commercial catches of marine fish and invertebrates. W. G. Doubleday and D. Rivard, Canadian Special Publication of Fisheries and Aquatic Sciences **66:** 165-170.

Sparre, P., H. Knudsen and S. Munch-Petersen (1977). Optimization of sampling programs for estimation of age distribution and fishing mortality.

Tanaka, S. (1953). Precision of age-composition of fish estimated by double sampling method using the length for stratification. Bulletin of the Japanese Society of Scientific Fisheries **19**: 657-670.

Table 1. UK(Scotland) - Neighbour sequences for cod areas and gears (see Figure 2).

Cod				
Area	Area Neighbour for gear 1-5	Area Neighbour for Gear 20	Gear	Gear Neighbour
1	4, 2	4	1	-
2	5, 6	-	2	3, 4
3	4, 8	4	3	2, 4
4	3, 8	-	4	3, 2
5	6, 2	-	5	3
6	5, 7	2, 5	20	3
7	6, 5	2, 5, 6		
8	4, 3	4		
9	1, 3	3		
10	6	2, 5, 6		
11	6	-		
12	10	2, 5, 6		
13	2			

Table 2. Comparison between mean bootstrap SOP and total landings estimated by the WGNSSK (ICES 2001).

Year	SOP bootstrapped countries	WG landings	%bootstrapped
1991	78879	102478	77%
1992	80334	114020	70%
1993	83670	121749	69%
1994	78932	110634	71%
1995	100500	136096	74%
1996	98064	126320	78%
1997	89736	124158	72%
1998	102282	146014	70%

Table 3. CV (%) of estimated catch numbers-at-age, combined data.

	1991	1992	1993	1994	1995	1996	1997	1998	Average
1	7.7	4.9	7.0	6.3	5.7	8.7	5.7	13.8	7.5
2	3.0	2.3	1.8	3.1	2.1	3.0	2.8	2.0	2.5
3	2.8	3.0	3.5	2.5	4.2	2.4	2.3	3.7	3.0
4	4.3	3.5	4.8	4.5	4.0	4.7	3.2	4.9	4.2
5	5.3	5.8	5.7	6.2	6.7	5.1	5.2	5.9	5.7
6	5.3	7.2	11.7	7.8	8.9	8.5	5.6	8.8	8.0
7	10.9	7.0	15.2	15.1	12.5	12.5	11.0	10.6	11.9
8	9.2	16.7	14.1	22.0	27.8	19.3	20.6	22.8	19.1
9	21.7	17.7	32.3	33.3	40.2	26.8	22.0	36.9	28.9
10	28.0	32.8	36.6	54.5	30.8	28.6	27.4	28.8	33.4
11+	37.5	33.5	58.2	56.8	22.1	43.0	53.5	62.5	45.9

Table 4. UK(England & Wales).
CV of numbers at age

country	EW
species	COD
area	IV

Average of cv_num	year								
age	1991	1992	1993	1994	1995	1996	1997	1998	Grand Total
1	8	17	8	12	9	10	9	18	11
2	6	6	4	5	4	4	4	2	4
3	4	4	5	4	5	4	4	4	4
4	5	4	4	7	5	5	4	4	5
5	6	8	6	9	9	8	5	7	7
6	6	10	13	12	12	15	9	11	11
7	14	12	22	18	14	18	17	16	17
8	13	20	24	29	18	23	26	25	23
9	30	26	27	45	41	34	38	34	34
10	43	42	48	89	31	53	69	51	53
11	40	55	75	47	39	54	36	47	52

Table 5. Netherlands.
CV of numbers at age

country	NL
species	COD
area	IV

Average of cv_num	year								
age	1991	1992	1993	1994	1995	1996	1997	1998	Grand Total
0	105		158			111		67	110
1	5	2	12	8	6	12	8	122	22
2	6	4	2	15	4	4	9	3	6
3	8	9	8	10	18	10	7	17	11
4	6	7	10	19	8	17	14	16	12
5	9	11	9	17	23	9	20	16	14
6	10	12	16	18	24	27	15	41	20
7	50	16	28	30	24	31	33	20	29
8	71		22	34	48	39	32	50	42
9		59	53	67	51	72	58	55	59
10		49	49	111	66	59	52	60	64

Table 6. Denmark.

CV of numbers at age

country	DK
species	COD
area	IV

Average of cv_num	year								
age	1991	1992	1993	1994	1995	1996	1997	1998	Grand Total
1	6	7	34	14	9	30	29	60	24
2	3	3	4	6	2	6	3	7	4
3	4	7	7	5	6	4	5	11	6
4	9	7	12	10	8	11	8	13	10
5	11	13	13	11	13	10	13	13	12
6	10	18	24	15	15	14	12	17	16
7	20	15	30	31	23	21	22	30	24
8	23	44	33	48	39	31	61	60	42
9	40	67	61	74	37	38	47	56	52
10	34	58	58	54		48	50		50
11	65	48	76	54		50	48	52	56

Table 7. Scotland (simple jack-knife analysis).

CV of numbers at age

country	SC
species	COD
area	IV

Average of cv_num	year								
age	1991	1992	1993	1994	1995	1996	1997	1998	Grand Total
1	14	14	18	16	17	22	13	20	17
2	5	4	3	7	4	6	6	4	5
3	5	6	7	4	6	4	4	6	5
4	11	6	9	5	7	6	5	7	7
5	10	11	10	8	9	9	8	9	9
6	9	11	15	10	13	13	8	9	11
7	14	13	22	17	21	23	13	12	17
8	15	18	17	22	25	27	24	23	21
9	24	17	25	18	25	18	23	22	21
10	22	19	28	27	27	29	17	26	24
11	20	24	24	27	26	23	32	30	26

Table 8. Mean correlation coefficient for catch at age from 1991-1998, for combined Danish, E&W, Dutch and Scottish bootstrapped estimates.

	1	2	3	4	5	6	7	8	9	10	11+
1	1.00	-0.02	-0.23	-0.17	-0.10	-0.06	-0.04	-0.03	-0.03	-0.01	0.00
2		1.00	-0.22	-0.27	-0.19	-0.09	-0.11	-0.06	0.00	-0.02	-0.01
3			1.00	-0.09	-0.11	-0.06	-0.01	-0.02	-0.02	0.00	0.01
4				1.00	0.03	-0.05	0.01	0.02	-0.01	-0.02	0.00
5					1.00	-0.01	0.00	-0.01	-0.01	0.01	-0.01
6						1.00	-0.06	-0.05	-0.03	0.02	-0.01
7							1.00	0.01	0.02	0.00	-0.04
8								1.00	0.02	0.06	-0.02
9									1.00	0.02	0.02
10										1.00	0.01
11+											1.00

Table 9. Intercept (a) and slope (b) estimates from a linear regression of log(Variance) on log(Mean) for combined Danish, UK(E&W), Dutch and Scottish bootstrapped estimates (ages 1-10)

Year	a	b	st error a	st error b
1991	-1.70	1.46	0.57	0.08
1992	-1.33	1.39	0.50	0.07
1993	-0.28	1.28	0.36	0.05
1994	-0.68	1.34	0.44	0.07
1995	-0.76	1.36	0.40	0.06
1996	-1.17	1.41	0.45	0.07
1997	-1.40	1.40	0.42	0.06
1998	-0.88	1.39	0.40	0.06
All years	-1.00	1.37	0.15	0.02

Figure 1. UK(E & W) - sampling areas for North Sea cod.

NORTH SEA (ICES DIVISION IV) ROUNDFISH SAMPLING AREAS

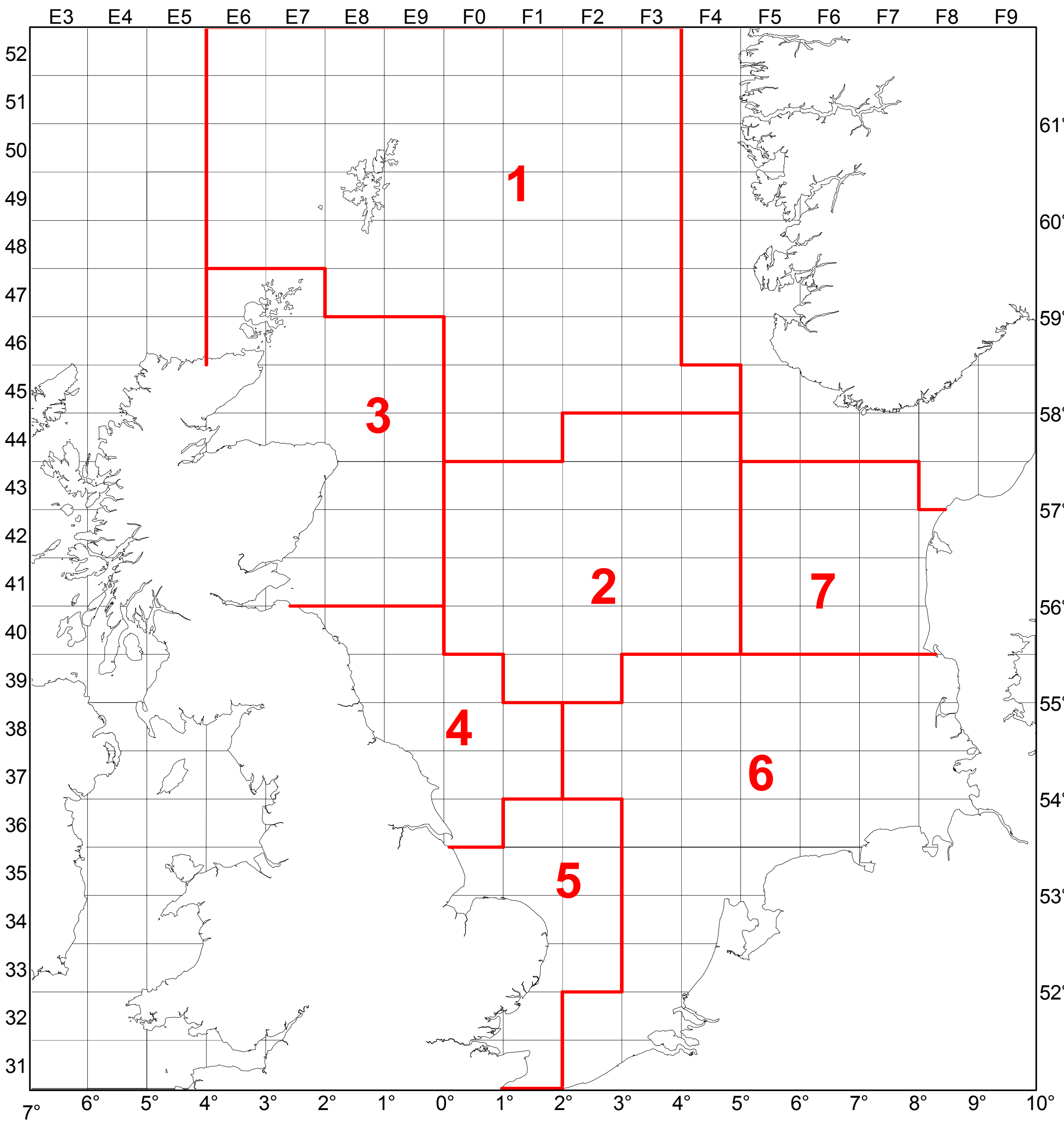
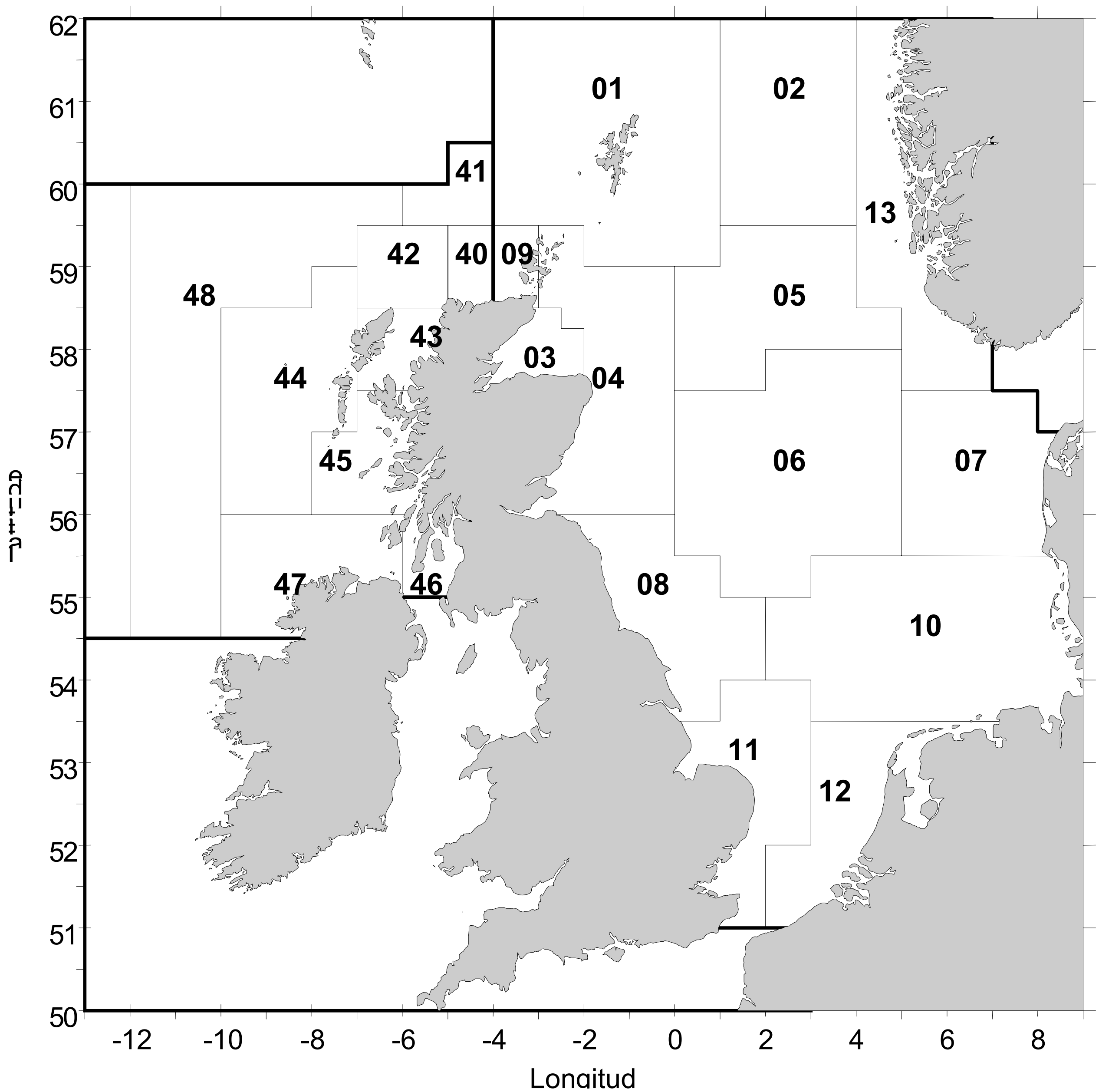


Figure 2. UK(Scotland) - sampling areas for market sampling in the North Sea and West of Scotland.



01
02
13
41
42
40
09
48
05
43
03
44
04
45
06
07
47
46
08
10
11
12
62
61
60
59
58
57
56
55
54
53
52
51
50
-12
-10
-8
-6
-4
-2
0
2
4
6
8
Longitud

Figure 3. Estimated catch numbers-at-age for 1991 to 1998, showing WG catch (line) and 1000 bootstrap estimates (points) from combination of UK(Scotland), Danish, Dutch and UK(E&W) bootstrap estimates.

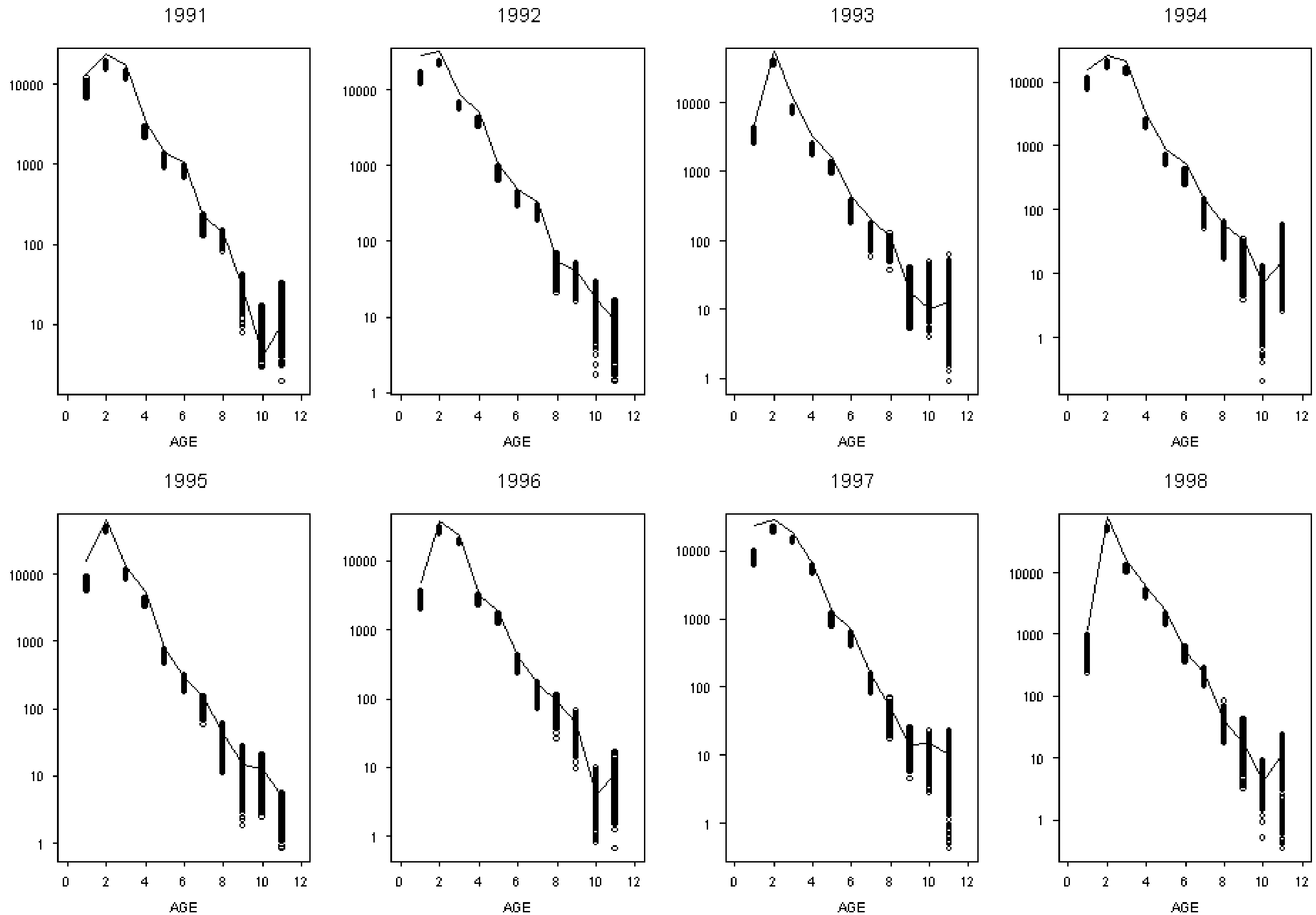


Figure 4. Estimated mean catch weights-at-age for 1991 to 1998, showing WG weights (line) and 1000 bootstrap estimates (points) from combination of UK(Scotland), Danish, Dutch and UK(E&W) bootstrap estimates.

