# Precision in catch at age data with regard to sampling design

Joël Vigneau and Stéphanie Mahévas

**Abstract**

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

## 1 Introduction

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

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

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

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

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

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

# 2 Method

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

## 2.1 Notations

We use the following notations:

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

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

**Strata 1 : quater**

**Strata 2 : harbours set**

**Strata 3 : commercial category or mtier**

The estimator of the landings at age $i$ is

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

and its variance estimator decomposed into three elements is:

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

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

## 2.2 Analytical variance formulation

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

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

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

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

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

### 2.2.1 Stratified sampling by commercial category

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

Some additionnal notations for the stratified sampling by commercial category :

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

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

The variance estimator of the total landings is the following.

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

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

**Variance of landings at length**

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

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

The estimator of the variance is

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

and from Cochran (1977),

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

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

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

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

**Variance of the proportion of landings at age $i$**

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

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

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

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

### 2.2.2 Stratified sampling by mtier

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

Let us introduce some additionnal notations :

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

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

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

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

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

The estimate of the landings of length $j$ is expressed as

$$\begin{aligned} \hat{D}_j &= \sum_{l=1}^{L} \hat{D}_{jl} & (3)\\ &= \sum_{l=1}^{L}\sum_{v=1}^{n_l} w_l \frac{d_{jlv}}{w_{lv}} & (4)\\ &= \sum_{l=1}^{L} w_l \sum_{v=1}^{n_l} \frac{\sum_{k=1}^{K} d_{jlvk}}{\sum_{k=1}^{K} w_{lvk}} & (5)\end{aligned}$$

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

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

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

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

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

**Variance of the proportion of landings at age $i$**

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

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

## 2.3 Analytical tool for samples exploratory analysis

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

1. given a strata $k$,

(a) given a length class $j$

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

(b) over all length classes

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

2. over all strata $k$

(a) given a length class $j$

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

(b) over all length classes

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

Note that $\Delta_{jv}$ matchs with the last part of equation 2.

### 2.4 Resampling method

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

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

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

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

set-up :

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

Bootstrap loop repeated at each iteration :

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

final calculation

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

### 2.4.2 Bootstrap of stratified sampling catch-at-length estimator

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

The algorithm for such a bootstrap is as follows :

set-up :

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

Bootstrap loop repeated at each iteration :

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

final calculation

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

### 2.4.3 Bootstrap catch-at-age estimator

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

The algorithm for such a bootstrap is as follows :

set-up :

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

Bootstrap loop repeated at each iteration :

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

final calculation

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

### 2.5 Simulation

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

set-up :

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

Simulation double loop

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

final graph

- draw a contour plot of the double loop precision matrix

## 3 Materials

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

## 3.1 Atlantic hake

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

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

Age sampling scheme combines different methodologies

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

## 3.2 Eastern Channel sole

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

# 4 Results

## 4.1 exploratory analysis of the samples

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

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

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

## 4.2 catch-at-length precision

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

## 4.3 Age length keys precision

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

### 4.4 catch-at-age precision

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

## 5 Results

### 5.1 exploratory analysis of the samples

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

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

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

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

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

## 5.2 catch-at-length precision

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

## 5.3 Age length keys precision

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

## 5.4 catch-at-age precision

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

# 6 discussion

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

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

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

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

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

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

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

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

Analytical variance at age decomposition (Absolute and relative)

