ICES WKSDFD 2005

ICES Advisory Committee for Fisheries Management

ICES CM 2005/ACFM:11

Report of the Workshop on Sampling Design for Fisheries Data (WKSDFD)

1–3 February 2005
Pasajes, Spain

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

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer
H.C. Andersens Boulevard 44-46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

Recommended format for purposes of citation:
ICES. 2005. Report of the Workshop on Sampling Design for Fisheries
Data, 1-3 February 2005, Pasajes, Spain. ICES CM 2005/ACFM:11, 78 pp.

Contents

1 General

1.1 Participants

Sondre Aanes, Norway	sondreaanes@imr.no
Richard Ayers, UK England	R.A.Ayers@cefas.co.uk
Ulrich Berth, Germany	ulrich.berth@ior.bfa-fisch.de
Otte Bjelland, Norway	otte.bjelland@imr.no
Liz Clarke, UK Scotland	e.d.clarke@marlab.ac.uk
Willem Dekker, Netherland	willem.dekker@wur.nl
Michele De Meo, Italy	demeo@irepa.org
Fabio Fiorentino, Italy	fabio.fiorentino@irma.pa.cnr.it
Ole Former, Denmark	ofo@dfu.min.dk
Dorleta Garcia, Spain	dgarcia@suk.azti.es
Hans Gerritsen, Ireland	hans.gerritsen@marine.ie
Marianna Giannoulaki, Greece	marianna@imbc.gr
Isabel González Herraiz, Spain	igonzalez@suk.azti.es
Maria Hansson, Sweden	maria.hansson@fiskeriverket.se
David Hirst, Norway	david.hirst@nr.no
Ane Iriondo, Spain	airiondo@suk.azti.es
Ernesto Jardim, Portugal	ernesto@ipimar.pt
Kapiris Kostas, Greece	kkapir@ncmr.gr
Mika Kurkilahti, Finland	mika.kurkilahti@rktl.fi
Sebastiaan Luyssaert, Belgium	Sebastiaan.Luyssaert@dvz.be
David Maxwell, UK England	d.l.maxwell@cefas.co.uk
Hilario Murua, Spain	hmurua@pas.azti.es
Kay Panten, Germany	kay.panten@ish.bfa-fisch.de
Jukka Pönni, Finland	jukka.ponni@rktl.fi
Iñaki Quincoces, Spain	iquincoces@suk.azti.es
Tiit Raid, Estonia	raid@sea.ee
Katja Ringdahl, Sweden	katja.ringdahl@fiskeriverket.se
Paz Sampedro, Spain	paz.sampedro@co.ieo.es
Norbert Schultz, Germany	norbert.schultz@ior.bfa-fisch.de
Josefine Svorin, Denmark	jsv@dfu.min.dk
Valentin Trujillo, Spain	valentin.trujillo@vi.ieo.es
Andrés Velasco, Germany	andres.velasco@ior.bfa-fisch.de
Joël Vigneau, France (Chairman)	Joel.Vigneau@ifremer.fr

1.2 Background and terms of reference

The present workshop is a follow up to WKSCMFD (Workshop on Sampling and Calculation Methodology for Fisheries Data) held in Nantes in February 2004. The WKSCMFD wrote in their report

"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."

The WKSCMFD (ICES 2004a) has in 2004 developed guidelines for sampling and estimation of precision related to the numerous biological parameters collected at the scale of Europe and for any stocks where information is collected. This is linked to the Regulation (EC) N°1639/2001 but has also a more general interest. The purpose of this new Workshop is to look at these analysis or estimations and use the experience gained from it to improve the guidelines.

PGCCDBS (ICES 2004b) has approved the minutes of WKSCMFD (ICES 2004a) and has proposed to ICES/ACFM that a workshop on sampling design for fisheries data, to be held in Pasajes (Spain) in February 2005. At the 92nd ICES Annual Science Conference, it was decided that the Workshop on Sampling Design for Fisheries Data [WKSDFD] (chair : Joël Vigneau, France) will meet in Pasajes, Spain, from 1 – 3 February with the following terms of reference :

a. analyse the estimates of precision of the basic fisheries assessment data by country;
b. on the basis of this analysis advise on sampling strategies including stratification and sampling effort.

1.3 Executive summary

The Workshop participants have much appreciated the good facilities, arrangements and support provided by AZTI, Pasajes and staff members.

The collection of data at the scale of Europe initiated by the DCR is aimed to improve the cooperation between countries and the quality of data collected for scientific use. For most of the data gathered, the collection makes use of a sampling process and thus it is necessary to deal with uncertainty and bias. The first step for quantifying the quality of a data set is to calculate the precision reached, but last year’s WKSCMFD (ICES 2004a) showed that a good precision was a necessary but not sufficient condition. For a given stock and disagregation level, the sampling scheme must certify that the data collected is representative of the population sampled.

Following the WKSCMFD statement that the only way to improve a sampling scheme was to analyse the data and the coefficients of variation, the single question addressed to this workshop was to advise on sampling strategies. The sampling strategy, or in more general terms, the sampling design, covers the complete process of conducting a survey, going from the description of the data collected, the selection of samples, the stratification, the method used to calculate the estimators, the sampling intensity to reach a given precision to end with the quality control of the data collected.

The first point of note was that only a small number of stocks have been analysed by very few countries. To go one step further than the overview tables of last year’s WKSCMFD, the information on stratification, the method used to estimate the age structure and the precision attained when available, are provided at a stock and country level. To complement the summary tables, a description of the different sampling methods is developed but this essential question, especially the elaboration of the age structure, demands a more precise attention for a specific forum.

The information required to begin a discussion on the broad issue of optimising the sampling design was not available. It has therefore been decided to propose an analysis of some case studies going from the exploratory analysis of the samples to the calculation of precision using different methods and different sampling designs. The comparison of methods and sampling designs are important preliminary steps to validate options taken by different countries.

Following last year’s workshop description of methods for calculating precision, a certain number of statistical tools have been developed or are under construction in different countries. Instead of letting all the countries cope with the difficult and long-running task of implementing a complete statistical tool by themselves, the group gave support to the idea of developing an “open source” common tool. The group draws attention to the important impact of such a process promoting convergence and improvement.

The only advice on sampling strategies, as requested by the principal term of reference, is to incite all the countries to investigate and analyse their data the same way as shown here with the case studies. Any data user can see with the lengthy table of Annex A, at a stock level, how the sampling is done and to what intensity. The two complementary workshops WKSCMFD and WKSDFD applied themselves to give the rules, whilst the harmonisation of sampling is to be done at a Regional and stock level, especially as métier sampling is to be initiated in the future.

Questioning specific issues of the broad process of sampling design remains at a high level among participants. The particularity of this workshop is to include different and complementary skills that enable rich and appreciated discussions. The choice of continuing to meet every year goes against the spirit of a one-off workshop and the issue of sampling design can only be a follow-up process. The group recommends therefore to continue the organisation of such a workshop every year but with more focus on the individual use of one common and integrated tool.

1.4 Precision criteria for vector-type estimators

Various types of parameters are to be estimated in module H and I. The definition of the precision criterion (DCR 1639-2001, section B.4) applies to a scalar type estimator, while many of the parameters are a vector by nature. For some vector-type estimators (weight and length by age, maturity and fecundity), the Regulation states that precision must be calculated for those elements of the vector, that correspond to specifically defined criteria:

DCR 1639-2001, section I.1.c.i – *"For stocks for which ages of individual fish can be read, average weights and lengths for each age must be estimated with a precision of level 3, up to an age such that cumulated landings for the corresponding ages account for at least 95 % of the national landings for the relevant stock."*

DCR 1639-2001, section I.1.c.i – *"For stocks for which age reading is not possible, but for which a growth curve can be estimated, average weights and lengths for each age must be estimated with a precision of level 2, up to an age such that cumulated landings for the corresponding ages account for at least 90 % of the national landings, for the relevant stock."*

DCR 1639-2001, section I.1.c.ii – *"For maturity and fecundity, precision of level 3 must be achieved within the age and/or length range, the limits of which correspond to a 20 % and 90 % of mature fish."*

For other vector-type estimators (catch and discards in numbers by age or length), a predefined sampling intensity was required in the original Regulation (DCR 1639-2001, appendix XV), which is replaced by a precision criterion (DCR 1581-2004, Annex I, section 3.a.i.b, replacing DCR 1639-2001, section H.1.b) from 2005 onwards. Additionally, the Guidelines for Technical Reports (STECF, 2005) call for precisions to be reported. However, neither Regulation DCR 1581-2004 nor the Guidelines spell out, how to apply a scalar criterion to a vector-type estimator.

Preliminary calculations (see section 5) have shown that estimates of catch in numbers by age achieve a high precision for the age classes dominating the catch, which are usually in the middle of the age range; for the youngest and oldest age classes, a low precision is found.

A number of options to solve the problem for vector-typed estimators spring to mind:

1. Apply the precision criterion to each age/length class separately. Clearly, this would place excessive emphasis on rare age/length classes. Conventionally, older age classes are grouped into a plus-group in assessments, minimising their influence on the overall assessment.
2. Define a new criterion along the lines spelled out in DCR 1639-2001 (section I.1.c.i). That is: apply the required criterion to the average precision achieved in the dominant age/length. Unlike DCR 1639-2001 (section I.1.c.i), a meaningful criterion for catch in numbers by age has to exclude youngest age classes too. A logical extension would then be to consider the average precision in the most abundant age classes, constituting, say, 90 % of the catch in total.
3. Apply the criterion to the average of those age/length classes that influence the stock assessment the most (e.g. sensitivity analysis). In VPA-type assessments, this could place more emphasis on the older age (greater length) classes, than the previous option. Clearly, this option makes sense in the long run, but it is not yet obvious what aspect of the assessment (stock abundance, fishing mortality, spawning stock, etc) is to be focused, or how to analyse the sensitivities.
4. Apply the criterion to the average precision achieved in a range of age classes, using the same range as the assessment working groups apply in averaging estimates of fishing mortality (e.g. F2-4). However, adopting this solution for age-structured estimates would not solve the problem for length-structured estimators.

The second option (applying criteria to the average of those length/age classes constituting the most abundant fraction of the catch) although not taking into account the correlation between lengths/ages, is consistent with the remainder of the DCR, and applicable to all vector-type estimators, and is therefore recommended. Clearly, the selection of a suitable range of length/age classes must be sustained over several years and reflect the use that is made of the data. A wide selection of the selected length/age range would assure robustness to altering exploitation levels and/or incidental strong or weak year classes and would not give the false impression that a fraction of the population is not worth sampling.

2 ToR a - Overview of national sampling programmes and procedures for estimating precision

For member states within the EU, regulation EC 1581/2004 states levels of sampling intensities and targets of precision within the sampling programmes. Statistical procedures and sampling programme design often varies from stock to stock, even within single countries. Full details of the sampling programmes and results will be reported by each country, following the Guidelines for Technical Reports (STECF, 2005); the information presented below is not considered to replace or stand in for the official reporting, and intended to highlight general patterns only.

A preliminary overview of samplings by country and stock is presented in detail in Annex A, and summaries of main characteristics are given below.

2.1 Overview of national sampling programmes

The precision targets of regulation EC 1581/2004 give rise to a fundamental shift in the way countries design their sampling programmes and have thereby raised a large number of questions regarding the methods to be used. Member States are in a process of changing the statistical treatment of data collected within their national programme, which is a rather slow process in comparison to the requirements of the EU Regulations.

2.1.1 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, which in itself is often dictated by every day practicalities. Tables 1 and 2 summarize the current sampling strategies, with details on principal methods of sampling (Table 2a), stratification (Table 2b) and location of the sampling (Table 2c). The information contained in these tables updates those in last years report (WKSCMFD, ICES 2004a), filling in additional information, and now focusing on the year 2004.

Table 1. - Overview of 2004 National Programs in respect to calculation of precision levels.

	Stocks included in Appendix XV (EC 1639/2001)	Bel	DK	UK En	Est	Fra	GFR	Gre	Ire	Ita	Lat	NL	Pt	Fin	UK Sc	Sp	Sw	No
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(1)	7	38 (1)	28	14	44	80	13(1)	25	40	16	44	38	17	25
	1.1.1 No of stocks sampled (aged based assessment)	18	28	35	6	21	21	8	32	40	4	10	12	14	22	19	12	20
	1.1.2 No of stocks sampled (length based assessment)	5	11	18	0	17	7	3	12	80	1	1	7	2	2	19	3	4
	1.2.1 No of stocks for which precision is reported (aged based assessment)	5	3	12(2)	0	3	0	8	0 (1)	1	0	7	6	0	0	19	3	6
	1.2.2 No of stocks for which precision is reported (length based assessment)	0	3	0	0	11	0	3	0	0	0	0	0	0	0	19	0	0
	1.3.1 Methods† for calculation of precision (aged based assessment)	A+B	A	A	-	A	-	A	-	A	-	B	A+B	-	-	A+B	B	BM
	1.3.2 Methods† for calculation of precision (length based assessment)	-	A	-	-	A+B	-	A	-	-	-	-	-	-	-	B	-	-
2. Discard sampling	2.1 No of target stocks sampled	13	20	79	0	18	23	35	30	72	1 (2)	0	17	2	30	16	8	0
	2.2 No of target stocks for which precision is reported	0	0	17	0	9	0	0	0	0	0	0	17	0	0	6	0	0
	2.3 Methods† for calculation of precision	-	-	A	-	A	-	A	-	-	-	-	A+B	-	B	GLM	-	-
3. Other biological parameters (SMALK)	3.1 Number of stocks sampled	2	24	45(3)	7	15	24	14	0 (2)	80	13 (3)	21	10	14	6	59	16	17
	3.2 Number of stocks for which precision is reported	0	0	c	0	15	0	14	0	0	0	0	10	0	0	59	0	0
	3.3 Methods† for calculation of precision	-	-	A	-	A	-	A	-	-	-	-	A+B	-	-	A+B	-	-

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

UK En-1: assessment definition of stock used or definition by species and ICES division or subdivision if no assessment
UK En-2: precision routinely supplied to assessments for 12 species but is available for all 35 age based stocks
UK En-3: Sex, weight and maturity. Precision reported for sex based ALKs, growth curves and maturity modelling
France-1: Tropical water stocks excluded
D-1: No of Stocks indicated for which precision calculation is planned but not yet done for 2003 data
IRL-1: Precision work to be commenced in 2004
IRL-2: No. of stocks to be determined following survey in Q1 2004
I-1: at now, biological samples are not routinely used for stock assessment
LTV-1: This includes 5 stocks (5 species) for which analytical assessment is performed, 4 species from Appendix XIII, and 3 species which are not listed in the Regulation
LTV-2: discards of cod, besides seal damaged salmon is being recorded in recent years
LTV-3: for 3 species only weight additionally to length and age is recorded.

Table 2a. Overview of 2004 sampling strategies for age/length composition of commercial landings.

	Stocks included in Appendix XV (EC 1639/2001)	Bel	DK	UK En	Est	Fra	GFR	Gre	Ire	Ita	Lat	NL	Pt	Fin	UK sc	Sp	Sw	No
1	No of stocks sampled (total)	23	39	*60*	7	38	28	14	44	80	*13*	*21*	*40*	16	*44*	38	17	25
2	No of stocks sampled for age composition	18	28	*35*	7	22	21	11	32	40	*12*	*15*	*21*	14	*22*	19	13	20
3	No of stocks sampled with ALK method†	18	28	*35*	0	21	21	8	32	40	*5*	*3*	*21*	10	*22*	19	0	17
3.1	No of stocks with lengths sampled from unsorted landings	12	20	*9*	0	2	28	2	28	na	*5*	*11*	*40*	14	*11*	0	0	17
	No of stocks with lengths sampled from market categories	11	15	*35*	0	19	3	6	16	na	*0*	*3*	*0*	0	*11*	19	0	0
3.2	No of stocks with otoliths sampled independently of length distribution	18	16	*35*	0	21	0	8	0	0	*4*	*3*	*21*	8	*0*	19	0	7
	No of stocks with otoliths sampled from length distribution	0	14	*35*	0	1	21	0	22	40	*1*	*12*	*0*	8	*22*	0	0	10
3.3	No of stocks with otoliths sampled from unsorted landings	7	20	*9*	0	2	21	2	6	0	*5*	*9*	*0*	10	*11*	0	0	17
	No of stocks with otoliths sampled from market categories	11	15	*35*	0	19	2	6	26	0	*0*	*6*	*0*	0	*11*	19	0	0
3.4	No of stocks with otolith samples stratified by length	7	28	*35*	0	19	21	8	32	40	*1*	*3*	*21*	8	*22*	19	0	7
	No of stocks with random otolith samples	11	0	*0*	7	0	0	0	0	0	*4*	*12*	*0*	8	*0*	0	0	10
4	No of stocks sampled with direct method‡	0	0	*0*	7	1	0	3	0	0	*7*	*12*	*0*	0	*0*	0	13	3
4.1	No of stocks with fish sampled from unsorted landings	0	0	*0*	7	0	0	2	28	0	*7*	*9*	*0*	14	*0*	0	6	3
	No of stocks with fish sampled from market categories	0	0	*0*	0	0	0	1	16	0	*0*	*12*	*0*	0	*0*	19	7	0
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	0

Table 2b. - Overview of 2004stratification level of sampling strategies for age/length compositions of commercial landings.

	Stocks included in Appendix XV (EC 1639/2001)	Bel	DK	UK En	Est	Fra	GFR	Gre	Ire	Ita	Lat	NL	Pt	Fin	UK Sc	Sp	Sw	No
1	No of stocks sampled (in total)	23	34	60	7	38	28	14	44	80	13	25	40	16	44	38	17	25
2	No of stocks stratified in time (total)	23	34	60	7	35	28	0	44	80	13	25	40	16	30	38	17	21
	No of stocks stratified by quarter	22	34	10	7	30	28	0	35	46	8	25	0	16	0	38	16	20
	No of stocks stratified by time unit shorter than quarter	1	0	60	0	5	0	0	9	0	3	3	40	0	30	0	1	1
	No of stocks stratified by time unit longer than quarter	0	0	0	0	0	0	13	0	34	2	0	0	0	0	0	0	0
3	No of stocks stratified in space (total)	23	34	60	7	2	28	11	44	80	6	0	40	16	30	38	17	22
	No of stocks stratified by ICES/FAO division	23	12	0	0	0	24	14	5	0		21	0	0	0	0	0	0
	No of stocks stratified by ICES subdivision / GFCM geographical sub-area	23	22	60	7	2	4	0	39	80	4	21	24	16	0	31	17	4
	No of stocks stratified by space unit smaller than ICES subdivision / GFCM geographical sub-area	1	0	5	0	0	0	0	9	0	2	0	24	0	30	7	0	18
4	No of stocks stratified by gear (total)	23	6	45	7	13	12	6	44		5	0	40	10	30	38	10	15
5	Overstratification* :																	
	No of stocks for which this has not been analysed	18	34	0	7	15	28	8	0	0	0	0	40	16	0	-	17	23
	No of stocks for which this is experienced to be a problem	4	-	6	0	6	-	3	7	0	-		-	-	0	-	-	2
	No of stocks for which this is experienced NOT to be a problem	1	-	47	0	17	-	0	37	0	-	21	-	-	0	-	-	0

Table 2c. Overview of 2004 locations of sampling of commercial landings.

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

2.2 Magnitude and precision of the 2004 sampling programme

A preliminary overview of samplings by country and stock is presented in detail in Annex A

The Regulation applicable for 2004 (DCR 1639-2001) lists predefined sampling intensities for each stock, for the number of samples and the sample sizes, in relation to the magnitude of the landings by country. Figure 2.2.a and b shows this relation for the length-composition sampling; Figure 2.2.c and d show the same for the age-composition sampling.

The revised Regulation for 2005 (DCR 1581-2004) and the Guidelines for Technical Reports (STECF, 2005) call for precisions to be reported. Figure 2.2.e and f show the relationship between preliminary estimates of precision and the number of samples or the number of fish sampled, for length-composition sampling. Figure 2.2.g and h show the same for the age-composition sampling.

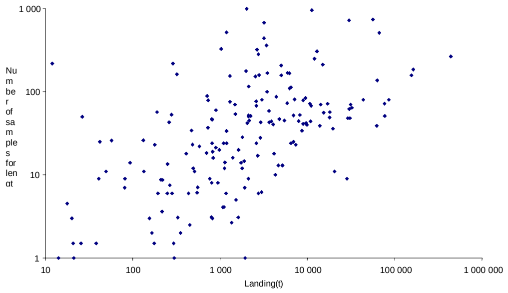


Figure 2.2.a Relation between the landings per country per stock and the number of samples taken for the length composition of the catch.

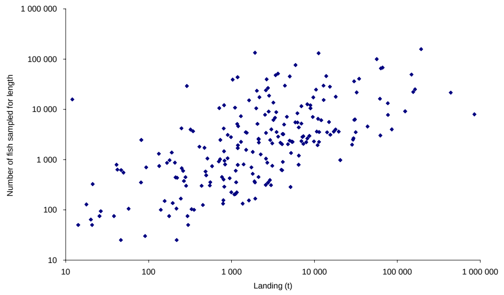


Figure 2.2.b Relation between the landings per country per stock and the number of fish sampled for the length composition of the catch.

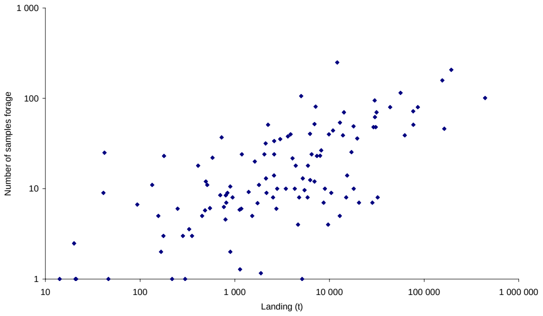


Figure 2.2.c Relation between the landings per country per stock and the number of samples taken for the age composition of the catch.

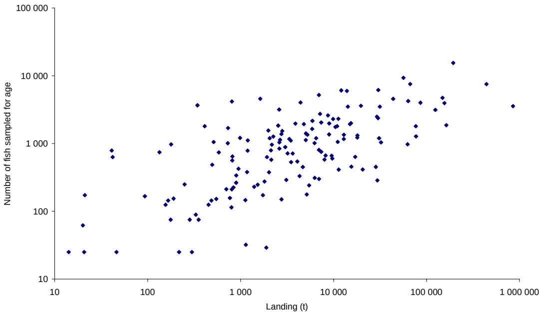


Figure 2.2.d Relation between the landings per country per stock and the number of fish sampled for the age composition of the catch.

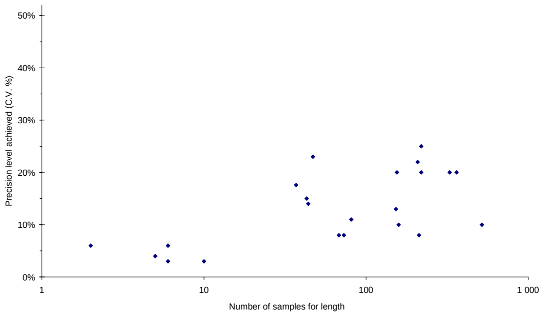


Figure 2.2.e Relation between preliminary estimates of precision and the number of samples taken for the length composition of the catch.

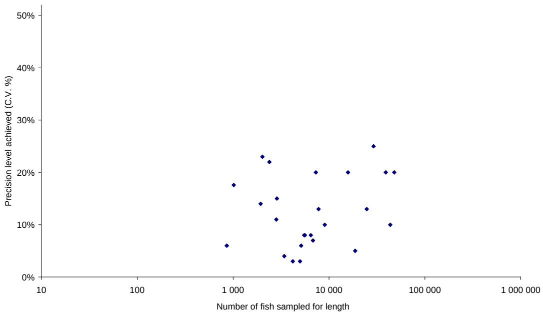


Figure 2.2.f Relation between preliminary estimates of precision and the number of fish sampled for the length composition of the catch.

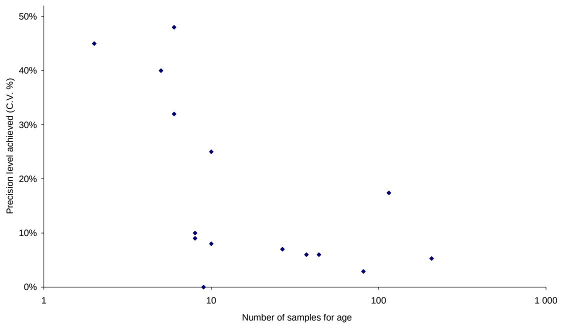


Figure 2.2.g Relation between preliminary estimates of precision and the number of samples taken for the age composition of the catch.

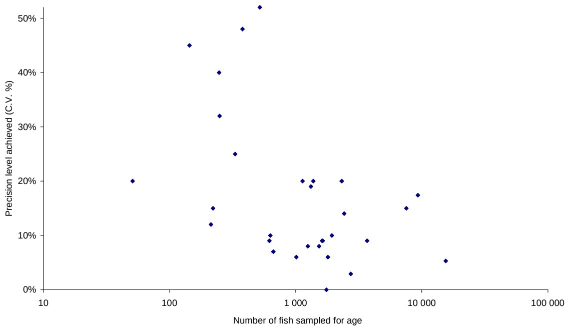


Figure 2.2.h Relation between preliminary estimates of precision and the number of fish sampled for the age composition of the catch.

3 ToR b - Sampling methods

Developing a sampling design to answer one question by carrying out a survey by one laboratory is relatively straightforward for an expert. Developing a sampling design for a multipurpose, multinational and various environment issue is more complex. The sampling design covers all the sampling process from the definition of the sampling units to the quality control of the data. Every step of the process is obviously variable from one country to another and even within one country. Instead of trying to disentangle the multiple combinations of particularities and propose a catalogue of solutions, the group has proposed to focus on the convergence points. As the question to answer is the same everywhere, estimating the length/age structure of the landings, the important further step is to consider the statistical methods used to collect the samples.

In the purpose of estimating the length or age structure of the total landings of one species/stock, the question arises on what is the choice of methods available. From section 2 compiling the strategies used in the different countries and from the literature, the choice appears to be limited to three principal methods. Taken from the book written by Quinn and Deriso (1999), we propose the following description

3.1 Simple random sampling

“Under a simple random sampling protocol, a simple random sample is taken from a population of individuals. The population may be a commercial catch of fish, a survey catch of fish, or an entire fish population.[…]It is assumed that fish are sampled independently and have the same probability of being sampled. Each fish in the sample is aged, and the resultant frequency of each age in the sample is calculated”

This method is referred in the summary table 2a of section 2 as “direct method”. The authors draw the attention on the fact that this method may be difficult to carry out in practice, because of the danger to unintentionally select certain components of the population. To ensure a representative sample of the population a particular attention must be given on the sampling plan and particularly on partitioning the sampled population into strata. This statement on the representativity of sampling is worth for every methods.

3.2 Two-stage random sampling

“Under two-stage sampling, two assumptions are necessary.

> 1. A simple random sample from the catch is taken in the first stage. The sample is classified into length intervals.
> 2. A random subsample from each length class is taken for aging in the second stage. The first well-known allocation for the age subsample is fixed allocation, where a constant number of fish are aged from each length class. The second is proportional allocation, where the number of fish aged is selected proportional to the length frequencies. […] Any other allocation can also be used. After ages are obtained, age frequencies are classified by length and age into an age–length key."

This method is referred in the summary table 2a of section 2 as "Stocks with otoliths sampled from length distribution"

3.3 Separate length and age samplings

"Subsampling data for age are frequently not available, but there is often the need to apply age–length keys to length frequency data to obtain age composition estimates"

This method is referred in the summary table 2a of section 2 as "Stocks with otoliths sampled independently of length distribution"

3.4 Discussion

There may often be factors, such as time, area, gear… which divide the population into sub-populations (groups/strata) and we may expect the age composition to vary among the different sub-populations. This has to be accounted for, at the moment of drawing a sampling design in order to obtain a representative sample of the population, by adopting a stratified sampling. The stratified sampling can be applied to each of the method described above, the corresponding formulae being applicable to a single strata and the combination of all the strata calculated as described by Cochran (1977).

The objective of the sampling is of particular importance at the moment of drawing up a sampling plan. The example discussed in section 5 shows that the precision decrease with the scarcity of the length/age classes. In VPA-type assessment, estimation of F at the oldest true age (before the plus group) is used to start the backward recursion, and therefore it would be wise to avoid too much imprecision on the older ages.

Another discussion was to consider whether the method was different by sampling at sea or under auction. A paper by Aanes and Pennington (2003) considers the age composition of the commercial catch of Northeast Arctic cod from a sample of clusters of commercial fishing trips. From Cochran (1977), it is said that the two-stage sampling with units of unequal sizes is a natural extension of the one-stage sampling with cluster units of unequal sizes.

One of the principal outcome of the Aanes and Pennington paper is that fish sampled from the same trip (i.e. from a "cluster") tend to be more similar in age than those in the total catch. Sampling few units from a large number of clusters is then preferable than taking samples of all the units of a few clusters This consideration is very important at the moment of removing otoliths in the purpose of building an ALK by the method of separate length and age samplings. It is clear that the otoliths has to come from different locations and different time of the year to be as representative as possible of the sampled population.

More work needs to be done in how to elaborate and handle age–length keys. The different formulas to implement related to the method used and the consequence on the final estimation needs to be discussed in a further workshop.

4 Available tools

Tools are currently being developed by some institutes or research groups to investigate the sampling and to calculate precision in the biological data. Here is a presentation of the tools used during the Workshop.

4.1 Casa

- **Main objectives**
  This tool called Casa is documented in WD "Notes on R package "casa" (catch-at-age sampling analysis) by E. Jardim, P. Sampedro and V. Trujillo", allows to estimate catch-at-age and precision levels of length and age sampling in terms of coefficients of variation.
- **Methods**
  Methods applied are: analytical and non-parametric bootstrap approaches that were described in the report of WKSCMFD (2004). In the future, it is planned to include other new approaches, *e.g.*: Bayesian methods.
- **Environment**
  This package is developed in R environment with S4 classes (objected oriented program).
- **Advantages and disadvantages**
  The use of this tool is quite automatic, though it requires knowledge of programming and statistics. On the other hand, the input data needs to be set-up in the right format in advance.
- **Other applications**
  Casa could be used as a tool for quality control and to optimise sampling design.

4.2 Excel spreadsheet

- **Main objectives**
  The principal objective of the Excel spreadsheet is to decompose the estimator used for exploratory analysis and the precision estimation of a length structure.
- **Methods**
  The formulas used are described in a Working Document available in annex of 2004 WKSCMFD (Vigneau and Mahevas, 2004)
- **Environment**
  The template is a Microsoft Excel spreadsheet.
- **Advantages and disadvantages.**
  The main advantage is that the description of the method is very didactic.
  The disadvantage comes from the non-dynamic particularity of the spreadsheet
- **Other applications**
  The dynamic use of this method is developed in the S-plus package described below.

4.3 S+ package

- **Main objectives**

  The package proposes different modules that allows

  Importation of data

  Exploratory analysis

  Single strata precision calculation using analytical method

  Complex strata precision calculation using bootstrap calculation

  The main objective of the package is to investigate in the sampling data collected and propose a precision estimation by length or age class, a precision estimator between and within strata using the 90% rule (see section 1.5).

This package has been used to explore the data of the two case studies analysed in section 5.

- **Methods**

  The length data analysed are samples composed of number at length and total weight sampled. Various descriptive variables are included in the data set like gear, métier, date, area, harbour, to allow different combination of strata.

  For one given species, the software is designed to analyse:

  Sampling of the vessel catch sorted by commercial categories.

  Sampling of unsorted vessel catch.

  Sampling of commercial categories.

  The formulas used are described in a Working Document available in annex of 2004 WKSCMFD (Vigneau and Mahevas, 2004).

  The age data analysed are Age–Length keys. The formulas used are those developed by Kimura (1977).

- **Environment.**

  The software has been developed in a Guide User Interface (GUI) with S-Plus version 6.1.

- **Advantages and disadvantages.**

  The advantage of such a tool is to help the people in charge of sampling to scrutinise the data. It allows the search of outliers, errors, possible bias, empty cells, and distortion of sampling intensity per strata and proposes disaggregated results per strata.

  The exploratory analysis, with the GUI environment is very dynamic and easy to use.

  The disadvantages are:

  The impossibility to run the software out of S-Plus environment and with version anterior to 6.1.