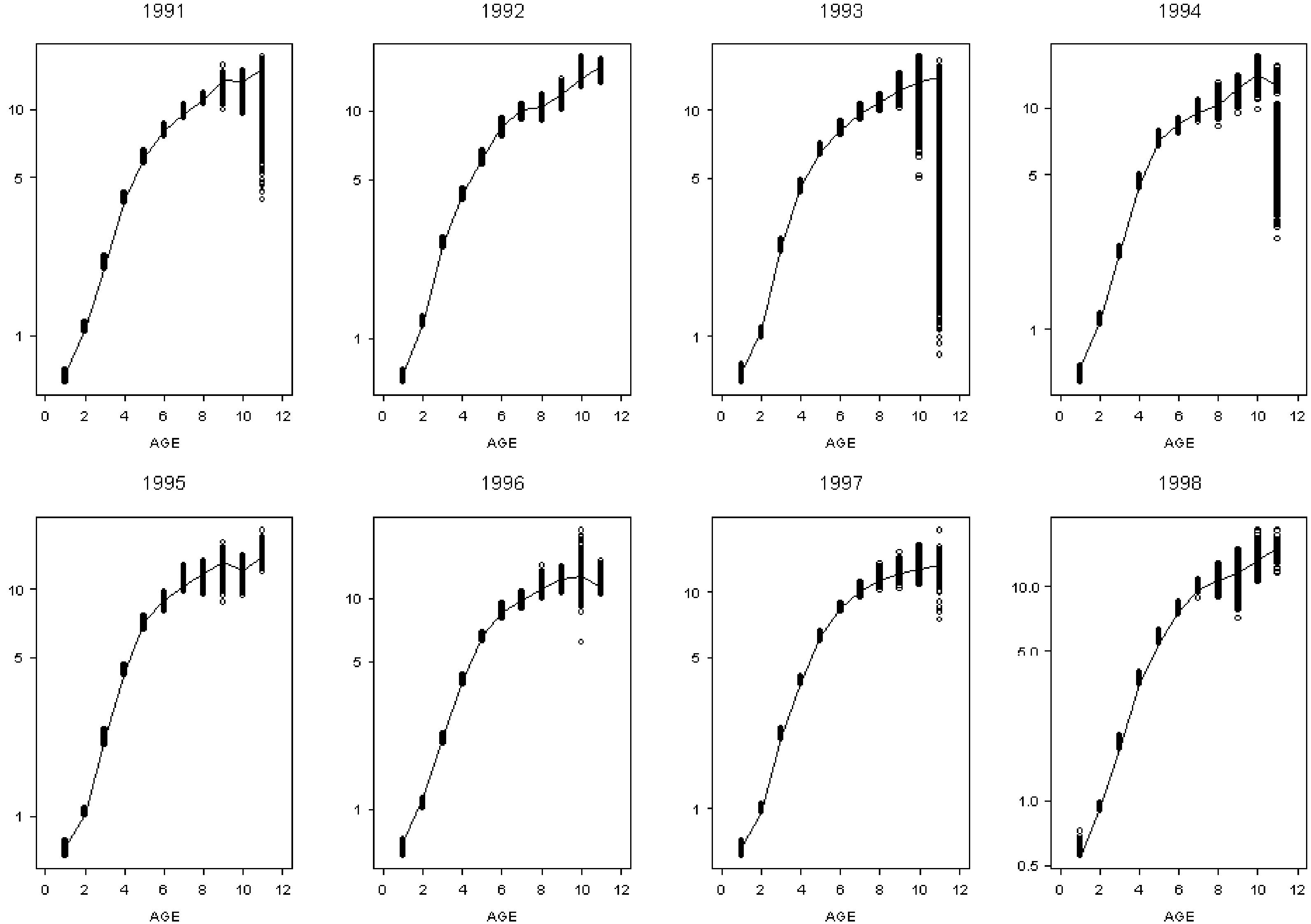


Figure 5. Histograms of bootstrap estimates of raised international catch numbers-at-age.

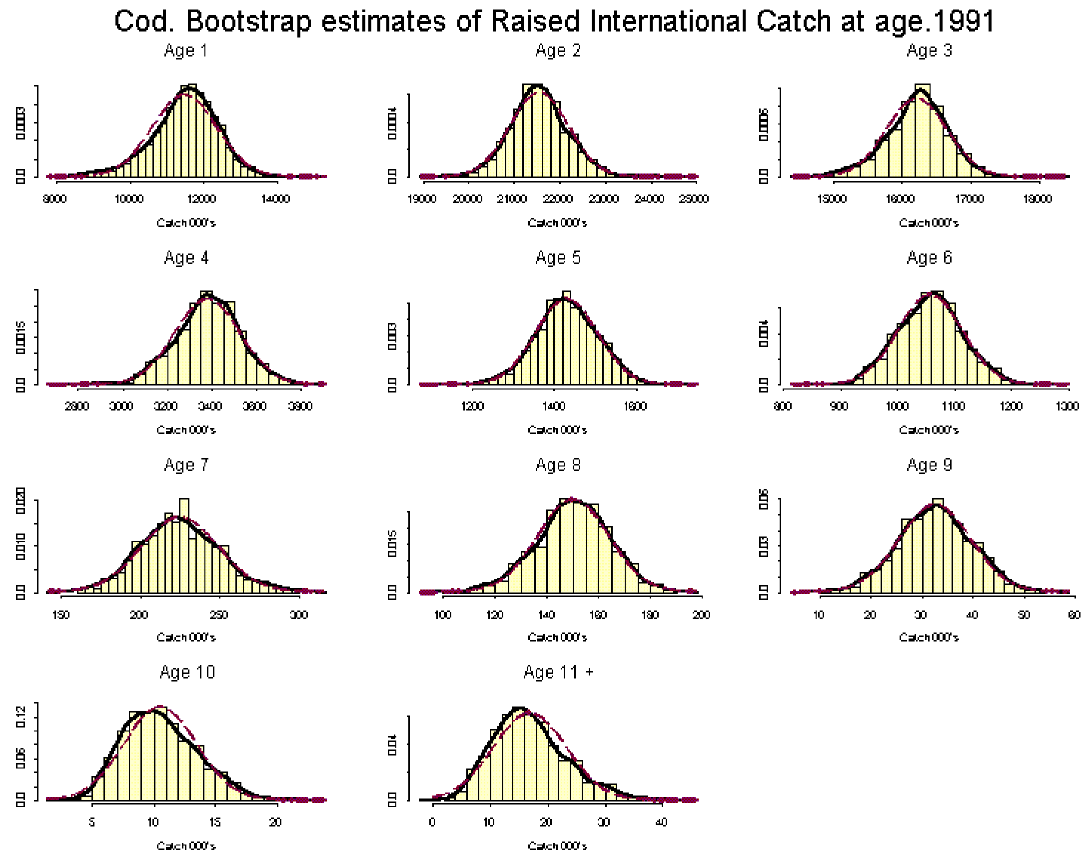


Figure 6. Histograms of bootstrap estimates of raised international catch numbers-at-age.

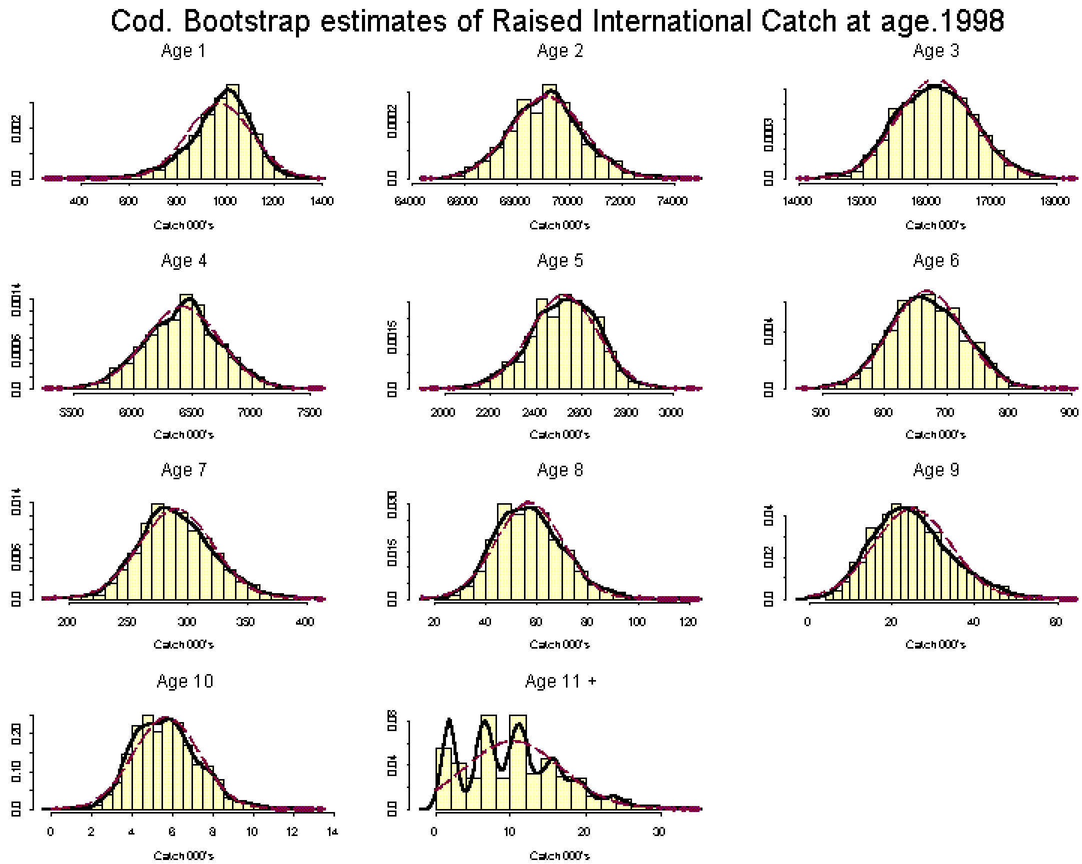
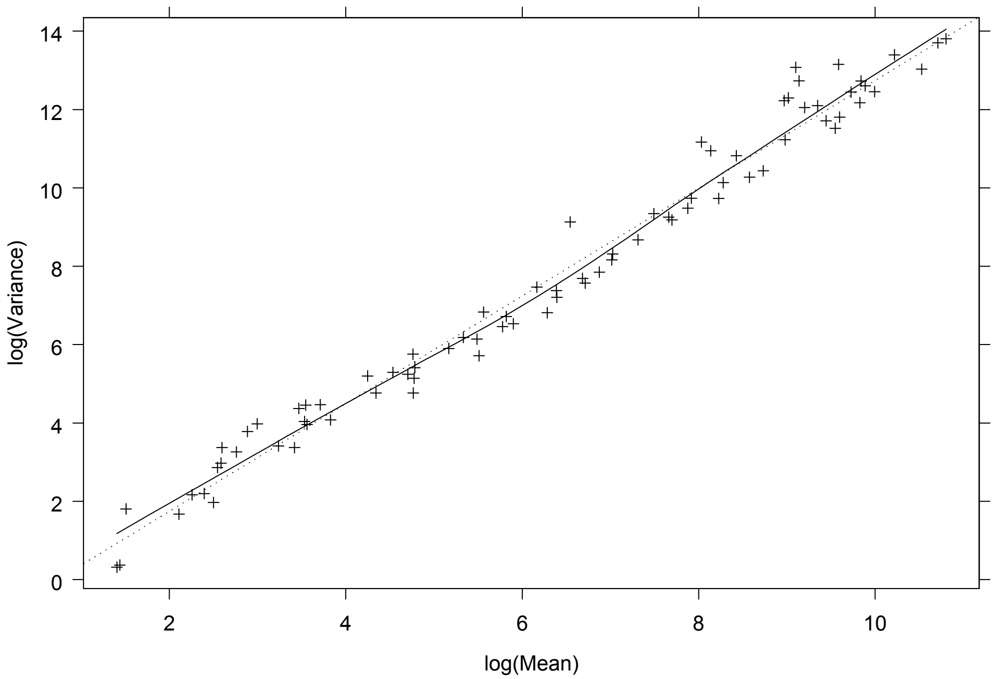


Figure 7. Log(variance)-log(mean) plots of combined UK(Scotland), Danish, UK(E&W) and Dutch bootstrapped estimates (ages 1-10, years 1991 to 1998).



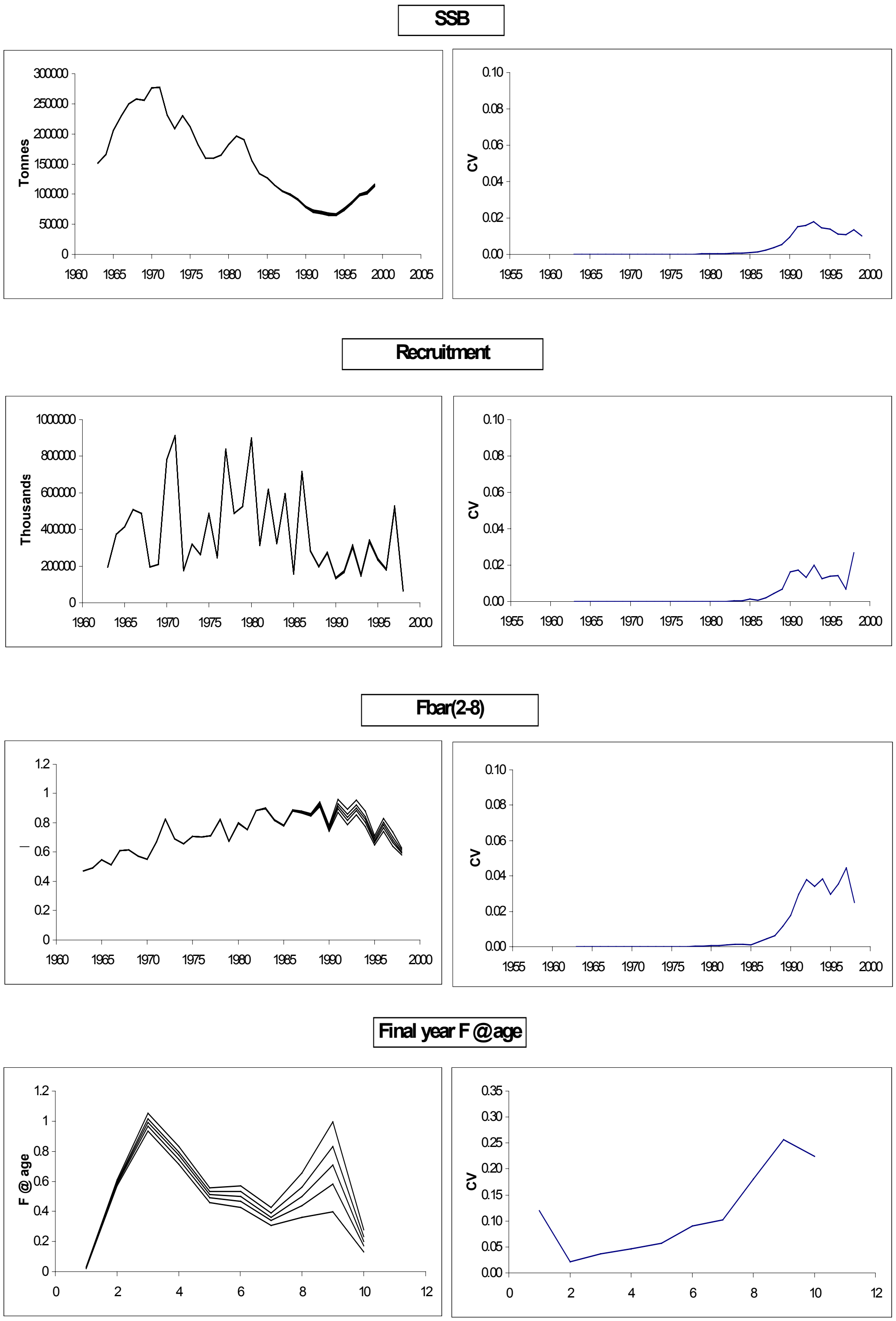


Figure 8. The 5,25,50,75,95th percentiles of Fbar (2-8), recruitment at age 1, SSB and F at age in the 1998 resulting from fitting the 1999 ICES WG XSA model to 1000 bootstraps of the North Sea cod catch at age data for the years 1991 - 1998. Tuning fleet CPUE data constant.

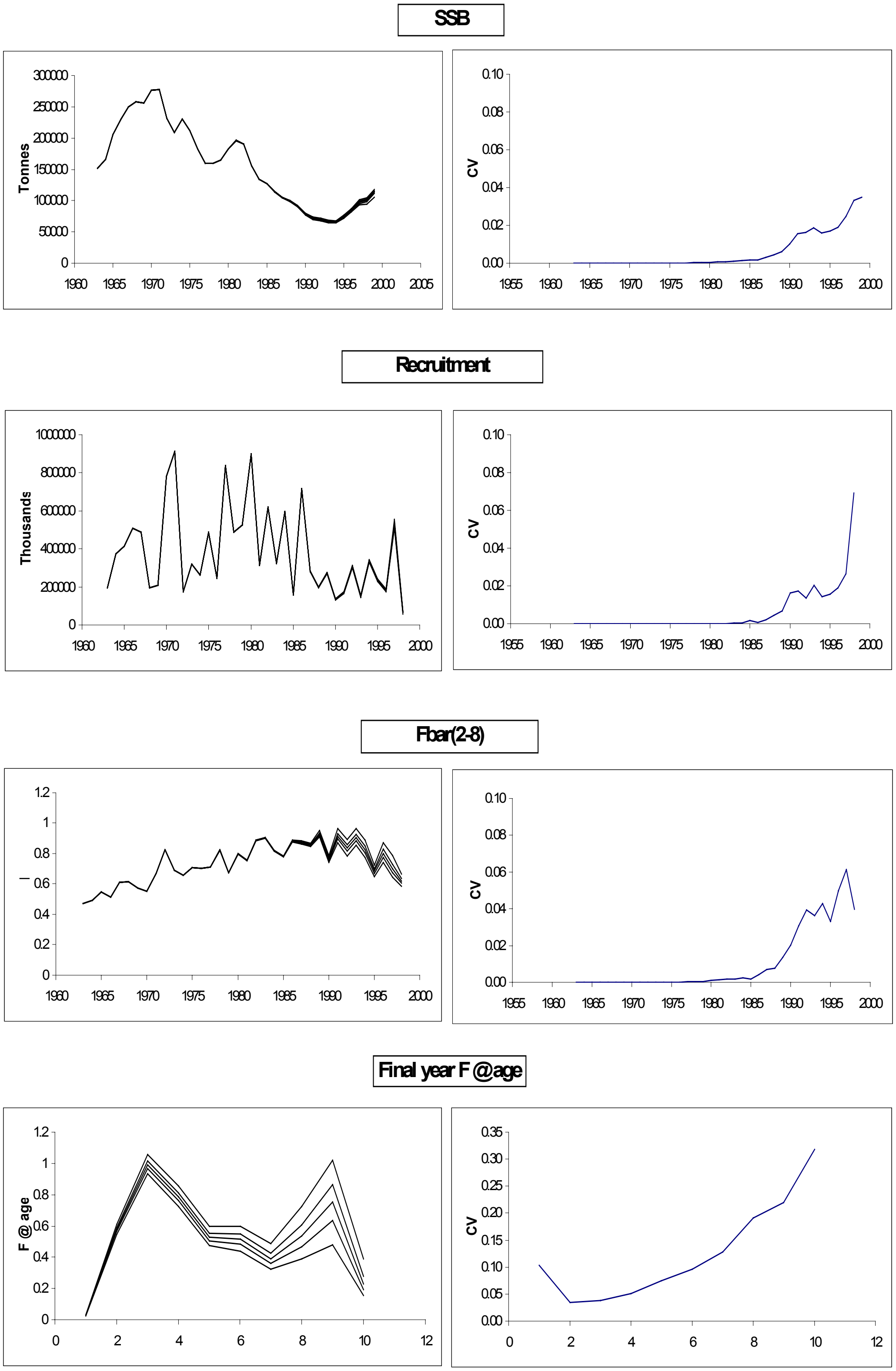


Figure 9. The 5,25,50,75,95th percentiles of Fbar (2-8), recruitment at age 1, SSB and F at age in the 1998 resulting from fitting the 1999 ICES WG XSA model to 1000 bootstraps of the North Sea cod catch at age and commercial fleet tuning data for the years 1991 - 1998. Survey tuning CPUE data constant.

