RELATÓRIOS CIENTÍFICOS E TÉCNICOS

SÉRIE DIGITAL

REPOT OF THE WORKSHOP ON SAMPLING EFFORT FOR BIOLOGICAL PARAMETERS (WKSEBP)

Cristina Silva, Manuela Azevedo, Corina Chaves, Rui Coelho, Ana Maria Costa, David Dinis, Sandra Dores, Ana Cláudia Fernandes, Patrícia Gonçalves, Pedro G. Lino, Hugo Mendes, Teresa Moura, Cristina Nunes, Melinda Oroszlányová, Daniel Pinto e Maria do Carmo Silva

2017

17

RELATÓRIOS CIENTÍFICOS E TÉCNICOS DO IPMA – SÉRIE DIGITAL

Destinam-se a promover uma divulgação rápida de resultados de carácter científico e técnico, resultantes da actividade de investigação e do desenvolvimento e inovação tecnológica nas áreas de investigação do mar e da atmosfera. Esta publicação é aberta à comunidade científica e aos utentes, podendo os trabalhos serem escritos em Português, Francês ou Inglês.

Edição
IPMA
Rua C – Aeroporto de Lisboa
1749-007 LISBOA
Portugal

Edição Digital
Anabela Farinha

As instruções aos autores estão disponíveis no sitio web do IPMA
http://ipma.pt
ou podem ser solicitadas aos membros do Corpo Editorial desta publicação

Capa
Conceição Almeida

ISSN
2183-2900

Workshop on Sampling Effort for Biological Parameters - WKSEBP

IPMA, 18-20 April 2017

programa nacional de amostragem biológica

Citation:

Silva C, Azevedo M, Chaves C, Coelho R, Costa AMC, Dinis D, Dores S, Fernandes ACF, Gonçalves P, Lino PG, Mendes H, Moura T, Nunes C, Oroszlányová M, Pinto D, Silva MC, 2017. Report of the Workshop on Sampling Effort for Biological Parameters (WKSEBP), IPMA, Lisbon 18 – 20 April. ***Relat. Cient. Téc. do IPMA (http://ipma.pt)***, ***nº 17***, **55 p + 4 Anexos**.

Report of the Workshop on Sampling Effort for Biological Parameters (WKSEBP)

Cristina Silva, Manuela Azevedo, Corina Chaves,
Rui Coelho, Ana Maria Costa, David Dinis, Sandra Dores,
Ana Cláudia Fernandes, Patrícia Gonçalves, Pedro G. Lino,
Hugo Mendes, Teresa Moura, Cristina Nunes,
Melinda Oroszlányová, Daniel Pinto and Maria do Carmo Silva

IPMA - Divisão de Modelação e Recursos da Pesca (DivRP) Av. Brasília, 1449-006 Lisboa

RESUMO

Título: Relatório do Workshop sobre o Esforço de Amostragem para a Estimação de Parâmetros Biológicos

O "Workshop on Sampling Effort for Biological Parameters (WKSEBP)", presidido pelas investigadoras Manuela Azevedo e Cristina Silva (IPMA), decorreu no IPMA-Algés, de 18 a 20 de Abril de 2017, para analisar o actual Programa Nacional de Amostragem Biológica (PNAB/Data Collection Framework) com o objectivo de otimizar o esforço de amostragem e melhorar a precisão na estimação de parâmetros biológicos. Foram analisados e discutidos vários métodos e abordagens que resultaram num conjunto de recomendações para trabalho futuro. Foram desenvolvidos quatro casos de estudo centrados nos seguintes temas: 1) amostragem de comprimentos em campanhas de investigação, 2) amostragem de comprimentos em lota, 3) estimação de chaves de idade-comprimento e 4) determinação de ogivas de maturação. Em cada caso de estudo foi analisado e discutido o número de amostras, o tamanho efetivo de cada amostra e a precisão na estimação dos parâmetros biológicos.

Palavras chave: amostragem por comprimentos, chave comprimento-idade, ogiva de maturação, tamanho efectivo da amostra.

ABSTRACT

The Workshop on Sampling Effort for Biological Parameters (WKSEBP), chaired by Manuela Azevedo and Cristina Silva (IPMA) met in Lisbon, 18 – 20 November 2017, to focus on the analysis of the current Portuguese sampling designs under PNAB/DCF (Programa Nacional de Amostragem Biológica/Data Collection Framework) with the aim to optimize the sampling effort and improve precision on the estimation of biological parameters. Several approaches and methodologies were discussed and guidelines for future work were recommended. Four case-studies were analyzed focusing on: 1) survey sampling for length, 2) at-market sampling for length, 3) estimation of age-length key and 4) maturity ogive. The number of samples, the effective sample size and the precision in parameters estimation were discussed in each case-study.

Key words: length sampling, age-length key, maturity ogive, effective sample size

REFERÊNCIA BIBLIOGRÁFICA

Silva C, Azevedo M, Chaves C, Coelho R, Costa AMC, Dinis D, Dores S, Fernandes ACF, Gonçalves P, Lino PG, Mendes H, Moura T, Nunes C, Oroszlányová M, Pinto D, Silva MC. 2017. Report of the Workshop on Sampling Effort for Biological Parameters (WKSEBP). ***Relat. Cient. Téc. do IPMA(http://ipma.pt)***, ***nº 17***, **55 p+ 4 Anexos**.

Table of contents

1 Introduction

1.1 Terms of reference

The **Workshop on Sampling Effort for Biological Parameters** (WKSEBP), chaired by Manuela Azevedo and Cristina Silva met in Lisbon, 18–20 April 2017, to focus on the analysis of sampling effort needed to estimate biological parameters with a certain precision. Data used in the analysis were collected under PNAB/DCF (Programa Nacional de Amostragem Biológica/Data Collection Framework) and the main objective was to optimize the number of samples to collect, in terms of time and costs saving. Four case-studies were presented and analyzed during the workshop: hake

- CS 1) Surveys sampling for assessing the precision of length-frequency estimates: hake (*Merluccius merluccius*) and horse mackerel (*Trachurus trachurus*).
- CS 2) At-market sampling for landings length composition: hake commercial size categories.
- CS 3) Sampling for ALKs: blue-whiting (*Micromesistius poutassou*).
- CS 4) At-market sampling for maturity ogive: mackerel (*Scomber scombrus*) and hake.

1.2 Background

The Workshop was organized within the scope of the National Biological Sampling Programme - PNAB/DCF.

1.3 Conduct of the meeting

The workshop participants made available several sets of data and scripts prepared in advance to the meeting, as well as several presentations (Annex 4) which subsequently formed the basis of the workshop's investigations and discussions during the week.

The following speakers presented the talks indicated:

<u>Case Study 1</u>

P01 – Melinda Oroszlányová: Assessing the precision of length-frequency estimates (Following the paper of M. Pennington, 2002)

<u>Case Study 2</u>

P02 – Manuela Azevedo: At-market sampling for landings length composition: hake commercial size categories case-study. M. Azevedo, C. Silva

<u>Case Study 3</u>

P03 – Patrícia Gonçalves: Sampling for ALKs – blue-whiting case-study

<u>Case Study 4</u>

P04 – Ana Maria Costa: Southern mackerel, *Scomber scombrus*, mature ogive. A. Costa, C. Nunes, M. C. Silva.

During the workshop the participants were divided into four subgroups, according to the case-studies:

Subgroup 1 – Effective sampling size for biological parameters in research surveys.

Subgroup 2 – Effective sampling size in at-market sampling for size-categories.

Subgroup 3 – Sampling for ALKs.

Subgroup 4 – Sampling for maturity ogive.

1.4 Structure of the report

The structure of the report is as follows:

Section 2 describes the work developed during the workshop, related to CS 1.

Section 3 describes the work developed during the workshop, related to CS 2.

Section 4 describes the work developed during the workshop, related to CS 3.

Section 5 describes the work developed during the workshop, related to CS 4.

Section 6 presents the main conclusions and recommendations.

2 Case Study 1: Survey length sampling

2.1 Introduction

One way to predict the status of a fish stock is to conduct an at-sea survey directed to the stock of interest. IPMA's at-sea surveys collect data to estimate the abundance (and relative abundance) of fish stocks and the relative frequency of population characteristics such as length, age, etc. It is of great importance to have precise and unbiased estimates; at the focus of this section is the study of precision, in particular of length-frequency estimates.

The aim of the exercise conducted for Case Study 1 is to optimize the number of individuals that need to be measured at surveys in order to determine the structure of the population regarding its length composition. Pennington *et al*. (2002) assesses the precision of length-frequency distributions estimated from trawl-survey samples, and shows that the effective sample size estimated using Kish's design effect (Cochran, 1977) can be much smaller than the currently applied sample size. Such an optimal sample size can be derived from data sets that satisfy certain criteria. The main assumption of Pennington *et al*. (2002) is that fish caught together (i.e. at a given survey station) tend to have more similar characteristics (e.g. length), than those caught randomly from the entire population. This would imply that fish caught together will contain less information about the population length distribution than fish sampled randomly from the general population. In this case, the effective sample size for the estimate of the population length-frequency distribution could be much smaller than the number of fish sampled with the current design. In other words, the sample mean estimated from randomly measured fish from the population should have the same precision as the population mean estimated from fish measured with the currently applied method.

The objective of the present study is to estimate the effective sample size and test for the precision of mean length estimates, based on the methodology of Pennington *et al*. (2002), using data of two species, the European hake - *Merluccius merluccius* (HKE) and the Atlantic horse mackerel – *Trachurus trachurus* (HOM), from the Portuguese Bottom Trawl Surveys.

2.2 Methods

In order to estimate the optimal sample size defined by Pennington *et al*. (2002), first, one needs to estimate the variance of the population length distribution and the variance of population mean length. If m_i fish are randomly sampled (or if all fish are sampled) at each station $(i = 1..n)$, and $M = \sum_{i=1}^{n} M_i$ is the total number of fish caught during the survey (where M_i is the actual or estimated number of fish caught at station i) and x_{ij} the length of the j th fish at station i, the variance of the population length distribution $\hat{\sigma}_x^2$ can be estimated using the following expression:

$$\hat{\sigma}_x^2 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m_i}\left({M_i}/{m_i}\right)\left(x_{i,j} - \hat{R}\right)^2}{M - 1} \quad (2.1)$$

In the above formula, $\hat{R}$ is the ratio estimator of the population mean length of Cochran (1977), given as

$$\hat{R} = \frac{\sum_{i=1}^{n} M_i \widehat{\mu_i}}{M} , \quad (2.2)$$

where $\widehat{\mu_i}$ is the estimate of the average length of fish caught at station i .

The variance of the population mean length can then be calculated as follows:

$$var(\hat{R}) = \sum_{i=1}^{n} \frac{\left(M_i/\overline{M}\right)^2 (\widehat{\mu_i} - \hat{R})^2}{n(n-1)} \quad , \tag{2.3}$$

where $\overline{M} = \sum_{i=1}^{n} M_i/_n$. Pennington *et al*. (2002) derives the estimate of the effective sample size by substituting (2.1) and (2.3) in

$$\frac{\widehat{\sigma}_x^2}{\widehat{m}_{eff}} = var\left(\hat{R}\right) . \tag{2.4}$$

As Pennington *et al*. (2002) relates the effective sample size to Kish's design effect $deff$ as follows:

$$deff = \frac{var(\hat{R})}{\widehat{\sigma}_x^2/_m} \quad , \tag{2.5}$$

where m is the number of fish sampled at random from the population, it can be written in the following form:

$$m_{eff} = \frac{m}{deff} . \tag{2.6}$$

Thus,

$$m_{eff} = \frac{m}{deff} = \frac{m}{\frac{var\left(\hat{R}\right)}{\sigma_x^2/_m}} = \frac{m * \frac{\sigma_x^2}{m}}{var\left(\hat{R}\right)} = \frac{\sigma_x^2}{var\left(\hat{R}\right)} \quad ,$$

and so

$$\widehat{m}_{eff} = \frac{\widehat{\sigma}_x^2}{var\left(\hat{R}\right)} . \tag{2.7}$$

Then, by selecting random samples from the total number of fish caught, and comparing the mean length estimated from these samples with the population mean length estimated by $\hat{\mathbf{R}}$ using (2.2), one can assess the precision of mean length estimates.

The case study of the present section considers Portuguese Autumn Groundfish Surveys (PT-PGFS Q4) carried out by IPMA in the past two years (2015 and 2016), and uses the data of HKE and HOM, to show whether Pennington's method to determine the optimal number of individuals to be sampled at surveys can be adopted.

The survey area is the Portuguese continental coast, covering the area extending from latitude 41°20' N to 36°30' N (ICES Area IXa), where 12 sectors are defined along the three main geographical zones as follows:

i. North (N): Caminha (CAM), Matosinhos (MAT), Aveiro (AVE), Figueira da Foz (FIG), Berlengas (BER);
ii. Southwest (SW): Lisbon (LIS), Sines (SIN), Vila Nova de Milfontes (MIL), Arrifana (ARR);
iii. South (S): Sagres (SAG), Portimão (POR), Vila Real de Santo António (VSA).

All of the above 12 sectors are further stratified by depth (1: ≤ 100m; 2: >100 and ≤ 200m; 3: >200 and ≤ 500m), defining the strata. The data used for the analyses consist, for each sampled station, of the number of fish caught, their average length, the number of fish measured, and the estimated or actual length of every fish caught. For each station, information on zone, sector and stratum were also made available.

Applying (2.7) to the four cases (HKE – 2015, HKE – 2016, HOM – 2015 and HOM – 2016), various scenarios emerged during the analyses. Namely, whether to consider the whole survey region, without taking into account any strata, or consider the stratification of the survey region by the 12 sectors, or consider the stratification of the survey region by zones (N, SW and S) and by depth (3 levels). While analyzing the above scenarios, several questions and doubts popped up, which are discussed below (see Section 2.4).

In order to show that the estimates obtained from fish sampled at random have the same precision of the sample mean obtained from the estimate based on the existing survey samples, 100 simulated distributions of $\hat{\bar{R}}$ for the different scenarios have been analyzed. The simulated estimates of the mean lengths were generated by randomly sampling (without replacement), from the total number of fish caught, a number of fish determined by the effective sample size formula (2.7).

2.3 Results

2.3.1 Exercise 1 – Considering the whole survey region, without taking into account any strata

2.3.1.1 Estimating the effective sample size

Table 2.1 presents the estimates of the effective sample size (denoted by ess_1) and summary statistics for assessing the precision of the estimated length distributions of HKE and HOM by year. When considering the whole survey region, the results indicate that for both HKE and HOM, the estimated effective sample size is very small compared to the number of sampled fish. This means that measuring only 1.5% of the samples in average would be sufficient in order to have the same precision of mean length estimates. The effective sample size per station was also calculated, however, its usefulness in this context, and whether it can be determined in such way is doubtful and discussed below (see Section 2.4).

Table 2.1. Data available and results for exercise 1 (considering the whole survey region, without taking into account any strata). See section 2.2 for further definitions.

Year	Species	Nº stations	Nº fish (total)	Nº sampled fish	$\hat{R}$ (cm)	var ($\hat{R}$)	$\hat{\sigma}_x^2$	ess_1	ess_1/Nº stations	(ess_1/sampled fish)*100%
2015	HKE	83	21950	10733	19.9	0.1	36.1	263	3	2.5%
2016	HKE	81	7290	5078	22.3	1.4	71.6	51	1	1.0%
2015	HOM	66	72808	6356	14.9	0.7	30.9	47	1	0.7%
2016	HOM	55	11248	2580	19.0	0.2	10.4	44	1	1.7%

2.3.1.2 Re-sampling

For this first exercise, the $\hat{\bar{R}}$ (values in Table 2.1) are identical to the mean of the re-sampling simulations (Figure 2.1).

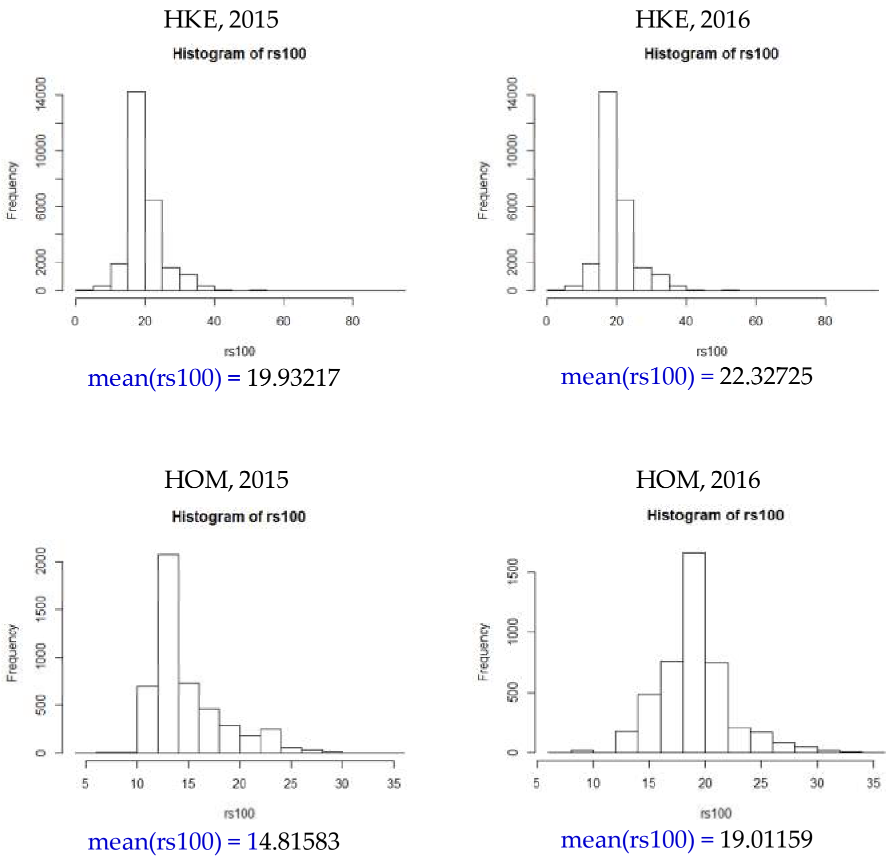


Figure 2.1. Results from the resampling simulations and respective mean for hake (top) and horse mackerel (bottom), for the years 2015 (left) and 2016 (right).

2.3.2 Exercise 2 – Considering the survey region stratified by sectors

2.3.2.1 Rationale

The mean length varies with sector in both species (Figures 2.2 and 2.3), being most evident for horse mackerel (Figure 2.3). For this reason, it was hypothesized that length distributions could be more precisely defined if an effective sample size was estimated by sector.

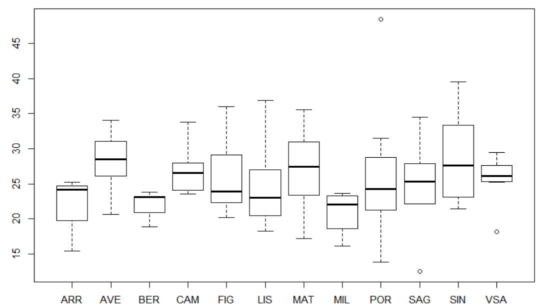


Figure 2.2. Interquartile range of hake total length (cm) by sector in 2016.

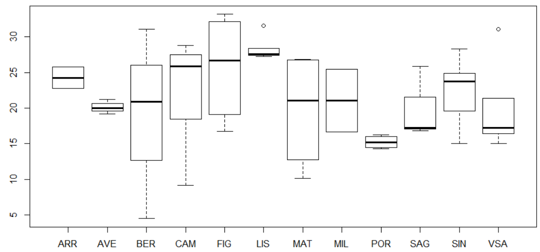


Figure 2.3. Interquartile range of horse mackerel total length (cm) by sector in 2016.

2.3.2.2 Estimating the effective sampling size

Table 2.2 contains the estimates of the effective sample size and summary statistics for assessing the precision of the estimated length distributions of HKE and HOM by sector, for each analyzed year. The results indicate that, similarly to exercise 1, the estimated effective sample size by sector is very small compared to the number of sampled fish, for both species. In the case of horse mackerel, very low effective sample sizes were estimated for some sectors despite the wide length range observed.

Table 2.2. Data available and results for exercise 2 (considering the survey region stratified by sectors). See section 2.2 for further definitions.

Year	Species	Sector	Nº stations	Nº fish (total)	Sampled fish	$\hat{R}$ (cm)	var ($\hat{R}$)	$\hat{\sigma}_x^2$	ess_1	ess_1/Nº stations	(ess_1/sampled fish)*100
2015	HKE	CAM	8	3735	1412	20.7	0.3	30.1	113	14	8.0%
2015	HKE	MAT	7	5197	1094	19.0	0.5	30.5	65	9	5.9%
2015	HKE	AVE	7	1675	918	19.9	0. 8	14.1	18	3	2.0%
2015	HKE	FIG	10	3308	2081	20.0	0.5	24.7	51	5	2.4%
2015	HKE	BER	6	1619	890	21.1	2.2	26.6	12	2	1.4%
2015	HKE	LIS	7	1220	862	17.5	0.5	29.4	54	8	6.3%
2015	HKE	SIN	8	927	811	19.2	4.2	100.6	24	3	3.0%
2015	HKE	MIL	6	594	594	18.8	3.6	66.3	18	3	3.1%
2015	HKE	ARR	5	338	338	18.2	3.1	54.2	18	4	5.2%
2015	HKE	SAG	5	455	354	24.5	0.7	61.1	86	17	24.3%
2015	HKE	POR	7	1544	494	21.4	0.4	21.2	51	7	10.4%
2015	HKE	VSA	7	1339	855	20.1	8.9	78.7	9	1	1.0%
2016	HKE	CAM	8	799	676	26.1	0.6	41.5	73	9	10.8%
2016	HKE	MAT	8	128	126	25.8	8.0	65.6	8	1	6.5%
2016	HKE	AVE	5	109	109	24.7	5.5	36.4	7	1	6.1%
2016	HKE	FIG	9	506	506	22.8	0.9	40.5	46	5	9.1%
2016	HKE	BER	5	668	436	21.8	3.2	71.1	22	5	5.1%
2016	HKE	LIS	9	986	769	21.5	1.3	51.6	40	5	5.3%
2016	HKE	SIN	10	506	455	26.0	2.5	71.3	29	3	6.3%
2016	HKE	MIL	4	368	368	19.8	5.5	69.9	13	3	3.5%
2016	HKE	ARR	3	214	214	18.4	8.9	47.7	5	2	2.4%
2016	HKE	SAG	5	167	123	24.3	2.1	78.5	38	8	31.0%
2016	HKE	POR	8	1499	544	16.8	4.6	83.2	18	2	3.4%
2016	HKE	VSA	7	1340	752	26.1	0.4	69.8	188	27	25.0%
2015	HOM	CAM	5	32166	787	12.6	<0.05	61.8	1839	368	233.7%
2015	HOM	MAT	6	2822	553	15.6	0.5	5.2	11	2	2.0%

Year	Species	Sector	Nº stations	Nº fish (total)	Sampled fish	$\hat{R}$ (cm)	var ($\hat{R}$)	$\hat{\sigma}_x^2$	ess_1	ess_1/Nº stations	(ess_1/sampled fish)*100
2015	HOM	AVE	8	1947	835	15.0	2.0	9.7	5	1	0.6%
2015	HOM	FIG	8	3966	403	17.8	<0.05	3.6	144	18	35.8%
2015	HOM	BER	2	30	30	20.6	31.7	61.9	2	1	6.7%
2015	HOM	LIS	6	469	466	14.5	3.4	39.2	11	2	2.4%
2015	HOM	SIN	6	10653	1003	17.0	3.0	16.8	6	1	0.6%
2015	HOM	MIL	4	161	161	21.8	5.7	28.8	5	1	3.2%
2015	HOM	ARR	5	2105	410	15.6	1.5	8.7	6	1	1.4%
2015	HOM	SAG	5	4240	524	17.7	2.4	20.4	9	2	1.6%
2015	HOM	POR	6	12689	667	16.6	0.3	11.6	37	6	5.6%
2015	HOM	VSA	5	1560	517	15.1	1.1	7.3	7	1	1.3%
2016	HOM	CAM	6	105	105	17.8	7.9	67.8	9	1	8.1%
2016	HOM	MAT	4	88	60	22.0	19.5	53.6	3	1	4.7%
2016	HOM	AVE	3	747	233	20.0	<0.05	2.3	166	55	71.2%
2016	HOM	FIG	8	5019	549	18.4	0.1	6.6	128	16	23.2%
2016	HOM	BER	3	24	24	22.3	7.6	36.9	5	2	20.4%
2016	HOM	LIS	5	274	203	27.7	0.1	4.2	86	17	42.2%
2016	HOM	SIN	9	2642	421	20.5	0.7	5.5	9	1	2.0%
2016	HOM	MIL	2	45	45	22.5	15.3	24.9	2	1	3.6%
2016	HOM	ARR	2	11	11	23.9	1.9	3.3	2	1	15.5%
2016	HOM	SAG	3	417	151	17.4	0.5	9.6	18	6	12.1%
2016	HOM	POR	4	355	355	15.37	0.2	3.2	22	6	6.3%
2016	HOM	VSA	6	1521	423	17.7	0.1	5.5	43	7	10.1%

2.3.2.3 Re-sampling

The re-sampling simulations showed that the means had values close to $\hat{R}$. Figure 2.4 illustrates the results of resampling for hake and horse-mackerel in sector CAM.

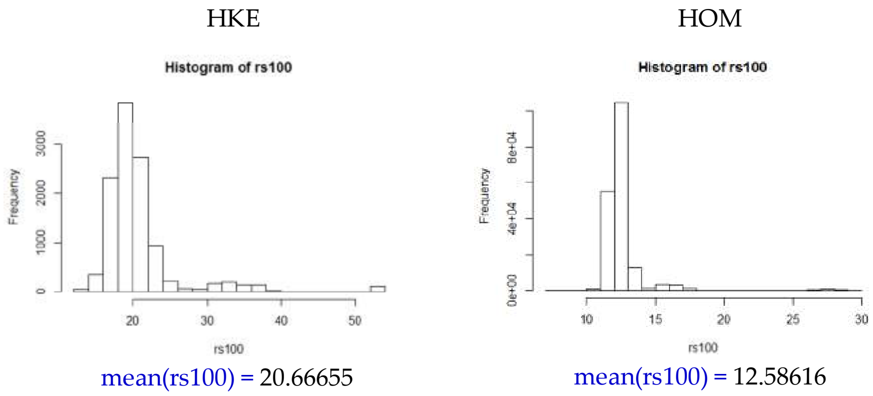


Figure 2.4. Results from the re-sampling simulations and respective mean for hake and horse mackerel in sector CAM in 2015.

2.3.3 Exercise 3 – Considering the survey region stratified by zones (N, SW and S) and by depth (3 levels)

2.3.3.1 Rationale

Mean lengths vary with sector, but also with depth (Figures 2.5 and 2.6). For these reasons, it was also tested whether length distributions could be more precisely defined when an effective sample size is estimated by area and stratum.

