ICES SGPIDS REPORT 2012

ICES ADVISORY COMMITTEE

ICES CM 2012/ACOM:51

REF. PGCCDBS

Report of the Study Group on Practical Implementation on Discard Sampling Plans

18–22 June 2012

ICES HQ, Copenhagen, Denmark

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

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

H. C. Andersens Boulevard 44–46
DK–1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

Recommended format for purposes of citation:

ICES. 2012. Report of the Study Group on Practical Implementation on Discard Sampling Plans, 18–22 June 2012, ICES HQ, Copenhagen, Denmark. ICES CM 2012/ACOM:51. 87 pp.

Contents

Executive summary

The Study Group on Practical Implementation of Discard Sampling Plans (SGPIDS; chaired by Edwin van Helmond, The Netherlands) met 18 June – 22 June 2012 in Copenhagen, Denmark. Sixteen participants representing 10 countries were present at the meeting, including the chair, Bram Couperus, of ICES WGBYC (Working Group on Bycatch of Protected Species). SGPIDS was proposed by ICES PGCCDBS (2010) in response to a request from the Regional Coordination Meeting for the North Sea and Eastern Arctic (RCM NS&EA; 2010) to foster the exchange of experience and expertise between experts on discard sampling, planning and implementation of PGCCDBS recommendations and ultimately synchronize coordination and data collection procedures of discard sampling between members states.

To address the terms of reference more efficiently the group split up into subgroups. The group regularly meet in plenary to discuss and synchronize output and results produced during subgroup sessions.

During the first meeting in 2011, the study group identified potential sources of bias within discard sampling programmes. Bias in vessel selection and sampling effort allocation were reported to be common to all national sampling programmes. In the attempt to improve data quality and reduce bias, the study group provides the practical tools to implement unbiased sampling frames, random vessel selection procedures and data quality indicators.

Since cooperation between member states for at–sea sampling schemes is strongly promoted (PGCCDBS 2012), the study group started to design its data collection framework at a regional level. This regional sampling frame is divided into national strata. These national strata could be further divided into sampling strata. Fundamental to these schemes is a move away from national ad hoc, quota based sampling schemes, where a sampling event was conditional on where and how a vessel has fished or was going to fish, to a random vessel selection process, where the sampling frame describes the population (of vessels) to be sampled. Besides the concept of randomized, unbiased sampling frames, the study group drafted a rough guide on how to implement such sampling scheme based on the experience of members states currently adopting the approach. Three steps are identified in the process of implementation: 1) Defining the frame and creating vessel drawlists. 2) Selecting a random vessel. 3) Monitoring and recording the selection process.

To synchronize data collection procedures of discard sampling between member states, the study group completed its first step in 2011: To make an inventory of different discard sampling methods used across Europe. In 2012, areas of overlap, where the same, or very similar fleets are sampled by different member states using different sampling methods, are evaluated and assessed. Based on three case studies the study group concluded that differences in methods are caused by a mixture of restraints, i.e. logistic, financial, cultural, historical or practical. Also dissimilar research objectives were identified as an important bottle neck for synchronizing sampling methods between member states.

To be able to evaluate and compare performance levels of (national) sampling programmes, one reporting standard is essential. The study group agreed to further develop standardized reporting during the next meeting in 2013.

Equal to 2011, the study group provided an updated summary of the current sampling programmes in European waters. Compared to 2011, the study group was able to include additional countries such as Germany and Poland.

1 Introduction

The results of discard sampling programmes play an increasing role in stock assessments and fisheries management. The quality of the discard data as well as uniformity of the data between countries play a vital role in the usability of these data. The Study Group on Practical Implementation of Discard Sampling plans (SGPIDS) is essential to allow standardization and harmonization of discard sampling plans and to provide a platform for the exchange of expertise on discard sampling practices for the next years.

1.1 Supporting information

Priority:	**Essential**
Scientific justification:	Currently all Member States collect data of discard pratices under the Data Colletion Framework (DCF) of the European Commision. This DCF sets out precision levels by métier which need to be met by the different members states. Generally resources available and other practical constraints limit the number of samples and, conseqently, precision leversl are not met. SGPIDS notes that in order to meet the precision level requirements members states unwillingly bias their sampling programmes, i.e. to collect data ot the higest possible numbers of trips, institutes only collaborate with skippers who are willing to take observers on board. To examine whether the precision requirements of the programme are met, SGPIDS suggest a different approach. An approach with focus on the quality of the sampling programmes itself (representative sampling), rather then excessively increasing sampling levels just to meet (unrealistic) presision levels. In pursuit of standardized discard sampling between counties it is important that practical differences between programmes and possible improvements are identified. At within–trip level, it is important that bias and variability associated to different sampling protocols is investigated. Comparison of results of different methods used eventually lead to the most appropriated sampling protocols in discard sampling on board commercial vessels of various fisheries. Potential sources of bias within sampling programmes were identified during the first meeting of the study group. Bias in vessel selection and sampling effort allocation are reported to be common to all national sampling programmes. Providing the practical tools to define appropriate sampling frames, vessel selection procedures and reporting programme outputs will contribute to reduction of bias and ultimately standardize discard sampling programmes between Member States.
Resource requirements:	Participants should bring descriptions of sampling procedures to the meeting. Reports of discard results on a national level. Additional resources required to undertake additional investigations regarding on board sampling techniques (i.e. age–length–keys, lengt–weight relations, discard data at haul level, ect.)
Participants:	Scientists managing discard sampling schemes or projects, either under or outside DCF, within European waters.
Secretariat facilities:	Scientists managing discard sampling schemes or projects, either under or outside DCF, within European waters.
Financial:	None
Linkages to advisory committees:	ACOM
Linkages to other committees or groups:	PGCCDBS, RCMs, WGBYC, WKPICS1.
Linkages to other organizations:	None

1.2 Terms of Reference

The **Study Group on Practical Implementation of Discard Sampling Plans** (SGPIDS), chaired by Edwin van Helmond, the Netherlands, will meet 18–22 June 2012 in ICES Headquarters, Copenhagen, Denmark to:

1) Develop and define quality standard levels for discard sampling programmes e.g. recording refusal rates, sampling coverage (spatial and temporal distribution), self–sampling validation procedures;

2) Identify appropriate on board sampling techniques; evaluate the effect of different on board sampling protocols (e.g. different usage of age–length–keys, sampling unsorted catch vs. landings and discard separately, sample size and raising procedures to haul level, usage of length–weight–relations, systematic sampling vs. census sampling, etc.);

3) Identify practical improvements to define sampling frames (i.e. based on effort/landings, etc.);

4) Develop statistically sound and practical tools to implement vessel selection procedures (including registration of refusal rates);

5) Develop standardize reporting of results of sampling designs (case studies: reports of discard results on a national level);

1.3 Adoption of the agenda and terms of reference

The adopted agenda of the Study Group on Practical Implementation of Discard Sampling plans is presented in Annex 2 of this report.

Due to time constraints, considering the other terms of reference and the number of participants, terms of reference (5) was only briefly addressed during one of the plenary sessions (see Annex 2 for agenda). Therefore, the study group recommends to review standardize reporting of results of sampling designs in more detail during the next SGPIDS meeting.

2 Quality standard levels for discard sampling programmes (ToR1)

During the first meeting, SGPIDS identified potential sources of bias within sampling programmes. Bias in vessel selection and sampling effort allocation were reported to be common to all national sampling programmes. Providing the practical tools to define appropriate sampling frames, vessel selection procedures will contribute to reduction of bias.

Instead of developing a kind of " traffic–light–indicator–score–card" about the performance of our sampling (WKACCU), the study group concentrated on improving data quality and reducing bias by creating a practical hand–out on how to set up randomized sampling schemes. In the description the focus is primarily on on–board observer sampling, but the suggested indicators would work equally good for self–sampling programs with a random selection of vessels.

2.1 Regional approach to at-sea sampling

Regional cooperation and task sharing have been strongly promoted within the present Data Collection Framework (DCF). The main way to support a regional approach has been through the regional coordination meetings (RCMs), established by the Commission. So far these meetings have been successful in collating national meta data on the performance of fisheries and sampling to a regional level. This has led to a more common understanding of fisheries and sampling within a region. It is foreseeable that the emphasis of a regional approach in data collection will be even stronger in the revised DCF since the reform of the Common Fisheries Policy will include greater regionalisation of fisheries management. Stocks and fisheries are further usually assessed at a regional/international level. Regional data collection programmes need to be documented and based on statistically sound survey methods to allow for quality control and transparency in the data collection–assessment–advice process. In a regional programme the countries are considered as strata. Sampling strata are then defined at the national level (Figure2.1).

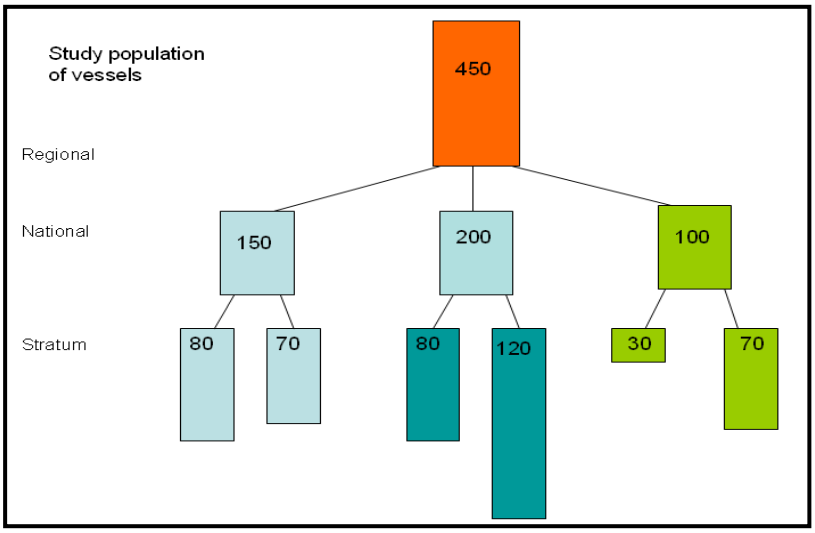


Figure 2.1. A schematic overview on how a regional data collection programme could be designed. The study population is identified at the regional level. Countries are considered as strata. The population in each country strata is divided into sampling strata.

A regional data collection programme requires that the study population is identified at the regional level and that a regional sampling frame is established. The regional sampling frame is then divided into national sampling frames. The national sampling frames could further be divided into sampling strata if this is suitable. Quality indicators would be assessed for each sampling strata and add up through the hierarchy to the national and regional level with appropriate weighting.

2.2 Sampling terms: populations, study populations, stratification, national sampling plans and (simple) quality indicators

2.2.1 Total population and study population at Regional and national levels

Following the definitions outlined at (ICES, 2011) the total population at the regional level would comprise all the vessels operating in the region. Within that total regional population of vessels there is a smaller study population comprised of all the vessels in the regional list that were suitable for sampling.

The regional population would be obtained by combining the national lists of registered vessels. The study population is required to be defined first at the national level based on knowledge of the vessels and the sampling capabilities at the national level. Typically a large number of small vessels may be excluded at this point. The study population at the regional level is the combination of the study populations at the national level.

2.2.2 Sampling strata at the national level

The national study population of vessels would be divided into a number of mutually exclusive sampling frames. Each vessel from the national study population of vessels would be in one, and only one, sampling frame and that sampling frame is specific to a particular sampling stratum. The allocation of vessels to strata is typically done to produce more homogeneous sampling groups to reduce the variance in estimates (ICES 2011).

A vessel would remain within that stratum on the long term basis but the vessels lists would also require to be updated, probably on an annual basis, so that vessels new to the fishery could be added to a sampling stratum and decommissioned vessels could be removed from the stratum. It would generally not be the case that a vessel would be added or removed from a list because it happened to be fishing in a different location for a period of time (see Section 3.1 for further discussion on this point).

The fishing trips undertaken by the vessels in the stratum during a time period, typically a year, form the study population of trips. This is the primary sampling unit in which we are interested in sampling.

A temporal sampling stratification, which is often the quarter for a continuous fishery, but can be otherwise defined for a seasonal fishery, would typically be used to manage the sampling effort devoted to covering the fishery over the year.

The population of trips (if it were known in advance) could be envisaged as a grid with vessels in the sampling frame along one axis and time along the other axis (Figure 2.2). Vessel may or may not be fishing in a particular time, different vessels would have different number of trips over the course of the year and trip duration is also very often different both for a single vessel over time and between different vessels. Each cell in the grid can be considered as a potential sampling opportunity but

because of the unpredictable nature of fishing trips, means that the sampling population of trips is dynamic, sample sizes cannot be predicted in advance and which cells in the sampling grid actually have trips cannot be predicated either. How samples can be selected in this situation is discussed in Section 2.2.4).

2.2.3 Allocation of sampling effort

The national sampling plan would allocate a number of trips to be sampled in a particular stratum over the year, (possible broken down by quarter or over some other temporal stratification).

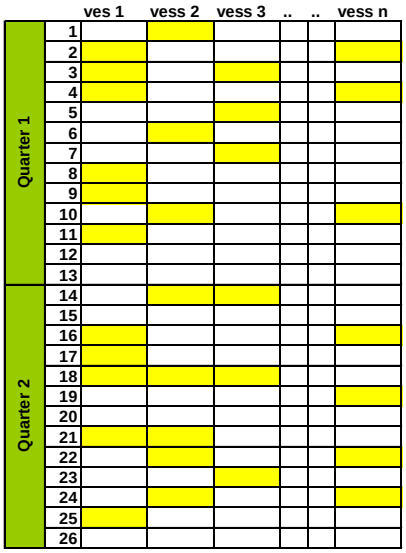


Figure 2.2. Example of a sampling lattice of vessels active over time, first 2 quarters of a year, trips identified as a potential sampling opportunity (yellow marked cells).

2.2.4 Selecting vessels and trips

The essential part of the selection of the vessel from the sampling frame is that the list is randomized prior to the selection, and then that a contact protocol (see below) is followed. The randomization is *the* key element in the selection process that ensures that if we operated in an ideal world, each vessel has an equal probability of selection and that the samples collected subsequently are representative of the study population in the frame. Representative samples lead to unbiased estimators of the population parameters of interest.

As we do not operate in an ideal world, documenting the contact process allows quantifying the extent to which our realized samples may differ from the ideal.

There are three basic approaches to set up a randomized list:

- Simple random sample without replacement
- Random sample with replacement
- Selection with probability proportionate to size (PPS)

Briefly, random sampling without replacement is where a member of the study population can be selected only once, random sampling with replacement allows a sampling unit to be selected any number of times. Both these methods ensure unbiased estimates of population parameters and allow the calculation of a sample variance though the estimation processes differ slightly. The practicalities of sampling without replacement can be more exacting than sampling with replacement.

Sampling with probability proportionate to size (PPS) is the practice of weighting selection probabilities according to some auxiliary variable that is a measure of the "size" of the sampling units. For example, the number of trips undertaken, or the landed tonnage, from the previous year could be used to make the selection probability of the vessels that are most active within a fleet more likely than the less active vessels. PPS sampling has similar effects to stratification by size and the allocation of different sampling fractions to different stratum.

These approaches are further outlined in WKMERGE (2010).

2.3 Quality indicator 1; Reporting of population and sampled population

The first suggestion for a quality indicator is the reporting of the number of unique vessels in the total population, study population and realized samples – and number of trips in the total population, study population planned samples and realized samples.

These relatively simple Figures give a good overview of the coverage of the sampling in respect to the population and planned sampling.

The Figures could be reported in the form of a Table described in Figure 2.3. With small modifications this Table could easily be uploaded to a regional database to enable standardized reporting of the Figures, see Sections 2.7 and 5.

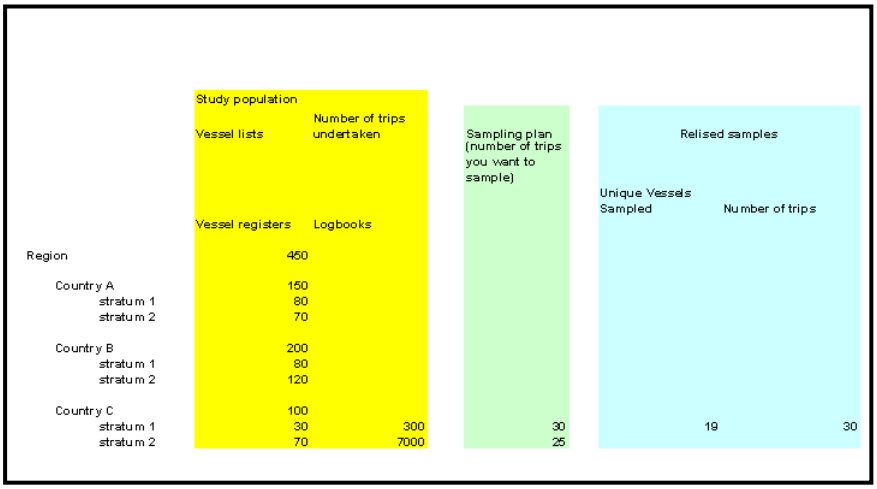


Figure 2.3. An example of a quality indicator: The number of unique vessels in the total population, study population and realized samples.

2.4 Quality indicator 2: Non–response and refusal rates

The determination of non–response and refusal rates is an essential component of a vessel selection scheme. Each selection of a vessel from our randomized lists leads to a contact attempt. The result of that contact attempt can be classified in one of two ways: It is either successful and the sample of data was obtained, or it was a non–response. The non–response rate quantifies the extent to which our random selection was available to be sampled. Some of the non–responses we obtain will be classified as "refusals", following the accepted sampling terminology. The refusal rate quantifies the extent to which we are unable to obtain samples from our study population because they will not provide it, for whatever reason.

These rates can be defined for different time intervals but for the DCF purpose it seems most appropriate to determine these rates annually. They should be calculated separately for each distinct major stratum of the fleet (e.g. trawler, gillnetters and beam trawlers).

The **non–response rate** is defined here as the proportion of all attempted contacts that ultimately failed to provide a sample, for whatever reason.

The **refusal rate** is the proportion of vessel skippers who, having been successfully contacted, ultimately failed to allow the observer to go on board to obtain the sample.

A less emotive term that may be of some use is the **success rate** which is simply 1–refusal rate.

All non–responses as a whole is interesting when you are looking at bias. If the non–responses have characteristics similar to the study population, e.g. engaged in the same fishery on the same fishing grounds etc., then even a high non–response rate may not cause bias in the estimates. If the non–responses however do practice a distinct differently fishery than the rest of the study population, then it will bias the estimates. The similarity between the non–response and study population, together with the potential bias caused by the non–response, can be analysed with the method suggested in Sections 2.5.1 and 2.5.2.

This leads on to a standard sampling protocol for contacting the vessels; so that we can obtain the numbers we need to calculate a standard non response rate and a refusal rate. Inspired by the existing terminology (Cochran, 1977), the non–responses

have been subcategorized into 4 groups, which will allow a better analyses of the non–response, together with a better understanding of the factors causing the non–responses rate.

2.4.1 Standard Contact Protocol

To ensure that there is a common approach to the calculation of non–response rates and refusal rates, we propose that there exists a standard sampling protocol that includes a number of stages. This protocol would be applied to each of the vessels selected from the sampling frame for a particular stratum and over the time period defined by the temporal sampling stratification.

Prior to this contact process, the vessel list is randomized according to one of the methods set out above, and the outcome of this process is that an observer obtains a trip on a specific named fishing vessel or that no sample was obtained from that specific named vessel during the temporal period set out in the stratification.

In most of what follows, the term vessel will be used as the focus of the contact and we assume that the skipper (captain) of the vessel is the individual who has the final say as to whether an observer accompanies the vessel on a particular trip. There may be instances where this is not the case and the contacts and negotiations are with vessel owners. This will be determined by the particular situation at the national level but the process does not differ in either respect. We also envisage that this process is applicable for obtaining a self–sample from a vessel participating in a self–sampling scheme. We assume that the person who is attempting to make the contact is either the observer themselves or a trip organizer who would then allocate the trip to an observer.

The standard protocol would involve the following steps:

- Can the vessel be contacted?
- Is the vessel suitable as far as the sampler is concerned?
- Is the contact attempt answered?
- Was the vessel sampled?

These stages are set out in Figure 2.4. If at each stage the answer is “yes”, then the contact attempt is successful, the observer gets the trip on the vessel and a sample is obtained. If the answer is “no” at any stage, then the contact attempt is classified as one of four possible categories:

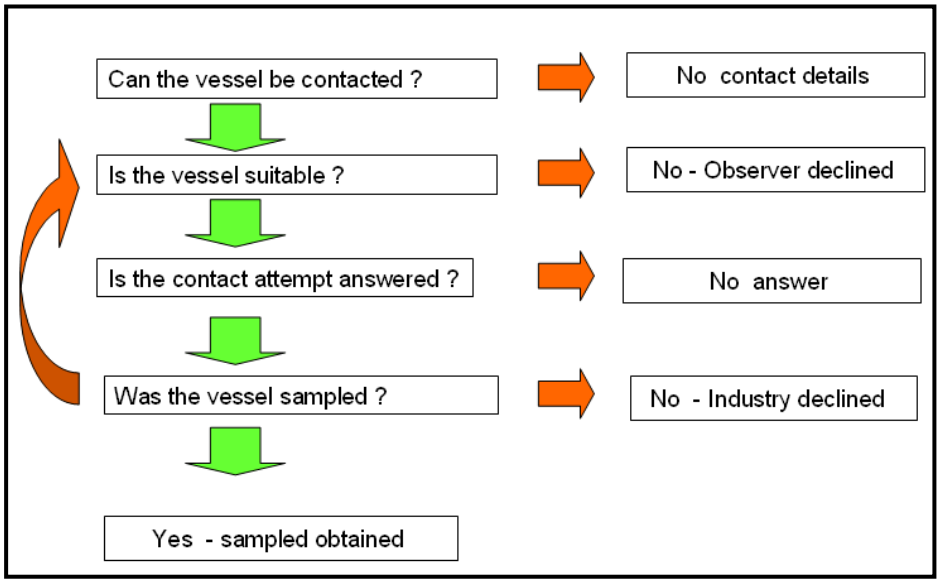


Figure 2.4. Scheme of a standard contact protocol; given a specific named vessel generated from a randomized list; the process would be followed to attempt to place an at–sea observer on a vessel.