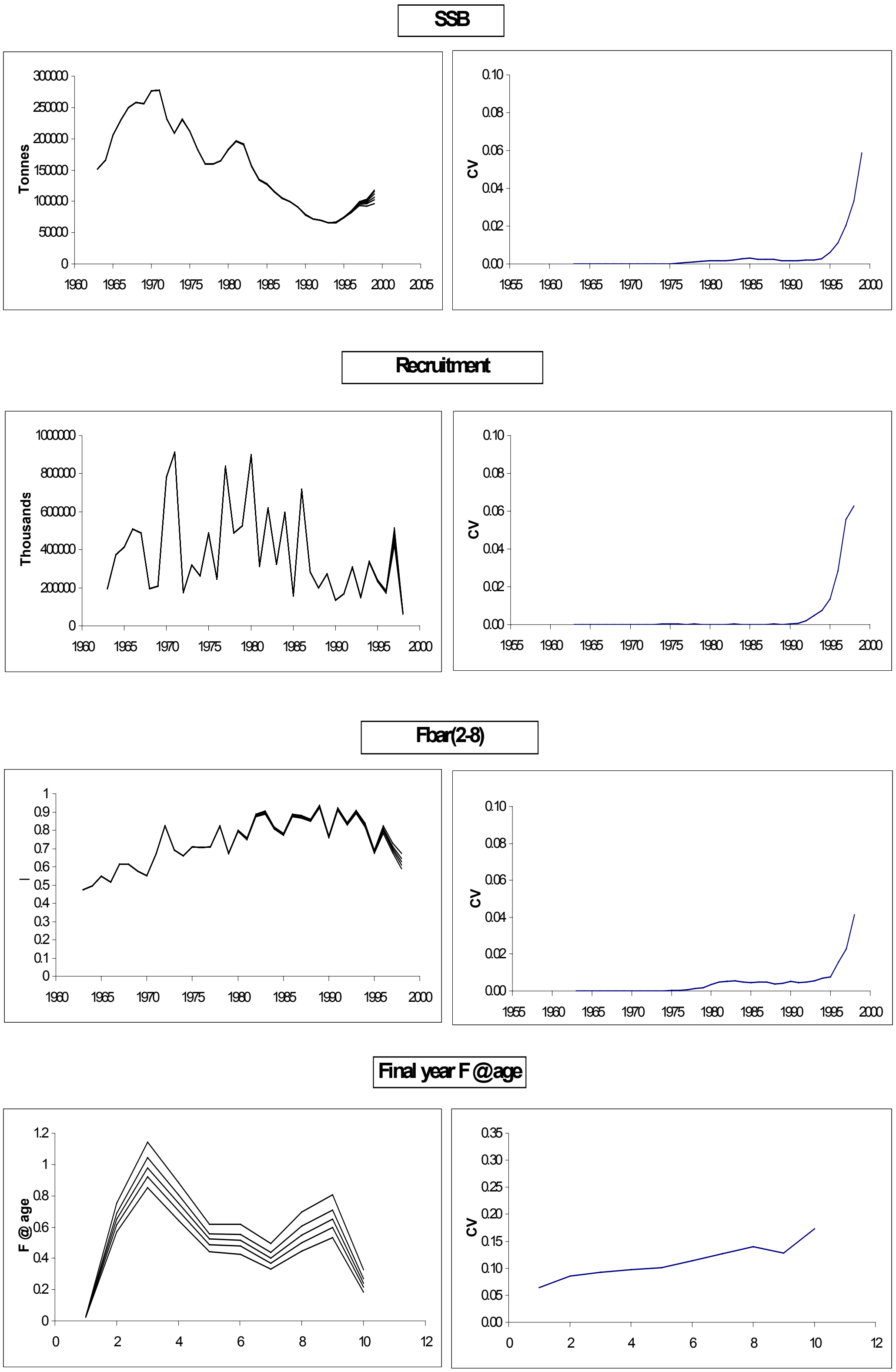


Figure 10. The 5,25,50,75,95th percentiles of Fbar (2-8), recruitment at age 1, SSB and F at age in 1998 resulting from fitting the 1999 ICES WG XSA model to 1000 non-parametric bootstraps samples of the North Sea cod survey and commercial fleet catch per unit effort data. Catch at age data constant.

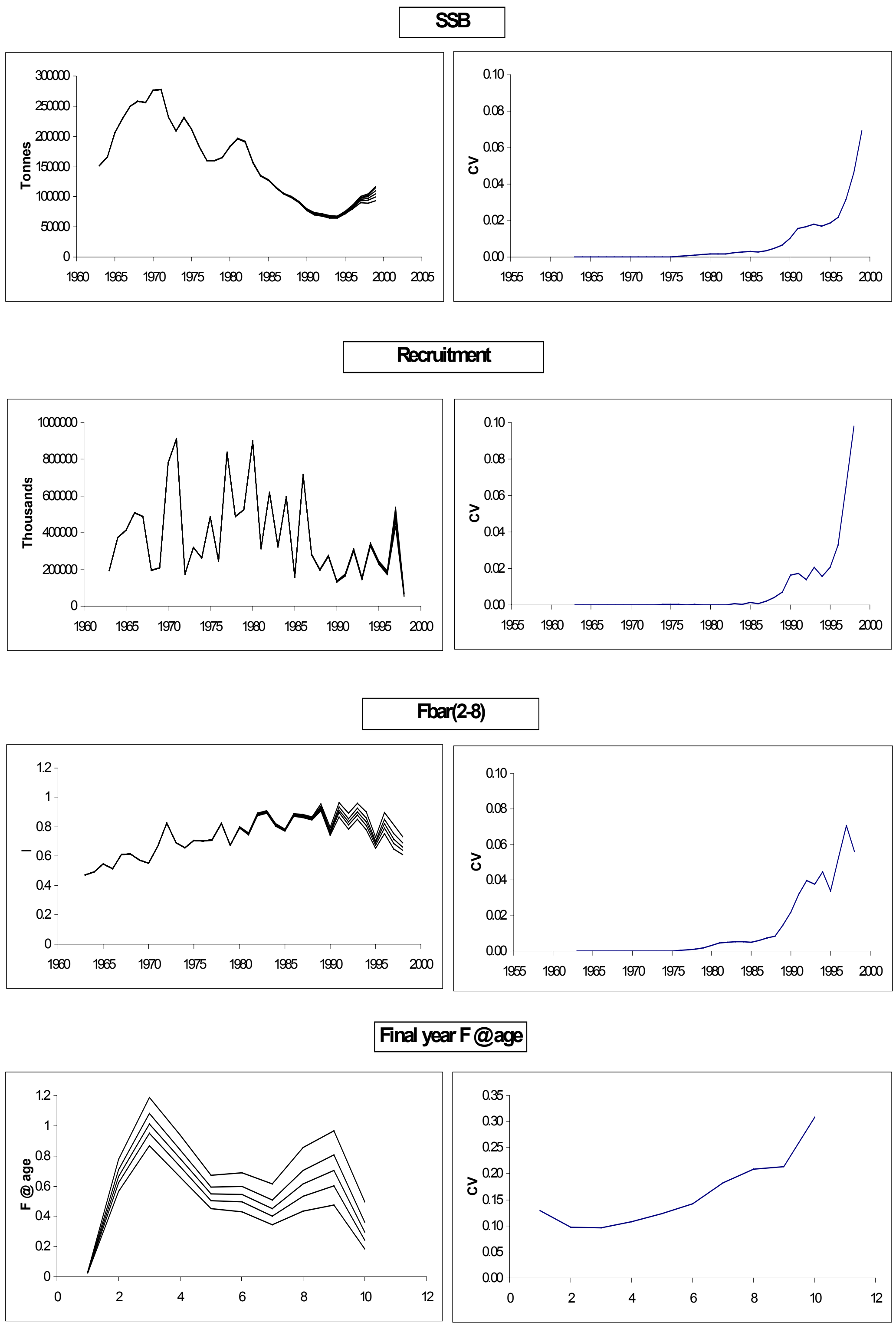


Figure 11. The 5,25,50,75,95th percentiles of Fbar (2-8), recruitment at age 1, SSB and F at age in 1998 estimated by fitting the 1999 ICES WG XSA model to 60 non-parametric bootstraps of each of 1000 assessments derived from the bootstrap samples of the North Sea cod catch at age and fleet tuning data for the years 1991 - 1998.

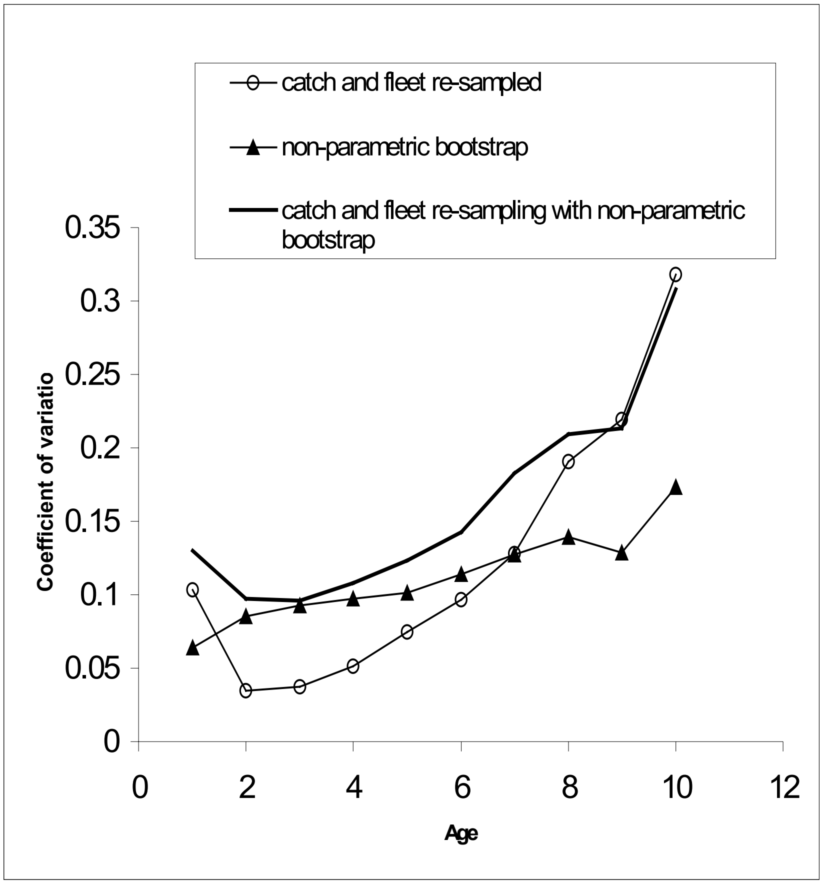


Figure 12. Estimated coefficients of variation of the North Sea cod fishing mortality at age in the final assessment year derived from three bootstrap procedures.

Draft ICES Paper

The precision of international market sampling for North Sea herring and its influence on assessment.

E. J. Simmonds, C. L. Needle, , H Degel, S Flatman, C. M. O'Brien, M. A. Pastoors, A. P. Robb and M. Vinther

E. J. Simmonds FRS Marine Laboratory, PO Box 101 Aberdeen UK. [tel 44 1224 295366 fax 44 1224 295511 Email simmondsej@marlab.ac.uk] C. L. Needle, and A P Robb, FRS Marine Laboratory, PO Box 101 Aberdeen UK.[tel 44 1224 876544 fax 44 1224 295511]
S Flatman and C. M. O'Brien, CEFAS Pakefield Road,Lowestoft, Suffolk NR33 OHT [Tel: 44 1502 562244 Fax 44 1502 513865]
M. A. Pastoors RIVO Postbox 68 1970 AB IJmuiden, The Netherlands [Tel +31 255 564646, fax +31 255 564644, email: m.a.pastoors@rivo.dlo.nl]
H. Degel and M. Vinther: DIFRES, Charlottenlund Slot, 2920 Charlottenlund, Denmark [tel: +45 33 96 3300, fax: +45 33 96 3333, e-mail: mv@dfu.min.dk]

Abstract

Market sampling is a key source of data for catch-at-age-based assessment. Little has been documented about the influence of potential error in these data on the precision of assessments and the management information they produce. This paper presents the results of a study of the precision of North Sea herring fish market sampling carried out by the UK, Denmark and the Netherlands. Data from eight years of market sampling were analysed to obtain the precision of estimated numbers-at-age in the catch. The market sample data was then used to estimate 1000 realisations of the international catch-at-age and mean weights-at-age in the catch. Three methods of estimating the variability of missing catch data were used and three options for the catch-at-age matrices were computed. These base datasets were utilised to obtain 1000 assessments conditional on the ICA (Integrated Catch-at-age Analysis) model. From the outcome of these assessments the influence of the market sampling programmes on the management of the stock are presented as 95% confidence intervals on the main management parameters (recruitment, SSB, F0-1 and F2-6). In addition, the influence of missing data is estimated. The implications of our conclusions on the requirements from a market sampling programme are discussed.

Introduction

The catch at age matrix forms a major part of the assessment of many fish stocks. For stocks with international shares of catch the sampling schemes are often diverse and the task of assembling the catch at age matrix is a complex and time consuming process. The influence of this data on the assessment is rather poorly studied. To the knowledge of the authors, there are no comprehensive studies of the precision of international market sampling programmes and their implications for fisheries management advice. The papers published on this issue either deal either with the potential effects of theoretical uncertainty in basic data on the advice provided (Pope and Gray 1983; Pope 1988; Pelletier and Gros 1994; Coggins and Quinn 1998), or on the estimation techniques and results of the analysis of uncertainty in the basic data itself (Tanaka 1953; Johnston et al. 1975; Pope and Knights 1975; Kimura 1977; Sparre et al. 1977; Gavaris and Gavaris 1983; Smith and Maguire 1983; ICES 1994; Reinert and Lewy 1998). In this paper, we present an attempt to combine these two approaches. This paper describes the analysis of catch at age data on North Sea herring from 1991 to 1998 inclusive and their use in the assessment. We wanted to combine results from different sampling programs to arrive at total international estimates of catch numbers at age and their associated variances, we could have follow two routes:

1. attempt to combine the raw sampling data, calculate appropriate age-length keys (ALK) and raise the sampling data to the total international landings. In this way the variances of the procedure could be directly calculated (Gavaris and Gavaris 1983; Smith and Maguire 1983) or obtained from bootstrap analysis. A pre-requisite for this approach is that the sampling procedures (strata) are harmonized so

that samples can be freely exchanged. This harmonisation is difficult to obtain from data already collected independently by different countries with different sampling and data storage methods.

2. use bootstrap or jackknife techniques to generate an a certain number of realisations of national age compositions and weights at age. Then combine these national realisations as an assessment working group would have done with the data for a single year, delivering a number of realisations of the international age composition. These are then fed into an stock assessment program to arrive at bootstrapped stock estimates. This approach has been followed in this study.

We present first the results from studies of national market sampling programmes for estimating catch at age of North Sea herring for the period 1991 to 1998. Market sample data from the major fishing countries for these species have been collated at minimum aggregation level and used to generate 1000 national and then international replicates for use in bootstrapped assessments, whereby the assessment procedure was kept the same as used in the most recent ICES working group (ICES 2000).

Materials and methods

Three sets of national data were used to provide 1000 replicate market sampling for three nations, Netherlands, Denmark and Scotland. The sampling methods are different for each nation, and consequentially the methods for deriving the 1000 replicates were also different. These are described by nation below. Data from the Norwegian market sampling program on North Sea herring was kindly made available by IMR Bergen. However, it was found that the implementation of the bootstrap method for these data was difficult due to sparse coverage and a need for complex fillin rules, similar to the Scottish data (see below). However, this requires intimate knowledge of the fishery which was not available to the authors. Therefore, the Norwegian herring market sample data have not been used in the analysis presented here. The implications of this and the implications of other missing data are included in these studies by examining different methods for allocation of un-sampled catch. The results from this part of the investigation indicate that this is not a major problem (see below).

The methodology for each national analysis is different, and are described below.

National analysis of The Netherlands data

The method followed a generic method developed at CEFAS UK. The process of manually allocating unsampled catches to sampled strata has not been taken into account, so the re-sampling process only operated on the already existing temporal and spatial stratification. Bootstrapping the catch at age data was carried out at the vessel level using SAS code to replicate the raising calculations carried out by the Market sampling system as closely as possible. The market samples were stratified by quarter, fleet and area. The original data were extracted from the database which holds length and biological sample data, along with combined and raised processed data. Firstly, using the sample number and vessel codes in the database, two lists were formed: the boat-trips from which age samples and length samples had been taken. Each list was sampled with replacement to form a new list – the *bootstrap*. The *bootstrap* length and age samples were then comprised of the data from the boat-trips included in the new lists. Catch-at-age and weight-at-age estimates were calculated from the bootstrap length and age samples. This bootstrap procedure was repeated 1000 times for each period of interest.

The algorithm for the bootstrap procedure used is as follows:

Set-up:

- read in original length, age and weight data.
- create a list of unique identifiers for sampling units – here used vessel code and sample number.

- calculate values that will not change with each bootstrap sample – commercial weight totals and numbers of samples.

Bootstrap loop which is repeated for a 1000 iterations.
- set seed for random number generator.
- form bootstrap length sample by resampling length data.
- calculate length distributions (LD) and analytical variance due to length sampling for bootstrap length sample using appropriate stratification and length groups (5cm cod, 2cm plaice).
- set seed for random number generator.
- form bootstrap age sample by resampling age data.
- calculate age-length key (ALK) and analytical variance due to ageing for bootstrap age sample using appropriate stratification and length groups.
- calculate age-length distribution (ALD) and analytical variance from LD and ALK.
- calculate mean length within each length group and parameter for length-weight relationship.
- calculate mean weight-at-age from ALD and length-weight relationship.
- append the estimates from this iteration to the output file.