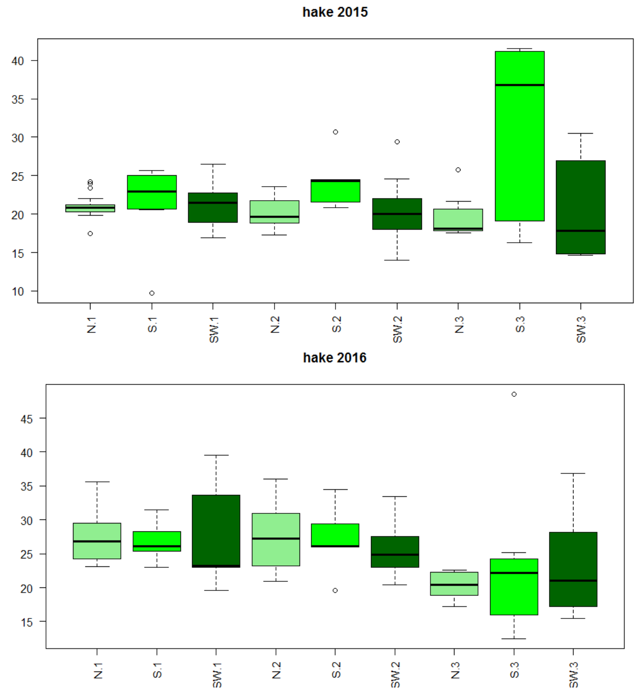


Figure 2.5. Interquartile range of hake total length (cm) by area and stratum in 2015 and 2016.

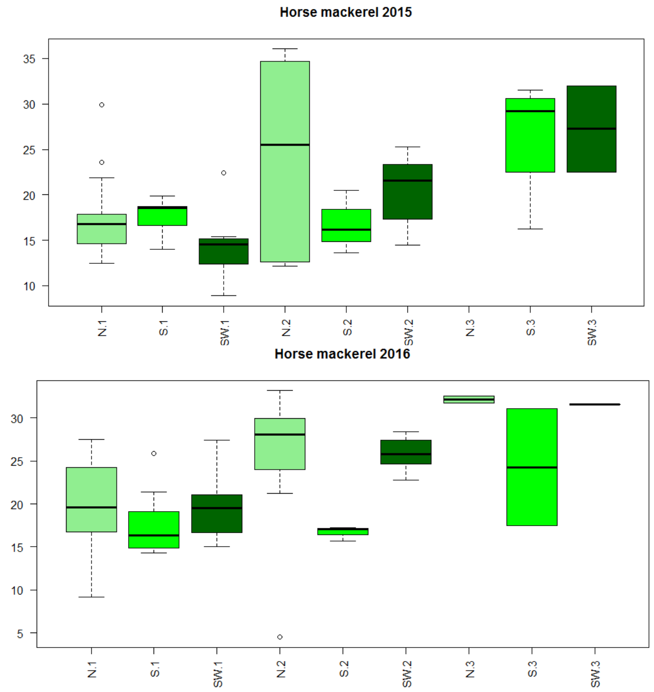


Figure 2.6. Interquartile range of horse mackerel total length (cm) by area and stratum in 2015 and 2016.

2.3.3.2 Estimating the effective sample size

When stratifying the survey region, the optimal sample sizes for the estimates of the length composition of HKE followed the same pattern, with proportions from the samples that would be sufficient to measure varying between 0.8% and 26.2%. In the case of HOM, the results show a different pattern, with more extreme effective sample size estimates in certain strata, reaching 733.3% of the sampled fish that should be measured. It might be due to negative intra-haul correlation, what can push the effective sample size above the number of fish sampled (Cochran, 1977). Although, this is rare for trawl surveys (Pennington *et al.*, 2002).

Table 2.3. Data available and results for exercise 3 (Considering the survey region stratified by zones (N, SW and S) and by depth (3 levels)). See section 2.2 for further definitions.

Year	Species	Zone	Depth	Nº stations	Nº fish (total)	Sampled fish	$\hat{R}$ (cm)	var ($\hat{R}$)	$\hat{\sigma}_x^2$	ess_1	ess_1/Nº stations	(ess_1/sampled fish)*100%
2015	HKE	all	all	83	21950	10733	19.9	0.1	36.1	263	3	2.4%
2015	HKE	N	all	38	15533	6395	20.0	0.2	27.2	161	4	2.5%
2015	HKE	N	1	17	5848	2474	20.5	0.03	11.0	316	19	12.8%
2015	HKE	N	2	13	7801	2691	19.8	0.6	41.3	68	5	2.5%
2015	HKE	N	3	8	1884	1230	18.9	0.4	16.7	39	5	3.2%
2015	HKE	SW	all	26	3079	2605	21.3	1.6	71.2	46	2	1.8%
2015	HKE	SW	1	8	2838	1277	20.3	1.7	42.9	26	3	2.0%
2015	HKE	SW	2	5	434	390	26.4	5.2	41.0	8	2	2.0%
2015	HKE	SW	3	6	66	66	31.1	26.1	186.5	7	1	10.9%
2015	HKE	S	all	19	3338	1733	20.0	0.2	59.2	351	19	20.3%
2015	HKE	S	1	8	382	382	18.7	1.0	20.7	21	3	5.6%
2015	HKE	S	2	12	2085	1611	18. 6	1.0	66.5	70	6	4.3%
2016	HKE	all	all	81	7290	5078	22.3	1.4	71.6	51	1	1.0%
2016	HKE	N	all	35	2211	1853	24.0	0.7	54.9	79	2	4.3%
2016	HKE	N	1	16	857	734	26.1	0.5	33.7	67	4	9.2%
2016	HKE	N	2	13	834	599	24.6	0.5	79.6	157	12	26.2%
2016	HKE	N	3	6	520	520	19.6	0.3	23.5	82	14	15.8%
2016	HKE	SW	all	26	2074	1806	21.4	6.4	92.6	15	1	0.8%
2016	HKE	SW	1	5	591	374	20. 9	1.5	51.4	35	7	9.3%
2016	HKE	SW	2	13	935	884	24.1	2.0	57.0	29	22	3.3%
2016	HKE	SW	3	8	548	548	19.5	4.5	79.8	18	2	3.2%
2016	HKE	S	all	20	3006	1419	24.0	0.7	82.5	119	6	8.4%
2016	HKE	S	1	8	785	566	26.1	0.7	45.9	64	8	11.3%
2016	HKE	S	2	5	872	302	23.7	5.5	86.5	16	3	5.2%
2015	HOM	N	all	29	40931	2608	13.4	0.4	47.5	123	4	4.7%
2015	HOM	N	1	20	10376	2240	16.2	0.4	6.9	19	1	0.8%
2015	HOM	N	2	9	30556	368	12.5	<0.05	65.6	2699	300	733.3%
2015	HOM	N	3	0								
2015	HOM	SW	all	21	13388	2040	16.8	1.6	17.2	11	1	0.5%
2015	HOM	SW	1	8	8661	796	14.7	<0.05	4.2	122	15	15.4%
2015	HOM	SW	2	4	4718	1240	20.5	0.8	19.5	26	6	2.1%

Year	Species	Zone	Depth	Nº stations	Nº fish (total)	Sampled fish	$\hat{R}$ (cm)	var ($\hat{R}$)	$\hat{\sigma}_x^2$	ess_1	ess_1/Nº stations	(ess_1/sampled fish)*100%
2015	HOM	SW	3	4	7	4	25.1	14.6	62.2	4	1	107.5%
2015	HOM	S	all	16	18489	1708	16.7	0.3	13.8	41	3	2.4%
2015	HOM	S	1	8	7004	1174	18.1	0.6	14.4	23	3	2.0%
2015	HOM	S	2	11	11443	492	15.8	0.1	11.0	83	8	16.8%
2015	HOM	S	3	2	42	42	27	6.3	28.7	5	2	11.0%
2016	HOM	all	all	55	11248	2580	19.0	0.2	10.4	44	1	1.7%
2016	HOM	N	all	24	5983	971	18.6	0.2	8.4	50	2	5.1%
2016	HOM	N	1	14	5863	879	18.5	0.1	6.6	72	5	8.1%
2016	HOM	N	2	8	115	87	27.2	1.0	15.8	17	2	19.1%
2016	HOM	N	3	2	5	5	31.9	0.1	0.8	14	7	278.0%
2016	HOM	SW	all	18	2293	680	21.2	1.0	10.0	10	1	1.5%
2016	HOM	SW	1	6	2391	398	20.2	0.5	5.8	11	2	2.9%
2016	HOM	SW	2	11	570	271	25.3	0.6	4.7	7	1	2.7%
2016	HOM	SW	3	1	11	11	31.6		2.5		0	0.0%
2016	HOM	S	all	13	2973	929	17.3	0.1	6.6	71	6	7.7%
2016	HOM	S	1	8	711	487	17.1	0.6	12.3	22	3	4.4%
2016	HOM	S	2	3	422	278	16.6	0.2	2.6	11	4	3.8

2.3.3.3 Resampling

A number of fish determined by the effective sample size per station was randomly selected from each station from the total number of fish caught. Since the number of fish measured per station was sometimes less than the number one would have to select in the re-sampling process, the random selection was done with replacement in such cases, and then the random samples were merged. In these cases, the resulting simulations showed a slight difference (0.9 to 3.8 cm) in the mean length obtained by re-sampling from all stations versus the mean length obtained by re-sampling per station (see Figure 2.7).

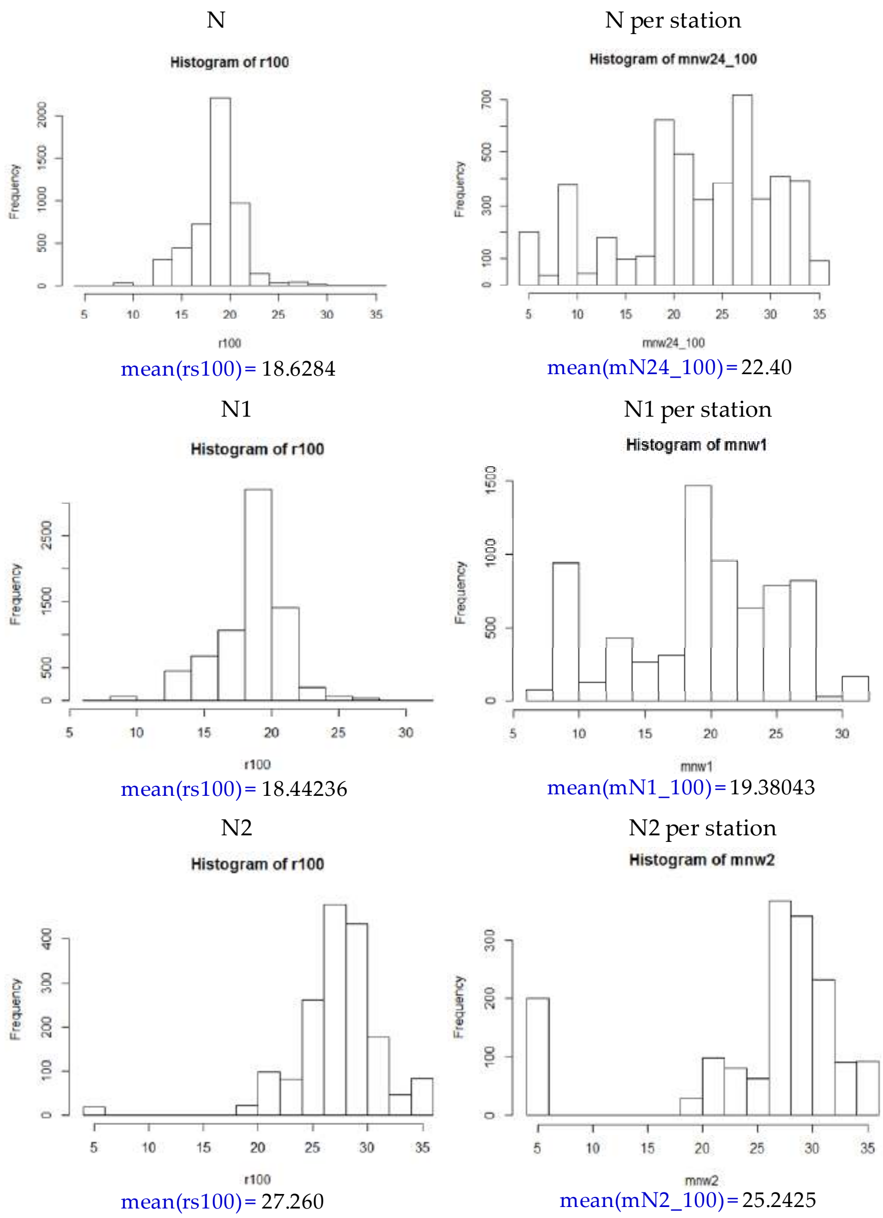


Figure 2.7. Results from the resampling simulations and respective mean for horse mackerel in 2016 by area: North, North per station, North1, North1 per station, North2 and North2 per station.

2.4 Discussion and future work

Three main types of exercises have been conducted regarding survey fish length sampling. The first one focuses on estimating the optimal sample size for one species at a time per survey. Since the number of stations at the surveys might vary from year to year, as well as the total number of fish caught and the number of fish sampled, it means that the optimal sample size may also vary from year to year. Thus, it might not be enough to estimate the effective sample size only from one year and apply it for future years, but it shall be estimated from a group of surveys of the same type. Therefore, it might be preferred to estimate the optimal sample size for instance for the last few years (at least 5-10), and calculate their average. Then, this average optimal sample size could be used as a reference at the forthcoming surveys.

The abundance of a particular species in a survey region can be variable. This motivated the second type of exercise, in which the optimal sample size was estimated for each of the 12 pre-defined sectors ("water cuboids") along the Portuguese coast. Such stratification might imply that the estimated optimal sample size can be much smaller or much greater in some sectors than the estimated optimal sample size without stratifying the data per sectors. Similarly, the third type of exercise, where the survey area is divided into three major zones (N, SW and S) with three depth strata in each, faces the same implications of stratifying the survey region.

The question is: which method is the best to determine the optimal sample size for a particular group (cluster) of fish, i.e. per station. While looking for the answer, one has to keep in mind that, as Pennington *et al*. (2002) reminds, besides variable fish density, very small effective sample sizes, and so rather imprecise estimates of length distributions might be due to fish being more similar in a haul, i.e., at a station, than fish in the general population. If fish of similar length tend to be caught together, with the increased variance, the effective sample size can drastically decrease (because of intra-haul correlation, see Cochran, 1977).

A very low effective sample size per station implies that in order to significantly improve survey precision, fish should be sampled from as many locations as possible. This could improve overall survey efficiency without increasing survey cost. If more locations were sampled (with shorter hauls), the total number of fish caught would be less in average, but the estimates of fish density (abundance) would be more precise, and the resulting fish samples would be more representative of the whole population. However, changing survey design accordingly could lead to the loss of information for less abundant species.

For some species it might be preferred to stratify the catch at a station, e.g. for small, medium and large size categories. In this case, a random sample would be chosen from each stratum, and the stratified mean length ($\widehat{\mu_i}$) would be estimated. For other sampling schemes at a station, the variance of the population length distribution is calculated differently from (1), based on the frequency of fish in each length bin (f_k) and the bins' midpoints (y'_k):

$$\widehat{\sigma}_x^2 = \frac{\sum_{k=1}^{L} f_k \left(y'_k - \widehat{R}\right)^2}{M-1} \quad . \tag{2.8}$$

Further investigation should be performed to explain the extreme effective sample size estimates observed for HOM in certain strata (exercise 3) and to evaluate whether length distributions could be more precisely defined. Future work is also needed to check whether this methodology can be applied to all species, considering fish distribution patterns. Finally, the influence of the change in the sampling effort (using effective sample size) on the raising procedure and on the estimation of the total length composition should be evaluated.

2.5 References

Cochran WG, 1977. Sampling techniques. 3rd ed. John Wiley and Sons, New York, NY, 428 p.

Pennington M, Burmeister LM, Hjellvik V, 2002. Assessing the precision of frequency distributions estimated from trawl-survey samples. *Fish. Bull.* 100: 74–80.

3 Case Study 2: Hake at-market length sampling

3.1 Introduction

The current at-market sampling design is a stratified multistage design, with [auction * day] as the Primary Sampling Unit (PSU). It is stratified by fleet (or métier), auction and quarter (Figure 3.1). Following the DCF requirements (EU, 2016a), less significant fleets are not sampled (e.g. dredges, beach-seines) and sampling effort is based on number of trips. Annual sampling effort is fixed by the DCF National Sampling Plan (EU, 2016b) which sets the number of trips to be sampled for each fleet (≈ métier). Sampling effort is allocated to auctions and quarters proportionally to the most recent year's landings.

For each fleet, the visit dates in each [auction * quarter] are spread somewhat systematically throughout the quarter in a way that covers all week days of that fleet activity.

In every [auction * visit_date], observers attempt to sample a predefined number of vessel sale events, which are haphazardly selected from a list of all landings awaiting auction. This list includes the name of each vessel and the commercial species, commercial category and weight of each of landed box. Each vessel sale event generally corresponds to the landings of one fishing trip. A minor proportion of vessel sale events may not be present in the selection list when sampling starts.

From each selected trip, the observers aim to sample boxes from every landed species and commercial size category. Within each size category, the observers select 1 box randomly. Since 2014, when there are very few fish from a species inside the box, observers take more boxes until the length composition of the size category is well defined. Also, when different species are present within a box, observers sample them all. This sampling design, referred to as "trip-based" design, is linked to the concurrent sampling.

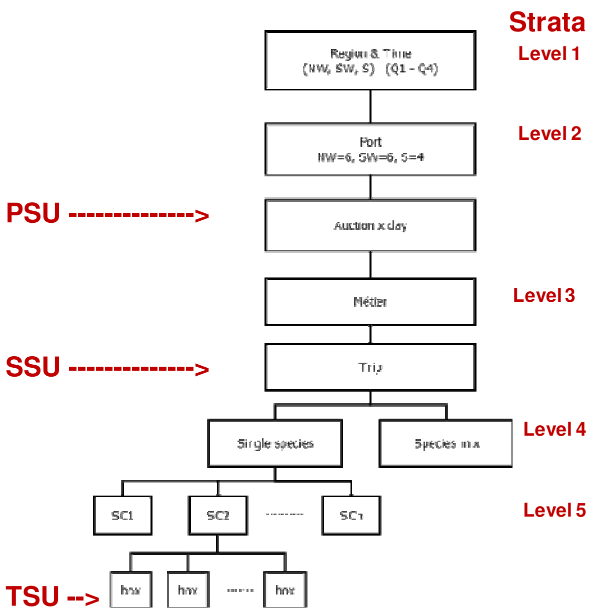


Figure 3.1. Current at-market sampling design (Azevedo *et al.*, 2016)

Some difficulties may constrain this sampling strategy, namely vessels arriving to port after the auction has started, with large amounts of landings/species/categories meaning no time to sample the complete trip. (e.g.: OTB_DEF). Also, some commercial species may not be available for sampling if they have previously been included in a fixed sale contract. Sometimes, when observers do not have time to sample all commercial species, the more important species (for stock assessment, with TAC, etc) are selected.

The aim of this case-study is to analyze the effective sampling size to estimate the annual length composition of hake landings, based on the length distribution by size category obtained from sampling size categories. Modeling length distribution using commercial size categories was first presented in the Workshop on Sampling Design and Optimization of (Azevedo *et al.*, 2014), using horse-mackerel data and later its application in a horse mackerel focused pilot plan was further developed in Azevedo *et al.*, 2016 and ICES, 2017. In this case-study, the sampling effort and the effective number of length measurements by sample in a size category sampling design for hake are discussed. This sampling design, referred herein after as "size category-based" design, is a "species focus" sampling scheme.

The analysis was carried out based on the year 2013 sampling data. The main auction markets sampled were grouped in three zones: NW (Póvoa do Varzim, Matosinhos, Aveiro and Figueira da Foz); SW (Peniche, Sesimbra and Sines); and S (Portimão, Olhão and Vila Real de Santo António). In these ports, trips from several metiers were sampled, particularly bottom otter trawl trips targeting either demersal fish or crustaceans and trips from multi-gear vessels, some of them identified as using gill or trammel nets and longliners (P02).

The dataset comprises length data for hake, recorded in 365 sampled trips with hake landings (positive trips), corresponding to 720 samples and 5748 length measurements. Table 3.1 summarizes the total number of trips sampled, the number of samples and the number of individuals measured by size category and zone. In 2013, the maximum number of sampled hake individuals per size category box was around 30. There were, however, several sampled trips where the number of hake landed by size category was very low, especially in size categories that include the larger individuals (T1 and T2) (see Table 3.1).

The analyses were performed using the statistical environment R (R Core Team, 2017).

Table 3.1. Summary statistics of hake at-market sampling in 2013: total number of trips per zone, number of samples and number of individuals measured by zone and size category (SC).

zone	SC	# trips	# samples	# indiv
NW	T1		39	184
	T2		39	161
	T3		54	373
	T4		93	708
	T5		48	325
	Total	146	273	1751
SW	T1		25	155
	T2		32	257
	T3		46	314
	T4		57	441
	T5		22	195
	Total	106	182	1362
S	T1		16	76
	T2		28	187
	T3		78	814
	T4		82	861
	T5		61	697
	Total	113	265	2635
TOTAL		365	720	5748

3.2 Analysis of data

3.2.1 Size categories (SC)

In the auction, landings of hake and of other species are sold in size categories. In the case of hake there are 5 size categories from T1 (largest fish) to T5 (smallest fish), which have different market prices. Although the size category classification is established by EU regulation (Council Regulation (EC) No. 2406/96 of 26 November 1996), its application may differ by zone. Figure 3.2 presents the 2013 hake landings by size category and zone. As shown, in 2013 the size categories with highest landings were T4 and T5 (small sizes) and mainly from the NW and SW. In S zone, low landings were recorded for the largest fish, T1 and T2. Figure 3.3 shows the overall mean length (± 1 sd) in each size category and its variability among zones for 2013.

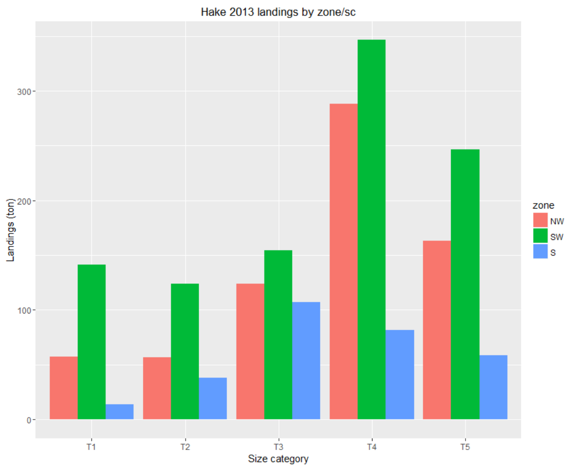


Figure 3.2. Portuguese hake landings by size category and zone.

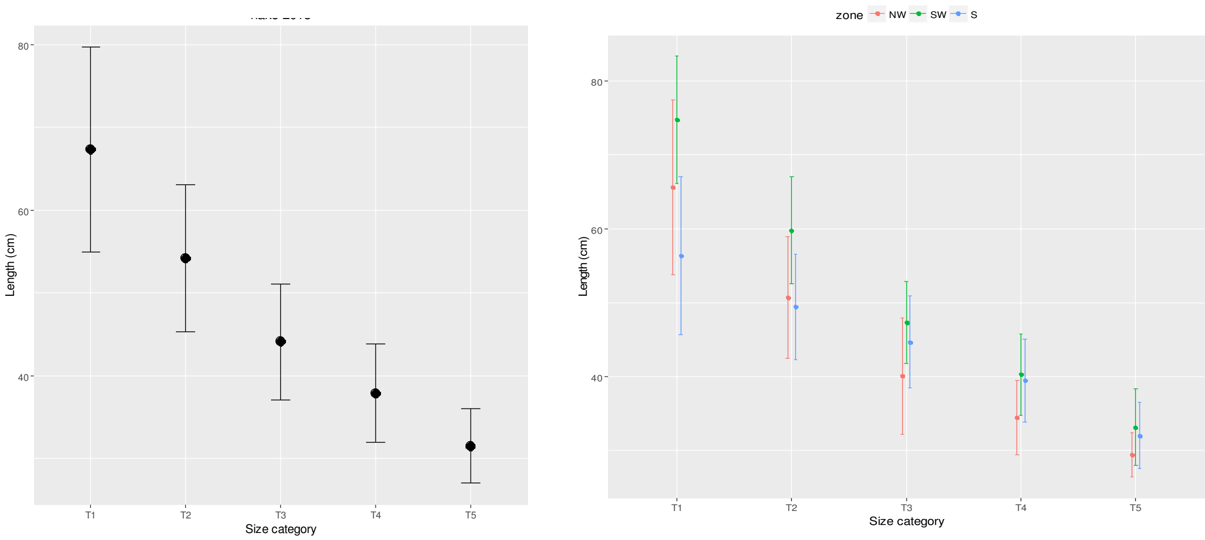


Figure 3.3. Mean length (± 1 standard deviation) by size category in 2013. Left – all samples combined; right – samples by zone.

To estimate the required sampling effort (number of positive trips) to characterize the size categories, two approaches, the "trip-based" and "size category-based" designs, were developed and described in the following sections. Given that the mean length by size category may vary among zones (Figure 3.3), the analyses were carried out considering this factor.

3.2.2 Trip-based design

The first set of analyses aimed at exploring the number of (positive) trips necessary to characterize the length composition by size category.

Considering that no major changes are made to the current trip sampling design, simulations were performed by zone, i.e. a number *k* of whole trips are randomly sampled in each zone and all the existing size categories and length measurements from each sampled trip were used for the analysis. It is noted that the trip landings may not include all size categories. The simulation scheme is presented in Figure 3.4.

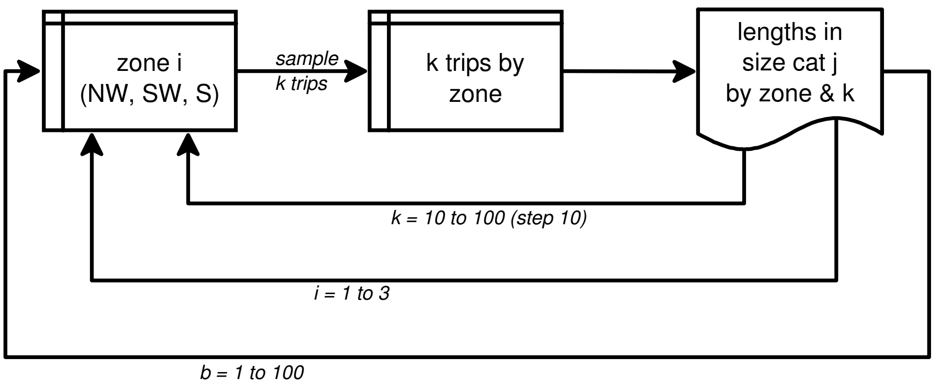


Figure 3.4. Simulation scheme for the trip-based design 1. *i* – levels of zone; *k* – number of trips to sample; *b* – number of re-samples)

The simulations were performed with and without trip replacement for the three zones, with *k* = 10:100 (step 10) and 100 re-samples. As shown in Table 3.1, the number of trips with hake landings, sampled in 2013 in each zone (NW = 146, SW = 106, S = 113), is greater than the maximum value of *k*.