Can the vessel be contacted?

In the simplest sense is there a telephone number for the skipper, the vessel owner, the fishery organization to which the vessel might belong that allows the vessel skipper to be contacted so a trip could be a requested. Addresses or e–mail can be equally applicable, face to face contact is also applicable but the crucial part of this stage is that the observer or trip organizer has the information available to be able to contact the randomly selected vessel.

If there are no contact details the contact attempt is recorded as "No Contact Details". The reason for having no contact details can be further subdivided into: observer decline (e.g. the trip organizer or observers being unable to establish and maintain an up–to–date telephone list); or industry decline (because the skipper does not provide his mobile number despite attempts undertaken by the fishery research institute).

Is the vessel suitable?

'Suitable' is used in a very broad sense and covers all situations when the vessel is categorized by the observer or the trip organizer as suitable for sampling. The critical point about this category is that *the observer or the trip organizer* is the one who ultimately decides not to select the vessel. If the vessel is not suitable, the contact attempt is recorded as "Observer declined"

It is not uncommon that a vessel is considered to be suitable before the skipper was contacted but unsuitable after the skipper has been contacted. This situation is catered for by the feedback loop. Thus, if the observer or trip organizer considers the vessel to be unsuitable after the skipper has been contacted (based on a changed perception of the situation), then the vessel is again recorded as "Observer declined"

There are numerous reasons why a vessel would not be categorized as suitable by those responsible for arranging an at–sea sampling trip. Concrete issues when this process is put to practice are discussed in detail in Section 3.

Is the contact attempt answered?

This we define as a situation where an attempt was made to contact the skipper or vessel but there was no response. This is the situation where contact details were known but (repeated) telephone calls or e–mails or other written correspondence were not answered.

If the skipper did not respond to the contact attempt, it is recorded as "No Answer"

Was the vessel sampled?

If the skipper does not allow the observer on board during a trip and so a sample cannot be obtained, the contact attempt is recorded as "Industry declined". The attempt to get a trip on a particular vessel can retard for several reasons and temporal limits may be established to consider a sequence of contact attempts to the same vessel as "Industry decline". For example, when several phone calls did not result in an observer trip despite the vessel having gone out fishing, the attempt may be categorized as "Industry decline" at the end of a quarter.

If the skipper does allow the observer to join the vessel and a sample is obtained, then the contact is recorded as "Sample obtained".

2.4.2 Contact log and the calculation of non–response and refusal rates

To be able to collect the information required to determine non–response and refusal rates, the phone calls need to be documented in an appropriate way. Examples of phone call documentation schemes are given in the presentations of Germany–Baltic and Scotland in this report (Annex 6 and 7) and of Denmark and France in WKPICS.

A phone call documentation list can have a central role in the determination of non–response and refusal rates. Experience from participants also suggests that one or two responsible trip organizers facilitate the selection of vessels and the documentation of the efforts and results rather than having several self–organized observers.

The analysis of the phone call documentation list should provide data for completing a contact summary Table of the type shown in Figure 2.5. Each row in the contact summary Table is the final result of 1 to many attempted contacts to a vessel that had been chosen randomly from a vessel list where an observer or trip organizer attempted to get an observer on board.

Sampling strata identifier	GER_OTB_Q2_2012	target number of trips = 35

contact order based on the randomised list	vessel	no contact details	Observer declined	no answer	Industry declined	sample
1	vess-034	1				
2	vess-056	0	1			
3	vess-321	0	0	1		
4	vess-934	0	0	0	1	
5	vess-098	0	0	0	0	1
6	…					
100		**5**	**30**	**10**	**20**	**35**
149						
150						

date	Trip Identifier	aux var 1	aux var 2

Non-response rate	100-35/100	0.65
Refusal rate	20/55	0.36

Figure 2.5. Vessel contact summary Table for a given stratum (OTB trawlers).

Explanations of the vessel contact summary Table (Figure 2.5):

Column 1: Random number after which the vessel to be contacted was chosen

Column 2: Vessel name (encrypted in the final version)

Column 3: Insert 1 if contact details of the vessel are NOT available, i.e. currently you don´t know the telephone number (e.g. because you are still completing your phone number list, confidentiality reasons of the industry); otherwise insert 0 and go to column 4

Column 4: Insert 1 if the vessel could be contacted but was not contacted based on the decision of the observer or the trip organizer. Or if the vessel was called and a trip was declined based on the decision of the observer or trip organizer. Otherwise insert 0 and go to column 5

Column 5: Insert 1 if you were unable to contact the vessel despite a reasonable amount of contact attempts within a reasonable time interval; otherwise insert 0 and go to column 6

Column 6: Insert 1 if you were able to contact the vessel but the conversations did not result in a trip (or self–sample); otherwise insert 0 and go to column 7

Column 7: Insert 1 for a successful contact which resulted in an observer trip or a self–sample.

The attempted contacts are recorded in the list. Once the planned number of samples is achieved or at the end of the year, the proportions of the different contact categories can be easily calculated.

A worked out example of the use of a vessel contact summary Table (Figure 2.5) illustrates the described procedure. For simplicity, we assume that our target number of samples to obtain is 35, and that we achieve this after contacting 100 vessels from a stratum that contained 150 vessels. Note, that probably several hundred phone calls were required to get these results. Summing up the 1´s in each column gives the proportion of each contact category. In our example, 5 of the randomly selected vessels could not be called due to lack of contact details. 30 were not called because they had been excluded or were excluded after being called based on observer or trip organizers judgment. 10 were called several times but never answered the calls. On 20 of the vessels we were unable to get a trip because the skipper declined (refusal) and for 35 of the vessels allowed to observe a trip and a sample was successfully obtained.

The non–response rate in this example is calculated as:

sum of the sums of columns 3–6 (i.e. 5+30+10+20)/100 = 0.65;

or: (100 vessels with attempted contact)–(35 vessels sampled)/100 = 0.65.

The refusal rate in this example is calculated as:

(20 vessels which refused attempts/55 vessels which were successfully contacted) = 0.36

2.5 Quality indicator 3 (Final quality indicators): Measurement of 'goodness-of-fit'

2.5.1 Comparisons of realized samples with sample populations

The final quality indicators envisaged for at–sea sampling are relatively simple and easily interpretable comparisons of various metrics in the total and study populations with the same metric in the realized samples. If the sampling is probability–based and the non–response rates low, then the data obtained from the samples will be proportional to that in the populations sampled.

For example Figure2.6 shows comparisons of the official landed weights (filled columns) with the landed weight sampled; the lines show the relative sampling activities for a haddock fishery (left) and a hake fishery (right). The comparisons are between gear type (top), ICES area (middle) and quarter (lower).

It can be seen that the sampling of the haddock fishery is generally more proportionate to the landings than the hake fishery, the latter, for example, has no sampling of a particular gear type (longlines) which is responsible for half the landings and ICES area 4 is disproportionately sampled in relation to ICES area 6. The study group suggest these kinds of plots can potentially be used to evaluate the sampling programs at the level of strata, country and region, and note that such a "relative value" function has been developed under the COST project.

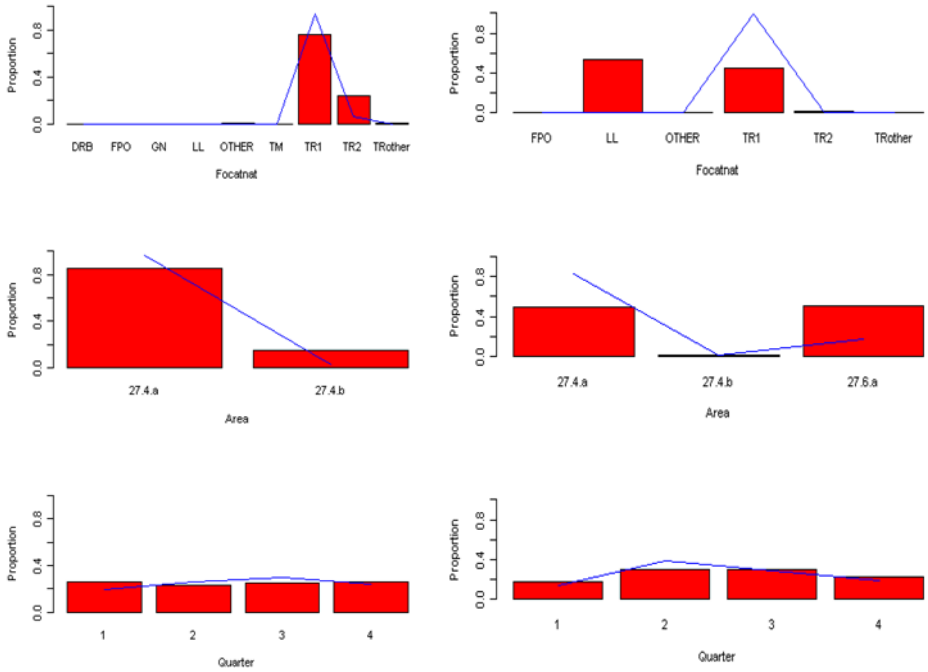


Figure2.6. Example of proportional comparisons of landed weight (bars), and landed weight sampled (lines) for gear types (top), ICES area (middle) and quarters (lower) for a haddock fishery (left) and a hake fishery (right).

Other examples could include VMS plots of vessel activity and discard observations.

To enable such goodness-of-fit indicators to be comparable across national sampling schemes (i.e. on a regional level), it is however necessary that the same variables are used. The most appropriate variables for comparisons of at–sea sampling are likely to be number of sampled trips, compared with the total number of trips, and these should be limited to a time, space and technical comparison. It is also necessary that the resolution of these variables (e.g. if the time variable is measured in months, quarters of some other seasonal division), should be consistent from one nation to another *within* a stock or fishery, though it is recognized that the most appropriate metrics may be quite different *between* different stocks or fisheries.

It also has to be appreciated that to make a meaningful comparison the resolution of such goodness-of-fit plots should be at finer scale than the original sampling stratification (e.g. if a population of trips happen to be confined to a single ICES area, then by definition so will all the samples obtained from such a population and thus a plot at the resolution of ICES area would be uninformative; the resolution would need to be at the level of statistical rectangle). For most of the regional fisheries the study group have encountered ICES areas (lowest level for the strata), quarters and métiers are appropriate resolutions for the comparison. But if a data is confined to a single dimension then the comparison should be made at a finer level, e.g. if there is only one quarter it would be relevant to stratify by month.

A single "goodness-of-fit" statistic that could be used to quantify differences for this limited range of comparisons would be of considerable benefit. Suitable statistics are likely to be non–parametric (hence no requirement to make assumptions about the distribution of the underlying data) and need to be appropriate to the measurement

scale of the data (i.e. ordinal of interval scales). Kruskal–Wallis tests, spearman rank correlations, $\chi 2$ tests are possible candidates but the study group felt that such considerations would benefit from input from expert groups such as WKPICS (recommendations).

Additionally, the details of such "goodness-of-fit" quality indicators need to be appropriate to the uses to which they are going to be put. For example, the use of relative standard errors (CVs) for auditing within DCR but not as yet as part of the stock assessment process are a case in point. SGPIDS consider that while the basic principles of the quality indicators are clear-cut, the details and their application encompass wider considerations. It is also noted that PGCCDBS 2012 recognized the need for a revision of the WKACCU scorecard approach and that a WKACCU 2 workshop may be a more appropriate forum in which to finalize such details. The RCMs are likely to be able to provide expert advice on appropriate metrics for different fisheries and stocks.

2.5.2 Bias due to exclusion of objects: Comparison of total population and study population

Given that there is the information necessary to differentiate the total population and study population as envisaged by quality indicator 1, (Section 2.3), then it should be possible to undertake comparisons between the characteristics of the total population and the same characteristics of the study population. Such comparisons would potentially be similar in form as the "goodness-of-fit" comparisons for realized samples outlined in Section 2.5.1. The variables to be compared would depend on what auxiliary variable information was available (i.e. the coverage and composition of logbook or sales slip data). Put simply, if the study population is substantially smaller than the total population and the study population differs significantly from the total population, then considerable care would need to be exercised in making inferences from the study population to the total population. Again the study group felt that such considerations would benefit from input from expert groups such as WKPICS. (recommendations).

2.5.3 Bias due to Non-response

Given that there is the information necessary to differentiate the non–responding Section of the study population from the responding part of the study population, as envisaged by quality indicator 2 (Section 2.4), then the extent of non–response bias can potentially be assessed. Again such comparisons would potentially be similar in form as the "goodness-of-fit" comparisons for realized samples outlined in Section 2.5.1. but with the census data (logbook and sales slips) differentiated according to the responding Section of the study population and the non–responding Section of the study population.

Non–response bias results if the non–response rate is high and the non–responders differ in important respects from the responders. Again if such a situation is found to be the case, considerable care needs to be exercised in making inferences from the sample data. Specifically with at–sea sampling for discards, if high discarding rates are correlated with high non response rates then there exists the potential for serious non–response bias.

2.6 Implications for the Regional Database

By developing reports for the quality indicators in a regional database it will make them easily available for different end–users and secure a standardized way of reporting all the suggested quality indicators (see Section 5). Furthermore the reporting could be done at several predefined levels e.g. regional, national or strata depending on the needs for different end–users.

To report all the suggested quality indicators, development of the regional database is needed. This development includes:

Update of the exchange format to

- allow for identification of sampling frames (e.g. sampling frame code in the different tables)
- add variables that are needed to identify sampling frames (e.g. number of vessels)
- add new tables on planned sampling (by sampling frame) and on non–responses (standardized contact log)

Produce new types of plots and reports.

2.6.1 Identification of the sampling frame in sampled data as well as in data on effort and landings

A code for sampling frame/strata is needed in the exchange format to be able to compare at the level of strata, since it may not be possible to identify the strata at the level of aggregation defined by the current format. Even if it is possible, it would make the processing and reporting much more simple. The code should be present in all exchange format tables.

2.6.2 Reporting of population and sampled fleet

The important measurements are the number of unique vessels in the total population, study population and realized samples – and number of trips in the total population, study population planned samples and realized samples.

If a code for sampling frame/strata is added to the current exchange format, it should be easy to identify the total population, study population and realized samples, but information about number of vessels in the official statistics and planned trips is missing. The needed information could be made available by adding the number of vessels to the CE Table (effort data in the exchange format). 'Number of Vessels' will have to be a decimal number, so it is possible to summarize the Figures by region, country and strata over the different strata. There is also a need to upload a small Table about planned sampling. For the purpose of the quality indicators suggested in this report, the following variables are needed in such a table:

Country, sampling frame/strata, planned trips

2.6.3 Reporting of Non–response

Needed – A code for sampling frame/strata in all present tables. The present tables are CS (sampled data), CL (landing data) and CE (effort data). Upload of the suggested contact log (new table).

The non–response and refusal rate could easily be calculated and tabulated at the requested level of interest.

2.6.4 Sample coverage

Needed – A code for sampling frame/strata in all present tables (CL, CE and CS).

Able to make coverage plots and statistic of the sampled population vs. study population at all levels.

2.6.5 Bias due to exclusion from the study population

Needed – a code for sampling frame/strata in all present tables (CL, CE and CS).

The excluded part will have no coding for sampling frame/strata and thereby permitting comparisons against the study population by e.g. area and quarter.

2.6.6 Bias due to non-response

Needed : Upload of the suggested contact log for reporting of non–response. To compare the non–response with the study population it should be possible to identify the pool of trips which have caused the non–response and refusal.

Assuming an exchange format with aggregated data (as present); this could be done by tagging the aggregated data in CL and CE with a code or tagging the data in the non–response Table with the needed information area, statistical rectangle, quarter, month, fishing activity etc.

Assuming an exchange format with disaggregated data; adding tripId to CE, CL, CS and contact log, so a direct linkage is possible

Able to make coverage plots and statistic of the non–response vs. study population at all levels.

3 Practical improvements to define sampling frames (ToR3); statistical sound and practical tools to implement vessel selection procedures (including registration of refusal rates) (ToR4)

In 2011 SGPIDS described in detail the sampling programs of the then contributors. The focus of the study group then was to review differences, provide guidance and to define minimum standards in procedures and protocols. Programs at the time had been designed to meet national precision levels and minimum targets for defined métiers. Although the focus of the study group has not changed, the results of catch sampling workshops promoting statistically sound methodology (WKPRECISE, WKMERGE and WKPICS including SGPIDS) has meant that these sampling programs have been evolving to improve on their statistical robustness. PGCCDBS 2012 provided a clear road map to regional coordination and a focus for developing statistically sound sampling schemes. Fundamental to these schemes is a move away from national ad hoc, quota based sampling schemes where a sampling event was conditional on where and how a vessel had fished or was going to fish. Since 2011 some member states have revised their sampling schemes and moved towards a random approach. Table 3.1 provides a summary of the current sampling programmes and is an update of the descriptions in SGPIDS 2011. The Table includes additional countries such as Germany and Poland.

In line with the Road Map described in PGCCDBS the study group has drafted a rough guide to developing sampling frames and vessel selection procedures using the experience of some member states currently adopting this approach.

Fundamental to a random vessel selection process is the sampling frame which describes the population to be sampled. The concept of sampling frames is described in Section 2.2. The sampling frame provides a list of vessels or trips which can be identified and sampled with a known probability. Following the key quality indicators described in that Section, this guide consists of three steps – 1. Defining the frame and creating drawlists; 2. Selecting a random vessel; 3 Monitoring and recording the selection process.

Table 3.1 summarizes the study group contributors use of sample frames, drawlists and contact procedures. In the Table the number of strata provides a reference to how each member state is currently categorizing their vessel activities and allocating their sampling effort. The periodicity of these strata is a reference to the sampling periods and the frequency that draw–lists may be updated. The Section relating to the selection process indicates whether draw–lists of vessels are created and vessels randomly drawn from them. Whether the vessels are selected by the observer or allocated by a coordinator shows how contact procedures may differ between member states. The contact procedures and what information may be recorded during this process is also summarized in this table.

Table 3.1. Summary of sampling frames, selection processes and contact procedures by participating member state.

		Sampling Frames			Selection process				Monitoring			
Country	**Sampling technique**	**Strata**	**Strata update**	**Primary sampling unit (PSU)**	**Drawlist**	**Random selection**	**Allocated or observer selected**	**Shared lists**	**Contact log**	**No. of calls**	**Range of responses**	**Remarks**
Example	*Self /Observer*	*Number of lists*	*Frequency of strata update*	*Vessel/Trip*	*Y/N*	*Y/N*	*Allocated/ Observer*	*Y/N*	*Y/N*	*Y/N*	*Y/N*	
England	*Observer*	*~24*	*Quarter*	*Vessel*	*Y*	*Y*	*Observer*	*Y*	*Y*	*Y*	*Y*	*Drawlists currently based on grouped métiers. Vessels listed by predominant activity previous year.*
Scotland	*Observer*	*3*	*Quarter*	*Vessel*	*Y*	*Y*	*Observer*	*N*	*Y*	*N*	*Y*	
Denmark	*Observer*	*13*	*Quarter*	*Trip*	*Y*	*Y*	*Observer*	*Y*	*Y*	*N*	*Y*	
Poland	*Observer*	*1*	*Quarter*	*Trip*	*N*	*N*	*Allocated*	*N*	*Y*	*N*	*Y*	*Vessel is randomly selected by coordinator*
Sweden	*Observer*	*13*	*Quarter*	*Trip*	*Y*	*Y*	*Allocated*	*Y*	*Y*	*N*	*Y*	*Drawlist currently based on group métiers and areas*
Spain (random selection trial)	*Observer*	*1*	*Annual*	*Trip*	*Y*	*Y*	*Allocated*	*N*	*Y*	*Y*	*Y*	*Started this year as a pilot*

												procedure
Spain (Mediterranean)	*Observer*	*9*	*Annual*	*Trip*	*N*	*Y/N*	*Observer*	*N*	*N*	*N*	*N*	
Spain (Atlantic)	Observer	7	Annual	*Trip*	*N*	*Y/N*	*Allocated and Observer*	N	N	N	N	
Spain (Basque)	*Observer*	*4*	*Annual*	*Trip*	*Y*	*N*	*Observer*	*N*	*N*	*N*	*N*	
Germany (Baltic)	*Observer*	*~8*	*Monthly*	*Vessel*	*Y*	*Y*	*Allocated*	*N*	*Y*	*Y*	*Y*	
Germany (Baltic)	*Self sampling*	*~8*	*Monthly*	*Vessel*	*Y*	*Y*	*Allocated*	*N*	*Y*	*Y*	*Y*	
Germany (N. Sea + Atlantic)	*Observer*	*~8*	*Annual*	*Vessel*	*Y*	*N*	*Allocated and Observer*	*Y*	*Y*	*Y*	*Y*	
Netherlands	*Observer*	*7*	*Annual*	*Trip*	*N*	*N*	*Allocated*	*N*	*N*	*N*	*N*	*Currently no documented random vessel selection procedures in place.*
Netherlands	*Self sampling*	*5*	*Annual*	*Vessel*	*N*	*N*	*Allocated*	*N*	*N*	*N*	*N*	*Vessels are selected to participate in reference fleet.*
Belgium	*Observer*	*8*	*Annual*	*Trip*	*N*	*N*	*Observer*	*N*	*N*	*N*	*N*	
Portugal	*Observer*	*7*	*Annual*	*Trip*	*N*	*N*	*Allocated*	*N*	*N*	*N*	*N*	*Keep record of receptive vessels.*

3.1 Step 1: Defining the sampling frame and drawlists

The sampling frame defines the population you want to sample. As described in Section 2.2 the frame covers the population and the period over which they are to be sampled. The population can be split into sub sets but if so, the distinctions between members should be clear, consistent and predictable.