  The bootstrap of the combination of the length structure and the age–length key respecting the process of the age structure elaboration for assessment purpose gives erratic results.

- **Other applications.**

  In IFREMER (France) where the S-Plus package is used, the sampling data is centralised in a database that is a data management warehouse. The S-Plus package is based on a special extraction of this centralised database. The future development of the centralised database will be the inclusion of all the modules of the S-Plus package in a press-button like software.

4.4 Comparing ALKs

- **Main objectives**

  This tool has been presented to this Group as a WD: “A simple model to compare ALKs by Hans Gerritsen”.

  This tool allows to compare ALKs and it can be used to fulfil missing values.

- **Methods**

  An ALK is a multinomial data set, which can be expressed as binomial data in the form of pairs of consecutive age classes. This allows a binomial GLM to be fitted to the data. The model might also be expanded to take into account the changes over time in the ALK by fitting the same data using a GAM but with the factor ‘time’ as a smooth term. This would improve precision and remove the necessity to stratify aged data by quarter.

- **Environment**

  A routine to fit and check the model is developed in the R-environment but the method is simple and it could easily be performed using any statistical package. An excel spreadsheet is also available which fits the logistic curves to ALK data and can be used to fill in missing data but cannot be used to compare ALKs.

- **Advantages and disadvantages**

  The main advantages are the simplicity and robustness of the model and transparent assumptions. The main disadvantage is that many age classes are represented twice in the dataset (as the age classes in consecutive pairs overlap). Therefore the degrees of freedom of the null model will be overestimated. The binomial approach is therefore not entirely appropriate.

- **Other applications**

  Another application is to check ALKs and how they might be combined.

4.5 Bootstrap method for biological parameters INBIO

- **Main objectives**

  The main objective of this tool (called INBIO) is to estimate uncertainty of some biological parameters as growth (age vs. length and age), maturity (by length and age), sex-ratio and length-weight relationship curves using an automatic procedure. Documented in WD “A simple tool to calculate biological parameters by P. Sampedro, M. Sainza and V. Trujillo”

- **Methods**

  The methods used to determine the uncertainty of biological parameters is non-parametric bootstrap on different models and fits, namely:

  - Growth at age (vs. length and weight):

    von Bertalanffy. Non-linear estimation w. minimum least squares (Gauss-Newton (G-N)).

  - Maturity (length and age):

    GLM. Logistic function. Binomial errors w. maximum log-likelihood fit.

  - Length – weight relationship:

    Standard. Non-linear estimation w. minimum least squares (G-N).

  - Sex-ratio (length and age):

    No model. Percentage by length and age. Cubic spline to plot.

  The statistics adopted for each estimate were: median and coefficient of variation. For all the estimates has been plotted their probability profiles (“density’s functions”) and some plots of model’s residuals (model error) to check visually: autocorrelation, homo/hetero-cedasticity, outliers and/or extreme values, linearity and normality.

- **Environment**

  The routine has been developed in R environment.

- **Advantages and disadvantages**

  The main advantages of this tool are:

  - Give estimates of uncertainty levels (precision and bias) for main parameters.
  - Non dependent of any assumption on statistical distribution (non-parametric).
  - Easy to implement and well-known:

    It can be used as a tool for quality control.

    It can be used to optimise sampling designs and sampling intensities.

It can be used for exploratory model analysis.

  - Allow us to compare uncertainty levels between Countries/Institutes.

  One disadvantage of this tool could be the knowledge on methods used.

- **Other applications**

  Other applications of this tool could be the development of quality control and optimisation of sampling design.

4.6 Economic information

- **Main objectives**

  The goal of this tool (called IREPA software) is to estimate the total production of fish, total value of fish production and the average producers' price by species and by area of catches. It is more documented in WD "Italian sample survey for landings statistics: links with the biological sampling of landings by M. De Meo".

- **Methods**

  The program is divided in three main Modules:

  - Module 1: Data Control. It is employed to minimize non-sampling errors.
  - Module 2: Sampling Procedure. It uses the proportional probabilities to size methodology. Sample size in each strata uses Bethel's procedure and it selects units using Hanurav-Vijayan's algorithm (for references, see DeMeo, 2005).
  - Module 3: Estimates and Precision level. This module uses the Horvitz-Thompson estimator, allowed by Hanurav-Vijayan's algorithm, to estimate total production of fish and its total value. It is applied an analytical approach using the Sen-Yates-Grundy model, to calculate the coefficient of variation of the former estimates.

- **Environment**

  This software has been developed in Statistic Visual Basic. This programming language is included and it runs in STATISTICA v. 6 (StatSoft).

- **Other applications**

  This tool, with some modifications, can be applied to the biological sampling too.

4.7 Bayesian analysis

- **Main Objectives**

  To analyse any sampling design in order to estimate catch-at-age and to give a correct measure of precision.

- **Methods**

  A Monte Carlo Markov Chain sampler written in C++ with a Splus interface. The program implements a Bayesian hierarchical modelling approach described first in Hirst *et al.* (2004) and further developed since.

- **Advantages and disadvantages**

  The model can estimate catch-at-age and precision for any known sampling scheme, without the need for filling in missing cells or merging strata. Age reading errors can also be accounted for. There is no need to modify the sampling scheme in order to make a bootstrap useable.

  The main disadvantage is that it is difficult for a non-statistician to understand the methodology, and to evaluate the assumptions.

- **Environment**

  A C++ program within Splus. The Splus interface has been partially rewritten in R in order to improve availability.

Many Countries/Institutes have developed their own applications to manage fisheries' samplings. As some examples of different kinds of tools developed, it is presented:

4.8 Other tools implemented

CEFAS

An important step in assessing a sampling scheme is to compare the distributions of the samples and the landings.

CEFAS is developing a flexible internet mapping server, called iSEA, to display and query its fisheries and environmental data.

The first version of iSEA is available to use on the CEFAS website at: www.cefas.co.uk/isea and an overview article is available at www.cefas.co.uk/coastmap/coastmap6.pdf. This article is reproduced as a WD titled: "CEFAS iSEA Mapping server". A second version of iSEA is due for release in April 2005. This will have improved displays and summaries, variable levels of access for different users and the ability to view the data used to produce each display.

In the future, by linking the CEFAS sampling and landing databases to the mapping server it will be a useful tool to study how representative sampling is in space and time. Also, as iSEA is designed for use on a range of datasets, it could include international data and be used by other countries.

IEO

IEO has developed its own application called SIRENO (Integrated System of Natural Oceanic Resources) for storage and processing of the information from whole fishing activity of Spanish fleets operating in European Atlantic waters.

This software now enables the management of all basic information in fisheries research (landings, efforts, market samplings, biological samplings, observers on board, *etc.*) and gives us all the information processed for the most relevant strata and aggregation level (species, different data periods, gears, areas) including precision levels using analytical and bootstrapping approaches. The reports generated from this application are made in ASCII format and they are easy to import into the most usual commercial software suites.

SIRENO is based on a relational database environment/platform. The software has been developed with ORACLE's tools. It can be executed in Internet/Intranet environments. The database is located in a server computer, under the operative system Unix (Application Sever) and Windows NT (Database Server), i.e.: one computer works as program server and the other is used as a database server. In these servers, there are programs and necessary software to put them automatically in the web. Both equipments are located at IEO's headquarter in Madrid. Users in coastal laboratories, they only need to have access to internet and use a standard I nternet browser. The safety of the program is established with an access policy controlled by a firewall.

Tools	Objectives	Methods	Coded	Advantages	Disadvantages	Availability
Casa	Catch at age estimates and precision	Analytical Bootstrap	R (S4,obj. oriented)	Well structured Fast and Flexible	Learning curve Slightly tough	Open Source GPL 2
Add-on Excel	Exploratory analysis and precision to lengths	Analytical	MS Excel	Didactic	Non-dynamic	
S+ pack.	Exploratory analysis and precision estimates on length and age	Analytical Bootstrap	S+	Visual EDA	Need some more development	Source free shareware
Comp. ALKs	Comparison ALKs Filling missing data	GLM	R	Relatively simple and robust	Binomial approach not entirely appropriate	Open Source GPL 2
INBIO	Main biological estimates and precision and bias	Non-linear (G-N) GLM Cubic spline	R	Simple, Fast and Easy	Methods' understanding	Open Source GPL 2
IREPA	Estimates of total production and total value	Analytical	Statistica Visual Basic	Easy and Friendly	Complex formulation	Proprietor Software
Bayes	Bayesian analysis for catch-at-age and precision	Hierarchical model	Splus, C++	Can analyse and sampling design	Complex theory	Proprietory software

Table 4.1. Summary of presented tools.

4.9 Discussion

The question of the tool is more important than one could think at a first sight. It can even turn to be a central issue in the future. During this workshop, demonstration has been made that running appropriate software is essential to avoid bias, learn about one's own sampling and improve the sampling design It is likely that it would be a waste of time, money and energy if every country was developing its own software for analysing its own data. The convergence of all the sampling methods (see section 2 and 3) into a short list of options will help a lot in using a common tool.

The group has come to the point that the best way to deal with this problem was to develop an open source software on a common platform using the outcomes of WKSCMFD (ICES, 2004a) and the outcomes of this workshop. This software would be a package containing different methods (freeware, shareware, open source, hidden source, …) able to work on the existing sampling designs and covering all the needs. The first step of this process would be to agree on a common data format and start to implement the data frame. Ernesto Jardim (Portugal) has proposed to initiate this concept and build the necessary framework of such a software.

If the use of this kind of tool can be made in such a workshop mixing statisticians and people aware of sampling coming from each of the Member States, the development of the methods have to be made intersessionally. The group draws the attention that letting people develop their own methods for this common platform alone in their lab without any collaboration is not the best solution. A framework between very few statisticians has to be found.

5 Case studies

5.1 Exploratory analysis

5.1.1 Introduction

After the data has been collected, some investigation needs to be carried out to control the quality. Using the S+ package (see section 3) two case studies are analysed, the hake in ICES Division VIIIa and IXc sampled in 1999 and the Baltic cod sampled in 2003.

The first idea is to control the adequation between the sampling intensity and the landings. This analysis can help to improve future sampling for the same purpose by optimising the sampling allocation and point out some possible weakness in the current sampling plan.

The second analysis gives the within strata percentile boxplot of the number of individuals sampled per unit, the sample weight per unit and the individual weight per sample. From the sampling theory stating that each sample must be representative of the whole population sampled, it can be informative for the sampling manager to have a quick look at the median of the number of individuals per sample. In case of a commercial category sampling, the boxplot of the individual weight per sample shows the adequate splitting and the magnitude of overlapping of the categories.

The following step is to draw the samples with the method given in a Working Document by Vigneau and Mahevas (WKSCMFD, 2004a). This method resumes one sample by one point and is able to point out possible outliers, possible misallocation of strata and enables to search for pattern or patches in the sampling that would show a problem in the stratification. This real exploratory of the data ca be done for a given range or the total range of length classes, a given number or the total number of strata, the samples can be sorted according to any available covariable and coloured according to any second covariable. This flexibility of use implies to settle the problem to address before running the analysis. The search of an outlier is done by running the analysis on all the length classes and all the strata, which always corresponds to the first intuitive run. Pointing out differences between two gears in catching small individuals will be another run with only the smallest lengths and the two gears to consider, *etc.*

The precision of a length structure can be resumed with one value (see section 1.4) making it possible to graph the CV's per strata. This graph is informative to check the homogeneity of the sampling and points out potential overstratification and its resulting problems of poor precision or poor sampling in some of the strata.

5.1.2 Exploratory analysis for hake 1999

Length data are from Portuguese and Spain based on market sampling and correspond to the entire landings from the stock. Sampling was carried out by gear category and numbers at age were estimated by vessel. Two gear categories were used: 1) trawls and 2) gill netters, long liners and multigears merged in one category called artisanal fleet.

Sampling intensities vs. landings distribution (Figure 5.1) shows that the sample sizes were representative of the landings.

In the case where the sampling is done by gear categories or métier and the landings are sorted by commercial categories, there is the need to sample each of the landed categories of the vessel in order to estimate the length composition of the sampling unit which corresponds to the sampled trip. Doing so, the resulting sample corresponds to the total landings of the vessel and leads to much heterogeneity in the number of individuals per sample. It is the case for hake and it is shown in Figure 5.2.

The plots of delta values can be studied by looking at the individual values and the general patterns. Figure 5.3 shows that there is a very low delta value (bottom right corner of the graph), regarding an artisanal haul that contains much less fish than expected from the average distribution. The length distribution (Figure 5.4) shows that this is a haul that contained much bigger fish than expected. Omitting this observation and re-analysing data might show differences between the two gear categories.

Re-analysis without the outlier observation is described in Figures 5.5 and 5.6. The pattern of observations in Figure 5.5 shows little evidence of changes in length distribution by quarter as the values are centered on zero in all quarters. Figure 5.5 clearly shows a difference in the distribution of delta values for trawls and artisanals. Artisanals are likely to catch bigger fish than trawls and this effect is stronger in the third quarter.

The CV estimations by quarter, gear and area shows a very homogeneous sampling (Figure 5.6). The CV are approximately at the same level, showing that the sampling effort was suitably distributed between all the strata.

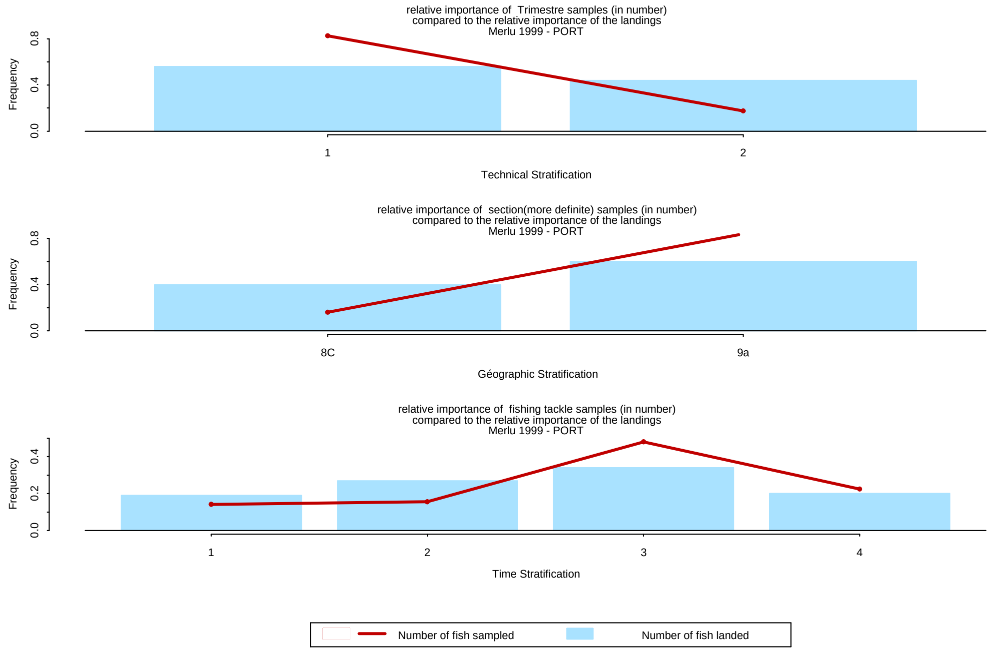


Figure 5.1 Sampling intensities vs. landings distribution.

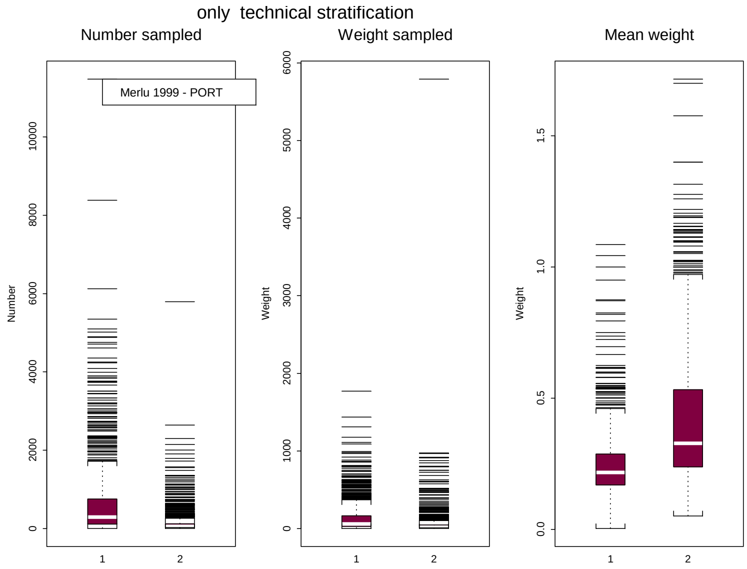


Figure 5.2 Distribution of the number of fish sampled, weight sampled and mean weight per fish in a sample.

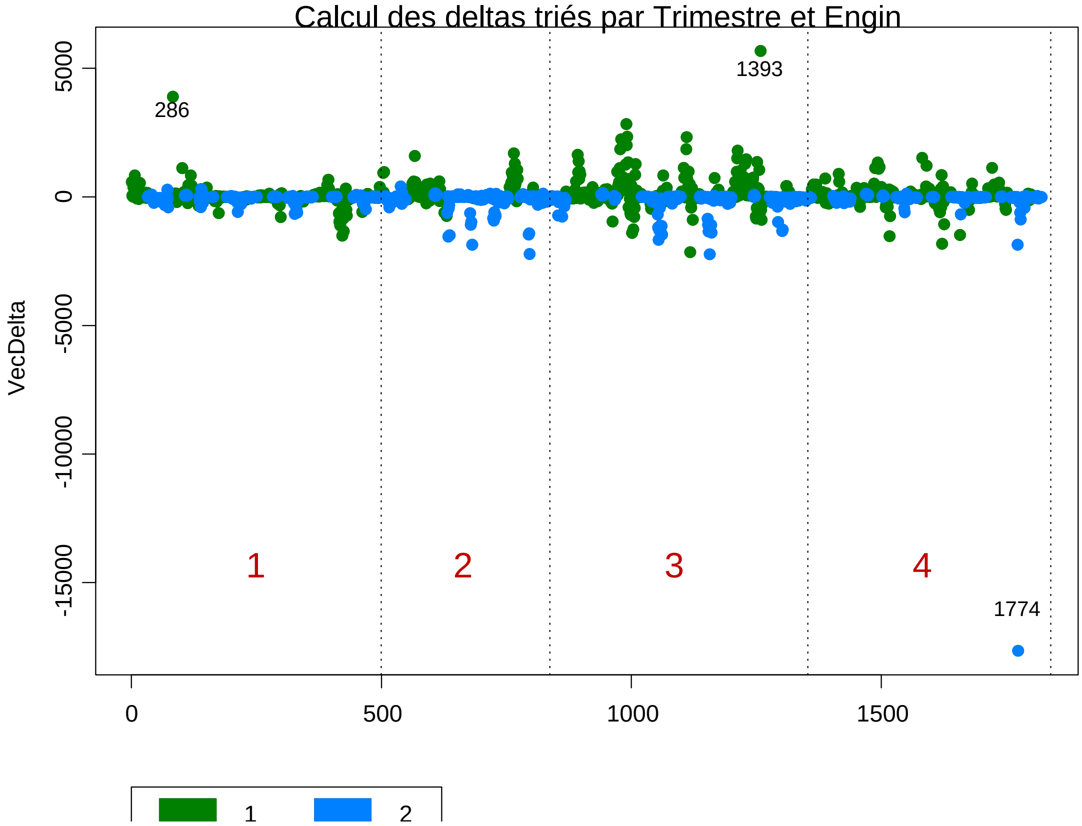


Figure 5.3 Delta value estimations for each haul, by quarter (1-4), coloured by gear (1: trawls, 2: artisanals).

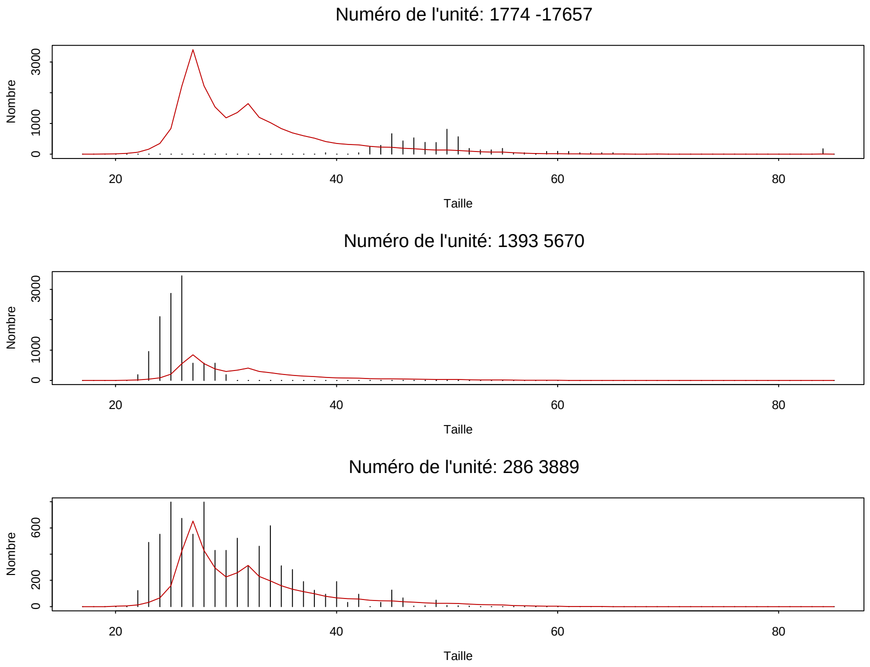


Figure 5.4 Observed (bars) and expected (red line) length distributions for observations with largest delta values.

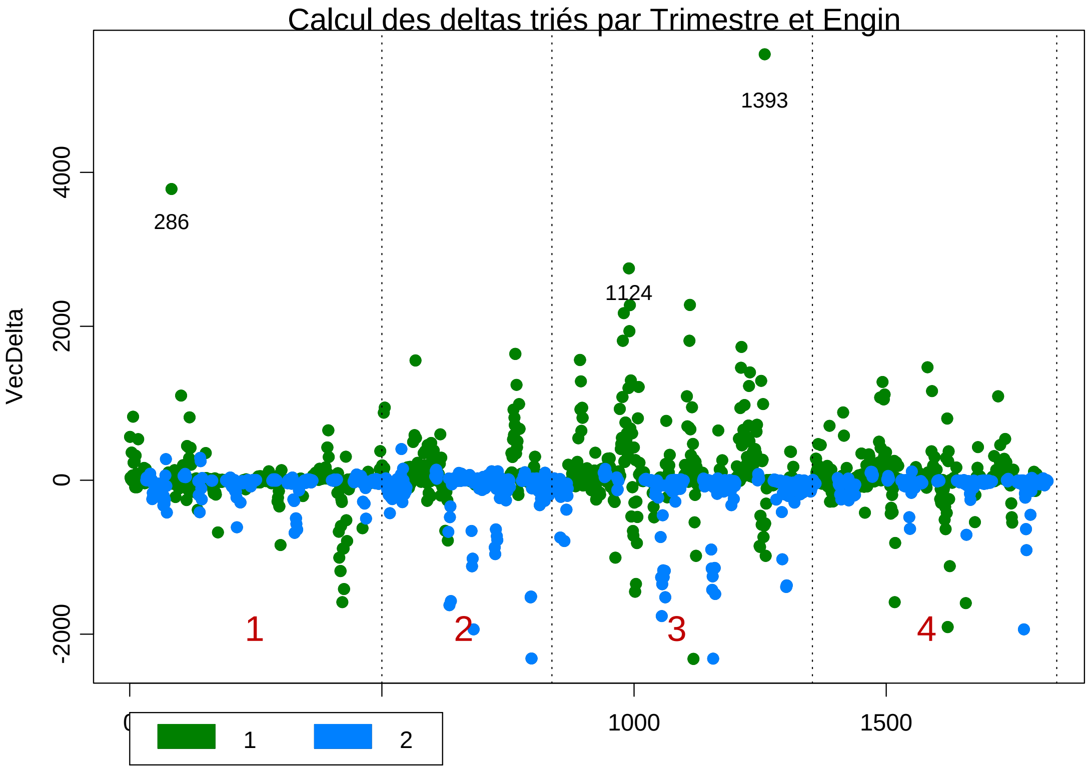


Figure 5.5 Delta value estimations for each haul (without the outlier observation) by quarter (1-4) and coloured by gear (1: trawls, 2: artisanals).

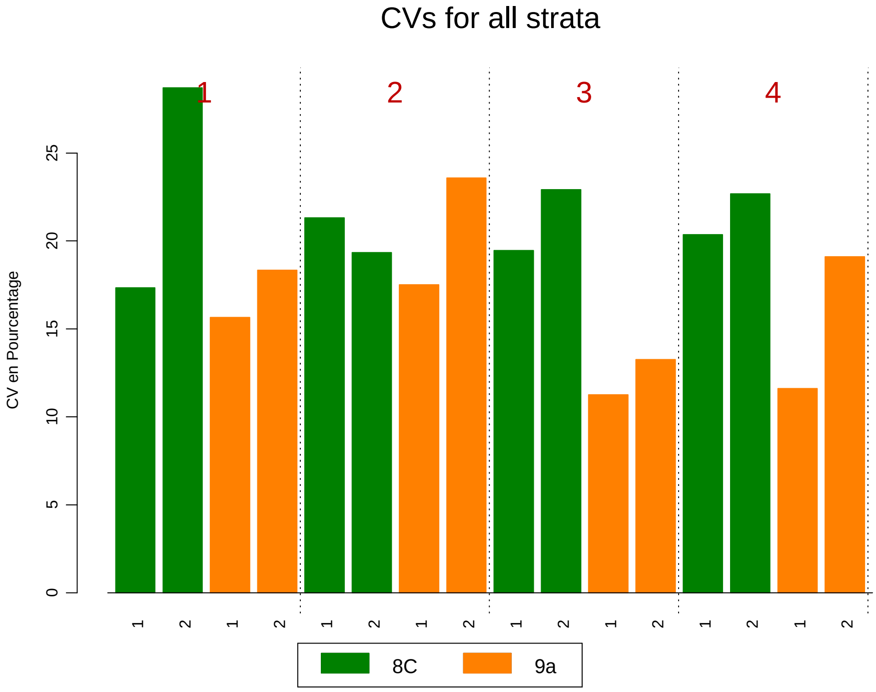


Figure 5.6 CV estimations as % percentage by quarter (1-4), gear (1: trawls, 2: artisanals) per area (ICES division: 8C, 9a) for all strata.

5.1.3 Exploratory Analysis for Baltic cod 2003 - Length samples

An exploratory analysis of length sampling for Baltic cod 2003 was carried out using length data from Swedish on-board sampling with haul as the sampling unit. Sampling is carried out on commercial fishing vessels targeting cod in the Baltic Sea. Within the Baltic, cod is separated into two stocks, one eastern stock (subdivision 25-32) and one western stock (subdivision 22-24). The Swedish sampling is stratified by quarter, area and gear. Area is defined as subdivision (23, 24, 25, 27) and gear defined as otter trawls (OTB), gillnets (GNS) and longlines (LLS). On each trip sampling is carried out on a haul-to-haul basis, both discards and the retained part of the catch is measured. The retained part of the catch is sampled by size category (1-5), while the discards are treated as one category.

Figure 5.7 shows the sampling programme was consistent across the three gear types with the same numbers and weights of landed fish sampled. Figure 5.8 shows sample sizes of discards increased from gillnets, to longlines, to otter trawls. This reflects the amounts discarded by these gears. The mean weight of sampled fish is the same for the three gears but is higher for discards than landed fish.

The plots of delta values can be studied by looking at individual values and general patterns. Figures 5.9 to 5.11 highlight a small number of OTB hauls that contained many more fish than expected from the average distribution. Checking the data set showed these were not data entry errors. They were genuine observations with unusually large catches of small fish. There is also one value with a very low delta value (bottom left corner of graphs), this is a haul with total weights recorded which was not sampled.

The pattern of observations in Figure 5.9 shows little evidence of changes in length distribution by quarter as the values are centred on zero in all quarters. There is no noticeable difference in the delta values for landings and discards, this is surprising because discards have a smaller mean length. On average, within a haul, mean length is 13 cm less for discards (range – 4 to 28 cm). There is a small range in delta for the discards due to a combination of a small length range and less weight being discarded than landed. It is likely that the difference of sampling weight between discards and landings explain the fact that the points corresponding to the discards in Figure 5.9 are all close to 0. Figure 5.10 shows some differences between the gears which support stratification by gear. OTB had lower mean length (larger delta values) and GNS higher mean length.

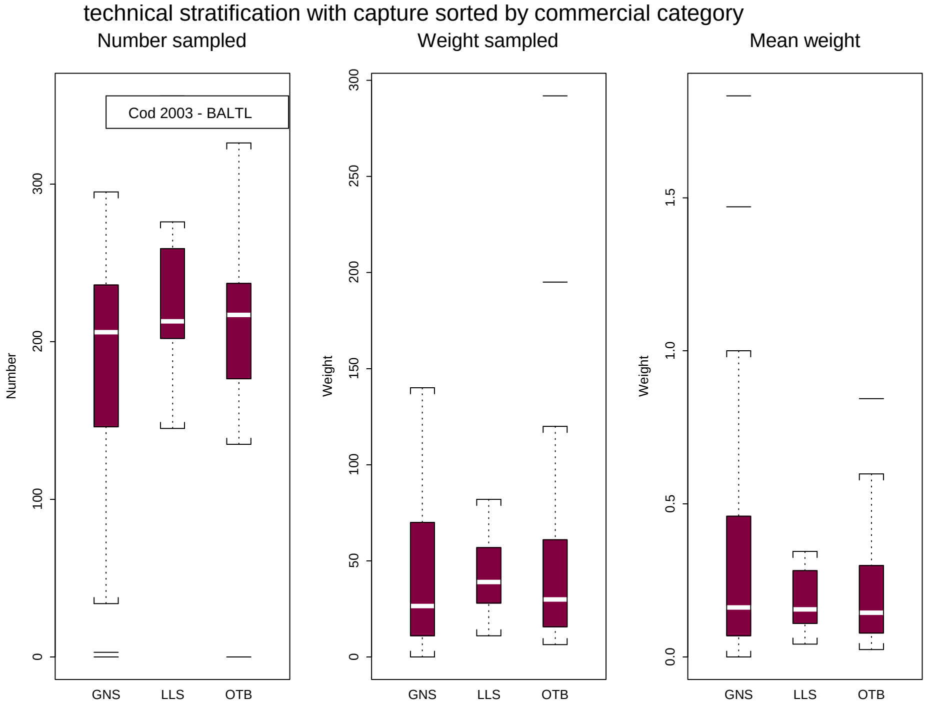


Figure 5.7 Landed fish. Distribution across hauls of the number of fish sampled, weight sampled and mean weight of sampled fish.

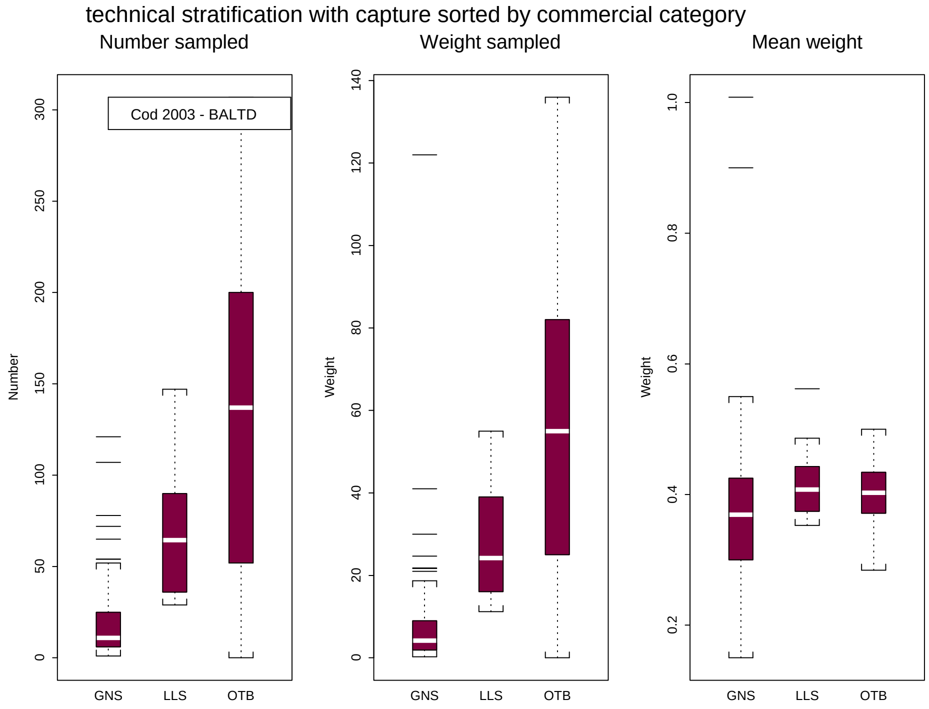


Figure 5.8 Discarded fish. Distribution across hauls of the number of fish sampled, weight sampled and mean weight of sampled fish.