The expected number of samples by size category (minimum and maximum) and zone for the simulated *k* trips (Table 3.2) is below or well below the number of sampled trips in 2013, since not all size categories are landed in each trip and some size categories are less frequent in some zones.

Table 3.2. Trip-based design. Minimum and maximum number of samples by size category and sampled trips (*k* from 10 to 100, step 10) in each zone, from simulations with 100 re-samples with replacement.

		Min					Max				
Zone	k	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5
NW	10	0	0	0	3	0	6	5	7	9	7
	20	0	1	2	6	0	9	10	12	17	12
	30	3	3	3	14	4	15	13	18	25	18
	40	3	5	9	18	6	18	18	24	33	21
	50	7	5	11	24	7	21	21	27	41	25
	60	8	10	14	27	13	24	26	29	47	31
	70	9	10	15	34	14	33	34	35	53	35
	80	12	14	15	43	17	31	32	42	61	38
	90	13	13	25	44	19	38	33	44	66	39
	100	17	12	26	51	21	43	38	48	79	45
SW	10	0	0	1	1	0	5	6	9	8	7
	20	0	1	4	6	0	10	11	15	16	9
	30	2	3	7	8	2	14	15	19	22	11
	40	4	6	10	13	2	16	20	29	28	16
	50	6	8	10	17	3	18	25	30	35	17
	60	6	5	15	21	5	22	29	35	43	24
	70	8	14	22	28	6	25	36	42	47	22
	80	10	16	24	32	8	32	33	44	53	30
	90	12	18	26	36	9	32	39	50	63	29
	100	11	18	28	43	12	34	45	54	66	29
S	10	0	0	2	4	2	4	6	10	10	10
	20	0	0	9	8	5	7	9	19	20	16
	30	0	2	14	15	11	9	14	28	26	21
	40	1	5	19	22	12	12	17	35	35	30
	50	2	5	28	29	20	13	20	45	45	35
	60	3	8	30	35	24	16	22	50	52	43
	70	3	9	40	38	24	20	26	58	60	47
	80	5	9	44	46	34	18	28	64	66	53
	90	5	13	52	56	37	21	32	73	75	60
	100	8	13	58	53	41	22	34	82	82	66

For each size category j in each zone i, the mean length, standard deviation (*sd*) and coefficient of variation (*CV*) were computed for each re-sample and *k*. As expected, the results are similar when sampling is carried out with or without replacement. Figure 3.5 shows the variability of the mean length and CV for the sampling with replacement.

For size categories T3, T4 and T5, which are the size categories with higher landings, the variability in the mean length does not change much for $k \geq 40$ in all zones. The size categories T1 and T2 which have lower landings in the NW and S, show decreasing variability in the mean length, from 10 to 60 trips.

For $k \geq 40$ the CVs were, on average, below 15% for all size categories in SW, for T2 to T5 in S and for T4 and T5 in NW. The highest CVs and variability was observed in T1 in S for k between 10 and 70 trips, which reflects the low number of samples of T1 in the dataset (Table 3.1).

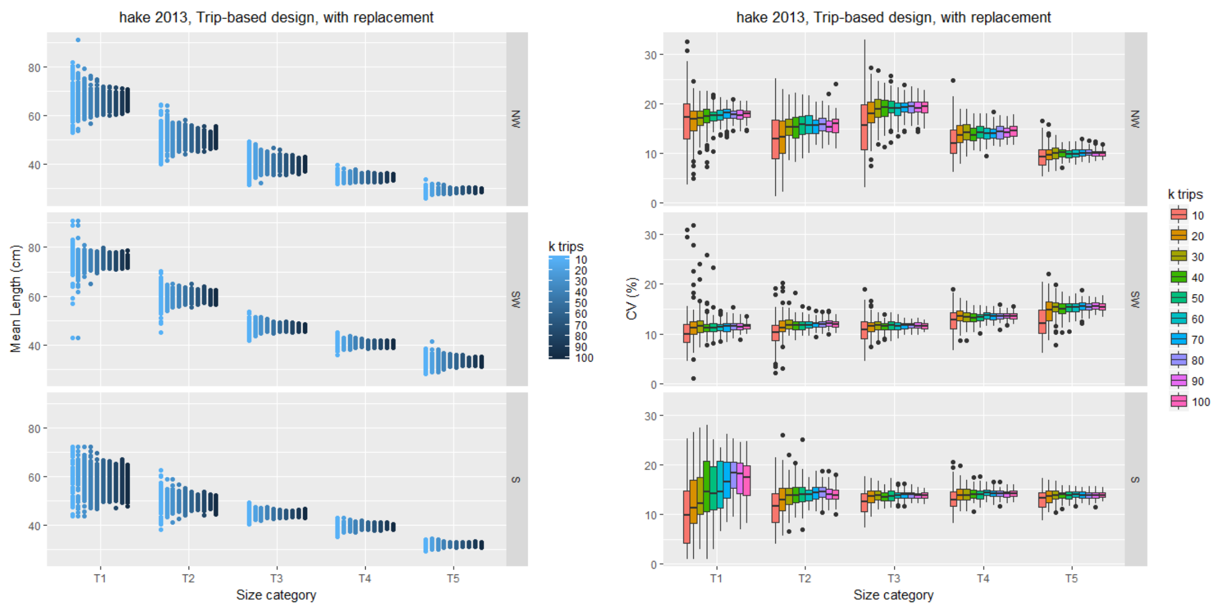


Figure 3.5. Trip-based design. Mean length (left) and length CV (right) by size category and number of sampled trips (*k* from 10 to 100, step 10) by zone, from simulations with 100 re-samples with replacement.

The second set of simulations aims to estimate the effective sample size (number of fish to be measured) by size category, to characterize its length distribution.

Considering the same trip-based design, how many fish shall we take from a size category in each sampling event to characterize its length distribution? In this case, sampling is performed in two steps: 1) the trips are sampled by zone as in the previous scheme, and 2) from each category of each sampled trip, *n* lengths are sampled. Considering the number of individuals that were measured from each trip/size category combination in the 2013 dataset (≈ 1 to 30), this second step was carried out with replacement. Again, as in the previous simulation, if one category is missing in the trip, this category will not be sampled. The simulation scheme is presented in Figure 3.6.

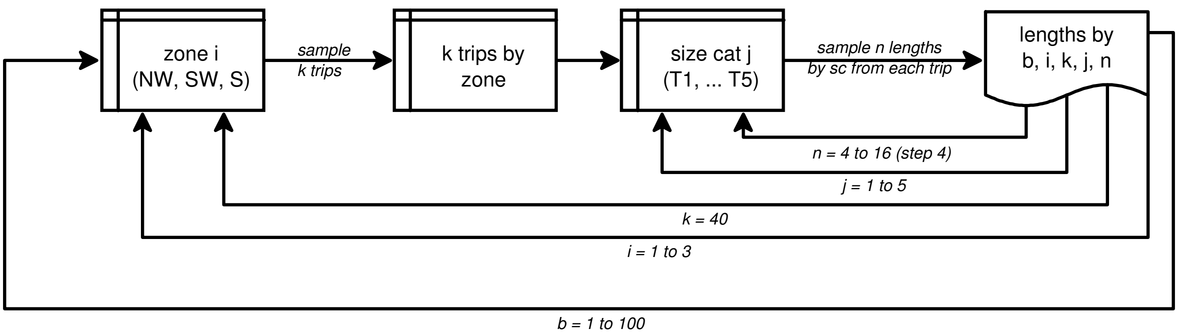


Figure 3.6. Simulation scheme for the trip-based design 2. *i* – levels of zone; *k* – number of trips to sample; *j* – size category levels; *n* – number of lengths to sample; *b* – number of re-samples.

Simulations were performed by zone for *k* = 40 trips (representing a reduction around 70% of the 2013 sampling effort, Table 3.1) and *n* from 4 to 16 fish (step 4). Since not all trips landed all size categories, the number of samples by size category for 40 trips varied according to Table 3.2, though the simulated number of fish measured in each size category sampled was achieved since the re-sample of n was performed with replacement.

The results indicate that the variability of the estimated mean length and its CV by size-category are very similar for all n analyzed, in each zone (Figure 3.7). However, as noted above, the achieved total sample size by size category and zone was well below the intended (*i.e.* 160 for $n = 4$, 320 for $n = 8$, 480 for $n = 12$ and 640 for $n = 16$).

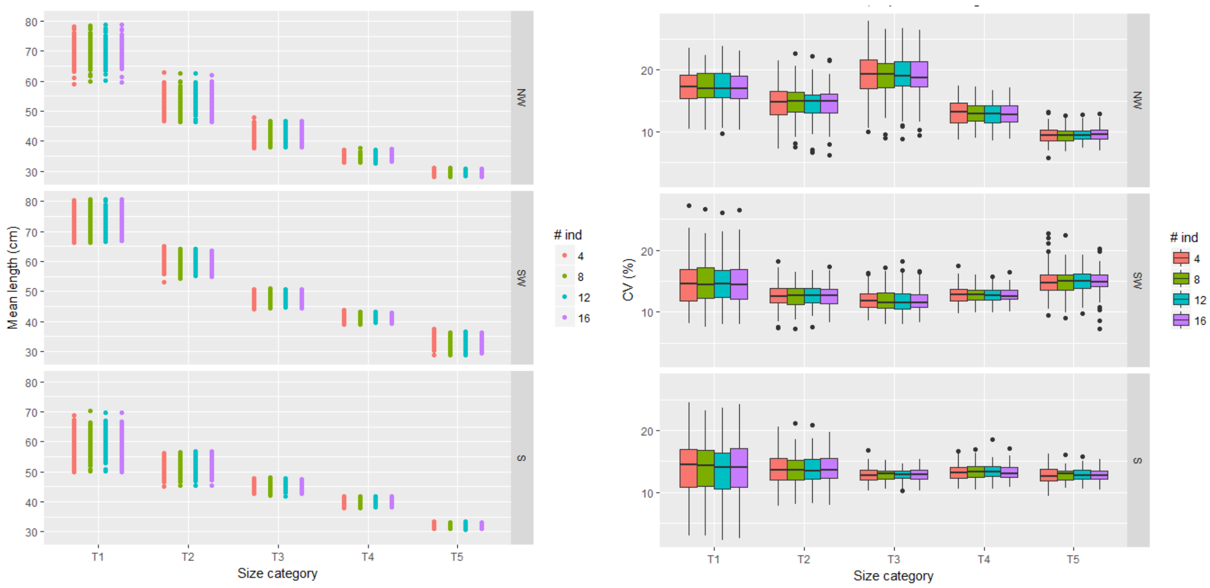


Figure 3.7. Trip-based design. Mean length (left) and its CV (right) with number of sampled individuals n (# ind from 4 to 16, step 4) by size category and zone. Simulations for k = 40 trips by zone, with 100 re-samples with replacement.

3.2.3 Size category-based design

In this approach, the trips are ignored and for each zone and size category, s samples of size n are taken.

For the analysis the samples (fish lengths) may be obtained using two approaches:

i) Randomly generate the samples assuming a normal distribution characterized by the mean length and standard deviation by size category and zone;

ii) Randomly sample from a pool of all lengths measured in 2013 in each size category and zone.

This sampling scheme, with the two variants, is presented in Figure 3.8.

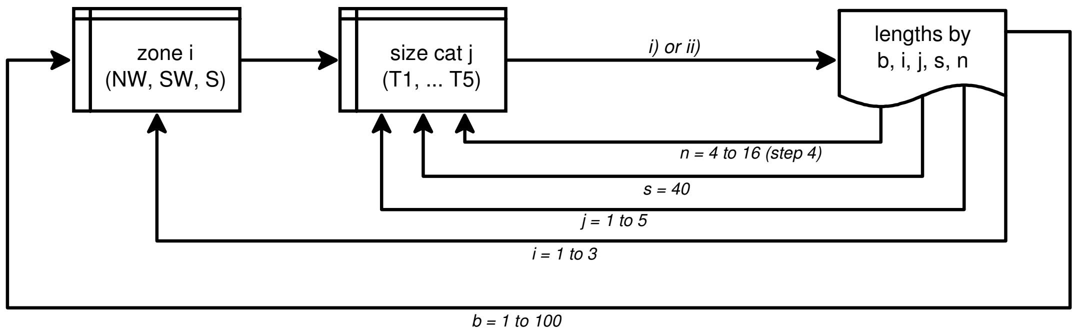


Figure 3.8. **Simulation scheme for the size category-based design. i) and ii) represent two different sampling procedures. *i* – zone levels; *j* – size category levels; *s* – number of samples: *n* – number of lengths to sample; *b* – number of re-samples.**

As expected, both variants give similar results:

- Stable mean length within each [zone * size category] combination, no matter the size of *n*.
- No trend observed in CV with the increase of *n* in each [zone * size category] combination.

Figure 3.9 shows the results for the variant ii). Sampling only 4 fish by size category in 40 samples corresponds to 160 fish to be measured by zone and a total annual of 2400 fish (160 fish x 3 zones x 5 sc). The results suggest a reduction of around 60% in the number of fish measured in 2013 (Table 3.1).

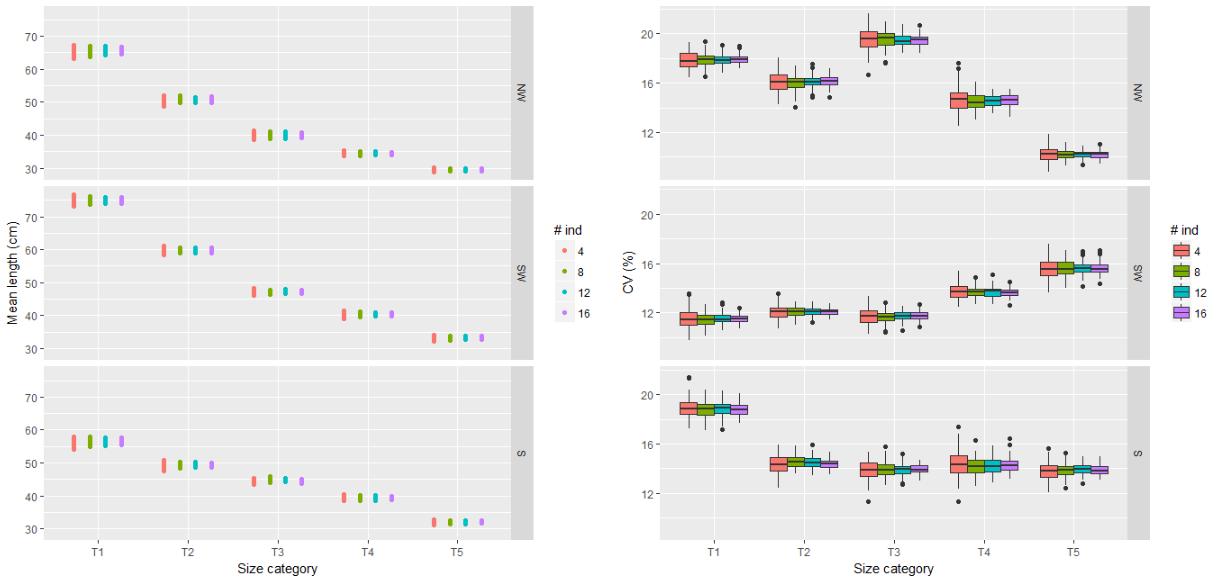


Figure 3.9. **Size category-based design. Mean length (left) and its CV (right) with number of sampled individuals *n* (# ind from 4 to 16, step 4) by size category. Simulations for 40 samples by zone, with 100 re-samples with replacement.**

3.3 Discussion and future work

Considering the current trip-based sampling design, in the case of hake landings, the simulations using 2013 data suggest that the number of trips to be sampled, for the characterization of size categories length distributions, may be reduced to around 40 in each zone. Table 3.2 illustrates the minimum and maximum number of samples by size category and *k*, in each zone, for the simulations performed with 100 re-samples with replacement. For example, with *k* = 40 trips, there is a chance of obtaining only 1 sample of size category T1 in S, which is clearly not enough to estimate the mean length for that category with acceptable precision. Increasing *k* to 50 trips, the expected minimum number of samples doubled, though it is still low (2 samples). The simulations for the effective sample size of length measurements (n) by size category were carried out considering the trip-based design (k = 40 trips) and the alternative size category-based design (s = 40 samples) from n = 4 to 16 fish. The size category-based design is not limited by the sampled trips, i.e. the current sampling level "*métier*" and sampling unit "trip" (Figure 3.1) are removed. Although the size categories classification may show some variability among zones/ports, in the same auction market no differences among gears are expected.

The comparison of mean lengths and *CV* obtained from the two designs (see Figures 3.7 and 3.9) shows that the two designs have similar mean lengths per size category but the mean length has lower variability in almost all size categories with the size category-based design. This is due to the fact that in a "size-category" sampling design observers will haphazardly select boxes by size category from a list of all boxes by size category awaiting auction. Therefore, in an auction visit, as the aim is not to sample a number of trips by *métier* but a number of individuals by size category, the probability of getting samples by size category increases. Also, if the number of individuals in a size category box is not enough, it is possible to complete the required number taking fish from another box, provided that more than one box of that category was landed. In the case of the trip-based design the observer will sample only the size categories present in the randomly selected trips (2 or 3 trips) and will be limited to the number of fish that is present in each size category box. If this number is low (e.g. T1), the size category may present larger variability of the mean length.

Even though the results from the size category approach suggest that n = 4 fish would provide reasonable precision levels for the mean length of the most representative size categories, the analysis should be extended to other years to investigate the consistency in the variability of the mean length by size category within and among zones, before its implementation.

Moreover, the analysis should also look at the required number of [auction * visits] by zone to accomplish either k = 40 trips (trip-based design) or s = 40 samples by size category (size category-based design). In the first case, it means computing the probability of trips with hake landings in an auction day and, for a positive trip, the probability of occurrence of each size category in the landings of hake; in the second case, taking into account the probability of occurrence of each hake size category in an auction day. Decision on the number of samples by size category will be a trade-off between accepted levels of precision and costs associated. For example, in 2013 the landings of T1 in zone S were low and the lowest among zones. This means that achieving s = 40 in T1 in S would likely require an enormous number of [auction * visits] with unacceptable sampling costs and without significantly improving the precision of the annual landings length composition.

It is recommended to perform another type of analysis with prior definition of acceptable precision levels (*e.g. CV* ≤ 12%) for the more represented size categories in landings. The maximum number of samples and required [auction * visits] by zone to accomplish the target precision (e.g. in NW: s = 30 for T4 and s = 20 for T5, hence, maximum = 30) will be used to compute the precision of the less represented categories in the landings (T1, T2 and T3). Our results suggest that, in this case, the precision of the less represented size categories in landings will be ≥ 20% *CV*, which may well be acceptable given the low contribution of the landings of these size categories to the total landings length composition.

3.4 References

Azevedo M, Silva C, Vølstad JH, Prista N, Alpoim R, Moura T, Figueiredo I, Dias M, Fernandes AC, Lino PG, Felício M, Chaves C, Soares E, Dores S, Gonçalves P, Costa AM, Nunes C. 2014. Report of the Workshop on Sampling Design and Optimization of fisheries data. *Relat. Cient. Téc. do IPMA* (http://ipma.pt), nº 2, 79 pp.

Azevedo M, Silva C, Vølstad JH, 2016. Modelling length distribution by commercial size category to estimate species catch length composition for stock assessment. ICES Annual Science Conference, Riga – Latvia, 19-23 September 2016. ICES CM 2016/O:170.

EU, 2016a. Commission Implementing Decision (EU) No. 2016/1251 of 12 July 2016 adopting a multiannual Union programme for the collection, management and use of data in the fisheries and aquaculture sectors for the period 2017-2019. *Official Journal of the European Union* L 207: 113 – 177.

EU, 2016b. Commission Implementing Decision (EU) No. 2016/1701 of 19 August 2016 laying down rules on the format for the submission of work plans for data collection in the fisheries and aquaculture sectors. *Official Journal of the European Union* L 260: 153 – 228.

ICES. 2017. Report of the Working Group on Commercial Catches (WGCATCH), 7-11 November 2016, Oostende, Belgium. ICES CM 2016/SSGIEOM:03. 142 pp.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

4 Case Study 3: Blue-whiting age-at-length sampling

4.1 Introduction

Estimating ages from hard structures for a large number of fish is very time-consuming, whereas measuring the length of a large number of fish is usually relatively faster and simple. The age structure for a large number of fish can be estimated by the relationship between age and length for a relatively small subsample of fish and then applying an age-length key (ALK) to the entire sample of fish. The selection of the subsamples to be used to construct the age-length key could be random (i.e. the number of specimens aged from each length category proportional to the number in each length category) or fixed (i.e. a constant number of specimens aged from each length category) (Kimura, 1977). Currently, the most common method of subsampling is to create length-groups of 10-mm, 25-mm, or 1mm lengths and collect age data structures from a fixed number of fish per length-group. The ages of fish in the subsample are then estimated by various methodologies, and statistics such as mean length and variance are computed for each age group represented in the subsample (Betolli and Miranda, 2001).

In most of the cases, a major problem is to achieve the exact fixed number of fish aged by length, in order to guarantee a compromise between the time-spent ageing and the final data quality on the ALKs. Due to the difficulty on comparing and assessing ALKs quality, the evaluation of age-length estimates is usually done based on the growth models. In fisheries, the most widely used growth model is the von Bertalanffy model, derived in 1938 by von Bertalanffy and based on simple physiological arguments. This model assumes that the growth rate of a fish declines with size, the change in the fish growth rate (*dl/dt*) is described by (Eq. 4.1):

$$\frac{dl}{dt} = K(L\infty - t) \qquad (4.1)$$

where t is time, l is length, K is the growth rate and $L\infty$ is the asymptotic growth at which growth is zero. The von Bertalanffy growth model equation (Eq. 4.2) is given by integrating equation 4.1. and described as:

$$lt = L\infty\left(1 - e^{-K(t-t0)}\right) \qquad (4.2)$$

The main goal of this case study was to develop an algorithm to define the minimum fixed number of fish by length class that should be used to construct the age-length key. Length and age data (determined from otolith readings) for blue-whiting (*Micromesistius poutassou*) from year 2004 and 2008 were used as test cases.

The algorithm, the statistical analyses and the plots were performed using the statistical environment R (R Core Team, 2017).

4.2 Age-length Keys

4.2.1 Ageing sampling procedure

In the case of blue-whiting captured off the Portuguese coast, the landings in each port were sampled monthly (or more recently, quarterly) for collecting biological parameters. From each sample, a subsample composed of 10 fish by length class was selected for biological sampling. From each fish of the subsample, length, weight, sex and maturity stage were recorded and the otoliths were collected for posterior ageing. Before 2011, all the collected otoliths were aged. Since 2011, a random otolith selection by quarter of a fixed number of 10 by length class (5 females and 5 males) was made (Figure 4.1) and only those were aged.

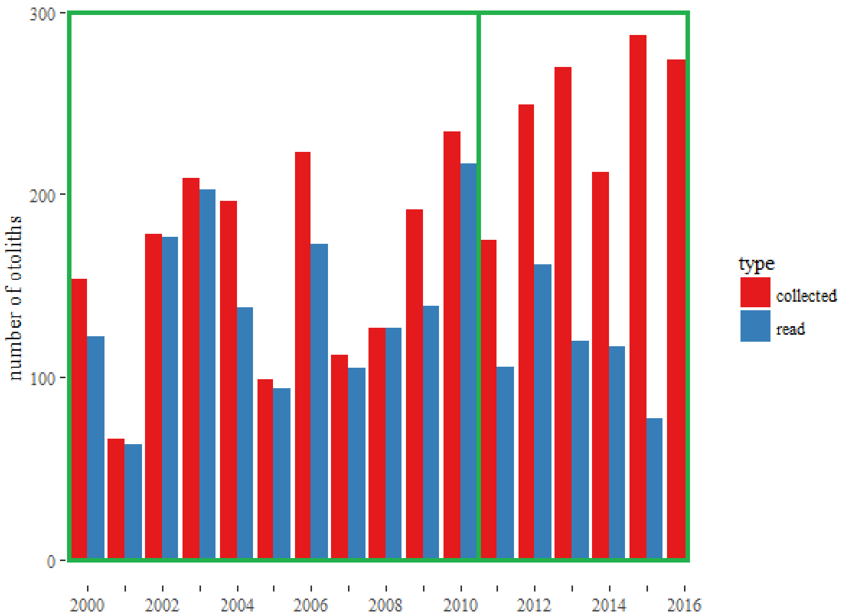


Figure 4.1. Number of blue-whiting otoliths collected (red) and aged (blue) by year, from 2000-2016, for the Portuguese coast. The green rectangles show the transition in the total number of aged otoliths: before 2011, all the collected were aged; after 2011, a random sample of 10 by length-class and quarter were aged.

4.2.2 Simulations

The following sampling scenarios were tested: the sampling period (quarter, semester, year) (Section 4.2.2.1) and the fixed number of otoliths to read by length class (1, 2, 4, 5, 10, 20, 30, 40, 50, 100) (Section 4.2.2.2).

4.2.2.1 Periodicity

The blue-whiting sampling data from 2004 were used to test whether the fixed number of 10 otoliths by length class (5 males and 5 females) uniformly distributed by port, should be collected by quarter (the current sampling scheme), by semester or by year. A total of 100 sampling simulations, without replacement, were performed in order to test each of those scenarios.

The available number of fish aged by length class in 2004 is shown in the histogram below (Figure 4.2).

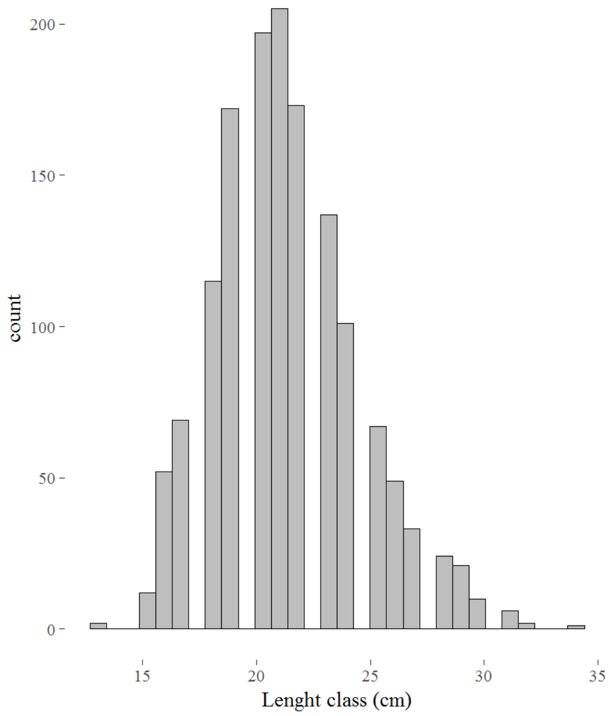


Figure 4.2. Number of blue-whiting otoliths aged by length class (cm) in 2004 (in total n=907).