In this instance we are concerned with national fisheries. At the most basic level, the only clear certainties are vessel and time. The vessel characteristics may limit the vessels to distinct fisheries or areas of operation and these if some clearly defined groups do not need to be sampled, they can be excluded.

3.1.1 How do I create a sampling frame?

The basic data sources for defining the sampling frame are:

- Current Vessel register
- Historic logbook data

The national vessel register provides a list of vessels and this, at its simplest, forms the overall sampling frame. The vessels physical characteristics from this and persistent fishing activity are contained in the logbook data and can be used to classify/stratify the vessels into distinct groups or sub–frames.

3.1.2 Stratification of the sampling frame

Following a thorough analysis of historic fleet activity and the characteristics of the vessels, the total population might be stratified into several sampling frames, based on some consistent and predictable characteristics. The sampling frames should be kept as simple as possible, and the structure needs to consider the end product as the raising procedures are dependent on the sampling frames.

Examples of characteristics that could be considered for stratification:

- Port of operation
- Length of trip
- Vessel characteristics
- Area of activity.
- Gear/gear groups
- Target species
- Quarter or month

Two examples of frames and vessel lists currently used – the Scottish currently have 3 lists covering 3 different components of the demersal fleet – Inshore vessels, offshore vessels and offshore Nephrops vessels. German–Baltic vessels were stratified by target species, subdivision, vessel length class, gear type and month using logbook information from the previous three years (use of SQL). Vessels in each stratum were ranked by their relative share to the stratum landing and those within the 90% threshold of cumulative landings were listed. For each month, vessel lists for the important strata are compiled (known from the landing patterns). From these lists, vessel owners are contacted, asked for observer trips or self–samples and each single phone call is documented.

These lists are the drawlists or reference lists that are used throughout the year to select vessels from.

If a group of vessels fall into a very clear category and do not need to be sampled, they can be excluded from a frame of vessels that does need sampling. Examples of these groups of vessels include pelagic seiners or very small boats <8 m LOA.

3.1.3 Exclusion of vessels

Some of the vessels inside the sampling frame might be practically inaccessible. If easily categorized, they could be excluded from the sampling frame at the outset, or if selected they can be chosen not to be sampled. If this forms part of the selection process, any exclusion will need to be registered (see step 3 below). If they are not in the sampling frame, they will be excluded from any raising procedure to that frame, but if included in the sampling frame, but rejected, they can still be included in the raising procedure. If included within the frame any potential bias may be more easily qualified.

3.1.4 Concerns

In the simplest case, all a nations vessels would be in the same sampling frame and by randomly selecting the vessels, the sampling would be representative for the entire nations' fisheries. Post stratification would then provide the information for individual métiers as required. However, in most instances there would not be enough resources to sample enough vessels to ensure sufficient coverage of all fisheries.

A common fear among member states about this process, is that vessels selected randomly from a broadly defined sampling frame will 'waste' effort on métiers they are not obliged to sample or that would be more adequately sampled. When moving away from the métier based sampling programmes, member states will still need to ensure that the allocated sampling effort reaches those parts of the fishery that significantly contribute to the discards. Some countries are currently dealing with this by constructing sampling frames defined by fisher métier, and allocating the sampling effort to the most 'important' métiers. However, the métier approach does not fulfil the criterion that strata should be stable over time, the mesh sizes, gear types and species groups that define a métier will change for a single vessel over time – vessels move between métiers. Since many vessels are active in more than one métier, these sampling strata are overlapping and therefore bias cannot be measured.

To illustrate the problem, the following example was discussed during the study group. Consider two different gears, A and B, these are used in the same area by the same vessels (Figure 3.1). The vessels can shift gear at any time, and in a non–predictable way. Most of the fishing activity is carried out with gear A, which generates low discards. Skippers are far more accommodating when vessels are using this gear and are therefore easier to negotiate a trip. Gear B generates very high discards, trips are less frequent and skippers are more reluctant to take observers on board (higher refusal rates) when using this gear.

If using a single random drawlist, the great majority (possibly all) of the observer trips could be carried out on vessels using gear A. But because of the higher discard rate overall from gear B (which also has a high variability of discards between trips) it is desirable to get a sufficient number of samples from this gear.

A	B
90% activity 8% discards Low refusal rate (100 tonnes discards)	10% activity 92% discards High refusal rate (900 tonnes discards)

Figure 3.1. A hypothetical scenario of 'wasted sampling effort', as an result of limited resources and random sampling. Gear A: generates low discards. Frequent trips. Low refusal rate. Gear B: generates very high discards, trips are less frequent. High refusal rate.

This problem is not uncommon and is currently being dealt with in a variety of ways. Some member states are stratifying by gear, using two different drawlists which overlap since most vessels appear in both strata. Another option is trying to deal with this by conditional sampling selection. All the vessels are in one drawlist but the same drawlist is used to sample gear A and B. A different draw is made for each gear and only those vessels are sampled that comply with the predetermined condition (A not B or B not A). The raising procedure, however, has to follow the conditions applied so it is better and for consistency that the condition is persistent and unlikely to change. Fishing area will persist but a mesh size difference for example would be unpredictable and liable to change.

Both processes complicate raising procedures and makes qualifying any biases difficult. If clear, persistent fleet characteristics cannot separate the vessels into distinct frames then it is necessary to sample the complete frame. There may be a seasonal aspect to the activity or affiliation to distinct ports which could help. You might consider increasing your sampling rate in a particular quarter to increase the chances of picking up the less frequent but significant part. Further advise on this issue is required from WKPICS.

3.1.5 Can the vessel list or selection process be weighted?

Currently some member states are using a weighted vessel list to ensure that the more active vessels are sampled. This is an example of unequal weighted probability sampling and needs to be accounted for in the raising procedure (ICES 2010) provides an example of the process. Other member states are excluding the less active vessels from their selection lists although these less active vessels still remain within the frame. This decision may be based on an assessed acceptable bias. WKPICS has been reviewing this issue and needs to provide clear guidance on best practice (recommendations).

3.1.6 With or without replacement?

Whether you decide to sample with or without replacement may depend on the number of vessels in your frame and an understanding of the within vessel and between vessel variation. Sampling with replacement allows the vessel selected to be available for selection within the same sampling period. This means the selection is more random but if the between vessel variation is greater than the within vessel

variation then sampling without replacement would give you a better measure of the variability across the fleet.

3.1.7 How frequently should these lists be updated?

Updating a list will capture new vessels to a fishery which otherwise might be excluded from frame and will get rid of vessels that are no longer relevant. How significant it would be to include these new vessels in the sampling frame would depend on the size of the fleet or number of vessels in the original frame, the sampling frequency and variability within the frame. It might be perceived to be an acceptable bias to ignore any new vessels and calculate the statistical impact at the end of the period.

Some member states are updating their frames on a quarterly or ad hoc basis using a comparison of historic data and the updated vessel registry. It may be statistically safer for the frame to not change over the period. Trying to account for change of ownership, change of skippers, new vessels and even changes of nationality may over–complicate the raising process. If necessary, these vessels may be more easily excluded during the selection process.

Overall, it seems reasonable to update the sampling frames and drawlists of vessels annually.

3.1.8 Regional planning

Although nationality may define a frame there are instances where a vessel of a member state may also occur in the drawlist of another member state – particularly with flag vessels and nomadic fleets that land into 'foreign' ports. Because of limited resources they might not fall into any member states frame and end up being ignored with the result that they get excluded from any raising procedure. Regional planning would improve on efficiencies in terms of sampling. In some instances setting up bilateral agreements to sample other member state's vessels is already happening. This might help improve on efficiencies but the probability of their being selected may not relate to the right frame, unless 'foreign' vessels are identified and allocated to the most appropriate frame. The RCMs and Regional database will go some way to improve on this.

3.2 Step 2 Random vessel selection

3.2.1 Randomization

Assuming the frame has been defined in step 1 – the list of vessels has been stratified into distinct groups. Quarterly targets might be defined for each group, probably weighted by the significance of those groups and the range of métiers those vessels cover. To sample one of those groups, a single vessel needs to be selected from that list randomly.

The tools currently used to do this ranges from programmes and functions in R to functions and VBA programmes in Excel and Access. For example, in Scotland R code is used to randomly order the groups providing a list of vessels that need to be approached sequentially until successful. For the next trip the code produces another randomly ordered list. The Danes use a neat system developed in VBA and Excel which provides the observer with a vessel to sample at the click of a button and the means of logging the observers success – if unsuccessful the button provides another vessel. The English system uses VBA in Access to provide a randomly shuffled list of

vessels in Excel which is shared over the national network by all observers. Each vessel on the spread sheet needs to be approached in turn and the successes recorded alongside.

These examples show the process of selecting a vessel or trip from a list. The other process is randomly selecting a vessel or trip from a grid explained in Section 2.2 (Figure 2.2). You have no control over which vessels are going to be active within a limited period. You can systematically select a period to sample an area and randomly select a vessel currently fishing within that period to approach.

These systems are evolving and systems are improving to include the means of recording successes.

3.2.2 Vessel selection

Most member states are currently using a list of vessels or groups of vessels as a reference for some form of selection. Member states select vessels from these lists in a number of ways. These methods range from:

a) Observers being provided with individual lists of vessels randomly sorted with the instruction to sample the vessel at the top of the list.

b) Vessels being randomly selected from the pool of vessels and trips being arranged by a coordinator and allocated to an observer.

c) A contact is selected from an expanding pool of fisheries organizations and vessel owners and they provide access to a vessel.

d) Using VMS to identify a fleet of vessels active in a particular management area within a limited period and randomly selecting a vessel from that. The vessel is then contacted to bring back a sample.

Those member states not already doing so are moving towards a random process of selecting vessels.

Some member states use a central coordinator to make the selection, contact the industry and then allocate that vessel to an observer as in the Swedish programme. In Scotland they have a coordinator creating individual drawlists and which are given to an observer who then has to chase the vessel.

There are a number of advantages to both processes – having a central coordinator saves the observer administrative 'downtime' and also provides an efficiency in that vessels can be allocated to an observer based on that observers location. Currently members states may, for efficiency, be using 3 or 4 regional drawlists allocated to individual observers based in and covering those regions. Having the option to appoint effort from a central draw could improve on efficiency and reduce non–responses and refusal rates.

Having the observers managing the selection and approaching the vessels could provide better expert screening during the contact procedure (see step 3; Section 3.3). Their local knowledge and experience may lead to better success at gaining access to vessels, e.g. seeing a vessel being painted on a slipway should suggest to the observer that it is not worth contacting that skipper. Therefore, if the former process of using a central coordinator is used it is recommended that the coordinators regularly exchange information with the observers.

As described in Section 2.2, the process of selection is that once a vessel has been identified then the skipper or vessel owner is contacted to arrange a trip. However, there may be a number of reasons why that vessel does not get sampled.

3.3 Step 3 Monitoring the selection process and recording successes

To sample a vessel either the skipper or the vessel owner needs to be contacted. Though obvious, it is important to note that this requires not only an updated list of active vessels but also constantly updated information about the address (including (mobile) phone contact) of their captains or owners. Recording the response or result of the contact is crucial to monitoring the success and quality of the sampling scheme. Although a number of countries have procedures in place for recording a 'refusal' rate, the study group agrees that the term needs to be better defined. The term could be considered contentious and there will be a number of reasons why a vessel is not available to sample beyond a refusal to cooperate. Most member states have schemes of recording this but there is a minimum number of generic reasons or responses needed to qualify the selection.

Section 2.4 highlights the importance of monitoring refusal rates and cataloguing calls and provides a Standard Contact Protocol. If this is likely to ever be contentious, it is worth considering formal or informal presentations to the industry of the sampling schemes and what information we are collecting and documenting.

Uwe Krumme presented an extensive analysis of the German–Baltic phone call–register where each call to an owners or individual vessel to arrange a trip or sample is logged (Annex 7). It clearly indicated the benefit of recording and, at the very least, highlighted the importance of defining the refusal rate. At WKPICS an analysis of the Danish system of registering successes was presented and being able to demonstrate to the industry the lack of cooperation over preceding years helped improve access to vessels in subsequent years.

To qualify the process and report on successes, as Section 2.4 describes, the selection and the outcome needs to be recorded. If unsuccessful, the next vessel on the random drawlist is approached or the vessel selection system is used to select a different vessel. Nevertheless, each selection needs to be accompanied with a comment that at least covers the five qualifiers listed in Section 2.4.1:

- 'No contact details'
- 'Observer declined'
- 'No answer'
- 'Industry declined'
- 'Sampled'

If the contact protocol comes to an end at any of the stages before the vessel is contacted (see Figure 2.4), the outcome still needs to be recorded against that vessel selection. Once the vessel has been contacted, there are various reasons why a vessel does not get sampled and these reasons also need to be recorded.

3.3.1 'No contact details'.

This means that within that sampling period the observer was unable to get the details of the owner or the skipper. This might occur if either or both the fishing industry and enforcement agencies are unwilling to provide contact details (e.g. confidentiality issues).

3.3.2 'Observer declined' (Not available/ Expert screening).

Prior knowledge of how the vessel is currently employed (e.g. at sea, on guard duty, under repair, skipper on holiday, no longer fishing) will mean, this vessel can be discounted without trying to make contact with the owner/skipper. This will include vessels that may be categorized in the short term as 'observer declined': vessels that can't be sampled because of safety/resources or practical reasons (the vessel may be single handed, there may be no living space for the sampler, the port the vessel is currently residing may be too far to travel to). In the Scottish sampling scheme all reasons are recorded. When your sampling protocol is set up as 'sampling without replacement', but your active drawlists still include these vessels then, if the vessel has already been sampled, and the vessel is selected again, will also be categorized as 'observer declined'.

When a skipper is willing to cooperate but an observer has concerns about the vessel and/or crew (e.g. safety reasons) and decides to not go, or the observer drops out for other reasons and can't be replaced, then this also needs to be recorded as 'observer declined'. Obviously, these 'refusals' need to be kept distinct from those measuring industry cooperation.

3.3.3 'No answer'.

This may follow a pre–defined number of attempts over a limited or the entire period.

3.3.4 'Industry declined' (Refusal Skipper/Industry)

Contact or an approach between an observer/trip coordinator and a skipper/vessel owner can result in a refusal to cooperate with the sampling scheme. Since there are many different reasons to refuse an observer on board, the study group agrees to establish sub–categories for the category 'Industry declined'. Decisions not to cooperate can be categorized in a wide variety of refusals: from a 'loud and clear' no ('hard no') to any kind of excuse ('soft no') or even a yes – but skipper consistently fails to deliver and is eventually categorized as a refusal. There may be a genuine reason rather than the determination not to cooperate, quantification of different categories may be useful to better understand the major reason for not getting on board. Sub–categories should be similar among member states to facilitate international comparisons.

In the case of "industry declined", this might be further categorized into e.g.

for active vessels:

hard NO,

soft NO e.g. call someone else,

for inactive vessels:

not fishing (e.g. not in the area (VMS data, www.marinetraffic.com), vessel in shipyard, oil job, training course, trip promised but cancelled due to illness of skipper etc).

Further categories may be added when needed.

Any additional sub–categories can be added as long as the overall distinction is maintained between who was responsible for the selected vessel not being sampled.

To facilitate the analysis of potential bias caused by the non–response, it is important that it is possible to link the non–responses to the potential trip we have not sampled. For this purpose vessel identifier and date have been included, together with a couple of auxiliary variable that a the national level will enable a link to e.g. logbook, VMS, sale slips or other sources of information. The resulting link could be put in the Table as the trip ID. These could also be the linkage to other tables in a regional database, see Section 2.6.

3.3.5 'Sampled'.

This is a register of the success of the approach and is only apparent after the vessel has been sampled rather than a skipper just saying yes.

It is important to ensure that all methods of approach are recorded. This detail might be included in the contact log but could include – e–mail, quayside approaches, even facebook.

The number of calls can be used as indicator of the effort required to gain access to a vessel but could also be used as an index of cooperation.

4 Identification of appropriate on board sampling procedures: 3 case studies (ToR2)

Historically, sampling of commercial vessels is organized on a national level: national fleets are sampled by observers using sampling protocols, which are developed on a national level. Although similar métiers can be recognized on an international level, the fisheries of each country have its specific features that result from culture and tradition. The same accounts – up to a certain level – for the port and governmental administrations and logistics. These international differences eventually resulted in different sampling protocols by country. Even in situations where commercial fleets operate across borders and similar nationally defined métiers are active in the same area, significant differences in sampling protocols between countries can be identified.

It is SGPIDS aim to harmonize or identify methods of best practices discard sampling protocols on board commercial vessels between member states. In the process of reaching its goal SGPIDS will identify standards of best practise for discard sampling on board commercial vessels. These standards eventually result in the most appropriate on–board–sampling–procedures for different fisheries. Adoption of these sampling procedures by member states will, in time, result in more standardized discard sampling at an international, c.q. EU, level.

The first step of the standardization process is to make an inventory of the different sampling methods used between the member states across Europe (ICES 2011). Tables 4.1–4.5 of the 2011 SGPIDS report give a complete, updated, overview of the sampling methods used by fleet segment or métier by eleven member states. Secondly, areas of overlap where similar métiers or fleets are sampled by different sampling schemes; should be identified and assessed. On board sampling procedures of different countries are in general very difficult to compare. Hence sampling methods of "similar métiers" in different countries can only be assessed with a thorough understanding of the fishery methods and the background on a very detailed level.

The group decided to focus on case–studies from fisheries where there is overlap in the sampling programmes. This is particular the case in neighbouring countries, where vessels of one country may fish under the flag of the other and target the same species assemblages. The group agreed to focus on three case studies and to make a description of the (national) on board sampling protocols, reflect on the pros and the cons, analyse the motives and identify possible bias and/or 'bottle necks' on different sampling levels: haul, trip and fleet.

4.1 Comparison of sampling procedures

The identification of on board sampling procedures has been addressed by comparing cases where different countries are sampling the same métiers. Jens Ulleweit and Edwin van Helmond presented a working paper (Annex 5) on a case study where (by accident) a German and a Dutch observer were sampling the same vessel and trip – an unique event in the history of on board discards sampling in Germany and the Netherlands and a unique opportunity to compare protocols. The case was used as a template for the comparison of two further sampling schemes with potential overlap: the sampling of Belgium and Dutch beam trawlers and the sampling of Spanish and Portuguese demersal trawlers in the ICES Division VIIc, IXa. In addition, the sam-

pling of the Dutch freezer trawlers by Mauritanians operating off the coast of Mauritania was added to the German–Dutch case study.

4.1.1 The German – Dutch freezer trawler case study

The sampling on board German and Dutch freezer trawlers is described in a working paper by Edwin van Helmond and Jens Ulleweit (Annex 5). The Sampling on board the same freezer trawlers operating off the coast of Mauritania by Mauritanian observers is described in Annex 8.

In this study case, differences in sampling protocol already appear in the very first level of onboard sampling: at the haul level. Two different strategies are described on how to take a sub sample at haul level: While Germany and Mauritania (up to 2005) collect (sub)samples from the catch which is later sorted by the crew into landing and discard ("cutting knife" approach), Dutch and Mauritanian observers (from 2005 onwards) collect a discard sample after the catch is sorted. These differences in sampling protocol, at the haul level , will have an effect on catch and discard estimates in the following steps (trip and fleet level).

In case of the biological sampling, the background for representative sampling in the Dutch protocol is historical (25 random selected fish from the catch): in the past in the market sampling program crew members were asked to take a random sample by species, by week and by subarea in a reference fleet. These samples were directly raised to the catch of the fleet. All sampling of pelagic species still uses the same method, but in the DCF sampling program the representative sample is now used to make an age–length key to split larger LF sample to age.

Another clear difference is based on the objectives of the different sampling strategies. For the Mauritanians and Dutch, sampling of the large fish and dolphins is an important issue, while monitoring of large bycatches in the German sampling protocol is completely missing. However, it seems that in daily practice on board, Mauritanian– and Dutch observers have problems to live up to the protocol, as they are often occupied with the discards sampling on the lower deck. In addition, observers feel that their presence on the bridge, the best positions to observer large bycatches, during hauling is not always appreciated by the crew. Also, because of safety regulations, observers are not always able to check the catch on deck and – in Mauretania – where bycatch is trapped by the so called "shark–filter" is missed by the observer.

Table 4.1. Comparison of sampling protocols on board freezer trawlers in Netherlands, Germany and by Mauritania.