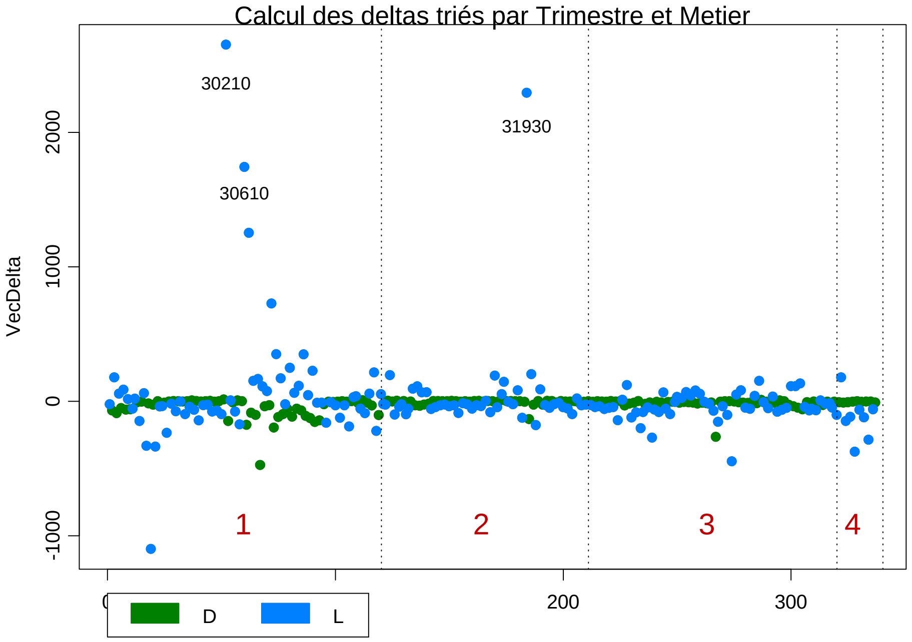


Figure 5.9 Delta value for each haul, by quarter (1 – 4), coloured by discards or landings. High values represent more fish observed than expected for the sample weight or smaller fish than the average distribution.

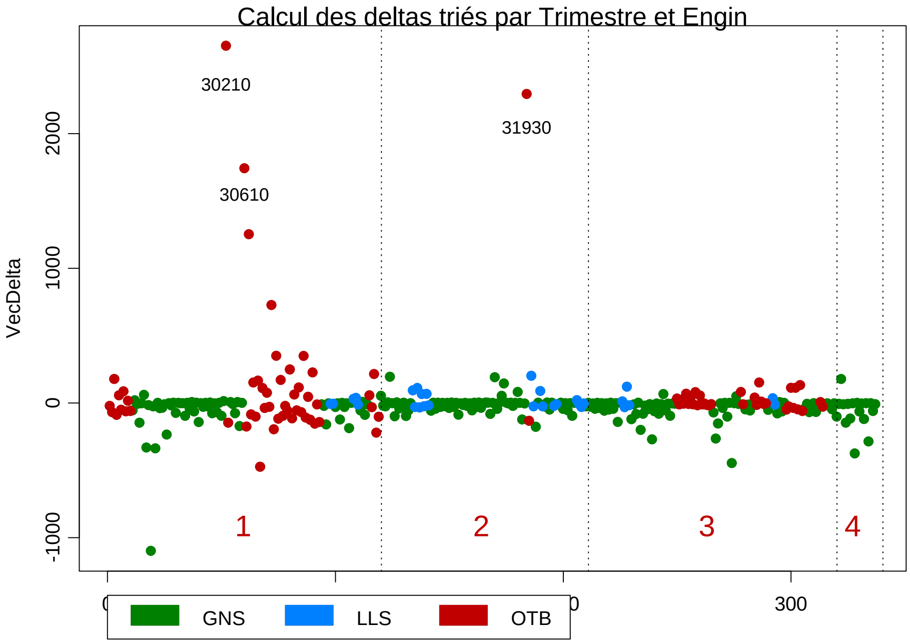


Figure 5.10 Delta value for each haul, by quarter (1 – 4), coloured by gear type.

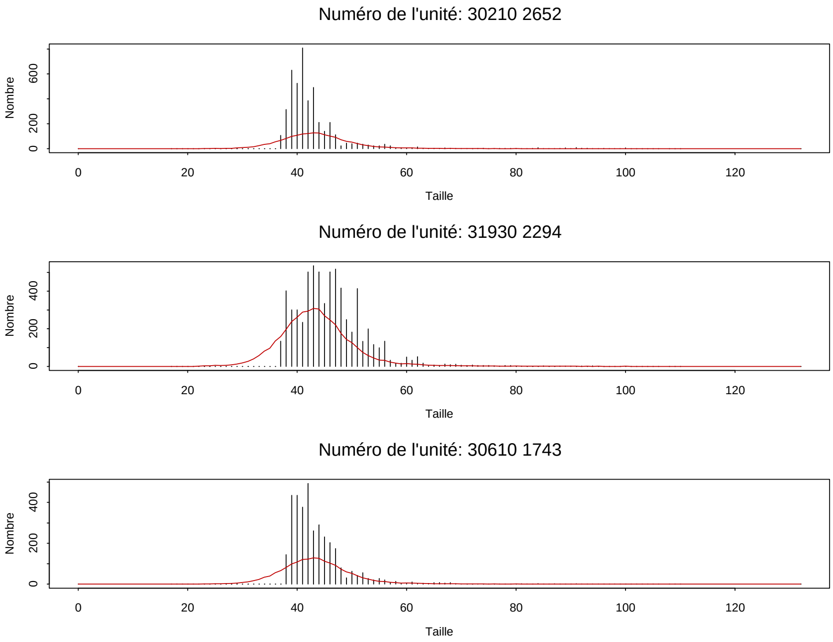


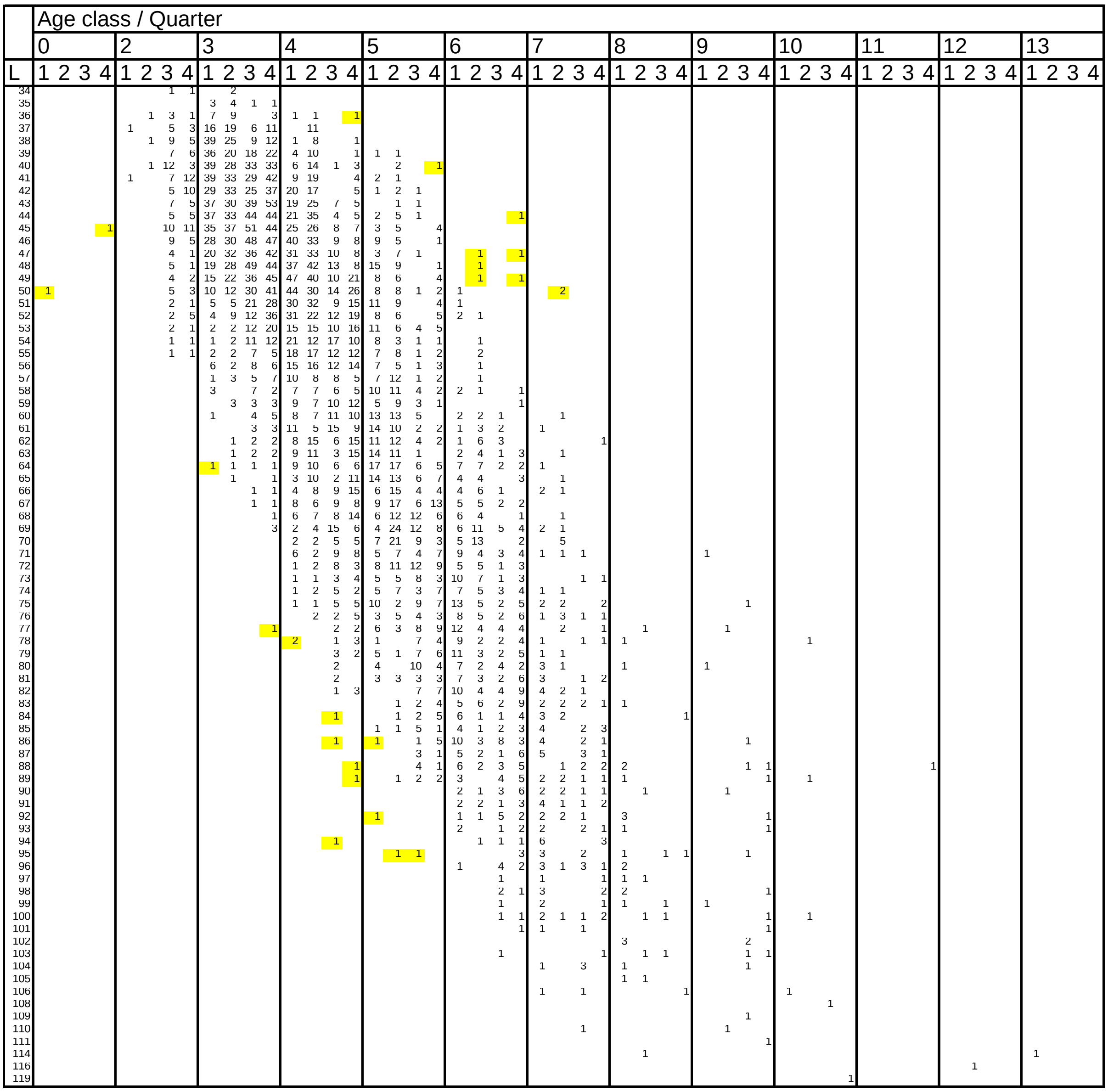
Figure 5.11 Observed (bars) and expected (red line) length distributions for observations with largest delta values.

5.1.4 Exploratory analysis – Baltic cod – aged data

The aged data in the Baltic cod data set was collected with simple random sampling within a market size category and not collected as an age length key, so differences between the areas and gear types can be explored using the tool described in the working document “A simple model for comparing age–length keys and filling in missing data”.

Removal of outliers

The data were summarised as sample numbers at age by quarter. In this way outliers can be detected and removed if they distinct from the main size range at age. The outliers here are to be seen as data that would bring extra noise and uncertainty to the model, not as errors in the data set. Figure 12 shows the outliers that were removed highlighted in yellow.



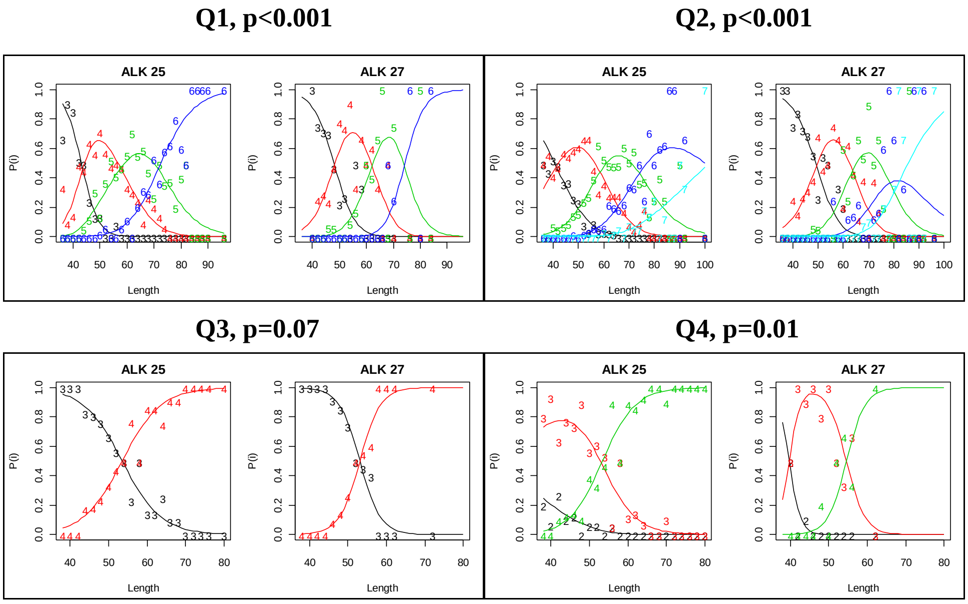
	Age class / Quarter																																																			
	0				2				3				4				5				6				7				8				9				10				11				12				13			
L	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
34							1	1		2																																										
35									3	4	1	1																																								
36						1	3	1	7	9		3	1	1		1																																				
37					1		5	3	16	19	6	11		11																																						
38						1	9	5	39	25	9	12	1	8		1																																				
39							7	6	36	20	18	22	4	10		1	1	1																																		
40						1	12	3	39	28	33	33	6	14	1	3		2		1																																
41					1		7	12	39	33	29	42	9	19		4	2	1																																		
42							5	10	29	33	25	37	20	17		5	1	2	1																																	
43							7	5	37	30	39	53	19	25	7	5		1	1																																	
44							5	5	37	33	44	44	21	35	4	5	2	5	1					1																												
45				1			10	11	35	37	51	44	25	26	8	7	3	5		4																																
46							9	5	28	30	48	47	40	33	9	8	9	5		1																																
47							4	1	20	32	36	42	31	33	10	8	3	7	1			1		1																												
48							5	1	19	28	49	44	37	42	13	8	15	9		1		1																														
49							4	2	15	22	36	45	47	40	10	21	8	6		4		1		1																												
50	1						5	3	10	12	30	41	44	30	14	26	8	8	1	2	1					2																										
51							2	1	5	5	21	28	30	32	9	15	11	9		4	1																															
52							2	5	4	9	12	36	31	22	12	19	8	6		5	2	1																														
53							2	1	2	2	12	20	15	15	10	16	11	6	4	5																																
54							1	1	1	2	11	12	21	12	17	10	8	3	1	1		1																														
55							1	1	2	2	7	5	18	17	12	12	7	8	1	2		2																														
56									6	2	8	6	15	16	12	14	7	5	1	3		1																														
57									1	3	5	7	10	8	8	5	7	12	1	2		1																														
58									3		7	2	7	7	6	5	10	11	4	2	2	1		1																												
59										3	3	3	9	7	10	12	5	9	3	1				1																												
60									1		4	5	8	7	11	10	13	13	5		2	2	1			1																										
61											3	3	11	5	15	9	14	10	2	2	1	3	2		1																											
62										1	2	2	8	15	6	15	11	12	4	2	1	6	3					1																								
63										1	2	2	9	11	3	15	14	11	1		2	4	1	3		1																										
64									1	1	1	1	9	10	6	6	17	17	6	5	7	7	2	2	1																											
65										1		1	3	10	2	11	14	13	6	7	4	4		3		1																										
66											1	1	4	8	9	15	6	15	4	4	4	6	1		2	1																										
67											1	1	8	6	9	8	9	17	6	13	5	5	2	2																												
68												1	6	7	8	14	6	12	12	6	6	4		1		1																										
69												3	2	4	15	6	4	24	12	8	6	11	5	4	2	1																										
70													2	2	5	5	7	21	9	3	5	13		2		5																										
71													6	2	9	8	5	7	4	7	9	4	3	4	1	1	1						1																			
72													1	2	8	3	8	11	12	9	5	5	1	3																												
73													1	1	3	4	5	5	8	3	10	7	1	3			1	1																								
74													1	2	5	2	5	7	3	7	7	5	3	4	1	1																										
75													1	1	5	5	10	2	9	7	13	5	2	5	2	2		2							1																	
76														2	2	5	3	5	4	3	8	5	2	6	1	3	1	1																								
77												1			2	2	6	3	8	9	12	4	4	4		2		1		1				1																		
78													2		1	3	1		7	4	9	2	2	4	1		1	1	1									1														
79															3	2	5	1	7	6	11	3	2	5	1	1																										
80															2		4		10	4	7	2	4	2	3	1			1				1																			
81															2		3	3	3	3	7	3	2	6	3		1	2																								
82															1	3			7	7	10	4	4	9	4	2	1																									
83																		1	2	4	5	6	2	9	2	2	2	1	1																							
84															1			1	2	5	6	1	1	4	3	2						1																				
85																	1	1	5	1	4	1	2	3	4		2	3																								
86															1		1		1	5	10	3	8	3	4		2	1							1																	
87																			3	1	5	2	1	6	5		3	1																								
88																1			4	1	6	2	3	5		1	2	2	2						1	1								1								
89																1		1	2	2	3		4	5	2	2	1	1	1							1		1														
90																					2	1	3	6	2	2	1	1		1				1																		
91																					2	2	1	3	4	1	1	2																								
92																	1				1	1	5	2	2	2	1		3							1																
93																					2		1	2	2		2	1	1							1																
94															1							1	1	1	6			3																								
95																		1	1					3	3		2		1		1	1			1																	
96																					1		4	2	3	1	3	1	2																							
97																							1		1			1	1	1																						
98																							2	1	3			2	2							1																
99																							1		2			1	1		1		1																			
100																							1	1	2	1	1	2		1	1					1		1														
101																								1	1		1									1																
102																													3						2																	
103																							1					1		1	1				1	1																
104																									1		3		1						1																	
105																													1	1																						
106																									1		1					1					1															
108																																							1													
109																																			1																	
110																											1							1																		
111																																				1																
114																														1																			1			
116																																														1						
119																																								1												

Figure 5.12

The aged dataset was analysed for differences between areas and gears. Data were available for four quarters, two gear-types (trawl and gillnet) and two areas (subdivision 25 and 27).

5.1.4.1 Differences between the areas:

For the gillnet data (Figure 5.13) there were significant differences between the areas for all quarters (p<0.001; p<0001; p=0.07; p=0.01). In subdivision 27 there appear to be more young fish (age 3) at length than subdivision 25. This is quite consistent in all quarters, although data is sparse for Q3 and Q4, therefore the model could only be fitted over a few age classes.


Q1, p<0.001

Q2, p<0.001

Q3, p=0.07

Q4, p=0.01

Figure 5.13 Gillnet in subdivisions 25 vs. 27.

In quarter 1 (Figure 5.14) there was no difference between the areas for trawl gear (p=0.59), for the other quarters there was no trawl data available for subdivision 27.

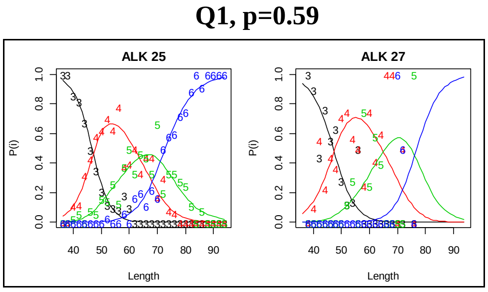
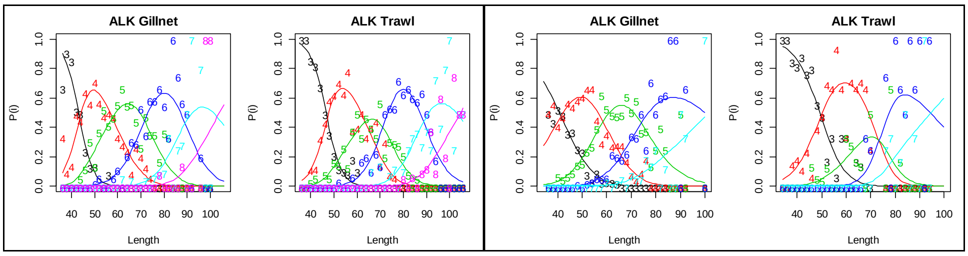

Q1, p=0.59

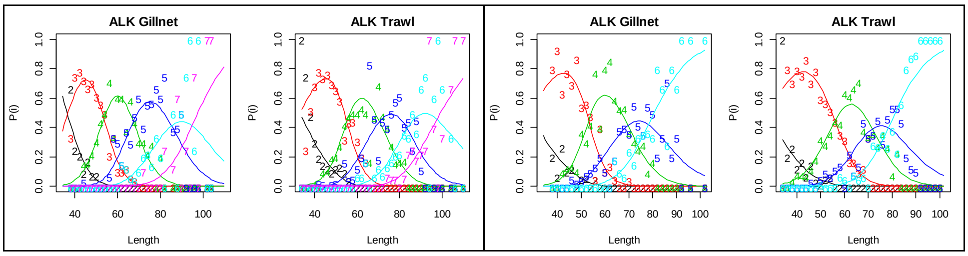
Figure 5.14 Trawl gear in subdivisions 25 vs. 27.

5.1.4.2 Differences between gear types:

In area 25 there were significant differences between the gear-types for the first two quarters (p<0.001; p<0.001) with more young fish at length in the trawl catches. (e.g. at in Q2 fish of 40 cm in the trawl catches are nearly all age 3, while fish of the same length in the gillnet catches are around 50% age 3 and 50% age 4). In the last two quarters there are no significant differences. The differences in the first two quarters might be related to spawning activity where the gillnet gear would catch more mature fish or target spawning areas more specifically than the trawl gear. Alternatively, there could be a vertical behaviour component where older fish might be found in other depth layers than younger fish. However, this effect is not seen in area 27.


Q1, p<0.001

Q2, p<0.001


Q3, p=0.26

Q4, p=0.26

Figure 5.15 Area 25, gear-types gillnet and trawl.

In area 27, no significant differences between the gear-types was found in Q1 (p=0.91; no data for trawl gear in other quarters).

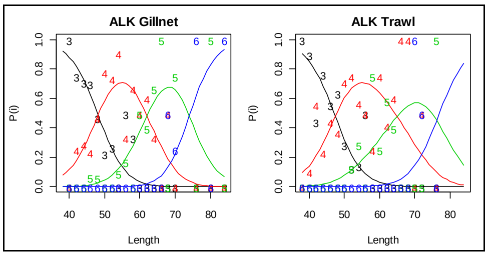

Q1, p=0.91

Figure 5.16 Area 27, gear-types gillnet and trawl.

5.1.4.3 Discussion on the method

The question of comparing ALK is essential especially in the purpose of combining them to have more information on the poorly sampled length classes. With this method developed by Gerritsen (2005), ALKs from different gear-types appeared to be significantly different, which is against expectation if the same population is being sampled. In this case gear-type was probably a proxy for area and the different gear types would have sampled different segments of the population.

A multinomial approach would be appropriate here and there is no need to reduce the data to a binomial form. (The binomial approach over-estimates the degrees of freedom (of the null model) as age classes are represented twice in the binomial dataset and is therefore not entirely appropriate. The assumptions of the model are based on the length distribution of the age groups (normality and equal variance). Therefore the model is based on length-given-age and not on age-given-length. However it works on proportions (more or less) and therefore the data themselves do no need to be normally distributed

The GAM appears to be a good approach to take the time element into account as the current stratification into quarters is arbitrary and does not take into account the growth during the quarter and does not utilise the strong correlation with previous and subsequent quarters

It is possible to prove that two ALKs are different but it is impossible to prove that they are not. In any case the tool could still be used as an index of similarity between ALKs.

5.2 Precision estimation

In order to inform the choice of sampling design, we require knowledge of how the different sampling designs affect precision. Simulation trials are one method of estimating the effect of the sampling design. Another potentially informative approach is to compare precision estimates from different countries, which have been calculated for different sampling designs using different estimation methods, to try to identify the causes of any differences. However, few precision estimates were available for this meeting.

Therefore, as a first step towards comparing precision resulting from different sampling designs, we attempted to compare precision estimates calculated using three different methods, for a case study stock: Baltic cod, using data collected by Sweden in 2003, described below.

Substantial effort was required to format the data into the correct format for the analysis packages and we were not able to perform this for more than one data set in the time available

5.2.1 Estimation Methods

The three methods considered were as described in WKSCMFD (ICES 2004a), and the packages used to apply these methods at WKSDFD were coded in R and Splus.

1. “analytical” method – based on sampling theory.
2. bootstrap method – case resampling.
3. Bayesian model based approach.

5.2.2 Analysis of Baltic cod data

The Baltic cod data consist of around 6000 age, length and weight samples taken as random samples from size classes within trips (*i.e.* no length stratification), and some 30000 length samples taken from a different set of size classes within trips. The length samples were taken at sea, and the ages at the port.

5.2.2.1 Sampling Theory (“Analytical”) Method

The sampling theory estimates were calculated using the stratified total estimator as described by Thompson (1992). The strata were defined by the combination of space, time, gear, and size category. The sampling unit was defined by combining date, vessel, and size category.

Results are given in section 5.2.2.4.

5.2.2.2 Bootstrapping Method

The CV was estimated by bootstrap resampling, using trip as the sampling unit. The strata were defined by the combination of space, time, gear, and size category. One thousand bootstrap resamples were taken, the catch-at-age estimated from each, and the CV for each age estimated as the CV of the bootstrapped catch-at-age estimates.

Results are given in section 5.2.2.4.

5.2.2.3 Bayesian Method

The data were analysed using the Bayesian hierarchical model of Hirst *et al.* (2004) and also using a bootstrap method. The analysis was done on the age samples alone, then on the full data set. The model accounts for variation between the sampling units (here the size category within the trips), and also age reading errors.

The results are shown in Tables 5.1-5.4 and Figures 5.1 and 5.2: For the three largest age classes, the mean CVs for the four analyses are

Age only	0.10
Age only, age errors	0.11
Age and length	0.09
Age and length, age errors	0.13

The main conclusions are:

1. Adding the length data to the analysis did not reduce the CV, but it did change the mean catch-at-age. The reason for this is that the two sets of samples appear to be sampling different populations. There is a higher proportion of large fish in the on-board length samples. Hence the catch-at-age of the older ages is increased when this data is included (and the catch-at-age of the younger fish is decreased accordingly). It would normally be expected that including more information would reduce the CV, but in this case the conflict between the two data sources adds uncertainty.

It is not possible to make a direct quantification of the effect of adding length samples to age samples in our case. This is because the two sources of data are to some degree incompatible. Clearly it makes a difference to the estimates of catch-at-age if the length samples are included, but it is not possible to say whether the estimates are better or worse. This can only be established by further investigation of the two sampling methods (*i.e.* why are more large fish sampled on board the boats?). Some possible causes are that the market sampling scheme over samples fish factories where small fish are more likely to be found, or that boats with a large number of large fish sometimes land their catch outwith the sampling area, or logistical problems to randomly choose boxes within piles of boxes at the auctions.

2. Including the possibility of age reading errors also increases the CV, as would be expected, and also has a large effect on the means. The catch of the largest age increases by 15% for the age-only data and by 10% when the lengths are used.

Including an estimate of age reading errors makes an important difference to the mean catch-at-age, and also to the CV. The implication of this is that more age samples are needed in order to reduce this error (or possibly otoliths could be read more than once). It is also apparent that it is important to take this source of error into account in the analysis.

5.2.2.4 Comparison of Methods

The comparison of CVs is only possible for the analysis of age-only data with no age errors. The means were very similar, but the bootstrap CVs were considerably smaller at older ages, whilst the analytical CVs were small for ages with the largest samples.

Age	2	3	4	5	6	7	8	9	10
Bootstrap	0.12	0.14	0.05	0.11	0.12	0.14	0.17	0.33	0.32
Bayesian	0.23	0.07	0.11	0.14	0.23	0.60	0.73	1.16	1.05
Samp theo	0.16	0.03	0.05	0.19	0.25	0.34	0.61	1.38	1.29

Note that these CVs are for estimates combined over gears and seasons, and are therefore not those that would be reported for stock assessment purposes. These would be for smaller strata, and would therefore be larger.

5.2.3 Conclusions

1. Although including length-only samples would usually reduce the CV to some extent, this does not happen in this example because the length-only data were collected in a different way to the age and length data. These two sampling methods do not appear to be sampling the same population, and so one or both must be unrepre-

sentative of the true catch. Until the reason for this discrepancy is established, there can be no correct analysis of the data.

2. The CVs are largest for the Bayesian method. It is known that bootstrapping tends to underestimate the variance, particularly when sample sizes are very small (as with the oldest age classes). The analytical method incorrectly assumes that the fish are all independent, and therefore does not take account of the large between sample variation in the data.
3. Age reading errors are likely to add a very significant bias to the catch-at-age estimates, though taking them into account does not appear to increase the CV greatly.

	Mean	Sd	CV	2.5%	97.5%
<2	30586.92	42545.12	**1.39**	4317.70	169749.9
2	690393.10	160569.42	**0.23**	446489.42	1061306.5
3	8069614.43	578659.45	**0.07**	6927288.66	9139905.5
4	3897206.75	422823.45	**0.11**	3095001.90	4765488.3
5	953665.45	137425.05	**0.14**	704645.19	1186470.0
6	193705.37	43941.61	**0.23**	135436.13	273534.8
7	69406.54	41865.44	**0.60**	36075.33	202455.6
8	45660.59	33375.48	**0.73**	21340.80	153441.6
9	35287.53	40976.58	**1.16**	7847.55	181666.4
10	38932.63	40866.04	**1.05**	15108.34	166657.0
11	23586.75	35121.92	**1.49**	3777.75	118650.6
12	35762.36	36128.34	**1.01**	11608.93	149911.2
>12	30155.70	47392.39	**1.57**	7677.09	177064.6
tot	14113964.13	555134.33	**0.04**	12942913.70	15249734.8

Table 5.1 Landed numbers-at-age, sd, CV, and CIs. Age only data, no age reading errors.

	Mean	Sd	CV	2.5%	97.5%
<2	24136.20	31041.70	**1.29**	3316.06	116876.10
2	50148.18	120050.69	**2.39**	4574.45	332288.19
3	9307900.76	611331.38	**0.07**	7971909.57	10296837.69
4	3700848.90	398140.04	**0.11**	2953896.97	4402425.98
5	645292.92	103314.32	**0.16**	487973.52	846702.20
6	155578.35	41015.22	**0.26**	102482.62	237959.53
7	49953.47	35332.34	**0.71**	23149.43	158012.13
8	35830.69	30343.06	**0.85**	11838.49	124358.52
9	18667.32	27463.40	**1.47**	4231.08	113526.84
10	27073.29	38594.88	**1.43**	6923.77	145813.33
11	16979.19	27505.07	**1.62**	2786.01	98679.31
12	22121.07	32668.62	**1.48**	5150.16	130496.23
>12	16491.57	29576.90	**1.79**	2309.31	91711.30
tot	14071021.91	428280.50	**0.03**	13216970.86	14873393.26

Table 5.2 Landed numbers-at-age, sd, CV, and CIs. Age only data, age reading probability of error of 0.1 in either direction (e.g. a fish of true age 5 has a 10% chance of being read as 4, and a 10% chance of being read as 6).

	Mean	Sd	CV	2.5%	97.5%
<2	20541.09	32512.35	**1.58**	4164.73	125649.9
2	505939.31	138368.97	**0.27**	304442.87	761733.0
3	5992103.84	424623.27	**0.07**	5189989.89	6724060.0
4	2578998.23	230540.65	**0.09**	2140967.46	2994980.3
5	692385.04	82103.33	**0.12**	549952.09	871089.3
6	168884.17	36915.47	**0.22**	115337.57	254504.0
7	155342.73	54286.17	**0.35**	85722.20	273520.2
8	128654.01	57189.66	**0.44**	56918.79	268193.3
9	45433.99	42390.92	**0.93**	11192.27	166059.6
10	48186.10	45060.93	**0.94**	15924.84	199209.1
11	31800.02	46977.78	**1.48**	5397.43	181949.8
12	47399.98	48759.15	**1.03**	6098.28	194931.5
>12	71160.59	42057.89	**0.59**	21855.90	184089.2
tot	10486829.11	460076.44	**0.04**	9778320.07	11422562.2

Table 5.3 Landed numbers-at-age, sd, CV, and CIs. Age and length data, no age reading error.