National analysis of Danish data

The Danish sampling procedure was changed in 1998. Until that time, practically all fish were aged and the raising of biological samples to total catch was made without considering length information, even though it exists. From 1998 only a part of the length-measured fish was aged such that the raising of samples includes length distributions and age-length keys for calculation of catch at age. This approach has been used for all years in the bootstrapping exercise for a more consistent approach. Even though it has been tried to mimic what has been done historically to produce catch at age data for ICES working groups, the catch at age numbers produce are probably slightly different.

The ICES herring assessment working group divides the total assessment area into smaller areas (IVaE, IVaW, IVb, IVc and IIIa) for storage of catch and sampling data. These areas were used in the stratification of the Danish catches as well. However, due to the large variation in the industrial herring catches, area IVb was divided further into sub-areas for this type of landings.

North Sea Herring landed for human consumption and for industrial purposes are sampled differently and are treated at two independent "species".

<u>Human consumption herring</u> are sampled from landings where the fishing grounds are known. Landings are not split up on size classes, such that the stratification includes the variables sub-area, year and quarter. Similar stratification is available for the official landings statistics. A number of strata with relatively small catches has not been sampled, and the same procedure, as used for cod and plaice, of copying samples from area was used. In cases of missing samples, samples representing area IIIaN were copied from a IIIaN sample in an earlier or later quarter, if possible, and only taken from a neighboring area if there were no IIIaN samples in the actual half-year. For other areas, samples were taken from a neighboring area within the same quarter. If there were no bordering areas, data were copied from another quarter within the half-year.

The amount of <u>industrial catches of herring</u> are determined from the total industrial landings and samples of the landings for species composition. The Ministry of Fishery makes this sampling and the split of total industrial landing on species. This process is not included in the bootstrap exercise, and the catch weight of industrial herring by quarter and area are considered exact, as for landings for human consumption. Both the Ministry of Fishery and DIFRES takes samples of industrial landings for the construction of catch at age data. Such samples do include all species as well, but these samples are all analyzed by DIFRES. All samples taken have approximately the same total weight and are assumed randomly sampled within the industrial fishery. A simple aggregation of the herrings sampled in a stratum does thereby represents the total industrial fishery targeting different species, both with respect to the proportion of herring in catch, and the size distribution of herring. The same procedure, as for human consumption herring, was applied when sampling had not covered all strata.

The preprocessing of data used for calculation of catch at age includes (using combinations of year, quarter, district and size class as strata):
1. Convert landings and samples weights to live fish weight
2. Ensure that at least one sample per stratum has been taken in all principal sampling districts. In cases of missing data, copy sampling data from the geographically nearest sampling district (and rename district name and create a new unique sample id.).
3. Move landings from the non-sampled district to the nearest sampled district stratum and sum landings. Do not include landings from ports in other countries.
4. Raise all the landings allocated to a district, such that they sum up to the total national landings weight. Do this raising by sea-area and quarter.

The data set for bootstrapping has now at least one sample for all strata (catch area, "district", year, quarter and size class), and the sum of all strata landings is equivalent to the total national landings. Each sample has measurement of individual fish, such that a length distribution, an age-length key, and a mean weight at age can be calculated.

Catch at age data normally produced to the ICES assessment split herring catches into two stocks (North Sea autumn spawners, and Baltic Sea spring spawners). This separation is based on additional data on number of vertebra and otolith structure, and has not been included in this study.

<u>Bootstrapping of the raising of samples to total landings</u>
The raising of samples to total catch weight includes the following steps (all steps are default done by stratum; human consumption and industrial herring: catch area, year and quarter):

1. Take a simple random sample of the available biological samples with replacement. The number of samples taken is equal to the number of available biological samples in the stratum. Calculate total weight of samples within a stratum from the individual sample weights actually selected.
2. Option a. For each of the resampled biological samples, take a random sample of the individual fish with replacement, of a size equal to the number of fishes within the sample. Re-calculate total weight of samples within a stratum from the individual fish weights (required) in the samples.
3. Calculate raising factor from strata sample total weight and total catch weight.
4. Create a length distribution as a simple sum of sampled fish. Raise length distributions by raising factors
5. Option b. Calculate proportion at age and mean weight at age for each length group from fishes included in the selected samples
5. Option c. Calculate proportion at age and mean weight at age for each length group from all available sample with a stratum.
6. Combine length distribution and proportion at age per length group to calculate catch at age, and mean weight and length at age.
7. Sum catches at age from all size classes and districts (cod and plaice only). Calculate mean weight and mean length, using catch at age numbers as weighting factor.
8. Output catch numbers, mean weight and mean length by catch area, quarter and age

The relatively few samples, in some case just one sample, per stratum makes bootstrapping of just samples pointless. Therefore resampling was extended with the resampling of individual fish, within a resampled biological sample (option a). This requires weight of individual fishes to calculate sample weight which were available for cod and plaice 1991-1998, and for human consumption herring 1991-1997. For these groups, the age information of just the resampled fish was used (option b).

The sampling level for industrial herrings 1991-1998 was reasonable high and there was no resampling done of individual fishes. This was furthermore impossible, as the fish were worked up by length group, and not individually. Age distribution was estimated from the resampled biological samples (option b).

The change in sampling methods from 1998 made it necessary to use the full set of available length-age information (option c) to convert length distribution into ages. It would of course have been possible to make a qualified guess on age for a fish length without a resampled age information. However, to simplify the programming the full set of age information was used for 1998.

National analysis of Scottish data

The data collected is aggregated to monthly based region and gear length distributions with age length keys. These data are collected from multiple samples, however as the data is combined before entry into the database, it is no longer possible to separate the individual samples at age and the data is treated as a series of length samples with associated age sampling. The total landings for the fleet are collected as a census by region, gear and month. The data can be thought of as estimates of 'data cells' where each cell is has a landing, and a length distribution and may have an age length key. The catch-numbers (N_{armg}) and catch-biomass (W_{armg}) at age are calculated as:

$$N_{armg} = \sum_l n_{lrmg} p_{larmg} * L_{rmg} / \sum_l n_{lrmg} w_{lrm}$$

$$W_{armg} = \sum_l n_{lrmg} p_{larmg} w_{lrm} * L_{rmg} / \sum_l n_{lrmg} w_{lrm}$$

where n_{lrmg} is the number sampled at length (l) by region (r), month (m) and gear (g), p_{larmg} is the proportion at age (a) for each length by region, month and gear. (for herring this is independent of gear), L_{rmg} is the landings by weight by region, month and gear, w_{lrmg} is the weight of an individual fish at length by region, month, derived from long weight length relationships (region and month dependant for herring, monthly for cod)

The p_{larmg} are calculated from the number of fish aged at each length. Both *p* and on occasion *n* may be missing for a particular month, region and gear. In this case the *p* are 'filled in' from another region, gear or adjacent month. These fill-in sequences are provided as a standard from the sampling program.

Three main methods were applied to try to estimate the precision of the Scottish market sampling scheme;
1. a simple jackknife (Efron and Tibshirani 1993) procedure with the use of fill-in rules for missing data,
2. a grouped bootstrap (Efron and Tibshirani 1993) where monthly and gear categories were combined to give a number of samples by region and quarter, these were then bootstrapped by group,
3. a weighted jackknife similar to the simple jackknife but weighing the probability of a data-cell being removed according to estimates of the probability of sampling the cell based on 8 years data.

In practice only the first method was used, the second method gave high CVs for all ages as would be expected. The third method provided smaller CVs but relied upon the calculation of the effective sample size, it is unclear how to calculate this factor for such a weighted resampling method, it was thought preferable to be conservative and assume that the simple jackknife would gave adequate results.

For all procedures the following initial set up was carried out
1. Obtain total catch for commercial catch per data cell L
2. Obtain a length frequency distribution per data cell LF
3. Obtain an age length key for those cells for which it was available ALK
4. Obtain a list of links between length keys to age length keys, for all data cells, using fill-in rules as required
5. Select data cells for removal randomly without replacement; for simple jackknife select with equal probability;
6. Create new data set with selected samples removed
7. Find new fill-ins for data cells without length or age length keys
8. Calculate the mean weight W of fish for each data cell using standard monthly, region dependent length weight relationship and length frequency LF.
9. Calculate the total number of fish N for each data cell from the total catch L and the mean weight W

10. Calculate the number at age N_a for each data cell using the total N, the length frequency LF and the age length key ALK
11. Calculate the mean weight at age for each data cell using the length frequency LF, weight at length W_l and the age length key ALK
12. Calculate the total numbers at age by summing the numbers at age per data cell
13. Calculate the total biomass at age by summing the numbers * mean weight at age per data cell

Following 1000 replications
1. Check that 1000 values have similar mean to original data
2. Calculate CV from mean and variance of 1000 replicates.
3. Correct the Jackknife estimates of CV for number of data cells and removed samples (Efron and Tibshirani, 1993)
4. For Jackknife inflate catch number at age of each replicate by scaling the replicates about the mean and setting the small number of negative observations (<1%) to zero.

The sampling, the procedure attempts to estimate data cells organised by month, area and gear. As the sampling is only partial, inevitably it is not possible to fill all of the cells where landings are reported in the year. In some cases no data is available at all, in others length keys only are available and age length key data must be supplied. The current method used is to assign length or more usually age/length key data from another cell. This process is in effect a step-wise spatial temporal based model, estimated by nearest neighbour method. The nearest neighbour is selected with a sequence of assignments from previous or following month based on the most similar areas. The presence of data in the same area gear cell is checked in previous and subsequent months, then in sequence (1^{st}, 2^{nd},3^{rd} etc.), until an alternative is found. (See Table 1 and Figure 1).

Analysis of internationally combined market data

The 1000 bootstrap replicates of mean weight and catch at age from Denmark, The Netherlands and Scotland were combined into 1000 replicates of international catch data. This fully sampled component constitutes on average 66% of the North Sea herring landings over this period. In addition to this fraction of the catch the area misreported data from VIa_{north}, is allocated to Scottish fleet and unsampled catches from English German and French fleets are usually raised by Netherlands samples in the Working Group, this increases the proportion of the catch covered by the sampling to 75% of the total. The major missing components are the remaining unallocated landings and the Norwegian catch discussed above. The bootstrapped components both underestimate and overestimate numbers at age because landings are both added and subtracted due to area misreporting, discards and catches of Baltic Spring Spawning herring in the North Sea.
To carry out the assessment the catch estimated from the bootstrap replicates had to be scaled to the WG catch. Three methods were used for this purpose:

- **Scn** Scaling to WG numbers at age by year and age dependant multiplicative factors.
- **Scb** Scaled to landings biomass, retaining bootstrap age structure but scaling with year dependant biomass scaling factors.
- **Miss** Difference between WG catch and mean bootstrapped replicated catch (positive or negative as necessary) was estimated. This missing catch at age by year was used to scale a simulated sampling scheme with the same CV and correlation at age as the Danish sampling scheme (but with uncorrelated with the Danish estimates).

Assessment of herring

The assessments carried out to study the effects of estimates of landings have been done using models, indices and procedures of the ICES Herring Assessment Working group (ICES 2000). The Integrated Catch at Age (ICA) model was used to assess the state of the stock (Patterson 1998).

Deterministic catch-numbers at age were available for the year range 1960 to 1990 (ICES 2000a, section 2.2) and bootstrapped numbers at age for the period 1991-1998. Also the bootstrapped mean weight at age was available for 1991-1998. All other data was the same as used in the assessment working group.
Survey indices:

- MIK 0-wr index. Available and used since 1977 as a recruitment index (ICES 2000a, section 2.3)
- Acoustic 2-9+ wr index. Available since 1989 (ICES 2000a, section 2.4)
- IBTS 1-5+ wr index. Separated into a 1 wr index (used since 1979) and a 2-5+ wr index (used since 1983). (ICES 2000a, section 2.3 and 2.6)
- Multiplicative larvae abundance index (MLAI). Available since 1973, used since 1979 as an SSB index (ICES 2000a, section 2.5).

Data from 2000 assessment were used for all other input parameters such as natural mortality, spawning proportions and proportion of mortality prior to spawning. Assessment of the stock was carried out by fitting the integrated catch-at-age model (ICA) including a separable constraint over a eight-year period(Patterson and Melvin 1996). Input parameters and model setup for the ICA assessments were taken from the 1999 assessment WG (ICES 1999). All catch data (within the separable period) where weighted with a weight of one. Each of the separate survey indices where also weighted with a weight of one, because errors were assumed to be correlated by age for both the acoustic survey and the IBTS (2-5+) index. The stock-recruitment model was weighted by 0.1 as in WG assessment, in order to prevent bias in the assessment due to this model component.

Results

The bootstrapped catch at age of North Sea herring 1991 to 1998 scaled to biomass (Scb – see above) are shown plotted with WG estimates of catch at age in Figure 2. This set shows the greatest deviation from the WG catch, for the other methods the mean catch at age is equal to the WG catch at age..

Catch at age distributions.

Histograms of the bootstrap estimates of catch at age, scaled to the total international landings, are shown in Figures 3 &4 for two arbitrarily chosen example years (1991-1998)**.** Superimposed on each histogram is a normal distribution with the same mean and variance as the data. These allow a visual inspection of how well normal distributions fit the data. The histograms show some departures from the normal distribution, mainly for the 0- and 1-ringers (industrial fishery) and for the older fish (due to the lack of data).
Overall, the normal distribution appears to be reasonable description of the data. This may be a result of the many stages of combining data involved in producing the international estimates. A plot of log variance log mean catch is shown in Figure 5.

Uncertainty and precision

The CV of catch numbers at age for the national and combined data set are presented in Figure 6. CV of the international catch numbers follow the same pattern as observed for the national data, with relatively higher CV on the very young and older age groups. As expected, the international CV are lower than the national CV. For both cod and plaice the CV of the most fished age groups are less than 5%,
The CV of the combined mean weight at are generally less than 20% for most age groups and about 5% for the dominant age groups The underlying correlation of catch numbers-at-age was estimated using the numbers-at-age obtained from the resampling of the market sampling data. The patterns of positive and negative correlation were similar across the years within a species and the mean correlation coefficients between estimates of catch numbers-at-age are given in the Table 2 The correlation between estimates at age is positive for ages 3 to 8 for herring. It appears from this analysis, that the process is dominated by groups of fish at older ages being landed together in groups, so the presence of a group of ages increases or decreases together. It is important that this type of correlation within the estimates of catch are dealt with correctly within the assessment and that the process inducing the correlation structure is understood.

Implications of uncertainty for stock assessment and management advice

North Sea herring management is based on SSB, F adult and F juvenile, with short term projections dependant on estimates of recruitment. The median and 95% intervals of these four parameters for the last few years of the assessment are shown for all three methods of combining the catch at age data in Figure 6 (juvenile 7a and adult fishing mortality 7b), Figure 8 (recruitment) and Figure 9 (SSB)
The Coefficient of variation on fishing mortality in the final year is 4% and 8% for adult and juvenile mortality respectively (Erreur! Source du renvoi introuvable.). The CV on recruitment is 4% and 2% for SSB due to the precision of the catch estimation. However, it must be remembered that these CVs are conditional on the estimate of total landings.
The differences in the assessment carried out in 1999 [ICES, 1999] and the median of the analysis presented here are small. For the terminal year (1998) the SSB is different by less than 3%, however, the difference in F was 13% and the difference in recruitment 16%. In all cases these differences are very small in the context of the intervals on the assessment from the analysis of historic uncertainty [Figure 2.8.13 in \ ICES, 1999]. It should be noted that for the analysis carried out here the model assumptions were different. The software used at the ICES WG was a special version of ICA which allowed two periods of separable constraint for juveniles (ages 0 & 1 wr) but only one for adults, (ages 2-8 wr). Here only the standard ICA was available in a form to allow multiple assessments and it was used with two separable periods for all ages. For SSB the difference was well inside the 95% intervals from the market samples, however, the differences in F and recruitment were just outside these intervals. This indicates that the model assumptions may be more important in estimating values for F and recent recruitment than the influence of the market sampling data for this stock.

Comparison of errors contributed by market sampling to uncertainty in the assessment for herring.

The contributions of the market sample data to the overall precision of the assessment may be estimated by comparing the coefficient of variation (CV) from the catch variation alone using the bootstrapped assessments with the estimates of historic uncertainty from the assessment carried out in the Herring Assessment WG in 1999. Three main parameters can be compared: spawning stock biomass (SSB) F adult, (F_{2-6}) and recruitment. In all cases the comparison is limited to the estimates made using data up to 1998 and for the years 1991 to 1998 the period over which bootstrap data is available for catch. The CV on the parameters estimated by variance – covariance method in ICA can be compared to the contribution due to the market sampling (see Table 3). The 95% intervals for the same parameters estimated from 1000 estimates can be compared in Figure 10 for F2-6 Figure 11 for SSB and Figure 12 for Recruitment.

The results for herring suggest that for this assessment the variability in SSB and Recruitment contributed by the market sampling is negligible with a CV of between one eighth to one thirtieth of the CV indicted by the historic uncertainty. The CV of the estimates of F_{2-6} suggest an interval of between half and one fifth if the historic CV is contributed by the market sampling programme. The relatively small contribution of the market sampling variability to the overall precision of the assessment may be the result of a rather well sampled fishery and specific to assessments which use market sampling only to construct the catch at age matrix and do not use commercial CPUE tuning indices.

DISCUSSION

The international sampling programmes appear to be delivering estimates of catch at age that are rather precise, with CV's of 6% for herring for the best estimated ages rising to about 30% at the older ages. While the precision of the best estimated ages is good, the current scheme is delivering much poorer CVs on older ages. Care must be taken to ensure that the importance of estimating both old and young year classes is fully understood. Based on the analysis of the histograms of the numbers at age, the normal distribution appears to be reasonable description of the catch at age data. However, this may be a result of the many stages of combining data involved in producing the international estimates and assumes independence among the national programmes. Negative correlations are observed between estimates of younger age classes, and positive correlations are found between estimates of older ages for the three species examined. The positive correlations at older ages are thought to be a property of the population distributions and the fisheries, older fish are caught and sampled in groups. In addition there is negative correlation between estimation of most of the old ages and most of the young ages. It is thought that this results from the above mentioned correlation in the estimates of older ages and the national raising procedures to total national catch. The mechanisms used to raise age structures to total catch result in a pattern of negative correlation between younger ages and all older ages.

The results of the analyses reported here are conditional on an accurate catch census, and do not yet include bootstrapped CPUE indices from commercial fleets (which are part of the market sampling programmes) because they are not used in herring assessment. These studies are suggesting that for the data sets examined the current levels of market sampling cause only small amounts of variability in assessment outputs for North Sea herring.

The relationship between the mean and variance of the numbers-at-age is fundamental to any future statistical modelling of catch numbers-at-age; as is the assumption of independence between numbers-at-age. The underlying relationship between mean-variance of catch numbers-at-age was investigated by considering the mean and variance of the numbers-at-age obtained from the resampling of the market sampling data and compared to the power relationship:

$$\text{variance}\{\text{bootstrapped numbers-at-age}\} = e^{a} \, . \, \text{mean}\{\text{bootstrapped numbers-at-age}\}^{b}$$

Relationships between mean and variance are observed for herring, slopes on the log variance-mean relationships are 1.7 for herring. Assessment models generally do not take this into account; changes to models or to weighting practices that would include these mean-variance relationships would be helpful. The apparent proportionality for the variance-mean relationship will facilitate the development of appropriate statistical models of catch-at-age that do not assume a log-normal distribution for catch-at-age.