LEVEL	NETHERLAND	GERMANY	MAURITANIA+ RIVO (BEFORE 2005)	MAURITANIA
Haul selection	All hauls *Con: in case the sampling of a haul is missed, this may cause a bias in the composition of the catch on trip level, because missed catches may differ from sampled catches (small, mixed)*	All hauls *Con: in case the sampling of a haul is missed, this may cause a bias in the composition of the catch on trip level, because missed catches may differ from sampled catches (small, mixed)*	All hauls *Con: in case the sampling of a haul is missed, this may cause a bias in the composition of the catch on trip level, because missed catches may differ from sampled catches (small, mixed)*	All hauls *Con: in case the sampling of a haul is missed, this may cause a bias in the composition of the catch on trip level, because missed catches may differ from sampled catches (small, mixed)*
Bridge data (information obtained from the crew)	Estimation and registration of total catch *–Con: Estimation from total catch is less accurate than getting information from the landings*	Get information of the total landing. (Some time information on total catch) *– Pro: Getting information from the landings more accurate than estimations from the total catch*	Estimation of the total catch *–Con: Estimation from total catch is less accurate than getting information from the landings*	Estimation of the total catch *–Con: Estimation from total catch is less accurate than getting information from the landings*
Sampling selection	Take a representative sample from the discard fraction. In order to get a representative sample, different sub–samples are taken repeatedly at different moments while processing the haul (Sample size: 20–25 Kg). *Pro: Unbiased sorting behavior information. Focused on the discarded fraction, provides more information on discarded part of the catch. Con: slower procedure, observer activity depend on the crew. More complex sampling protocol. For biological sampling,*	Take a sample of the unsorted catch. Sort all the fish species into landings and discards (based on the settings of the sorting machine) *Pro: faster procedure, more independence for the observer (not dependent to the crew dynamic)* *Con: sorting decided by the observer, possible bias compared with crew behavior.*	Take a sample from the different conveyor lines (after the sorting). One line per retained species and another line for the complete discards *Pro: Unbiased sorting behavior information . Different raising factors obtained for both sorting class.* *Con: slower procedure, observer activity depend on the crew. More complex sampling protocol*	Take a random sample of the entire catch (before the sorting). When sampling, the observer decided if a fish is retained or discarded (based on the sorting practiced of the crew) *Pro: faster procedure, more independence for the observer (not dependent to the crew dynamic)* *Con: sorting decided by the observer , possible bias compared with crew behavior.*

	another sampling must be drawn from the unsorted sample.			
Sampling procedure	Species Length measures: For fish species at lower cm. Species weight measures: Hanging balances.	Species Length measures: For fish species at lower cm. Species weight measures: Hanging balances.	Species Length measures: For fish species at lower cm. Species weight measures: Hanging balances.	Species Length measures: For fish species at lower cm. Species weight measures: Hanging balances.
Biological sampling	Age estimations of the unsorted catch: a . Take a representative sample of the unsorted catch, b. Otoliths from this sample are prepared and analysed. c. The sample of age analysis consists of 'sized' and 'undersized' fish. A sample consists of minimal 3 individuals per length class per area (ICES quadrant). *Pro:*	Age samples: During at least 2 hauls while the fish is measured 10 fish of each length class (sized and undersized) are separated and frozen for otolith analyses in the institute´s lab. This is done by ICES division. *Pro: Probability of damages on otholits reduced if extracted at the lab* *Con: lower geographical precision*		

	Con: More sampled effort needed that in case of getting the biological information from the discard sample. *Higher time consumption per sampled haul when otoliths analysed onboard*			
Raising to haul level	Discard percentage (dp): Discard weight at time *t*/retained weigh at time *t* Total Discard Estimation (D) =[dp* total catch] *Pro: Fast method. Systematic, not biased by human estimations.* *Con: Method relying on the assumption of the linear distribution of species within the total catch (Assuming length and species homogeneous distribution in the hold)*	From the sample: Discard Percentage per species (dps) Retained Percentage per species (rts) If species total landing available; Total Discard Estimation (D) =[Species total landing weight] *(dps/rts) If total catch available; total catch weight*[dps/100] *Pro: Fast method. Systematic, not biased by human estimations, two alternatives if there is a lack of information.* *Con: two alternatives may provide inconsistency in results*	Visually estimate of the contribution of every belt to the total in terms of percentage (lp) . Total catch per line obtained by: Total catch * ([lp]/100) *Pro:* *Con: highly biased by human estimations as it only use estimated values from both skipper and observer.*	Raising Factor (RF) for all fish. RF= [Total catch]/[weigth of random sample] For discarded species: Total Discard Estimation (D) =[Species weigth in sample]*RF *Pro: Fast method. Systematic, simplicity and easy to carry on in poor data situations.* *Con: biased by human skipper estimations of the total catch*
Marine Mammals/ Protected/ rare species	Systematic bycatch recordings of large animals by inspecting the catch when it is taken on board.		MM catch Recordings not systematic/parallel sampling for rare species. Specific raising factor	Systematic bycatch recordings of large animals by inspecting the catch when it is taken on board.

	Con: The observers are often occupied with the sampling of the catch in the factory, and they may not be informed of the moment when a new catch is taken on board	*Con: There is not a clear protocol for these protected and rare species.*	*Con: The observers occupied with the sampling of the catch in the factory,* *The crew is reluctant to let the observers inspect the catch: Animals retained in the shark filter are missed by the observers. There is no protocol for rare species.*

4.1.2 Comparison of Spanish and Portuguese demersal trawl protocols targeting mainly horse–mackerel

Explanations of the sampling scheme and the raising procedures are given in Annex 9. A summary of the on board protocols is given here.

Spanish protocol for sampling onboard of commercial OTB vessels:

The 'Spanish Discards Sampling Programme' follows the guidelines established in the ICES `Workshop on Discard Sampling Methodology and Raising Procedures' (2003). Observers–on–board programme is based on a hierarchical sampling design. Trip is considered as the Primary Sampling Unit (PSU). Fishing operations are stratified into métiers in order to achieve better sampling allocation. The métier herein considered for comparison with the Portuguese case is

OTB_MPD_>=55_0_HOM: trips targeting horse mackerel (*Trachurus trachurus*) operating in ICES VIIIc–IXa

Trip is quasi–randomly (cooperative or opportunistic) drawn from the total trips carried out at the time the sampling is planned. All hauls (census) are sampled in most of the cases. Physical covariates associated to the setting and hauling (date, time, hours fished, depth...) are recorded directly from the bridge. The skipper is asked to estimate the total catch once the codend is hauled on board. Catch sampling is split into the retained and discarded fractions after the commercial sorting. Samples for the discard fraction is obtained at the end of the conveyor belt at different moments during the sorting process (several boxes will comprise a single sample). The sampled weight depends on the discard composition, but usually is not larger than 20 kg. Information from all species or groups of species in the discard sample is collected. Length sizes are the target information for all fish species and Nephrops. The sampled species weight are estimated using the weight–length relationship when available, else, weight is obtained using hanging balances. Invertebrates, other than Nephrops, are counted and weighted by species or groups of species. Species weight in discard sample is raised to haul level using –[Total Catch Estimation (k)–Observed retained catch (k)]/Discard sample (k)– as raising factor. Otoliths are extracted from undersized individuals present in the discard fraction. The observer checks on deck or asks to the crew if any marine mammal is caught. Morphometrics are collected in those cases. When all hauls are sampled, the species discards at trip level are calculated as the sum of all hauls. If only a set of hauls were sampled, the mean discards per haul is estimated and raised to the total number of hauls.

Portuguese protocol for sampling onboard of commercial OTB vessels:

The Portuguese on–board sampling program, included in the EU DCR/NP, is based on a quasi–random sampling of cooperative commercial vessels between 12 and 40 meters long. The programme started in late 2003 and involves on–board sampling of several fisheries. These include, amongst other, vessels operating bottom otter trawl, deep–water set longlines, gill and trammelnets (of various mesh sizes) and purse-seine in ICES IXa. The bottom otter trawl fleet (OTB) is the most comprehensively sampled fleet, with a time-series dating back to end of 2003. For sampling purposes the Portuguese OTB fleet is split into two métiers, and the métier herein considered for comparison with the Spanish case is:

OTB_DEF_>=55_0_0 – a demersal fish fishery that operates cod–end mesh size 65–69 mm and >70 mm and targets horse–mackerel, cephalopods and other finfish.

As stated by EU DCR/NP (CR (EC) 199/2008; CD 2010/93/EU), the Primary Sampling Unit in the Portuguese on board sampling programs is the fishing trip. The Portuguese on–board sampling program targeting the OTB_DEF>=55_0_0 métier is based on a quasi–random sampling of trips from a set of cooperative vessels known to operate in the fishery. Haul selection is systematic (odd or even hauls) after random choice of the starting haul (first or second). Generally teams of two observers are deployed on board each vessel. The observers collect fishing effort information (date, time, hours fished) and other physical information directly from the bridge in every haul. In each haul selected for sampling, the observers follow the crew to the sorting deck to start the sampling. A sample of 2–3 boxes is taken from the unsorted catch. Average sample size per haul is ~21 kg. The sample is split into retained fraction and discard fraction following the fishers' criteria. Numbers and weights are obtained for all species in each of the fractions. Length frequency is obtained for all fish species, cephalopods and commercial crustaceans. The weights are usually taken using hanging balances, except when sea conditions are bad and weight–length relationships are used. Total discard weight is estimated as [total landings] * [sample discard weight]/[sample retained weight]. Biological sampling of discards (individual weights, otoliths and maturity) is done at the lab from samples collected onboard. Quarterly sampling targets per length class and sex are defined for a set of main commercial species. Biological sampling of the retained fraction is carried out after fish purchase and mostly for supplementing market samples. The observers check on deck or ask to the crew if any marine mammals or seabirds are caught. Trip total discards can be obtained but are not calculated routinely because the raising algorithms used in fleet level estimates are based on average discards per hour.

Table 4.2. Comparison of the on board protocols in Spanish and Portuguese demersal trawl fishery.

LEVEL	SPAIN	PORTUGAL
Haul selection	Depending on the trip duration, usually census *Pro: flexible, high coverage (almost 100%)* *Con: may be difficult to implement in multi–day trips*	**Haul selection** systematically (odd or even hauls) after random choice of starting haul (first or second) *Pro: statistically sound, easier to implement in multi–day trips* *Con: may miss areas with lower effort if vessel only makes one haul there*
Bridge data (information obtained from the crew)	Physical variables related to setting and hauling, Estimation of the total catch *Pro: enhances the acceptance by the industry of the reported data* *Con: no independent measurement (potential bias); sensitive to fishers/managers issues?*	Physical variables related to setting and hauling. Estimation of the total catch (since 2011)[1] *Pro: enhances the industry understanding of on board sampling work;* *Con: no independent measurement (potential bias); sensitive to fishers/managers issues?*

[1] This estimate is accessory. It is collected has a means to compare the industry estimates with the on–board sampling protocol estimates. It does not enter the discard calculations.

Sample selection	Discard sample: from the conveyor belt for discards at different moments during the sorting process. Retained sample: Random sample from the boxes for the retained species. *Pro: precise composition of the discards; independent/objective sampling* *Con: requires the sampling of retained fish from already sorted fish boxes; Final accuracy of discard estimates dependent on fisher's total catch estimate (see Bridge data.Final sample size of length samples, e.g. from discards may be dependent on discard rate.*	A sample is taken from the unsorted catch at different moments during the sorting process. Observer follow the fisher's criteria in the sorting of sample into retained and discard fractions. *Pro: easy; quantities always fit; final accuracy of discard estimates unbounded and independent of fisher's total estimate (see Bridge data)* *Con: "knife–edge" approach may obscure complex discard patterns, e.g. when fisher's keep some (but not all) of the fish from a specific size class. Final sample size of length samples from discards are dependent on within–species and between–species discard ratios.*
Sampling procedure	Species Length measures: For fish species and *Nephrops* at lower cm. Subsampling for abundant species with a clear unimodal length size distribution. Species weight measures: Weight–length key used when available. Otherwise hanging balances used *Pro: faster procedure* *Con: Weight length relationships must be fine–tuned (20% differences may occur at quarterly level)*	Species length measures: All species are measured except for some invertebrates. Generally all individuals are measured but subsampling of abundant species. Species weight measures: hanging balances except in harsh sea conditions (Weight–length relationships) *Pro: less assumptions (weight length relationship rarely used)* *Con: variability of hanging balances weights*
Raising to haul level	[Discard sample (k)]*([Total Catch Estimation (k)–Retained catch (k)]/Discard sample (k))	[total landings] * [weight of discard fraction]/[weight of retained fraction]
Marine Mammals/ Protected/ rare species	MM: Check on deck or ask to the crew if any marine mammal is caught. Morphometrics recorded. No systematic protocol for rare species collection *Con (MM): dependent on the cooperation of the crew; bycatch may be underestimated*	MM: Observers on deck in sampled hauls. Skippers are asked to inform the observers of incidental catch of cetaceans and seabirds in not sampled hauls. Observers instructed to record and bring ashore all unusual specimens *Pro (rare specimens): additional information may increase knowledge of rare species and/or bycatch.* *Con (MM): dependent on the cooperation of the crew; bycatch may be underestimated*

Comments: Like in the preceding case, different strategies are in use in the sampling of very similar métiers. Portuguese trip discard estimates are unbounded and independent of fisher's total catch estimation. Spanish trip discard estimates are bounded by fisher's total catch estimation. Portugal collects samples before the catches are sorted by the crew and sorts them across the fisher's criteria in each haul; Its sample size for, e.g. length frequency of discards depends on the number of fish discarded in the sample. Spain collects the discards sample after the catch has been sorted; Its

sample size for, e.g. length frequency of discards can be dependent on the number of fish discarded in the sample. In the determination of the discarded weights from each species Spain uses weight–length keys if available; Portugal relies mostly on hanging balances. However both countries may switch to each other's method, depending on the availability of weight–length keys and the sea state. In both sampling schemes there is dependence of the cooperation of the crew in the sampling of rare and endangered species.

4.1.3 Comparison of Belgian and Dutch beam trawl protocols

Belgian protocol for sampling onboard of commercial beam trawl vessels:

Selection of the vessels for the Belgian observer program is random but conditional on the cooperation of fishers ("quasi random"). Data are being collected at two levels: trip/vessel–related data (vessel name, trip number, fishing gear used, departure and return time,…) and haul–related data (number and date of haul, time and position of shooting and hauling, area, valid/invalid haul, haul sampled or not, remarks,…).

Every second haul is sampled by an observer to ensure that sampling takes place around the clock to reflect typical working conditions. The crew sorts the marketable fish from the conveyor belt and they store this retained part of the catch in baskets for the observer to sample later on (different species in different baskets). In the meantime, the observer is sampling the discarded fraction of the catch.

The observer sorts all the discarded species of commercial importance and determines the total weight for each species. For a selected set of species (*Solea solea*, *Pleuronectes platessa*, *Gadus morhua*, *Merlangius merlangus*, *Melanogrammus aeglefinus*, *Merluccius merluccius*, *Lophius piscatorius* , *Lophius budegassa*, *Psetta maxima* and *Scophthalmus rhombus*), the observer also takes length measurements. Length measurements for the different discarded ray species are estimated for males and females separately. Usually, the length of all individual fish in the discarded part of the tow is measured. Only if a species is extremely abundant, a smaller representative subsample is measured. The ratio of the total weight and the subsample weight is used to estimate the total number of discards per cm–size class per species in the sampled tow. The retained part of the catch is treated in the same way as the discarded part of the catch except for rays: the observer only takes a total weight of all rays together.

In each trip, otoliths from 5 fish per cm–size class per species per area, are collected for age reading. For the retained part of the catch, otoliths are taken of cod, whiting, hake and haddock (only one otolith from each specimen). In the discarded part of the catch both otoliths are collected from cod, whiting, hake, haddock, sole, plaice, turbot and brill.

Dutch protocol for sampling onboard of commercial beam trawl vessels:

Selection of the vessels is quasi–random and based on co–operative sampling. This means that co–operation of a skipper with the project is on voluntarily basis. For each discard sampling trip, one observer goes onboard a vessel, sampling at least 25% of the hauls.

After each haul, the marketable fish is sorted from the catch by the crew of the vessel on a conveyer–belt. From each sampled haul, a representative subsample of the discards was taken from the conveyer belt by the observers. All fish in the sub–sample were counted and measured. Benthic invertebrates were only counted. Total and

sampled volume of discards was recorded. In addition, sub–samples of the landed fish were measured, and total and sampled landings weight were recorded.

If possible, otoliths were collected from the commercial important discarded fish species (plaice, sole, dab) for age readings. All data were entered into a computer program on haul–by–haul basis and later transferred into a central database.

Sampling protocol per haul: (1) Volume estimation of total catch per haul together with the skipper (number of baskets). (2) Take sample of discards. The sample consists of one basket (35 kg). To get a representative sample, discards are taken at different moments from the conveyer belt when processing the haul. (3) Measuring discard sample: a. Sort all fish species, take length measurements and register total number by species and length class; b. Sort all benthos and register total number by species. (4) Measuring landings sample: a. Sample landings from target species (sole and plaice), 10–15 kg. Register total number by species and length class; b. Sample landings from non–target species (e.g. dab, turbot, brill, whiting, cod) 10–15 kg. Register total number by species and length class. (5) Age estimations of discards: a. Sample otoliths from most discarded commercial species (plaice, sole and dab); b. The sample of age analysis consists of undersized fish. A sample consists of minimal 3 individuals per length class per area (ICES quadrant). (6) Information on position, haul duration, wind direction, fishing depth en landed catch is collected in cooperation with the skipper for each haul. (7) Registration of total landings: information on total landings is collected from auction file at the end of the trip.

Table 4.3. Comparison of Belgian and Dutch protocols in the beam trawl fishery

LEVEL	BELGIUM	NETHERLANDS
Haul selection	One in two *Pro: sampling is representative for the trip; ensures equal distribution dark/light; high sample intensity; Statistically sound; systematic sampling* *Con: not flexible, high workload*	At least 25% of hauls *Pro: flexible* *Con: –low coverage for high between–haul–variation trips* *–not systematically drawn (sampling might not be representative for the trip; no equal distribution dark/light)*
Bridge information	position and time of haul	position and time of haul Skipper and observer agree on estimation of total catch *Pro: enhances the acceptance by the industry of the reported data* *Con: no independent measurement* Retained part of the catch by species *Pro: enhances the acceptance by the industry of the reported data* *Con: no independent measurement*
Sample selection	crew sorts marketable fish: observer determines total weight of all commercial species; subsample for length measurements of selection of species	Retained part of the catch: for a selection of species, at least 10 subsamples for length measurements are taken during the trip: in each haul one species is selected for sampling.

	Discards: total weight of all commercial species; subsample for length measurements of selection of species *Pro: larger sample size possibly resulting in more precise estimates* *Con: non–commercial invertebrates and non–commercial fish species information is missing*	Discards: summed sample of 5x 10 litre during processing: length measurements for fish and counting numbers by species in benthos fraction *Pro: additional information on non–commercial invertebrates and non–target fish species* *Con: No length frequency distribution on haul level of target species*
Biological sampling	5 fish per cm–size class per species per area for a selection of species (different selection set for landings and discards)	Retained part of the catch: no samples for age reading and other biological parameters Discards: otoliths are taken of 5 fish per cm–class per species per area (for 5 commercial species)
Raising to haul level	For discards and landings: ratio of total weight and subsample weight is used to estimate the total number of discards per cm–size class for selection of species. *Pro: Raising factors by species is better than unique raising factor obtained in a multispecies sample.*	For discards: raised to volume on a haul level. Total catch – total landings = total discards. For landings: Sampled weight is raised to total landing weight on trip level *Con: raising to haul level is not possible*
Marine Mammals/ Protected/ rare species	Protected/rare species: not included in protocol *Con: protected/rare species information is missing*	Protected/rare species: not included in protocol *Con: protected/rare species information is missing*

Comments: The Dutch sampling protocol describes that at least 25% of the hauls must be sampled. This is described as an absolute minimum to give guidance to observers on board commercial vessels. In general all (experienced) observers cover at least 80% of the hauls. A limited number of hauls are sampled during night-time, because of legal constrictions concerning maximum working periods for employees, however, this does not mean that no samples are taken after sun set.

The collection of otoliths in the Belgian on board sampling scheme is restricted to certain species, because retrieval of otoliths effects the quality and hence the value of the fish in some species.

Because a significant part of beam trawl catches consists of benthic invertebrates and non–commercial fish species the Dutch scheme includes additional sampling of these invertebrates and non–target fish species, and, therefore, creating the opportunity to take secondary objectives (i.e. create a complete overview of discarding in beam trawling) into account additional to DCF obligations.

4.2 Discussion

There are several differences in on board protocols in the three comparisons. A general issue is the estimation of the total catch vs. the measured logbook recorded (and checked) landings.

4.2.1 Haul selection

In the freezer trawler case, all hauls are sampled. In the Spanish – Portuguese comparison the Spanish protocol also targets the coverage of all hauls.

Differences regarding haul selection procedures are related to the expected trip duration (work load), and sampling method: in case of the Portuguese protocol, OTB_DEF trips are generally 1 day long; before 2011 observers sampled all hauls (census) and from 2011 onwards observers sample systematically (every other haul) with improved characterization of each haul (larger size of catch sample). In the Spanish protocol, a haul census is only possible for short trips: vessels involved exclusively in the target métier (monovalent vessels) perform mainly one day trips, whereas polyvalent vessels usually perform larger trips. It can be summarized as "logistic constraints", which may be caused by slight differences between the Spanish and Portuguese métiers.

In the Dutch beam trawl sampling protocol, at least 25% (although, see comments Table 4.3) of the hauls are sampled vs. 100% coverage in the Belgian protocol.

Sampling all hauls makes sampling representative of all fishing grounds and discard patterns in the fishing trip. However, there is a potential danger in trying to cover all hauls as some hauls may be preferably missed (e.g. the night hauls composition may be different) and such methodology may not be applicable in both longer and smaller trips. Systematic sampling (every other haul) provides a compromise but in small trips may lead to oversampling of some areas/discard patterns and under–sampling of others. Such oversampling/under–sampling will happen randomly and is thus statistically sound. However, it may increase variance and should be compensated with an improved number of observed trips. Another matter of concern is a tendency to reduced sampling of hauls presenting smaller catches (in all fisheries?). In the freezer trawler case missed catches are for example likely to be small, as these catches are quickly processed and therefore easily missed. In addition – in pelagic fishery – small catches are more likely to contain a mixture of species, because the targeted schools of fish were missed.