The 2004 ALK representation is shown in Figure 4.3.

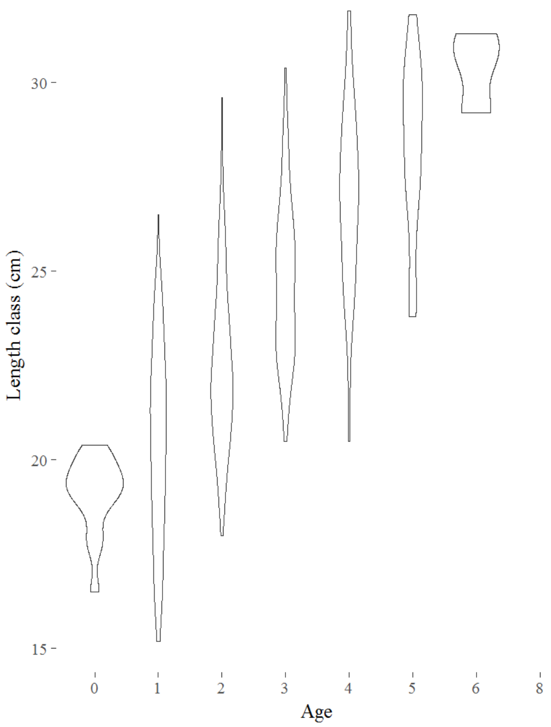


Figure 4.3. 2004 blue-whiting age-length distribution.

The von Bertalanffy growth model (4.2) was fitted to the age-length data obtained from the random sampling of 10 aged otoliths by length class (1 cm) and by quarter, by semester and by year and compared with the growth curve considering all the otoliths aged in 2004 (all) (Figure 4.4).

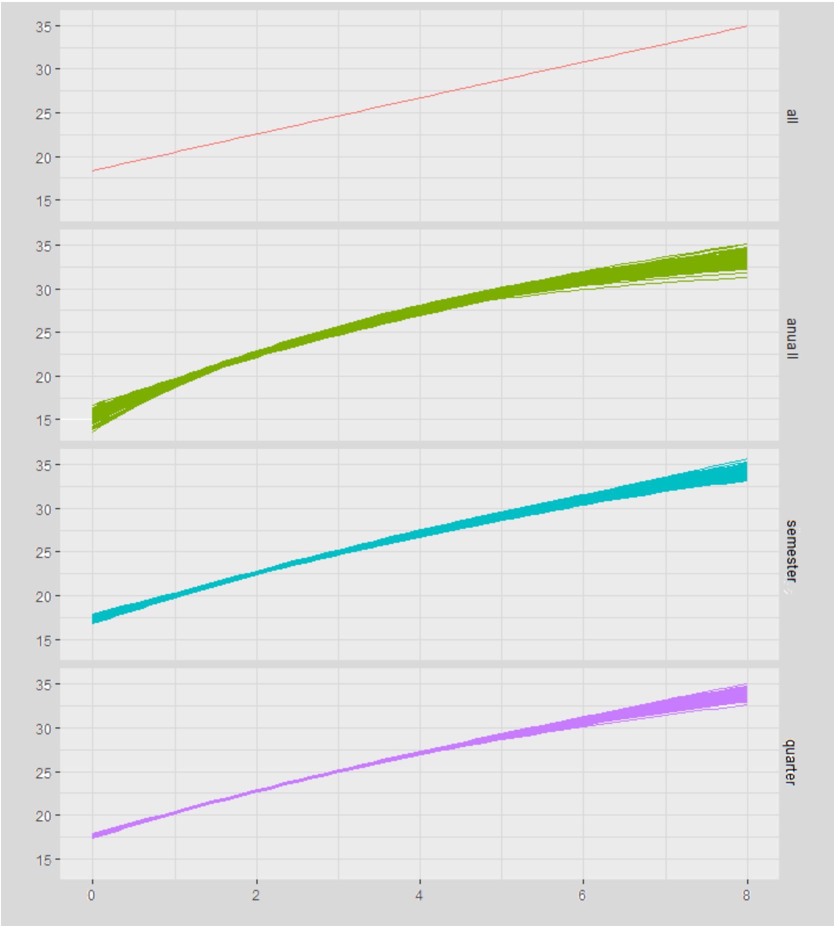


Figure 4.4. Blue-whiting growth curves from fitting von Bertalanffy growth model to all the 2004 data (red) (original data) and simulated samples of 10 otoliths by length class, using year (green), semester (blue) and quarter (violet) based selection. X-axis represents age and Y-axis represents length (cm).

The curves obtained by sampling a fixed number of otoliths selected by quarter and by semester were similar and also close to the curve when using all otoliths collected/read in 2004.

The age-length distribution from the 100 simulations based on quarter, semester and an annual sampling is shown in the next figure (Figure 4.5).

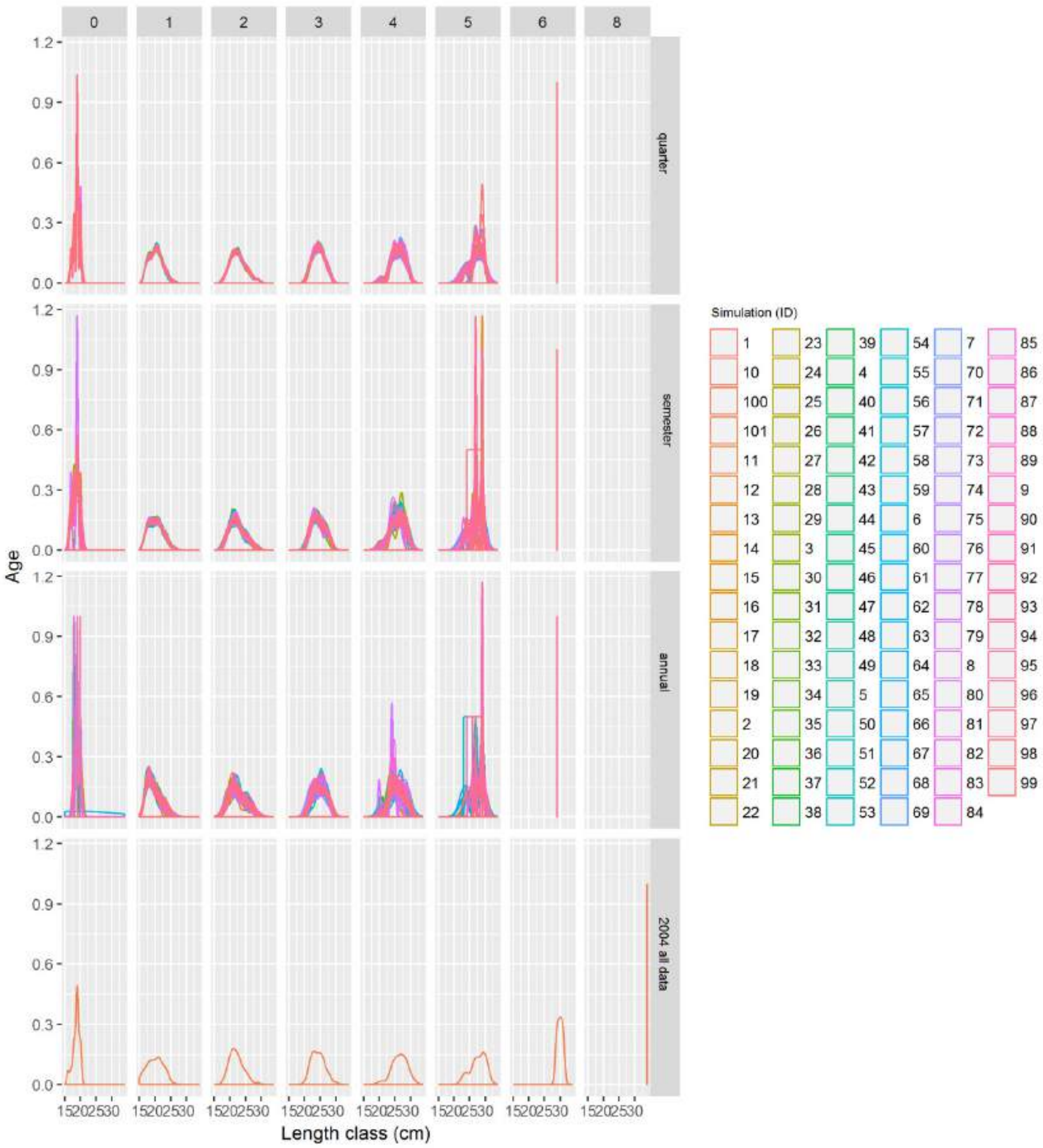


Figure 4.5. Blue-whiting length distribution by age from the 100 simulations (Simulation (ID)) of 10 otoliths per length class based on quarter, semester and annual sampling, and with all the 2004 available data (original data).

The parameters from the von Bertalanffy model (L_{inf}, k, t_0) obtained from the samples simulated by quarter, semester and year were also compared (Figure 4.6.).

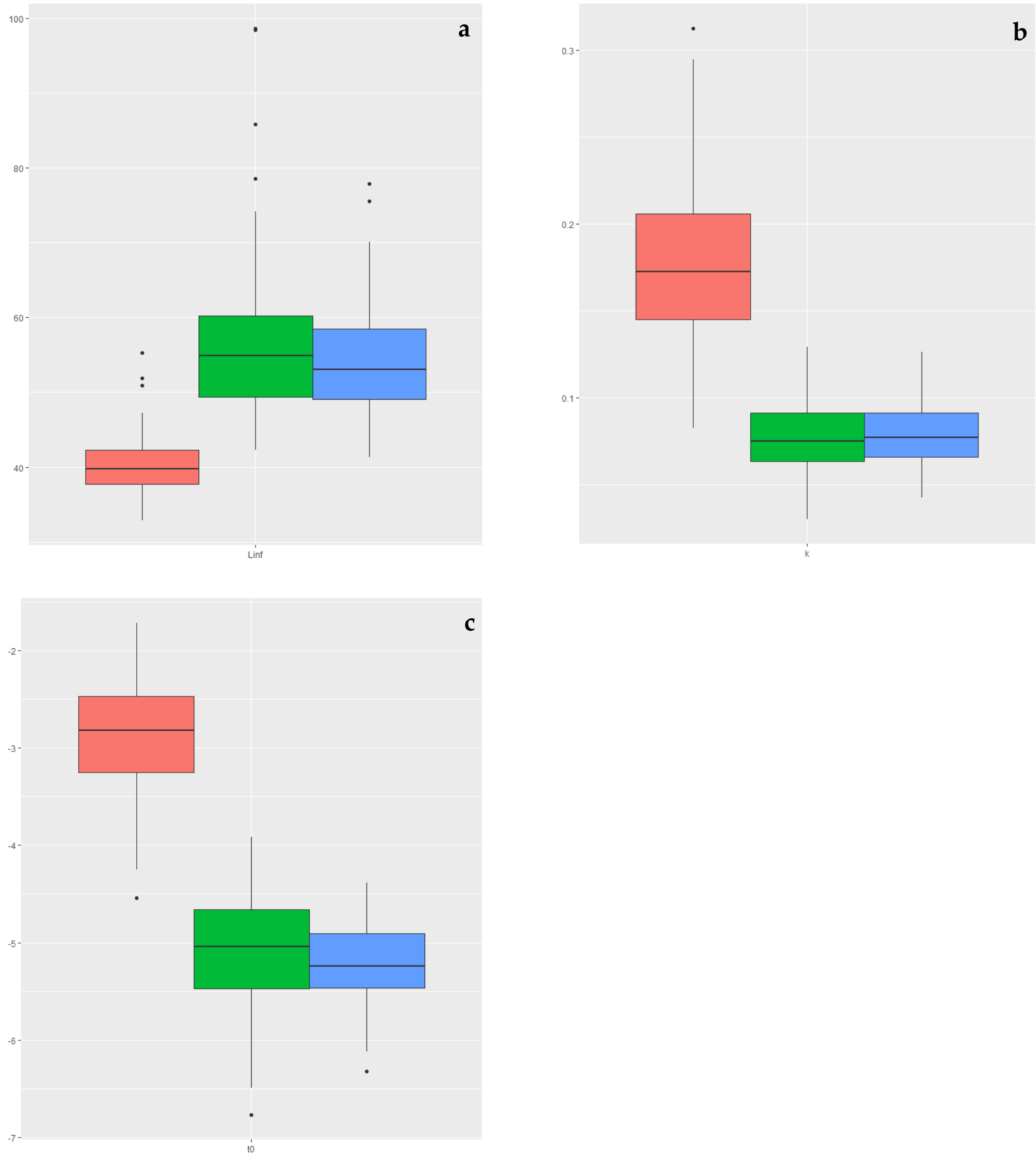


Figure 4.6. The parameters (L_{inf} (a), k (b) and t_0 (c)) from the von Bertalanffy growth model fitted to the 10-otoliths by length class data simulations, based on year (orange), semester (green) and quarter (blue) selection.

The figures above reveal similarities between the parameters obtained based on a quarterly and semester selection. Taking into account, the scenario of achieving the minimum number of aged fishes, the results support the change to the semester-based sampling period.

4.2.2.2 Number of otoliths by length class

The semester was then used as a base for testing the number of otoliths to read by length class, according to the results of the previous section (Section 4.2.2.1). The Portuguese blue-whiting sampling data of 2008 were used to evaluate what the minimum fixed number of otoliths should be aged by semester and length class (Figure 4.7), in order to guarantee that the growth model is still well fitted. A fixed number of 1, 2, 4, 5, 10, 20, 30, 40, 50 and 100 of otoliths by length class and semester were tested. In the length classes where these numbers of otoliths were not available, the total number of otoliths was used instead. A total of 100 simulations, resampling without replacement, were performed in order to test each of those effective sample sizes.

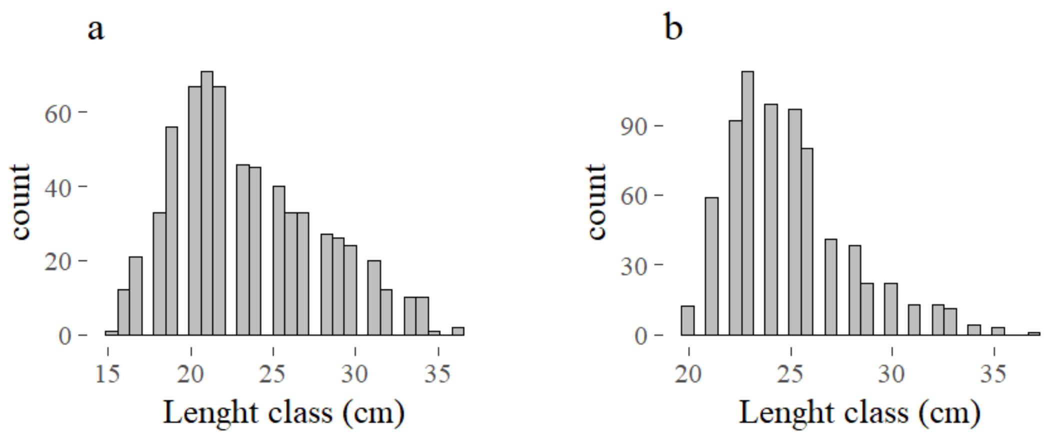


Figure 4.7. Number of blue-whiting aged otoliths by length class (cm) in 2008 and by semester. (a) 1st semester (total n=638) and (b) 2nd semester (total n=715).

The 2008 ALK representation is shown in Figure 4.8.

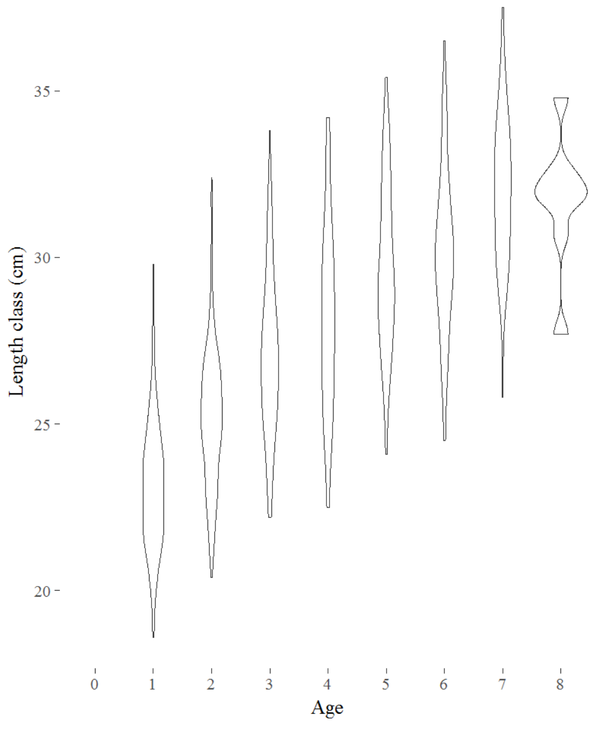


Figure 4.8. 2008 blue-whiting age-length distribution.

The age-length distribution based on the tested fixed number of otoliths by semester is presented in Figure 4.9.

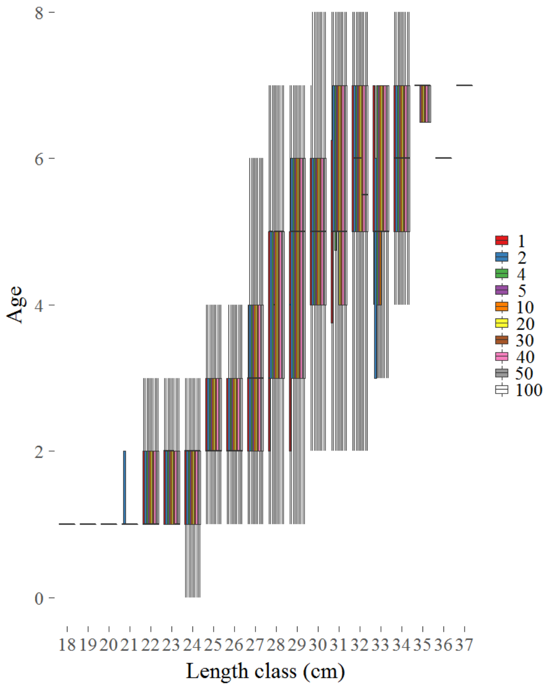


Figure 4.9. **Boxplots showing the length distribution by age (cm) by changing the fixed number of otoliths read by length class (1, 2, 4, 5, 10, 20, 40, 50, 100) and by semester.**

The von Bertalanffy growth model was fitted to the randomly sampled 1, 2, 4, 5, 10, 20, 30, 40, 50 and 100 aged otoliths per length class by semester and compared to the growth curve using all otoliths aged in 2008 (Figures 4.10 and 4.11).

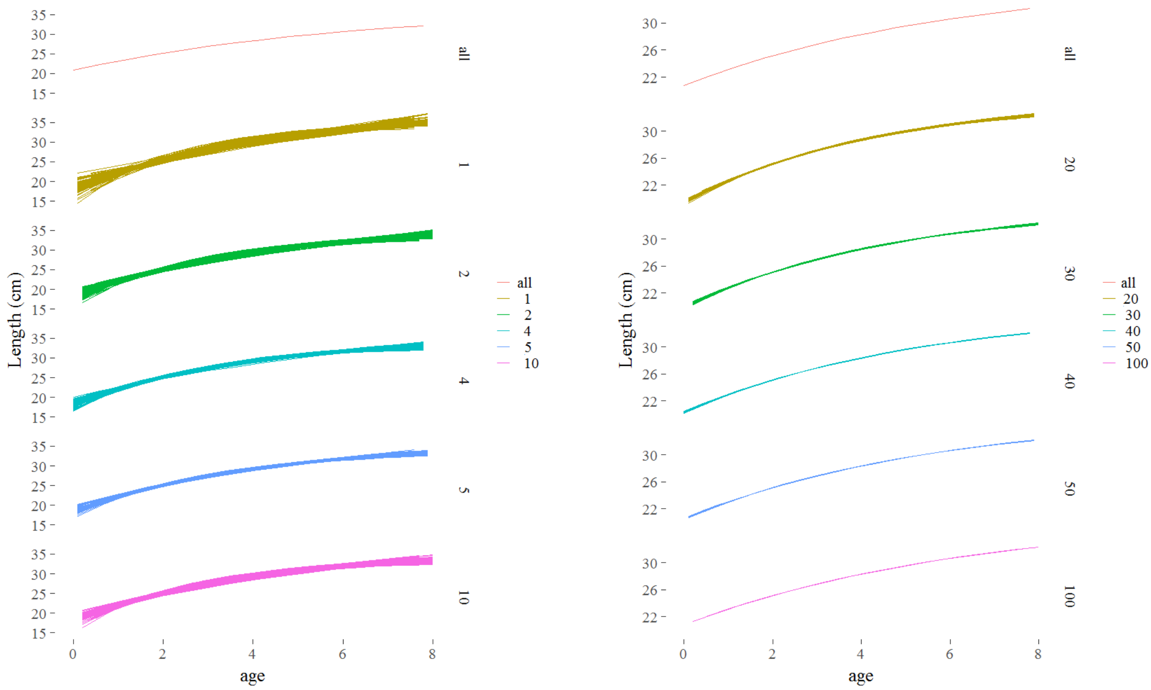


Figure 4.10. **Blue-whiting growth curves resulting from the von Bertalanffy growth model fitted to 100 simulations considering a fixed number of otoliths per length class by semester (1, 2, 4, 5, 10, 20, 30, 40, 50 and 100) compared to 2008 growth curve (original data).**

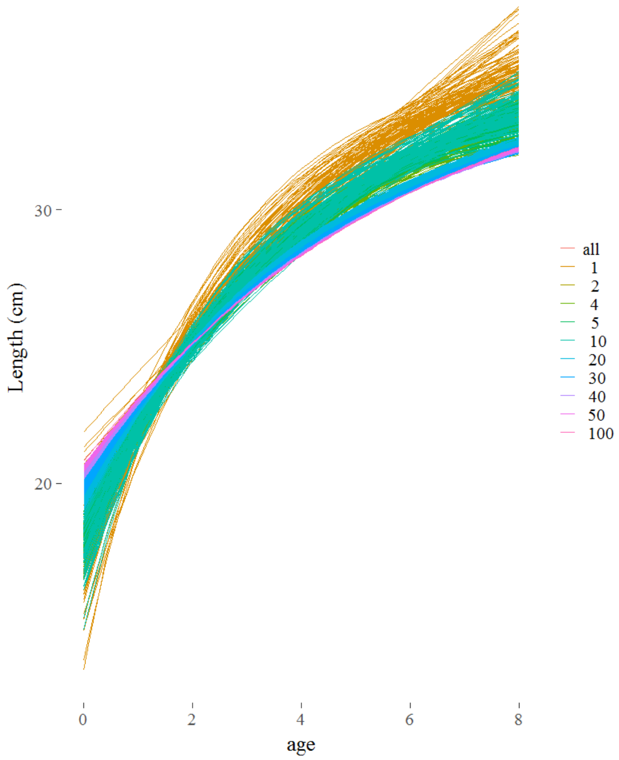


Figure 4.11. Blue-whiting growth curves resulting from the von Bertalanffy growth model fitted to 100 simulations considering a fixed number of otoliths per length class by semester (1, 2, 4, 5, 10, 20, 30, 40, 50 and 100) compared to 2008 growth curve (original data).

Figures 4.10 and 4.11 show the differences in the von Bertalanffy curve shapes according to the number of otoliths by length class, with two distinct groups, below and above 20 otoliths. Figure 4.10 also shows that in the cases where less otoliths were selected by length class, the curves present higher dispersion while an overlap is observed when more than 20 otoliths are sampled by length class.

The parameter values from the von Bertalanffy growth model fitted to the 2008 data were L_{inf} = 36.78, k = 0.16 and t_0 = -5.25. The values obtained from the simulations with a fixed number of otoliths are shown in Figure 4.12.

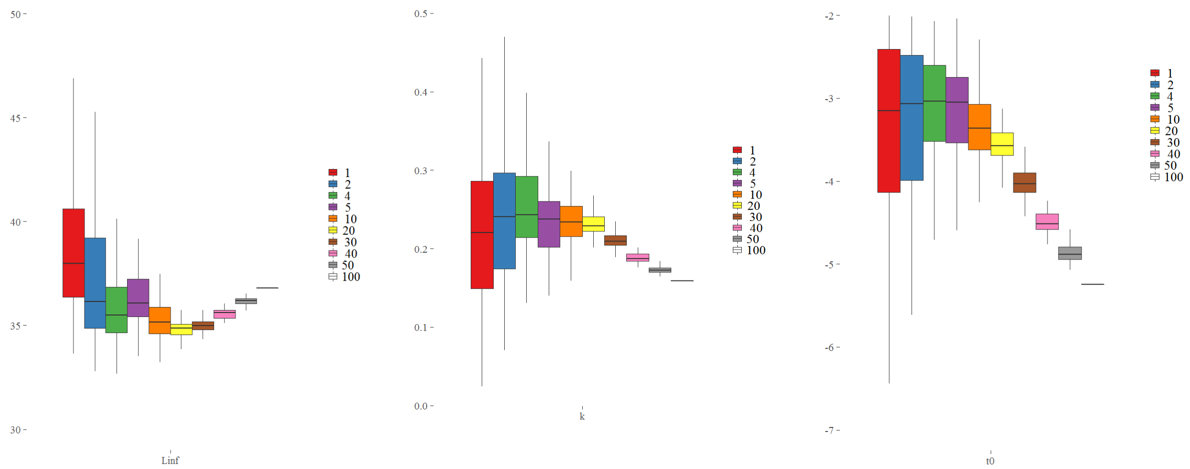


Figure 4.12. Variability of von Bertalanffy growth parameters (L_{inf}, k and t_0) for blue-whiting based on resampling a different number of otoliths per length class by semester.

The analysis indicates that with a fixed number of 30 otoliths per length class by semester, the growth curve is similar to the one obtained using all 2008 data showing low dispersion in parameter estimates (Figure 4.12).

The root mean squared error (RMSE) and the mean absolute percentage error (MAPE) from the von Bertalanffy growth parameter estimates were determined through (repeated) K-fold cross-validation, considering the scenarios described above and presented in Table 4.1. The prediction errors are small and very similar for a fixed number of 20 to 100 otoliths by length class.

Table 4.1. Prediction errors (RMSE and MAPE), estimated values and 95% confidence limits of von Bertalanffy growth parameters from cross-validation with fixed number of otoliths by length class (cm) by semester (2008 all data: L_{inf} = 36.78, k = 0.16 and t_0 = -5.25).