	Mean	Sd	CV	2,5%	97.5%
<2	15805.53	29969.17	**1.90**	617.55	114504.2
2	101900.46	68279.04	**0.67**	33915.90	268259.7
3	6645581.39	559692.20	**0.08**	5518571.55	7566424.7
4	2802135.33	395083.32	**0.14**	2115391.68	3665602.8
5	640698.81	117052.64	**0.18**	480625.76	921857.2
6	258659.87	73508.36	**0.28**	162535.46	398089.1
7	194530.68	86284.25	**0.44**	80606.41	393534.6
8	257593.28	81674.29	**0.32**	132347.46	436641.1
9	69592.81	42742.22	**0.61**	16817.06	162792.9
10	47431.03	40892.19	**0.86**	4787.28	160098.6
11	96101.97	32197.53	**0.34**	47008.64	170649.7
12	84598.94	43431.18	**0.51**	24248.39	188559.9
>12	182864.82	74875.47	**0.41**	77350.41	352817.0
tot	11397494.93	390350.03	**0.03**	10667019.51	12118525.8

Table 5.4 Landed numbers-at-age, sd, CV, and CIs. Age and length data, age reading probability of error of 0.1 in either direction.

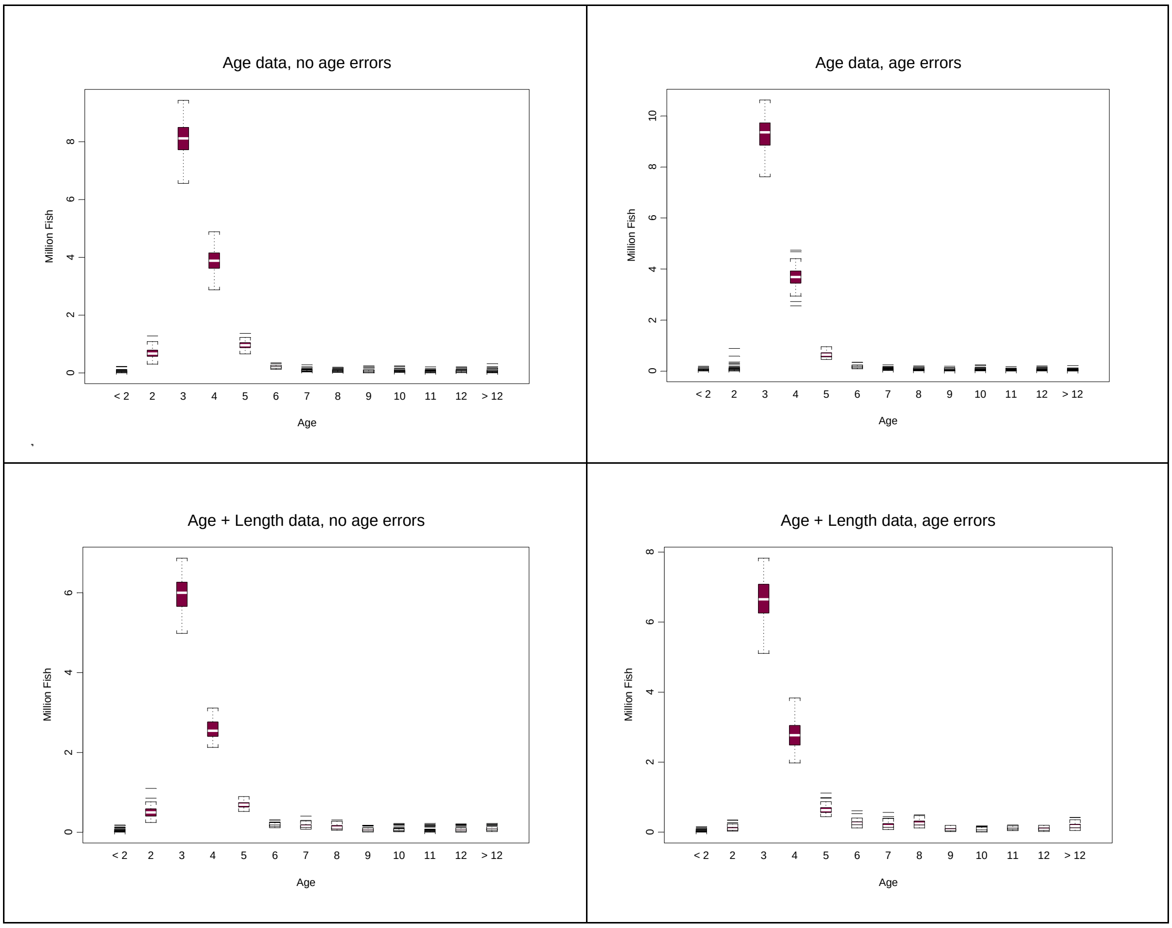


Figure 5.1 Precision at age for the Baltic cod age structure related to the scenario used and calculated by the Bayesian method.

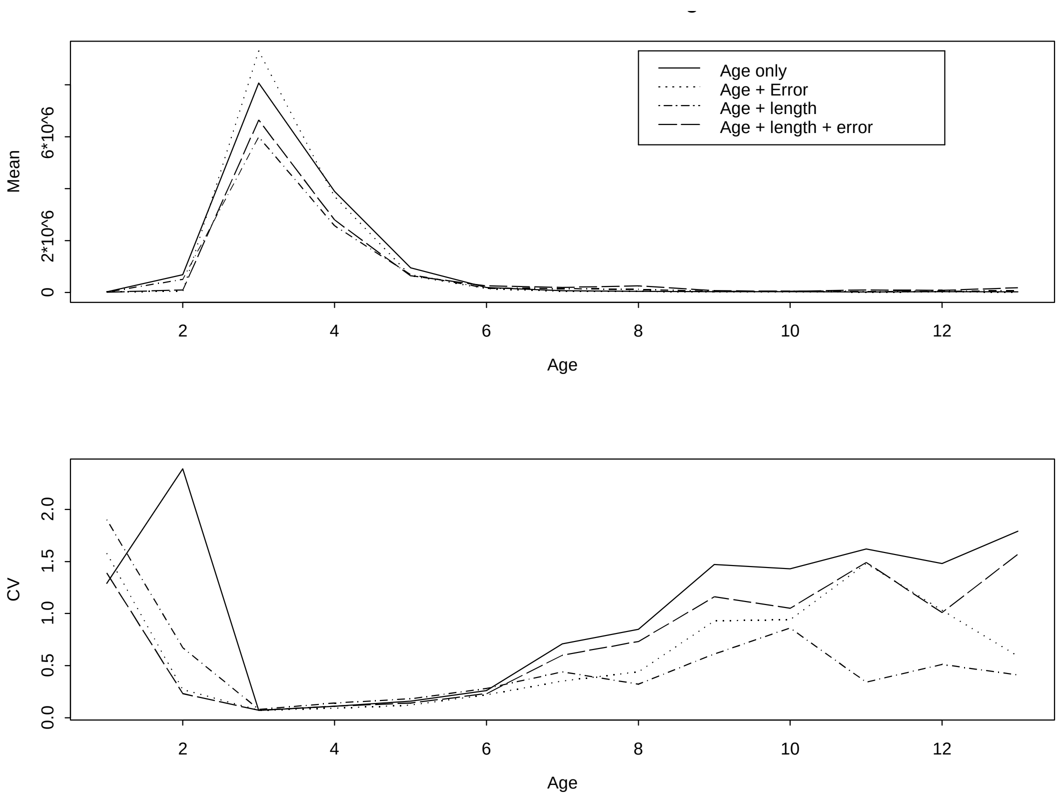


Figure 5.2 Mean and CV for the Baltic cod age structure related to the scenario used and calculated by the Bayesian method.

6 RECOMMENDATIONS

- More stocks need to be analysed in term of precision to improve the knowledge on the quality of the current sampling.
- It should be desirable to have a common dataset for some area to compare methodologies.
- The different existing tools should be assembled in a common "open source" tool based on a common data set format. The development of such a tool would contribute on a coherent, collaborative and more efficient work.
- A workshop involving only statisticians needs to be convened to address specific methodological issues like provide guidelines for elaborating and handling age length keys, calculate the precision related to the different sampling methods, …
- The problem of sampling design is a follow-up process that needs to be addressed in a long run project-like forum.
- The current workshop needs to shift to a user workshop involving tool developers and sampling managers in order to:

a. explore countries datasets prepared in a specified format using the tools currently developed.
b. propose improvements based on the outcomes of the exploration.

7 Summary of the working documents

WD01 - A simple model for comparing age–length keys and filling in missing data

Hans Gerritsen

A method is presented to model age–length keys (ALKs) using a Generalised Linear Modelling approach. The (multinomial) aged data are presented to the model in pairs of consecutive age classes so they take on a binomial form (each observation belongs to one of the two age classes). The models can be used for a number of purposes, in the simplest form they can estimate proportions at age for length classes for which no age data is available. It is also possible to test whether two or more ALKs are significantly different. And finally the models might be expanded to Generalised additive models that could take changes over time or in space into account which could improve precision and even estimate ALKs for periods or areas where no sampling took place.

WD02 - Italian sample survey for landing statistics: links with the biological sampling of landings

Michele De Meo

In considering the methodologies used in the collection and compilation of fishery statistics in the member countries of the European Economic Area it is essential to keep in mind the large variation in the structure of the fishing industries in those countries. In particular, in Italy (generally for the Mediterranean) there are a very large number of fishing vessels making landings in a large number of mainly small landing places and disposing of their products in markets of a less formal structure than in the other countries. In this situation, to collect accurate fishery data, it is impossible to observe the whole fleet, composed in Italy by about 15000 vessels, and the best system appears to be that of sample surveys.

A sampling procedure for biological parameters and related methods to estimate precision in length and age structure are affected by these considerations. In this context, the sample design for landings and effort data is to be considered the starting point for biological sampling.

IREPA has developed a sample design and a software to collect and manage landing statistics; following this procedure, a representative sample of the whole Italian fleet has been identified. Nevertheless this representative sample cannot be used to evaluate the biological composition of landings because it would be too expensive, in time and costs.

Consequently, it is suggested to select a sub-sample of the sampled vessels.

Length and age structure of the landings individuals will be recorded in some randomly selected fishing days.

Finally, the coefficients of variation will be calculated taking into account the statistical methodology for multi-stage sampling besides the guidelines produced by the WKSCMFD (ICES 2004a).

WD03 -A simple tool to calculate biological parameters uncertainty

Paz Sampedro, M. Sainza and V. Trujillo

Description of an automatic procedure to estimate uncertainty of some biological parameters: growth, maturation, sex-ratio and length-weight curves. A routine has been developed in R environment, which makes possible to fit the most usual models and to estimate the coefficient of variation for parameters by using the non-parametric bootstrap methodology. The approach is illustrated by an application for Southern Hake stock.

8 BIBLIOGRAPHY

Aanes, S. and M. Pennington (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.

Cochran, W. G. (1977). Sampling techniques, John Wiley and Sons.

DeMeo, M. (2005) Working document. “Italian sample survey for landings statistics: links with the biological sampling of landings”. WKSDFD, Pasajes.

Gerritsen, H. (2005) Working document. A simple model for comparing age–length keys and filling in missing data. WKSDFD, Pasajes

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

ICES (2004a). WKSCMFD. Workshop on Sampling Calculation and Methodology for Fisheries Data, 26-30 February 2004, Nantes, France. ICES CM 2004/ACFM:12.

ICES (2004b). PGCCDBS. Report of the Planning Group on Commercial Catch, Discards and Biological Sampling, 2-5 March 2004, Palma de Mallorca, Spain. ICES CM 2004/ACFM:13.

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

Quinn, T. J. and R. B. Deriso (1999). Quantitative fish dynamics: Oxford University Press. 542 pp.

Sampedro, P. M. Sainza and V. Trujillo (2005) Working Document. A simple tool to calculate biological parameters uncertainty. WKSDFD, Pasajes

STECF (2005). SGRN. Guidelines for the submission of Technical Reports on the National Data Collection Programs under Council Regulation (EC) N°1543/2000 and Commission Regulation (EC) N°1639/2001 as amended by Commission Regulation (EC) N°1581/2004. SEC (2005) 159. 50p.

Thompson, S. K. (1992). Sampling. New York: Wiley.

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

Annex A: Stock disaggregated sampling strategy summary and related precision

Country	Species		Area	Landings (t) 2004 or 2003	LENGTH - sampling								AGE - sampling								
					Stratification				Number samples	Number measured	Precision reached (CV)	Method used to calculate precision code	Stratification				Method used for taking age-samples	Number samples	Number aged	Precision reached (CV)	Method used to calculate precision code
	English	Scientific			Time interval, code	Space, code	Gear/metier, number of strata	Commercial category, number of strata					Time interval, code	Space, code	Gear/metier, number of strata	Commercial category, number of strata					
Bel	plaice	*Pleuronectes platessa*	IV	4303	Q1	-	1	4 & 5	10	4978	3%	A & B	Q1	-	1	4 & 5	ALK	10	331	25%	A & B
NO	Sandeel	*Ammodytes spp.*	IV,IIIa	29616	Q	S	1			2660			Q	S	1		ALK		286		
NL	Greater Argentine	*Argentuna silus*	VIa, VIIbc	2752	Q, S	D		-	6	347		B	Q, S	-	-	-	direct, full	6	150		B
Ita	Giant red shrimp	*Aristaeomorpha foliacea*	1.3	251	Q	S	1	2	14	675											
Ita	Giant red shrimp	*Aristaeomorpha foliacea*	2.2	1187	Q	S	1	2	34	1688											
Ita	Red shrimp	*Aristeus antennatus*	1.3	221	Q	S	1	2	9	434											
Ita	Red shrimp	*Aristeus antennatus*	2.2	133	Q	S	1	2	26	1306											
Gre	Bogue	*Boops boops*	3.1		Y	S	2	3	23	1150		A			2	3	ALK	20	500		
NO	Tusk	*Brosme brosme*	I-XIV	13373	Q	S							Q	S							
Fra	Edible Crab	*Cancer pagurus*	II, V, VI, VII (exc d), VIII, IX, X, XII, XI	6471	Q	-	2	-	113	4375			-	-	-	-	-				
UK	Edible crab	*Cancer pagurus*	All (Excl IV,VIId)		Q					27609											
UK	Edible Crab	*Cancer pagurus*	IV + VIId		Q					27549											
Ire	Edible Crab	*Cancer pagurus*	VI. VII			D				1358		B							0		
Ire	Herring	*Clupea harengus*	Celtic Sea		Q					8335		B	Q				direct		1123		B
NO	NSS herring	*Clupea harengus*	II	440891	Q	S	1		266	21451			Q	S	1		ALK	101	7521		
UK	Atlanto-Scandian Herring	*Clupea harengus*	IIa + V		Q					156							ALK		94		
Sw	Herring	*Clupea harengus*	IIIa	31457	Q	-	2	0	*70*	*3500*			Q	-	2	0	direct	70	3500		
Est	Herring	*Clupea harengus*	IIIb-d	6953	Q	S	2	-	52	5200		-	Q	S	2	1	Direct	52	5200		-
Fin	Herring	*Clupea harengus*	IIIb-d	63358	Q	S	3	-	137	65503			Q	S	3	-	ALK		4240		
Sw	Herring	*Clupea harengus*	IIIb-d(sd22-24)	9876	Q	S	2	0	*40*	*2300*			Q	S	2	0	direct	40	2300		
Sw	Herring	*Clupea harengus*	IIIb-d(sd25-29)	29190	Q	S	2	0	*48*	*2490*			Q	S	2	0	direct	48	2490		
Sw	Herring	*Clupea harengus*	IIIb-d(sd30)	5224	Q	-	2	0	13	1350			Q	-	2	0	direct	13	1350		
GFR	Herring	*Clupea harengus*	IIIc22-IIId24	13886	S	-	4		39	45908		A	S	-	4		ALK	39	5980		A
Est	Herring	*Clupea harengus*	IIId, Gulf of Riga	19647	Q	O, Gulf of	2	-	36	3600		-	Q	O, Gulf of	2	1	Direct	36	3600		-
Ire	Herring	*Clupea harengus*	Irish Sea		Q					0		B	Q				direct		112		B
GFR	Herring	*Clupea harengus*	IV	32337	Q	-	1		64	21762		A	Q	-	1		ALK	8	1040		A
UK	Herring	*Clupea harengus*	IV		Q	-				12217							ALK		3163		A
NL	Herring	*Clupea harengus*	IV + VIId	155628	Q, S	D		-	158	22163		B	Q, S	-	-	-	direct, full	158	3950		B
NO	Herring	*Clupea harengus*	IV,IIIa	124233	Q	S	1			9123			Q	S	1		Direct		3138		
Ire	Herring	*Clupea harengus*	Nor.Spring Spawning		Q					317		B	Q				direct		208		B
UK	Herring	*Clupea harengus*	VIa		Q					1740									459		
GFR	Herring	*Clupea harengus*	VIId	9684	Q	-	1		42	17329		A	Q	-	1		ALK	4	600		A
Ire	Herring	*Clupea harengus*	West of Ireland		Q					8515		B	Q				direct		616		B
Fin	Whitefish	*Coregonus lavaretus*	IIId	805	Q	S	2	-	47	4168			Q	S	2	-	Direct		4168		
Fra	Grenadier	*Coriphaenoides rupestris*	II, V, VI, VII (exc d), VIII, IX, X, XII, XI	9530	Q	-	1	2	84	7060			A	-	-	-	ALK		661		
UK	Roundnose Grenadier	*Coryphaenoides rupestris*	All NE Atlantic		Q					132							ALK		0		
NO	Lumpsucker	*Cyclopterus lumpus*	I-II				1														
Fra	Sea bass	*Dicentrarchus labrax*	II, V, VI, VII (exc d), VIII, IX, X, XII, XI	3200	Q	-	6	3	679	6136			Q	-	-	-	ALK		715		
UK	Sea Bass	*Dicentrarchus labrax*	All (Excl IV,VIId)		Q					3905							ALK		1076		
UK	Sea Bass	*Dicentrarchus labrax*	IV + VIId		Q					3791									857		
Fra	Sea bass	*Dicentrarchus labrax*	IV + VIId		Q	-	-	3	35	748			-	-							
Ita	Eledone	*Eledone cirrhosa*	1.3	1361	Y	S	1	2	3	133											
Ita	Eledone	*Eledone cirrhosa*	2.1	81	Y	S	1	2	7	350											
Ita	Eledone	*Eledone cirrhosa*	2.2	1110	Y	S	1	2	4	205											
Ita	Eledone	*Eledone moschata*	1.3	211	Y	S	1	2	9	441											
Ita	Eledone	*Eledone moschata*	2.1	2116	Y	S	1	2	9	450											
Ita	Eledone	*Eledone moschata*	2.2	1620	Y	S	1	2	3	154											
Ita	Anchovy	*Engraulis encrasicolus*	1.3	6237	Q	S	1		167	8351			Q	S	1		ALK	41	1013		
Ita	Anchovy	*Engraulis encrasicolus*	2.1	30044	Q	S	2		724	36178			Q	S	2		ALK	95	2370		
Ita	Anchovy	*Engraulis encrasicolus*	2.2	17092	Q	S	2		72	3594			Q	S	2		ALK	25	635		
Sp	Anchovy	*Engraulis encrasicolus*	IXa (Gulf of Cádiz)		Q	1	3	4		12166		-	Q	1	1	1	ALK		1390		A,B
Sp	Anchovy	*Engraulis encrasicolus*	VIII		Q	5	3	4		2072		-	Q	5	1	1	ALK		1249	8%	A,B

Country	Species		Area	Landings (t) 2004 or 2003	LENGTH - sampling								AGE - sampling								
	English	Scientific			Stratification				Number samples	Number measured	Precision reached (CV)	Method used to calculate precision code	Stratification				Method used for taking age-samples	Number samples	Number aged	Precision reached (CV)	Method used to calculate precision code
					Time interval, code	Space, code	Gear/metier, number of strata	Commercial category, number of strata					Time interval, code	Space, code	Gear/metier, number of strata	Commercial category, number of strata					
Fra	Anchovy	*Engraulis engrasicolus*	1.2	7360	S	-	-	4		2923			A	-	-	-	ALK		754		
Gre	Anchovy	*Engraulis engrasicolus*	2.2		Y	S	1		25	1250		A			1		ALK	13	325		A
Gre	Anchovy	*Engraulis engrasicolus*	3.1		Y	S	1		55	2750		A	Y	S	1		ALK	21	525		A
Fra	Anchovy	*Engraulis engrasicolus*	VIII	5873	Q	-	-	4	73	5506	8%	A	Q	-	-	-	ALK	8	1641	9%	A
Ita	Gurnard	*Eutrigla gurnardus*	1.3	21	Y	S	1	1	2	50			Y	S	1	1	ALK	1	25		
Ita	Gurnard	*Eutrigla gurnardus*	2.1	451	Y	S	1	1	3	125			Y	S	1	1	ALK	5	125		
Ita	Gurnard	*Eutrigla gurnardus*	2.2	46	Y	S	1	1	1	25			Y	S	1	1	ALK	1	25		
DK	Cod	*Gadus morhua*	22-24	7154	Q	S	2	5	81	2819	11%	A	Q	S	1	5	ALK	81	2737	3%	A
DK	Cod	*Gadus morhua*	25-26	10918	Q	S	2	5	44	1937	14%	A	Q	S	1	5	ALK	44	1801	6%	A
Fra	Cod	*Gadus morhua*	3Ps	3100	Q	-	-	-		750			A	-	-	-	ALK		290		
GFR	Cod	*Gadus morhua*	II	5921	Q	-	1		169	76038		A	Q	-	1		ALK	18	2159		A
NO	Cod	*Gadus morhua*	I-II	193544	Q	O, No 1	5		1014	157637			Q	O, No 1	5		O, No 2	207	15459	5%	BM, No 3
Sw	Cod	*Gadus morhua*	IIIaN	495	Q	-	1	5	12	486		B	Q	-	1	5	direct	12	486		B
DK	Cod	*Gadus morhua*	IIIaS	727	Q	S	2	5	37	1013	18%	A	Q	S	1	5	ALK	37	1011	6%	A
Sw	Cod	*Gadus morhua*	IIIaS	582	Q	-	1	5	22	740		B	Q	-	1	5	direct	22	740		B
Fin	Cod	*Gadus morhua*	IIIb-d	1163																	
Sw	Cod	*Gadus morhua*	IIIb-d(sd22-24)	2245	Q	-	2	5	51	1268		B	Q	-	2	5	direct	51	1268		B
Sw	Cod	*Gadus morhua*	IIIb-d(sd25-32)	12075	Q	-	2	5	250	6082		B	Q	-	2	5	direct	250	6082		B
GFR	Cod	*Gadus morhua*	IIIc22-IIId24	4404	S	-	5		87	29613		A	S	-	5		ALK	18	4029		A
GFR	Cod	*Gadus morhua*	IIId25-32	2598	S	-	5		77	23928		A	S	-	5		ALK	14	3164		A
GFR	Cod	*Gadus morhua*	IV	2134	Q	-	3		116	2192		A	Q	-	3		ALK	13	576		A
NL	Cod	*Gadus morhua*	IV	3886	Q			5	44	2166		B	Q	O	-	-	ALK	40	1970		B
UK	Cod	*Gadus morhua*	IV + VIId		Q	-				61575							ALK		14631		-
NO	Cod	*Gadus morhua*	IV,IIIa	5328	Q	O,No 4				2274			Q	O,No 4			ALK				
Ire	Cod	*Gadus morhua*	VI		Q	D				4568		B	Q				ALK		813		B
UK	Cod	*Gadus morhua*	VIa + VIb		Q	-				4342							ALK		1895		
Ire	Cod	*Gadus morhua*	VIIa		Q	S				2652		B	Q				ALK		842		B
UK	Cod	*Gadus morhua*	VIIa		Q					9560							ALK		771		
Fra	Cod	*Gadus morhua*	VIIb-k	5020	Q	-	2	5	208	2387	22%	A	Q	-	-	-	ALK		1124		
UK	Cod	*Gadus morhua*	VIIb-k (ex d)		Q					6236							ALK		1273		
Ire	Cod	*Gadus morhua*	VIIb-k,VIII,IX,X		Q	S				5027		B	Q				ALK		2506		B
Fra	Lobster	*Homarus gammarus*	II, V, VI, VII (exc d), VIII, IX, X, XII, XI	320	Q	-	2	-	163	3965			-	-	-	-	-				
Gre	Lobster	*Homarus gammarus*	3.1		Y	S	1	3	24	1200		A	Y	S	1	3	direct	24			A
UK	Lobster	*Homarus gammarus*	All (Excl IV, VIId)		Q					7091											
UK	Lobster	*Homarus gammarus*	IV + VIId		Q					19277											
Ire	Lobster	*Homarus gammarus*	VI. VII			D				50		B							0		
Ire	Orange Roughy	*Hoplostethus atlanticus*								1281		B					ALK		288		B
Pt	Four-spot megrim	*Lepidorhombus boscii*	9a	12	S	O(harbour)	7		219	15768	20%	1									
Sp	Four-spot megrim	*Lepidorhombus boscii*	VI, VII & VIIIab		Q	3	4	1		1788		-	Q	3	1	1	ALK				A,B
Sp	Four-spot megrim	*Lepidorhombus boscii*	VIIIc & IXa		Q	5	2	1	128	13890		-	Q	5	1	1	ALK		213	12%	A,B
UK	Megrim	*Lepidorhombus spp*	IV + VIId		Q					8909									324		
Ire	Megrim	*Lepidorhombus spp*	Vb,VI,,XII,XIV		Q	S				3006		B	Q				ALK		1121		B
Sp	Megrim	*Lepidorhombus spp*	VI, VII & VIIIab		Q	3	4	1		17764		-	Q	3	1	1	ALK		1315		A,B
Ire	Megrim	*Lepidorhombus spp*	VII		Q	D				15336		B	Q				ALK		1969		B
Fra	Megrim	*Lepidorhombus spp*	VII, VIIIab	3196	Q	-	-	5	441	13626			-	-	-	-	-				
Sp	Megrim	*Lepidorhombus spp*	VIIIc & IXa		Q	5	2	1		18162		-	Q	5	1	1	ALK		264		A,B
Sp	Megrim	*Lepidorhombus whiffiagonis*	VI, VII & VIIIab		Q	3	4	1	120	15976		-	Q	3	1	1	ALK		1315		A,B
UK	Megrim	*Lepidorhombus whiffiagonis*	VIa + VIb + Vb		Q					18989							ALK		1994		
Sp	Megrim	*Lepidorhombus whiffiagonis*	VIIIc & IXa		Q	5	2	1	127	4272		-	Q	5	1	1	ALK		51	20%	A,B

Country	Species		Area	Landings (t) 2004 or 2003	LENGTH - sampling								AGE - sampling								
	English	Scientific			Stratification				Number samples	Number measured	Precision reached (CV)	Method used to calculate precision code	Stratification				Method used for taking age-samples	Number samples	Number aged	Precision reached (CV)	Method used to calculate precision code
					Time interval, code	Space, code	Gear/metier, number of strata	Commercial category, number of strata					Time interval, code	Space, code	Gear/metier, number of strata	Commercial category, number of strata					
NL	Dab	*Limanda limanda*	IV	6960	Q			-	25	2332		B	Q	O	-	-	ALK	12	300		B
Pt	Squid	*Loligo spp*	9a	289	S	O(harbour)	7		219	29127	25%	1									
Ita	Squid	*Loligo vulgaris*	1.3	1079	Y	S	1	2	4	203											
Ita	Squid	*Loligo vulgaris*	2.1	435	Y	S	1	2	6	300											
Ita	Squid	*Loligo vulgaris*	2.2	554	Y	S	2	2	7	354											
Sp	Squid	*Loligo vulgaris*	VIIIc & IXa		Q	5	3	1		2510		-	Q	5	1	1	-				A,B
Ita	Black-bellied angler	*Lophius budegassa*	1.3	329	Q	S	1	3	3	103			Y	S	1	3	ALK	4	89		
Ita	Black-bellied angler	*Lophius budegassa*	2.1	266	Q	S	1	3	8	375			Y	S	1	3	ALK	0	0		
Ita	Black-bellied angler	*Lophius budegassa*	2.2	798	Q	S	1	3	3	155			Y	S	1	3	ALK	5	114		
Sp	Black-bellied angler	*Lophius budegassa*	VI, VII & VIIIab		Q	3	4	1	147	6701		-	Q	3	1	1	ALK		68		A,B
Sp	Black-bellied angler	*Lophius budegassa*	VIIIc & IXa		Q	5	7	1	279	3753		-	Q	5	1	1	ALK		221	15%	A,B
Ita	Anglerfish	*Lophius piscatorius*	1.3	20	Q	S	1	3	3	64			Y	S	1	3	ALK	2	62		
Ita	Anglerfish	*Lophius piscatorius*	2.1	38	Q	S	1	3	2	75			Y	S	1	3	ALK	0	0		
Ita	Anglerfish	*Lophius piscatorius*	2.2	218	Q	S	1	3	1	25			Y	S	1	3	ALK	1	25		
UK	Anglerfish	*Lophius piscatorius*	IV + VIId		Q					16300									871		
NO	Anglerfish	*Lophius piscatorius*	IV,IIIa,VI	790	Q	O,No 4	2			133											
Sp	Anglerfish	*Lophius piscatorius*	VI, VII & VIIIab		Q	3	4	1	149	8529		-	Q	3	1	1	ALK		526		A,B
UK	Anglerfish (Piscatorius)	*Lophius piscatorius*	VIa + VIb + Vb		Q					15669							ALK		587		
Sp	Anglerfish	*Lophius piscatorius*	VIIIc & IXa		Q	5	7	1	215	6391		-	Q	5	1	1	ALK		519	52%	A,B
GFR	Anglerfish	*Lophius spp*	IV	26	Q	-	1		50	94		A									
Ire	Anglerfish	*Lophius spp*	VI		Q	D				2827		B	Q				ALK		721		B
Sp	Anglerfish	*Lophius spp*	VI, VII & VIIIab		Q	3	4	1		15230		-	Q	3	1	1	ALK		594		A,B
Ire	Anglerfish	*Lophius spp*	VII		Q	D				25855		B	Q				ALK		1012		B
Fra	Angler	*Lophius spp*	VII, VIIIabd	12861	Q	-	-	6	2253	29675			Q	-	-	-	ALK	5	1168		
Sp	Anglerfish	*Lophius spp*	VIIIc & IXa		Q	5	7	1		10144		-	Q	5	1	1	ALK		740		A,B
Fra	Skipjack	*Lutjanus purpureus*	COPACO (WECAF)		S	-				9196			-	-	-	-	-				
Pt	Blue whiting	*M. poutassou*	9a	2652	S	O(harbour)	7		322	39606			Q	S	All		ALK		1134	20%	1
Ire	Blue Whiting	*M. poutassou*	Combined		Q					2161		B	Q						780		B
NL	Blue Whiting	*M. poutassou*	VI,VII,VIII	76670	Q, S	D		-	72	13240		B	Q, S	-	-	-	direct, full	72	1799		B
NO	Capelin	*Mallotus villosus*	I-II	190000	Q	O,No 4	1						Q	O,No 4	1		ALK				
NO	Haddock	*Melanogrammus aeglefinus*	I-II	56437	Q	O, No 1	5		744	99934			Q	O, No 1	5		O, No 2	115	9323	17.40%	BM, No 3
NO	Haddock	*Melanogrammus aeglefinus*	IV,IIIa	11387	Q	O,No 4				3536			Q	O,No 4			ALK				
GFR	Haddock	*Melanogramus eglefinus*	II	814	Q	-	1		46	12061		A	Q	-	1		ALK	7	563		A
Sw	Haddock	*Melanogramus eglefinus*	IIIa		Q	-			23	719			Q	-			direct	23	719		
GFR	Haddock	*Melanogramus eglefinus*	IV	1301	Q	-	2		76	2254		A									
UK	Haddock	*Melanogramus eglefinus*	IV + VIId		Q	-				98664									9650		-
Fra	Haddock	*Melanogramus eglefinus*	IV, IIIa	817	Q	-	-	3	3	288			Q	-	-	-	ALK		645		
Ire	Haddock	*Melanogramus eglefinus*	VI		Q	D				8171		B	Q				ALK		1287		B
UK	Haddock	*Melanogramus eglefinus*	VIa+VIb+VII {not d}+VIII+XII+XIV		Q					26654									3745		
Ire	Haddock	*Melanogramus eglefinus*	VII,VIII,IX,X		Q	D				13149		B	Q				ALK		3558		B
Fra	Haddock	*Melanogramus eglefinus*	VIIb-k	4788	Q	-	1	5	47	2022	23%	B	Q	-	-	-	ALK	8	1933		
GFR	Whiting	*Merlangius merlangius*	IV	278	Q	-	3		53	447		A									
NL	Whiting	*Merlangius merlangius*	IV	2052	Q			4	42	5139		B	Q	O	-	4	ALK	24	1200		B
UK	Whiting	*Merlangius merlangius*	IV + VIId		Q	-				76704									7994		
Fra	Whiting	*Merlangius merlangius*	IV+ VIId	8692	Q	-	-	4	34	2952			Q	-	-	-	ALK	7	2604		
Ire	Whiting	*Merlangius merlangius*	VI		Q	D				2275		B	Q				ALK		460		B
UK	Whiting	*Merlangius merlangius*	VIa + VIb + Vb		Q					6643							ALK		1503		
UK	Whiting	*Merlangius merlangius*	VII (ex a,d)		Q					9680							ALK		835		
Ire	Whiting	*Merlangius merlangius*	VIIa		Q	S				777		B	Q				ALK		234		B
UK	Whiting	*Merlangius merlangius*	VIIa		Q					3109									187		
Fra	Whiting	*Merlangius merlangius*	VIIb-k	1988	Q	-	1	4	178	10463			Q	-	-	-	ALK		1554		
Ire	Whiting	*Merlangius merlangius*	VIIb-k		Q					15548		B	Q				ALK		3012		B
Fra	Hake	*Merluccius merluccius*	1.2	2022	S	-	4	15	998	23426			A	-	-	-	ALK		376		
Ita	Hake	*Merluccius merluccius*	1.3	4074	Q	S	3	3	41	2026			Q	S	3	3	ALK	22	543		
Ita	Hake	*Merluccius merluccius*	2.1	2606	Q	S	2	3	68	3400			Q	S	2	3	ALK	34	842		
Gre	Hake	*Merluccius merluccius*	2.2		Y	S	2	3	15	750		A	Y	S	2	3	ALK	12	300		A
Ita	Hake	*Merluccius merluccius*	2.2	7978	Q	S	3	3	44	2221			Q	S	3	3	ALK	23	579		
Gre	Hake	*Merluccius merluccius*	3.1		Y	S	2	3	30	1500		A	Y	S	2	3	ALK	12	300		A