A limited number of hauls sampled during night-time and the lack of an instruction for a systematic sampling approach in the Dutch beam trawl protocol may cause a bias because the composition of the catches are known to differ between day en night (at night more sole; at day more plaice).

4.2.2 Bridge information

Location and weather state information is taken in all protocols from the bridge. In addition, in the German–Dutch Mauritanian case, the Germans take also the landings from the bridge. In the Belgian – Dutch comparison, the Dutch observer agrees with the skipper on the amount of the total catch and the retained part of the catch by species. In the Portuguese – Spanish comparison, physical variables related to setting and hauling are taken from the bridge. However, it is noticed that in all protocols where the observer "independently" estimates the amount of total or retained catch, it is most likely that the observer "calibrates" his result with the crews judgment and/or administration. An observer needs to have some reference and the tools to independently estimate the catch are often limited (e.g. plastic boxes). The use of an estimated total catch vs. the measured (and checked) logbook landings, may cause bias when the catches are raised to haul level (see raising to haul level).

4.2.3 Sample selection

Spanish observers collect samples after the catch was split into retained and discarded fractions, whereas the Portuguese observers collect samples directly from unsorted catch and use the crew's sorting criteria in each haul in the determination of the retained and discard fractions. Similarly, German and Mauritanian observers sort the total catch into a discarded and landed fraction by the observer. As put down in the description of the Mauritanian sampling protocol (Annex 8): this procedure seems to work "in general". However, more complex sorting behaviours, such as the crew keeping and discarding some specimens/species of the same length classes – for example slightly damaged or lean specimens – or the fishers discarding criteria changing throughout the haul – may be missed this way. In other words, the main drawback of this approach is that it tends to assume a knife–edged probability of discarding. Even if criteria change from haul to haul, at within–haul level this may be an oversimplification of the real sorting performed by the crew on the catch (that may be more haphazard) and lead to estimates of total or specific discard ratios that are biased. In contrast, the Spanish procedure gets more complex as the observer must perform as many samplings in retained fractions as the number of species retained (sorted on different boxes by the crew). It seems likely that in comparatively high discards rates, the method is straightforward. However, in case of low discards rates (e.g. 1–3%), discards may be completely missed. In addition, especially in low discards rates, the precision of the discards compositions is very low.

In the Belgian – Dutch beam trawler comparison, there is a striking difference in the sample selection. The time spent on the sampling of non–commercial fish species and invertebrate (benthos) species goes at the cost of the sampling of commercial species. As non–commercial species and invertebrates are not part of the DCF, it seems that the sampling protocol is created for other studies as well. In the Dutch protocol, in every sampled haul one commercial species is chosen for LF sampling.

4.2.4 Biological sampling

In the freezer trawler case, the Dutch representative sampling of the catch (25 random specimens; historical background) differs strikingly from the German sampling of 10 specimens per size class. The precision of this approach depends on the number of samples taken: this is not clear from the protocol as described here and is therefore not discussed here. However it is noticed that the samples are not used to directly raise to fleet level, but as age–length key, which may result in the under representation of larger and smaller size classes.

The biological sampling in the sampling of the beam trawler differs between the Belgian and Dutch. In the Dutch on board sampling biological samples are not taken in the retained part of the catch. The reason is that the Dutch have an extensive market sampling programme to collect biological information from landed catch, whereas in Belgium the DCF on board sampling scheme itself provides the market samples.

4.2.5 Raising to haul level

A difference between the protocols is that in the Spanish case study, the auxiliary variable used for raising samples to haul level relies on skippers evaluation of total catch while in the Portuguese case the auxiliary variable used in raising is the retained catch and the data is collected by the observer during the catch storage. Using the skippers data may include voluntary or involuntary bias (underestimation), which linearly affect the discard raising to haul level. In practice the method assumes

that the more landings the less discards and skipper's estimate puts an upper limit on total catch and, consequently, on discards . On the contrary, raising by retained catch, is independent from skipper's opinion (and hence provides for unbounded estimates that are less voluntary and involuntary biases) but it also implies that the more the landings the more the discards. Alternative methods for estimating the total catch have been tested, such as estimations from the filling percentage of the hole (Santos *et al*., 2008), another promising alternative is the estimations obtained using wires tension information when hauling the codend.

The bias created by the rough estimation of the total catch in the Dutch sampling scheme adds up to bias created by the estimation of the discards fraction. The found difference of 10% between the two observers estimates of the total catches on the freezer trawler trip may be caused by this uncertainties. Although this difference is considered to be small in the context of the trip, it may play a greater role when raised to fleet level, since the size of the catches in this fisheries are relatively large. However, it was noticed during the discussion on this topic during the meeting, that the Dutch observers are not completely oblivious concerning their knowledge of the landings: in fact they are well aware of the contents of the crew's logbook, and are encouraged to "calibrate" earlier estimates during the trip of the total catch to actual recorded landings. This is an example how comparison between protocols are hampered by "hidden" details in the protocol.

Differences are also spotted in the species weight measurement in the Spanish – Portuguese sampling scheme. The preferred tool in Spanish protocols is the length–weight relationship, whereas Portuguese only use this method under harsh weather conditions and mostly rely on hanging balance measurements for weight determination. Using the length–weight relationship could be seen as a way of saving time during the sampling process and avoiding the variability of the measurement of weight when using hanging balances under harsher weather conditions. This translates into larger discard samples per unit time and higher precision. The main drawback is the fact that species weight is estimated. Inadequate length–weight relationship may produce bias when estimating species weight in discard sample: This source of bias may be present in case of the Spanish protocol which uses yearly l/w relationships with no consideration on species seasonal l/w variation. In the Portuguese case, estimates may be sensitive to variability of hanging balance measurements (with ship movement) which may overall lead to less precise estimates.

4.2.6 Recording of Marine mammal, rare and endangered species

In the freezer trawler case, the recording of cetaceans is not part of the sampling protocol on in the German sampling scheme. The Dutch protocol includes cetaceans. The reason for this is that the required monitoring of cetaceans under EU Regulation 812 (EC 2004) is implemented in the DCF observer scheme. The sampling of other PETS (Protected, Endangered and/or Threatened Species) is described in the observer manual, but is labeled as "not mandatory", which leaves the possibility for the observer to not record this bycatch. Germany–North Sea/Atlantic does not have a separate observer program for the monitoring of under regulation 812/2004. (In contrast to Germany NS/Atlantic, Germany–Baltic has included this in the sampling protocol).

The Spanish and Portuguese protocols, includes recording of marine mammals in, at least, the sampled hauls. However, in unsampled hauls the protocols rely mostly on fisher's information.

In the Belgian – Dutch beam trawler protocols have no instruction concerning recording of PETS. Sampling of PETS is part of the Dutch manual, part is labeled "not mandatory". A problem with the monitoring of beam trawl is that bycatch rates are extremely rare, possibly due to the low vertical opening of the trawl. Even in case of a bycatch, the observers and crew tend to judge that specimens of mammals and birds in the catch must have been dead at the time when they ended up in the trawl (see also Section 6).

4.3 Summarizing comments

The protocols that have been compared here have differences. The extent of the differences is not always clear. A main topic in all sets of comparison is whether a catch sample is sampled prior to processing (i.e. sampling of catch and split in discarded and landed part of the catch by the observer) or whether a sample is taken after processing (i.e. independent samples of discards and landings after processing of the catch by the crew). Another important issue is the independence of discard estimates and total catch estimates.

Possibly splitting the catch by the observer is sometimes (in case of high discards rates) an accurate method. On small vessels it may sometimes be the only feasible approach. Under high discard situations, an estimate of the total discard of large hauls is difficult to get because there is mostly not enough space and time to collect and weigh or even to estimate the total discard. When discard rates are low the method lacks precision and discards of rare species may be missed. However, then experienced observers can easily take a discard sample in addition to the catch sample.

Independent estimates of discards, either obtained from sample or from the total catch from experienced observers, are probably more accurate than a subjective estimate from a skipper. On the other hand, skipper's expert judgment and life–experience training in catch estimation may, at instances, compensate this. Member countries are encouraged to investigate possible biases caused by this method.

Not having a haul selection protocol, does not ensure that the selected hauls are representative for the trip. A more systematic selection procedure (e.g. "one in two") is preferable.

As expected, the differences in the on board protocols are due to a mixture of logistic constraints (e.g. interaction with the crew, restrictions in handling of the catch because the value may become lower), legal constrictions (e.g. maximum allowed working periods of employees) and other causes (e.g. financial restrictions, unknown/traditional/historical reasons). However – surprisingly – the differences as found in the beam trawler comparison are probably caused by dissimilar research objectives between countries: the including collection of data for other studies/program vs. stringently follow DCF regulations – In the beam trawler case, the Dutch sampling includes invertebrates and non–commercial species apparently for the use in other programmes than the DCF.

PETS (Protected, Endangered and Threatened Species) sampling is carried out to a certain degree: marine mammals/dolphins in the Portuguese and Spanish demersal trawl sampling schemes. Cetaceans in the Dutch– and all large bycatches in the Mauritanian freezer-trawler sampling schemes. PETS sampling is not included in the DCF. However, in the context of future EC targets (Speech Damanaki on 210911 and EC, 2011), it is recommended to update discards sampling protocols (see also Section 6).

Many of the potential biases have not been quantified, so it is hard to compare methodologies. A joint project involving a set of mixed–country observer trips and full catch evaluation would be highly beneficial as it would provide concrete data on biases and, possibly, increase our knowledge of the uncertainty arising when pooling discard information of different countries. These studies potentially form a basis towards standardization of methodologies between member states. . These studies need to focus on causes of potential bias like skippers estimates vs. observer estimates, sample selection (catch vs. landings + discards), raising procedures, use of length–weight relationships vs. hanging balances.

Fishframe currently does not have reference on how estimates are obtained. With increased standardization, such information should be included as estimates are obviously biased because of different national sampling approaches.

5 Standardized reporting of results of sampling designs (ToR5)

Due to time constraints and a restricted number of participants, terms of reference (5) were only addressed during one of the plenary sessions (see Annex 2 for agenda).

Before the start of the meeting participants had the possibility to submit case studies or examples of reports used to communicate discard programme results at a national level. Three case studies or examples of reports were submitted:

- Denmark: draft version of the report on "Danish Sampling of Commercial Fishery – Overview with special attention to discards 2012 data".
- Netherlands: "Discard sampling of Dutch bottom–trawl fisheries in 2009 and 2010. CVO report: 11.008".
- Ireland: "Atlas of Demersal Discarding – Scientific Observations and Potential Solutions".

In plenary the three reports were evaluated. The submitted cases all present discard estimates by species at a national level; two of the reports present discard estimates at fleet level, one report presents estimates of average discard rates by hour. All three reports briefly report on bias indicators (mainly spatial coverage) and give only a rough idea of the performance levels of the sampling programmes.

An important discussion point was the dilemma on how to report to end-users: Write specific reports for a specific group of end-users e.g. managers, policy-makers, industry, etc. or write a general report useful for as many as end-users as possible. Raising procedures and potential differences caused by different procedures were discussed and the study group referred to the Workshop on Discard Raising Procedures (ICES 2007). Presenting average discard rate by effort unit (e.g. hour) or haul could be useful to compare discards of species by for example métiers or gear, however, the study group was cautious for direct comparison between estimates, when end-users not take the variability of average discard estimates into account, e.g. variance, number of samples taken, etc.

During the plenary session, the study group acknowledged that the current description of ToR 5 was not sufficient in capturing the general idea of a discard report: Presenting results of discard estimates at a fleet, métier or national level. The current context of the ToR only refers to results of sampling designs and, therefore, only refers to reporting of results that indicate the level of performance of sampling programmes, e.g. refusal rates, bias indicators, sampling coverage, etc.

Within the context of reporting the performance of sampling designs the study group agrees that: To be able to evaluate and compare performance levels of (national) sampling programmes, one reporting standard is essential. Because, the development of quality standard levels is still in progress, ToR (1), the study groups agrees to deal with standardized reporting of results of sampling designs in more detail during the next meeting.

6 WGBYC (Working Group on Bycatch of Protected Species)

WGBYC requested by means of a presentation by Bram Couperus to update the Table with running sampling schemes (Table 6.2) with information on the whether the protocols used are recording protected, endangered and/or threatened species (PETS).

Tasks of WGBYC include the review and collate bycatch estimates of protected species under EU regulation(s) and Protected, Endangered, Threatened Species (PETS). WGBYC was driven mainly by EC Regulation 812/2004 on cetacean bycatch, and by the Habitats Directive that requires bycatch monitoring of "Annex IV" species.

WGBYC and WGREV812 raised a lot of issues concerning Regulation 812 (ICES, 2011) which were passed via ICES to the commission. As a result, in 2011 the Commission reviewed the Regulation (EC, 2011) . In this review the main issues were mentioned: (1) the fact that Reg. 812 only dealt with Cetaceans and not with other endangered and protected species, (2) The areas and métiers to be monitored were inadequate, (3) inadequate requirements concerning the use of mitigation devices ("pingers"), and (4) the fact that the monitoring is expensive and some member states did not have the resources to implement the Regulation. The Commission did not see possibilities to amend the Regulation, but instead aims at the future implementation of monitoring of a broader range of PETS (instead of only cetaceans) in the DCF. In the presentation the members were asked to take this future development into account in the development and the design of sampling schemes and on board sampling protocols.

The group agreed that collection of additional data on PETS may result in a more efficient use of resources, as observer programmes are relatively expensive. Protocols can be adjusted in some cases to make routine discards sampling more usable for the monitoring of bycatch. The recording of large specimens, like dolphins and seals, should take place on haul level: the observer should observe the hauling of the net and the opening of the codend for each haul.

There are issues that have to be addressed for future successful implementation. This can be divided in issues that may be solved relatively easy (minor issues) and issue that are considered very important and hard to deal with, if possible at all (major issues).

6.1 Major issues

The DCF sampling schemes are not designed to estimate bycatch of endangered and protected species. Implementation of the monitoring is not just a matter of adjusting the on board sampling protocols. In order to estimate bycatch rates by ICES areas, sampling schemes have to be changed on the level of fleet segment and métier selection as well.

Bycatch data collected in current DCF schemes should be treated with care. Raising recorded bycatch to larger areas is often not possible. Third parties may use the data in a wrong way.

In some sampling schemes when recording of dolphins and seals by the observer, fishery crews may become less cooperative, because the bycatch of some PETS – in particular harbour porpoises and dolphin species – draws attention of the public and add to a negative image of the fishery.

Adequate sampling of comparatively rare fish species of small size (e.g. shads) is difficult to implement, possibly it involves sampling of the whole catch instead of taking a sub sample. It also requires flexible sampling, depending of the catch, which is hard to achieve on commercial vessels. Thorough sampling of rare fish species may be impossible due to a lack of available resources.

6.2 Minor issues

Protocols should include a list of rare species that should be recorded during trips. These species should have a code in the institute database and code lists should be available to the observer who enters the data in the database. Currently lot of PETS data are lost due to the fact that there is no reference code for the national database!

Following from the above, it is also important to have reference codes for international databases.

Identification of rare bycatch is often a problem, because observers are not familiar with all the species involved. This can be dealt with, by limiting the number of species, accept identification by group (for example in both shad species which are difficult to distinguish), collection of specimens for further investigation ashore, provision of identification guides and by taking pictures. An extensive manual with clear instructions and which includes a Section for the identification of rare species is very important.

Rare species are often considered to have been dead already during the bycatch. This seems to happen often in the sampling on board beam trawlers where observers assume that is impossible to catch a large, fast swimming animal like a harbour porpoise, because the low vertical opening of the trawl.

The group emphasized that a clear list of PETS is required. In situations where it is possible to sample more than a (few) basket(s), this may give the observer a clue which species the catch should be scanned for. A list is also required in a number of sampling schemes where only a selection of (commercial) species is recorded. Table 6.1 presents such a list of species. It is based on the species which are protected under national law and should be protected under the Bird – and habitat directive.

Table 6.1. Species that should be monitored after (future) integration of Protected, Endangered, Threatened Species (PETS) sampling in DCF sampling schemes.

Group	**Identification by...**
Cetaceans	species
Seals	species
Birds	species
Turtles	species
Shads	2 species: *Alosa alosa*; *Alosa fallax*; (or Alosa sp.)
Lampreys	2 species: *Lampetra fluviatilis*; *Petromyzon marinus*
Sturgeon	1 species: *Acipenser oxyrinchus*

Table 6.2. Sampling schemes for Protected, Endangered, Threatened Species (PETS).

				cetaceans	seals	birds	turtles	shads		lampreys		sturgeon	
Member–state	**Type**	**Métier(s) or fleet segment(s)**	**Vessels**	**haul level**	**haul level**	**haul level**	**haul level**	**haul level**	**sample level**	**haul level**	**sample level**	**haul level**	**sample level**
BE	onboard	TBB_DEF	all	N	N	N	N	N	N	N	N	N	N
BE	self–sampling	TBB_DEF	VIIf,g	N	N	N	N	N	N	N	N	N	N
DK	onboard	all métiers in DCF	all	Y	Y	Y	Y	N	Y	N	Y	N	Y
DK	self–sampling	GNS	all										
EN	onboard	all métiers in DCF	all										
ES	onboard	OTB; PS	Mediterranean	Y	Y	N	N	N	Y	N	Y	N	Y
ES	onboard	GTR; LLS	Mediterranean	Y	Y	N	N	N	Y	N	Y	N	Y
ES	onboard	OTB	Atlantic (long trips)	Y	N	N	N	N	Y	N	Y	N	Y
ES	onboard	OTB	Atlantic (short trip)	Y	Y	N	N	N	Y	N	Y	N	Y
ES	onboard	PS	Atlantic	Y	Y	N	N	N	Y	N	Y	N	Y
ES	onboard	PTB	Atlantic	Y	N	N	N	N	Y	N	Y	N	Y
ES	onboard	GNS	Atlantic	Y	Y	N	N	N	Y	N	Y	N	Y
FR	onboard	all métiers in DCF	all										
IRE	onboard	Demersal fleet	all										
IRE	onboard	Pelagic fleet	all										
IRE	self–sampling	Nephrops	all										
NL	onboard	OTB; TBB_DEF; SSC	validate ref..fleet	N	N	N	N	N	Y	N	Y	N	Y
NL	onboard	TBB_CRU	all	N	N	N	N	N	Y	N	Y	N	Y
NL	onboard	OTM; PTM	all in EU waters	Y	N	N	N	N	Y	N	Y	N	Y
NL	self–sampling	OTB; TBB_DEF;SSC	Reference fleet	N	N	N	N	N	Y	N	Y	N	Y
NOR	self–sampling	demersal fleet (High seas)	Reference fleet										
NOR	self–sampling	demersal fleet (Coastal)	Reference fleet										
PT	onboard	OTB_CRU; OTB_DEF	all										
PT	onboard	GTR_DEF; GNS_DEF; LLS_DWS	all										
PT	onboard	PS_SPF	all										
PT	onboard	TBB_CRU	all										
SCO	onboard	OTB_DEF; OTT; PTB; OTB_CRU; SSC	all										
SCO	onboard	OTB_CRU; OTT_CRU	all										
SWE	onboard	OTB_DEF; OTB_CRU	all	Y	Y	Y	Y	N	Y	N	Y	N	Y
SWE	self–sampling	GNS+GTR+LLS	all	N	N	N	N	Y	Y	Y	Y	Y	Y
PL	onboard	all métiers in DCF: OTB, GNS, LLS,OTM	selection										
GER (Baltic)	onboard	trawlers	all	Y	Y	Y	Y	N	Y	N	Y	N	Y
GER (Baltic)	onboard	GNS	all	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
GER (Baltic)	self–sampling	all métiers in DCF	all	N	N	N	N	N	Y	N	Y	N	Y

References

Amending Regulation (EC) No 88/98. COM(2011) 578. E. Commission.

Cochran, W.C. 1977. Sampling techniques. Wiley, New York, U.S.A.

EC (2004). Council Regulation (EC) No 812/2004 of 26.4.2004 laying down measures concerning incidental catches of cetaceans in fisheries and amending Regulation (EC) No 88/98.

EC (2011). Communication from the Commission to the European

EC (2011). Communication from the Commission to the European Parliament and Council on the implementation of certain provisions of Council Regulation (EC) No 812/2004 laying down measures concerning incidental catches of cetaceans in fisheries and amending Regulation (EC) No 88/98. COM(2011) 578. E. Commission.

ICES. 2008. Report of the Workshop on Methods to Evaluate and Estimate the Accuracy of Fisheries Data used for Assessment (WKACCU), 27–30 October 2008, Bergen, Norway. ICES CM 2008\ACOM:32. ICES, Copenhagen.

ICES. 2010. Report of the Joint ICES – STECF Workshop on methods for merging me–tiers for fishery based sampling (WKMERGE), 19–22 January 2010, Copenhagen, Denmark. ICES CM 2010/ACOM:40. ICES, Copenhagen.

ICES. 2011. Report of the Working Group on Bycatch of Protected Species (WGBYC 2011). ICES CM 2011/ACOM:26: 75.

ICES. 2011. Report on the Workshop on Practical Implementation of Statistical Sound Sampling Programmes. In prep.

ICES. 2012. Report of the Planning Group on Commercial Catches, Discards and Biological Sampling (PGCCDBS 2012), 30 January–3 February 2012, Rome, Italy. ICES

CM 2012/ACOM:50. 163 pp.

