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 $\mathrm{V\hat{a}r}(\hat{X})$ be the estimate of the variance of $\hat{X}$.

The estimate of the coefficient of variation (CV) of $\hat{X}$ is defined as: $\mathrm{C\hat{V}}(\hat{X}) = \dfrac{\sqrt{\mathrm{V\hat{a}r}(\hat{X})}}{\hat{X}}$

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.