Country	Species		Area	Landings (t) 2004 or 2003	LENGTH - sampling								AGE - sampling								
	English	Scientific			Stratification				Number samples	Number measured	Precision reached (CV)	Method used to calculate precision code	Stratification				Method used for taking age-samples	Number samples	Number aged	Precision reached (CV)	Method used to calculate precision code
					Time interval, code	Space, code	Gear/metier, number of strata	Commercial category, number of strata					Time interval, code	Space, code	Gear/metier, number of strata	Commercial category, number of strata					
Pt	Hake	*Merluccius merluccius*	9a	1920	S	O(harbour)	7		1272	133560			A	S	All		ALK		630	10%	1
Fra	Hake	*Merluccius merluccius*	IIIa, IV, VI, VII, VIIIab	10452	Q	-	5	6	1078	24724	13%	A	Q	-	-	-	ALK	9	1756	-	-
UK	Hake	*Merluccius merluccius*	IV + VIId		Q					3798									0		
Ire	Hake	*Merluccius merluccius*	Vb,VI,VII,XII,XIV		Q	S				7225		B	Q				ALK		1111		B
UK	Hake	*Merluccius merluccius*	VIa + VIb		Q					15861							ALK		717		
Sp	Hake	*Merluccius merluccius*	VIIIc & IXa		Q	5	10	1	316	34019		-	Q	5	1	1	ALK		2426	14%	A,B
Sp	Hake	*Merluccius merluccius*	VI, VII & VIIIab		Q	3	4	1	354	54402		-	Q	3	1	1	ALK		1821		A,B
UK	Blue Whiting	*Micromesistius poutassou*	I - X		Q					1131							ALK		423		
NO	Blue whiting	*Micromesistius poutassou*	I-X,XII,XIV	851395	S	S	2			7948			S	S	2				3560		
Sp	Blue whiting	*Micromesistius poutassou*	VIIIc & IXa		Q	5	10	1		19107		-	Q	5	1	1	ALK		619	9%	A,B
GFR	Dab	*Microstomus kitt*	IV	712	Q	-	3		89	10614		A									
NL	Lemon Sole	*Microstomus kitt*	IV	949	Q	-		3	8	425		B	Q	-	-	3	direct,	8	425		B
UK	Lemon Sole	*Microstomus kitt*	IV + VIId		Q					9390									323		
NO	Blue ling	*Molva dypterygia*	I-XIV	547	Q	S							Q	S							
UK	Ling	*Molva molva*	All NE Atlantic		Q					444							ALK		103		
NO	Ling	*Molva molva*	I-XIV	13952	Q	S							Q	S							
Ire	Ling	*Molva molva*			Q					264		B	Q				ALK		260		B
Ita	Red mullet	*Mullus barbatus*	1.3	1804	Q	S	2	2	28	1418			Q	S	2		ALK	11	275		
Ita	Red mullet	*Mullus barbatus*	2.1	3 018	Q	S	2	1	80	4000			Q	S	2		ALK	35	885		
Gre	Red mullet	*Mullus barbatus*	2.2		Y	S	1	3	12	600		A	Y	S	1	3	ALK	10	250		A
Ita	Red mullet	*Mullus barbatus*	2.2	8205	Q	S	2	2	53	2630			Q	S	2		ALK	27	665	7%	A
Gre	Red mullet	*Mullus barbatus*	3.1		Y	S	2	3	25	1250		A				3					
GFR	Red mullet	*Mullus barbatus*	II	82	Q	-	1		9	2453		A									
GFR	Red mullet	*Mullus barbatus*	XII, XIV,	3634	Q	-	1		59	51337		A	Q	-	1		ALK	38	713		A
Fra	Stripped red mullet	*Mullus surmuletus*	II, V, VI, VII (exc d), VIII, IX, X, XII, XI	1500	Q	-	1	2	54	1561			-	-	-	-	-				
Ita	Stripped red mullet	*Mullus surmuletus*	1.3	895	Q	S	2	2	21	1063			Q	S	2	2	ALK	11	264		
Ita	Stripped red mullet	*Mullus surmuletus*	2.1	93	Q	S	2	2	14	700			Q	S	2	2	ALK	7	167		
Gre	Stripped red mullet	*Mullus surmuletus*	2.2		Y	S	2	3	6	300		A	Y	S	2	3	ALK	6	150		A
Ita	Stripped red mullet	*Mullus surmuletus*	2.2	2120	Q	S	2	2	52	2600			Q	S	2	2	ALK	32	794		
Gre	Stripped mullet	*Mullus surmuletus*	3.1		Y	S	2	3	20	1000		A	Y	S	2	3	ALK	6	150		A
Fra	Stripped red mullet	*Mullus surmuletus*	IV + VIId	2700	Q	-	-	5	17	870			-	-	-	-	-				
Ita	Nephrops	*Nephrops norvegicus*	1.3	472	Q	S	1	2	34	1714											
Ita	Nephrops	*Nephrops norvegicus*	2.1	1484	Q	S	1	2	70	3500											
Ita	Nephrops	*Nephrops norvegicus*	2.2	2124	Q	S	1	2	50	2521											
Ire	Nephrops	*Nephrops norvegicus*	Aran Grounds, FU17			O (FU)				5800		B							0		
Ire	Nephrops	*Nephrops norvegicus*	Celtic Sea FU 20-22			O (FU)				4800		B							0		
Sw	Nephrops	*Nephrops norvegicus*	IIIaN		Q	O			38	10824			Q	O							
Sw	Nephrops	*Nephrops norvegicus*	IIIaS		Q	O			14	4958			Q	O							
Ire	Nephrops	*Nephrops norvegicus*	Ireland NW Coast, FU 18			O (FU)				200		B							0		
Ire	Nephrops	*Nephrops norvegicus*	Ireland SW & SE Coast, FU 19			O (FU)				1400		B							0		
Ire	Nephrops	*Nephrops norvegicus*	Irish Sea East,FU 14			O (FU)				0		B							0		
Ire	Nephrops	*Nephrops norvegicus*	Irish Sea West, FU 15			O (FU)				14400		B							0		
NL	Nephrops	*Nephrops norvegicus*	IV	1102	Q	FU5		4	24	10798		B	-	FU5	-	-	-				B
UK	Nephrops	*Nephrops norvegicus*	IV + VIId		Q	FU				101738									0		
Ire	Nephrops	*Nephrops norvegicus*	Landings outside WG FU's, L			O (FU)				0		B							0		
Ire	Nephrops	*Nephrops norvegicus*	Landings outside WG FU's, M			O (FU)				0		B							0		
Ire	Nephrops	*Nephrops norvegicus*	Porcupine Bank, FU16			O (FU)				3200		B							0		
UK	Nephrops	*Nephrops norvegicus*	VI		Q	FU				99937											
Sp	Nephrops	*Nephrops norvegicus*	VI & VII		Q	2	2	1		10172		-	Q	2	1	1	-				A,B
UK	Nephrops	*Nephrops norvegicus*	VII (EX d)		Q	FU				3829											
Fra	Nephrops	*Nephrops norvegicus*	VII (unités fonctionnelles)	3658	Q	FU	-	2	43	2860	15%	A	-	-	-	-	-				
Sp	Nephrops	*Nephrops norvegicus*	VIII & IX		Q	5	2	1		12490		-	Q	5	1	1	-				A,B
Fra	Nephrops	*Nephrops norvegicus*	VIII, IX (unités fonctionnelles)	2839	S	FU	-	2	1183	18730	5%	A	-	-	-	-	-				
Gre	Octopus	*Octopus vulgaris*	2.2		Y	S	1	1	4	100		A	Y	S	1	1	direct	4			A
Gre	Octopus	*Octopus vulgaris*	3.1		Y	S	1	1	6	150		A	Y	S	1	1	direct	6			A
Pt	Octopus	*Octopus vulgaris*	9a	1183	S	O(harbour)	7		518	43512	10%	1									
Sp	Octopus	*Octopus vulgaris*	VIIIc & IXa		Q	5	3	1		6675		-	Q	5	1	1	-				A,B
Pt	Flying squid	*Ommastrephidae*	9a	147	S	O(harbour)	7				25%	1									

Country	Species		Area	Landings (t) 2004 or 2003	LENGTH - sampling								AGE - sampling								
					Stratification								Stratification								
	English	Scientific			Time interval, code	Space, code	Gear/metier, number of strata	Commercial category, number of strata	Number samples	Number measured	Precision reached (CV)	Method used to calculate precision code	Time interval, code	Space, code	Gear/metier, number of strata	Commercial category, number of strata	Method used for taking age-samples	Number samples	Number aged	Precision reached (CV)	Method used to calculate precision code
Pt	Auxialliary seabream	*P. acarne*	9a	1033	S	O(harbour)	7		328	39032	20%	1									
Ita	Bream	*Pagellus erythrinus*	1.3	547	Y	S	2	2	6	305			Y	S	2	2	ALK	6	152		
Ita	Bream	*Pagellus erythrinus*	2.1	14	Y	S	1	2	1	50			Y	S	1	2	ALK	1	25		
Ita	Bream	*Pagellus erythrinus*	2.2	803	Y	S	2	2	8	400			Y	S	2	2	ALK	8	211		
NO	Northern prawn	*Pandalus borealis*	I-II	34653	Q	O,No 4	1			40627											
Sw	Pandalid shrimp	*Pandalus borealis*	IIIaN	2155	Q		1	2					Q		1	2					
NO	Northern prawn	*Pandalus borealis*	IV,IIIa	3987	Q	O,No 4	1			628											
NO	King crab	*Paralithodes camtschaticus*	I	820			1														
Ita	White shrimp	*Parapenaeus longirostris*	1.3	823	Q	S	1	2	19	950											
Ita	White shrimp	*Parapenaeus longirostris*	2.1	50	Q	S	1	2	11	550											
Ita	White shrimp	*Parapenaeus longirostris*	2.2	10673	Q	S	1	2	72	3606											
Fra	Scallop	*Pecten maximus*	VIId	7360	A	-	1	1	23	2050			Q	-	1	1	D	23	2050		
Fra	Shrimp	*Penaeus subtilis*	COPACO (WECAF)		S	-				4156			-	-	-	-	-				
Fin	Perch	*Perca fluviatilis*	IIId	984	Q	S	2	-	20	2804			Q	S	2	-	ALK		1209		
Est	Perch	*Perca fluviatilis*	IIId, Gulf of Riga	512	Q	O, Gulf of	2	-	11	1050		-	Q	S	2	1	Direct	11	1050		-
GFR	Flounder	*Platichtys flesus*	IIIb-d	1626	S	-			20	15141		A	S	-			ALK	20	4581		A
Est	Flounder	*Platichtys flesus*	IIId	410	Q	S	2	-	18	1800		-	Q	S	2	1	Direct	18	1800		-
Fin	Flounder	*Platichtys flesus*	IIId	42	Q3,Q4	S	1	-	25	629			Q3,Q4	S	1	-	Direct	25	629		
Sw	Plaice	*Pleuronectes platessa*	IIIaN	179	Q	-	1	4	23	973			Q	-	1	4	direct	23	973		
Sw	Plaice	*Pleuronectes platessa*	IIIaS	134	Q	-	1	4	11	744			Q	-	1	4	direct	11	744		
GFR	Plaice	*Pleuronectes platessa*	IV	3461	Q	-	2		100	8785		A	Q	-	2		ALK	10	1107		A
NL	Plaice	*Pleuronectes platessa*	IV	43743	Q			4	80	4560		B	Q	O	-	4	direct,	80	4560		B
UK	Plaice	*Pleuronectes platessa*	IV + VIId		Q					17705									1878		
Ire	Plaice	*Pleuronectes platessa*	Vb,VI,XII,XIV		Q	S				2307		B	Q				ALK		427		B
Ire	Plaice	*Pleuronectes platessa*	VIIa		Q	S				6827		B	Q				ALK		532		B
UK	Plaice	*Pleuronectes platessa*	VIIa		Q					2816							ALK		334		
Fra	Plaice	*Pleuronectes platessa*	VIId	2536	Q	-	-	4	153	7789	13%	A	Q	-	-	-	ALK	8	1837		
Fra	Plaice	*Pleuronectes platessa*	VIIe	195	Q	-	-	4	6	136			-	-	-	-	-				
UK	Plaice	*Pleuronectes platessa*	VIIe		Q					14724							ALK		2167		
Bel	plaice	*Pleuronectes platessa*	VIIfg	166	Q1	-	1	4 & 5	2	858	6%	A & B	Q1	-	1	4 & 5	ALK	2	144	45%	A & B
Ire	Plaice	*Pleuronectes platessa*	VIIfg		Q	S				6835		B	Q				ALK		412		B
UK	Plaice	*Pleuronectes platessa*	VIIfg		Q					8248							ALK		895		
GFR	Saithe	*Pollachius virens*	II	2168	Q	-	1		45	17411		A	Q	-	1		ALK	9	961		A
NO	Saithe	*Pollachius virens*	I-II	148431	Q	O,No 4	5			49312			Q	O,No 4	5		O, No 2		4725		
GFR	Saithe	*Pollachius virens*	IV	8965	Q	-	1		79	10505		A	Q	-	1		ALK	10	1983		A
UK	Saithe	*Pollachius virens*	IV + VIId		Q					12995									5226		
Fra	Saithe	*Pollachius virens*	IV, IIIa	15008	Q	-	-	4	212	5596	8%	A	Q	-	-	-	ALK	8	1933	10%	A
NO	Saithe	*Pollachius virens*	IV,IIIa	61713	Q	O,No 4				16262			Q	O,No 4			ALK				
Fra	Saithe	*Pollachius virens*	Vb, VI, XII, XIV	3493	Q	-	1	4	168	3529			Q	-	-	-	ALK		531		
Ire	Saithe	*Pollachius virens*	VI		Q	D				432		B	Q				ALK		495		B
UK	Saithe	*Pollachius virens*	VIa + VIb + Vb		Q					1872							ALK		1061		
Ire	Saithe	*Pollachius virens*	VII,VIII,IX,X		Q	D				1994		B	Q				ALK		473		B
GFR	Turbot	*Psetta maxima*	IIIc22-IIId25	21	Q	-			1	325		A	Q	-			ALK	1	173		A
GFR	Turbot	*Psetta maxima*	IV	261	Q	-	2		43	595		A									
NL	Turbot	*Psetta maxima*	IV	2607	Q			7	24	1042		B	Q	O	-	7	direct,	24	1042		B
UK	Turbot	*Psetta maxima*	IV + VIId		Q					336									0		
Ita	Rays spp.	*Raja clavata*	1.3	216	Y	S	1	2	4	106											
Ita	Rays spp.	*Raja clavata*	2.1	91	Y	S	1	2	1	30											
Ita	Rays spp.	*Raja clavata*	2.2	294	Y	S	1	2	2	75											
Ita	Rays spp.	*Raja miraletus*	1.3	18	Y	S	1	2	5	128											
Ita	Rays spp.	*Raja miraletus*	2.2	26	Y	S	1	2	2	75											
Fra	Ray	*Raja spp.*	II, V, VI, VII (exc d), VIII, IX, X, XII, XI	2900	A	-	-	2	28	390			-	-	-	-	-				
Sp	Rays	*Raja spp.*	All areas		Q	8	7	1		0		-	Q	8	1	1	ALK				A,B
UK	Skates & Rays	*Raja spp.*	All NE Atlantic		Q					2028							ALK		0		

Country	Species		Area	Landings (t) 2004 or 2003	LENGTH - sampling								AGE - sampling								
					Stratification								Stratification								
	English	Scientific			Time interval, code	Space, code	Gear/metier, number of strata	Commercial category, number of strata	Number samples	Number measured	Precision reached (CV)	Method used to calculate precision code	Time interval, code	Space, code	Gear/metier, number of strata	Commercial category, number of strata	Method used for taking age-samples	Number samples	Number aged	Precision reached (CV)	Method used to calculate precision code
NL	Rays spp.	*Raja spp.*	IV	811	Q	-		4	24	1459		B	-	-	-	-	-				B
UK	Skates & Rays	*Raja spp.*	IV + VIId		Q					2447									0		
Ire	Ray /Skate	*Raja spp.*								680		B							0		
NO	Gr. Halibut	*Reinhardtius hippoglossoides*	I-II	8215	Q	O,No 4	5			12615			Q	O,No 4	5		O, No 2				
GFR	Greenland Halibut	*Reinhardtius hippoglossoides*	XIV	5033	Q	-	1		158	45307		A	Q	-	1		ALK	106	1409		A
Fin	Salmon	*Salmo salar*	IIIb-d	343	Q	S	4	-		3666			Q	S	4	-	Direct		3666		
Sw	Salmon	*Salmo salar*	IIIb-d	706	Q	S	2	9					Q	S	2	9?					
Ita	Bonito	*Sarda sarda*	1.3	994	Q	S	2	1		224											
Ita	Bonito	*Sarda sarda*	2.1	140	Q	S	1	1		100											
Ita	Bonito	*Sarda sarda*	2.2	1157	Q	S	2	1		222											
Fra	Sardine	*Sardina pilchardus*	1.2	15539	Q	-	-	4		3118			A	-	-	-	ALK		454		
Ita	Sardine	*Sardina pilchardus*	1.3	6259	Q	S	1		110	5492			Q	S	1		ALK	12	311		
Ita	Sardine	*Sardina pilchardus*	2.1	12916	Q	S	3		307	15364			Q	S	3		ALK	54	1346		
Gre	sardine	*Sardina pilchardus*	2.2	0	Y	S	1		28	1400		A	Y	S	1		ALK	26	650		A
Ita	Sardine	*Sardina pilchardus*	2.2	5451	Q	S	2		45	2258			Q	S	2		ALK	10	241		
Gre	sardine	*Sardina pilchardus*	3.1	0	Y	S	1		57	2850		A	Y	S	1		ALK	16	400		A
Pt	Sardine	*Sardina pilchardus*	9a	66528	S	O(harbour)	2		514	67848			Q	O(zones)	All		ALK		7542	15%	1
Fra	Sardine	*Sardina pilchardus*	VIIIabd	11073	Q	-	2	4	68	6462	8%	A	Q	-	-	-	ALK		1058		
Sp	sardine	*Sardina pilchardus*	VIIIc & IXa	0	Q	5	3	2		23161		-	Q	5	1	1	ALK		3687	9%	A,B
Ita	Mackerel	*Scomber scombrus*	1.3	1137	Y	S	1		14	353			Y	S	1		ALK	1	32		
Ita	Mackerel	*Scomber scombrus*	2.1	485	Y	S	2		23	578			Y	S	2		ALK	6	144		
Ita	Mackerel	*Scomber scombrus*	2.2	1890	Y	S	2		15	367			Y	S	2		ALK	1	29		
Pt	Mackerel	*Scomber scombrus*	9a	2749	S	O(harbour)	8		283	26602			Q	S	All		ALK		1380	20%	1
UK	Mackerel	*Scomber scombrus*	II -IX		Q					15979							ALK		3573		
Fra	Mackerel	*Scomber scombrus*	IIa, IIIbcd, IV	1934	Q	-	-	3	1	167			-	-	-	-	-				
GFR	Mackerel	*Scomber scombrus*	IV	4660	Q	-	2		13	7110		A	Q	-	2		ALK	4	451		A
UK	Mackerel	*Scomber scombrus*	IV	0	Q					14485									4297		
NL	Mackerel	*Scomber scombrus*	IVac, VIId	17986	Q, S	D		-	49	3975		B	Q, S	-	-	-	direct, full	49	1225		B
NO	Mackerel	*Scomber scombrus*	I-X	163406	Q	O,No 4	1		186	25050			Q	O,No 4	1		ALK	46	1864		
Ire	Mackerel	*Scomber scombrus*	NEA	0	Q					2029		B	Q				direct		986		B
GFR	Mackerel	*Scomber scombrus*	VI, VII	15394	Q	-	1		56	28219		A	Q	-	1		ALK	14	2000		A
NL	Mackerel	*Scomber scombrus*	VI, VIIabcefghjk, VIII	62495	Q, S	D		-	39	3013		B	Q, S	-	-	-	direct, full	39	975		B
Sp	Mackerel	*Scomber scombrus*	VIIIc & IXa	0	Q	5	3	1		11625		-	Q	5	1	1	ALK		1620	9%	A,B
GFR	Brill	*Scophtalmus rombus*	IV	57	Q	-	2		26	105		A									
NL	Brill	*Scophtalmus rombus*	IV	1189	Q			4	24	783		B	Q	O	-	4	direct,	24	783		B
UK	Brill	*Scophtalmus rombus*	IV + VIId	0	Q					292									0		
NO	Golden redfish	*Sebastes marinus*	I-II	6955	Q	O,No 4	5			11565			Q	O,No 4	5		O, No 2		805		
NO	Deep-sea redfish	*Sebastes mentella*	I-II	1193	Q	O,No 4	1			1925			Q	O,No 4	1		O, No 2		1111		
Ita	Cuttle fish	*Sepia officinalis*	1.3	1920	Q	S	2	1	7	350											
Ita	Cuttle fish	*Sepia officinalis*	2.1	4155	Q	S	2	1	18	900											
Ita	Cuttle fish	*Sepia officinalis*	2.2	2994	Q	S	1	1	6	310											

Country	Species		Area	Landings (t) 2004 or 2003	LENGTH - sampling								AGE - sampling								
					Stratification								Stratification								
	English	Scientific			Time interval, <u>code</u>	Space, <u>code</u>	Gear/metier, <u>number</u> of strata	Commercial category, <u>number</u> of strata	Number samples	Number measured	Precision reached (CV)	Method used to calculate precision <u>code</u>	Time interval, <u>code</u>	Space, <u>code</u>	Gear/metier, <u>number</u> of strata	Commercial category, <u>number</u> of strata	Method used for taking age-samples	Number samples	Number aged	Precision reached (CV)	Method used to calculate precision <u>code</u>
Pt	Cuttle fish	*Sepia officinalis*	9a	1298	S	O(harbour)	7		155	7285	20%	1									
Sp	Cuttle fish	*Sepia officinalis*	VIIIc & IXa	0	Q	5	3	1		1211		-	Q	5	1	1	-				A,B
Ire	Migratory Sharks	*Shark-like Selachii*		0						7		B							0		
Ita	Sole	*Solea vulgaris*	1.3	353	Y	S	1	1	2	100			Y	S	1	1	ALK	3	75		
Ita	Sole	*Solea vulgaris*	2.1	836	Y	S	2	1	16	800			Y	S	2	1	ALK	9	225		
Bel	sole	*Solea vulgaris*	IV	1527	Q1	-	1	7 & 9	5	3413	4%	A & B	Q1	-	1	7 & 9	ALK	5	247	40%	A & B
GFR	Sole	*Solea vulgaris*	IV	898	Q	-	1		60	3102		A	Q	-	1		ALK	2	338		A
NL	Sole	*Solea vulgaris*	IV	14230	Q			5	70	3496		B	Q	O	-	5	direct,	70	3496		B
UK	Sole	*Solea vulgaris*	IV + VIId		Q					25359									3488		
Ire	Sole	*Solea vulgaris*	Vb,VI,XII,XIV		Q	S				1798		B	Q						0		
Ire	Sole	*Solea vulgaris*	VIIa		Q	S				1936		B	Q				ALK		212		B
UK	Sole	*Solea vulgaris*	VIIa		Q					2985							ALK		210		
Ire	Sole	*Solea vulgaris*	VIIbc		Q	S				1457		B	Q				ALK		515		B
Bel	sole	*Solea vulgaris*	VIId	1171	Q1&Q4	-	1	7 & 9	6	5122	6%	A & B	Q1&Q4	-	1	7 & 9	ALK	6	379	48%	A & B
Fra	Sole	*Solea vulgaris*	VIId	2805	Q	-	-	5	159	9031	10%	B	Q	-	-	-	ALK	10	1533	8%	A
Fra	Sole	*Solea vulgaris*	VIIe	190	Q	-	-	5	57	1383			A	-	-	-	ALK		154		
UK	Sole	*Solea vulgaris*	VIIe		Q					11933							ALK		1535		
Ire	Sole	*Solea vulgaris*	VIIfg		Q	S				2572		B	Q				ALK		128		B
UK	Sole	*Solea vulgaris*	VIIfg		Q					8540							ALK		863		
Ire	Sole	*Solea vulgaris*	VIIhjk		Q	S				2604		B	Q				ALK		381		B
Bel	sole	*Solea vulgaris*	VIIIab	249	Q2	-	1	7 & 9	6	4191	3%	A & B	Q2	-	1	7 & 9	ALK	6	249	32%	A & B
Fra	Sole	*Solea vulgaris*	VIIIab	3345	Q	-	4	6	1419	6795	7%	A	Q	-	-	-	ALK		1167		
Gre	Picarel	*Spicara smaris*	2.2		Y	S	1		3	150		A	Y	S	1		direct	3			
Gre	Picarel	*Spicara smaris*	3.1		Y	2	2		10	50		A	Y	S	2		direct	10			
Sw	Sprat	*Spratus spratus*	IIIa	6488	Q	-	1	0	*24*	*1200*			Q	-	1	0	direct	24	1200		
Est	Sprat	*Spratus spratus*	IIIb-d	30219	Q	S	1	-	62	6163		-	Q	S	1	1	Direct	62	6163		-
Fin	Sprat	*Spratus spratus*	IIIb-d	8949	Q	S	2	-	41	12051			Q	S	2	-	ALK		1363		
GFR	Sprat	*Spratus spratus*	IIIb-d	28455	S	-			9	1981		A	S	-			ALK	7	452		A
Sw	Sprat	*Spratus spratus*	IIIb-d	85862	Q	-	1	0	*80*	*4000*			Q	-	1	0	direct	80	4000		
Fin	Pikeperch	*Stizostedion lucioperca*	IIId	732	Q	S	2	-	79	2460			Q	S	2	-	ALK		1691		
Est	Pikeperch	*Stizostedion lucioperca*	IIId, Gulf of Riga	41	Q	O, Gulf of	2	-	9	789		-	Q	S	2	1	Direct	9	789		-
Ita	Tuna	*Thunnus alalunga*	1.3	2584	Q	S	1			313											
Ita	Tuna	*Thunnus alalunga*	2.1	245	Q	S	1			167											
Ita	Tuna	*Thunnus alalunga*	2.2	4083	Q	S	1			613											
Gre	Albacore tuna	*Thunnus alalunga*	3.1		Y	S	1		All specimens of the landings are measured												A
Fra	Tuna	*Thunnus alalunga*	Atlantic, Indian, Pacific oceans	2926	Q	-	-	3	43	2445			-	-	-	-	-				
Ire	Albacore Tuna	*Thunnus alalunga*	Combined							200		B							0		
Ita	Tuna	*Thunnus thynnus*	1.3	3093	Q	S	4			2107											
Ita	Tuna	*Thunnus thynnus*	2.1	207	Q	S	1			868											
Ita	Tuna	*Thunnus thynnus*	2.2	1203	Q	S	5			4610											
Fra	Tuna	*Thunnus thynnus*		6442	A	-	-	-		788			-	-	-	-	-				
Gre	Blue fin tuna	*Thunnus thynnus*	3.1		Y	S	1		All specimens of the landings are measured												A
Ita	Mediterranean horse mackerel	*Trachurus mediterraneus*	1.3	702	Y	S	1	1	18	916			Y	S	1	1	ALK	8	212		
Ita	Mediterranean horse mackerel	*Trachurus mediterraneus*	2.1	283	Y	S	1	1	6	300			Y	S	1	1	ALK	3	75		
Ita	Mediterranean horse mackerel	*Trachurus mediterraneus*	2.2	766	Y	S	3	1	9	450			Y	S	3	1	ALK	6	157		
Gre	Mediterranean horse mackerel	*Trachurus mediterraneus*	3.1		Y	S	1	3	11	550		A			1	3	ALK	10	250		
Ita	Horse Mackerel	*Trachurus trachurus*	1.3	1400	Y	S	1	1	16	804			Y	S	1	1	ALK	9	230		
Ita	Horse Mackerel	*Trachurus trachurus*	2.1	1124	Y	S	1	1	12	600			Y	S	1	1	ALK	6	146		
Ita	Horse Mackerel	*Trachurus trachurus*	2.2	1737	Y	S	3	1	14	700			Y	S	3	1	ALK	7	173		
Gre	Horse mackerel	*Trachurus trachurus*	3.1		Y	S	1	3	7	350		A			1	3	ALK	10	250		
Pt	Horse mackerel	*Trachurus trachurus*	9a	11242	S	O(harbour)	8		960	130560			Q	S	All		ALK		2318	20%	1
GFR	Horse Mackerel	*Trachurus trachurus*	IV	5149	Q	-	2		13	285		A	Q	-	2		ALK	1	177		A
NL	Horse Mackerel	*Trachurus trachurus*	IVa, VI, VIIabcefghjk, VIII	77041	Q, S	D		-	51	7749		B	Q, S	-	-	-	direct, full	51	1275		B
NL	Horse Mackerel	*Trachurus trachurus*	IVbc,VIId	30827	Q, S	D		-	48	6286		B	Q, S	-	-	-	direct, full	48	1200		B
NO	Horse mackerel	*Trachurus trachurus*	I-X	20514	Q	S	1		11	975			Q	S	1		ALK	7	413		
GFR	Horse Mackerel	*Trachurus trachurus*	VI, VII	18058	Q	-	1		57	17847		A	Q	-	1		ALK	10	1320		A
Sp	Horse Mackerel	*Trachurus trachurus*	VI, VII & VIIIab		Q	3	3	1		0		-	Q	3	1	1	ALK				A,B
Sp	Horse Mackerel	*Trachurus trachurus*	VIIIc & IXa		Q	5	3	1		21333		-	Q	5	1	1	ALK		1319	19%	A,B
Ire	Horse Mackerel	*Trachurus trachurus*	Western		Q					4632		B	Q				direct		1245		B
Ita	Gurnard	*Trigla lucerna*	1.3	177	Y	S	1	1	2	75			Y	S	1	1	ALK	3	75		
Ita	Gurnard	*Trigla lucerna*	2.1	156	Y	S	3	1	3	150			Y	S	3	1	ALK	5	125		
Ita	Gurnard	*Trigla lucerna*	2.2	299	Y	S	1	1	1	50			Y	S	1	1	ALK	1	25		
NO	Norway pout	*Trisopterus esmarki*	IV,IIIa	11357	Q	S	1			2244			Q	S	1		ALK		412		
Pt	Norway pout	*Trisopterus luscus*	9a	3405	S	O(harbour)	7		362	47916	20%	1									

Country	Species		Area	Landings (t) 2004 or 2003	LENGTH - sampling								AGE - sampling								
	English	Scientific			Stratification				Number samples	Number measured	Precision reached (CV)	Method used to calculate precision code	Stratification				Method used for taking age-samples	Number samples	Number aged	Precision reached (CV)	Method used to calculate precision code
					Time interval, code	Space, code	Gear/metier, number of strata	Commercial category, number of strata					Time interval, code	Space, code	Gear/metier, number of strata	Commercial category, number of strata					
Sp	Pouting	*Trisopterus luscus*	VIIIc & IXa		Q	5	10	1		1598		-	Q	5	1	1	ALK		602		A,B
Ita	Swordfish	*Xiphias gladius*	1.3	4222	Q	S	4			3199											
Ita	Swordfish	*Xiphias gladius*	2.1	47	Q	S	1			619											
Gre	Swordfish	*Xiphias gladius*	2.2		Y	S	1			181											A
Ita	Swordfish	*Xiphias gladius*	2.2	4127	Q	S	1			3234											
Gre	Swordfish	*Xiphias gladius*	3.1		Y	S	1			1202											A
Fra	Swordfish	*Xiphias gladius*	Atlantic, Indian, Pacific oceans	1800	Q	-			12	520			-	-	-	-	-				

Footnotes bellow the table, numbered as: country code - number. E.g. NL-1, NL-2, DK-1,DK-2, IT-1, etc. So each country has its private number range.