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

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

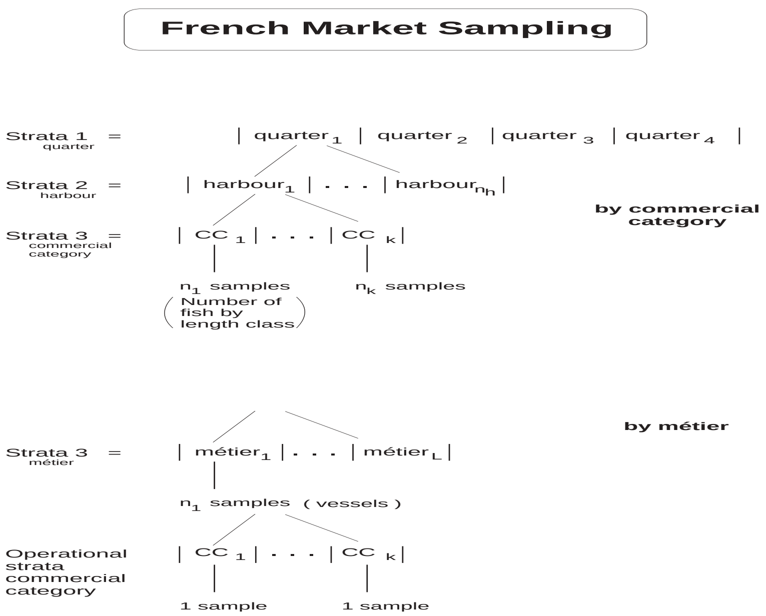


Figure 1: *French market sampling*

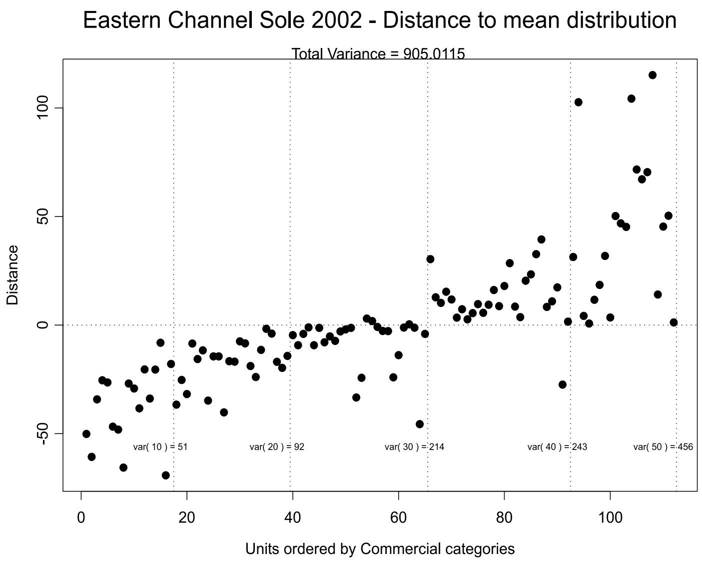


Figure 2: *Sole VIID 2002 - Distance to mean distribution. Samples ordered by commercial category*