				Linf			k			t_0		
	Data	RMSE	MAPE	Median	95% LL	95% UL	Median	95% LL	95% UL	Median	95% LL	95% UL
	2008			36.86	33.99	44.24	0.16	0.09	0.23	-5.26	-7.59	-3.99
Fixed number of otoliths by length class (cm) and by semester	1	24.7	23.99	36.38	35.79	37.07	0.29	0.27	0.32	-1.87	-2.06	-1.69
	2	24.1	22.15	34.05	33.68	34.54	0.34	0.32	0.36	-1.62	-1.75	-1.49
	4	23.8	23.16	33.31	33.05	33.55	0.37	0.35	0.38	-1.57	-1.67	-1.48
	5	23.9	22.97	33.42	33.19	33.65	0.37	0.35	0.38	-1.57	-1.64	-1.47
	10	23.7	23.07	32.54	32.39	32.69	0.39	0.38	0.40	-1.53	-1.59	-1.47
	20	23.6	22.79	32.32	32.21	32.43	0.38	0.37	0.39	-1.66	-1.71	-1.61
	30	23.5	22.76	32.34	32.23	32.45	0.35	0.34	0.36	-1.96	-2.00	-1.91
	40	23.4	22.59	32.58	32.46	32.72	0.32	0.31	0.33	-2.24	-2.29	-2.19
	50	23.4	22.41	33.01	32.87	33.14	0.29	0.28	0.29	-2.54	-2.59	-2.48
	100	23.4	22.28	34.21	34.31	34.75	0.22	0.22	0.23	-3.48	-3.48	-3.33

Note: Confidence intervals of the VB model parameters (L_{inf}, k, t_0) were obtained by bootstrap of the mean centered residuals. A total of 1000 datasets were generated by resampling.

4.3 Discussion and future work

The results obtained for 2004 blue-whiting seem to indicate that a random sample selection of 10 otoliths by length class and by semester produces a growth curve similar to the curve based on a quarterly based random sample. However, this is not so clear from the analysis of 2008 data, which indicates a minimum number of 30 otoliths per length class by semester based on the similarity of growth curves and low parameter estimates dispersion. These results seem to be in line with Kimura (1977), which states that small increases in the age sample will likely increase the accuracy of an age-distribution more effectively than relative large increases in the length sample.

Using sardine as a case study, Azevedo *et al*. (2014) show that by taking an age random subsample (*i.e.* with the number of specimens aged from each length category proportional to the number in each length category), similar age-length distributions are obtained when the number of aged fish is reduced from 10 to 1 for each of the subsamples collected along the year. Similar results were obtained in a study using Pacific Ocean perch and Pacific cod as case studies (Kimura, 1977). In all the mentioned studies, the fixed number is selected on a sample/haul basis. The same principle is proposed by Aanes and Vølstad (2015). Based on simulations, they show that the collection of subsamples of one fish per 5 cm length bin (10 fish total) per haul or trip, in length-stratified samples, is sufficient and nearly as efficient as a random subsample of 20 fish.

It is important to state that this blue-whiting case study was primarily designed to apply and test the algorithm. Therefore, the results obtained should be regarded as preliminary and no changes should be made in the current sampling based on the current study at this stage.

Two approaches must be tested: (i) change the algorithm in order to take an age random subsample proportional to the length distribution by period (quarter, semester, annual); and (ii) the simulations shall be repeated using trips/ports as sample units instead of the time period. Moreover, the algorithm should be applied to a larger data set, considering a minimum of 10 years. Taking into account this new approach, the number of otoliths per length class and by period will be tested and evaluated. The subsequent results from this application shall be used to produce the inputs to stock assessment and the impact of the changes evaluated in terms of blue-whiting population structure results.

The authors believe that it will be possible to perform the correct and necessary changes in the sampling effort, i.e., to reduce the sampling effort and still obtain accurate growth estimates of the blue-whiting Portuguese component of the population and that the same principle is valid and could be applied to other fish species.

The algorithm, the statistical tests and the plot codes developed in R (R Core Team, 2017) will be made available as a tool to be applied to other fish species.

4.4 References

Aanes S, Vølstad JH, 2015. Efficient statistical estimators and sampling strategies for estimating the age composition of fish. *Canadian Journal of Fisheries and Aquatic Sciences*, 72(6), pp. 938–953.

Azevedo M, Silva C, Vølstad JH, Prista N, Alpoim R, Moura T, Figueiredo I, Dias M, Fernandes AC, Lino PG, Felício M, Chaves C, Soares E, Dores S, Gonçalves P, Costa AM, Nunes C. 2014. Report of the Workshop on Sampling Design and Optimization of fisheries data. *Relat. Cient. Téc. do IPMA* (http://ipma.pt), nº 2, 79 pp.

Bettoli PW, Miranda LE, 2001. Cautionary note about estimating mean length at age with sub-sampled data. *North American Journal of Fisheries Management*, 21: 425–428.

Kimura DK, 1977. Statistical assessment of the age-length key. *J. Fish. Res. Board Can.* 34: 317–324.

R Core Team, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

5 Case Study 4: Mackerel and hake maturity ogive sampling

5.1 Introduction

To collect data for the estimation of maturity ogives the spawning season, which is the period with higher proportion of actively spawning individuals (Murua *et al.*, 2003), must be known and in the case of species with indeterminate fecundity the best sampling time coincides with the peak of spawning activity (Dominguez-Petit *et al.*, 2017). Data can be collected during scientific surveys, on board of commercial vessels and from landings (market samples). However, market samples are often biased, due to the lack of individuals below the minimum landing size, which is the case of mackerel, the present case study.

The sampled number must be representative of the population and the sampling design needs to ensure a good coverage of the whole length range.

There are a number of studies that provide extensive review of several methods currently used to estimate fecundity in marine species in relation to their reproductive strategy (e.g. Murua *et al.*, 2003; ICES, 2008). In this study we evaluate the methodologies used to estimate maturity ogives, using data from hake and mackerel collected by PNAB/DCF (Programa Nacional de Amostragem Biológica/Data Collection Framework). We analyzed the effective sample size per length class and the implications in model performance.

5.2 Maturity sampling

5.2.1 Samples collection

Sampling should be carried out over the entire stock spatial distribution (including juvenile and adult areas). Preferably, the samples must be obtained in several ports and with different fishing gears, since the size of the captured specimens depends on the mesh size of the net. In species landed in commercial size categories (T1 to Tn categories), like mackerel and hake, information from landings given by the auctions helps to direct the sampling effort, so that it covers all size categories.

Sampling for the maturity ogive curve must have a greater effort on the length and/or age group within the transition from immature to mature individuals (Murua *et al.*, 2003), and to avoid misidentification between maturity stages, sampling must be done during the spawning season to reduce the macroscopic sampling error.

A record of the number of females sampled by length class must be kept along the sampling season, to allow a balanced sampling over the entire length range and avoid the acquisition of too many samples.

For the purpose of this contribution, mackerel data from 2011, 2012 and 2013 and also 2010 hake maturity data were analyzed.

Mackerel data

Historical analysis of the PNAB/DCF sampling data showed that individuals larger than 17 cm in length are very scarce; only 55 fish smaller than 17 cm were caught by the purse seiners in Peniche and Matosinhos, in June and July of 2005, 2007, 2010 and 2016 and out of the spawning season.

Commercial catches are mainly directed to larger sizes, so one should try to identify the port of landing and the fishing gear where smaller individuals may occur. However, there are not many records of landings of small individuals. The large Eastern Atlantic mackerel stock has its southernmost distribution limit in Portuguese waters and the observed length gaps in samples might reflect the stock structure and availability of individuals. Will this be a problem of sampling design or these fish do not inhabit our waters? However, length sampling gaps shall be covered through surveys of opportunity (e.g. IBTS and pelagic acoustic surveys) or discards data.

Because mackerel data showed some length classes with poor representation (e.g. 21-24cm) we combined female data from 2011, 2012, 2013 presented in Figure 5.1 and Table A.1, to test for optimal sample size, assess the implications of unbalanced sampling coverage in maturity ogive performance and explore the consequences of combining data and ogives from different years.

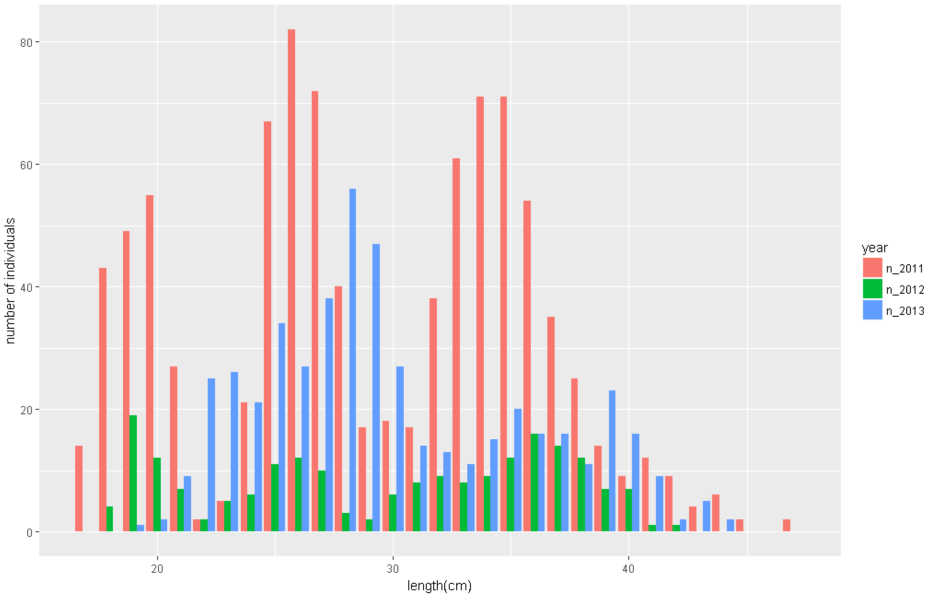


Figure 5.1. Mackerel samples by length class in 2011, 2012 and 2013.

<u>**Hake data**</u>

Because mackerel data showed some length classes with poor representation and to avoid potential problems in aggregating data from several years, we also analyzed the hake combined female and male maturity data from 2010. This dataset provided a good sampling coverage (1744 individuals) over the observed length range to evaluate the impact of different sample sizes per length class on the precision and variability of maturity ogives (Figure 5.2 and Table A.2).

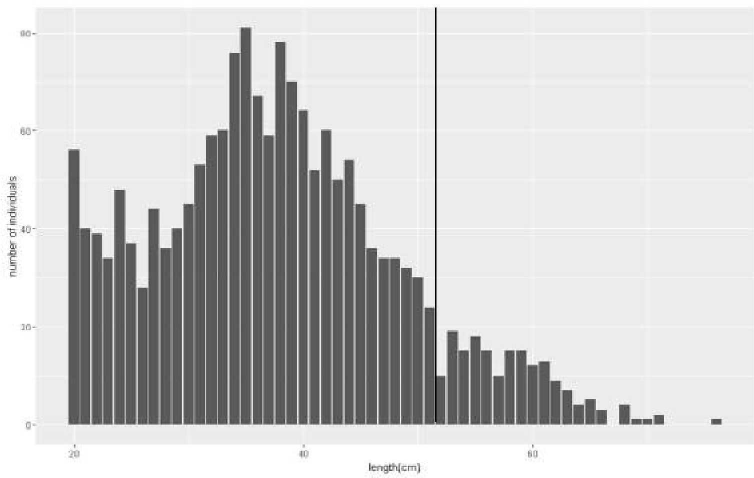


Figure 5.2. Hake maturity samples by length from 2010 (vertical line indicates the upper limit for which the resampling simulation analysis was performed, see section 5.3)

5.2.2 Macroscopic vs. microscopic identification

Distinguishing between mature and immature females is of great importance, because misidentification of these two stages may yield over- or underestimates of spawning stock biomass. Macroscopic maturity evaluation is a rapid and inexpensive manner of determining the reproductive status and allows for many fish to be assessed in the field (Tomkiewicz *et al*., 2003). However, this is often difficult to achieve in practice because maturation progresses continuously and the external aspect of the gonads may lead to doubts in the assignment of the maturity stage. This issue can be tackled with histological observations, though this process is more expensive and more time consuming.

5.3 Maturity ogive estimation

Typically, in the assessment of commercial species, maturity is assumed to be length (or age) dependent. The proportion of mature fish, *p*, increases with size (or age) and may be approximated by a logistic function of the type:

$$p = \frac{1}{1 + e^{(\alpha - \beta X)}} \tag{5.1}$$

However, this relationship is not linear (primarily due to the constraint that individuals are only either immature $p = 0$ or mature $p = 1$). Thus, the more convenient and robust option is to transform the logistic model into a linear equation. The required procedure transforms the response variable to *log(p/1-p),* which is called the *logit* function. With this transformation, a linear model is formed with:

$$logit(p) = \alpha + \beta X \tag{5.2}$$

With the appropriate data on mature fish (count data or proportion data) it is possible to estimate the parameters α e β. Rather than choosing parameters that minimize the sum of squared errors like in ordinary regressions, estimation in a logistic regression chooses parameters that maximize the likelihood of sample values. The logistic regression was fit in the software R (R Core Team, 2014) with the *glm* function using the binomial as the response variable distribution and a *logit* link.

To assess the precision and estimate the confidence intervals (CI) of the maturity ogive, it is more robust to use bootstrapping rather than the normal procedures (e.g. using the standard deviation of the parameter estimates). In our case study, bootstrapped precision estimates were originated from 100 replicates and 95% CI for the predicted probability of maturation at each length class was computed from each bootstrap sample by locating the 2.5^{th} and 97.5^{th} percentiles of the distribution of the bootstrapped samples. The resulting maturity ogives for mackerel in 2011 and hake in 2010 with the estimated 95% confidence intervals are shown in Figure 5.3.

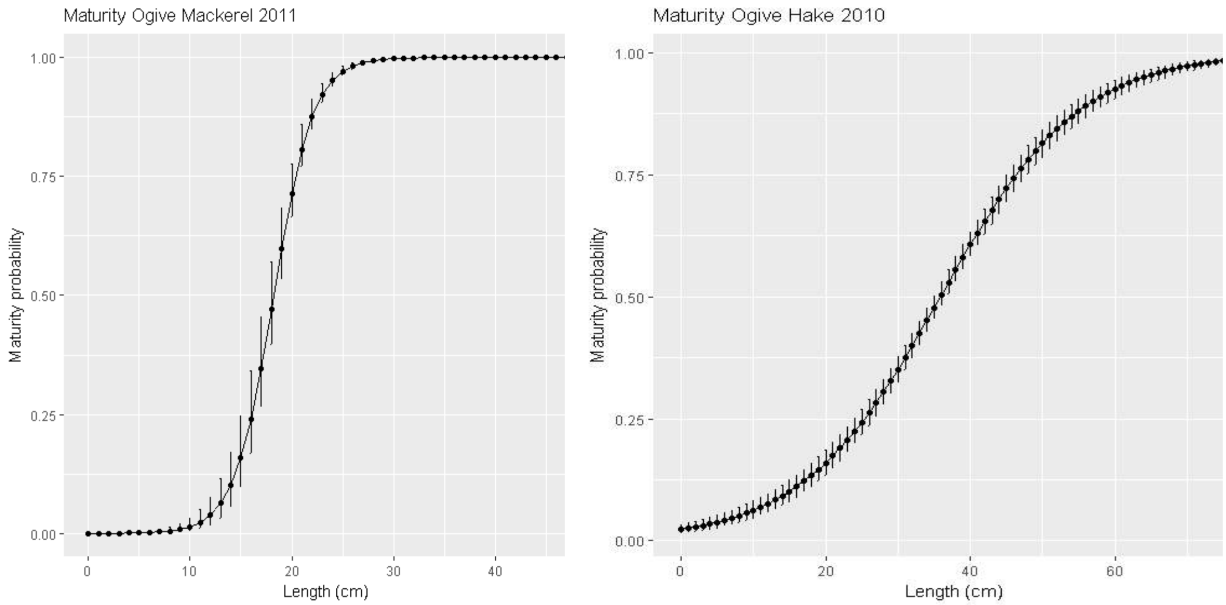


Figure 5.3. Mackerel (left) and hake (right) maturity ogives with estimated 95% confidence interval.

As shown in Figure 5.3 the hake's ogive has a slow (maturation) and constant growing curve and mackerel exhibits a fast (maturation) growing curve. Hake have a long range of sizes in which the individuals gradually mature. For this species the sampling effort should cover all length classes in which some individuals are still immature and others are maturing. To have an improved ogive for mackerel we should increase the sampling effort at the short range size in which individuals mature for the first time.

As the majority of commercial species in Portuguese waters have a spawning season in the first semester (with no age-0 individuals) it is proposed that the proportion of mature fish at each age class a, p_a, is derived from the length dependent maturity ogive with:

$$p_a = \frac{\sum_L n_l \cdot p_l \cdot q_{l,a}}{\sum_L n_l \cdot q_{l,a}} \qquad (5.3)$$

where L is the number of length classes, n_l is the number of individuals sampled at each length class, p_l is the proportion mature at length-class and $q_{l,a}$ is the proportion of individuals in each length-class that are in age-class a from the annual ALK.

5.4 Effective sample size and sampling variability on maturity ogive performance

Much discussion and statistical approaches are abundant in what is an adequate effective sample size to properly model biological parameters in fish populations. Particularly, the use of too small samples may lead to statistically non significant and inaccurate results, which may not properly reflect the reality. The objective in this section is to evaluate the impact of different sample sizes on the precision and variability of the maturity ogive using both hake and mackerel case study data. Optimal sample size is particularly discussed for the hake case study as in the mackerel a broader discussion is opened up regarding the implications of unbalanced sample coverage and the consequences of aggregating maturity ogives for stock assessment.

5.4.1 Hake case study (testing for sample size)

As stated before and because mackerel data showed some unbalanced sampling across years and length classes, we decided to test for effective sample size using the combined female and male hake data from 2010, which had sufficient sampling coverage of individuals across all length classes. To evaluate the impact of different sample sizes on the precision and variability of the maturity ogive we resample at each length class, n=20, n=15, n=10 and n=5 observations where more than 20 individuals per length class were observed (Figure 5.2). For each simulation the probability of maturation was estimated as in equation 5.2.

Figure 5.4 shows the boxplot of the prediction difference (Δp) distribution between the "true" proportion of mature at length – estimated from the full data set – and the estimates from the 100 simulations at each sample size. Each boxplot characterizes the vector of 100 Δp (or errors) at each length class and each has the median value of the distribution, the minimum and maximum values and the 25th and 75th percentiles of the error distribution, also showing the outliers assuming a normally distributed error. From simple visual inspection the error distributions of n=20, n=15 and n=10 models were quite similar with a slight increase in deviation as sample size decreases. Unexpectedly, models using n=15 had greater dispersions across several length classes comparatively to n=10. Models using a sample size n=5 exhibited a much wider distribution spread with some outliers surpassing the 20% difference in maturation probability. This means that sampling at this size introduced much lower prediction accuracies than the remainder of models.

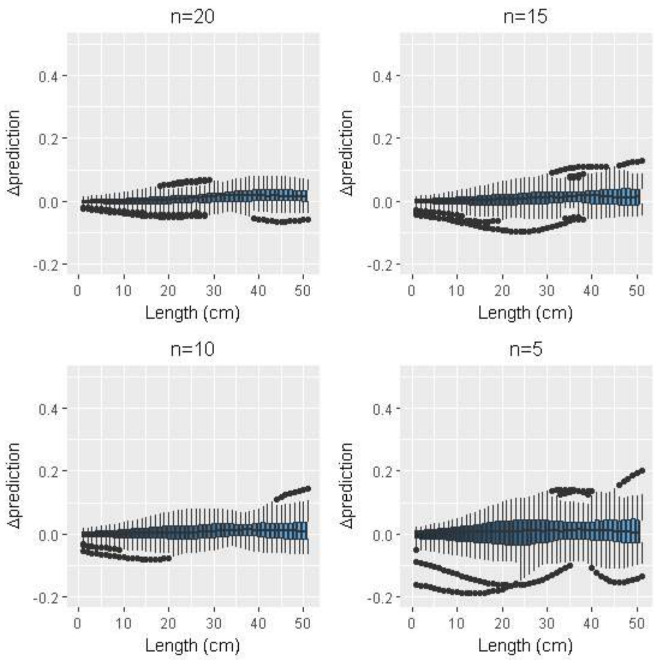


Figure 5.4. Boxplot of the prediction difference distribution between the "true" proportion of mature (estimated from the full data set) and the GLM estimates from 100 simulations. Each boxplot represents the error/Δprediction distribution at different sample sizes, n = 20, n=15, n=10 and n=5 by length class.

To quantify the impact of resampling at different sample sizes the mean absolute error (MAE) and the root mean square error (RMSE) were computed across the *L* resampled length classes to evaluate overall model precision and performance:

$$MAE = \sum_{i=1}^{L} |\Delta p_i| \tag{5.4}$$

$$RMSE = \sqrt{\sum_{i=1}^{L} \Delta p_i^2} \tag{5.5}$$

Both statistics were averaged over the 100 simulated models at each sample size and the results are shown in Table 5.1. As expected, model performance decreases as sample size reduces, more clearly when resampling at n=5. For comparative purposes, the length at 50% maturity was estimated from equation 5.2 as,

$$L_{50} = -\frac{\alpha}{\beta}$$

The mean L_{50} estimates at each sample size showed very similar results with a slight increase in variability as sample sizes decreases, yet again the variability at L_{50} was lower with n=10, comparatively to n=15.

Table 5.1. Root mean square error (RMSE) and mean absolute error (MAE) averaged from the 100 simulations at different sample sizes. For comparative purposes L_{50} estimates and standard deviation (sd) are also shown.

Sample size by length class	RMSE	MAE	L_{50} (sd)
20	0.15	0.93	36.5 (0.91)
15	0.18	1.12	36.5 (1.16)
10	0.19	1.17	36.4 (1.05)
5	0.29	1.88	36.3 (1.70)

5.4.2 Mackerel case study (testing for sampling variability)

The mackerel maturity ogive was estimated for the years 2011, 2012, 2013 and for all years combined data. The choice between using annual estimates or aggregating values across years may give different estimates. In addition to the differences in the L_{50} estimation, statistically significant differences were also found among the maturity ogive parameters from 2013 to both 2011 and 2012 and combined data (Figure 5.5). Looking at the distribution of individuals and the estimated maturity probabilities (Figure 5.5 and Table A.1), a lack of sampled individuals in 2011 and 2012 was observed in the upper growing part of the curve (length classes 22-24cm). On the other hand, 2013 maturity sampling did not cover enough small individuals (17-21cm) in the lower growing part of the curve. The maturity ogive estimate from the pooled data was very similar to the 2011 and 2012 due to a very large number of observations in these years on other (and probably less crucial) length classes. The information from the individuals sampled in 2013 for the upper growing part of the curve is being down weighted relative to the large amount of individuals observed on the remainder length classes. These results show that the appealing option of combining data from several years to cover sampling gaps can be hampered by unbalanced sampling coverage.

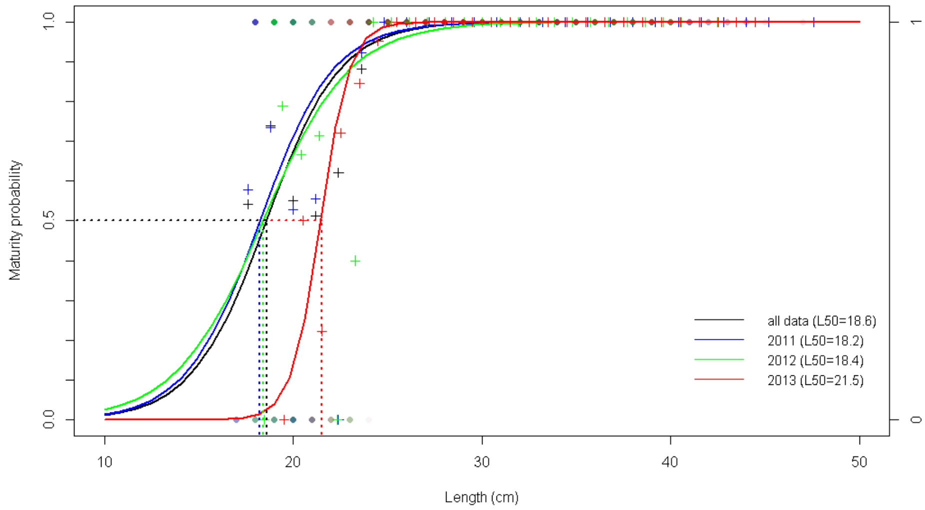


Figure 5.5. Maturity ogive for 2011, 2012, 2013 and combined data. For comparative purposes $L_{50\%}$ is also shown for each year.

When trying to repeat the simulation analysis on the effective sample size per length class to the combined mackerel maturity data we observed some inconsistent results (not shown) in the maturity ogive performance in the growing part of the curve. This is particularly evident in binary type responses/models where the use of i) too small samples, ii) unbalanced sampling and iii) non-existent or very low number of positive events may lead to low predictive power of the model and inaccurate estimations (see Hosmer *et al*., 2013).

For comparative purposes and to evaluate the implications for assessment, we estimated from equation 5.3 (using the Portuguese ALK) the mackerel maturity-at-age for 2011, 2012, 2013 and all years combined. As in the length based analysis, Table 5.2 shows that there are clear differences between 2013 and the other years in some of the age-classes, particularly in young ages that usually have many individuals. Consequently, a minor variation in the proportion of mature may imply a large variation in total SSB.

Table 5.2. Estimated maturity-at-age for 2011, 2012, 2013 and combined data

Age	2011	2012	2013	pooled
0	0.4	0.34	0.15	0.39
1	0.95	0.91	0.87	0.95
2	0.97	0.96	0.98	0.98
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12+	1	1	1	1

We simulated a fisheries population to illustrate the difference in SSB estimates, when using the 2013 maturity ogive or the combined ogive based on the available years, which is an often used option to reduce the observed effect of sampling variability. Assuming a population with recruitment distributed lognormally (with μ=6.91 and σ=0.99), equivalent to a mean recruitment of 1000 individuals with a CV = 10, we propagated the stock over the course of a 13-year period from 2001 to 2013 with M = 0.12 and F-at-age as estimated by ICES (2016) for the NE Atlantic mackerel. To better assess the implications of using different maturity ogives and not to introduce extra variability, SSB was estimated using the observed weight-at-age in 2013.

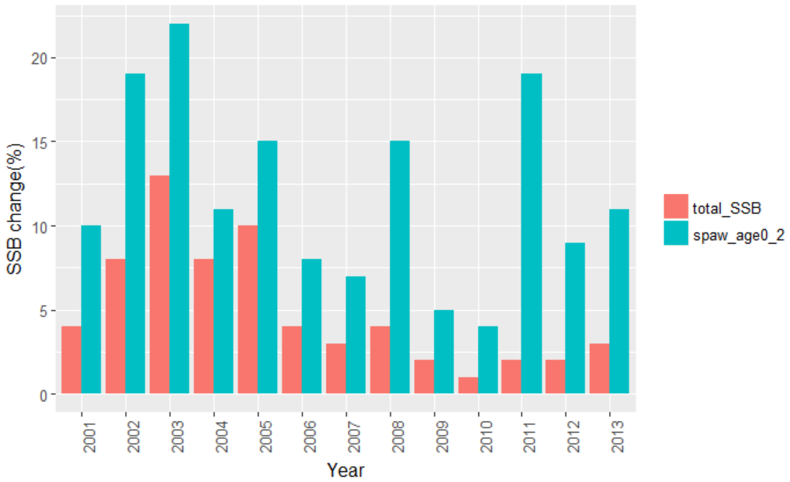


Figure 5.6. Percentage differences in estimated SSB and spawners at ages 0-2 by using the 2013 maturity ogive and the combined ogive.