While the precision of this well-sampled fisheriy appears to be rather good, no attempt has been made to check whether the international sampling is representative. It is particularly important if sampling methods are changed that care is taken to ensure that sampling covers the whole fishery.

3.1 References

Cleveland, W. S. and S. J. Devlin (1988). "Locally-weighted regression: an approach to regression analysis by local fitting." Journal of the American Statistical Association **83**: 596-610.

Coggins, L. G. and T. J. Quinn (1998). A simulation study of the effects of aging error and sample size on sustained yield estimates. Fishery stock assessment models. F. Funk, T. J. Quinn, J. Heifetzet al**:** 955-975.

Darby, C. D. and S. Flatman (1994). Virtual population assessment: version 3.2 (windows/dos) user guide.

Gavaris, S. and C. A. Gavaris (1983). Estimation of catch at age and its variance for groundfish stocks in the Newfoundland region. Sampling commercial catches of marine fish and invertebrates. W. G. Doubleday and D. Rivard, Canadian Special Publication of Fisheries and Aquatic Sciences. **66:** 178-182.

ICES (1994). Report of the workshop on sampling strategies for age and maturity, Copenhagen, 3-9 february 1994. **ICES C.M. 1994 / D:1 ref: G,H,J**.

ICES (1999). Report of the Herring assessment working group for the area south of 62, Copenhagen, 15-24 march 1999. **ICES C.M. 1999 / ACFM:12**.

ICES (2000a). Report of the Herring assessment working group for the area south of 62, Copenhagen, 14-23 march 2000. **ICES C.M. 2000 / ACFM:10**.

ICES (2000b). Report of the working group on the assessment of demersal stocks in the North Sea and Skagerak. Copenhagen, 11-20 October 1999. **ICES C.M. 2000 / ACFM: 7**.

ICES (2001). Report of the working group on the assessment of demersal stocks in the North Sea and Skagerak. Copenhagen, 3-12 October 2000. **ICES C.M. 2001 / ACFM: 7**.

Johnston, C., M. D. Nicholson and J. A. Pope (1975). The precision of estimated age compositions.

Kimura, D. K. (1977). "Statistical assessment of age-length-key." Journal of the Fisheries Research Board of Canada **34**: 317-324.

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

Patterson, K. R. and G. D. Melvin (1996). Integrated catch at age analysis. Version 1.2.

Pelletier, D. and P. Gros (1994). "Assessing the impact of sampling errors on model-based management advice: an application to the Northwest Atlantic." Canadian Journal of Fisheries and Aquatic Sciences **48**: 2129-2139.

Pope, J. G. (1988). Collecting fisheries assessment data. Fish population dynamics**:** 63-82.

Pope, J. G. and D. Gray (1983). An investigation of the relationship between precision of assessment data and the precision of Total Allowable Catches. Sampling commercial catches of marine fish and invertebrates. W. G. Doubleday and D. Rivard, Canadian Special Publication of Fisheries and Aquatic Sciences. **66:** 151-157.

Pope, J. G. and B. Knights (1975). Sources of variation in catch at age data and the optimal use of age reading effort.

Reinert, R. and P. Lewy (1998). Sampling errors in estimated catch-at-age data in the Danish human consumption sampling program for the North Sea.

Sheather, S. J. and M. C. Jones (1991). "A reliable data-based bandwidth selection method for kernel density estimation." Journal of the Royal Statistical Society series B **53**: 683-690.

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

Smith, S. J. and J. J. Maguire (1983). Estimating the variance of length composition samples. Sampling commercial catches of marine fish and invertebrates. W. G. Doubleday and D. Rivard, Canadian Special Publication of Fisheries and Aquatic Sciences. **66:** 165-170.

Sparre, P., H. Knudsen and S. Munch-Petersen (1977). Optimization of sampling programs for estimation of age distribution and fishing mortality.

Tanaka, S. (1953). "Precision of age-composition of fish estimated by double sampling method using the length for stratification." Bulletin of the Japanese Society of Scientific Fisheries **19**: 657-670.

Venables and Ripley (1997). Modern Applied Statistics with S-Plus.

Wedderburn, R. W. M. (1974). "Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method." Biometrika **61**: 439-447.

3.2 Results

Table 1 Scottish neighbour sequences for herring areas (see: Figure 1)

Area	Area Neighbour
1	2, 8 ,3
2	1, 3, 8
3	4, 2, 1
4	3, 5, 6
5	4, 11, 6
6	5, 4, 11
7	1, 9, 8
8	9, 2, 1
9	8, 7, 1
10	9, 13, 7
11	5, 12, 6
12	11, 6, 14
13	10, 9, 7
14	12, 6, 11

Table 2 Mean correlation coefficient for North Sea herring catch at age from 1991-1998, for combined Danish, Dutch and Scottish bootstrapped estimates.

	0	1	2	3	4	5	6	7	8	9+
0	1.00	-0.29	-0.07	-0.06	-0.05	-0.02	0.01	-0.01	0.00	0.00
1		1.00	-0.22	-0.10	-0.07	-0.05	-0.03	-0.04	0.00	0.00
2			1.00	0.02	-0.37	-0.34	-0.30	-0.29	-0.26	-0.31
3				1.00	-0.02	-0.24	-0.31	-0.21	-0.17	-0.23
4					1.00	0.19	0.06	0.02	0.00	-0.04
5						1.00	0.29	0.20	0.08	0.06
6							1.00	0.22	0.15	0.31
7								1.00	0.18	0.37
8									1.00	0.22
9+										1.00

Table 3 Comparison of Estimated Coeficient of Variation (CV) for management parameters $F_{2\text{-}6}$, SSB and Recuitment for 1991 to 1998. Estimated from 1999 assessment using ICA estimates of historic uncertainty and estimates of the contribution of the markert sampling variability from bootstraped assessments.

		1991	1992	1993	1994	1995	1996	1997	1998
F2-6	As Hist Unc	41%	41%	40%	41%	43%	40%	42%	45%
	Market	3%	3%	9%	8%	18%	9%	7%	8%
SSB	As Hist Unc	128%	182%	136%	203%	39%	27%	20%	23%
	Market	3%	4%	4%	8%	3%	3%	3%	4%
Recruits	As Hist Unc	175%	293%	138%	71%	47%	40%	38%	40%
	Market	5%	3%	3%	3%	4%	4%	6%	4%

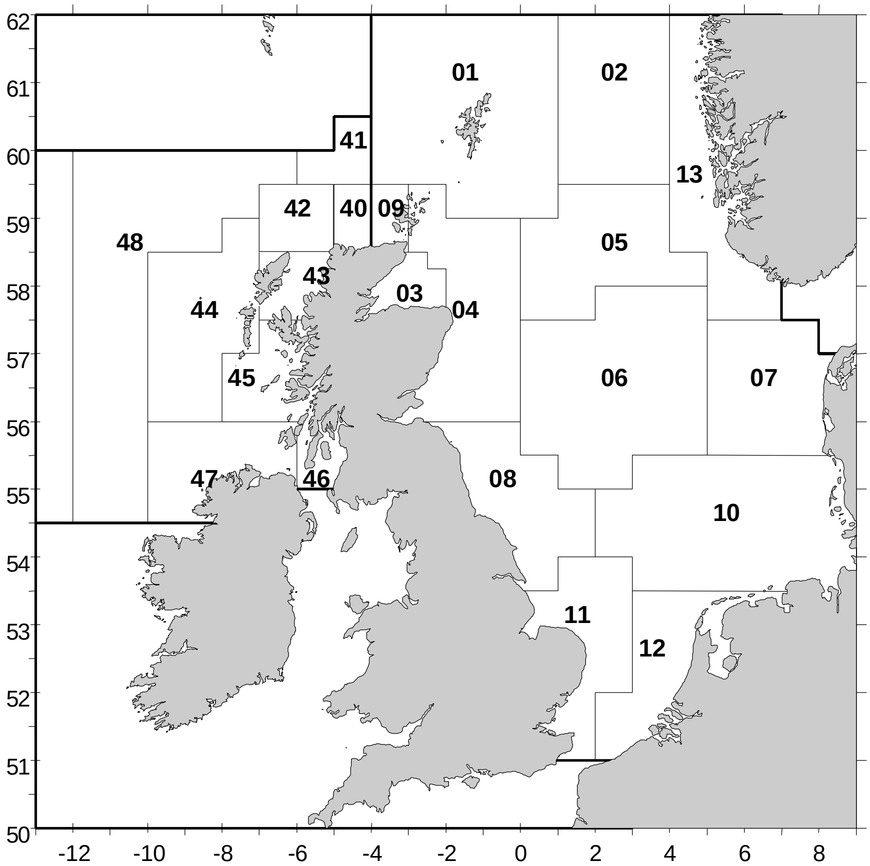


Figure 1 Scottish herring sampling areas.

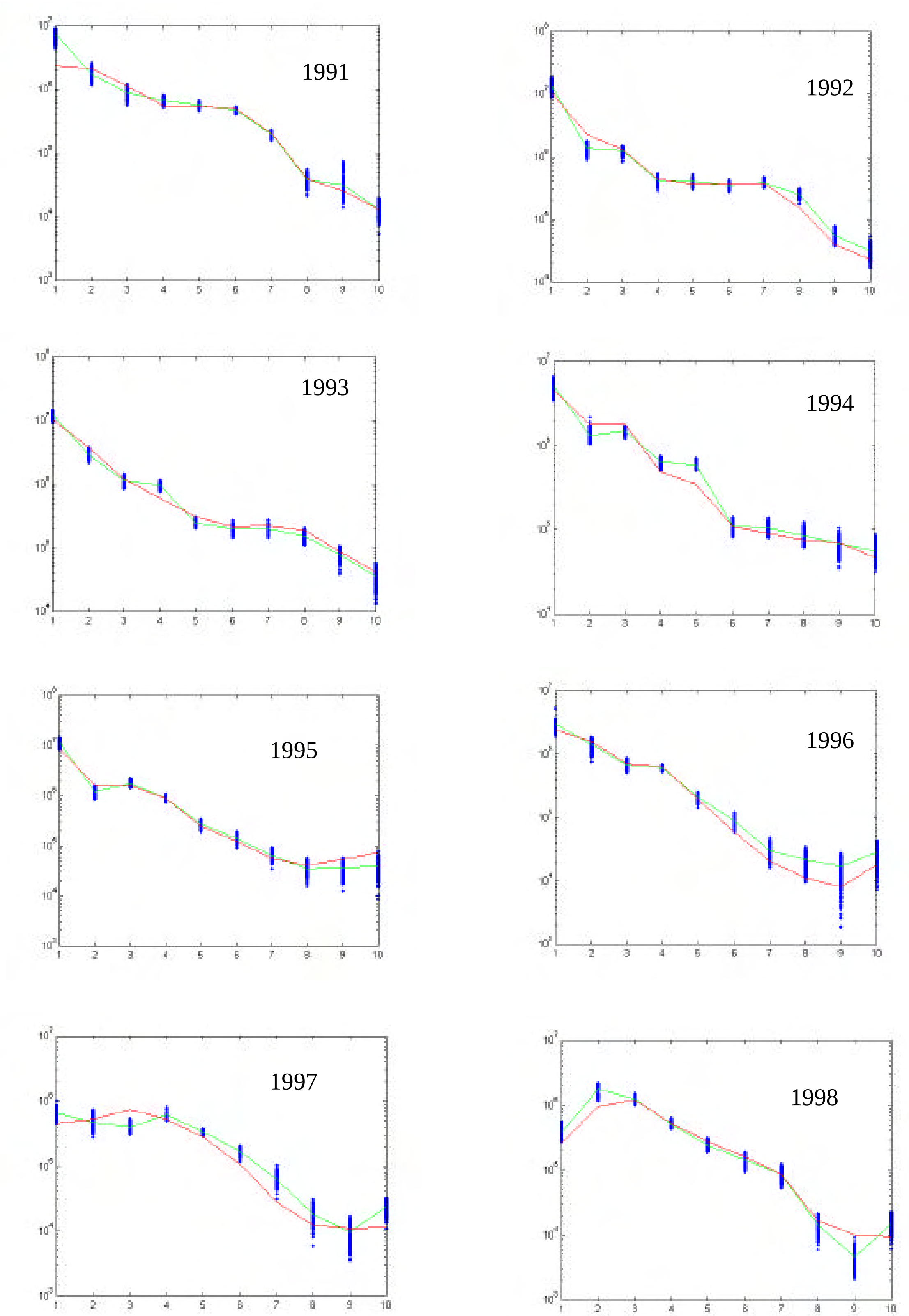


Figure 2. North Sea herring catch in number at age (wr+1) from 1991 to 1998, showing WG catch (red), boostrap mean catch (green) and boostrap values (blue) for catch scaled to biomass of landings (Scb). For other methods (Scn and Miss) mean numbers at age for WG are equal to mean numbers at age in the bootstrap.

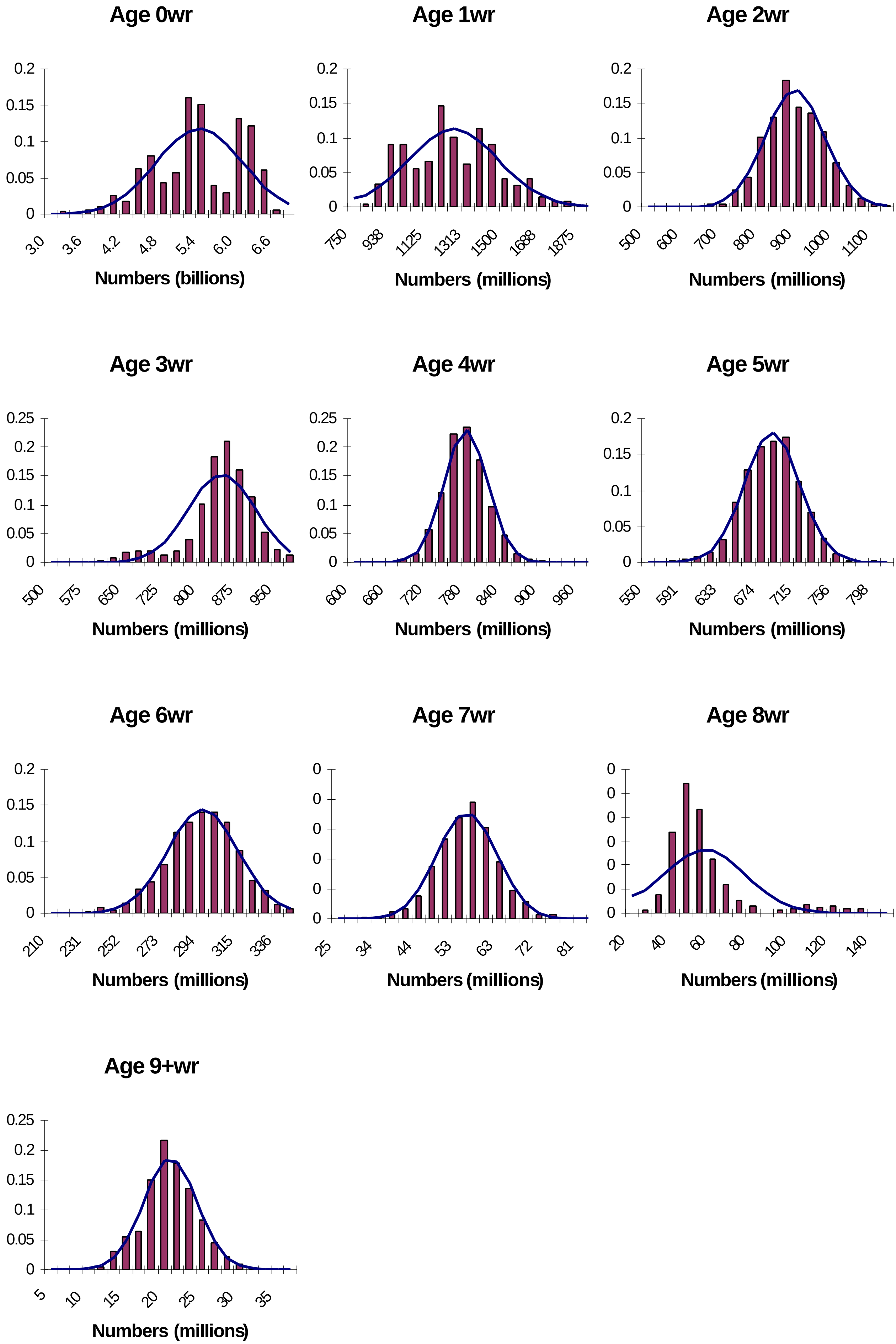


Figure 3. Histograms of international catch at age (wr) from 1991. Lines show a normal distribution with same mean and variance.

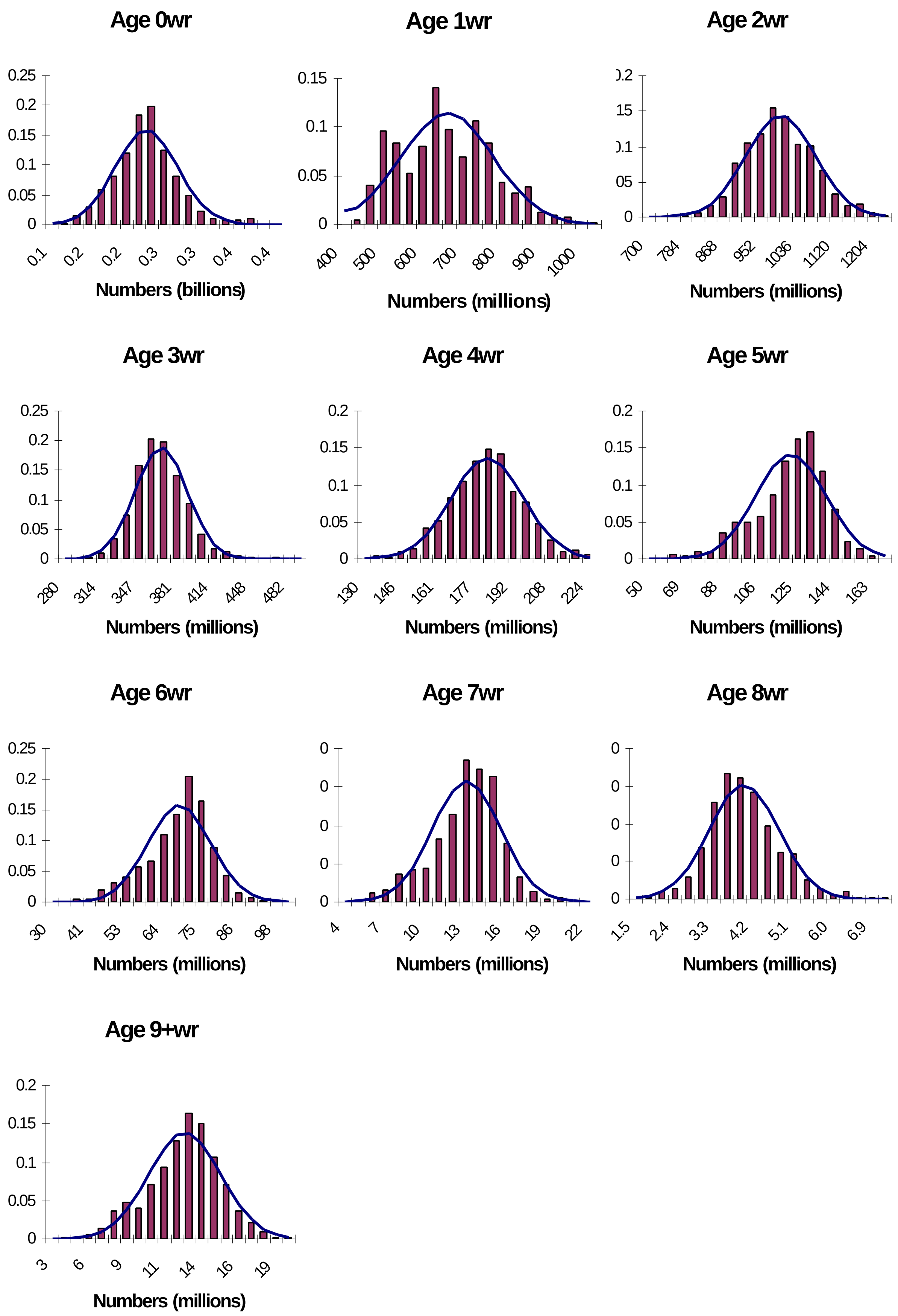


Figure 4. Histograms of international catch at age (wr) from 1998. Lines show a norrnal distribution with same mean and variance.