NO- 1: The spatial strata follows the ICES subdivisions except for subdivision IIa which is divided into 6 statistical regions given by the Norwegian Directorate of Fisheries

NO- 2: The samples consists mostly of direct age samples, but a minor part is samples of ages stratified by length groups. More specifically, 47 of 207 and 16 of 99 for cod and haddock, respectivGrey, are samples of ages stratified by length groups. The method used for estimation uses direct estimates when they are available and a modGreled age length rGreationship for the other

NO- 3: The Bayesian hierarchical modGre is not officially implemented as the method used for calculating precision, but will be implemented for most of the Norwegian stocks during 2005.
NO-4: Similar to No 1, but the precise information is not presently available.

FIN-1: Landings are for year 2003
FIN-2: Finland has derogation for cod sampling. It is assumed that catch composition is similar than for Swedish and Danish cod fisheries.

IRL-1: information from 2003 sampling, empty cGrels not available during WG, could be provided later

E-1: 2003 data
E-2: Only ICES stocks are listed
E-3: Only prGreiminary analytical CVs are listed

S-1: O=functional unit

I-1 Landings reported apply to 2003
I-2 C.V. calculated as an abundance weighted mean over age groups for 2002 and 2003

NL-1: sub = age sampling by cc; full = age sampling from unsorted catch

Annex B: Working documents

Working Document to WKSDFD – San Sebastian 1-3 February 2005

A simple model for comparing age-length keys and filling in missing data, with an application to Celtic Sea cod (*Gadus morhua*)

Hans Gerritsen
Marine Institute, Galway Technology Park, Parkmore, Galway, Ireland.

Abstract

A method is presented to model Age-Length Keys (ALKs) using a Generalised Linear Modelling approach. The (multinomial) aged data are presented to the model in pairs of consecutive age classes so they take on a binomial form (each observation belongs to one of the two age classes). The models can be used for a number of purposes, in the simplest form they can estimate proportions at age for length classes for which no age data is available. It is also possible to test whether two or more ALKs are significantly different. And finally the models might be expanded to Generalised additive models that could take changes over time or in space into account which could improve precision and even estimate ALKs for periods or areas where no sampling took place.

Introduction

Most fisheries stock assessments are based on estimated numbers of fish per age class.

Sampling for age data usually takes place on a length-stratified basis, where a target number of samples is collected for each length class. Additionally, a larger random sample is taken to obtain a length frequency of the catch or landings. To obtain numbers at age, the aged sample is raised up to the total length frequency using an age-length key (ALK), which consists of the proportions at age each for length class. The length-stratified sampling strategy ensures fish from a large size range are represented in the aged sample while optimising the cost of age sampling.

In order to raise the aged data to the total length frequency, the ALK needs to have data for all size classes for which length frequency data are available. There is currently no standardised method for filling in missing data; some form of running average is usually applied. When data are sparse, this approach is unsatisfactory. To my knowledge no model exists that can estimate missing values in an ALK. Another problem is that it is difficult to compare two ALKs. Two different gear types might have catches with different size frequencies, but if the proportions of the age classes at length are not statistically different, they might be combined to increase precision. Existing methods to test differences between ALKs are either awkward to interpret or inappropriate for most data. Hayes (1993) and Horbowy (1998) both suggested comparing each proportion at age for each length class using multiple Fisher's or Chi squared tests. Although the application of these tests is straightforward, the interpretation of the results is not, as there are as many p-values as the number of age and length classes that are considered. When a large number of statistical tests are applied, the likelihood of type I error (false positive) increases. This is usually dealt with by reducing the p-value at which differences are considered significant (Hayes, 1993) or only accepting differences when a certain number of tests return significant results (Horbowy, 1998). However, reducing the likelihood of type I error necessarily increases the likelihood of type II error (false negative) which, in this case, is even more serious that type I error as one would not like to conclude that two ALKs are not different when in fact they are. Dwyer et al. (2004) took a different approach and suggested applying a two-dimensional Kolmogorov-Smirnov test. This approach relieves many of the problems of previous methods but requires the age data to be sampled randomly. Their method might therefore useful to compare age-structures of populations but not ALKs, which are by definition non-random samples from the population. The present paper presents a method to fit a generalised linear model (GLM) to length-at-age data by expressing the proportions at age in pairs of age classes so they take on a binomial form and a logistic model can be applied. The model can be used to predict missing values or to test whether two or more ALKs are significantly different using a single statistical test.

Method

In an ALK, the sample numbers at age by length class are converted to proportions at age by length class. As fish in one length class could belong to a number of age classes, it is difficult to formulate a model to describe these proportions at length. Therefore, the data are expressed in pairs of consecutive age classes so each observation can only belong to one of the two age classes, which allows a binomial model to be fitted. For each pair of age classes the proportions at length are given by:

$$(1) \qquad P_{i,i+1,j} = \frac{n_{i,j}}{n_{i,j} + n_{i+1,j}}$$

where $n_{i,j}$ is the number of fish at length j that are age i and $n_{i+1,j}$ is the number of fish at length j that are age $i+1$. The proportions $P_{i,i+1,j}$ tend to decrease gradually from unity to zero with increasing length and can be described by the logistic function:

$$(2) \qquad \mathrm{logit}(P_{i,i+1,j}) = \beta_0 + \beta_1 \cdot j$$

where j is the fish length, and β_0 and β_1 are the parameters to be fitted. A binomial GLM with a logit link function can be fitted with ‘length’ as a continuous variable, ‘age’ as a factor and a ‘length·age’ interaction to describe $P_{i,i+1,j}$. The interaction term effectively results in a model with a separate curve for each pair of age classes. When fitting the logistic models, the assumptions are made that the underlying age classes have normal length distributions and equal variance.

In order to convert the modelled proportions $P_{i,i+1,j}$ back to an ALK, eq. 1 can be re-written to provide proportions at each age. However, when the number of age classes becomes large, the solution becomes quite long-winded. A simpler approach is to set the numbers at length of the youngest age class at an arbitrary (positive non-zero) value and calculate the numbers at length of the next age class, relative to those of the first: with parameters $P_{i,i+1,j}$ and $n_{i,j}$ given, the values of $n_{i+1,j}$ can be found from eq. 1. These values can now be used to calculate those of the next age class and so on. An ALK can then be constructed from these relative numbers length, although the numbers at length themselves remain arbitrary.

The modelled ALK can be used to fill in expected proportions at age for length classes that have missing data or alternatively, the entire modelled ALK can be used to assign age classes to the length-frequency data. Standard errors of the estimated numbers at age were obtained by bootstrapping the numbers at age in each length class. For every bootstrap replication, this would yield a re-sampled ALK to which the model was applied. A total of 1000 replications were performed for each estimate. The length frequency data was presumed to be known without error so the standard error given here only reflects the variability due to the aged data.

To test the difference between two or more ALKs, terms like gear-type, area or year, can be introduced into the model (along with all second and third level interactions).

It can then be tested if the extra term improves the original model by evaluating the difference in residual deviance between the two nested models against the difference in degrees of freedom of the models, using the Chi-square test (Collet, 2003). If the model is significantly improved by including the extra term, this indicates that there are significant differences between the gear types, areas or years.

The GLM was applied to data taken from the routine port-sampling programme carried out by the Marine Institute in 2003. Random samples of landed fish were measured and length-stratified samples were aged. Ages were allocated according to the number of hyaline rings visible in the otoliths. Samples of Celtic Sea cod (*Gadus morhua*) from the second quarter of 2003 were used to illustrate the method. In ICES area VIIg (eastern Celtic Sea; Fig. 1), cod were caught with three major gear types: beam trawl, otter trawl and Scottish seine. In area VIIj (western Celtic Sea) otter trawl and seine data were available but the seine data only covered two age classes and was therefore omitted from the analysis. Otter trawl gear included both single and twin rigs.

Results

Table 1 shows the length-at-age data for the different gear types and areas. As there were very few a one-year-old fish and fish aged five or older, these age classes were omitted from the analysis. The data in Table 1 were grouped in 2cm length classes for conciseness but the results were nearly identical to the ungrouped data. Even when data are sparse, as in the present case, it is not necessary to group data in larger length classes as the GLM is fitted to the binary data set (success/failure) rather than the binomial data (proportions). This also makes the interpretation of nested models more straightforward. Leverage, standardised deviance residuals and the Cook statistic were checked for extreme values (Collet, 2003). Some observations were identified as influential but all of them fell within the size range expected from previous years and adjacent quarters in ICES area VII, so no observations were excluded.

The first set of models was fitted to data from all gear-types in both areas. Model 1 includes only the terms ‘length’, ‘age’ and their interaction (Table 1), in other words the same model is applied to all gear types and areas. In model 2 the factor ‘stratum’ was introduced. The strata consisted of the different gear types and areas: BeamVIIg; OtterVIIg; SeineVIIg and OtterVIIj. Model 2 is represented graphically in Fig 2. Table 2 shows that inclusion of the factor ‘stratum’ resulted in a significant reduction of the residual deviance (Chi-square test; $p = 0.04$). This indicates that there were significant differences between the strata. To further examine these differences, the effect of stratum was tested for two strata at a time for all possible combinations. The p-values of these analyses of deviance tests (Table 3) might be interpreted as an index of similarity between two strata. Table 3 shows that data from beam trawls in VIIg were significantly different from all other strata. Therefore beam trawl data were removed from the data set and the significance of including the factor ‘stratum’ was tested again: without beam trawl data the difference was no longer significant ($p = 0.41$; models 3 & 4; Table 2) suggesting that it is unlikely that there are differences between the strata in this dataset.

The combined ALKs resulted in very similar parameter estimates to the ALKs from single strata and the modelled ALKs also produced similar parameter estimated to the raw data (Fig 3.). In other words, it appeared that no bias was introduced by either using combined ALKs or using modelled ALKs. Using modelled ALKs instead of the raw data did not seem improve the precision (Fig 3.), while combining ALKs did reduce the standard errors except for age classes with very few samples.

Discussion

The fit of a logistic curve to $p_{i+1,j}$ appears to be quite robust to the assumptions of normality and equal variance. In age classes where growth is rapid, the variance of the length distribution might differ from one age class to the next, but when the difference in means is large, the fit still tends to be good, regardless of the variances. Conversely, when growth is slow, there is little reason for the spread of the length distribution to change, unless mortality or migration are strongly dependent on both size and age. If size selection, e.g. gear selectivity, takes place on all age classes in the same way, the proportions of the age classes at length remain the same and the ALK will not be influenced. For this reason it is usually assumed that the gear-type should not influence the ALK, differences between gear-types found here are likely to reflect spatial differences in the use of the gears; beam trawls are used on different grounds than other trawl types and might catch a different component of the population.

It was found that values with high leverage could be quite influential; removing the largest three-year-old from the beam-trawl data, increased the p-value for the difference between model 1 and model 2 (Table 2) from 0.04 to 0.14, so the decision whether gear-types and areas could be combined, might rest on one observation. In the present case there was no strong argument to remove this observation, as it fell within the size range expected from other years and surrounding areas. However this does highlight the need for a thorough check for outliers and evaluation of observations that might be influential. Influential observations would result in somewhat conservative results: one might conclude that there are significant differences between ALKs based on one or few extreme observations, but one would not conclude that ALKs can be combined if in fact they are different. In the present case, it appears that beam trawl data might not be very different from other gears although differences appeared to be significant. On the other hand, if data are very sparse, there might not be enough information to suggest that there are differences between ALKs when in fact the underlying distributions might be quite different. One might suggest a minimum number of overlapping length classes for each pair of age classes. In the extreme case that there is no overlap between age classes (i.e. when growth is very rapid) it would be advisable to exclude these age classes from the analysis. Although it is possible to fit a logistic curve, the slope would tend towards infinity and the assumptions of the model would be violated.

Despite these caveats, the model seems to produce ALKs that are not biased against the raw data and could therefore provide a standardised method for filling in data for missing length classes. It also provides an objective test to compare ALKs and precision levels could be improved by combining data from different areas or gear types. The current method might also be expanded by modelling the change of the parameters over time, so rather than grouping the data in time periods like quarters; one could model the data of an entire year (or longer). As growth, mortality and migration are not likely to follow a linear pattern, a Generalised Additive Model (GAM) might be the most appropriate approach. A GAM with a smooth term for the sample date could be fitted for each pair of length classes and a modelled ALK would then be available for any moment in time. One might even estimate ALKs for periods where no sampling took place. As the GAM takes the autocorrelation over the entire time period into account, the precision of these models would be improved over the ALKs based on single quarters. A similar approach might be taken on a spatial, rather than temporal basis.

In conclusion: the GLM approach appears to have the potential to become a useful tool to produce estimates for missing length classes, objective comparisons between ALKs and might be developed further to model changes over time or space.

References

Collet, D. (2003) Modelling binary data, 2nd edn. Chapman & Hall/CRC.

Dwyer, K., Koen-Alonso, M., & Walsh, S.J. (2004) Finding the magical minimum sample size: a computer-intensive approach to minimize re-aging effort to construct age-length keys for yellowtail flounder. NAFO SCR, Doc 04/09, 1-4.

Hayes, D.B. (1993) A statistical method for evaluating differences between age-length keys with application to Georges Bank haddock, *Melanogrammus aeglefinus*. Fish. Bull., 91, 550-557.

Horbowy, J. (1998) Comparison of age-length keys of Baltic cod derived from polish commercial and research data. Fish. Res., 36, 257-266.

Figures

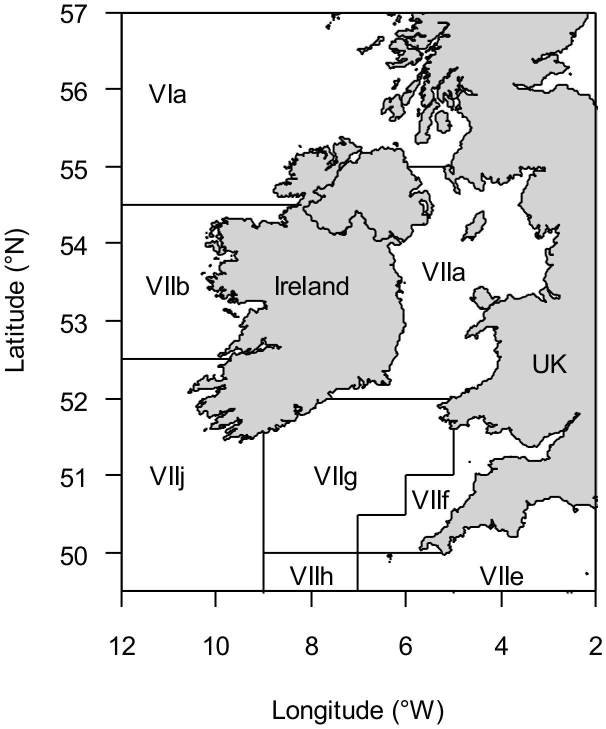


Fig. 1. ICES areas around Ireland. The present samples were taken from areas VIIg and VIIj.

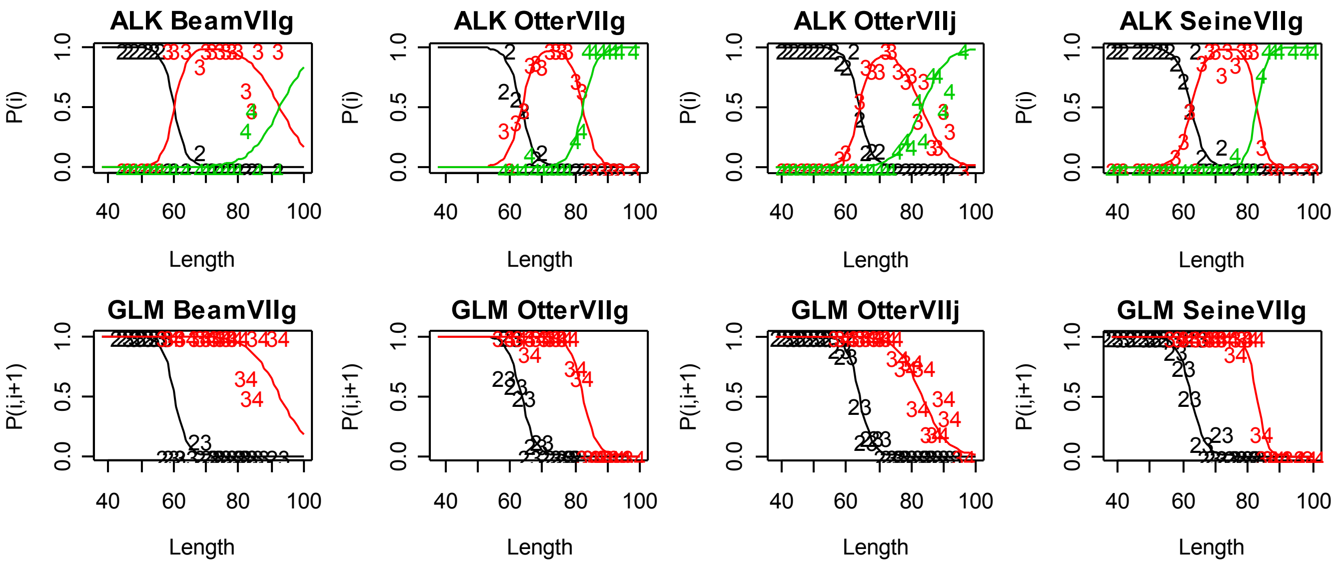


Fig 2. Observed and predicted ALKs ($P_{i,j}$; top row) and observed values of $p_{i,i+1,j}$ and predicted logistic curves (bottom row) from Model 2 in Table 2. Coloured numbers in the plots refer to age classes or pairs of age classes at length (e.g. 23 refers to $P_{2,3,j}$).

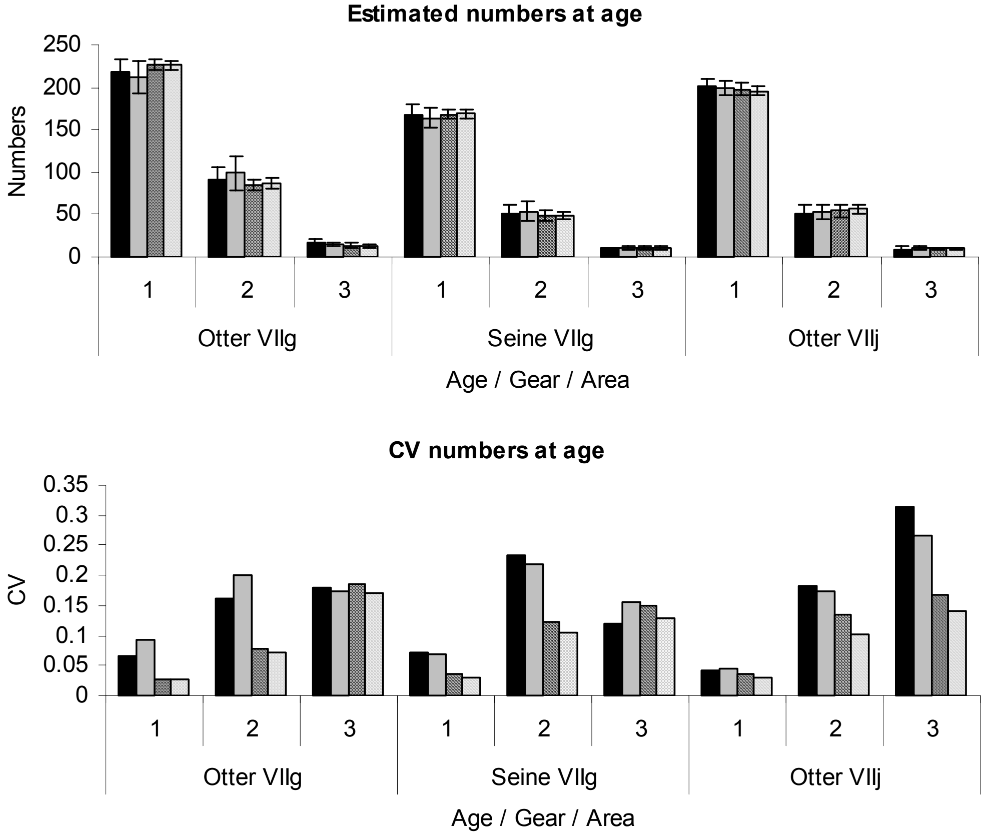


Fig. 3. Estimated numbers at age using single ALKs based on raw data (solid black bars), single modelled ALKs (solid grey bars), combined ALKs using raw data (hashed black bars) and modelled combined ALKs (hashed grey bars) with their standard errors (top panel). The bottom panel shows the Coefficients of Variation of these estimates. Length frequencies were never combined.

Tables

Table 1. Sample numbers at age for the three gear types used in ICES areas VIIg and VIIj during the second quarter of 2003.

Age	Beam VIIg			Otter VIIg			Seine VIIg			Otter VIIj		
Length	2	3	4	2	3	4	2	3	4	2	3	4
38							1					
40							2			5		
42							3			14		
44	4									24		
46	7						10			31		
48	7						6			22		
50	8						10			17		
52	7						7			22		
54	6						4			14		
56	8						2			10		
58		1		2	1		8	1		9	1	
60		1		10			3	1		6	1	
62				3	2		1	1		4		
64		5		2	2		1			3	4	
66					7	1	1	8		1	6	
68	1	6		1	9			8		1	5	
70		6		2	12			9		1	5	
72		9			12		1	4			6	
74		8			14			7			10	
76		6			12			7	1		10	
78		1			11			3			9	2
80		1			3	1		5			4	3
82		2	1		4	2		3			5	1
84		1	1			4		1	4		3	7
86		2				1			4		2	1
88						2			1		1	8
90						2			2		1	4
92		1				2					1	1
94						1			2			2
96												
98						1			1			1
100									1			

Table 2. Analyses of deviance of nested models.

Model	Terms	Df	Dev	ΔDf	ΔDev	P
Strata: BeamVIIg - OtterVIIg - SeineVIIg - OtterVIIj						
Null		918	1206.79			
model 1	length·age	3	313.14			
model 2	length·age·stratum	15	291.67	12	21.47	0.04
Strata: OtterVIIg -SeineVIIg - OtterVIIj						
Null		768	1012.91			
model 3	length·age	3	270.58			
model 4	length·age·stratum	11	262.3	8	8.29	0.41

Note: The models were fitted to data from the second quarter of 2003. The terms included all main effects and interactions. Df indicates the degrees of freedom (null and model), Dev is the deviance (null and residual), ΔDf is the difference in degrees of freedom, ΔDev is the difference in deviance and P is the significance level of the Chi-square test.

Table 3. P-values (Chi-square test) for analysis of deviance as an indication of similarity between the strata (gear type / area).

Stratum	BeamVIIg	OtterVIIg	SeineVIIg	OtterVIIj
BeamVIIg	1	0.02	0.03	0.01
OtterVIIg		1	0.87	0.46
SeineVIIg			1	0.14
OtterVIIj				1

Italian sample survey for landing statistics: links with the biological sampling of landings

Michele De Meo
IREPA Onlus – Via San Leonardo, trav. Migliaro, 84100 Salerno – Italy
demeo@irepa.org

Introduction

In considering the methodologies used in the collection and compilation of fishery statistics in the member countries of the European Economic Area it is essential to keep in mind the large variation in the structure of the fishing industries in those countries. In particular, in Italy (generally for the Mediterranean) there are a very large number of fishing vessels making landings in a large number of mainly small landing places and disposing of their products in markets of a less formal structure than in the other countries. In this situation, to collect accurate fishery data, it is impossible to observe the whole fleet, composed in Italy by about 15000 vessels, and the best system appears to be that of sample surveys.

A sampling procedure for biological parameters and related methods to estimate precision in length and age structure are affected by these considerations. In this context, the sample design for landings and effort data is to be considered the starting point for biological sampling.

The present document describes the principal aspects of the sampling study, developed by IREPA and the Italian National Statistical Institute (ISTAT), for the estimate of the total production of fish, the total value of fish production and the average producers' price, by species and by area of capture, during each calendar month .

Data collection: the stratification and the questionnaire

The elementary analysis units are represented by the landings from vessels included in the sampled fleet. The list from which the sample units will be extracted comprises all the vessels making up the Italian fishing fleet. The sampling base, which is also official, is the Vessel Register (ALP) kept by the Directorate-General of Fisheries and Aquaculture of the Ministry of Agriculture and Forestry Policies. This sampling base excludes ocean-fishery vessels and tuna-fishing vessels belonging to APTS (Associazione Produttori Tonnieri Salernitani), which are recorded on a census base, due to small number of fishing units.

The whole Italian fleet is stratified by using three variables: the geographical area, the fishing techniques and the vessel size. The stratification is designed to maximise the homogeneity of the strata, using known parameters of the population, correlated with the target variables. By crossing these three variables, different strata are defined. Independent samples are selected from each stratum.

The geographical variable is composed by the maritime regions in which the Italian coastline is lawfully divided. In some cases, regions have been sub-divided so that data can be aggregated also per FAO statistical division.

The survey follows a CAI procedure: *computer assisted interviewing.* Two electronics questionnaires are used, namely an annual questionnaire for recording structural information and a weekly questionnaire for activity information. Data processing, for each vessel, is completely computerized.

	LOA < 12	LOA 12-18	LOA 18-24	LOA 24-40
Demersal Trawl	234	1488	890	389
Pelagic Trawl	5	34	32	56
Purse Seiner	37	120	82	64
Dredges	119	584	11	
Passive Gears	9155	427	13	
Polyvalent	528	103	10	
Beam Trawl	13	27	22	25
Long-lines	239	253	79	23

Tab. 1 – Italian fishing fleet by gears and length classes (September 2004).

Implementation of the sample survey: developed procedures and software algorithms

To collect and manage data, IREPA has developed a data managing software using the language Statistica Visual Basic with the software Statistica 6 – StatSoft. This data managing software is composed by three different module: the first (*Data Control*), including all the procedures for the control and the correction of data, the second (*Sampling Procedure*), characterised by procedures developed to estimate the optimum sample allocation and to select sample units, the third (*Estimates and Precision level)*, composed by the procedures to estimates totals and relative coefficients of variation.

Data Control

In the first module, procedures with the aim of recognize and treat *errors* in the data collected have been developed, in order to guarantee final results of certain quality standards and to minimize non sampling errors.

First of all, the study has been structured in such a way that it is possible to control the data already during the collection phase (i.e. during the interview). Data collectors use a specific software for data input; its insertion, can thus only occur if the different logical coherence bonds existing between the measured variables are respected, bonds following a specific *Incompatibility plan*.

Once the data are input in the IREPA server, the records undergo control and correction; in this phase, the scope of controlling the data is to analyse and treat partial non-response errors (PNE) and complete non-response errors (CNE).

With regards to the correction phase, as for the treatment of PNE and CNE, seen the actual possibility of gathering the correct information with a new interview, it was decided that this would be the best strategy to estimate the wrongful and missing variables.

In general, for data control, the sampling units are grouped in appropriate study domains, in order to obtain groups of vessels that are homogenous with respect to the variables observed (catches and prices of different species). The study domains, in the majority of cases, are represented by the strata included by the sampling strategy. Within each single study domain, for each variable, a first control consists in the graphical analysis of the values measured in specific weeks selected for the current year. In an analysis of this type, a value "too distant" from the rest of the data would be seen as suspicious, and it would lead to a second interview of the unit; the "distance" from the remaining points would be evaluated taking into account specific threshold values, established by following time series, and on the basis of auxiliary information relative to the considered period.

With regard to landings, a second analysis is conducted to highlight excessive and suspicious variations from the previous year value; in particular, in such a case, the data will be "cleansed" from the effect of the variable "fishing_days". A similar control is used for the prices observed for individual vessels.

Sampling Procedure

This module includes the procedures for the estimates of the sample size and the selection of the units.

The sampling strategy adopted uses the selection, without replacement, of sampling units by following the *PPS methodology (probability proportional to size*; sample units (vessels) are selected with non-constant selection probability, but proportional to a size measure. The use of such sampling plan and its application, as opposed to simple random sampling, is justified by the intention to use the information provided by the auxiliary variable. This auxiliary variable would obviously have to be known for all the units of the population studied, and it has to be "linked" to the unknown variable that the study is trying to estimate. In statistical terms, the link translates as a "direct relation" between the known auxiliary variable and the targeted variable. The objective of the utilization of the information provided by the auxiliary variable is to obtain a better estimate; the "stronger" this direct relation would be, the smaller the variability of the estimator (i.e. the variance), and the more precise the estimate. In the theoretical extreme situation of exact proportionality, the estimator would have zero variance and it would assume, in every sample, the exact total to estimate.

In this case, the known auxiliary variable is Length Overall (Loa), which use as auxiliary variable was preceded by an explorative analysis which confirmed the hypothesis of proportionality between LOA on one side, and catches and incomes on the other.

For the estimate of the sample size, it was developed the procedure from Bethel (1989); through this, the optimum sample size of the sample for a stratified and multivariate sampling study is estimated. More precisely, some threshold values are fixed (bonded) for the coefficient of variation of the total estimator and for each of the variables observed; due to the precision bonds required at each stratum, the internal sample size for every one is estimated by identifying the "minimum cost" sample. The cost C is defined as:

$$C = c_0 + \sum_{h=1}^{H} c_h n_h$$

where c_0 represents a fixed cost correlated with the organisation of the sampling, c_h represents the sampling costs of any single unit in the stratum h^{th} (h=1 … H), while n_h (n_h=1 … N_h) represents the number units extracted from the h^{th} stratum.

Finally, there is the possibility of establishing a minimum sample size (input as absolute and relative data) to be observed within each stratum; this situation occurs when Bethel's procedure leads sample sizes that, even if respecting the precision bonds fixed for the coefficient of variation of the total estimator, would result too low and of difficult acceptance (for example, variance estimate problems, in case of optimum sample sizes equal to one, may arise). Generally, such limits are fixed to 4 units, in absolute terms, and to 3,5%, in relative terms.

For the selection of the sampling units, the Hanurav-Vijayan's algorithm was implemented (1968); this choice is linked to the different advantageous property of this procedure, described in the next paragraph.

Estimates and Precision levels

This module includes the procedures for the estimates of totals and relative coefficients of variation.

The Hanurav-Vijayan's algorithm allows us to select sample units with a selection probability proportional to the auxiliary variable utilized (LOA) and to "easily" calculate the joint selection probabilities. In this way, it is possible to estimate the total production of fish and the total value of fish production, for each strata, using the Horvitz-Thompson estimator (1952); it is also possible to estimate, with an analytical approach, the variance of these estimates and, consequently, the coefficient of variation, by applying the Sen-Yates-Grundy formula (1953).

For each variable, the estimate of the total is:

$\hat{Y}_h = \sum_{i=1}^{n_h} \frac{y_{hi}}{\pi_{(h)i}}$, where h is the stratum index of the sampling unit;

$\pi_{(h)i} = \frac{n_h LOA_{(h)i}}{LOA_{(h)}}$ is the probability of inclusion of first order for the units i of the stratum h,

where $LOA_{(h)i}$ is the overall length of the unit i of the stratum h, while $LOA_{(h)} = \sum_{i=1}^{n_h} LOA_{(h)i}$.