Jardim, E., Prista, N, Fernandes, A.C., Silva, D., Ferreira, A.L.,, Abreu, P., Fernandes, P., 2012. Manual de procedimentos a bordo: arrasto de fundo com portas. Relat. Cient. Téc. Inst. Invest. Pescas Mar, nº 55, 20 p. + Anexos. (http://www.inrb.pt/fotos/editor2/ipimar/rct.serie_digital/relno55_final.pdf)

Parliament and Council on the implementation of certain provisions of Council Regulation (EC) No 812/2004 laying down measures concerning incidental catches of cetaceans in fisheries.

Santos, J., Victor Lopez, Gonzalo R. Mucientes & Nélida Perez. 2008. Non–invasive methods for fishing discard estimations onboard industrial trawlers. XV Simpósio Ibérico de Estudos de Biologia Marinha,8–12 September 2008. Poster Nº28, theme 6.

Annex 1 : List of participants

Name	Address	Phone/Fax	E-mail
Bram Couperus	IMARES PO Box 68 1970 AB Netherlands	Phone + 31 25 55 64 6444	bram.couperus@wur.nl
Edwin van Helmond (Chair)	IMARES PO Box 68 1970 AB Netherlands	Phone + 31 25 55 64 6444	edwin.vanhelmond@wur.nl
Jon Elson	CEFAS Lowestoft Pakefield Rd. NR33 0HT Suffolk, UK	Phone: +44 152 52 45 11	jon.elson@cefas.co.uk
Josefine Egekvist	DTU Aqua Charlottenlund Slot Jægersborg Alle 1 2920 Charlottenlund Denmark	Phone +45 33 96 34 38	jsv@aqua.dtu.dk
Kisten Birch Haakansson	DTU Aqua Charlottenlund Slot Jægersborg Alle 1 2920 Charlottenlund Denmark	Phone + 45 35883037	kih@aqua.dtu.dk
Nuno Prista (note: via web conference)	IPIMAR Avenida de Brasilia 1449–006 Lisbon Portugal	Phone + 351 213027185	nmprista@ipimar.pt
Ryszard Grzebielec	NMFRI Ul. Kollataja 1 81–332 Gdynia Poland	Phone: + 48 587356226	rysiek@mir.gdynia.pl
Juan Santos	IEO Vigo Cabo Estai Canido 36200 PO Box 1552 Vigo Spain	Phone: + 34 986492111	juan.santos@vi.ieo.es
Jon Ruiz Gondra	AZTI SUK Txatxarramendi Ugartea E – 48395 Sukarrieta Spain	Phone: +34 9460 29400	jruiz@azti.es
Krzysztof Radtke	NMFRI Ul. Kollataja 1 81 – 332 Gdynia Poland	Phone: +48 58 7356 223	radtke@mir.gdynia.pl
Uwe Krumme	JHVTI vTI OSF Rostock Germany	Phone: + 49 3818116148	uwe.krumme@vti.bund.de

Sofia Carlshamre	SLU DAR IMR SE – 45321 PO Box 4 Lysekil Sweden	Phone: +46 52318772	sofia.carlshamre@slu.se
Katja Ringdahl	SLU DAR IMR Turistgatan 5 453 30 PO Box 4 Lysekil Sweden	Phone: +46 10 478 4043	Katja.Ringdahl@slu.se
Sofie Vandemaele	ILVO Ankerstraat 1 8400 Oostende Belgium	Phone: +32 59 569883	Sofie.Vandemaele@ilvo.vlaanderen.be
Peter Clark	Marine Scotland FRS Marine Lab PO box 101 Aberdeen AB119DB, UK	+ 44 1224295400	p.clark@marlab.ac.uk
Alastair Pout	Marine Scotland FRS Marine Lab PO box 101 Aberdeen AB119DB, UK	+ 44 1224 876544	a.pout@marlab.ac.uk

Annex 2: Agenda

SGPIDS – Study Group on Practical Implementation of Discard Sampling plans

Agenda 18 June – 22 June 2012

Monday (18/6)	13:00 – 15:00	Introduction: ToR's, agenda, organize sub-groups
	15:00 – 15:30	Break
	15:30 – 17:30	Subgroups
Tuesday (19/6)	09:00 – 10:30	Plenary: WGBYC
	10:30 – 12:30	Plenary: presentations Uwe Krumme and Alastair Pout
	12:30 – 13:30	Lunch
	13:30 – 17:30	Subgroups
Wednesday (20/6)	09:00 – 10:00	Plenary: presentations Jon Ruiz
	10:00 – 12:30	Subgroups
	12:30 – 13:30	Lunch
	13:30 – 16:00	Subgroups
	16:00 – 17:30	Plenary: group A and C
Thursday (21/6)	09:00 – 10:30	Plenary: group B and ToR 5
	10:30 – 12:30	Subgroups
	12:30 – 13:30	Lunch
	13:30 – 17:30	Subgroups
Friday (22/6)	09:00 – 10:30	Plenary: discuss ToR's 2013
	10:30 – 12:00	Review draft report
	12:00	End meeting

Terms of Reference for 2013

The **Study Group on Practical Implementation of Discard Sampling Plans** (SGPIDS), chaired by Edwin van Helmond, The Netherlands, will meet in SLU DAR IMR, Lysekil, Sweden, 24–28 June 2013 to:

e) Review implemented discard sampling frames and vessel selection procedures;

f) Evaluate the quality indicators for discard sampling programmes, as defined at SGPIDS 2012;

g) Develop and define quality indicators for discard on board sampling protocols for commercial vessels;

h) Review the reporting of results of national sampling designs (and discard estimates) and references to regional databases;

i) Continue to collaborate with WGBYC on integrating data on protected species with relevant discard survey data.

Supporting Information

Priority	
Scientific justification	The coordination and planning of discards sampling is part of the tasks of PGCCDBS and more regionally of the Regional Co–ordination Meetings (RCMs). However, these groups lack expertise, scope and time to deal with the practical aspects of discard sampling. This meeting can build upon the outcome of WKDRP, WKEID, WKACCU, WKPRECISE, WKMERGE and WKPICS with regard to the tools and methodology used to analyse discard data and its sampling bias.
Resource requirements	Particpants should bring descriptions of samping protocols to the meeting. Resources, i.e. case studies, working documents and/or published work, are required to study on board sampling techniques to define appropiate quality indicators (ToR c). Reports of results of sampling designs (and discard estimates) (ToR d).
Participants	The Group is normally attended by some 10–20 members and guests.
Secretariat facilities	Meeting facilities, including sharepoint and secretarial support.
Financial	No financial implications.
Linkages to advisory committees	ACOM
Linkages to other committees or groups	WKPICS, RCMs, WGBYC, PGCCDBS.
Linkages to other organizations	EC (DCF reform 2012–2013).

Annex 3: Recommendations

Recommendation	Adressed to
1. The construction and use of age–length keys (ALKs) and length–weight keys (LWKs) is a critical stage in estimating numbers-at-age for stock assessment working groups. However, bias and error in the application of ALK and LWKs, is poorly understood and a rather neglected topic in the raising of discard data. Many questions were raised during the SGPIDS meetings: what is the bias introduced by the use of survey–based and landings–based ALKs as a proxy to discard ALKs, what is the spatial and temporal resolution of the ALKs used by Member States, how should ages be combined (i.e. as a weighted or unweighted sample), what are the consequences of relying on LWKs, which in some cases, date back up to 30 years? Additionally, many of these issues apply to the raising of landings data as well. Therefore, SGPIDS strongly recommend PGCCDBS to address the problems at some wider forum in the near future. A new platform (workshop/study group) may be created for the exchange of expertise on ALKs and LWKs in order to improve the quality of data used in stock assessments. The corrected methodology for using ALKs and LWKs must be applicable by 2014 with the implementation of the reformed DCF (2014–2020).	PGCCDBS
2. Simple "goodness-of-fit" statistics that could be used to interpret comparisons of various metrics in total and study populations, for (at–sea) sampling schemes, would be a considerable advantage for the development of good quality indicators for discard samling programmes. Suitable statistics are likely to be non–parametric (hence no requirement to make assumptions about the distribution of the underlying data) and need to be appropriate to the measurement scale of the data (i.e. ordinal of interval scales). Kruskal–Wallis tests, spearman rank correlations, $\chi 2$ tests are possible candidates but the study group felt that such considerations would benefit from input from expert groups such as WKPICS.	WKPICS2
3. Similar to the previous recommendation the study group would benefit from input from an expert group, such as WKPICS, to identify and quantify bias caused by exclusion of objects. For example in comparissons of total population and study population.	WKPICS2
4. Currently some members states are using a weighted vessel list to ensure that the more active vessels are sampled. This is an example of unequal weighted probability sampling and needs to be accounted for in the raising procedure. Other member states are excluding the less active vessels from their selection lists although these less active vessels still remain within the frame. This decision may be based on an assessed acceptable bias. WKPICS has been reviewing this issue and needs to provide clear guidance on best practice.	WKPICS2

Annex 4: German– Dutch case study

A case study on different sampling procedures for the sampling of the pelagic freezer trawler fleet between the Netherlands and Germany

Jens Ulleweit and Edwin van Helmond

Introduction

The European Data Collection Framework (DCF) explicitly fosters regionally harmonized sampling schemes for commercial fishing activities. National sampling programmes encompass the collection of métier– and stock–based variables and are co–ordinated between countries in order to ensure an efficient use of employed sampling effort by task–sharing. Several workshops have addressed the harmonization of data collection and data processing across research institutions in Europe and worldwide. In practical terms, however, there are often differences in sampling procedures between countries that are largely being ignored but may have a significant effect on the estimation of input parameters for stock assessment.

More recently, PGCCDBS advised that SGPIDS should evaluate differences and possible data deviations caused by different national sampling schemes within the same fishing fleet and should aim to synchronize on–board sampling protocols.

The European pelagic freezer trawler fleet is regularly sampled both by German and the Netherlands. A comparison of the sampling procedures for this fleet is shown in this working document. Furthermore, taking advantage of the accidental double manning of a German– flagged but Dutch–owned pelagic freezer trawler with observers of both countries in January 2010, results derived from data collected by different sampling procedures for the same trip are presented and discussed. These findings were derived from a poster presentation for the Fishery Dependent Information Conference in 2010 (Stransky et al. 2010).

The sampled fishery

The European pelagic freezer trawler fleet consists of large vessels from approx. 70 to >120m length with 3200 to >11000 hp. Depending on fishing season these vessels are targeting small pelagic species like herring, mackerel, horse mackerel and blue whiting for human consumption. The catch of a single haul can easily reach more than 100 t which is pumped on–board and can be stored in refrigerated seawater tanks until processing. The catch is first automatically sorted and manually quality controlled. After sorting the whole fish is deep–frozen in 20kg portions. Holding capacities are from 1200 to >5000 tonnes. The vessels are manned with up to 30 crew members.

The sampling methods

The following Table shows the sampling methods by the Netherlands and Germany according to the national observer sampling protocols:

Table 1. List of sampling procedures in the Netherlands (left panel) and Germany (right panel

Netherlands	Germany
1) Estimation and registration of total catch 2) Estimation of discard percentage. 3) Take a sample of the unsorted catch (total sample size: 20–25 kg). This sample includes landings and discards. In order to get a representative sample, sub–samples are taken repeatedly at different moments while sorting the haul. 4) Take a sample of discards (total sample size: 20–25 kg). In order to get a representative sample, different sub–samples are taken repeatedly at different moments while processing the haul. 5) Measuring catch sample: Sort all the fish species and take length and weight measurements for each species. Register the total number by species and length class. 6) Measuring discard sample: Sort all the fish species and take length and weight measurements for each species. Register the total number by species and length class. 7) Age estimations of the unsorted catch: a. Take a sample of the unsorted catch. b. Otoliths from this sample are prepared and analysed. c. The sample of age analysis consists of 'sized' and 'undersized' fish. A sample consists of minimal 3 individuals per length class per area (ICES quadrant). 8) Registration of total landings: Information on total landings is collected at the end of the trip **Protocol of estimating the discard percentage** 1) Take weight sample of discards from the gutter over a certain time period. 2) Take weight sample of catch from conveyer belt over the same time period as the discard sample. 3) Calculate discard percentage from the proportion between the two samples	Check the presetting of the sorting machines for the target species and determination of non–marketable species for the determination of the discards Take a sample of the unsorted catch If the haul is stored in different tanks the sample should be taken from different storage tanks. The total sample size is depending on the target species e.g. herring > 50kg, mackerel >200kg.This sample includes landings and discards. Measuring catch sample: Sort all the fish species into landings and discards components referring to the settings of the sorting machine and take length and weight. Age samples: During at least 2 hauls while the fish is measured 10 fish of each length class (sized and undersized) are separated and frozen for otolith analyses in the institute´s lab. This is done by ICES division. **Estimating the discard percentage** Example: The subsample of 69,04kg contains: 54kg landings of horse mackerel = 78,2% 4,4kg discards of horse mackerel = 6,4% 10,64% discards of mackerel = 15,4% **Estimation of the discard proportion** The observer will get information on the haul size form the ship´s crew on the total landings. Example: Overall landings horse mackerel = 50000kg = 78,2%, therefore the proportion of horse mackerel discards is (50000*6,4/78,2) = 4072kg, the proportion of mackerel discards is (50000*15,4/78,2) = 9852kg. In case the information from the crew is related to the total catch (=landings plus discards) = 50000kg = 100%, the proportion is calculated in the following way:

Example: The sample is taken over a time period of 30 seconds. This results in: – A weight sample of the discards of 2 kg – A weight sample of the catch of 26 kg The percentage discards is calculated by taking de ratio between the discard sample and catch sample: – Percentage discards = (2kg / 26 kg) *100 ≈ 8%	Landing horse mackerel = (50000*78,2/100) = 39100kg Discard horse mackerel = (50000*6,4/100) = 3200kg Discard mackerel = (50000*78,2/100) = 7700kg

Comparison of sampling methods

In January 2010 a German flagged but Dutch owned freezer trawler was accidently manned by observers of both countries. Sampling was conducted according to the national observer instructions (see Table 1). The duration of the trip was 20 days. The vessel was operating in ICES Divisions VIa, VIIb, VIId and VIIj. From 31 conducted hauls during the trip, 24 were sampled by both observers.

The average catch per haul was 49.5 t (range 14–125 t) raised by the German observer and 54.5 t (range 12–150 t) raised by the Dutch observer. The raised landings and discard estimates by ICES Division show minor to moderate differences between both countries (Table 2).

Looking at the species distribution, differences in occurrence can be detected for the infrequently caught species. This can be explained as the observers were not taking their samples at the exactly same place and time. One rare species was determined differently (*Diplodus sargus* or *Pterycombus brama*).

Clearer differences can be noted for the length distributions (Figures 2 and 3, note different scales for landings and discards). Although both graphs for the retained mackerel in Div. VIIb (Figure 3) are showing the same main peak, the German data show more smaller length classes in the landings fraction and the Dutch data show a more widened discard length distribution (Figure 2). These differences can be explained by the incongruent sampling by the observers but are also resulting from the multiple discard recording by the Dutch observer *vs*. the single discard recording by the German observer.

Conclusion

The Dutch and German sampling methods are quite similar. However, while minor to moderate differences were found in catch estimates by weight and numbers and length distributions for the sampled trip, these could play a greater role when raising data to the whole fleet or fisheries.

In view of a possible synchronization of sampling procedures the Dutch method of multiple discard sample takings during the processing of one haul is more effective than the German method.

For stock assessment purposes and for regional co–ordination and synchronization of commercial fisheries sampling, the methodology for catch sampling of all countries

and for all fishing métiers should be documented in detail, in order to detect largely deviating sampling schemes.

Reference:

Christoph Stransky, Jens Ulleweit, Edwin van Helmond, Kay Panten, Harriet van Overzee (2010) Reading the 'small print': Differences in sampling schemes for commercial fishing activities between two countries, Poster to the Fishery Dependent Information Conference, Galway, Ireland, 23–26 Aug 2010

Table 2. Weight and numbers of the total catch, landings, discards and sample by country, ICES Division and species

Country	Division	Species	total Kg	total N	landings Kg	landngs N	discards Kg	discardsN	sample Kg	sample N
D	6A	Scomber scombrus	265000	897006	265000	897006	0	0	439	1424
NL	6A	Scomber scombrus	318000	1095990	318000	1095990	0	0	316	1044
D	7B	Capros aper	531	11989	0	0	531	11989	3	64
NL	7B	Capros aper	1084	24990	0	0	1084	24990	6	113
D	7B	Melanogrammus aeglefinus	856	2212	0	0	856	2212	3	7
NL	7B	Melanogrammus aeglefinus	1876	4680	0	0	1876	4680	0	28
D	7B	Merlangius merlangus	60	188	0	0	60	188	0	1
NL	7B	Merlangius merlangus	211	349	151	0	60	349	1	3
D	7B	Merluccius merluccius	92	461	0	0	92	461	0	1
NL	7B	Merluccius merluccius	158	876	0	0	158	876	0	2
NL	7B	Micromesistius poutassou	14	263	0	0	14	263	0	2
D	7B	Scomber scombrus	246819	884889	246402	882110	417	2779	411	1409
NL	7B	Scomber scombrus	300633	1114711	299976	1109923	658	4788	325	1138
D	7B	Trachurus trachurus	145530	707065	144520	700134	1010	6931	248	1228
NL	7B	Trachurus trachurus	159150	799840	158952	798782	198	1057	132	662
D	7D	Clupea harengus	247256	1503909	246911	1500972	344	2937	407	2481
NL	7D	Clupea harengus	220282	1328814	220030	1326589	252	2224	346	2094
NL	7D	Diplodus sargus	117	467	0	0	117	467	1	2
NL	7D	Loligo	50	71	0	0	50	71	1	1
D	7D	Merlangius merlangus	1680	2898	0	0	1680	2898	2	3
NL	7D	Merlangius merlangus	61	-			61		1	
D	7D	Mullus surmuletus	83	613	0	0	83	613	0	2
NL	7D	Mullus surmuletus	44	271	0	0	44	271	1	2
D	7D	Pleuronectes platessa	99	292	0	0	99	292	0	1
NL	7D	Pleuronectes platessa	102	-	102		0		0	
D	7D	Pterycombus brama	300	2632	0	0	300	2632	0	2
D	7D	Scomber scombrus	4808	29111	1100	3801	3708	25310	36	140
NL	7D	Scomber scombrus	4116	16514	2741	11038	1375	5476	48	201
D	7D	Trachurus trachurus	211916	1427909	208832	1405197	3084	22712	454	2871
NL	7D	Trachurus trachurus	146310	995882	145723	991605	587	4277	282	1757
NL	7J	Capros aper	1	13	0	0	1	13	0	1
D	7J	Loligo	52	724	0	0	52	724	0	4
NL	7J	Loligo	24	250	0	0	24	250	2	19
NL	7J	Melanogrammus aeglefinus	9	118	0	0	9	118	1	9
NL	7J	Merlangius merlangus	1	13	0	0	1	13	0	1
D	7J	Merluccius merluccius	30	28	0	0	30	28	2	2
NL	7J	Merluccius merluccius	30	26	0	0	30	26	2	2
D	7J	Micromesistius poutassou	25	121	0	0	25	121	0	1
D	7J	Scomber scombrus	36008	120821	33552	110092	2456	10729	124	434
NL	7J	Scomber scombrus	48399	165061	44286	156810	4113	8251	145	485
D	7J	Trachurus trachurus	125857	584944	125415	581468	442	3476	242	1142
NL	7J	Trachurus trachurus	105201	488932	104978	487726	223	1206	245	1152

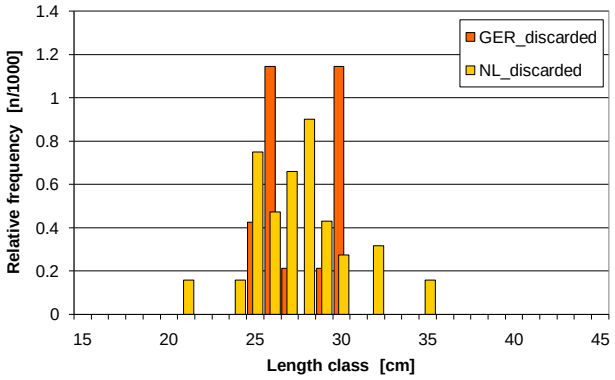


Figure 2. Mackerel in ICES Div. VIIb – Discards, length distributions by country

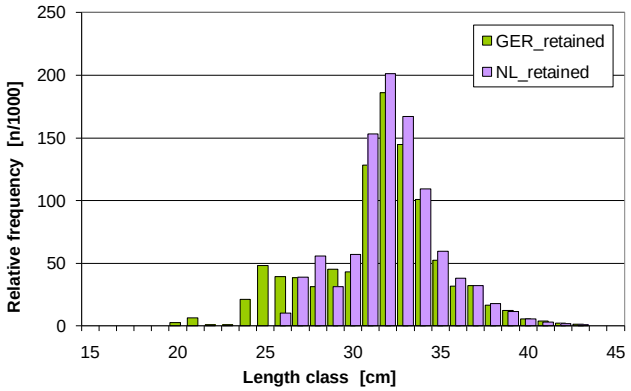
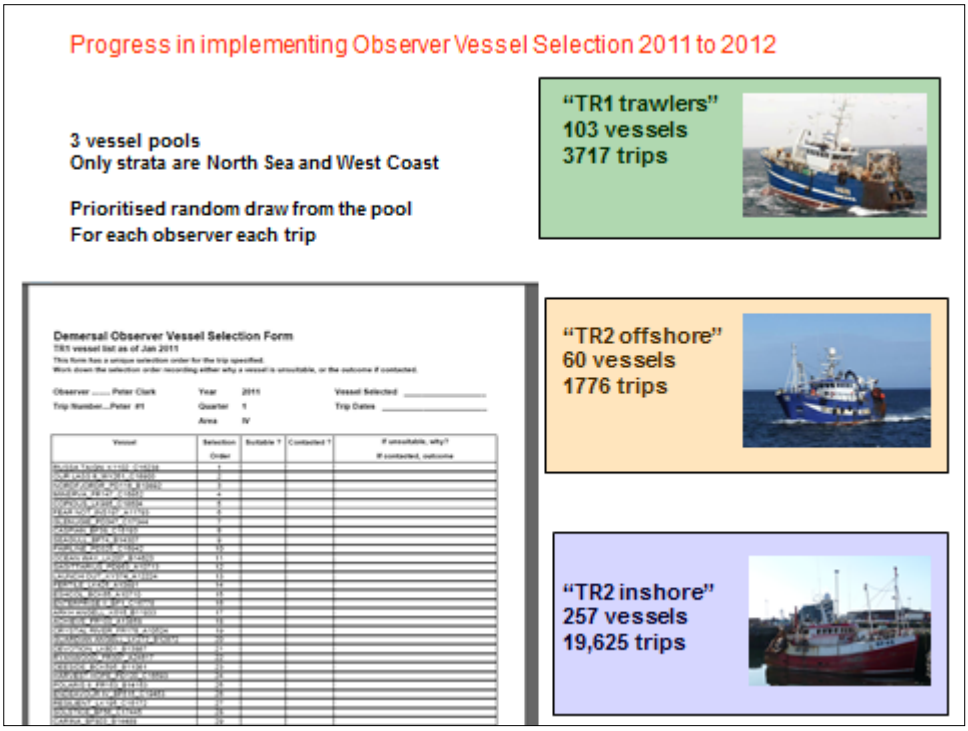


Figure 3. Mackerel in ICES Div. VIIb – Landings, length distributions by country

Annex 5: Progress in implementing Observer Vessel Selection 2011 to 2012 (Marine Scotland Science – Marine Laboratory)

This presentation documents the progress achieved as part of an on–going process, as presented to the study group. It is not as a final definitive statement, or estimate thereof, and as such no data shall be cited without prior approval of the author.