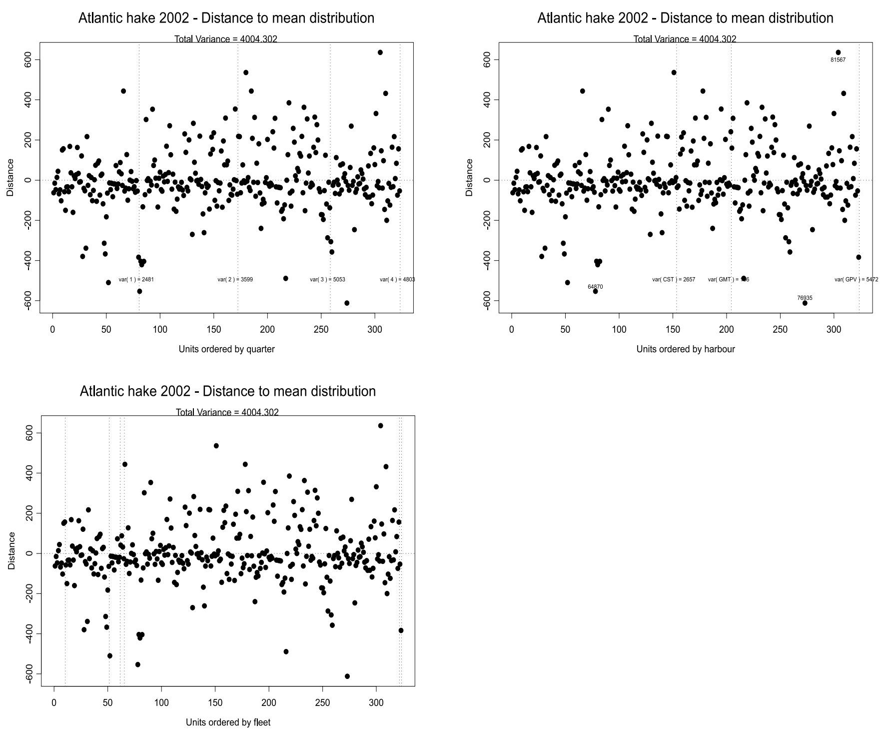


Figure 3: *Atlantic hake 2002 - Distance to mean distribution*

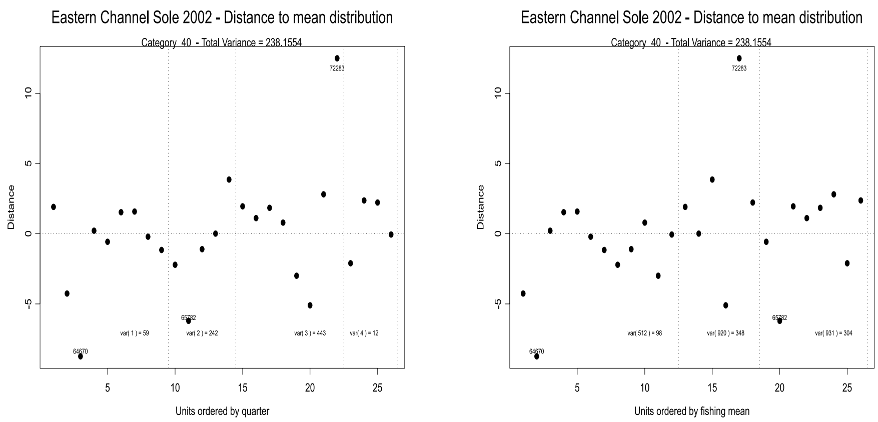


Figure 4: *Sole VIID 2002 - Distance to mean distribution for one commercial category. Samples ordered by a) quarter b) fishing mean*

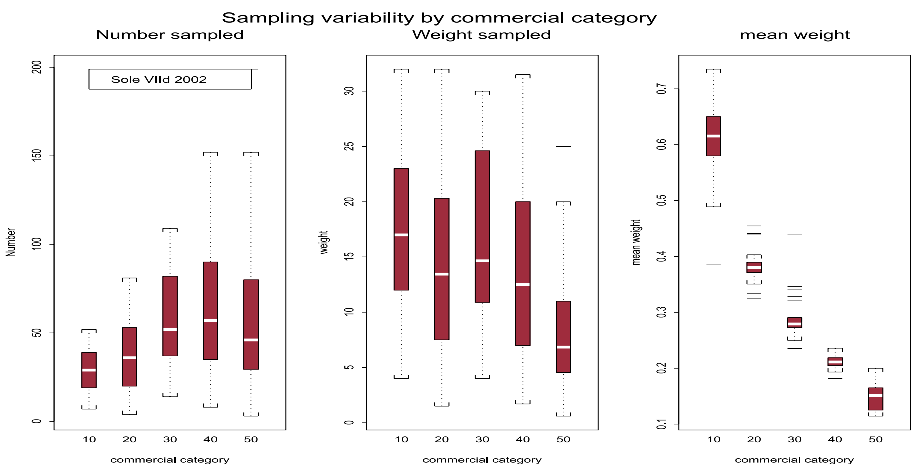


Figure 5: *Sole VIID 2002 - Sampling variability a) in number of fish measured by sample, b) sample weight and c) mean weight of a fish by sample*