Obviously, by doing the summatory of each stratum, an estimate of the total of the variable Y is obtained*:*

$$\hat{Y} = \sum_{h=1}^{H} \hat{Y}_h = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \frac{y_{hi}}{\pi_{(h)i}}$$

The estimate of the mean, similarly to the estimate of the totals, is given by $\hat{\bar{Y}}_h = \frac{1}{N_h} \sum_{i=1}^{n_h} \frac{y_{hi}}{\pi_{(h)i}}$, for the individual stratum h, and $\hat{\bar{Y}} = \frac{1}{N} \sum_{h=1}^{H} N_h \hat{\bar{Y}}_h = \frac{1}{N} \sum_{h=1}^{H} \sum_{i=1}^{n_h} \frac{y_{hi}}{\pi_{(h)i}} = \frac{\hat{Y}}{N}$ for the variable total Y.

For the estimate of the total variance the Sen-Yates-Grundy formula is used:

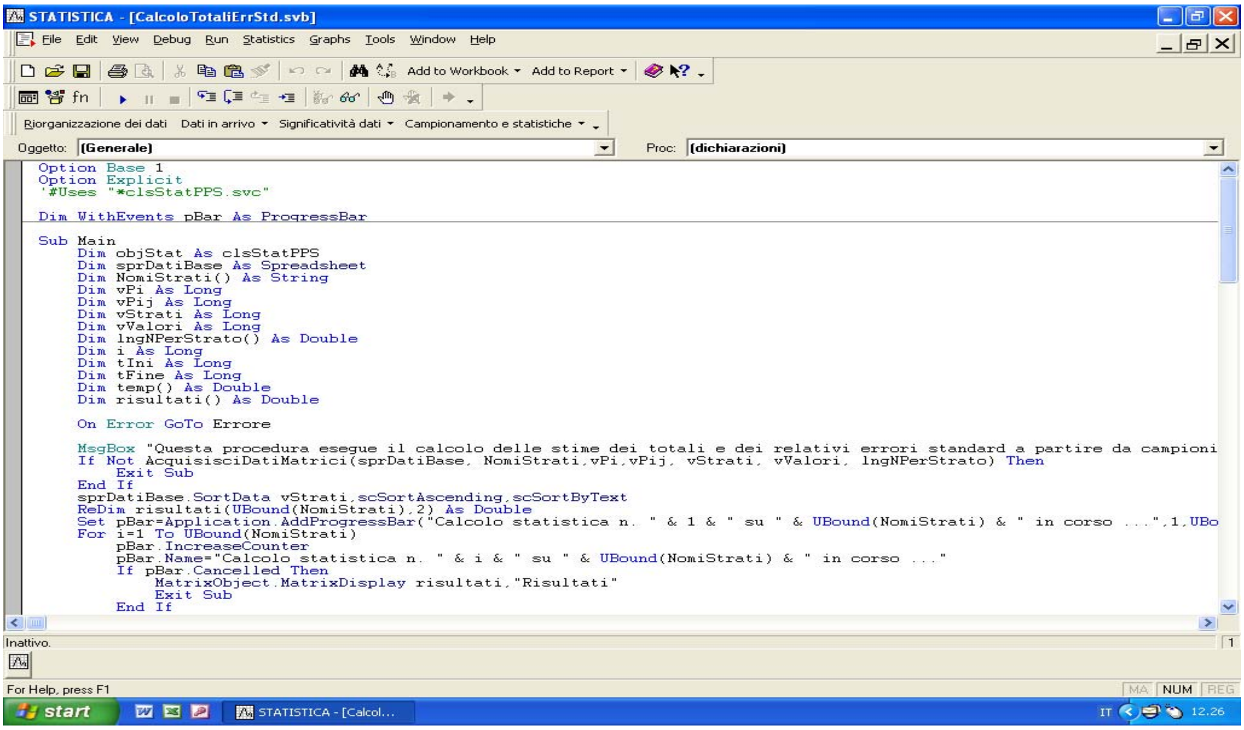
$\hat{\sigma}^2(\hat{Y}_h) = \sum_{i=1}^{n_h} \sum_{j>i}^{n_h} \left(\frac{\pi_{(h)i}\pi_{(h)j}}{\pi_{(h)ij}} - 1 \right) \left(\frac{y_{hi}}{\pi_{(h)i}} - \frac{y_{hj}}{\pi_{(h)j}} \right)^2$, for the individual stratum h, while, having obtained the sample from H independent selections in each stratum, the total variance is obtained from the sum of the internal variances at each individual stratum:

$$\hat{\sigma}^2(\hat{Y}) = \sum_{h=1}^{H} \hat{\sigma}^2(\hat{Y}_h) = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j>i}^{n_h} \left(\frac{\pi_{(h)i}\pi_{(h)j}}{\pi_{(h)ij}} - 1 \right) \left(\frac{y_{hi}}{\pi_{(h)i}} - \frac{y_{hj}}{\pi_{(h)j}} \right)^2 .$$

The estimate of the total and the estimate of the variance, for each single stratum and for all other strata, are used for the estimate of the coefficient of variation of the "estimator of the total":

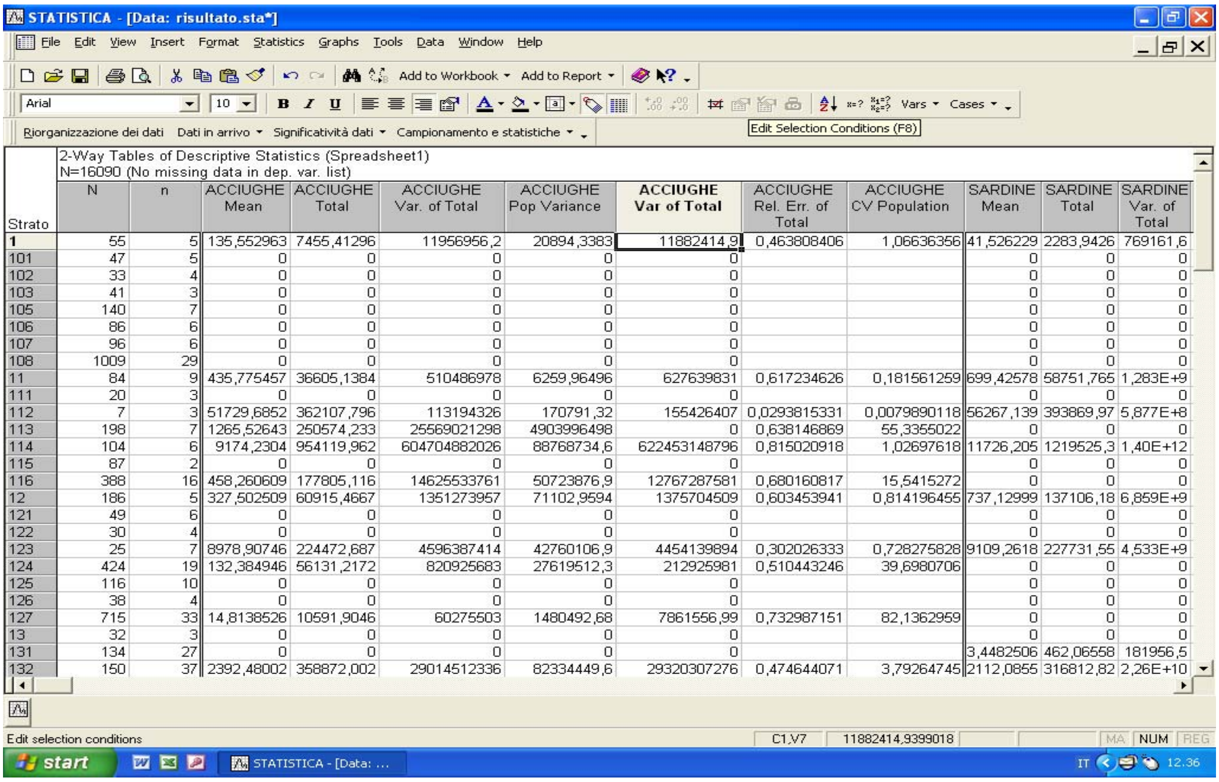
$$\overline{CV}_h = \frac{\hat{\sigma}\left(\hat{Y}_h\right)}{\hat{Y}_h} \quad \text{and} \quad \overline{CV} = \frac{\hat{\sigma}\left(\hat{Y}\right)}{\hat{Y}} .$$

This index is taken into consideration to assess the "goodness" of the obtained estimates.



```
Option Base 1
Option Explicit
'#Uses "*clsStatPPS.svc"

Dim WithEvents pBar As ProgressBar

Sub Main
    Dim objStat As clsStatPPS
    Dim sprDatiBase As Spreadsheet
    Dim NomiStrati() As String
    Dim vPi As Long
    Dim vPij As Long
    Dim vStrati As Long
    Dim vValori As Long
    Dim lngNPerStrato() As Double
    Dim i As Long
    Dim tIni As Long
    Dim tFine As Long
    Dim temp() As Double
    Dim risultati() As Double

    On Error GoTo Errore

    MsgBox "Questa procedura esegue il calcolo delle stime dei totali e dei relativi errori standard a partire da campioni
    If Not AcquisisciDatiMatrici(sprDatiBase, NomiStrati,vPi,vPij, vStrati, vValori, lngNPerStrato) Then
        Exit Sub
    End If
    sprDatiBase.SortData vStrati,scSortAscending,scSortByText
    ReDim risultati(UBound(NomiStrati),2) As Double
    Set pBar=Application.AddProgressBar("Calcolo statistica n. " & 1 & " su " & UBound(NomiStrati) & " in corso ...",1,UBo
    For i=1 To UBound(NomiStrati)
        pBar.IncreaseCounter
        pBar.Name="Calcolo statistica n. " & i & " su " & UBound(NomiStrati) & " in corso ..."
        If pBar.Cancelled Then
            MatrixObject.MatrixDisplay risultati,"Risultati"
            Exit Sub
        End If
```

Tab. 2 – IREPA software: Sen-Yates-Grundy formula and Horvitz-Thompson estimator implemented in Statistica Visual Basic.



2-Way Tables of Descriptive Statistics (Spreadsheet1)
N=16090 (No missing data in dep. var. list)

Strato	N	n	ACCIUGHE Mean	ACCIUGHE Total	ACCIUGHE Var. of Total	ACCIUGHE Pop Variance	ACCIUGHE Var of Total	ACCIUGHE Rel. Err. of Total	ACCIUGHE CV Population	SARDINE Mean	SARDINE Total	SARDINE Var. of Total
1	55	5	135,552963	7455,41296	11956956,2	20894,3383	11882414,9	0,463808406	1,06636356	41,526229	2283,9426	769161,6
101	47	5	0	0	0	0	0			0	0	0
102	33	4	0	0	0	0	0			0	0	0
103	41	3	0	0	0	0	0			0	0	0
105	140	7	0	0	0	0	0			0	0	0
106	86	6	0	0	0	0	0			0	0	0
107	96	6	0	0	0	0	0			0	0	0
108	1009	29	0	0	0	0	0			0	0	0
11	84	9	435,775457	36605,1384	510486978	6259,96496	627639831	0,617234626	0,181561259	699,42578	58751,765	1,283E+9
111	20	3	0	0	0	0	0			0	0	0
112	7	3	51729,6852	362107,796	113194326	170791,32	155426407	0,0293815331	0,0079890118	56267,139	393869,97	5,877E+8
113	198	7	1265,52643	250574,233	25569021298	4903996498	0	0,638146869	55,3355022	0	0	0
114	104	6	9174,2304	954119,962	604704882026	88768734,6	622453148796	0,815020918	1,02697618	11726,205	1219525,3	1,40E+12
115	87	2	0	0	0	0	0			0	0	0
116	388	16	458,260609	177805,116	14625533761	50723876,9	12767287581	0,680160817	15,5415272	0	0	0
12	186	5	327,502509	60915,4667	1351273957	71102,9594	1375704509	0,603453941	0,814196455	737,12999	137106,18	6,859E+9
121	49	6	0	0	0	0	0			0	0	0
122	30	4	0	0	0	0	0			0	0	0
123	25	7	8978,90746	224472,687	4596387414	42760106,9	4454139894	0,302026333	0,728275828	9109,2618	227731,55	4,533E+9
124	424	19	132,384946	56131,2172	820925683	27619512,3	212925981	0,510443246	39,6980706	0	0	0
125	116	10	0	0	0	0	0			0	0	0
126	38	4	0	0	0	0	0			0	0	0
127	715	33	14,8138526	10591,9046	60275503	1480492,68	7861556,99	0,732987151	82,1362959	0	0	0
13	32	3	0	0	0	0	0			0	0	0
131	134	27	0	0	0	0	0			3,4482506	462,06558	181956,5
132	150	37	2392,48002	358872,002	29014512336	82334449,6	29320307276	0,474644071	3,79264745	2112,0855	316812,82	2,26E+10

Tab. 3 - IREPA software: estimates and precision levels.

Final remarks and links with the biological sampling of landings

As described above, a one-stage stratified sample design is used to estimate landings data. The sample units consist in vessels in the official Vessel Register.

Using the Bethel's procedure it is possible to estimate the minimum sample size of vessels that ensure the precision bonds fixed for the coefficient of variation of the total estimator and, applying the Hanurav-Vijayan's algorithm to select sample units, we are able to estimate the coefficient of variation (that is, the precision levels) with an analytical approach (the Sen-Yates-Grundy formula).

Following this procedure, a representative sample of the whole Italian fleet has been identified. Nevertheless this representative sample cannot be used to evaluate the biological composition of landings because it would be too expensive, in time and costs .

Consequently, it is suggested to select a sub-sample of the sampled vessels.

Length and age structure of the landings individuals will be recorded in some randomly selected fishing days.

Finally, the coefficients of variation will be calculated taking into account the statistical methodology for multi-stage sampling besides the guidelines produced by the WKSCMFD (Nantes – 26-30 January 2004).

Appendix

Tab. A – Sampled vessels in 2005.

Fishing techniques	Geographical area	Vessel size	Number of sampled vessels
Beam trawl	ABRUZZO	LOA 18-24	4
Beam trawl	E. ROMAGNA	unspecified	6
Beam trawl	F.V.GIULIA	LOA 12-18	4
Beam trawl	MARCHE	unspecified	12
Beam trawl	VENETO	unspecified	8
Demersal trawl	ABRUZZO	LOA 12-18	5
Demersal trawl	ABRUZZO	LOA 18-24	21
Demersal trawl	ABRUZZO	LOA 24-40	8
Demersal trawl	CALABRIA IONICA	LOA 12-24	4
Demersal trawl	CALABRIA TIRRENICA	LOA 12-24	4
Demersal trawl	CAMPANIA	LOA 12-18	4
Demersal trawl	CAMPANIA	LOA 18-24	5
Demersal trawl	E. ROMAGNA	LOA < 12	4
Demersal trawl	E. ROMAGNA	LOA 12-18	6
Demersal trawl	E. ROMAGNA	LOA 18-24	16
Demersal trawl	F.V.GIULIA	LOA 12-18	4
Demersal trawl	LAZIO	LOA 12-18	4
Demersal trawl	LAZIO	LOA 18-24	13
Demersal trawl	LIGURIA	LOA 12-18	6
Demersal trawl	LIGURIA	LOA 18-24	11
Demersal trawl	MARCHE	LOA 12-18	8
Demersal trawl	MARCHE	LOA 18-24	22
Demersal trawl	MARCHE	LOA 24-40	19
Demersal trawl	MOLISE	LOA 12-24	4
Demersal trawl	MOLISE	LOA 24-40	5
Demersal trawl	PUGLIA IONICA	LOA 12-18	6
Demersal trawl	PUGLIA NORD	LOA < 12	4
Demersal trawl	PUGLIA NORD	LOA 12-18	19
Demersal trawl	PUGLIA NORD	LOA 18-24	35
Demersal trawl	PUGLIA NORD	LOA 24-40	22
Demersal trawl	SARDEGNA	LOA 12-18	11
Demersal trawl	SARDEGNA	LOA 18-24	5
Demersal trawl	SARDEGNA	LOA 24-40	7
Demersal trawl	SICILIA EST	unspecified	4
Demersal trawl	SICILIA NORD	LOA 12-18	4
Demersal trawl	SICILIA NORD	LOA 18-24	4
Demersal trawl	SICILIA SUD	LOA 12-18	42
Demersal trawl	SICILIA SUD	LOA 18-24	30
Demersal trawl	SICILIA SUD	LOA 24-40	81
Demersal trawl	TOSCANA	LOA 12-18	13
Demersal trawl	TOSCANA	LOA 18-24	12
Demersal trawl	VENETO	LOA < 12	10
Demersal trawl	VENETO	LOA 12-18	31
Demersal trawl	VENETO	LOA 18-24	16
Dredges	ABRUZZO	LOA 12-18	5
Dredges	CAMPANIA	unspecified	4
Dredges	E. ROMAGNA	LOA 12-19	8
Dredges	F.V.GIULIA	unspecified	5
Dredges	LAZIO	unspecified	4
Dredges	MARCHE	LOA 12-18	12
Dredges	MOLISE	unspecified	4
Dredges	PUGLIA NORD	LOA 12-18	4

Dredges	VENETO	LOA 12-19	10
Passive gears	ABRUZZO	LOA < 12	15
Passive gears	CALABRIA IONICA	LOA < 12	16
Passive gears	CALABRIA TIRRENICA	LOA < 12	11
Passive gears	CAMPANIA	LOA < 12	36
Passive gears	E. ROMAGNA	LOA < 12	19
Passive gears	F.V.GIULIA	LOA < 12	19
Passive gears	LAZIO	LOA < 12	18
Passive gears	LAZIO	LOA 12-18	4
Passive gears	LIGURIA	LOA < 12	18
Passive gears	MARCHE	LOA < 12	31
Passive gears	MOLISE	LOA < 12	4
Passive gears	PUGLIA IONICA	LOA < 12	18
Passive gears	PUGLIA IONICA	LOA 12-18	4
Passive gears	PUGLIA NORD	LOA < 12	22
Passive gears	SARDEGNA	LOA < 12	61
Passive gears	SARDEGNA	LOA 12-18	7
Passive gears	SICILIA EST	LOA < 12	16
Passive gears	SICILIA EST	LOA 12-18	4
Passive gears	SICILIA NORD	LOA < 12	36
Passive gears	SICILIA NORD	LOA 12-18	4
Passive gears	SICILIA SUD	LOA < 12	32
Passive gears	SICILIA SUD	LOA 12-18	4
Passive gears	TOSCANA	LOA < 12	20
Passive gears	TOSCANA	LOA 12-18	4
Passive gears	VENETO	LOA < 12	18
Passive gears	VENETO	LOA 12-18	9
Pelagic trawl	E. ROMAGNA	unspecified	13
Pelagic trawl	F.V.GIULIA	LOA 12-18	4
Pelagic trawl	MARCHE	LOA 24-40	5
Pelagic trawl	PUGLIA NORD	LOA 24-40	4
Pelagic trawl	VENETO	unspecified	29
Polyvalent	CALABRIA IONICA	LOA < 12	4
Polyvalent	CALABRIA TIRRENICA	LOA < 12	4
Polyvalent	CAMPANIA	LOA < 12	5
Polyvalent	E. ROMAGNA	LOA < 12	4
Polyvalent	LIGURIA	LOA < 12	4
Polyvalent	MARCHE	LOA < 12	4
Polyvalent	SICILIA NORD	LOA < 12	9
Polyvalent	SICILIA NORD	LOA 12-18	4
Purse seiner	ABRUZZO	LOA 24-40	5
Purse seiner	CALABRIA TIRRENICA	unspecified	4
Purse seiner	CAMPANIA	unspecified	15
Purse seiner	F.V.GIULIA	unspecified	4
Purse seiner	LAZIO	unspecified	4
Purse seiner	LIGURIA	unspecified	8
Purse seiner	PUGLIA NORD	unspecified	4
Purse seiner	SARDEGNA	unspecified	4
Purse seiner	SICILIA EST	unspecified	4
Purse seiner	SICILIA NORD	unspecified	4
Purse seiner	SICILIA SUD	unspecified	24
Purse seiner	TOSCANA	unspecified	5
Long-line	CALABRIA IONICA	LOA 12-18	4
Long-line	CALABRIA TIRRENICA	unspecified	4
Long-line	CAMPANIA	LOA < 12	4
Long-line	LIGURIA	unspecified	4
Long-line	MARCHE	LOA 12-18	4
Long-line	PUGLIA IONICA	LOA < 12	4
Long-line	PUGLIA IONICA	LOA 12-18	4

Long-line	PUGLIA NORD	LOA 12-18	4
Long-line	SICILIA EST	LOA < 12	4
Long-line	SICILIA EST	LOA 12-18	8
Long-line	SICILIA EST	LOA 18-24	4
Long-line	SICILIA NORD	unspecified	4
Long-line	SICILIA SUD	LOA < 12	4
Long-line	SICILIA SUD	LOA 12-18	4
Long-line	SICILIA SUD	LOA 18-24	4
TOTAL			**1311**

References

Barcaroli, G., D'Aurizio L., Luzi O., Manzari, A., Pallara, A., (1998) Metodi e software per il controllo e la correzione dei dati, Documenti Istat.

Benedetti R., Espa G., Piersimoni F. (2001) Available methods, techniques and software for survey data editing Casear Conference, Rome, 7 June 2001

Bethel J. (1989) Sample allocation in multivariate surveys, Survey Methodology, 15, 47-57

Brewer, K.R.V., Hanif, M., (1983) Sampling with unequal probabilities, Springer-Verlag, New York.

Chaudhuri, A. (1978) On Estimating the Variance of a Finite Population, Metrika, 25, 65-76.

Cochran, W.G. (1961) Sampling Techniques, Wiley, New York.

Margarucci, E., R., CONCORD v.1.0 – Manuale utente e aspetti metodologici, Documenti Istat.

Deville J.C., Särndal C.E. (1992) Calibration Estimators in Survey Sampling, Journal of the American Statistical Association, 87, 376-382.

EUROSTAT (2002) The collection and compilation of fish catch and landing statistics in member countries of the European economic area

EUROSTAT (2000): Assessment of quality in statistics. Item 4 of the agenda "Standard quality report". Luxemburg 4-5 Aprile 2000

Fox, D.R. (1989), Computer Selection of Size-Biased Samples, The American Statistician, 43(3), 168 -171

Hanurav, T.V. (1967) Optimum utilization of auxiliary information: πps sampling of two units from stratum, Journal of the Royal Statistical Society, B-29, 374-391.

Horvitz, D.G., Thompson, D.J., (1952) A generalization of sampling without replacement from a finite universe, Journal of the American Statistical Association, 47, 663-685.

IREPA. (2001) Osservatorio economico sulle strutture produttive della pesca marittima in Italia. Franco Angeli Eds. Milano, Italy, 312 p., July 2001 (in Italian). Other years available.

IREPA, ISTAT (2001). Statistical sampling design for the estimation of "quantity and average price of fishery products landed each calendar month in Italy by Community and EFTA vessels" (Reg. CE n. 1382/91 modified by Reg. CE n. 2104/93) – working paper of the Working Group "Fishery Statistics ", Joint Eurostat/ICES meeting Luxembourg, 18-19 February 2002

Maselli, M., Signore, M., Panizon, F., (1989) Manuale di tecniche di indagine – Il sistema di controllo della qualità dei dati, Vol. 6 – ISTAT.

Pagliuca D., GENESES v3.0 – Manuale utente e aspetti metodologici, Documenti Istat.

Sabatella E., Manzi C., DeSantis A. (2002) *An harmonized statistical information system for the Italian fishery management*, EAFE (European Association of fisheries economists) Conference, Faro, March 2002

Särndal, C.E., Swensson, B., Wretman J. (1992) *Model Assisted Survey Sampling,* Springer Verlag, New York.

Sen, A. R. (1953), *On estimating of variance in sampling with varying probabilities,* Journal of the Indian Society of Agricultural Statistics, 5, 119-127.

StatSoft, Inc. (2001), *STATISTICA System Reference,* StatSoft, Inc.

Vijayan, K. (1968) *An Exact πps Sampling Scheme - Generalization of a Method of Hanurav*, Journal of the Royal Statistical Society, B-30, 556-566.

Yates, F., Grundy, P.M. (1953) *Selection without replacement from within strata with probability proportional to size*, Journal of the Royal Statistical Society, B-15, 253-261.

A simple tool to calculate biological parameters' uncertainty

P. Sampedro[1], M. Sainza and V. Trujillo[2]

1. IEO, BP.130, 15001. A Coruña (Spain); 2. IEO, BP.1552, 36200 Vigo (Spain)

Abstract

We describe an automatic procedure to estimate uncertainty of some biological parameters: growth, maturation, sex-ratio and length-weight curves. A routine has been developed in R environment, which makes possible to fit the most usual models and to estimate the coefficient of variation for parameters by using the non-parametric bootstrap methodology. The approach is illustrated by an application for Southern Hake stock.

Keywords: Bootstrap, Growth, Hake, Length-Weight relationship, Maturity, Sex-ratio, Uncertainty.

Introduction

The initial intention of this work has been to produce a simple tool to fulfil with the commitment of the EU Regulation (Reg. No 1639/2001) in relation to precision levels of so called "other biological parameters" under module I of the cited Regulation.

For the purpose of this paper biological means: growth (age vs. length and age), maturity (by length and age), length-weight relationship and sex ratio. Biological parameters which are of fundamental importance to assess and advice for fishery managements

Due to the complexity and huge variability of different species, it is quite difficult to develop a common approach and for that it was decided to implement standard models and fits (really, not at all). For some cases, this procedure will not be adequate (e.g.: crustaceans), though it can be used as an exploratory tool to check some assumptions and behaviours of parameters.

To estimate biological parameters and their uncertainties it has been employed resampling techniques which are often used for estimating different measures of uncertainty for any statistics (Efron, 1979). In the last years many investigators have chosen the bootstrap technique to determinate the uncertainty of biological parameters (Brodziak & Mikus, 2000; Bullock *et al.*, 2004; Gros *et al.*, 1987; Meyer *et al.*, 1986; Moguedet & Pérez, 1988; Roa, 1999). The goal of this technique is the easiness of implementation and the non reliance on normal theory.

The fundamental and critical assumption of bootstrapping developed by Efron (1979) is that the observed data are representative of the underlying population.

The main advantages of this tool are:

- Give us estimates of uncertainty levels (precision and bias) for main parameters,
- Non dependent of any assumption on statistical distribution (non-parametric),
- Easy to implement and well-known:
  - It can be used as a tool for quality control,
  - It can be used to optimize sampling designs and sampling intensities,
  - It can be used for exploratory model analysis,
- Allow us to compare uncertainty levels between Countries/Institutes.

Material and Methods

As an example (case study), it has been applied this approach for the southern stock of hake (2002-2004 biological sampling data). The data required consisted in length (in cm), weight (in g), age (in year), maturity (coded as: 0 for immature; 1 for mature), sex (coded as: 1 for males, 2 for females, 3 for unsexed). The results can be obtained by sex or sex combined.

It was chosen the non-parametric bootstrap method. This procedure consists in generate B bootstrap samples, in our case 1000 replicates, by resampling with replacement the original data. Then all statistics for each parameter can be calculated from each bootstrap sample.

The statistics adopted for each estimate were: median and coefficient of variation (relative variation, estimation error). For all the estimates has been plotted their probability profiles ("density's functions") and some plots of model's residuals (model error) to check visually: autocorrelation, homo/hetero-cedasticity, outliers and/or extreme values (maybe, observation errors?), linearity and normality.

Models and fits adopted were:

- Growth at age (vs. Length & Weight):
  von Bertalanffy. Non-linear estimation w. minimum least squares (G-N)

- Maturity (Length & Age):
  GLM. Logistic function. Binomial errors w. maximum log-likelihood fit.

- Length - Weight Relationship:
  Standard. Non-linear estimation w. minimum least squares (G-N)

- Sex-ratio (Length & Age):
  No Model. Percentage by length and age. Cubic spline to plot

Results

We show, as an example, results for mixed sexes for Southern Hake stock. See Tables 1-6 and Figures 1-6.

Conclusions

This tool is simple and powerful and it can be used for many purposes and give us an easy and quick way to produce estimates for the main biological parameters and their associated variability and bias.

References

Brodziak, J. & R. Mikus. 2000. Variation in life history parameters of Dove sole, *Microstomus pacificus,* off the coast of Washington, Oregon, and northern California. *Fishery Bulletin* 98: 661-673.

Bullock, S. H., Turner, R.M., Hastings, J.R., Escoto-Rodríguez, M., Ramírez Apud López, Z. & J.L. Rodríguez Navarro. 2004. Variance on size-age curves: Bootstrapping with autocorrelation. *Ecology*, 85(8): 2114-2117.

Davison, A.C. & D.V. Hinkley. 1999. Bootstrap methods and their application. Cambridge University Press, 575 pp.

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

Gros, P., Laurec, A. & H. Dupouy. 1987. Estimation des parameters de croissance apport des techniques de reechantillonnage (Bootstrap & Jackniffe) et des statistiques robustes (Médianes et modes). *ICES Comité des statistiques,* CM 1987/D: 10, 25pp.

Meyer, J.S., Ingersoll, C.G., McDonald, L.L. & M.S. Boyce. 1986. Estimating uncertainty in population growth rates: Jackknife vs Bootstrap techniques. *Ecology*, 67(5): 1156-1166.

Moguedet, P. & N. Pérez. Estimation of Megrim (*Lepidorhombus whiffiagonis*) growth parameters, for males and females, from the ICES division VII: fitting to the V. Bertalanffy model using, resampling techniques, as well as several adjustable central values (mean, median and mode length at age). ICES CM 1988/G: 9. 18 pp.