	Q1	Q2	Q3	Q4	T
no. of trips	16	13	14	14	57
missing sheets	3	1	1	3	8
sheets need more work	2	4	4	4	14

57 observer trips
35 selection sheets could be used.
38.5% of trips Observers are not filling in the forms

Quarter	Vessel	No. on list	Observer	Not Contacted	1	2	3	4	5	6	7	8	9	10	Contacted	1	2	3	4	5	6	7	8	9	10	11
1	Harvester PD 95	5	DDS	2								1		1	6			2			3					1
1	Boy Andrew LH 3	23	JW	22	2		1		16	1		2			1											1
1	Harvester A 565	17	PC	15	1		1		10	1			2		2				1							1
1	Rebecca Jansen	0	JM																							
1	Carina B503	23	IP	22			1		21						1											1
1	Ocean Harvest P	3	DDS												3									1	1	1
1	Fair Morn V 9A10	0	JW																							
1	Arcturus UK39	4	DDS												4	1					2					1
1	Celestial Dawn B	0	PS																							
1	Silver Lining III T	7	DG	6					1	2			2	1	1											1
1	Spitfire LH 107	8	RW												8	1			6							1
1	Endeavor IV BF 3	15	MG	14			4		7	1	2				1											1
1	Seagull V B274	75	IP	71			2		67		1	1			4					2				1		1
1	Ardent II INS 127	31	CD	30		1	8	1	4	4	9	2	1		1											1
1	Caledonia SC123	0	JW																							
1	Renown PR245	0	PC																							

New list of summarised outcomes for vessel selection form 08/12/2011

Not Contacted	Contacted
1 No longer/Not in fleet	1 No answer
2 Used this quarter	2 Direct no
3 Vessel unsuitable	3 Indirect no
4 Persistent definite refusals	4 Incorrect area
5 Incorrect area	5 Incorrect target
6 Incorrect target	6 At sea
7 At sea	7 Not fishing
8 Not fishing	8 Unsuitable logistics
9 Excessive travel required	9 Yes but trip did not take place
10 No contact information	10 Yes but trip not sampled
	11 Yes trip sampled

Table 1. Summarised selection form review data Q1 - Q4 2011

Reason No.	Not Contacted	No.	%	Reason No.	Contacted	No.	%
1	No longer/Not in fleet	51	9.11	1	No answer	7	6.19
2	Used this quarter	13	2.32	2	Direct no	1	0.88
3	Vessel unsuitable	59	10.5	3	Indirect no	10	8.85
4	Persistent definite refusals	10	1.79	4	Incorrect area	13	11.5
5	Incorrect area	285	50.9	5	Incorrect target	4	3.54
6	Incorrect target	29	5.18	6	At sea	7	6.19
7	At sea	56	10	7	Not fishing	13	11.5
8	Not fishing	14	2.5	8	Unsuitable logistics	19	16.8
9	Excessive travel required	16	2.86	9	Yes but trip did not take place	4	3.54
10	No contact information	27	4.82	10	Yes but trip not sampled	0	0
				11	Yes trip sampled	35	31
	Total	560	100		Total	113	100

560/673 83% selections ruled out by Observers

285/673 42% for "Wrong Area"

Simple stats from contaced vessels	%
Sampled Successfully (11/total c *100)	31
Refusal Rate (1+2+3+10/total c *100)	15.9
Unavailable to sample (4+5+6+7+9/total c *100)	36.3
Observer preference (8/totalc *100)	16.8
Total	100

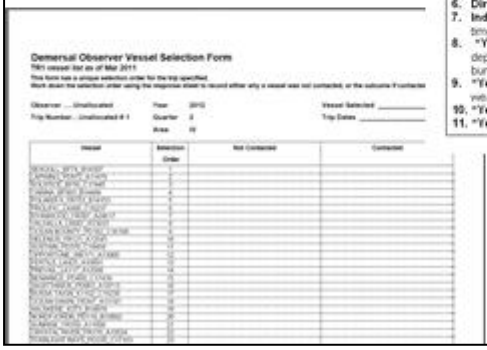
Revised form for 2012

Not Contacted Responses:

1. **No longer / Not in Fleet** – decommissioned, sunk, sold, registered elsewhere
2. **Used this quarter**
3. **Vessel unsuitable** – boat in poor condition, too old or small, skipper/crew issues (explain how this is known)
4. **Persistent definite refusals** – vessel won't take scientists, history of non cooperation
5. **Incorrect area** – ruled out by observer on basis of ICES area being fished (explain how this is known)
6. **Non discard fishery** – ruled out by observer on basis of vessel operating in a non discard fishery (explain how this is known)
7. **At sea** – fishing, oil job (explain how this is known)
8. **Not fishing** – on slip, tied up for repairs, days at sea, quota, weather (explain how this is known)
9. **Excessive travel required** (give explanation)
10. **No contact information**

Contacted Responses:

1. **No answer** – phoned but no answer
2. **At sea** – fishing, oil job
3. **Not fishing** – on slip, tied up for repairs, days at sea, quota, weather
4. **Incorrect area** – vessel not fishing in ICES area required
5. **Non discard fishery** – vessel using pots/creels targeting crustaceans or dredging for scallops
6. **Direct "No"** – skipper or agent said no
7. **Indirect "No"** – skipper or agent gave other reason for declining e.g. no room, not this time but another time
8. **"Yes" but unsuitable logistics** (for observer) – excessive travel required, unsuitable departure date at the time vessel was contacted, genuinely no working room or spare bunk on vessel
9. **"Yes" but trip did not take place** – broke down at harbour, vessel didn't sail due to weather
10. **"Yes" but trip not sampled** – left without observer
11. **"Yes" trip sampled**



Issues

Why observers were not filling in the forms

The number of vessels being ruled out without being contacted

wrong area – stratification

Reliable calculation of a refusal rate

tight well defined selection protocol

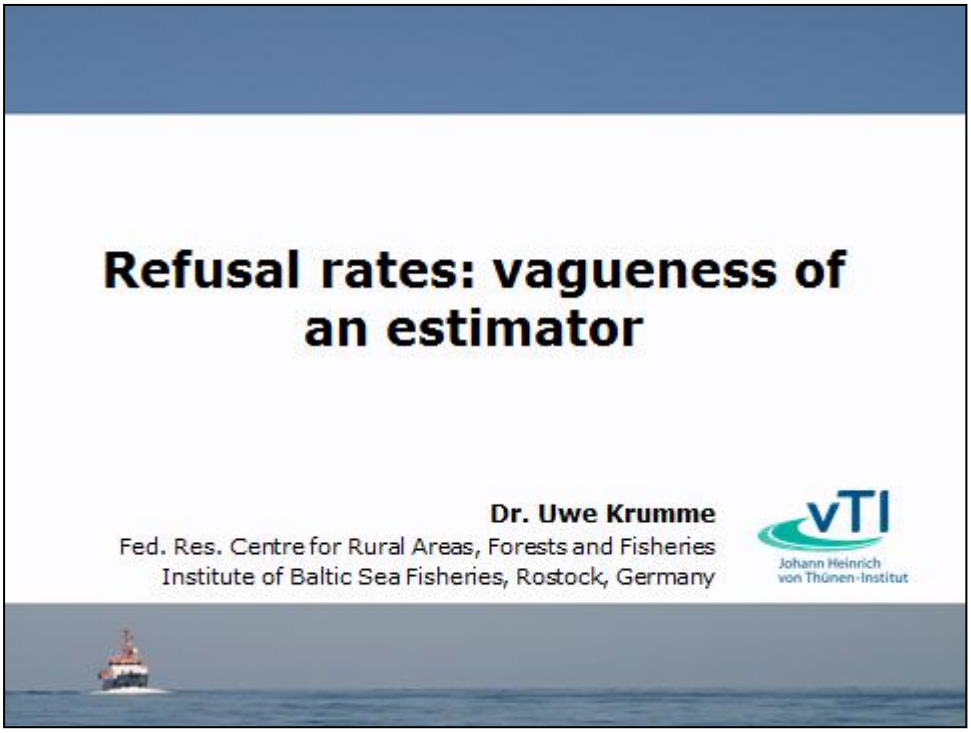
Advantages

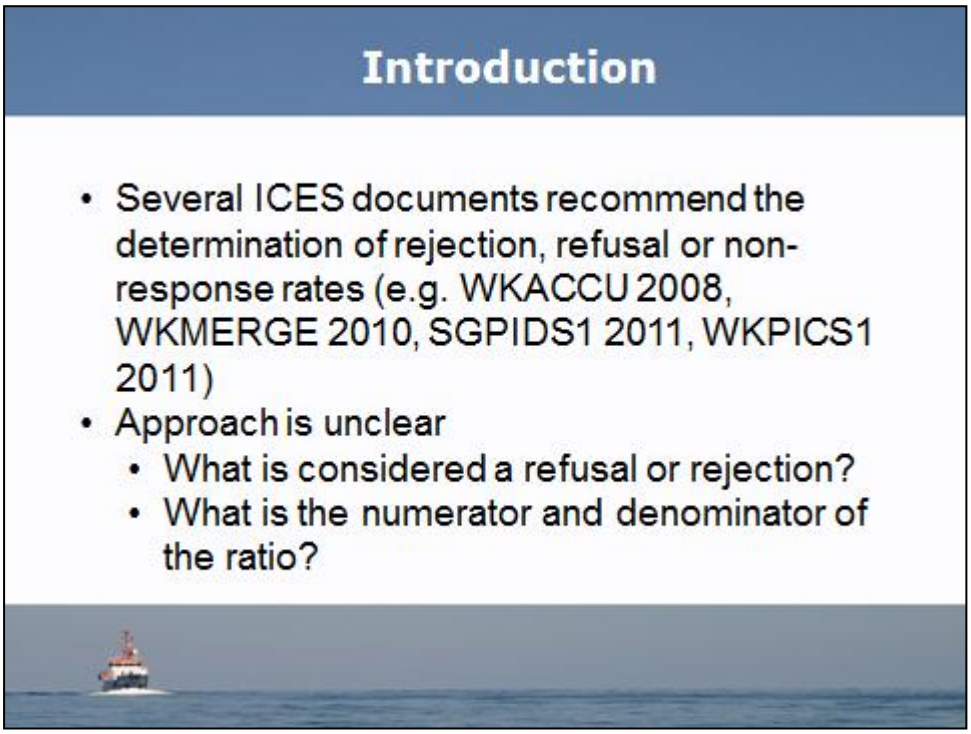
Know what is going on ! Selection process is documented

If nothing else sampling coverage has improved

	Demersal Observer Trips	Unique vessels sampled	%of trips on a previously sampled vessel
2009	57	43	24.6
2010	58	45	22.4
2011	57	53	7.0

Annex 6: Refusal rates: vagueness of an estimator





Case study: Baltic/Germany

- DCF sampling in SD22, 24, 25, 28, 29
- Until 2011: Stratified opportunistic sampling scheme
- Since 2012: Random vessel selection scheme
- Record keeping of DCF phone calls since 2008

Random vessel selection scheme

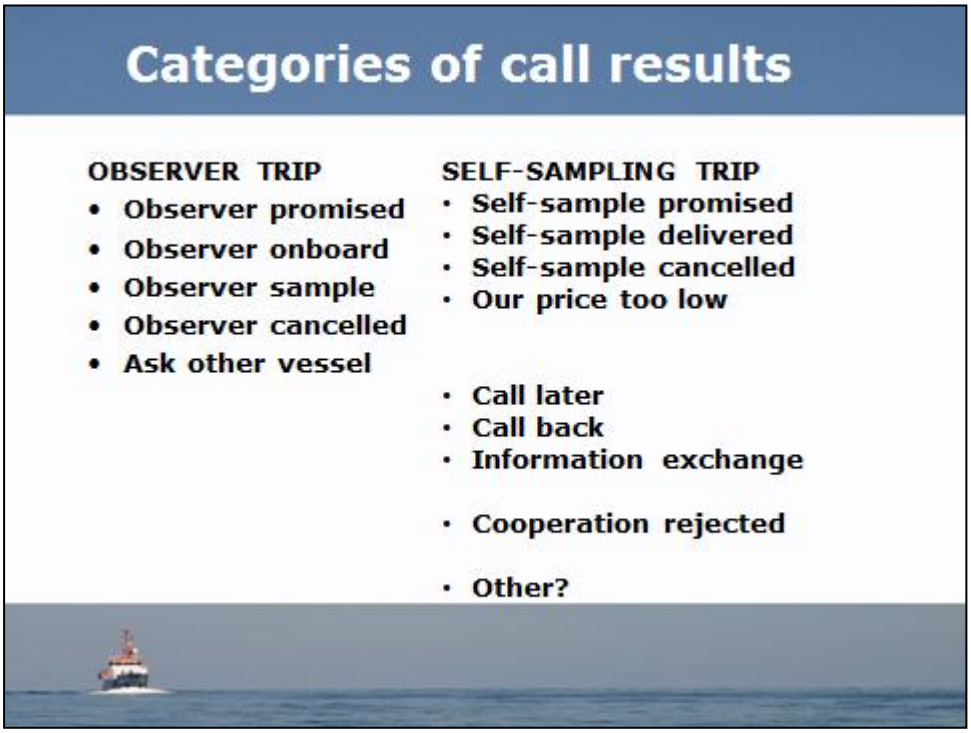
- all vessels were stratified by target species, subdivision, vessel length class, gear type and month (2008-10)
- Vessels in each stratum ranked by their relative share to the stratum landing
- Those within the 90% threshold of cumulative landings were listed
- Vessel lists compiled according to strata for each month
- From these lists, vessel owners are contacted randomly or based on expert judgment, and the phone calls are documented

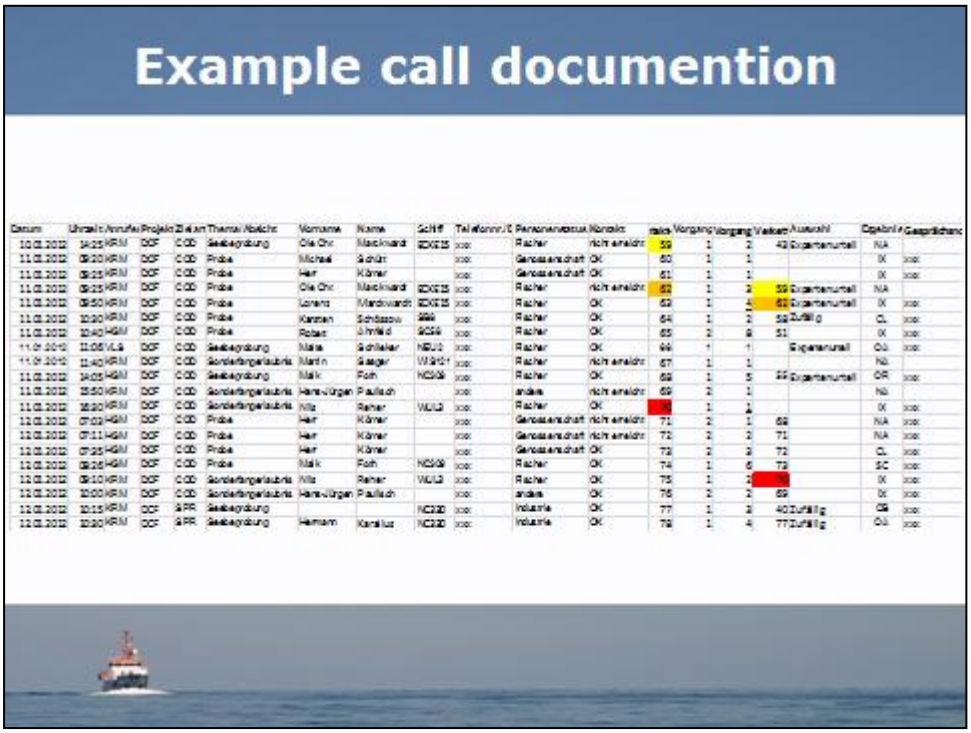
Changes with improved sampling scheme

- List of about 60 → 140 fishing vessels with known fishing patterns and contact information
- More choice/alternatives given unpredictibilities in the fisheries
- Better coverage in space and time
- Previous sampling scheme already targeted most of the main fishing vessels

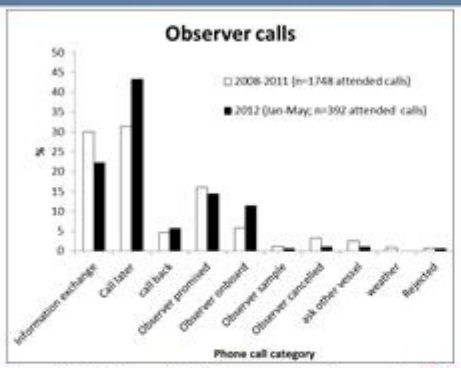
Header of call documentation list

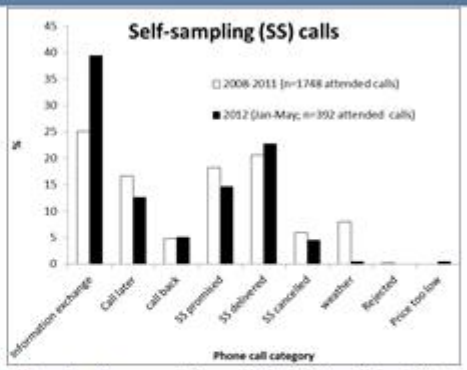
- Date, Time
- Caller (abbreviations of OSF staff members)
- Project (usually dcf)
- Target species (cod, her, spr)
- Topics/Intention (onboard observer, self-sample, other)
- Surname of target subject
- Name of target subject
- Vessel (e.g. SAS111)
- Phone number/E-Mail-address
- Status of the person (fisher, producer organization, industry, other)
- Contact (OK or not attended)
- Contact no. (consecutive numbering)
- Contact identifier (Contact number with target subject to get a trip or self-sample)
- Progress (Number of successive contacts to "Contact identifier" and "target subject")
- Chain of phone calls (phone call refers to Contact no. X)
- Selection (random, expert judgment)
- Result of phone call (several categories)
- Memo (details on phone call)


Categories of call results
OBSERVER TRIP
• Observer promised
• Observer onboard
• Observer sample
• Observer cancelled
• Ask other vessel
SELF-SAMPLING TRIP
• Self-sample promised
• Self-sample delivered
• Self-sample cancelled
• Our price too low
• Call later
• Call back
• Information exchange
• Cooperation rejected
• Other?