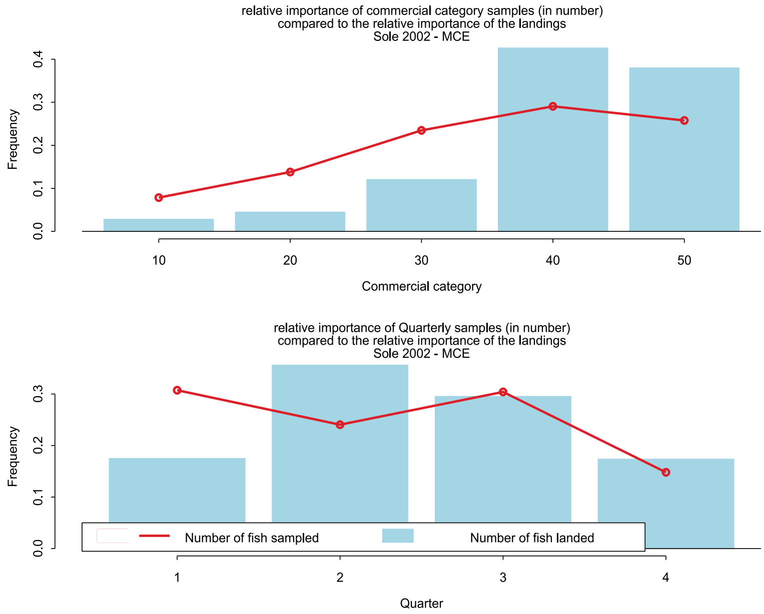


Figure 6: *Sole VIID 2002 - Relative importance of sampling effort against relative importance of landings by a) commercial category b)quarter*

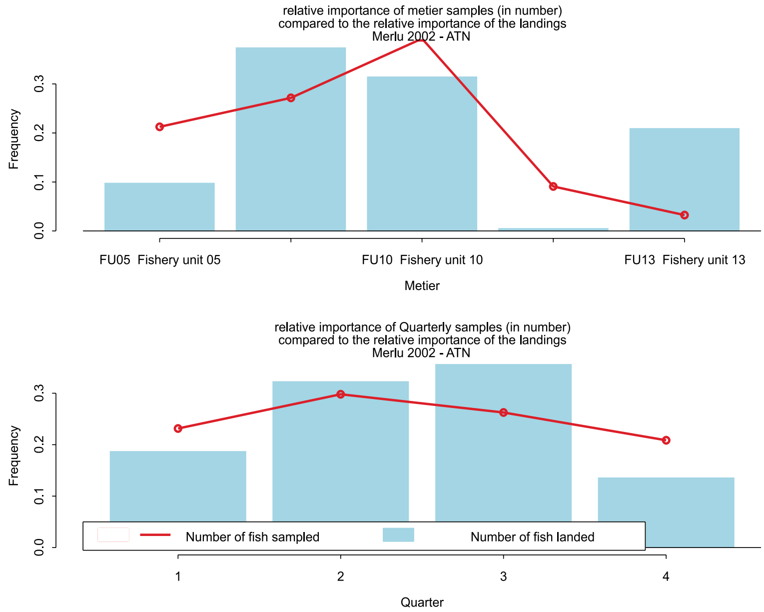


Figure 7: *Atlantic hake 2002 - Relative importance of sampling effort against relative importance of landings by a) fishery unit b) quarter*