Roa, R., Ernst, B & F. Tapia. 1999. Estimation of size at sexual maturity: an evaluation of analytical and resampling procedures. *Fishery Bulletin* 97: 570-580.

```
Species: Southern Hake
Sex: mixed
ORIGINAL DATA: von Bertalanffy
Nonlinear regression model:  model:  length ~ L * (1 - exp(-k * (age - t0)))
data:  parent.frame()
L              K               T0
85.0783325    0.1405240     -0.6200726
residual sum-of-squares:  62466.45

ESTIMATED PARAMETER
Original          Boot     CV boot
Linf  85.07833     85.10914     0.024
k        0.140524      0.1406       0.041
t0     -0.620073     -0.619285   0.058
```

Table 1. Estimates of growth parameters' uncertainty

```
Species: Southern Hake
 ORIGINAL DATA: RESULTS OF GLM
Call:
glm(formula = mat ~ len, family = binomial(link = logit), data = mad.tal.dat)
Deviance Residuals:
   Min       1Q   Median       3Q      Max
-2.3524  -0.7237  -0.4130   0.7194   2.9390
Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -5.585703   0.162287  -34.42   <2e-16 ***
Length         0.134759   0.004325   31.16   <2e-16 ***
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
(Dispersion parameter for binomial family taken to be 1)
Null deviance:       5988.2  on 4885  degrees of freedom
Residual deviance:  4593.0  on 4884  degrees of freedom
AIC: 4597;    Number of Fisher Scoring iterations: 4
ESTIMATED PARAMETERS
       Original    Boot      CV boot
B0     -5.5857  -5.5827     0.027
B1       0.1348    0.1346     0.031
L50  41.4495   41.4726    0.008
```

Table 2. Estimates of maturity parameters' uncertainty at length

```
Species: Southern Hake
 ORIGINAL DATA: RESULTS OF GLM
Call:
glm(formula = mat ~ age, family = binomial(link = logit), data = mad.eda.dat)
Deviance Residuals:
   Min       1Q   Median       3Q      Max
-2.9741  -0.9382  -0.4987   1.1260   2.3662
Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -3.09405    0.15939  -19.41   <2e-16 ***
Age           0.71472    0.03947   18.11   <2e-16 ***
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

(Dispersion parameter for binomial family taken to be 1)
   Null deviance: 2714.5  on 2005  degrees of freedom
Residual deviance: 2266.2  on 2004  degrees of freedom
AIC: 2270.2;  Number of Fisher Scoring iterations: 4
ESTIMATED PARAMETERS
   Original     Boot    CV boot
B0   -3.0941   -3.1052    0.052
B1    0.7147    0.7175    0.055
A50   4.3290   4.3214     0.019
```

Table 3. Estimates of maturity parameters' uncertainty at age

```
Species: Southern Hake
ORIGINAL DATA: RESULTS OF NLS
Nonlinear regression model
  model:  weight ~ a * Length^b
   data:  parent.frame()
        a           b
0.005424983 3.068325537
 residual sum-of-squares:  3797810

 ESTIMATED PARAMETERS
   Original      Boot         CV boot
a  0.005425   0.005427        0.076
b 3.068326   3.068588         0.007
```

Table 4. Estimates of length-weight relationship parameters' uncertainty

Species: Southern Hake

PERCENTAGE OF FEMALES BY LENGTH

Length	%Orig	%Boot	CV-boot	n
14	100.0	100.0	0.000	1
15	25.0	25.0	0.865	4
16	50.0	50.0	0.414	6
17	33.3	33.3	0.424	12
18	46.7	46.7	0.099	120
19	42.9	42.9	0.087	163
20	44.3	44.3	0.078	194
21	49.8	49.8	0.073	207
22	49.2	49.2	0.073	197
23	52.0	52.0	0.068	204
24	49.4	49.4	0.075	166
....	......	....	....	......
....	......	....	....	......
70	100.0	100.0	0.000	1
76	100.0	100.0	0.000	1

Table 5. Estimates of sex-ratio's uncertainties at length

PERCENTAGE OF FEMALES BY AGE

Age	%Orig	%Boot	CV-boot	n
0	35.3	35.3	0.327	17
1	53.2	53.2	0.066	205
2	49.9	49.9	0.044	499
3	48.4	48.6	0.045	535
4	65.7	65.7	0.038	364
5	79.0	79.0	0.036	214
6	82.6	82.6	0.043	121
7	86.0	86.0	0.060	43
8	89.5	89.5	0.079	19
9	100.0	100.0	0.000	3
10	100.0	100.0	0.000	3

Table 6. Estimates of sex-ratio's uncertainties at age

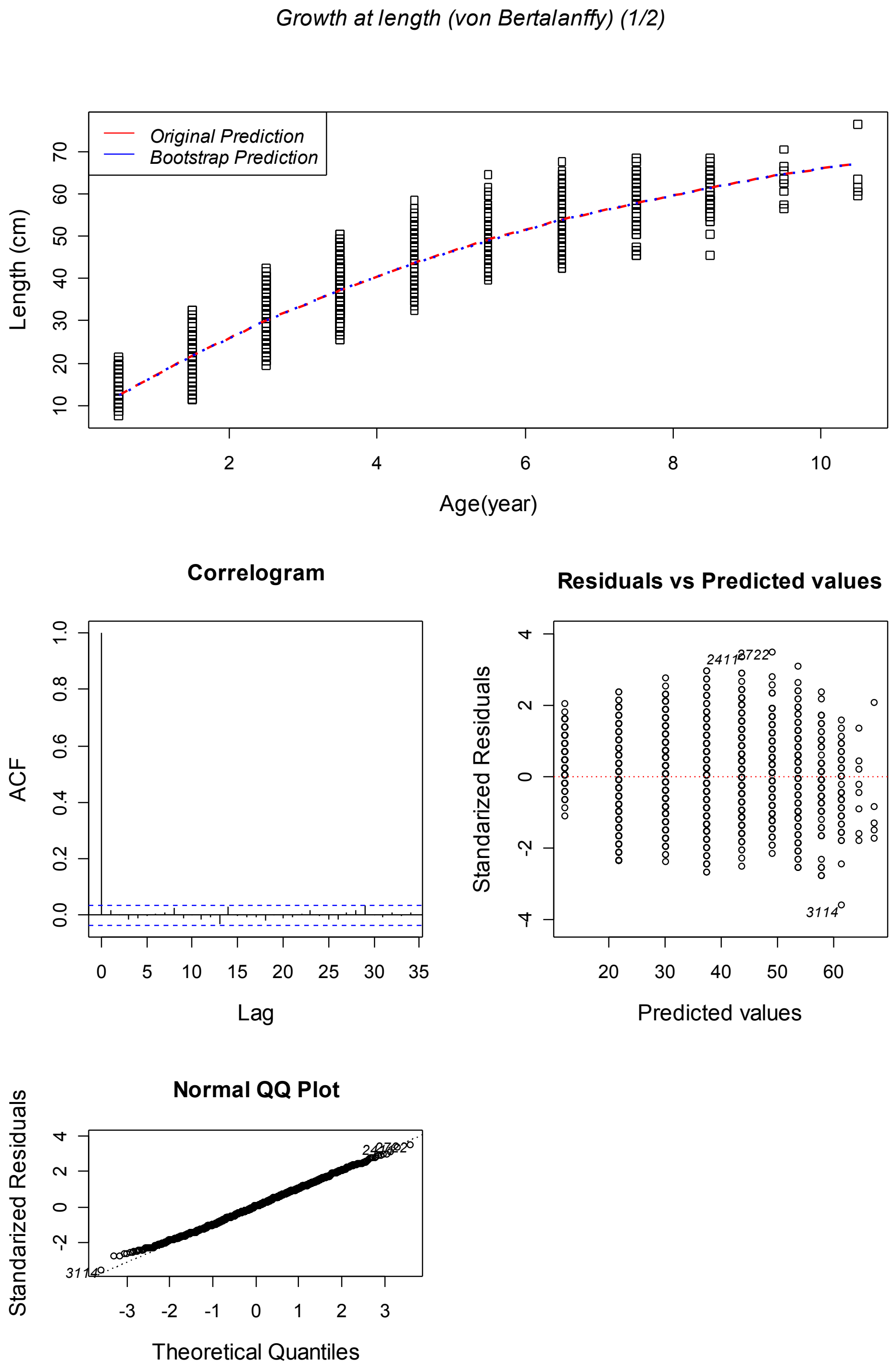


Figure 1. Plot of deterministic and stochastic models and residuals' patterns for growth.

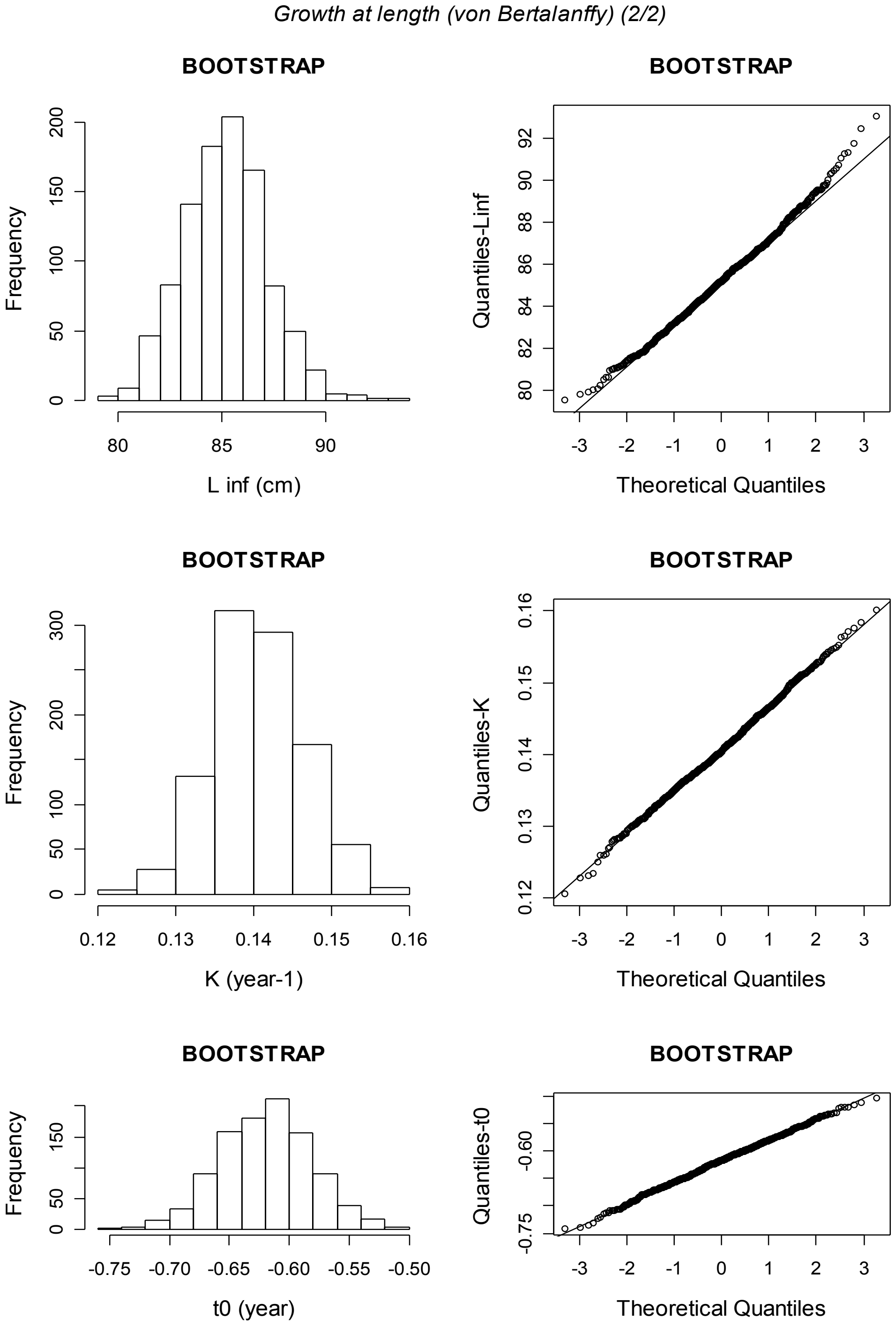
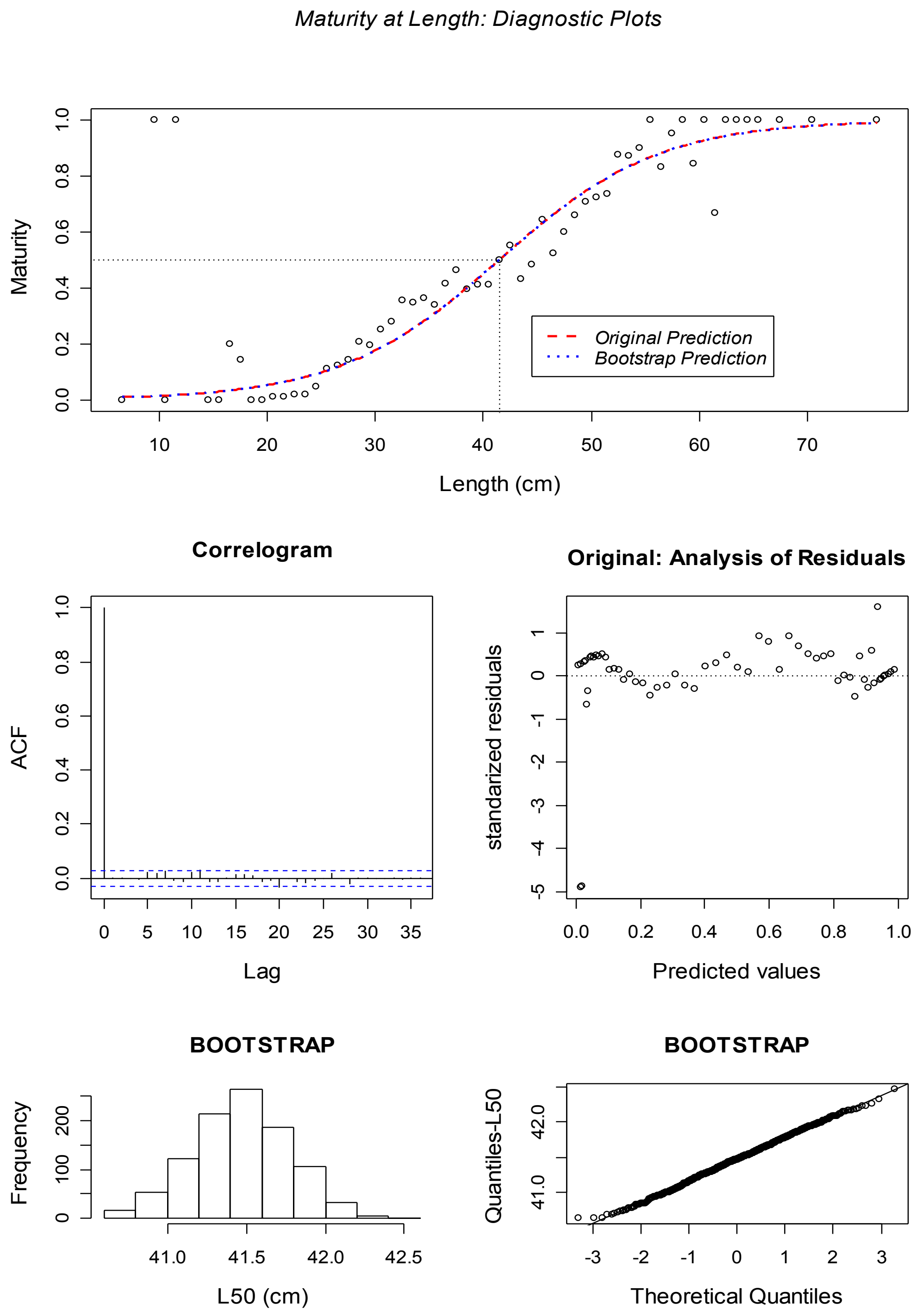


Figure 2. Plot of probability profiles (density's functions) and q-q plots for growth parameters.



Maturity at Length: Diagnostic Plots

Correlogram

Original: Analysis of Residuals

BOOTSTRAP

BOOTSTRAP

Figure 3. Plot of deterministic and stochastic models, residuals' patterns for maturity at length and probability profiles (density's function) and q-q plot for L_{50}.

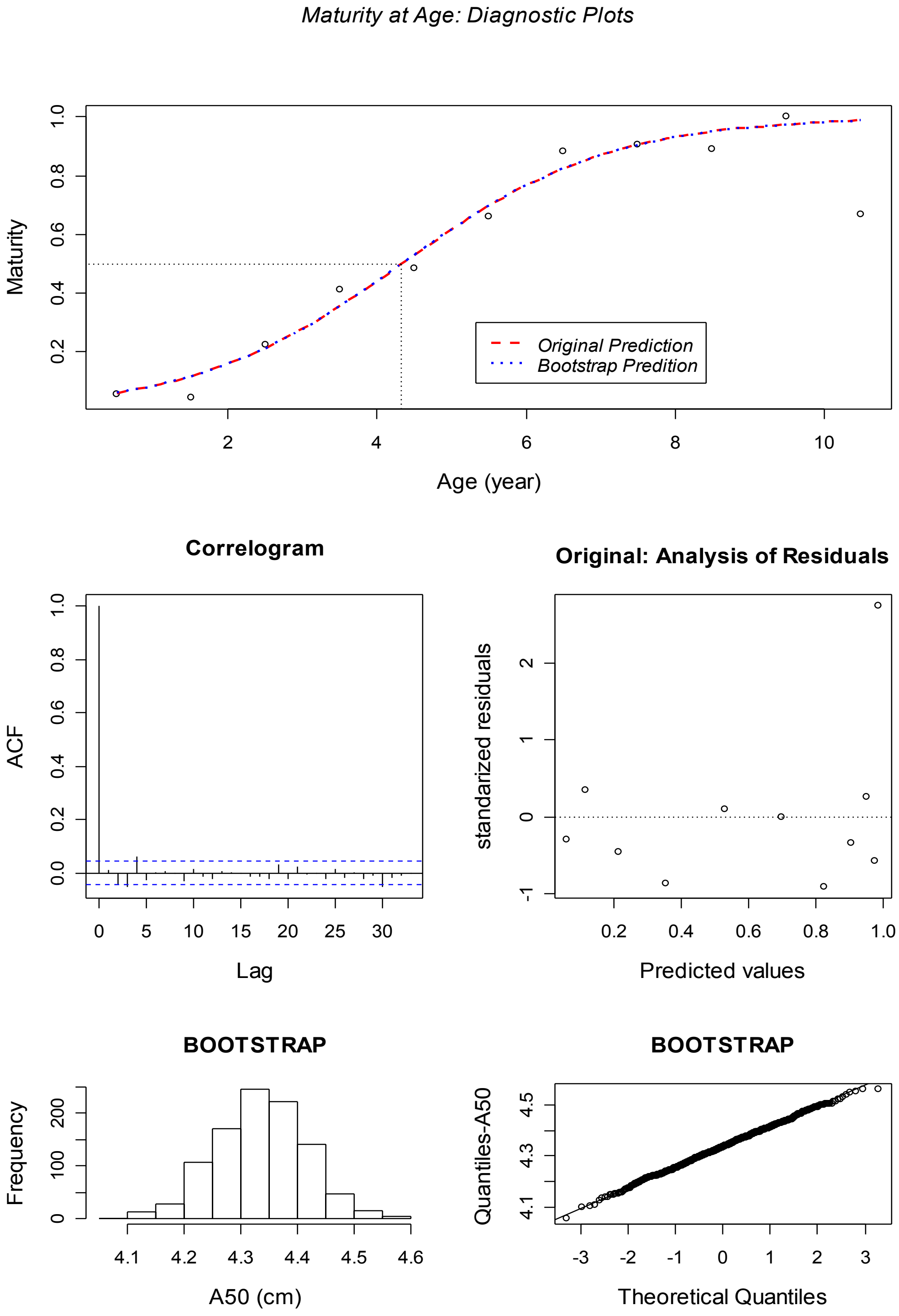


Figure 4. Plot of deterministic and stochastic models, residuals' patterns for maturity at age and probability profiles (density's function) and q-q plot for A_{50}.

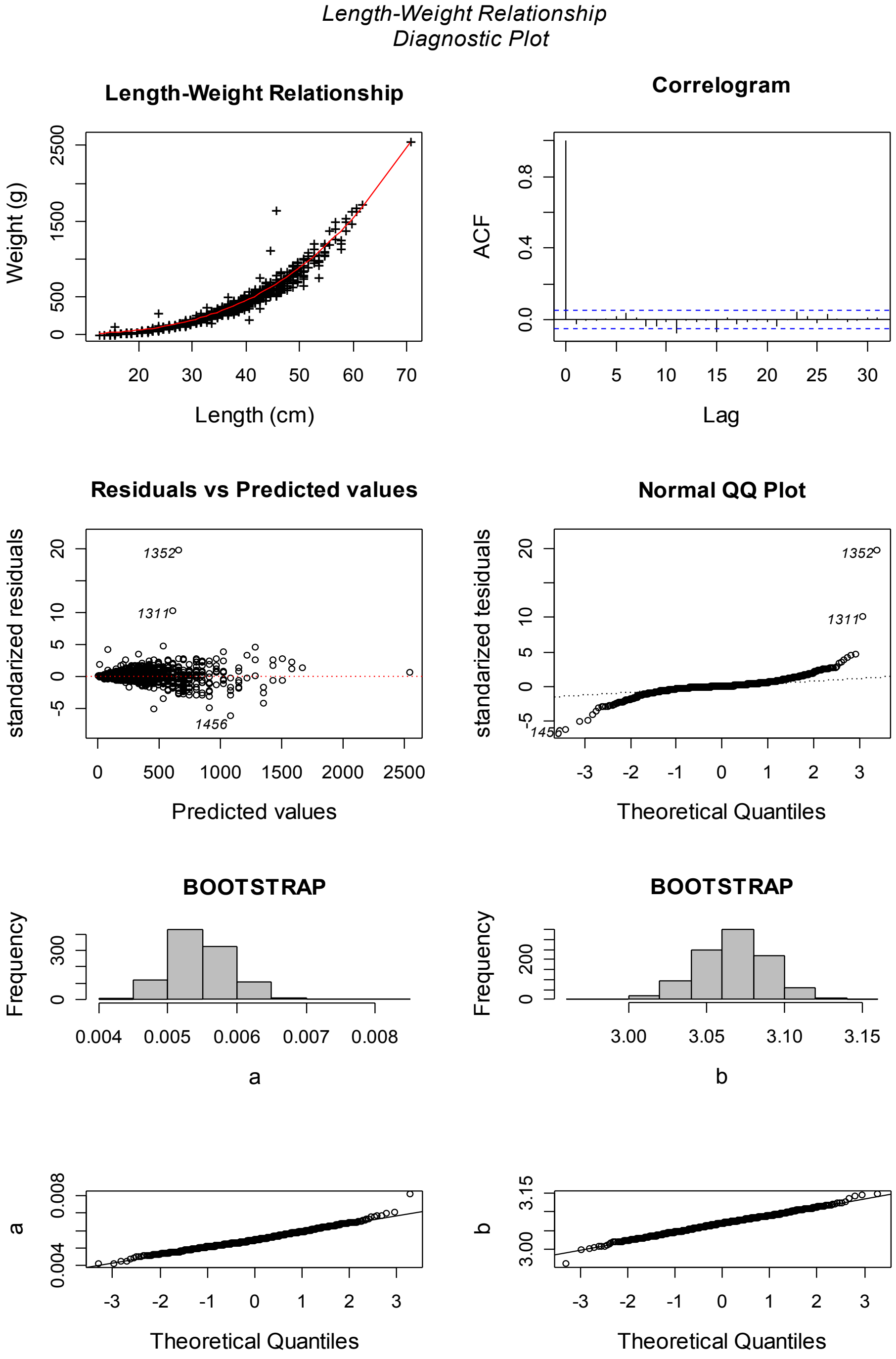
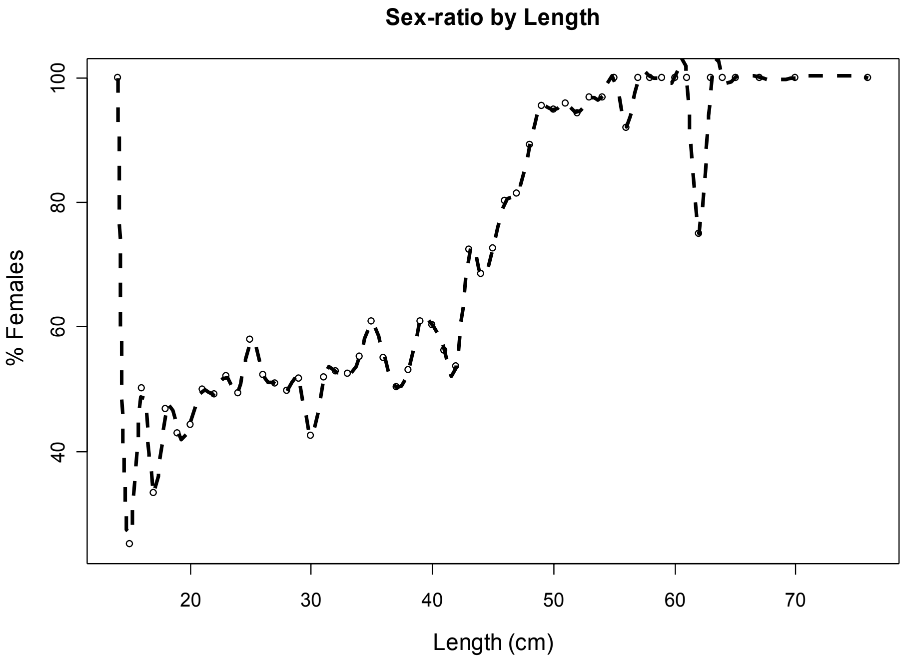


Figure 5. Plot of deterministic and stochastic models, residuals' patterns for length-weight relationship and probability profiles (density's function) and q-q plot for parameters estimated.



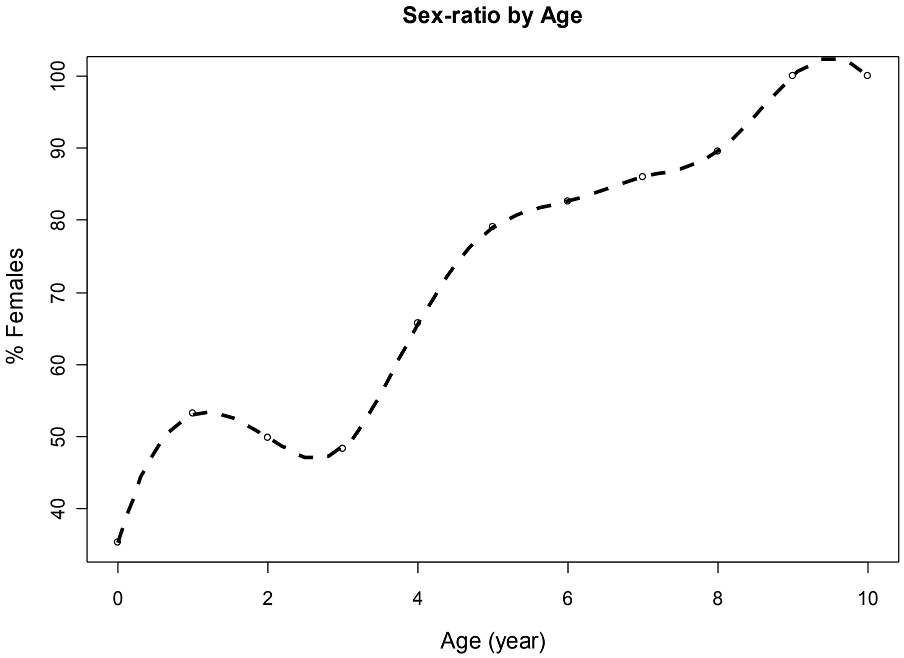


Figure 6. Plot of sex-ratio percentages at length and age.

Annex C: List of participants

NAME	ADDRESS	PHONE/FAX	EMAIL
Sondre Aanes	Institute of Marine Research P.O. Box 1870 Nordnes N-5817 Bergen Norway	T +47 55 238 500 F +47 55 238 531	sondre.aanes@imr.no
Richard Ayers	CEFAS Lowestoft Laboratory Lowestoft Suffolk NR33 0HT United Kingdom	T +44 1502 524 210	r.a.ayers@cefas.co.uk
Ulrich Berth	Bundesforschungsanstalt f. Fischerei Institut für Ostseefischerei An der Jägerbäk 2 D-18069 Rostock-Marienehe Germany	T +40 389 05 113 F +40 389 05 261	ulrich.berth@ior.bfa-fisch.de
Otte Bjelland	Institute of Marine Research P.O. Box 1870 Nordnes N-5817 Bergen Norway	T +47 55 238 500 F +47 55 238 531	otte.bjelland@imr.no
Liz Clarke	Fisheries Research Services Marine Laboratory P.O. Box 101 375 Victoria Road Aberdeen AB11 9DB United Kingdom	T 01224 876544 F 01224 295511	e.d.clarke@marlab.ac.uk
Willem Dekker	Animal Sciences Group, Wageningen UR Nederlands Instituut voor Visserijonderzoek P.O. Box 68 1970 AB IJmuiden Netherlands	T +31 255 564 712 F +31 255 564 644	Willem.Dekker@wur.nl
Michele De Meo	IREPA ONLUS Via S.Leonardo Trav. Migliaro 84131 Salerno Italy	T +39 089 338 978 F +39 089 330 835	demeo@irepa.org
Fabio Fiorentino	Consiglio Nazionale Delle Ricerche (CNR) Instituto per l'Ambiante Marino Costiero (IAMC) Sezione di Mazara del Vallo Via Luigi Vaccara, 61 – 91026 MAZARA DEL VALLO (TP) Italy	T +39 0923 948 966 F +39 0923 906 634	Fabio.fiorentino@irma.pa.cnr.it

NAME	ADDRESS	PHONE/FAX	EMAIL
Ole Folmer	Danish Institute for Fishery Research Charlottenlund Slot DK-2920 Charlottenlund Denmark	T +45 33 96 34 69 F +45 33 96 33 33	ofo@dfu.min.dk
Dorleta Garcia	AZTI Txatxarramendi ugartea, z/g 48395 Sukarrieta (Bizkaia) Spain	T +34 94 602 94 00 F +34 94 687 00 06	dgarcia@suk.azti.es
Hans Gerritsen	Marine Institute Galway Technology Park Parkmore Galway Ireland	T 00353 91 780368 F 00353 91 730470	hans.gerritsen@marine.ie
Marianna Giannoulaki	Greece		Marianna@imbc.gr
Isabel Gonzalez Herraiz	Fundación AZTI Fundazioa Dpto. Investigación Marina-Itsas Ikerketa Saila Txatxarramendi ugartea z/g 48395 SUKARRIETA (Bizkaia) Spain	T +34 94 602 94 00 F +34 94 687 00 06	igonzalez@suk.azti.es
Maria Hansson	National Board of Fisheries Institute of Marine Research Box 4 SE-453 21 Lysekil Sweden		maria.hansson@fiskeriverket.se
David Hirst	Norwegian Computing Center Strathweltie, Tarland, Aboyne AB34 4YS Norway	T +44 1339 881 262	david.hirst@nr.no
Ane Irondo Arrillaga	AZTI (Spain) Itsas Ikerketa Saila Txatxarramendi ugartea, z/g 48395 SUKARRIETA (Bizkaia) Spain	T 94 602 94 00 F 94 687 00 06	airiondo@suk.azti.es
Ernesto Jardim	Portugal IPIMAR Avenida de Brasilia P-1449-006 Lisbon Portugal		enesto@ipimar.pt

NAME	ADDRESS	PHONE/FAX	EMAIL
Kostas Kapiris	Hellenic Center for Marine Research Institute of Marine Biological Resources Agios Kosmas Helliniko 167 77 Athens Greece	T 210 – 9822557 F 210 – 9811713	kkapir@ncmr.gr
Mika Kurkilahti	Finnish Game and Fisheries Research Institute FGFRI Turku game and Fisheries fesearch. Itäinen pitkäkatu 3, FIN-20520 Turku Finland	T +358 20 575 11 F +358 20 575 16 89	mika.kurkilahti@rktl.fi
Sebastian Luyssaert	CLO Sea Fisheries Department Ankerstraat 1 B-8400 Oostende Belgium	T +32 59 342 262 F +32 59 330 629	sebastiaan.luyssaert@dvz.be
David Maxwell	David Maxwell CEFAS Lowestoft Laboratory Lowestoft Suffolk NR33 0HT United Kingdom	T +44 1502 524 328	d.l.maxwell@cefas.co.uk
Hilario Murua	AZTI Foundation Herrera kaia, Portualdea z/g 20110 Pasaia (Gipuzkoa) Spain		hmurua@pas.azti.es
Kay Panten	Bundesforschungsanstalt f. Fischerei Institut für Seefischerei Palmaille 9 D-22767 Hamburg Germany	T +49 4038 905 108 F +49 4038 905 263	kay.panten@ish.bfa-fisch.de
Jukka Pönni	Finnish Game and Fisheries Research Institute (FGFRI), Kotka Unit Sapokankatu 2 48100 Kotka Finland	T +358 205 751 894 F +358 5 234 4040	jukka.ponni@rktl.fi
Inaki Quincoces	AZTI Txatxarramendi Ugartea, z/g 48395 Sukarieta Bizkaia Spain	T +34 94 602 94 00 F +34 94 687 00 06	iquincoces@suk.azti.es
Tiit Raid	Estonian Marine Institute of the University of Tartu Mäealuse 101 12618 Tallinn Estonia		raid@sea.ee

NAME	ADDRESS	PHONE/FAX	EMAIL
Katja Ringdahl	National Board of Fisheries Institute of Marine Research Box 4 SE-453 21 Lysekil Sweden		katja.ringdahl@fiskeriverket.se
Paz Sampedro	Instituto Español de Oceanografía C. O. A Coruña, Muelle de Animas s/n, 15001 A Coruña, Galicia, Spain	T (+34) 981 20 53 62 F (+34) 981 22 90 77	paz.sampedro@co.ieo.es
Norbert Schultz	Bundesforschungsanstalt f. Fischerei Institut für Ostseefischerei An der Jägerbäk 2 D-18069 Rostock-Marienehe Germany		norbert.schultz@ior.bfa-fisch.de
Josefine Svorin	Danish Institute for Fisheries Research Charlottenlund Slot DK-2920 Charlottenlund Denmark	T +45 33 96 34 38 F +45 33 96 33 33	jsv@dfu.min.dk
Valentin Trujillo	Inst. Español de Oceanografía Centro Oceanográfico de Vigo Cabo Estay - Canido Apdo 1552 E-36200 Vigo Spain	T +34 986 492 111 F +34 986 492 351	valentin.trujillo@vi.ieo.es
Andres Velasco	Institut für Ostseefischerei Rostock (IOR) Institute for Baltic Sea Fisheries An der Jägerbäk 2 18069 Rostock Germany	T + 49 (0) 381 810 269 F + 49 (0) 381 810 445	andres.velasco@ior.bfa-fisch.de
Joël Vigneau Chair	IFREMER Avenue du Général de Gaulle B.P. 32 14520 Port-en-Bessin France	T +33 231 515 600 F +33 231 515 601	joel.vigneau@ifremer.fr