Table 5.3. Simulated mackerel stock and the difference in total SSB and spawners at ages 0-2 (spaw0-2) by using the 2013 maturity ogive (ogive_13) and the combined ogive (ogive_all).

age / year	2001*	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
0	525.7	1587.7	3416	2155.1	6185.9	984.1	3298.8	3235.6	392.9	1046.1	1797.3	808.2	1888.4
1	449.3	449.3	1359.7	2928.4	1849.4	5297.7	843.6	2825.1	2773.8	336.9	896.8	1540.8	692.9
2	375.7	374.9	377.6	1149.4	2478	1561.8	4496.3	716.7	2402.6	2356.6	286.5	763.5	1311.7
3	302.4	302.4	302.4	303.3	927.1	2018.7	1290.3	3744.4	597.5	1998.8	1960.5	237.9	632.6
4	220.7	220.5	225.4	223.6	227.2	713.4	1572.2	1009.9	2939.6	470	1575.5	1554.6	187.3
5	143.6	143.4	147.9	154	158	163.3	518	1146.2	731.9	2128.2	344	1162.5	1145.9
6	87.9	86.4	86.4	90.7	99.1	105.3	107.9	346.5	777.6	498.5	1467.1	240.7	812.7
7	48.9	49	47.4	48.7	54	59.7	63.4	67.9	219.9	502.8	324.3	969.7	160.6
8	27.4	25.2	23.4	23.8	28.1	32	31.7	34.9	41	136.2	311.8	207	608.5
9	15.3	14.1	12	11.7	13.8	16.7	17	17.5	21.1	25.4	84.4	199	129.9
10	8.6	7.9	6.7	6	6.8	8.2	8.9	9.4	10.6	13	15.7	53.9	124.9
11	4.8	4.4	3.8	3.4	3.5	4	4.3	4.9	5.7	6.5	8.1	10	33.8
12	2.7	2.5	2.1	1.9	2	2.1	2.1	2.4	2.9	3.5	4.1	5.2	6.3
SSB (ogive_all)	468.7	503.4	718.4	1111.9	1566.3	2214.7	2455.2	2725.1	2863.5	2588.7	2357.5	2163.1	1911.6
SSB (ogive_13)	451.0	462.6	624.9	1024.3	1405.6	2119.9	2371.5	2615.4	2816.5	2561.2	2305.8	2124.2	1860.8
ΔSSB%	4	8	13	8	10	4	3	4	2	1	2	2	3
spaw0-2 (ogive_all)	178.7	216.2	431.3	821.4	1091.7	1263.7	1286.8	743.5	1020.8	632.9	276.9	457.0	481.7
spaw0-2 (ogive_13)	161.0	175.3	337.9	733.8	931.0	1168.9	1203.1	633.7	973.8	605.4	225.2	418.0	430.9
ΔSSB%	10	19	22	11	15	8	7	15	5	4	19	9	11

* initial population numbers in 2001 were estimated from $N_{i+1,2001}=N_{i,2001}*\exp(-z_{i,2001})$

Figure 5.6 shows the differences in total SSB and spawners at ages 0-2 by using the 2013 maturity ogive and the combined ogive. Total SSB estimates had differences ranging from 1% to 13% with the highest values observed between 2002 and 2005, reaching 13% and 10% in 2003 and 2005, respectively. These differences were mostly driven by the different spawners estimated at younger ages, amplified by the occasional strong year-classes observed in this period. The estimated spawner fraction weight at ages 0-2 achieved a mean difference of 12% over the entire period, fluctuating between 4%-22% (Table 5.3).

The results shown here indicate that, if in a given period we estimate the proportions of mature-at-age using all the available data, an option which could seem more appealing to reduce the observed effect of sampling variability, the SSB estimates for that period may suffer an artificial increase. In particular, as we saw earlier, this effect might occur when there is a strong evidence of unbalanced sampling among years and could be further amplified in pelagic species with high recruitment variability.

5.5 Discussion and future work

There are a number of studies that provide extensive review of several methods currently used to estimate fecundity data in marine species in relation to their reproductive strategy (e.g. Murua et al., 2003 ; ICES, 2008). However, some recommendations could be derived from the present study.

In order to obtain a good maturity ogive, it is vital to have an appropriate sampling of the lengths at which the individuals mature. A balanced sampling coverage of maturity data in the available length classes was proven to be particularly important. Moreover, binary type models between mature and immature are particularly sensitive to the use of small and unbalanced samples leading to a low predictive power of the model. Care should also be taken to non-existent or very low number of positive events in larger length classes.

As the majority of commercial species in Portuguese waters have a spawning season in the first semester (with no age-0 individuals) it is also proposed that the proportion of mature fish at each age should be derived from the length dependent maturity ogive.

When starting a species-specific maturity analysis for the first time, it will probably be necessary to collect a large sample size. After this initial intensive study, the sampling strategy can be refined by length, age, season and area to reduce the sampling effort. The analysis performed during this study on the effective sample size for maturity ogive estimation suggests that, in a trade-off between model performance and sampling effort, a sample of 10 females per length class could be appropriate, being aware of the number of positive events in the growing part of the logistic curve and the specific reproductive strategy.

Differences between years in maturity estimates can be caused by many factors: i) sampling variability, ii) uncertainty in macroscopic maturity stage determination, iii) changes or gaps in the population coverage (Portuguese waters are the mackerel stock south boundary) and iv) natural changes in sexual maturation caused by e.g environmental factors and fisheries induced evolution. To reduce the effect of sampling variability it may be appealing to use average values based on several years rather than annual estimates, however, averages may be less responsive if genuine annual changes occur, using average values for maturity can smooth out observed annual variations but can lead to substantial deviations between the assumed and actual values if multiyear trends are present. If, in a given set of years, the corresponding maturity ogives seem to have changed, SSB estimates for that period may suffer an artificial decrease or increase that will have consequences in the advice for management (Murta et al., 2011) that could be further amplified especially in fast maturity species with occasional strong year classes.

From the present study, we were able to identify some strategies for future work:

- As sampling for maturity ogive must have greater effort in length classes within the transition from immature to mature individuals, an initial evaluation should be done of each species-specific maturation strategy to better adjust the sampling strategy. Concurrently, evaluating the historical landings by size at different ports will allow a good coverage of needed individuals and to define an objective sampling effort.
- As the majority of commercial species in Portuguese waters have a spawning season in the first semester (with no age-0 individuals) it is recommended that the proportion of mature fish at each age class is derived from the length dependent maturity ogive by applying the annual ALK.

- A record of the number of females sampled by length class should be done along the sampling season, to allow a balanced sampling over the entire length range and avoid the acquisition of too many samples.
- Estimate the maturity ogives, based on microscopic identification to reduce the error, through different seasons and evaluate seasonal fluctuations.
- Calibrate macroscopic results with microscopic results.

5.6 References

Domínguez-Petit R, Anastasopoulou A, Cubillos L, Gerritsen HD, Gonçalves P, Hidalgo M, Kennedy J, Korta M, Marteinsdottir G, Morgado C, Muñoz M, Quincoces I, Saínza M, Thorsen A e Vitale F, 2017. Chapter 3: Maturity. *In* Domínguez-Petit R, Murua H, Saborido-Rey F and Trippel E (eds.). Handbook of applied fisheries reproductive biology for stock assessment and management. Vigo, Spain. Digital CSIC.

(http://digital.csic.es/bitstream/10261/87787/3/Dominguez-Petit_2017_Maturity.pdf)

Hosmer DW, Lemeshow S, 2000. Applied Logistic Regression. New York: Wiley, 2000 (2nd edition).

ICES, 2008. Report of the Workshop on Maturity Ogive Estimation for Stock Assessment (WKMOG) ICES CM 2008/ACOM:33. 3-6 June 2008 Lisbon, Portugal.

ICES, 2016. Working Group on Widely Distributed Stocks (WGWIDE). ICES CM 2016/ACOM:16. 31August-6 September 2016. ICES HQ

Murta AG, Costa AM, Gonçalves P, 2011. Variability in the maturity ogives of horse mackerel estimated with microscopic criteria. WD to the ICES Working Group on Anchovy and Sardine (WGANSA), 24-28 June 2011, Vigo, Spain, 6 pp.

Murua H, Kraus G, Saborido-Rey F, Witthames PR, Thorsen A, Junquera S, 2003. Procedures to estimate fecundity of marine fish species in relation to their reproductive strategy. *J. Northw. Atl. Fish. Sci.*, **33**: 33-54.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Tomkiewicz J, Morgan MJ, Burnett J, Saborido-Rey F, 2003. Available information for estimating reproductive potential of Northwest Atlantic groundfish stocks. *J. Northwest Atl. Fish. Sci.*, **33**: 1-21.

Table A.1. Mackerel sampling and observed proportion immature (prop_imat) and mature (prop_mat) in 2011, 2012 and 2013 by length class

Length	n_2011	prop_imat	prop_mat	n_2012	prop_imat	prop_mat	n_2013	prop_imat	prop_mat	n_total
17	14	1.00	0.00	NA	NA	NA	NA	NA	NA	14
18	43	0.23	0.77	4.00	1.00	0.00	NA	NA	NA	47
19	49	0.27	0.73	19.00	0.21	0.79	1.00	1.00	0.00	69
20	55	0.47	0.53	12.00	0.33	0.67	2.00	0.50	0.50	69
21	27	0.44	0.56	7.00	0.29	0.71	9.00	0.78	0.22	43
22	2	1.00	0.00	2.00	1.00	0.00	25.00	0.28	0.72	29
23	5	0.40	0.60	5.00	0.60	0.40	26.00	0.15	0.85	36
24	21	0.00	1.00	6.00	0.00	1.00	21.00	0.05	0.95	48
25	67	0.00	1.00	11.00	0.00	1.00	34.00	0.00	1.00	112
26	82	0.00	1.00	12.00	0.00	1.00	27.00	0.00	1.00	121
27	72	0.00	1.00	10.00	0.00	1.00	38.00	0.00	1.00	120
28	40	0.00	1.00	3.00	0.00	1.00	56.00	0.00	1.00	99
29	17	0.00	1.00	2.00	0.00	1.00	47.00	0.00	1.00	66
30	18	0.00	1.00	6.00	0.00	1.00	27.00	0.00	1.00	51
31	17	0.00	1.00	8.00	0.00	1.00	14.00	0.00	1.00	39
32	38	0.00	1.00	9.00	0.00	1.00	13.00	0.00	1.00	60
33	61	0.00	1.00	8.00	0.00	1.00	11.00	0.00	1.00	80
34	71	0.00	1.00	9.00	0.00	1.00	15.00	0.00	1.00	95
35	71	0.00	1.00	12.00	0.00	1.00	20.00	0.00	1.00	103
36	54	0.00	1.00	16.00	0.00	1.00	16.00	0.00	1.00	86
37	35	0.00	1.00	14.00	0.00	1.00	16.00	0.00	1.00	65
38	25	0.00	1.00	12.00	0.00	1.00	11.00	0.00	1.00	48
39	14	0.00	1.00	7.00	0.00	1.00	23.00	0.00	1.00	44
40	9	0.00	1.00	7.00	0.00	1.00	16.00	0.00	1.00	32
41	12	0.00	1.00	1.00	0.00	1.00	9.00	0.00	1.00	22
42	9	0.00	1.00	1.00	0.00	1.00	2.00	0.00	1.00	12
43	4	0.00	1.00	NA	NA	NA	5.00	0.00	1.00	9
44	6	0.00	1.00	NA	NA	NA	2.00	0.00	1.00	8
45	2	0.00	1.00	NA	NA	NA	NA	NA	NA	2
46	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
47	2	0.00	1.00	NA	NA	NA	NA	NA	NA	2
Total	942	-	-	203	-	-	486	-	-	1631

Table A.2. Hake sampling and observed proportion immature (prop_imat) and mature (prop_mat) in 2010 by length class (combined males and females)

length(cm)	n_samples	prop_imat	prop_mat	length(cm)	n_samples	prop_imat	prop_mat
20	56	0.91	0.09	49	32	0.31	0.69
21	40	0.90	0.10	50	30	0.17	0.83
22	39	0.87	0.13	51	24	0.17	0.83
23	34	0.82	0.18	52	10	0.20	0.80
24	48	0.81	0.19	53	19	0.16	0.84
25	37	0.73	0.27	54	15	0.00	1.00
26	28	0.89	0.11	55	18	0.06	0.94
27	44	0.70	0.30	56	15	0.07	0.93
28	36	0.72	0.28	57	10	0.10	0.90
29	40	0.65	0.35	58	15	0.07	0.93
30	45	0.49	0.51	59	15	0.07	0.93
31	53	0.53	0.47	60	12	0.08	0.92
32	59	0.47	0.53	61	13	0.00	1.00
33	60	0.45	0.55	62	9	0.00	1.00
34	76	0.46	0.54	63	7	0.00	1.00
35	81	0.44	0.56	64	4	0.00	1.00
36	67	0.52	0.48	65	5	0.00	1.00
37	59	0.63	0.37	66	3	0.33	0.67
38	78	0.40	0.60	67	0	NA	NA
39	70	0.43	0.57	68	4	0.00	1.00
40	64	0.45	0.55	69	1	0.00	1.00
41	52	0.40	0.60	70	1	0.00	1.00
42	60	0.35	0.65	71	2	0.00	1.00
43	50	0.46	0.54	72	0	NA	NA
44	54	0.39	0.61	73	0	NA	NA
45	45	0.24	0.76	74	0	NA	NA
46	36	0.36	0.64	75	0	NA	NA
47	34	0.29	0.71	76	1	0.00	1.00
48	34	0.15	0.85				
				Total	1744	-	-

6 Main Conclusions and Recommendations

The analysis to optimize the number of individuals that need to be measured at surveys (Case Study 1) was performed for the Portuguese bottom trawl demersal surveys in 2015 and 2016, based on the methodology of Pennington *et al*. (2002), using data of two species, the European hake - *Merluccius merluccius* (HKE) and the Atlantic horse mackerel – *Trachurus trachurus* (HOM). Simulations were performed considering 1) the whole survey region (without taking into account any strata), 2) the survey region stratified by sectors and 3) the survey region stratified by zones (N, SW and S) and by depth (3 levels). In general, the results indicate that for both HKE and HOM, the effective sample size to estimate the mean length is very small compared to the number of sampled fish. In exercise 3) more extreme effective sample size estimates were observed for HOM in certain strata, what may require a further analysis of whether length distributions could be more precisely defined. Future work should extend these analyses to other survey species, using a longer period and use the estimated average optimal sample size as a reference in the upcoming surveys. It is recommended to evaluate how the change in sampling effort (using the effective sample size) affects the raising procedure and the total length composition estimated by area (zone, sector or stratum).

Case Study 2 aimed to analyze the sampling effort and the effective sample size to estimate the species annual landings length composition, based on two designs: "trip-based" and "size category-based". HKE 2013 length at-market sampled data by size category were used in the simulations. The results suggest that the number of trips to be sampled, for the characterization of size categories length distributions, may be reduced to around 40 in each zone. The results from the size category-based design suggest that $n = 4$ fish would provide reasonable precision levels for the mean length for the most representative size categories. Decision on the number of samples by size category will be a trade-off between accepted levels of precision and costs associated. It is recommended to perform another type of analysis by defining *a priori* acceptable precision levels (e.g. $CV \leq 12\%$) for the more represented size categories in landings and compute the precision of the less represented categories in the landings and corresponding sampling costs. The methodology should be applied to other species landed in size categories, given the likely significant reduction in sampling effort, effective sample size and sampling costs.

Case Study 3 focused on growth parameters and age-length key, using length and age data (determined from otolith readings) for blue-whiting (*Micromesistius poutassou*) collected from at-market sampling in 2004 and 2008. An algorithm was applied to define the minimum number of fish by length class that should be used to construct the age-length key. Different scenarios were simulated, varying the sampling period and the fixed number of otoliths to be read by length class. Preliminary results suggest a reduction in the sampling effort without affecting the quality of the growth estimates. For blue-whiting, it is recommended a minimum of 30 otoliths per length class by semester considering the similarity of the estimated growth curve with the one obtained with all 2008 data. This algorithm should be applied to a larger data set, considering a minimum of 10 years, and the impact in the population numbers-at-age for stock assessment evaluated.

In case Study 4, mackerel and hake maturity data were analyzed to test for optimal sample size, assess the implications of unbalanced sampling coverage in maturity ogive performance and explore the consequences of combining data and ogives from different years. Results on the effective sample size for maturity ogive estimation suggests that, in a trade-off between model performance and sampling effort, a sample of 10 females per length class could be appropriate. A balanced sampling coverage of maturity data in the observed length range was proven to be crucial for mature and immature binary type models which are more sensitive to the use of unbalanced samples, reducing the predictive power of the model. Moreover, if in a given set of years the corresponding maturity ogive seems to have changed by natural causes or deficient sampling, the appealing option of using maturity data from several years to reduce the observed effect of sampling variability, results in different SSB estimations that will have consequences in the advice for management. To improve some of the mentioned issues, it is recommended that a record of the number of females sampled by length class is kept along the sampling season, promoting a balanced sampling over the entire length range. To reduce the maturity ogive bias and variability from observation errors, it is also recommended to calibrate macroscopic results with microscopic results.

Annex 1: List of participants

Manuela Azevedo (Chair) IPMA - Algés mazevedo@ipma.pt	Cristina Silva (co-Chair) IPMA- Algés csilva@ipma.pt	Corina Chaves IPMA- Algés cchaves@ipma.pt
Rui Coelho IPMA- Olhão rcoelho@ipma.pt	Ana Maria Costa IPMA- Algés amcosta@ipma.pt	David Dinis IPMA- Algés david.dinis@ipma.pt
Sandra Dores IPMA- Algés sdores@ipma.pt	Ana Cláudia Fernandes IPMA- Algés acfernandes@ipma.pt	Patrícia Gonçalves IPMA- Algés pgoncalves@ipma.pt
Pedro Lino IPMA-Olhão plino@ipma.pt	Hugo Mendes IPMA- Algés hmendes@ipma.pt	Teresa Moura IPMA- Algés tmoura@ipma.pt
Melinda Oroszlányová IPMA- Algés melinda.oroszlanyova@ipma.pt	Daniel Pinto IPMA- Algés daniel.pinto@ipma.pt	Maria do Carmo Silva IPMA- Algés mcarmo@ipma.pt
Cristina Nunes IPMA- Algés cnunes@ipma.pt		

List of invited participants:

Catarina Maia IPMA - Algés cmaia@ipma.pt	Mónica Inácio IPMA- Algés monica.inacio@ipma.pt	Bernardo Alcoforado IPMA- Algés balcoforado@ipma.pt

Annex 2: Meeting Agenda

WORKSHOP ON SAMPLING EFFORT FOR BIOLOGICAL PARAMETERS (WKSEBP)

IPMA, 18 – 20 April 2017

Day 1	
09:30 – 09:40	**MAzevedo** – Welcome message. Presentation of participants.
09:40 – 11:15	**MOroszlányová** - Assessing the precision of length-frequency estimates. Presentation/Discussion (data set/methodology/R script) & workplan
11:15 – 11:30	Coffee break
11:30 – 13:00	**MAzevedo/CSilva** – at-market sampling for landings length composition: hake commercial size categories case-study. Presentation/Discussion (data set/methodology/R script) & workplan
13:00 – 14:00	Lunch break
14:30 – 16:00	**PGonçalves** – Sampling for ALKs: blue-whiting case-study. Presentation/Discussion (data set/methodology/R script) & workplan
16:00 – 16:15	Coffee break
16:15 – 17:30	**AMCosta/CNunes/MCSilva** – at-market sampling for maturity ogive: mackerel case-study. Presentation/Discussion (data set/methodology) & workplan
17:30 – 18:00	Wrap up & setup of the working plan for Day 2

Day 2	
09:30 – 11:00	Working with case-studies/data sets
11:00 – 11:15	Coffee break
11:15 – 13:00	Working with case-studies/data sets
13:00 – 14:00	Lunch
14:00 – 16:00	Working with case-studies/data sets
16:00 – 16:15	Coffee break
16:15 – 18:00	Presentation of results and discussion

Day 3	
09:30 -11:00	Presentation of results and discussion
11:00 – 11:15	Coffee break
11:15 – 13:00	Discussion
13:00 – 14:00	Lunch
14:00- 16:00	Outline of the report & report writing
16:00 – 16:15	Coffee break
16:15 – 18:00	Report writing
18:00	End of the meeting

Annex 3: Glossary of terms

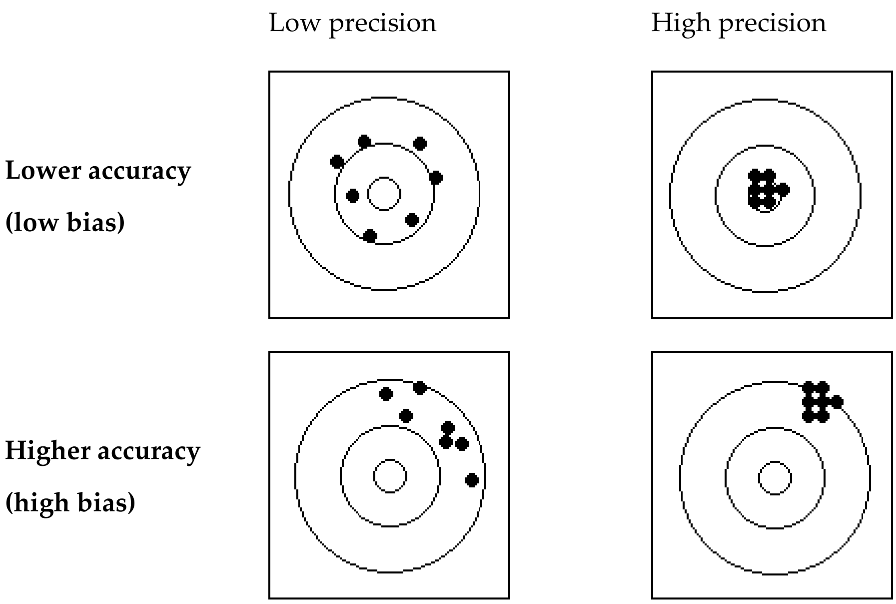
ACCURACY (exactidão)

An indicator of the closeness of an estimated value (e.g. population parameter) to the actual value. Accuracy refers to the closeness of computations or estimates to the exact or true values that the statistics were intended to measure.

BIAS (enviesamento)

A term which refers to how far the average statistic lies from the parameter it is estimating, that is, the error which arises when estimating a quantity. Errors from chance will cancel each other out in the long run, those from bias will not.

A measurement procedure or estimator is said to be biased if, on the average, it gives an answer that differs from the truth. The bias is the average (expected) difference between the measurement and the truth. For example, if you get on the scale with clothes on, that biases the measurement to be larger than your true weight (this would be a positive bias). The design of an experiment or of a survey can also lead to bias. Bias can be deliberate, but it is not necessarily so.



AGE-LENGTH KEY (chave de idade-comprimento)

The age structure for a large number of fish can be estimated by summarizing the relationship between age and length for a relatively small subsample of fish and then applying this summary to the entire sample of fish, which is named as an age-length key (ALK).

BOOTSTRAPING

In statistics bootstrapping is a method for estimating the sampling distribution of an estimator by resampling with replacement from the original sample.
(http://www.ices.dk/Lists/Glossary/DispForm.aspx?ID=290)

CONCURRENT SAMPLING (amostragem simultânea)

Sampling all or a predefined group of species that are simultaneously present in landings (or catches) of a certain fishing trip.

DESIGN EFFECT (efeito do delineamento)

A design effect (DEFF) is an adjustment made to find a survey sample size, due to a sampling method (e.g. cluster sampling, respondent driven sampling, or stratified sampling) resulting in larger sample sizes (or wider confidence intervals) than you would expect with simple random sampling (SRS). The DEFF tells you the magnitude of these increases.

The design effect is the ratio of the actual variance to the variance expected with SRS. It can more simply be stated as the actual sample size divided by the effective sample size (the effective sample size is what you would expect if you were using SRS). For example, let's say you were using cluster sampling. A DEFF of 2 means the variance is twice as large as you would expect with SRS. It also means that if you used cluster sampling, you'd have to use twice the sample size. (http://www.statisticshowto.com/design-effect/)

EFFECTIVE SAMPLE SIZE (tamanho efectivo da amostra)

An effective sample size (sometimes called an adequate sample size) in a study is the one that will find a statistically significant effect for a scientifically significant event. In other words, an effective sample size ensures that an important research question gets answered correctly. In order to achieve this, your sample must be the "right" size: neither too big nor too small. This is more of an art than a science. Note: The term "effective sample size" is also used in the calculation of design effects and has a much narrower definition; specifically, it's the sample size you would expect if you used simple random sampling. (http://www.statisticshowto.com/effective-sample-size/)

K-FOLD CROSS VALIDATION (validação cruzada com k-amostras)

Cross-Validation is a statistical method of evaluating and comparing learning algorithms by dividing data into two segments: one used to learn or train a model and the other used to validate the model. In typical cross-validation, the training and validation sets must cross-over in successive rounds such that each data point has a chance of being validated against. In k-fold cross-validation the data is first partitioned into k equally (or nearly equally) sized segments or folds. Subsequently k iterations of training and validation are performed such that, within each iteration a different fold of the data is held-out for validation while the remaining *k-1* folds are used for learning. (http://leitang.net/papers/ency-cross-validation.pdf)

MATURITY OGIVE (ogiva de maturação)

A distribution curve with the cumulative proportions of mature and immature individuals.

MEAN ABSOLUTE ERROR (erro absoluto médio)

The Mean Absolute Error (MAE) is the average of all absolute errors. Absolute Error is the amount of error in your measurements. It is the difference between the measured value and "true" value.

The formula is:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - x|$$

where n is the number of errors and $|x_i - x|$ are the absolute errors.

MEAN ABSOLUTE PERCENTAGE ERROR (erro percentual médio absoluto)

The mean absolute percentage error (MAPE), also known as mean absolute percentage deviation (MAPD), measures the accuracy of a method for constructing fitted time series values in statistics. It calculates the mean absolute percentage error (Deviation) function for the forecast and the eventual outcomes.

The mean absolute percentage error (MAPE) is defined as follows:

$$MAPE = \frac{100}{N} \times \sum_{i=1}^{N} \left| \frac{x_i - \hat{x}_i}{x_i} \right|$$

where x_i is the actual observations time series, $\hat{x}_i$ is the estimated or forecasted time series and N is the number of non-missing data points.
(http://www.spiderfinancial.com/support/documentation/numxl/reference-manual/forecasting-performance/mape)

PILOT PLAN (plano piloto)

Small scale preliminary study conducted in order to evaluate feasibility, time, cost, adverse events, and effect size (statistical variability) in an attempt to predict an appropriate sample size and improve upon the study design prior to performance of a full-scale research project.