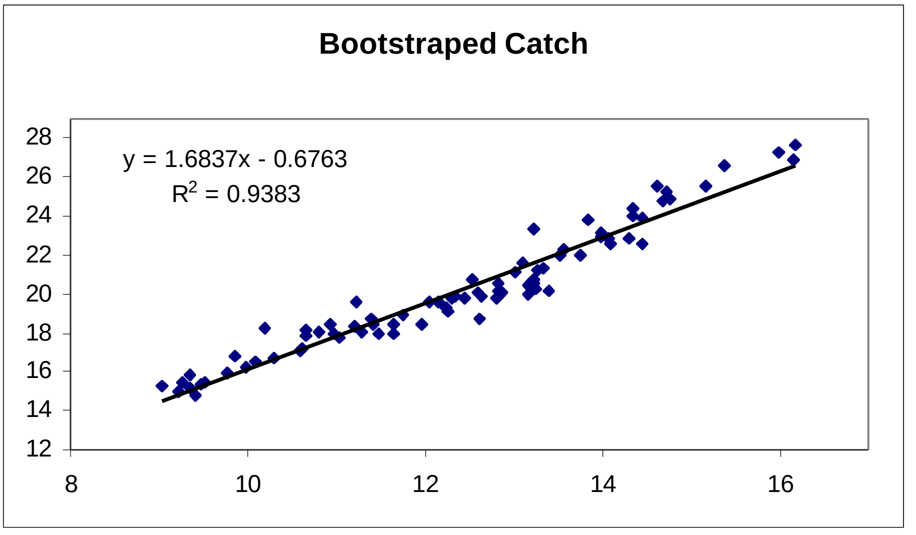


Figure 5 North Sea herring log(Mean)-log(Variance) plots of combined Danish, Dutch and Scottish bootstrapped estimates (ages 1-9). All years combined.

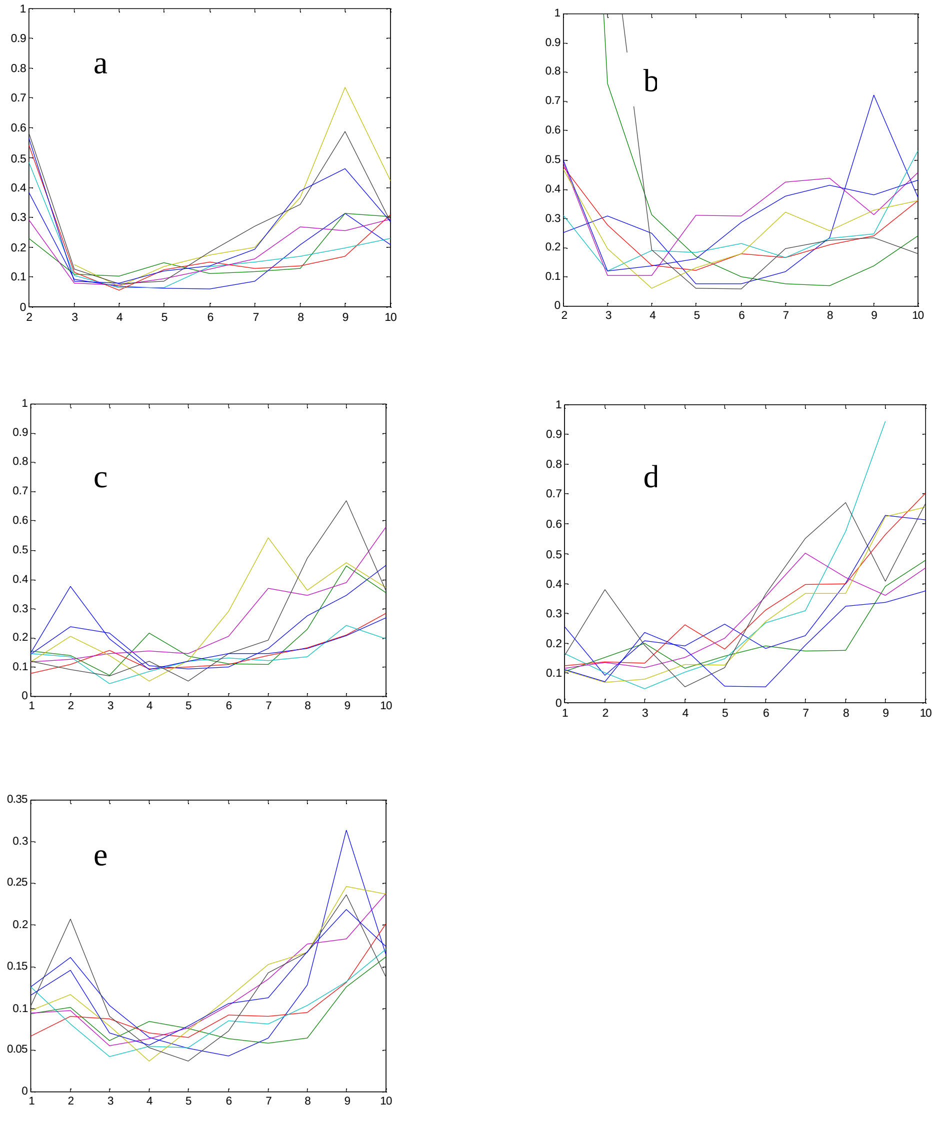


Figure 6 National CVs for North Sea herring catch at age (wr+1) for a) Netherlands, b) Scotland, c) Denmark in ICES area VI, d) Denmark in ICES area III. CVs at age for North Sea herring for 75% of combined international catch (e).

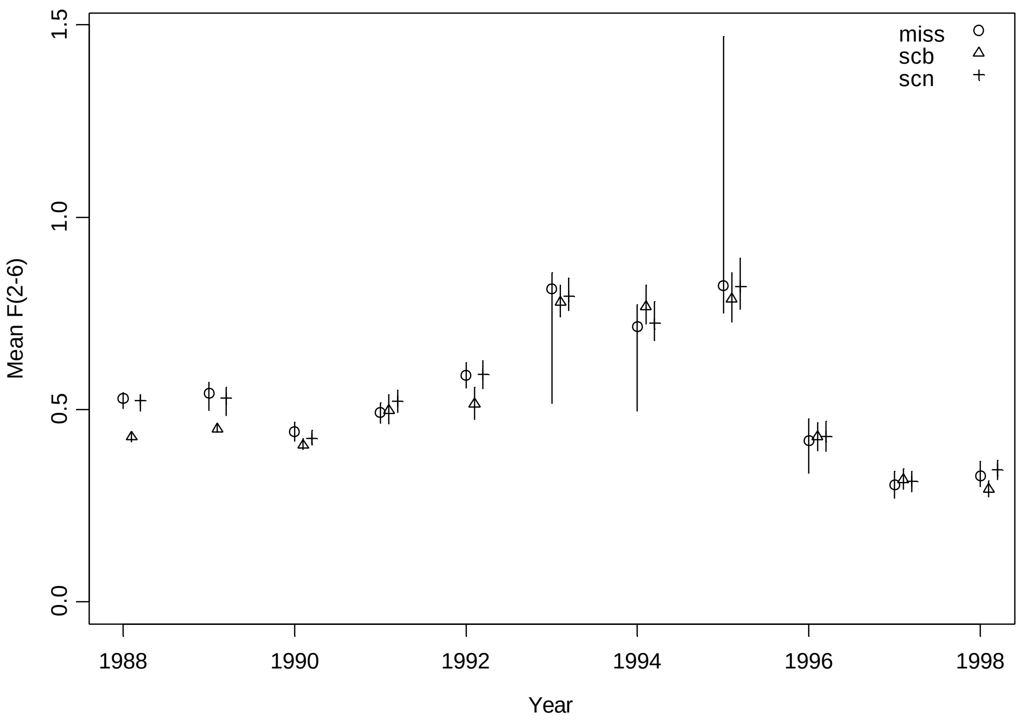


Figure 7a Estimated median and 95% intervals on estimated mean adult F (ages 2-6 wr) from 1988 to 1998 conditional on total landings and boostraped estimates of catch at age 1991-1998.

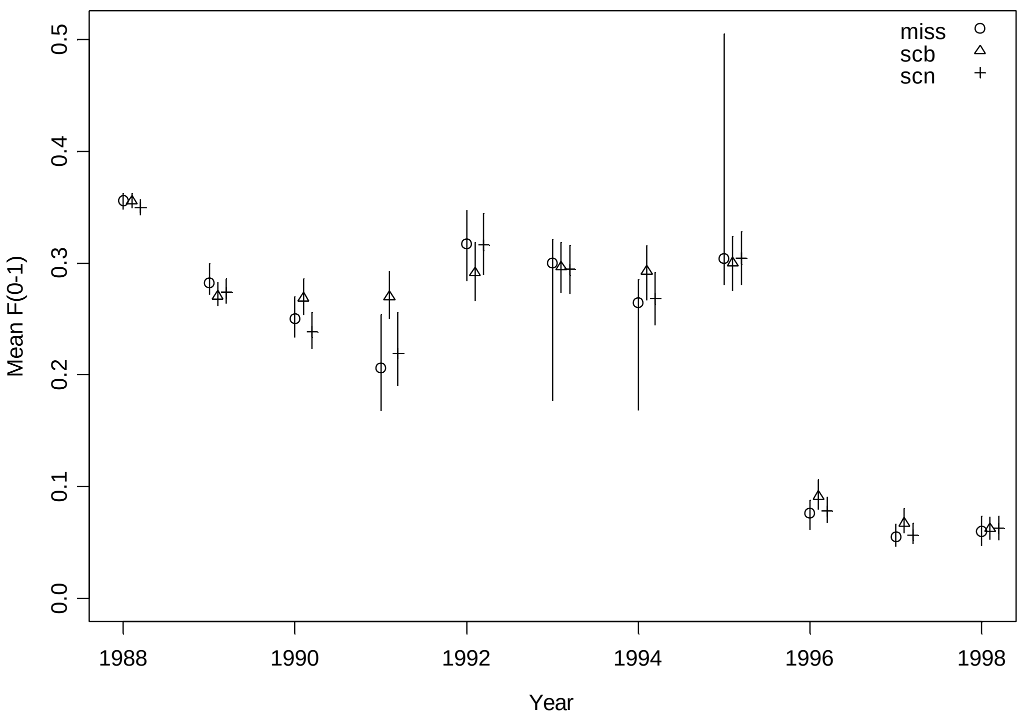


Figure 7b Estimated median and 95% intervals on estimated mean juvenile F (ages 0-1 wr) from 1988 to 1998 conditional on total landings and boostraped estimates of catch at age 1991-1998.

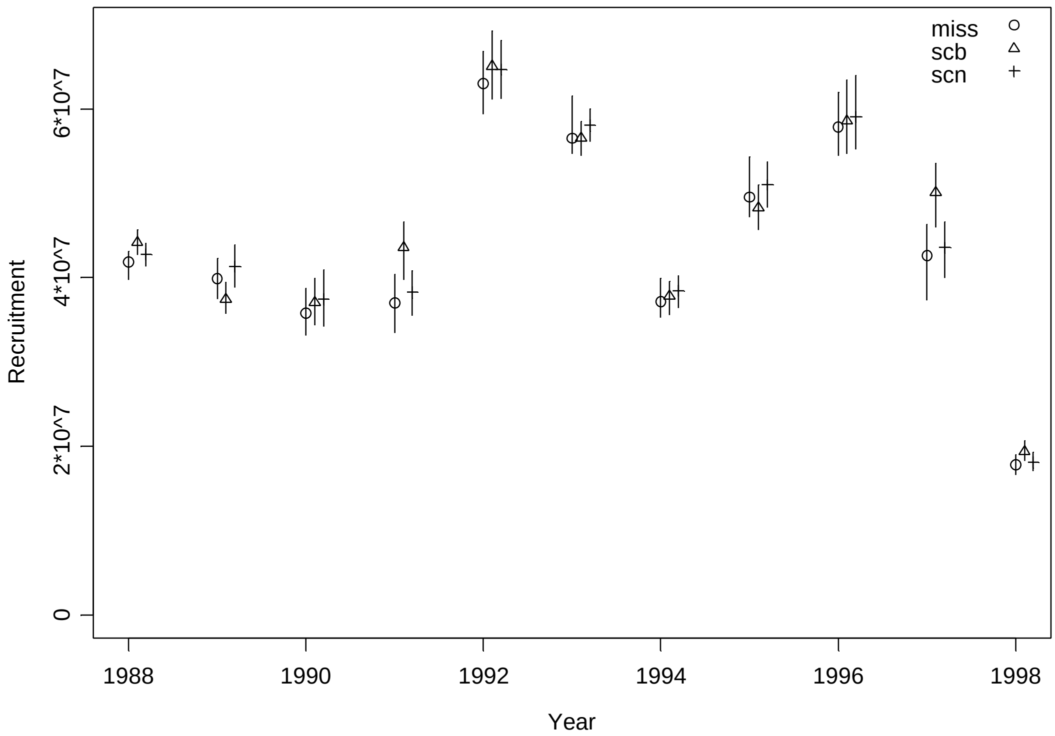


Figure 8 Estimated median and 95% intervals on North Sea hering recruitment 1988 to 1998 conditional on landings and bootstrapped estimated catch at age 1991-1998.

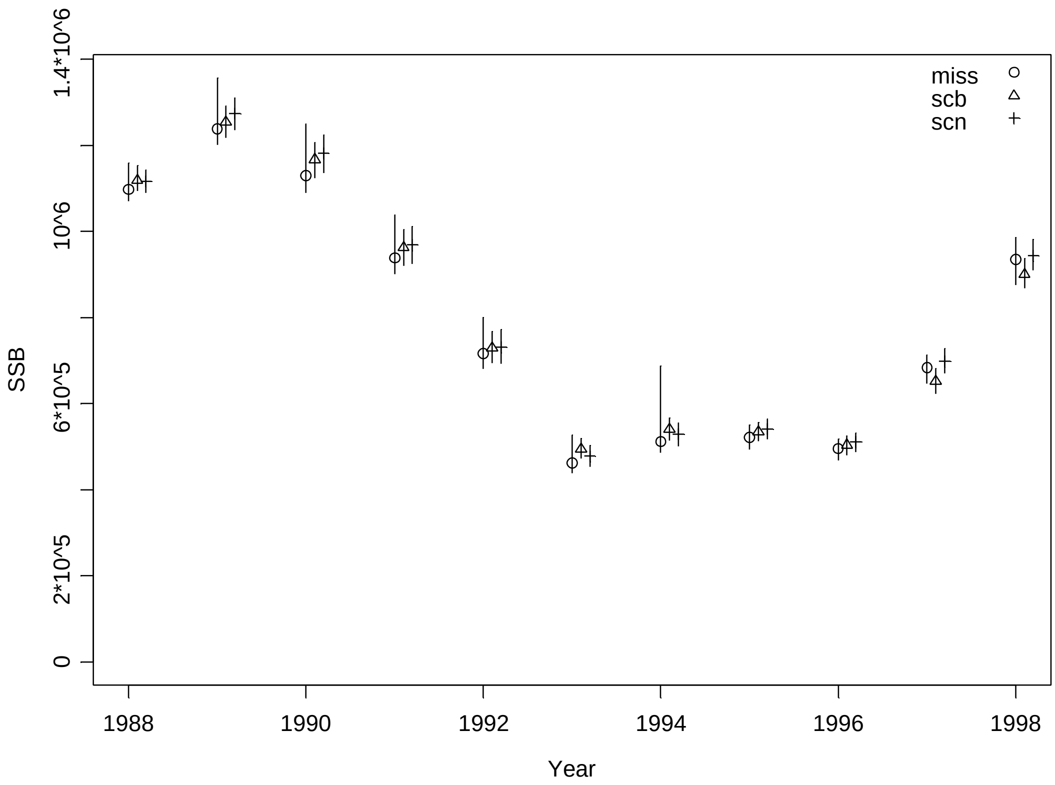


Figure 9 Estimated median and 95% intervals on North Sea hering spawning stock biomass (SSB) 1988 to 1998 conditional on landings and bootstrapped estimated catch at age 1991-1998.

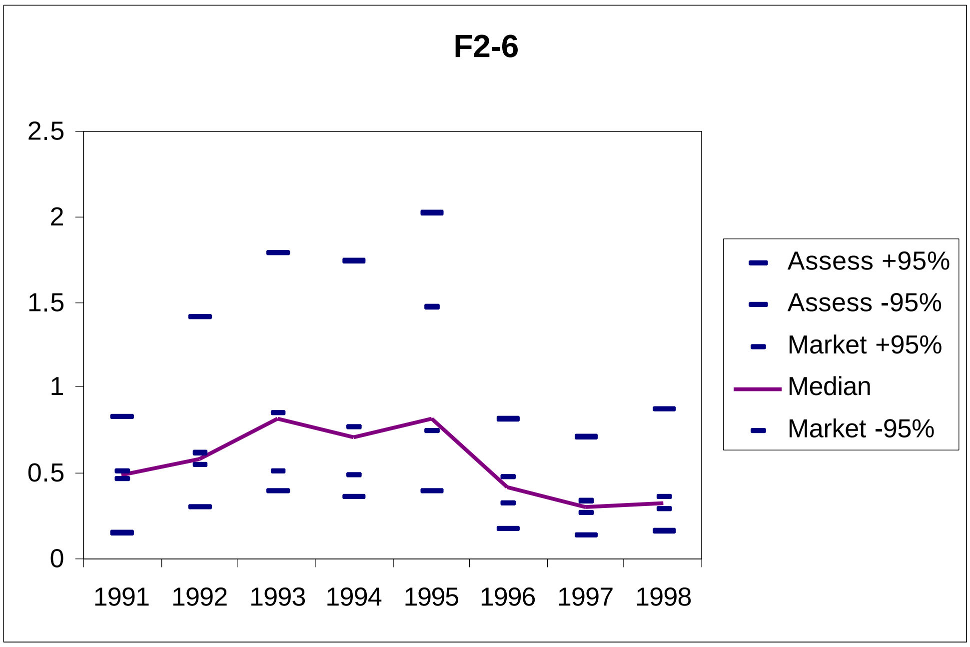


Figure 10 Median $F_{2\text{-}6}$ for herring 1991 to 1998 with 95% intervals estimated by variance – covariance historic uncertainty from the ICA assessment (Assess +-95%) and the contribution of market sampling data estimated by bootstrap assessments (Market +-95%).