Example call documention

Call results 2008-2011 vs. 2012





- Not attended (no connection established): stable over time (37-38% all calls)
- Calls made to get an observer onboard: more CL, OO
- Calls made to get a self-sample: more IX
- Only 4 calls with refused cooperation in 2012 (1 RJ, 1 OO, 1 NA, 1 CL)

Refusal rates 2012

- Herring samples bought at harbors: 0% (rejections after severe quota reductions in the past)
- Cod: In terms of the total number of different vessels contacted (Jan-May 2012): 4/69 did not cooperate (5,8%)
- Cod: all vessels on our list: 4/140 (2,9%)
- Cod: trawlers contacted: 4/46 (8,7%) – 13 trips, 9 SS
- Cod: trawlers on our list: 4/60 (6,7%)
- Cod: Gillnetters contacted: 0/23 (0%) – 10 trips, 2 SS
- Cod: Gillnetters on our list: 0/80 (0%)
- In terms of phone calls: <1%
- Other approaches to define a rejection rate…

Conclusion 1

- For Baltic/Germany, the number of uncooperative vessels is presently low
- Why? 4 main producer organisations which support us (relatively good relationship to the fisheries)
- Significance of uncooperative vessels for the sampling is low given the number of vessels we cannot sample, e.g. due to limitations in man power

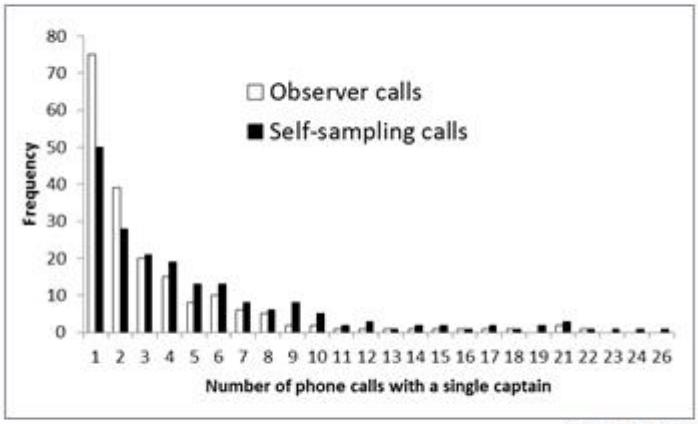
Conclusion 2

- Definition of numerator and denominator of a rejection rate is unclear (?/? ???)
- Informative value is low if the reasons for larger unsampled parts of the fleet root somewhere else (e.g. low overall sampling rates)
- Other estimators may be equally or more informative (e.g. number of different major vessels sampled from all major vessels in a national fleet stratum during a year)

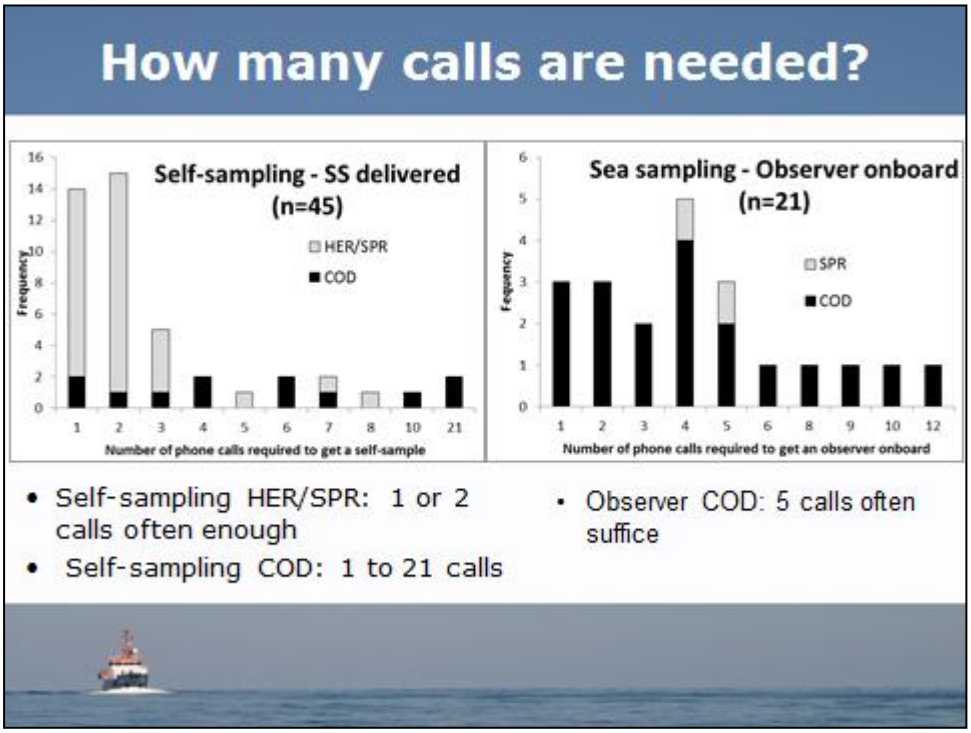
Conclusion 3

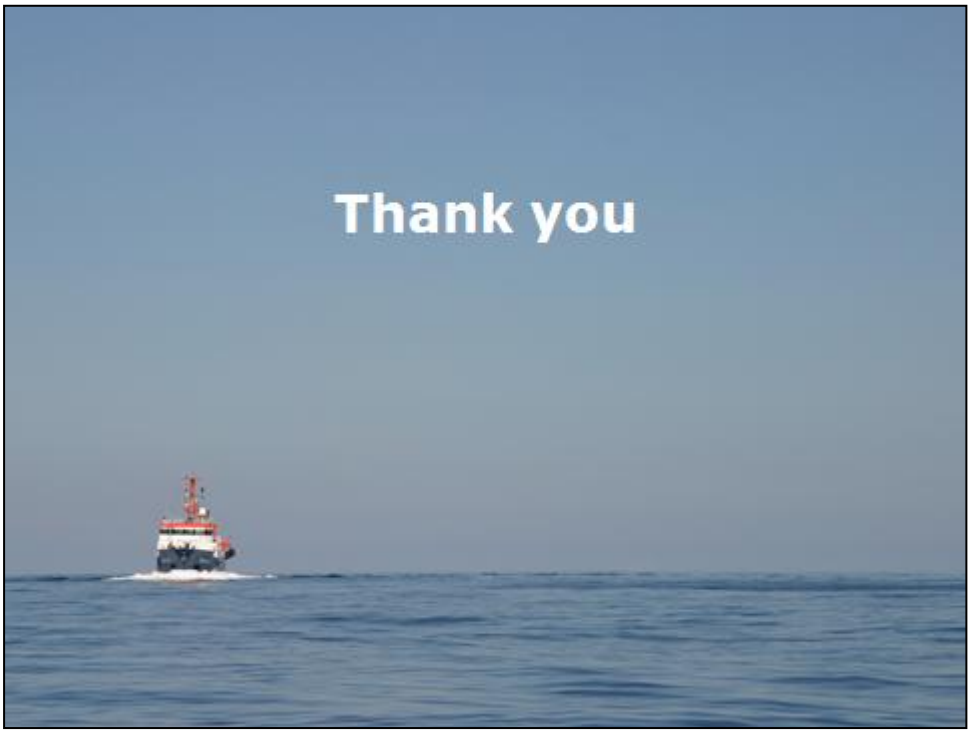
- Merits of phone call documention go beyond the parameter „refusal rate" (e.g. analyses of performance inhouse, sampling scheme, proof of efforts etc.)

Example 1



(data: Jan-May 2012)


How many calls are needed?
Self-sampling - SS delivered (n=45)
HER/SPR
COD
Frequency
Number of phone calls required to get a self-sample
Sea sampling - Observer onboard (n=21)
SPR
COD
Fequency
Number of phone calls required to get an observer onboard
• Self-sampling HER/SPR: 1 or 2 calls often enough
• Self-sampling COD: 1 to 21 calls
• Observer COD: 5 calls often suffice


Thank you

Annex 7: Estimating discards on board EU trawlers in Mauritania

Estimating discards on board EU trawlers in Mauritania

Ad Corten
12 June 2012

1. The handling of catches on board pelagic trawlers

The catch is normally taken on board by means of a fish pump that is attached to the cod-end while the net is still in the water. The contents of the cod-end are then pumped into storage tanks where they may remain for 1-2 days. From the storage tanks the catch is pumped to the working deck where the fish is first passed over a sieve that eliminates the smallest fish. Then the fish is passed to a conveyor line where the catch is sorted into different categories by the crew. One of the categories are the discards. This category is flushed overboard. The other categories are transported to the plate freezers.

Large fish or dolphins that cannot pass through the fish pump are removed by the crew from the cod-end and thrown directly overboard. Some vessels use a large meshed panel in front of the cod-end, the so-called "shark filter" that retains large fish and dolphins, and prevents them of blocking the fish pump. The animals retained by this panel can be released while the net is still in the water. Dolphins retained by the shark filter are normally dead, but sharks and other fish may still be alive.

Discards may thus occur at the following stages of the hauling and sorting process:

a) Large fish and dolphins that are retained by the "shark filter" may be released without being taken on board;
b) In the absence of a "shark filter", the large animals are taken on board en then discarded;
c) Small fish will be removed by the sieve at the beginning of the conveyor line and flushed directly overboard;
d) Fish of no commercial value are sorted out on the conveyor line and then flushed overboard.

The aim of a sampling programme is to obtain estimates of discards at each of the four steps mentioned above. The estimation of discards at steps (a) and (b) requires that observers go out on the main deck. This requires additional work, the approval of the skipper, and the use of safety precautions (hardhats).

In Mauritania, discard sampling programmes have been conducted in two different period: from 1999 - 2005, and from 2010 – present. The methods used during each period are discussed below.

2. Sampling methods in 1999 – 2005

During this period, the RIVO project organised an observer programme on board EU trawlers in Mauritania. This programme was organised in cooperation with the Mauritanian research institute CNROP (now called IMROP). Mauritanian technicians of the institute were trained by Dutch staff, and at least 12 trips were made each year. The observers (working in teams of two) normally stayed on board during the entire trip. As most of the vessels were operating from Las Palmas, observers often had to fly to Las Palmas to board the vessels.

The sampling protocol described that the observers had to take samples at the conveyor line at a point where the crew had already separated the catch into different categories. At this point, the different "lanes" on the conveyor belt contained different commercial categories (e.g. sardinella,

horse mackerel, mackerel) and also a category "discards". From each category, a sample of at least 20 kg had to be taken. To arrive at an estimate of the total catch, the percentage contribution of each category had to be estimated visually (e.g. 50% sardinella, 20% horse mackerel, 20% mackerel and 10% discards). These percentages were then applied to the estimate of the total catch provided by the skipper (e.g. 100 tons), which yielded estimates for the weight of each category in the catch (in this case 50 tons sardinella, 20 tons horse mackerel, 20 tons mackerel, and 10 tons discards). For each sample, a raising factor (RF) was then calculated as

$$RF = \frac{\text{total catch of the category}}{\text{weight of corresponding sample}}$$

The estimate for the discards was accurate as far as species composition and length distribution was concerned. However, the estimate of the absolute amount of discards depended on the visual estimation of the percentage of this category on the conveyor line.

Commercial species that occurred less frequently in the catch (e.g. bonito) were taken by the crew from the conveyor line and temporarily stored in baskets next to the conveyor line. At a later stage, these species were also frozen. The observers also took samples of these less frequent species and calculated a raising factor on the basis of the estimated total catch of the species and the weight of the sample.

Discards of large animals (sharks and dolphins) were sometimes recorded, and sometimes not. From the records of the observers, it was not always clear which of the hauls they had inspected for larger animals. During the first year of the programme, by-catches of dolphins were frequently reported, but the number of records decreased in subsequent years. It was not clear whether this was due to a reduction in by-catches, or to a slackening of the activity of the observers.

3. Sampling methods from 2010 - present

After the first sampling programme had come to an end in 2005 due to the termination of the RIVO project, the Mauritanian institute in 2009 decided to revive the observer programme. For this purpose, a number of observers was taken over from the fisheries inspection service DSPCM. These people had previously worked without any scientific supervision, and it took some time to train them properly. This time the observer programme was organised primarily by IMROP. The Dutch cooperation project merely provided advice and equipment.

In consultation with the observers and IMROP staff, it was decided to change the sampling procedure somewhat. The observers were of the opinion that sampling the catch after the crew had sorted it into categories was too complicated. They preferred to take a random sample of the entire catch when it was pumped from the storage tanks to the sorting deck. In this way, all fish in the sample can simply be raised to the total number in the catch by the following raising factor:

$$RF = \frac{\text{total catch as estimated by captain}}{\text{weight of the random sample}}$$

The problem of course is to estimate the discards on the basis of this random sample. This done by looking at the sorting practice of the crew. It is determined which species and length categories are discarded by the crew, and then these species and length categories in the random sample are also classified as "discards". In general, this system appears to work reasonably well. When the trawlers in

the beginning of 2012 were fishing on mixed concentrations of adult and juvenile sardine and sardinella, the observers reported discard percentages of up to 40%.

The observers are also asked to record the by-catch of large animals by inspecting the catch when it is taken on board. They are requested to record for which hauls they have inspected the by-catch of large animals. In this case, the reported by-catches of large animals can be related to a certain percentage of the total number of hauls, and quantitative estimates can be made of the total by-catch of this type of species.

In practice, the sampling of by-catches of large animals is rather limited for a number of reasons. The observers are often occupied with the sampling of the catch in the factory, and they may not be informed of the moment when a new catch is taken on board. In addition, the crew is reluctant to let the observers inspect the catch on the main deck because of the safety risk. Finally, there is the problem that animals retained in the shark filter are normally not taken on board, and therefore cannot be inspected by the observers.

In summary, the current sampling system probably estimates the discards of fish with reasonably accuracy. The by-catches of larger animals, however, are not yet recorded sufficiently accurate. The improvement of this aspect requires more training (and motivation) on the side of the observers, and more cooperation from the crews of the vessels.

Annex 8: Spanish and Portuguese discard sampling protocols

SPANISH DISCARD SAMPLING PROTOCOL ON OTB OPERATING IN VIIIc–IXa DIRECTED TO HORSE MACKEREL

Introduction

The sampling strategy and the estimation methodology used in the `Spanish Discards Sampling Programme' follows the guidelines established in the ICES `Workshop on Discard Sampling Methodology and Raising Procedures' (2003). Observers–on–board programme is based on a hierarchical sampling design, meaning that sampling scheme is carried out over nested sampling units. In the Spanish case, trip is considered as the Primary Sampling Unit (PSU).

Fleets stratification

Fishing operations are stratified into métiers in order to achieve better sampling allocation. This stratification is based on a complex mix of technical (gear and target species), spatial (fishing area), and temporal (quarter) covariates. Five métiers are defined for the Northern Spanish coastal bottom trawl fleet (ICES VIIIc and IXa divisions):

OTB_DEF _>=55_0_E: trips targeting a mixed of demersal species in East VIIIc.

OTB_DEF _>=55_0_W: trips targeting a mixed of demersal species in West VIIIc and North IXa.

OTB_MPD_>=55_0_HOM: trips targeting horse mackerel (*Trachurus trachurus*).

OTB_MPD_>=55_0_MAC: trips targeting mackerel (*Scomber scombrus*).

PTB_DEF_>=55_0_0: trips targeting blue whiting and hake (*Micromesistius poutassou, Merluccius merlucius, respectively*).

The métier herein considered for comparison with the portuguese case is

OTB_MPD_>=55_0_HOM: trips targeting horse mackerel (*Trachurus trachurus*).

Trip selection

Although not considered as Primary Sampling Unit, a list of fishing vessels known to participate in the target métier is used as proxy to obtain a sampling trip. Trip is quasi–randomly (cooperative or opportunistic) drawn from the total trips carried out at the time that will perform the sampling.

Haul selection

The haul selection depends on the total number of hauls expected to be carried out during the trip. Mainly vessels perform short trip duration (usually 1 day), in this case the observer is aimed to sample on all hauls, some hauls (1–2) can be left out from the sampling scheme during larger trips.

Within haul sampling protocol

Physical covariates

Observers collect physical covariates associated to the setting and hauling (date, time, hours fished, depth...) directly from the bridge. Hours fished is the period of time between the gear contact with the seabed and the beginning of haul manoeuvre.

Catch estimation

Once the codend is hauled onboard, the observer ask to the skipper about a visual estimation of the total cacth.

Sampling strategy

The observer follows the crew to the sorting deck to start the catch sampling. The sampling protocol is split into the catch fractions generated after the commercial sorting.

Sampling on discarded fraction

Samples for the discard fraction is obtained at the end of the conveyor belt at different moments during the sorting process. The sampled weight depends on the discard composition, but usually is not larger than 20 k. Information from all species or group of species in the discard sample is collected. Length size are the target information for all fish species and nephrops. Sub–sampling within the discard sample is carried out for abundant species with a clear unimodal length size distribution. In that case, length size distribution obtained from the subsample is raised to the total weight obtained by dynamometer. The sampled weight are estimated using the weight – length relationship when available, else, weight is obtained using hanging balances. Invertebrates others than nephrops are counted and weighted by species or group of species.

Biologically sampled (otholits) are extracted from undersized individuals present in the discard fraction. The aim is to fill the gaps from the otholit collection obtained in markets.

Sampling on retained fraction

length size measures on retained species is carried out at intervals of time between the discard sample collection or when the sorting process ends.The weigth of the retained fraction is obtained using the length–weigth relationship when all individuals are sampled, otherwise total weigth is estimated by raising the average weigth per box to the total number of retained boxes.

Marine mammals catch monitoring

The observer checks on deck or asks to the crew if any marine mammal are caught. Morphometrics are collected and recorded in those cases.

Raising discards to haul level

The discard Raising Factor (RF):

RF= [Total Catch Estimation (k)–Retained catch (k)]/Discard sample (k)

RF is used to raise the species discard sampling to haul level by:

D=species discard sample [k]*RF

Raising discards to trip level

When all hauls carried out during the trip are sampled, the species estimation of trip discards is the sum of discards estimated in all hauls. If only a set of hauls were sampled, the mean discards per haul is estimated and raised to the total number of hauls.

PORTUGUESE DISCARD SAMPLING PROTOCOL ON OTB OPERATING IN IXa DIRECTED TO HORSE MACKEREL

Introduction

The Portuguese on–board sampling program, included in the EU DCR/NP, started in late 2003 and involves on–board sampling of several fishing métiers. These include, amongst other, bottom otter trawl, deep–water set longlines, gill and trammelnets (of various mesh sizes) and purse-seines. From these, the bottom otter trawl fleet (OTB) constitutes the most comprehensively sampled fleet. The procedures used to collect data on board and raise discard data from samples to fleet level discards are described in Prista et al. (2011), amongst other.

Fleets stratification

The Portuguese on–board sampling program of the OTB fleet is based on a quasi–random sampling of cooperative commercial vessels between 12 and 40 meters long. For sampling purposes the OTB fleet is split into two métiers:

- OTB_CRU_>=55_0_0 – a crustacean fishery that operates cod–end mesh sizes 55–59mm and >70mm targeting deep–water rose shrimp, Norway lobster and blue whiting. This fishery takes place in the SW and S coast of Portugal.
- OTB_DEF_>=55_0_0 – a demersal fish fishery that operates cod–end mesh size 65–69mm and >70mm and targets horse–mackerel, cephalopods and other finfish.

The métier herein considered for comparison with the Spanish case is

- OTB_DEF_>=55_0_0 – a demersal fish fishery that operates cod–end mesh size 65–69mm and >70mm and targets horse–mackerel, cephalopods and other finfish.
-

Trip selection

The EU DCR/NP (CR (EC) 199/2008; CD 2010/93/EU) establishes fishing trip as the sampling unit to be used by at–sea discard sampling programmes. The Portuguese on–board sampling programme targeting the bottom otter trawl fleet is based on a quasi–random sampling of trips from a set of cooperative vessels known to operate in each fishery. Annual sampling targets are fixed for each fishery, namely 12 trips in the OTB_CRU fishery and 27 trips in the OTB_DEF fishery. Recent work indicates these sampling levels are lower than the ones necessary to achieve DCF precision goals (ICES 2012). However, human and financial constraints along with the need to provide coverage to multiple métiers have so far prevented an increase in sampling levels even if substantial efforts were put into team optimization and improving the contact with fishers.

Haul selection

Previous to 2011, observers aimed at a complete census of all hauls made in each trip. From 2011 onwards, haul selection changed to systematically (odd or even hauls) after random choice of the starting haul (first or second).

Within haul sampling protocol

Physical covariates

Observers collect fishing effort information directly (date, time, hours fished). Hours fished is the period of time between the gear contact with the seabed and the beginning of haul manoeuvre. Observers also register environmental information from the skipper (GPS coordinates, depth, bottom type, etc.)

Catch estimation

Catch volume is estimated independently from skipper's opinion. It is obtained from the relative proportion between discards: retained weight in a sample from catch and raised by total landings (see detailed calculations below)

Sampling strategy

The observers follow the crew to the sorting deck to start the catch sampling. A sample from catch is taken that is generally composed of 2 or 3 boxes. Sample collection is carried out just before the sorting area and spread throughout the catch/sorting period in order to reduce within haul biases in species and size composition. Average sample size per haul is ~21 kg. The sample is split into retained fraction and discard fraction according to the fisher's criteria. Fisher's criteria has been observed to vary between hauls, with some species and/or sizes being retained in some hauls and discarded in others.

Sampling on discarded fraction

The sample size of the discard fraction depends on the relative proportion of discards in the catch sample. Data is collected from all species in the discard fraction: Numbers and weights are obtained for all species in the fraction and length frequency is obtained for all fish species, cephalopods and commercial crustaceans. Subsampling within the discard fraction is carried out for abundant species with a clear unimodal length size distribution. In that case, length size distribution obtained from the subsample is raised to the total weight of sample. All weighs are taken using hanging balances. When sea conditions are harsh, weight–length relationships are used.

Biological sampling (individual weights, otoliths and maturity) is done at the lab from samples collected onboard. Quarterly sampling targets per length class and sex are defined for each species.

Marine mammals catch monitoring

The observers check the deck for marine mammals in all sampled hauls. Morphometrics are collected and recorded in such cases. In unsampled hauls, the observers only ask to the crew if any marine mammal was caught.

Raising discards to haul level

Total discards at haul level are calculated as:

[total landings in haul] * [discard fraction weight]/[retained fraction weight]

Haul catch (total): Haul discards + total landings in haul

Haul discards (per species): Raising factor is total catch (estimated)

[total catch]*[weight of species retained fraction]/[total weight of catch sample], where "total weight of sample" does not encompass trash and inorganic debris (empty shells, rocks)

Raising discards to trip level

Total discards are not routinely calculated at trip level. Rather, in each haul species discards per hour are computed by dividing total discards of species by the hours fished. Then, an estimate of average species discards per hour is computed from all sampled hauls using a weighted mean (using haul duration as weighing factor). This estimate is used in fleet level calculations.