POPULATION (população)

A collection of units being studied. Units can be people, places, objects, epochs, drugs, procedures, or many other things. Much of statistics is concerned with estimating numerical properties (parameters) of an entire population from a random sample of units from the population. In fisheries, a group of fish of one species which shares common ecological and genetic features. The stocks defined for the purposes of stock assessment and management do not necessarily coincide with self-contained populations.

PRECISION (precisão)

It is a measure of how close an estimator is expected to be to the true value of a parameter. Precision is usually expressed in terms of imprecision and related to the standard error of the estimator. Less precision is reflected by a larger standard error (or a larger coefficient of variation).
https://store.fmi.uni-sofia.bg/fmi/statist/education/statlib/glossary/index.htm

RANDOM (aleatório)

Random selection is where each member of the population has an equal chance of selection. Haphazard means that a person picks items, presumably trying to emulate randomness. However, that person's choice could easily be biased i.e. not truly random.

RESAMPLING (reamostragem)

Resampling is the method that consists of drawing repeated samples from the original data samples. Resampling may be performed with or without replacement.
http://www.statisticssolutions.com/sample-size-calculation-and-sample-size-justification-resampling/

ROOT MEAN SQUARED ERROR (desvio padrão empírico (amostral) generalizado)

The RMSE of a parameter estimator is the square-root of the mean squared error (MSE) of the estimator. In symbols, if $\hat{\theta}$ is an estimator of the parameter θ, then

$$RMSE(\theta) = \sqrt{E[(\hat{\theta} - \theta)^2]}$$

The RMSE of an estimator is a measure of the expected error of the estimator. The units of RMSE are the same as the units of the estimator.

SAMPLE (amostra)

A sample is a collection of units from a population.

SAMPLING (amostragem)

Sampling is a process used in statistical analysis in which a predetermined number of observations are taken from a larger population; it helps to make statistical inferences about the population. The methodology used to sample from a larger population depends on the type of analysis being performed. It can be simple random, stratified, cluster, multistage or systematic. Below, definitions for some methods are given:

- **SIMPLE RANDOM SAMPLING (amostragem aleatória simples)**

  Simple random sampling is the basic sampling technique where we select a group of subjects (a sample) for study from a larger group (a population). Each individual is chosen entirely by chance and each member of the population has an equal chance of being included in the sample. Every possible sample of a given size has the same chance of selection; i.e. each member of the population is equally likely to be chosen at any stage in the sampling process.

- **STRATIFIED SAMPLING (amostragem estratificada)**

  In stratified random sampling the whole population is divided into subpopulations, called strata. A sample is selected using a random design within each stratum. Stratified sampling is usually applied to biological sampling of the landings and in scientific surveys.

- **MULTISTAGE SAMPLING (amostragem em multi-etapas)**

  Multistage sampling is a combination of the various methods previously mentioned. At each stage, there is a random selection of the sampling units, which can be elements or clusters.

(http://www.fao.org/docrep/009/a0198e/A0198E00.htm)

SAMPLING DESIGN (delineamento amostral)

The sampling design of a scientific survey refers to the statistical techniques and methods adopted for selecting a sample and obtaining estimates of the survey variables from the selected sample. The sample design provides information on the target and final sample sizes, strata definitions and the sample selection methodology. (https://stats.oecd.org/glossary/detail.asp?ID=3852)

SAMPLING EFFORT (Esforço de amostragem)

Number of samples to be collected. Sampling effort can be optimized to achieve the precision levels required for a certain estimate, taking into account the costs of sampling and the variance of the samples/strata.

SUM OF SQUARED ERRORS (soma de erros quadrados)

SSE is the sum of the squared differences between each observation and its group mean. It can be used as a measure of variation within a cluster. If all cases within a cluster are identical the SSE would then be equal to 0. Its formula is

$$SSE = \sum_{i=1}^{n}(x_i - \bar{x})^2$$

where n is the number of observations x_i is the value of the i^{th} observation and $\bar{x}$ is the mean of all the observations. (https://hlab.stanford.edu/brian/error_sum_of_squares.html).

<u>Other references used:</u>

EU Commission Decision (2010/93/EU) of 18 December 2009, adopting a multiannual Community programme for the collection, management and use of data in the fisheries sector for the period 2011-2013 (notified under document C(2009) 10121)

Cadima EL, Caramelo AM, Afonso-Dias M, Conte de Barros P, Tandstad MO, de Leiva Moreno JI. 2005. Sampling methods applied to fishing science: a manual. *FAO Fisheries Technical Paper*. No. 434. Rome, FAO. 88p. (http://www.fao.org/docrep/009/a0198e/A0198E00.htm)

Statistics Glossary (https://store.fmi.uni-sofia.bg/fmi/statist/education/statlib/glossary/index.htm)

https://www.stat.berkeley.edu/~stark/SticiGui/Text/gloss.htm

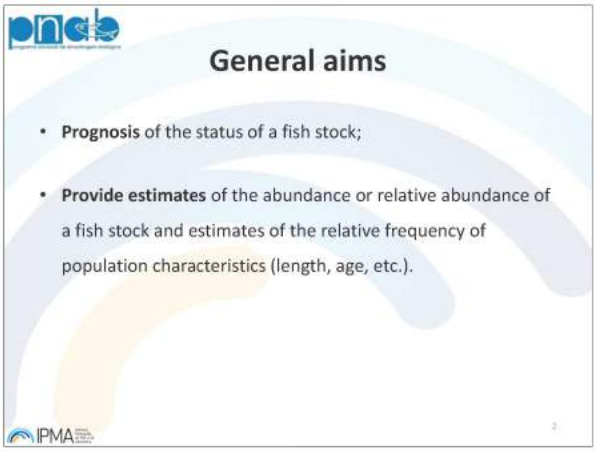
http:// http://ec.europa.eu/eurostat/statistics-explained/index.php/Thematic_glossaries

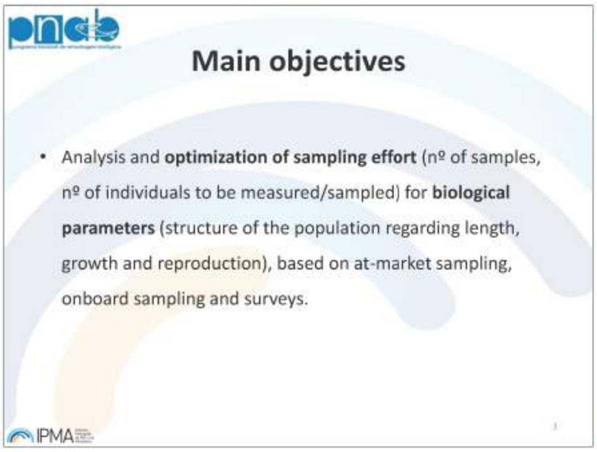
http://www.stats.gla.ac.uk/steps/glossary/sampling.html

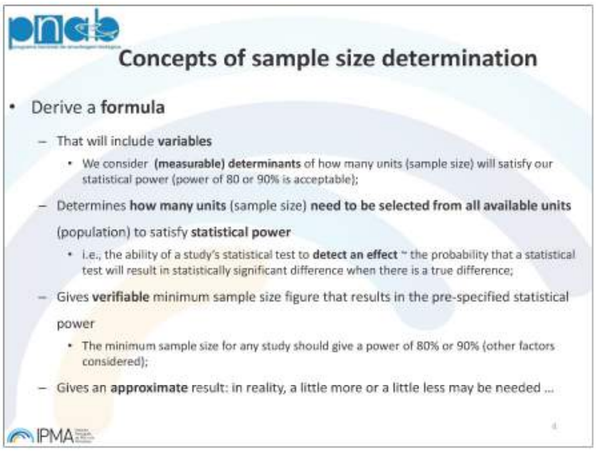
Annex 4: Presentations

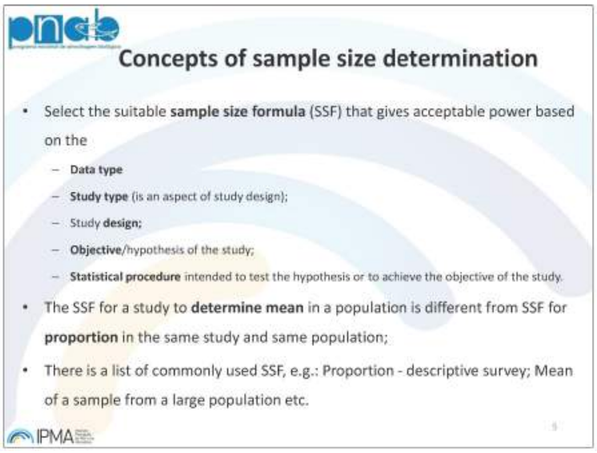
Annex 4: Presentations

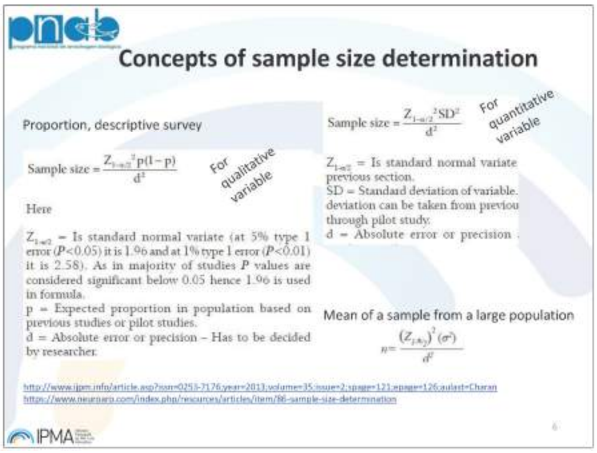
Case Study 1, P01 – Assessing the precision of length-frequency estimates (Following the paper of M. Pennington, 2002). M Oroszlányová.











Concepts of sample size determination

- Example:
  - Implicit hypothesis (i.e., not stated but implied):
    - Determine the mean length, age etc. of a population by calculating the mean length, age etc. of its sample;
  - Implicit null hypothesis:
    - There is no significant difference between the mean lengths, ages etc. of the sample and the respective population.

Concepts of sampling

- Prognosis accuracy(survey based assessment) > Prognosis accuracy(catch based assessment)
- Advantage(survey based assessment):
  - Uncertainties associated with survey estimates can be studied and quantified;
  - Based on such research, survey methods and stock assessments can be improved.

Concepts of sampling

- **Fish caught together tend to have more similar characteristics** (length, age etc.), **than those in the entire/general population**.
- Therefore, fish collected together/in **clusters** will contain **less information** about the population length (age etc.) distribution than fish sampled randomly.
- The **effective sample size** for estimates of the frequency distribution of a population characteristic (length, age etc.) can, therefore, be **much smaller** than the number of fish sampled.

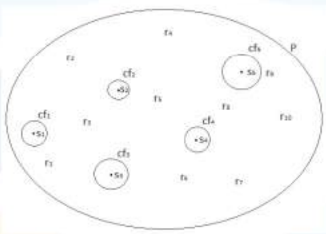
Concepts of sampling

- **Effective sample size:**
  - "Nº of fish that one would need to sample at random to obtain the **same information** on length contained in the cluster samples", i.e., the sample mean would have the same **precision** as an estimate based on a sample of *n* clusters.

Concepts of sampling

- **Precision:**
  - How concentrated are the values around the mean;
  - The reciprocal of variance (small variance → high precision);
  - Precision(random sample mean) = Precision(sample mean/clusters).
- In surveys, reduced haul duration can increase the number of stations and the number of sampled individuals, and so the **effective sample size**, hence the **precision** of length (age etc.) frequency estimates will **increase**.

Concepts of sampling



P – population (entire)
s – stations (survey) (at market: trips)
cfi – cluster of fish (caught)
ri – random fish

~ - similarity
F – fish
c – cluster
r – random
I – information
E – estimate (length, age, ...)

ess(E)

$$\sim(F_c \in cf_i) > \sim(F_r \in P) \Rightarrow I(\#F_c(E) \in cf_i) < I(\#F_r(E) \in P) \Rightarrow \#F_r(E) \ll \#F_c(E)$$

$$\downarrow duration(haul) \rightarrow \uparrow stations, \uparrow samples \rightarrow \uparrow ess(E) \Rightarrow \uparrow precision(E)$$

Concepts of sampling

Cluster sampling

- A sample of n clusters;
- In surveys, one cluster from each station;
- At-market, homogeneous landed size categories per trip.

Non-cluster sampling

- No clusters;
- In surveys, other sampling schemes at a station;
- At-market, non-homogeneous landed size categories per trip.

IPMA

Estimation of effective sample size

i. Estimate the **population mean fish length** (ratio estimator) (1) and its **variance** (2) based on the clusters of fish caught at n stations.

ii. Estimate the **variance of the population length distribution**

- **Cluster sampling** (3), i.e., if fish are randomly selected at each station (or if all fish are measured);
- **Non-cluster sampling** (4), i.e., for other sampling schemes at a station.

iii. **Effective sample size**

- **Cluster sampling:** (3)/(2);
- **Non-cluster sampling:** (4)/(2).

IPMA

Estimation of effective sample size

(1) $$\hat{R}=\frac{\sum_{i=1}^{n} M_i\hat{\mu}_i}{\sum_{i=1}^{n} M_i}$$

Where M_i = the number of fish caught (either actual or estimated) at station i; and $\hat{\mu}_i$ = an estimate of the average length of fish at station i.

(2) $$\mathrm{var}(\hat{R})=\sum_{i=1}^{n}\frac{(M_i/\overline{M})^2(\hat{\mu}_i-\hat{R})^2}{n(n-1)}$$

Where $\overline{M}=\sum_{i=1}^{n} M_i/n$

(3) $$\hat{\sigma}_x^2=\frac{\sum_{i=1}^{n}\sum_{j=1}^{m_i}(M_i/m_i)(x_{i,j}-\hat{R})^2}{M-1}$$

Where $M=\sum_{i=1}^{n} M_i$ is the total number of fish caught during the survey; and $x_{i,j}$ = the length of the j^{th} fish at station i.

(4) $$\hat{\sigma}_x^2=\frac{\sum_{k=1}^{L} f_k(y'_k-\hat{R})^2}{M-1}$$

Where f_k = the frequency of fish in the k^{th} length bin; and y'_k = the bin's midpoint.

IPMA

Estimation of effective sample size

(5) $$\frac{\hat{\sigma}_x^2}{\hat{m}_{eff}}=\mathrm{var}(\hat{R})$$

An estimate of the effective sample size can be derived by substituting the estimates from (2) and (3) or (4).

(6) $$deff=\frac{\mathrm{var}(\hat{R})}{\sigma_x^2/m}$$

The effective sample size is related to Kish's design effect (*deff*).

(7) $$m_{eff}=m/deff$$

Therefore, the effective sample size can be written in this form.

$$m_{eff}=\frac{m}{deff}=\frac{m}{\frac{var(\hat{R})}{\sigma_x^2/m}}=\frac{m*\frac{\sigma_x^2}{m}}{var(\hat{R})}=\frac{\sigma_x^2}{var(\hat{R})}\;\Rightarrow\;\hat{m}_{eff}=\frac{\hat{\sigma}_x^2}{var(\hat{R})}$$

IPMA

Example – Cluster sampling

	station	num_fish_caught	av_length	num_fish_random
	1	100	23.14	100
	2	50	22.80	50
	3	80	23.12	70
	4	100	23.37	80
	5	70	22.70	50
sum		400		350

Year	Species	Nº stations	Nº fish (total)	Sampled fish	R_hat (cm)	var_R_hat	sig_1	ess_1	ess_1/Nº stations	(ess_1/sampled fish)*100%
2016	SP1	5	400	350	23.1	0.01	3.74	265.2	53	75.8 %
	...									

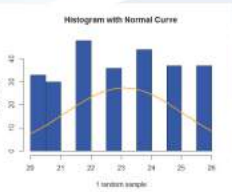
ess_1= sig_1/var_R_hat

1) During the 2016 autumn survey 400 fish (e.g. HKE) were caught (in 5 stations), 350 were measured(sampled), and the effective sample size was 265.2 fish or 75.8% of the total number measured.
2) If we randomly select 265 fish (out of 400) and calculate the average length, it should be 23.1 cm.
3) We can repeat the process of random selection (e.g. 30 times) to see the distribution.

IPMA

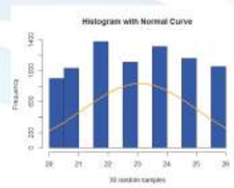
Example – Cluster sampling

Empirical distribution



Mean = 23.083 cm

1 random sample (265 fish out of 400)



Mean = 23.082 cm

30 random samples (30x265 fish out of 400)

IPMA

Next steps

- Develop the **methodology** and **scripts** for a case-study, and apply them to other species…
- Analyze **costs** …

Example 1
Input - Format

5 stations, 400 fish caught, 350 fish sampled, given average length per station (of all fish caught)

- **Input_1** (Excel or CSV):
  - List (Nº) of stations;
  - Nº of fish caught in each station;
  - Average length of (all) fish per station (in cm);
  - Nº of measured (random sample or all fish are measured?) fish per station.
- **Input_2** (Excel or CSV):
  - Length (in cm) of each measured (individual) fish in every station.

	station	num_fish_caught	av_length	num_fish_random
	1	100	23.14	100
	2	50	22.80	50
	3	80	23.12	70
	4	100	23.37	80
	5	70	22.70	50
sum		400		350

station	rand_length
1	23
1	21
1	24
1	23
1	24
1	23
1	25
1	24
1	22
…	…

```
inp_1 <- read.csv(file="input_1.csv",head=T,sep=",") # example input files – for only 1 species!
inp_2 <- read.csv(file="input_2.csv",head=T,sep=",")
```


Example 1
R-codes

library(dplyr)

1. Ratio estimator of mean length (population mean fish length in cm): **R_hat**

```
inp_1 <- inp_1 %>% group_by(station, num_fish_caught, av_length, num_fish_random) %>%
summarize(Mm=(num_fish_caught*av_length))
R_hat <- sum(inp_1$Mm)/sum(inp_1$num_fish_caught)   # expression (1) of the reference; (23.1)
```

2. Estimated variance of R_hat

```
var_R_hat <- sum(((inp_1$num_fish_caught/(sum(inp_1$num_fish_caught)/length(inp_1$station)))^2*(inp_1$av_length-
R_hat)^2)/(length(inp_1$station)*(length(inp_1$station)-1))) # expression (2) of the reference; (0.01)
```

3. Estimated variance of the population length distribution: sig_1
(Sampling scheme: fish are randomly selected at each station or all fish are measured)

```
inp_2$R_hat <- rep(R_hat, each = length(inp_2$rand_length))
inp_2 <- inp_2 %>% group_by(station, rand_length, R_hat) %>% mutate(x=((rand_length - R_hat)^2))
inp_2a <- inp_2 %>% group_by(station) %>% summarize(sumx=(sum(x)))
sig_1 <- sum(inp_1$num_fish_caught/inp_1$num_fish_random*(inp_2a$sumx))/(sum(inp_1$num_fish_caught)-1) #
expression (3) of the reference; (3.74)
```

4. **Effective sample size** (ess)

```
ess_1 <- sig_1/var_R_hat    # expression (5) of the reference; (265.2)
```

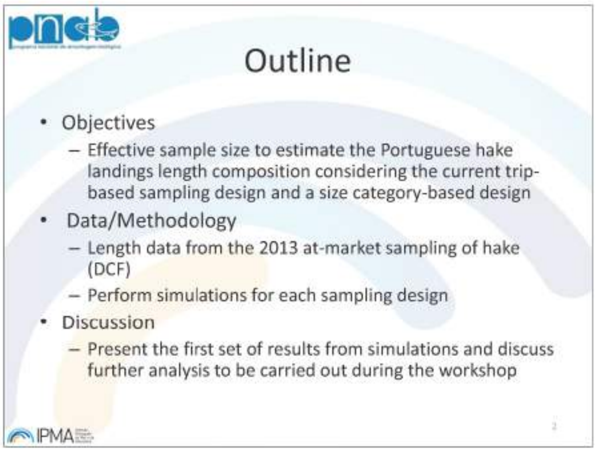

Example 1
R-codes

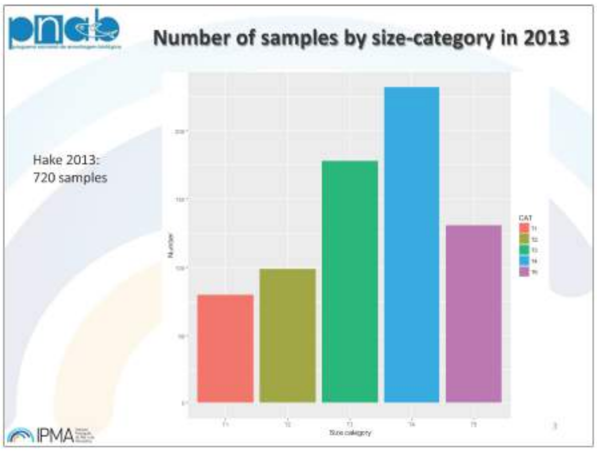
5. **Empirical distribution** (randomly select 265x (out of 400) and calculate the average length (should be 23.1 cm); repeat 30 x to see the distribution)

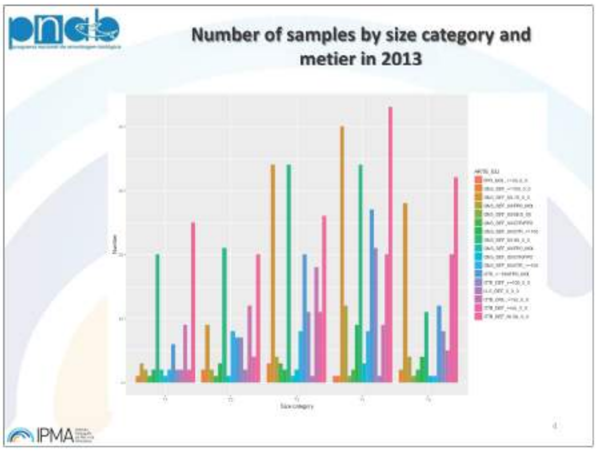
```
set.seed(100)
s1 <- sample(20:26, 100, replace = T)
s2 <- sample(20:26,  50, replace = T)
s3 <- sample(20:26,  80, replace = T)
s4 <- sample(20:26, 100, replace = T)
s5 <- sample(20:26,  70, replace = T)

S <- c(s1, s2, s3, s4, s5)
```

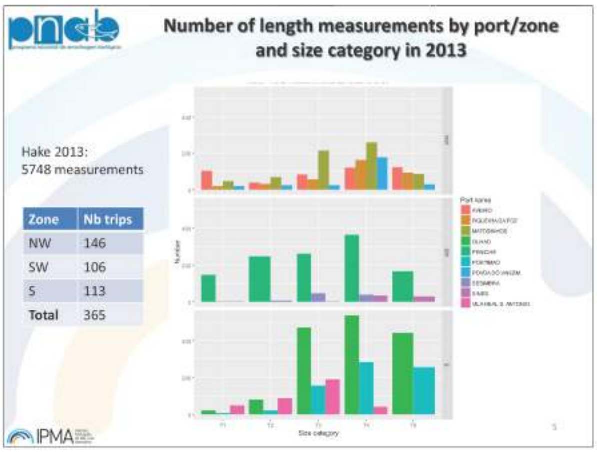
```
set.seed(100)
rs1 <- sample(S,265,replace=F)
hist(rs1)
mean(rs1)  (23.08302)

set.seed(100)
rs30 <- replicate(30, sample(S,265,replace=F))
hist(rs30)
mean(rs30) (23.08201)
```

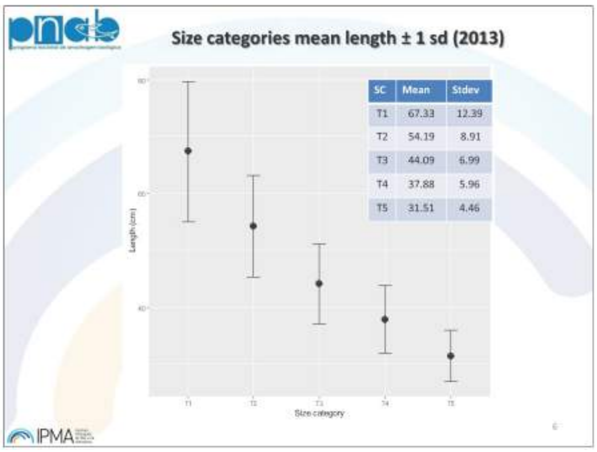
P02 – At-market sampling for landings length composition: hake commercial size categories case-study. M. Azevedo, C. Silva.