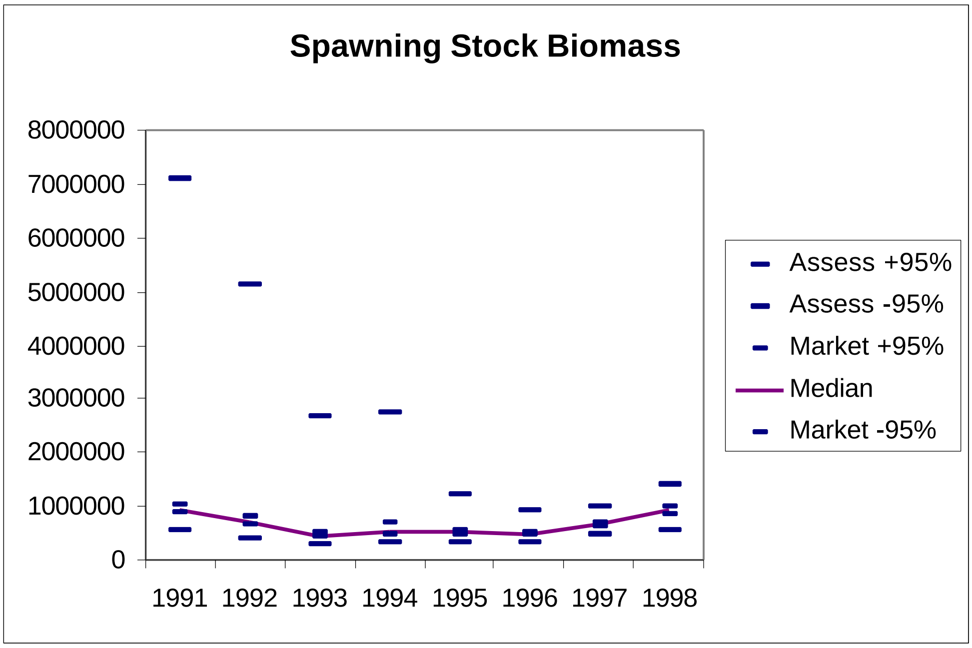


Figure 11 Median Spawning Stock Biomass for herring 1991 to 1998 with 95% intervals estimated by variance – covariance historic uncertainty from the ICA assessment (Assess +-95%) and the contribution of market sampling data estimated by bootstrap assessments (Market +-95%).

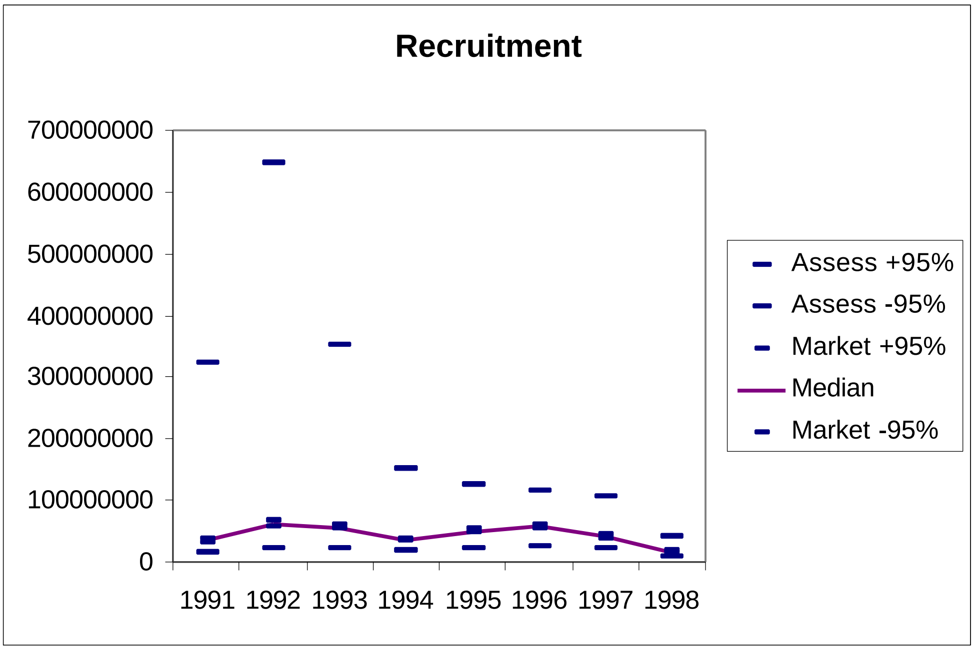


Figure 12 Median Estimates of recruitment (0wr) for herring 1991 to 1998 with 95% intervals estimated by variance – covariance historic uncertainty from the ICA assessment (Assess +-95%) and the contribution of market sampling data estimated by bootstrap assessments (Market +-95%).

Sampling Uncertainties of the Main Fish Caught on the Spanish Atlantic Coast

Sampedro P.[1] and V. Trujillo[2]

1. BP.130, 15001. A Coruña (Spain); 2. IEO, BP.1552, 36200 Vigo (Spain)

Abstract

Two different methods were applied to calculate the uncertainties related to sampling procedures and estimates of age composition for the main southern stocks of fish in Spanish Atlantic Iberian waters for 1999: one based on an analytical approach using the delta method and another based on simulation techniques using non-parametric bootstrap. These methods give estimates of age structure and its precision for stocks, in this case for main pelagic and demersal species. The procedures are a very useful tool to analyse the precision of estimates for the whole stock, quality control, and can even be used for the optimisation of the different sampling schemes. The statistic used in both methods is the coefficient of variation (CV), and in both cases results show that coefficients of variation for the most important reference ages give relatively good precision for estimates under the present sampling network. In spite of this, there is enough variation to indicate that there is room to optimise the actual sampling scheme. The authors support the bootstrap method as there are fewer constraints to the analysis.

Keywords: Anchovy, Sardine, Mackerel, Horse Mackerel, Blue Whiting, Four-Spot Megrim, Megrim, Black Anglerfish, White Anglerfish, Hake, Iberian Peninsula, Delta method, Bootstrap, Uncertainty, Coefficient of Variation, Sampling.

Introduction

Catch-at-age is a common parameter used to evaluate exploited fish stocks, and thus it is important to obtain good estimates of these values. Various methods have been used to calculate the uncertainty of these estimates. The first papers that tackle the problem of uncertainty of biological sampling, undertaken in fish markets and the laboratory, date back to the end of the seventies. In 1978, M. R. Holden used coefficients of variation by age class to evaluate the precision of sampling levels, and these were reported in an internal report by S. Flatman (1990). Coefficients of variation (CV's) can offer a form of quality control for information gathered in sampling nets, and may also be used as an entry for the assessment of the population dynamics of these species.

Historically several methods have been used to calculate CV's. Within the analytical method family, the delta method has been used to estimate the optimum number of otoliths per age class in cod stocks (*Gadus morhua*)(Baird and Maguire, 1983). The general procedure for the delta method is described by Flatman (1990). More recently, the delta method has been used to find the optimum levels at which to sample sardines (*Sardina pilchardus*) (Jardim, 1999). Being parametric, this method makes certain assumptions, violation of which can have implications for the results. An alternative is to use a non-parametric simulation method that does not rely on these assumptions. The bootstrap is one such non-parametric method (Elfron and Tibshirani, 1993). This technique has been used previously to calculate catch by age class in fisheries (Pelletier and Gros, 1991; ICES 2000a, EMAS 2000).

Both of these methods also allow detection of critical variation points, a feature that renders them powerful tools for establishing quality control criteria for national sampling networks that are involved in estimating length and age classes of catches (National Sampling Programmes (NSP))(Pestana et al., 1998; FIEFA, 2000; EMAS, 2000; SAMFISH 2000). Likewise, they allow comparison of relative variation between the different national sampling schemes when compiling international data, and the values can be used in quantitative optimisation procedures for national and international networks.

Comparison of both methods has been applied to the main commercial fish species submitted to fishing regulation that are captured in the Atlantic waters of the Iberian Peninsula. They have been used for both pelagic and bathypelagic species – anchovy (*Engraulis encrasicolus*), sardine (*Sardina pilchardus*), mackerel (*Scomber scombrus*), horse mackerel (*Trachurus trachurus*), blue whiting (*Micromesistius poutassou*) – and demersal species – four spotted megrim (*Lepidorhombus boscii*), common megrim (*Lepidorhombus whiffiagonis*), black anglerfish (*Lophius budegassa*), white anglerfish (*Lophius piscatorius*) and hake (*Merluccius merluccius*).

Materials and methods

Data

The data used in this study primarily originate from the sampling information network ("Red de Información de Muestreo" or RIM, FIEFA) that the Spanish Institute of Oceanography (IEO) has in the Atlantic rim fish markets, and from biological sampling carried out during 1999 by the laboratories involved. Basically, the information used refers to landings carried out at the different ports (in tonnes), length class distributions and the contribution of the age-length keys for the species concerned obtained from otolith readings (Table 1). The assumption is made that there is no error in this information, and therefore that the analysis reflects the "internal variation" of the sampling system chosen.

The sampling scheme developed, both for the length and age classes, is stratified random sampling taking into account the following parameters: gear, space (statistical division of the International Council for Exploration of the Seas (ICES)), time (trimester) and the factor length class, for age sampling.

Coefficient of variation (CV) was the chosen statistic to measure the relative variation of individuals by age classes, and is defined as: the relationship between error and the corresponding mean. This statistic includes both variation arising from sampling of lengths and from assignment of ages. The CV is therefore a powerful yet simple tool from which to determine the precision of the estimates and allows a quality control of any stratum considered.

Methods

Two methods were used to estimate CV's for age classes (CVa):

1. Analytical method based on the delta method (Flatman, 1990; Thompson, 1992; Jardim et al., 2001). This is a parametric method, where the CVa is defined in relation to the number of individuals per age class, which is based on the breakdown of total variation of individuals by age class into two components of variation: one, due to variation of ages through lengths (CVaf) and the other, due to variation of lengths through ages (CVlf).
2. Numerical method based on bootsrap simulations (Efron and Tibshirani, 1993; Jardim *et al.*, 2001). This is a non-parametric procedure , where the CV's in this work are defined in relation to the median (this being the most robust centralisation statistic which can also give us an initial idea of bias in the estimates). To obtain the breakdown of the corresponding CVaf and CVlf each of the components has in turn been fixed (sampling of length or age), and bootstrap replicates of components whose variation is to be determined have been generated. The number of individuals per length class has been obtained from 40,000 bootstrap replicas (200X200 replicates from length samples and length-age codes). In order to carry out these simulations an algorithm has been created in the R environment (Ihaka and Gentlemann, 1996).

Results

In this paper we have combined stratus in order to give a global perspective of the uncertainty associated with the main fish species. We have been consistent with the sampling design throughout the entire internal analysis process. In the end we have opted for selecting the same age classes used by the ICES Work Evaluation Groups (ICES 2000b; ICES 2001a,b) for the species considered (known as “real age classes”). We have therefore excluded the first age classes not completely recruited for some species, and for groups of species we have excluded all ages included in the “plus” group.

Relatively high CV values at distribution tails should not be of concern (Typical “U” shape), as there can be a scaling effect for less represented ages. For this reason, the so called “reference ages” have been studied in more depth (Highlighted in bold in Tables 2 and 3). These are of particular relevance, both from the point of view of the fishery (these are the most relevant ages for landings), and from the point of view of evaluation and management of these resources. It is these age classes that are used to represent the historical tendencies of the fishery exploitation rates, and are also used to establish adequate exploitation levels to be recommended for the management of these resources.

Comparative results for pelagic and bathypelagic species obtained using both of these methods are shown in Table 2 and Figure 1. Both associated components of variation are provided in order to show the level of influence they may have on the total CV.

In the case of anchovy, sardine and mackerel, CV’s obtained for reference ages using both methods are always lower than 20%, and most of the values estimated by the bootstrap method are below 10%. On the other hand, horse mackerel with reference ages between 0-12 have high CV values, only lower than 20% for ages 2-5 when the delta analysis method is used, and ages 4-10 if the bootstrap method is used. There is a considerable difference between results for the two methods in this species. In the case of blue whiting there are also differences between the two methods which occur for all reference ages.

Comparative results for demersal species obtained using both of these methods are shown in Table 3 and Figure 2. In the case of both species of megrim, with the exception of reference age 2, CVs values are mostly below 20%. For black anglerfish, values for reference ages determined using both methods are under 18%. For white anglerfish , differences between the two methods are considerable with values estimated for most reference ages using the delta method, with the exception of ages 7-8, being extremely high. Finally, values for hake found using both methods are below 20%, with those using the bootstrap method all being below 12%.

Conclusions

There are clear differences between methods when estimating numbers of individuals for pelagic and epipelagic species, and this is fundamentally due to the different way amplification functions are applied to the methods. There are also differences in the rate of uncertainty for the age groups calculated using the two methods.

In the case of anchovy, sardine and mackerel the CV's derived from bootstrap for most reference ages are low and equivalent to those found using the delta method. On the other hand, horse mackerel have high CV values and show big differences between methods, and highlight the variation present in the ages used in the tests. In the case of blue whiting, CV values are high in the older age classes, with differences between the methods in all ages.

Differences between methods are not as clear in demersal species. It may be that these species better meet the assumptions of the delta analysis method. For both megrim species (except for age 2), black anglerfish and hake values estimated for references ages using both methods are below 20%. On the other hand, for white anglerfish most ages are highly variable, especially the early ages, and all variation is higher using the delta method.

Looking at reference ages for all species grouped, with the exception of horse mackerel, blue whiting and white anglerfish , CV levels are below 20%. We can therefore conclude that, in general, levels of uncertainty associated with the main ages of species captured by the fleet are acceptable. If we look at the relative contribution to the variance by age class obtained by sampling lengths or ages, we find no clear pattern amongst methods or species. We must therefore use individual species analyses in each case.

Due to the fact that in some cases differences between methods have been found, we must conclude that there must be a careful choice of method used to calculate uncertainties for age classes. We consider the non-parametric bootstrap method to be the most adequate as there are no assumptions to be made, a factor that makes this method easier to apply internationally. On the other hand, this method requires more computational request and replicate samples must be of high quality.

These forms of analysis should be routinely implementes throughout the process of obtaining essential parameters in population dynamics, and we should stop considering cath-at-age matrices to be exact. Knowing the level of uncertainty, as well as allowing us to estimate the levels of error of information obtained, will allow the establishment of quality controls and the optimisation of sampling schemes.

Acknowledgements

We would like to thank all those involved in gathering the data: Informers and samplers at the fish markets, boat building associations, port authorities, other associations, and colleagues from the IEO involved in the biological studies, mostly in the interpretation of growth of species which is the focus of this study. We thank Maria Sainza for her collaboration in those aspects relating to growth. We are also grateful to our colleagues from research institutes in the Basque country (AZTI) and Portugal (IPIMAR) for their contribution of information. We particularly thank our friend and colleague E. Jardim without whose help this work would not have been possible.

References

- Baird, J. 1983. *Sampling commercial catches of marine fish and invertebrates. In* A method to select optimum numbers for aging in a stratified approach., pp. 161–164. Ed. by W.G. Doubleday and D. Rivard. Canadian Special Publication of Fisheries and Aquatic Sciences, 66.

- Cochran, W. 1960. Sampling Techniques. Statistics. John Wiley and Sons, INC, New York.

- Efron, B. and Tibshirani, R. 1993. An Introduction to the Bootstrap. Chapman and Hall, New York.

- EMAS 2000. EMAS Project. Final report, DGXIV / CE.

- FIEFA 2000. Providing a Framework to Improve the Assessment of the Main Demersal and Pelagic Fisheries in Western European Waters - Study Contract 97/059. Final report, DGXIV / CE.

- Flatman, S. 1990. *The evolution of biological sampling targets and the precision of estimated age compositions with reference to demersal fish sampling in ICES Divisions VIIa, VIIe and VIIf and g*. Technical report, Ministry of Agriculture, Fisheries and Food. Directorate of Fisheries Research. Fisheries Laboratory Lowestoft (England).

- Gavaris, S. and Gavaris, C. 1983. *Sampling commercial catches of marine fish and invertebrates. In* Estimation of catch at age and its variance for groundfish stocks in the Newfoundland region., pp. 178–182. Ed. by W.G. Doubleday and D. Rivard. Canadian Special Publication of Fisheries and Aquatic Sciences, 66.

- ICES 2000a. Report of the Study Group on Market Sampling Methodology. ICES CM 2000/D:1. 58 pp.

- ICES 2000b. Report of the Northern Pelagic and Blue Whiting Fisheries Working Group. ICES CM 2000/ACFM:16. 307 pp.

- ICES 2001a. Report of the Working Group on the Assessment of Mackerel, Horse Mackerel, Sardine and Anchovy. ICES CM 2001/ACFM:6. 546 pp.

- ICES 2001b. Report of the Working Group on the Assessment of Southern Shelf Demersal Stocks. ICES CM 2001/ACFM:5. 786 pp.

- Ihaka, R. and Gentleman, R. 1996. R: *A language for data analysis and graphics*. Journal of Computational and Graphical Statistics, 5(3):299–314.

- Jardim, E. 1999. *Análise da amostragem de sardinha (Sardina pilchardus) em 1996 e 1997 - Cálculo dos coeficientes de variação do numero de indíviduos capturados por classes de idade*. Relatórios Científicos e Técnicos 37, IPIMAR, Lisboa.

- Jardim, E., Trujillo, V and Sampedro, P. ICES 2001.

- Pelletier, D. and Gros, P. 1991. Assessing the impact of sampling error on model-based management advice: comparison of equilibrium yield per recruit variance estimators. Canadian Journal of Fisheries and Aquatic Sciences, 48: 2129–2139.

- Pestana, G., Pereira, J., Trujillo, V., Santiago, J., Browne, J., and Nichols, J. 1998. Improvement of Data Collection for Stocks Assessment in ICES Sub-Areas VI, VII, VIII, IX and X - Study Contract 94-013. Final report, DGXIV / CE.

- SAMFISH 2000. Improving Sampling of Western and Southern European Atlantic Fisheries Study Contract 99/009. Final report, DGXIV / CE.

- Smith, S. and Maguire, J. 1983. *Sampling commercial catches of marine fish and invertebrates. In* Estimating the variance of length composition samples, pp. 165–170. Ed. by W.G. Doubleday and D. Rivard. Canadian Special Publication of Fisheries and Aquatic Sciences, 66.

- Thompson, S. 1992. Sampling. Statistics. John Wiley and Sons, INC, New York.

Table 1.