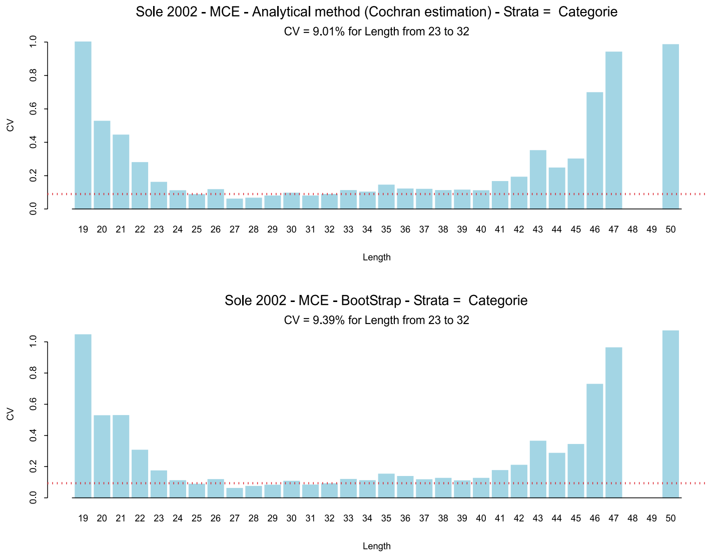


Figure 8: *Sole VIId 2002 - Precision in length distribution a) analytical b) bootstrap*

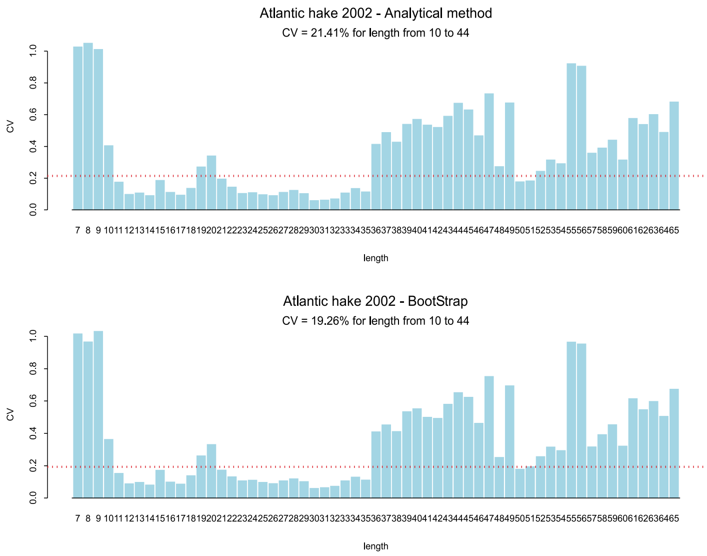


Figure 9: *Atlantic hake 2002 - Precision in length distribution a) analytical b) bootstrap*