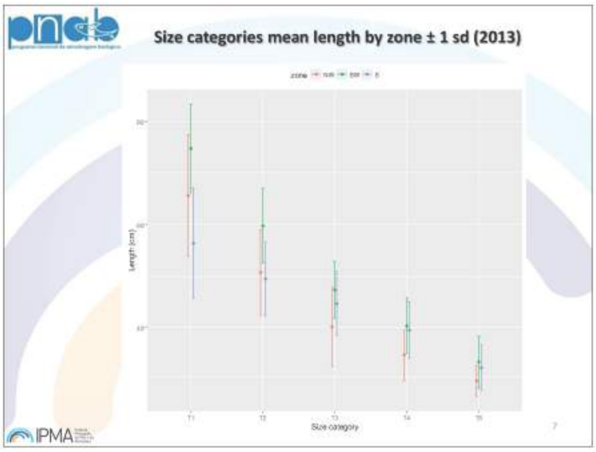


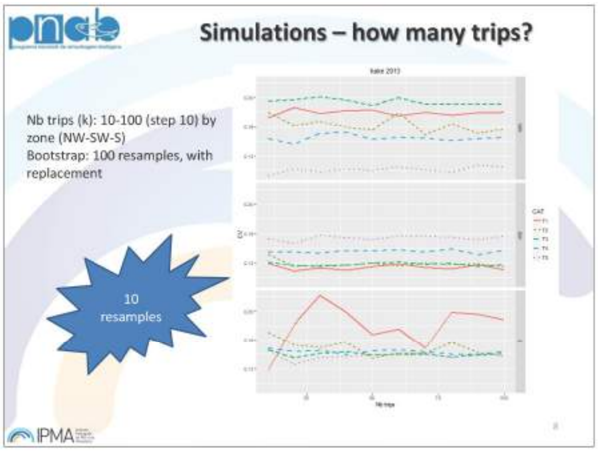


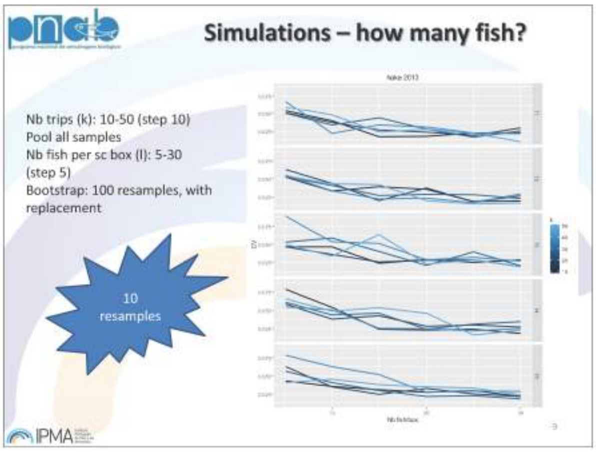




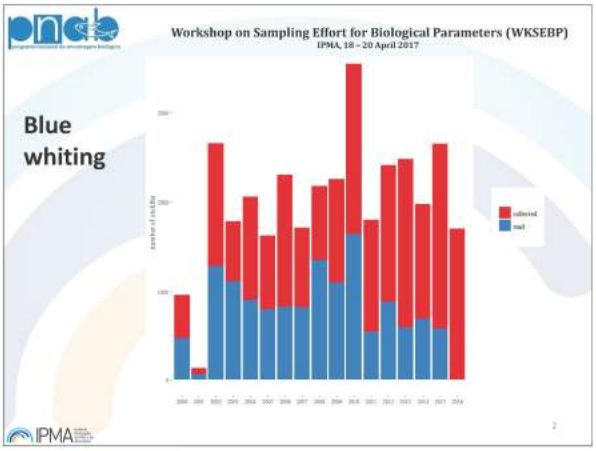


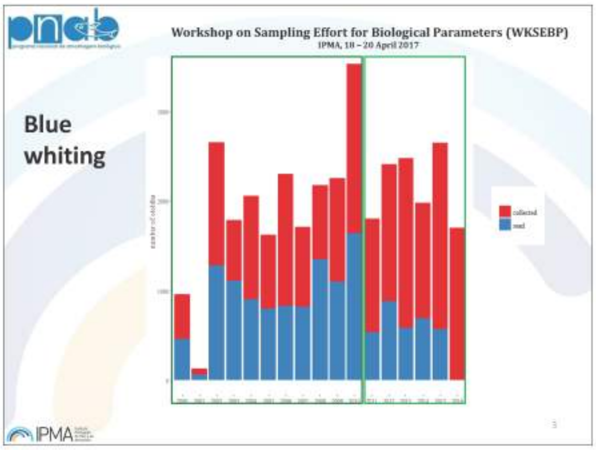

pnab
Size categories mean length by zone ± 1 sd (2013)
Length (cm)
T1
T2
T3
T4
T5
Size category
IPMA
7

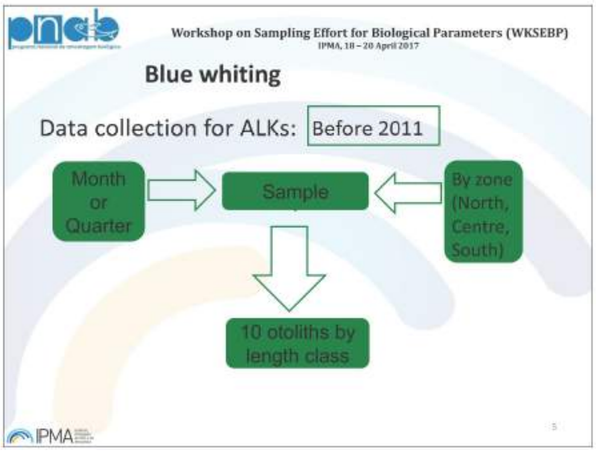

pnab
Simulations – how many trips?
Nb trips (k): 10-100 (step 10) by zone (NW-SW-S)
Bootstrap: 100 resamples, with replacement
10 resamples
hake 2013
CV
CAT
Nb trips
IPMA
8

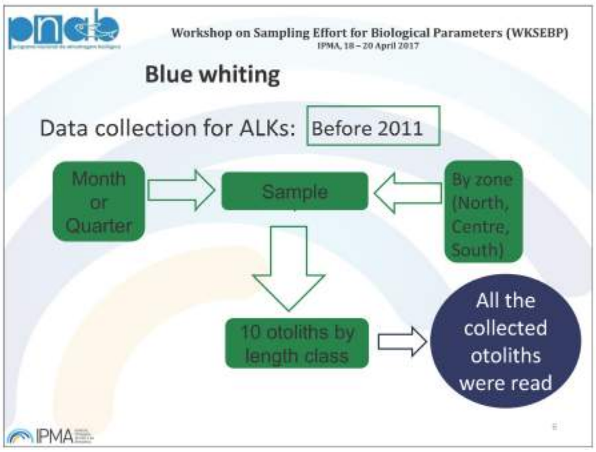

pnab
Simulations – how many fish?
Nb trips (k): 10-50 (step 10)
Pool all samples
Nb fish per sc box (l): 5-30 (step 5)
Bootstrap: 100 resamples, with replacement
10 resamples
hake 2013
CV
Nb fish/box
IPMA
9

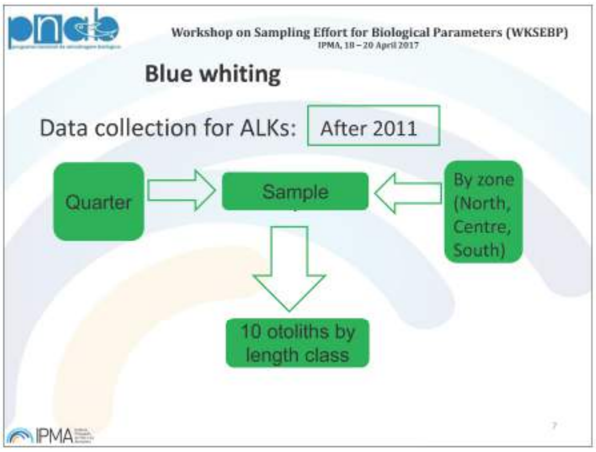
P03 – Sampling for ALKs – blue whiting case-study. Patrícia Gonçalves.

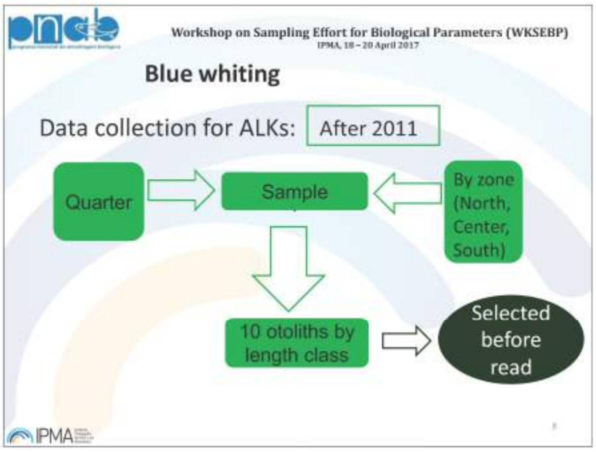


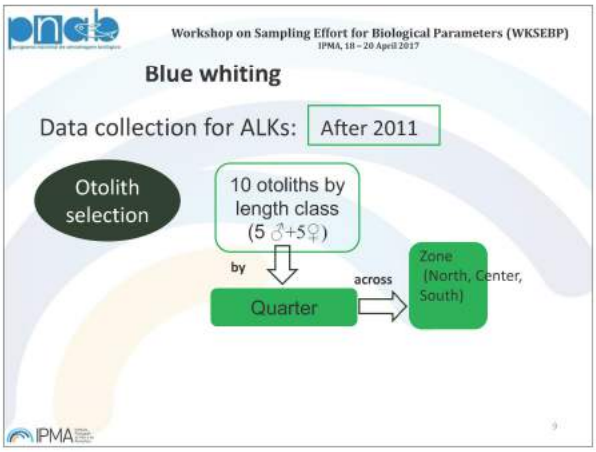


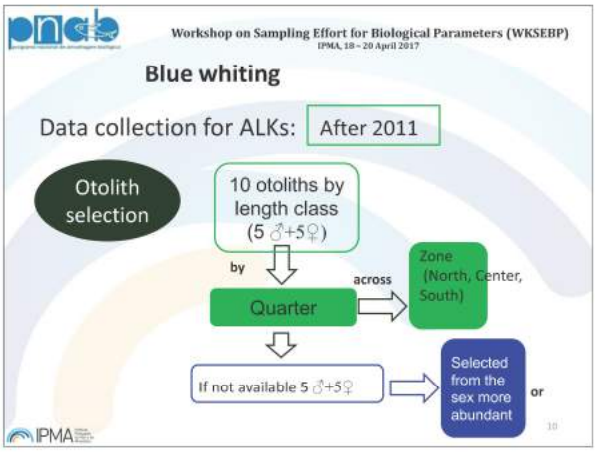


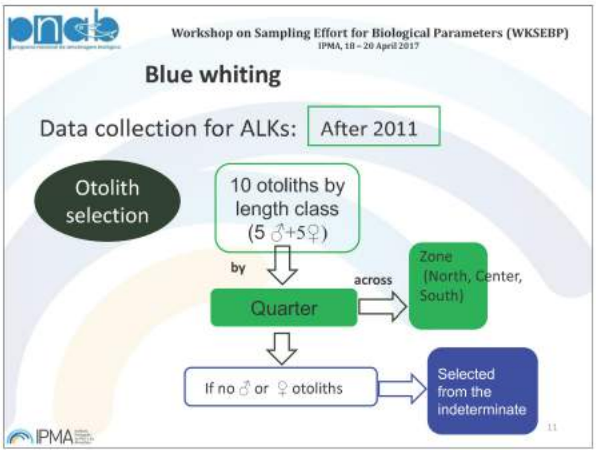


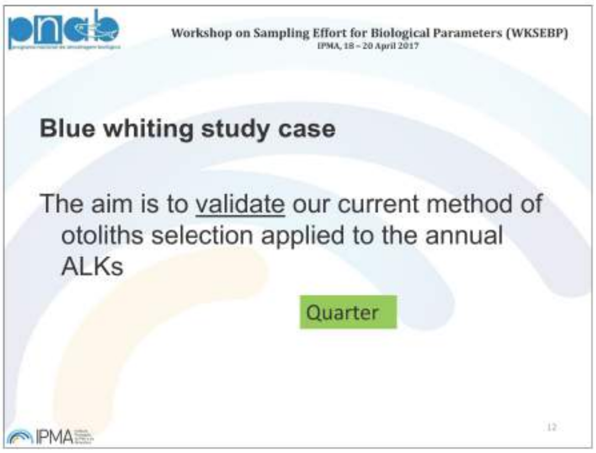

Workshop on Sampling Effort for Biological Parameters (WKSEBP)
IPMA, 18 – 20 April 2017
Blue whiting
Data collection for ALKs:
After 2011
Quarter
Sample
By zone (North, Centre, South)
10 otoliths by length class
IPMA
7


Workshop on Sampling Effort for Biological Parameters (WKSEBP)
IPMA, 18 – 20 April 2017
Blue whiting
Data collection for ALKs:
After 2011
Quarter
Sample
By zone (North, Center, South)
10 otoliths by length class
Selected before read
IPMA
8


Workshop on Sampling Effort for Biological Parameters (WKSEBP)
IPMA, 18 – 20 April 2017
Blue whiting
Data collection for ALKs:
After 2011
Otolith selection
10 otoliths by length class (5 ♂+5♀)
by
Quarter
across
Zone (North, Center, South)
IPMA
9


Workshop on Sampling Effort for Biological Parameters (WKSEBP)
IPMA, 18 – 20 April 2017
Blue whiting
Data collection for ALKs:
After 2011
Otolith selection
10 otoliths by length class (5 ♂+5♀)
by
Quarter
across
Zone (North, Center, South)
If not available 5 ♂+5♀
Selected from the sex more abundant
or
IPMA
10


Workshop on Sampling Effort for Biological Parameters (WKSEBP)
IPMA, 18 – 20 April 2017
Blue whiting
Data collection for ALKs:
After 2011
Otolith selection
10 otoliths by length class (5 ♂+5♀)
by
Quarter
across
Zone (North, Center, South)
If no ♂ or ♀ otoliths
Selected from the indeterminate
IPMA
11


Workshop on Sampling Effort for Biological Parameters (WKSEBP)
IPMA, 18 – 20 April 2017
Blue whiting study case
The aim is to validate our current method of otoliths selection applied to the annual ALKs
Quarter
IPMA
12

Blue whiting study case

Comparing the ALKs produced if:

a) All the otoliths collected were readed
b) Selection by quarter (10 otoliths/length class)
c) Selection by semester (10 otoliths/length class)
d) Selection by year (10 otoliths/length class)

Blue whiting study case

Years – 2004 and 2005
Length class - bewteen 15 and 34cm (2004)
between 17 and 29cm (2005)
Sex-ratio - 1:1 (same number males and females)
Number of otoliths – 10 otoliths by length class
(5♀+5♂)
Zone (N, C, S) – uniformly distributed

Blue whiting study case

Years – 2004 and 2005
Length class - bewteen 15 and 34cm (2004)
between 17 and 29cm (2005)
Sex-ratio - 1:1 (same number males and females)
Number of otoliths – 10 otoliths by length class
(5♀+5♂)
Zone (N, C, S) – uniformly distributed

by

Quarter Semester Annual

Results

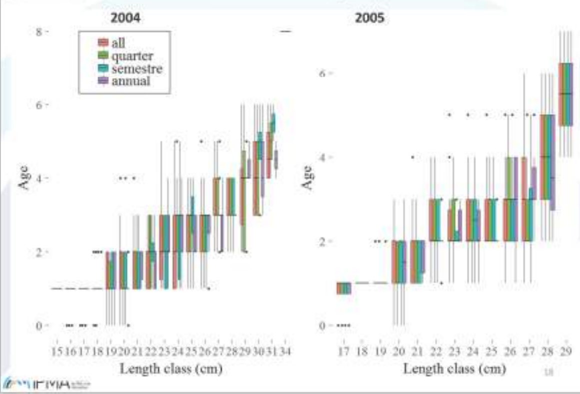
Number of otoliths read (total)

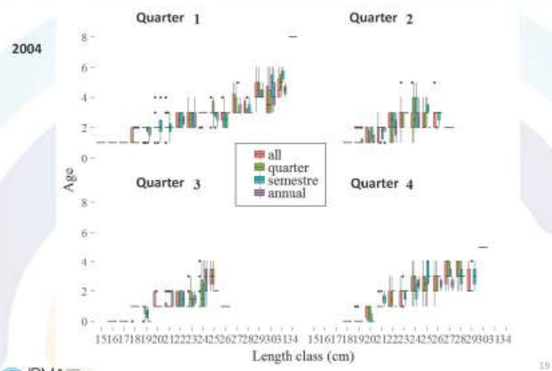
	All	Quarter	Semester	Annual
2004	907	347	198	123
2005	800	275	185	114

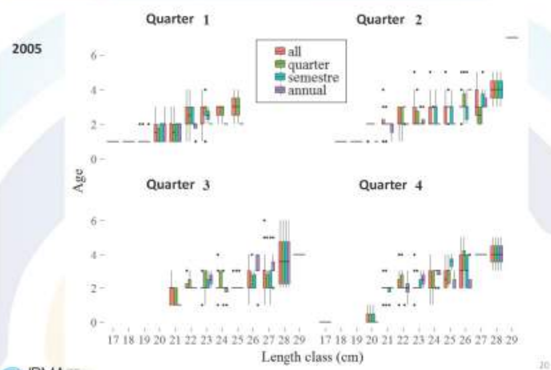
Results

Number of otoliths read (total)

	All	Quarter	Semester	Annual
		38%	22%	14%
2004	907	347	198	123
		34%	23%	14%
2005	800	275	185	114







Results

Mean age by length class (cm) #2004

LT	all	quarter	semester	annual
15	1.0000000	1.0000000	1.000000	1.000000
16	0.9583333	0.9090909	0.875000	1.000000
17	0.9629630	0.9090909	0.875000	1.000000
18	1.1509434	1.0000000	1.076923	1.100000
19	1.2638889	1.0454545	1.076923	1.111111
20	1.5851064	1.2647059	1.437500	1.100000
21	1.6967213	1.5945946	1.500000	1.700000
22	1.9256198	1.9210526	2.000000	1.800000
23	2.1000000	2.1351351	2.050000	2.400000
24	2.1511628	2.6176471	2.500000	2.700000
25	2.6666667	2.7777778	2.909091	2.428571
26	2.7708333	2.6818182	2.545455	2.571429
27	3.2758621	3.3529412	3.100000	2.714286
28	3.2777778	3.2222222	3.200000	3.000000
29	3.7500000	3.6000000	3.800000	4.333333
30	4.1428571	4.2000000	4.750000	4.333333
31	4.7500000	5.0000000	5.500000	4.500000
34	8.0000000	8.0000000	8.000000	8.000000

Mean age by length class (cm) #2005

LT	all	quarter	semester	annual
17	0.750000	0.750000	0.750000	0.75
18	1.000000	1.000000	1.000000	1.00
19	1.219512	1.090909	1.000000	1.10
20	1.528302	1.333333	1.384615	1.50
21	1.888889	1.766667	1.666667	1.70
22	2.274510	2.378378	2.250000	1.90
23	2.238095	2.378378	2.250000	2.30
24	2.341270	2.333333	2.550000	2.20
25	2.480392	2.687500	2.550000	2.10
26	2.804878	3.107143	2.800000	2.90
27	3.056604	2.571429	2.950000	3.30
28	3.818182	3.818182	3.818182	3.75
29	5.500000	5.500000	5.500000	5.50

Results

Mean age by length class (cm) #2004

LT	all	quarter	semester	annual
15	1.0000000	1.0000000	1.000000	1.000000
16	0.9583333	0.9090909	0.875000	1.000000
17	0.9629630	0.9090909	0.875000	1.000000
18	1.1509434	1.0000000	1.076923	1.100000
19	1.2638889	1.0454545	1.076923	1.111111
20	1.5851064	1.2647059	1.437500	1.100000
21	1.6967213	1.5945946	1.500000	1.700000
22	1.9256198	1.9210526	2.000000	1.800000
23	2.1000000	2.1351351	2.050000	2.400000
24	2.1511628	2.6176471	2.500000	2.700000
25	2.6666667	2.7777778	2.909091	2.428571
26	2.7708333	2.6818182	2.545455	2.571429
27	3.2758621	3.3529412	3.100000	2.714286
28	3.2777778	3.2222222	3.200000	3.000000
29	3.7500000	3.6000000	3.800000	4.333333
30	4.1428571	4.2000000	4.750000	4.333333
31	4.7500000	5.0000000	5.500000	4.500000
34	8.0000000	8.0000000	8.000000	8.000000

No significant differences

Mean age by length class (cm) #2005

LT	all	quarter	semester	annual
17	0.750000	0.750000	0.750000	0.75
18	1.000000	1.000000	1.000000	1.00
19	1.219512	1.090909	1.000000	1.10
20	1.528302	1.333333	1.384615	1.50
21	1.888889	1.766667	1.666667	1.70
22	2.274510	2.378378	2.250000	1.90
23	2.238095	2.378378	2.250000	2.30
24	2.341270	2.333333	2.550000	2.20
25	2.480392	2.687500	2.550000	2.10
26	2.804878	3.107143	2.800000	2.90
27	3.056604	2.571429	2.950000	3.30
28	3.818182	3.818182	3.818182	3.75
29	5.500000	5.500000	5.500000	5.50

No significant differences

Results

Mean age by length class (cm) #2004

LT	all	quarter	semester	annual
15	1.0000000	1.0000000	1.000000	1.000000
16	0.9583333	0.9090909	0.875000	1.000000
17	0.9629630	0.9090909	0.875000	1.000000
18	1.1509434	1.0000000	1.076923	1.100000
19	1.2638889	1.0454545	1.076923	1.111111
20	1.5851064	1.2647059	1.437500	1.100000
21	1.6967213	1.5945946	1.500000	1.700000
22	1.9256198	1.9210526	2.000000	1.800000
23	2.1000000	2.1351351	2.050000	2.400000
24	2.1511628	2.6176471	2.500000	2.700000
25	2.6666667	2.7777778	2.909091	2.428571
26	2.7708333	2.6818182	2.545455	2.571429
27	3.2758621	3.3529412	3.100000	2.714286
28	3.2777778	3.2222222	3.200000	3.000000
29	3.7500000	3.6000000	3.800000	4.333333
30	4.1428571	4.2000000	4.750000	4.333333
31	4.7500000	5.0000000	5.500000	4.500000
34	8.0000000	8.0000000	8.000000	8.000000

No significant differences

Mean age by length class (cm) #2005

LT	all	quarter	semester	annual
17	0.750000	0.750000	0.750000	0.75
18	1.000000	1.000000	1.000000	1.00
19	1.219512	1.090909	1.000000	1.10
20	1.528302	1.333333	1.384615	1.50
21	1.888889	1.766667	1.666667	1.70
22	2.274510	2.378378	2.250000	1.90
23	2.238095	2.378378	2.250000	2.30
24	2.341270	2.333333	2.550000	2.20
25	2.480392	2.687500	2.550000	2.10
26	2.804878	3.107143	2.800000	2.90
27	3.056604	2.571429	2.950000	3.30
28	3.818182	3.818182	3.818182	3.75
29	5.500000	5.500000	5.500000	5.50

No significant differences

? Error
? Precision

Blue whiting study case

Conclusions

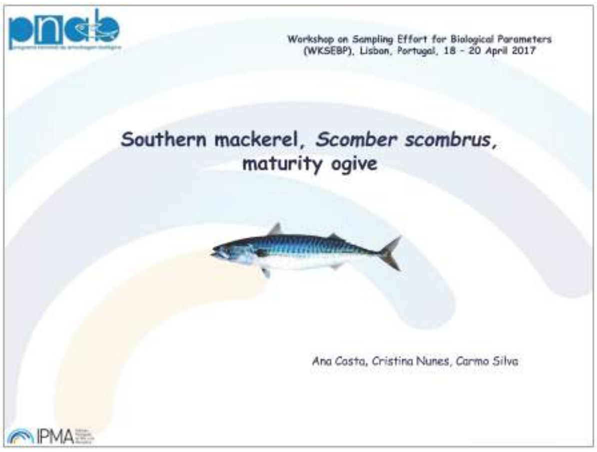
-no significant differences were observed between the ALKs using all the otoliths collected or just using the selected by quarter (10 otoliths per length class per sex);

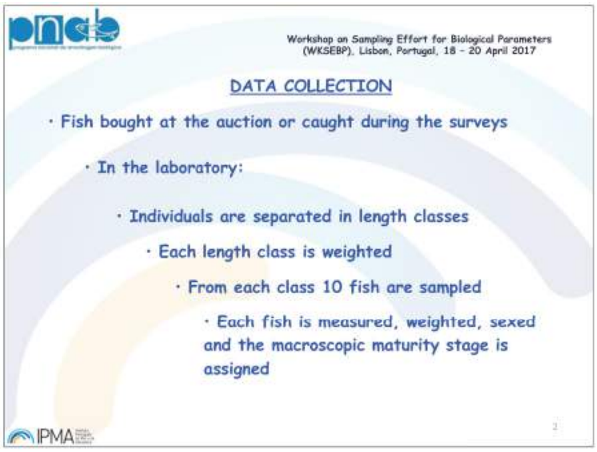
- these results validate the current otoliths selection method for BW.

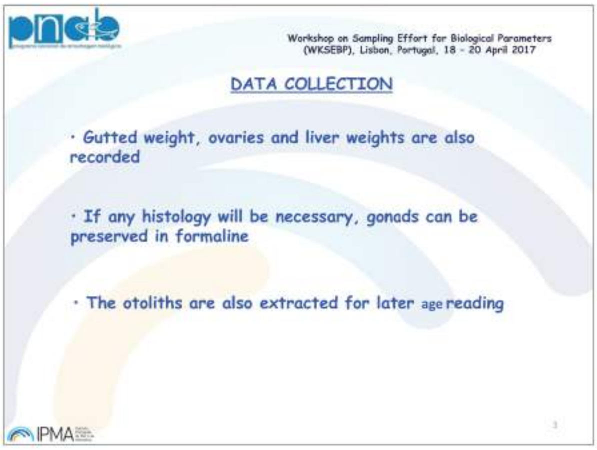
Next steps

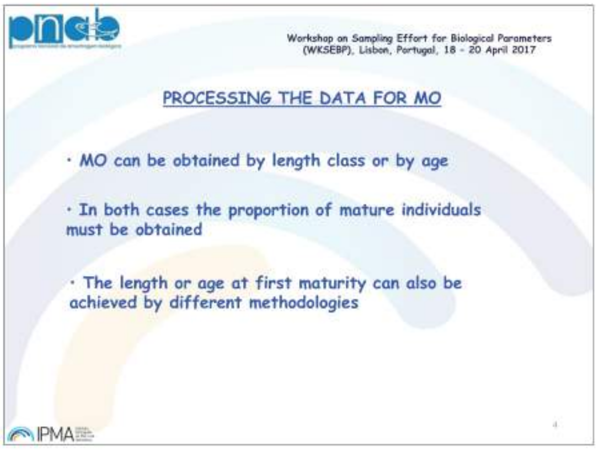
Determine precision and error….

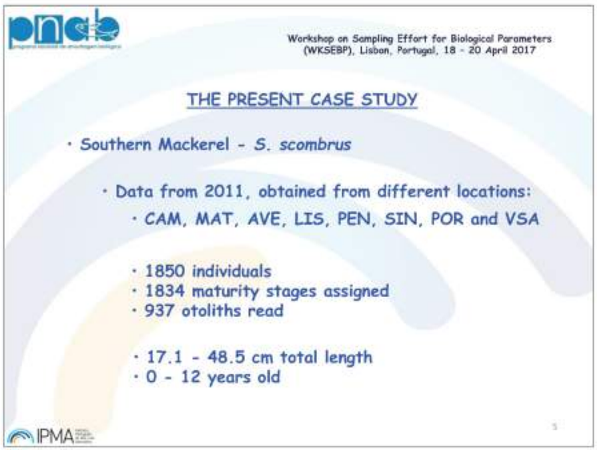
P04 – Southern mackerel, Scomber scombrus, mature ogive. A. Costa, C. Nunes, M. C. Silva.

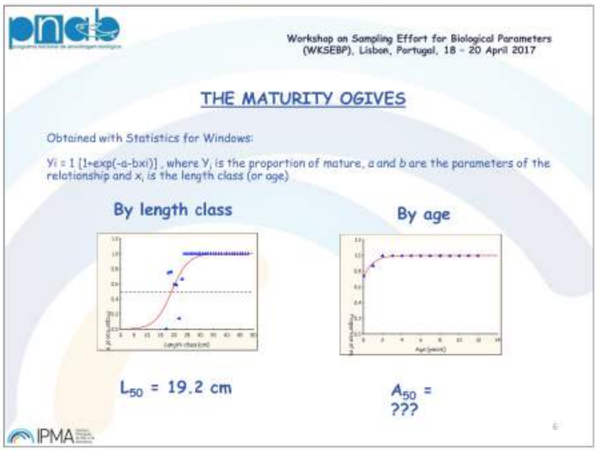












IMPORTANT ASPECTS TO KEEP IN MIND

- Data must be obtained during the spawning season in order to collect mature and imature individuals

- The maturity stages must be well described in the maturity scale of the species, so that there is no doubt in the assignement of each stage, in particular the identification of the imature individuals

- All length classes must be well covered, especially the ones of the middle part of the curve

MINISTÉRIO DA AGRICULTURA,
E DO MAR

IPMA
Instituto Português
do Mar e da Atmosfera