Species	Length samples	Measurements	Otoliths	Landings (ton)
PELAGICS				
Anchovy	60	6 392	7 719	8 699
Sardine	336	36 422	2 532	21 121
Mackerel	260	18 224	2 898	41 794
Horse mackerel	582	44 261	836	37921
Blue whiting	349	29 883	1 326	23 897
DEMERSALS				
Four-spot megrim	168	14 140	341	1011
Megrim	120	2 210	412	336
Black anglerfish	258	5 461	730	1 347
White anglerfish	235	4 418	611	1 583
Hake	562	37 084	784	4 247

Table 2. Comparison of the results obtained by Analytical Method and non-parametric Bootstrap, for pelagic and bathipelagic species. It is indicated for each age: catch-at-age (Na), coefficient of variation of numbers at age (CVa), coefficient of variation due age sampling (CVaf) and coefficient of variation due length sampling (CVlf).

		Analytical				Bootstrap			
Species	age	Na (10^3)	CVa	CVaf	CVlf	Na (10^3)	CVa	CVaf	CVlf
Anchovy									
	0	3 034	0.27	0.12	0.24	3 810	0.19	0.12	0.15
	1	**133 460**	**0.11**	**0.01**	**0.11**	**179 873**	**0.09**	**0.01**	**0.09**
	2	**176 658**	**0.05**	**0.01**	**0.05**	**314 755**	**0.01**	**0.01**	**0.01**
	3	**8 502**	**0.09**	**0.07**	**0.05**	**18 512**	**0.10**	**0.07**	**0.08**
	4	296	0.42	0.39	0.14	660	0.48	0.41	0.25
Sardine									
	0	41 531	0.28	0.03	0.28	21 857	0.53	0.03	0.53
	1	73 704	0.08	0.03	0.08	28 929	0.30	0.04	0.29
	2	**46 180**	**0.06**	**0.04**	**0.04**	**28 212**	**0.08**	**0.06**	**0.05**
	3	**43 448**	**0.06**	**0.04**	**0.03**	**44 866**	**0.07**	**0.05**	**0.04**
	4	**51 901**	**0.05**	**0.04**	**0.03**	**64 433**	**0.05**	**0.04**	**0.03**
	5	**34 345**	**0.06**	**0.05**	**0.03**	**45 013**	**0.06**	**0.05**	**0.03**
Mackerel									
	0	217 617	0.07	0.01	0.07	581	0.52	0.10	0.51
	1	53 414	0.11	0.06	0.10	5 653	0.17	0.11	0.14
	2	12 937	0.30	0.07	0.29	13 313	0.15	0.08	0.12
	3	10 947	0.19	0.06	0.18	25 763	0.08	0.06	0.05
	4	**15 935**	**0.11**	**0.04**	**0.10**	**38 488**	**0.06**	**0.04**	**0.04**
	5	**9 490**	**0.08**	**0.05**	**0.07**	**19 498**	**0.06**	**0.05**	**0.02**
	6	**11 406**	**0.08**	**0.04**	**0.07**	**22 038**	**0.06**	**0.05**	**0.03**
	7	**5 536**	**0.10**	**0.06**	**0.08**	**10 136**	**0.07**	**0.06**	**0.05**
	8	**2 370**	**0.12**	**0.09**	**0.08**	**3 890**	**0.12**	**0.11**	**0.06**
	9	1 477	0.14	0.11	0.10	1 956	0.14	0.11	0.09
	10	1 145	0.17	0.12	0.12	1 542	0.16	0.13	0.11
	11	891	0.19	0.13	0.13	1 057	0.19	0.14	0.13
	12	474	0.23	0.18	0.14	610	0.22	0.18	0.14
	13	284	0.43	0.22	0.37	346	0.29	0.28	0.15
	14	79	0.47	0.43	0.20	96	0.49	0.46	0.20
Horse mackerel									
	0	**11 684**	**0.59**	**0.14**	**0.58**	**10 477**	**0.26**	**0.11**	**0.23**
	1	**35 545**	**0.24**	**0.06**	**0.23**	**17 954**	**0.62**	**0.09**	**0.54**
	2	**38 351**	**0.14**	**0.11**	**0.08**	**13 119**	**0.73**	**0.12**	**0.55**
	3	**32 750**	**0.17**	**0.17**	**0.03**	**10 935**	**0.31**	**0.19**	**0.10**
	4	**48 240**	**0.11**	**0.11**	**0.03**	**32 721**	**0.16**	**0.10**	**0.07**
	5	**17 575**	**0.15**	**0.15**	**0.03**	**29 047**	**0.15**	**0.13**	**0.07**
	6	**4 262**	**0.20**	**0.19**	**0.06**	**22 829**	**0.16**	**0.15**	**0.07**
	7	**2 030**	**0.21**	**0.15**	**0.14**	**26 396**	**0.14**	**0.12**	**0.10**
	8	**1 185**	**0.25**	**0.11**	**0.22**	**25 320**	**0.14**	**0.11**	**0.12**
	9	**559**	**0.30**	**0.16**	**0.25**	**13 857**	**0.17**	**0.15**	**0.09**
	10	**279**	**0.48**	**0.18**	**0.45**	**7 999**	**0.21**	**0.20**	**0.14**
	11	**147**	**0.62**	**0.21**	**0.58**	**4 601**	**0.24**	**0.23**	**0.19**
	12	**91**	**0.68**	**0.28**	**0.62**	**3 496**	**0.26**	**0.25**	**0.20**
	13	49	0.97	0.33	0.92	988	0.30	0.27	0.33
	14	51	1.10	0.32	1.06	852	0.33	0.28	0.37
Blue whiting									
	0	12 717	0.23	0.12	0.19	27 276	0.31	0.16	0.25
	1	48 445	0.13	0.07	0.11	111 012	0.24	0.05	0.23
	2	88 381	0.07	0.06	0.04	101 484	0.20	0.07	0.19
	3	**117 455**	**0.05**	**0.04**	**0.03**	**77 869**	**0.16**	**0.09**	**0.14**
	4	**43 073**	**0.09**	**0.08**	**0.03**	**35 024**	**0.15**	**0.15**	**0.10**
	5	**23 999**	**0.11**	**0.10**	**0.04**	**26 510**	**0.19**	**0.16**	**0.15**
	6	**9 402**	**0.15**	**0.13**	**0.07**	**21 546**	**0.31**	**0.10**	**0.28**
	7	**2 135**	**0.30**	**0.27**	**0.11**	**5 630**	**0.42**	**0.23**	**0.28**
	8	472	0.54	0.49	0.22	1 514	0.81	0.40	0.36
	9	921	0.47	0.44	0.17	1 562	0.59	0.29	0.34

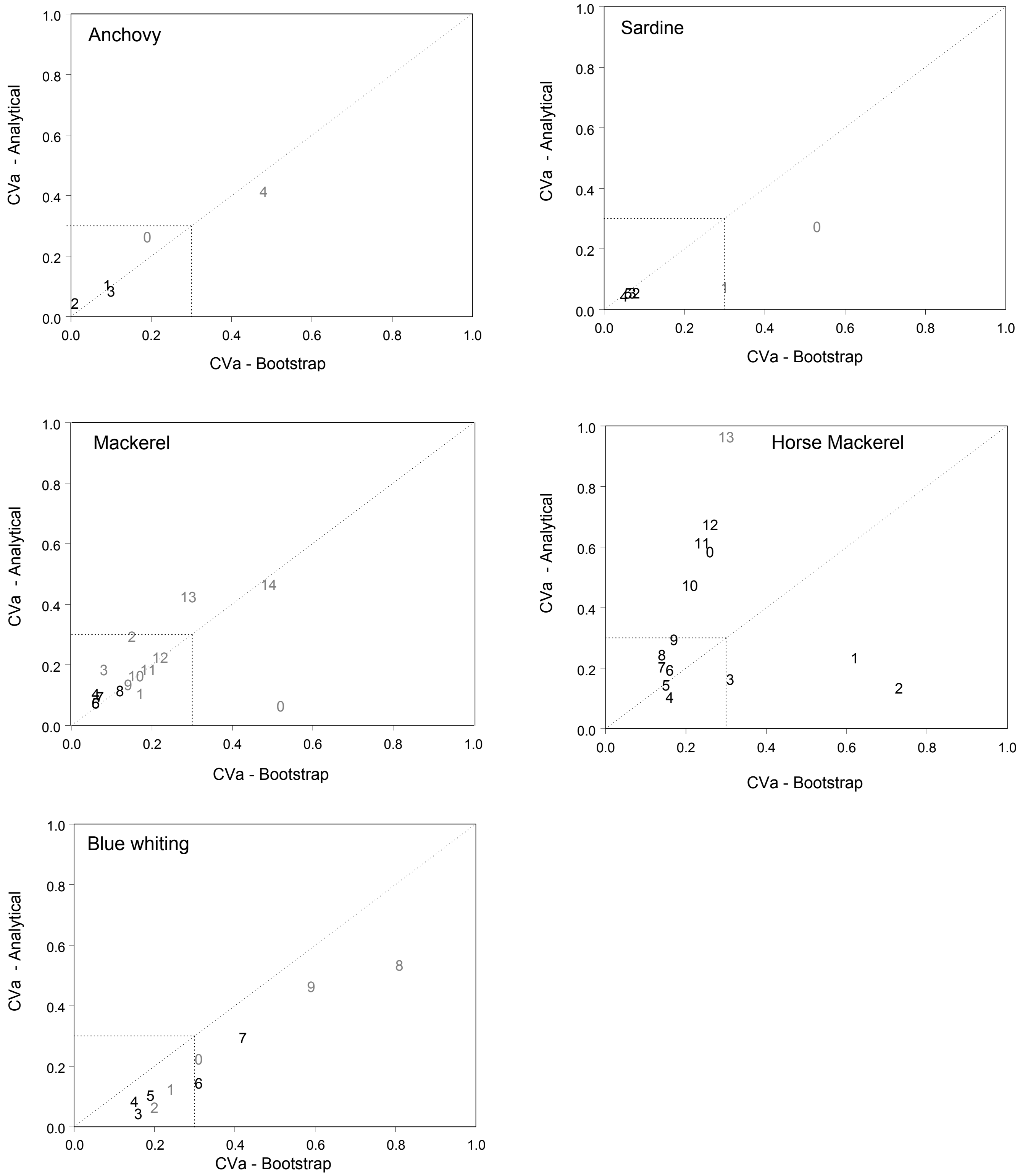
Table 3. Comparison of the results obtained by Analytical Method and non-parametric Bootstrap, for demersal species. It is indicated for each age: catch-at-age (Na), coefficient of variation of numbers at age (CVa), coefficient of variation due age sampling (CVaf) and coefficient of variation due length sampling (CVlf).

		Analytical				Bootstrap			
Species	age	Na (10^3)	CVa	CVaf	CVlf	Na (10^3)	CVa	CVaf	CVlf
Four-spot megrim									
	1	16	0.74	0.43	0.60	27	0.46	0.55	0.09
	2	**295**	**0.34**	**0.31**	**0.12**	**286**	**0.33**	**0.41**	**0.08**
	3	**2 038**	**0.14**	**0.13**	**0.05**	**1 926**	**0.14**	**0.14**	**0.06**
	4	**3 964**	**0.08**	**0.07**	**0.04**	**3 709**	**0.10**	**0.08**	**0.06**
	5	1 552	0.13	0.13	0.03	1 655	0.13	0.13	0.03
	6	238	0.31	0.31	0.05	280	0.33	0.32	0.07
Megrim									
	2	**204**	**0.34**	**0.18**	**0.29**	**205**	**0.29**	**0.18**	**0.22**
	3	**798**	**0.21**	**0.08**	**0.19**	**734**	**0.13**	**0.07**	**0.10**
	4	**621**	**0.17**	**0.10**	**0.14**	**498**	**0.11**	**0.10**	**0.05**
	5	448	0.16	0.11	0.12	324	0.13	0.12	0.05
	6	121	0.22	0.17	0.14	90	0.22	0.19	0.10
Black anglerfish									
	2	1	4.80	0.82	4.73	1	16.58	0.87	5.47
	3	10	0.86	0.29	0.81	16	1.51	0.46	1.08
	4	75	0.22	0.14	0.17	143	0.29	0.17	0.25
	5	156	0.14	0.12	0.08	300	0.20	0.13	0.14
	6	199	0.12	0.11	0.05	272	0.14	0.11	0.07
	7	**171**	**0.12**	**0.11**	**0.05**	**199**	**0.16**	**0.13**	**0.05**
	8	**111**	**0.12**	**0.12**	**0.04**	**93**	**0.17**	**0.14**	**0.07**
	9	**85**	**0.12**	**0.12**	**0.04**	**58**	**0.15**	**0.15**	**0.06**
	10	**63**	**0.12**	**0.11**	**0.05**	**39**	**0.14**	**0.13**	**0.09**
	11	**57**	**0.13**	**0.11**	**0.06**	**35**	**0.16**	**0.15**	**0.08**
	12	28	0.17	0.15	0.07	18	0.18	0.18	0.08
	13	10	0.24	0.23	0.08	7	0.28	0.24	0.14
White anglerfish									
	1	0	9.19	0.30	9.19	0	2.35	1.17	1.38
	2	1	9.27	0.27	9.27	0	1.30	0.84	1.00
	3	**6**	**3.91**	**0.20**	**3.90**	**8**	**0.42**	**0.34**	**0.17**
	4	**7**	**5.03**	**0.19**	**5.02**	**9**	**0.41**	**0.34**	**0.19**
	5	**8**	**1.60**	**0.10**	**1.60**	**8**	**0.28**	**0.16**	**0.21**
	6	**11**	**0.97**	**0.16**	**0.96**	**12**	**0.26**	**0.29**	**0.09**
	7	**21**	**0.30**	**0.14**	**0.26**	**20**	**0.20**	**0.20**	**0.08**
	8	**35**	**0.16**	**0.11**	**0.12**	**36**	**0.14**	**0.12**	**0.05**
	9	43	0.13	0.10	0.08	50	0.10	0.10	0.05
	10	38	0.13	0.11	0.06	41	0.13	0.11	0.05
	11	30	0.13	0.12	0.07	28	0.13	0.13	0.03
	12	27	0.14	0.11	0.08	21	0.15	0.15	0.03
Hake									
	1	1 551	0.14	0.10	0.10	1009	0.14	0.07	0.11
	2	**5 823**	**0.07**	**0.05**	**0.04**	**3091**	**0.12**	**0.05**	**0.10**
	3	**6 847**	**0.17**	**0.04**	**0.16**	**3804**	**0.08**	**0.04**	**0.07**
	4	**2 181**	**0.20**	**0.07**	**0.19**	**1064**	**0.12**	**0.07**	**0.09**
	5	**1 312**	**0.15**	**0.09**	**0.11**	**783**	**0.12**	**0.13**	**0.05**
	6	726	0.16	0.12	0.11	541	0.15	0.14	0.08
	7	295	0.21	0.16	0.12	265	0.22	0.19	0.11

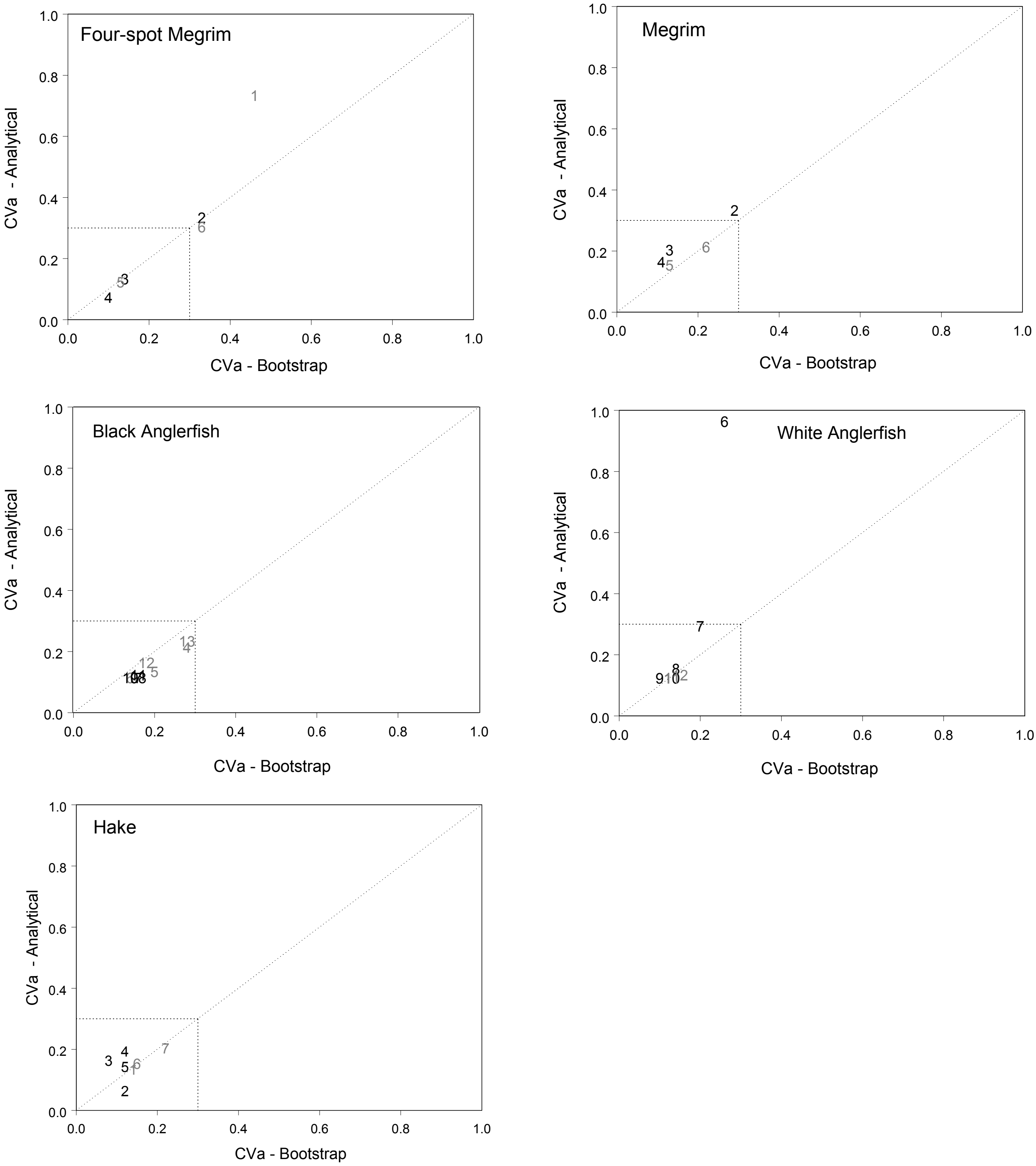
Figure Legends

Figure 1. Relationship bewteen coefficients of variation (CVa) obtained by analytical and bootstrap procedures for the pelagic and bathipelagic species. The bisecting line, where both estimates of CVa have same value, is represented. Dotted lines show CVa values of 0.3.

Figure 2. Relationship bewteen coefficients of variation (CVa) obtained by analytical and bootstrap procedures for the demersal species. The bisecting line, where both estimates of CVa have same value, is represented. Dotted lines show CVa values of 0.3.



Anchovy
Sardine
Mackerel
Horse Mackerel
Blue whiting
CVa - Analytical
CVa - Bootstrap
0.0
0.2
0.4
0.6
0.8
1.0


Four-spot Megrim
Megrim
Black Anglerfish
White Anglerfish
Hake
CVa - Analytical
CVa - Bootstrap
0.0
0.2
0.4
0.6
0.8
1.0