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

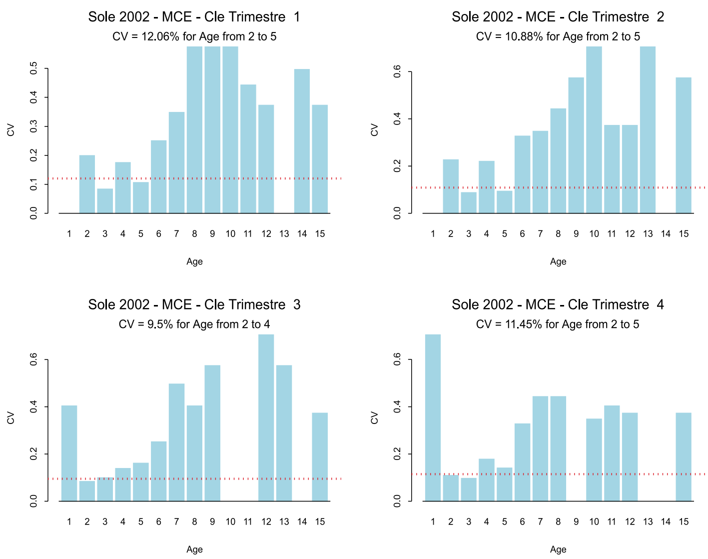


Figure 11: *Sole VIId 2002 - Precision in quarterly age length keys*

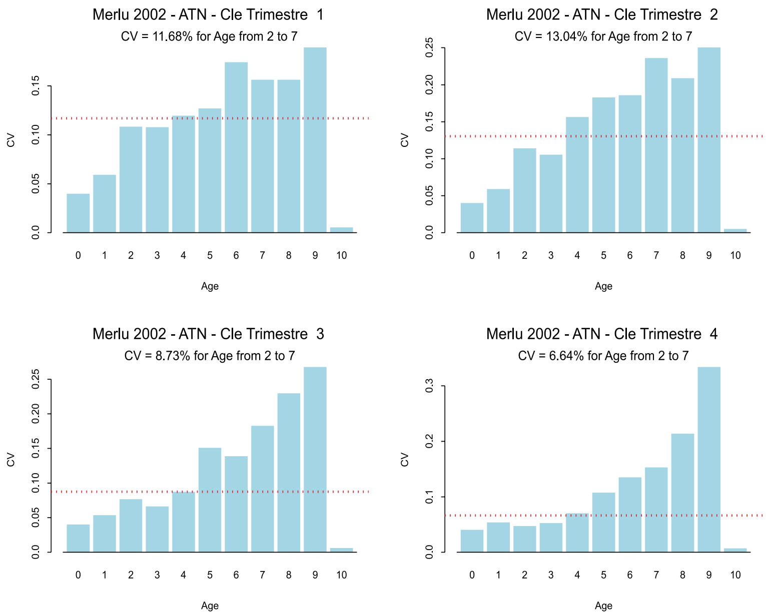


Figure 12: *Atlantic hake 2002 - CV at age in quarterly age length keys*

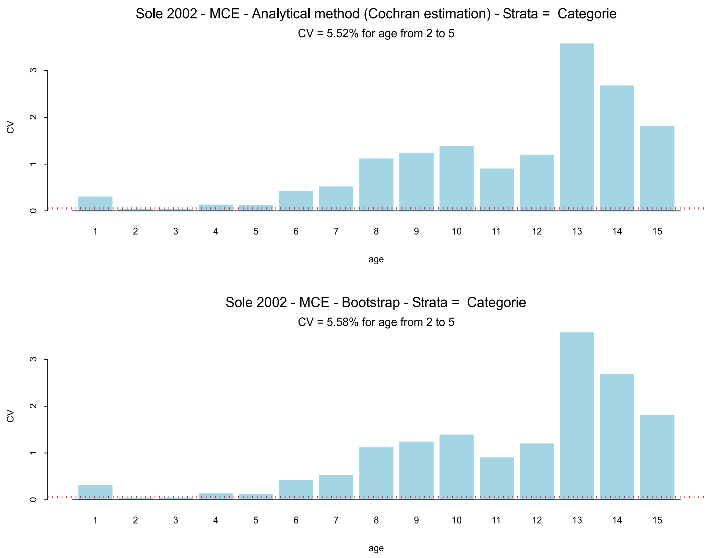


Figure 13: *Sole VIId 2002 - Precision in age distribution*

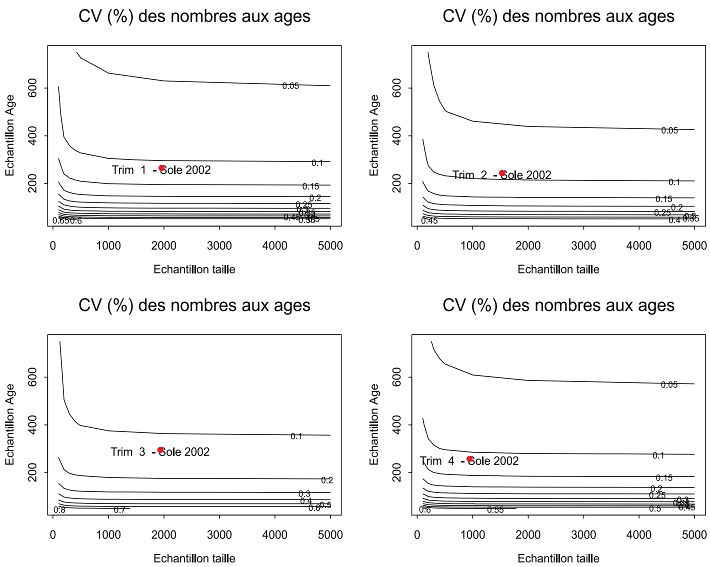


Figure 14: *Sole VIId 2002 - simulation of precision from different sampling effort*