Annex 3: Background document for response to special request regarding precision and bias based on RDBES format

Note: The following ad-hoc report will be added as an annex to the coming report of the Second Workshop on Estimation with the RDBES data model (WKRDB-EST2) 2020 meeting. As a result of being published before the full report, page numbers and headings will be subject to change.

This report aims to support EU member states in evaluating the statistical accuracy of their catch sampling data, where accuracy refers to the closeness of statistical estimates to their true values. Statistical accuracy is considered in terms of two components: precision and bias. Random uncertainties inherent in estimates due to sampling are described by precision, whilst systematic differences between the estimate and the true value are described by bias. Since this is a complex subject and sampling programmes are usually implemented differently in different countries the work presented relates only to national probabilistic sampling and design-based estimation. To use the code developed, member states will need to convert their national data to the commercial fisheries Regional Database & Estimation System (RDBES) data format.

The evaluation of data precision has been performed using two complementary techniques. For relatively simple sampling designs it is possible to use analytical functions to calculate the precision (or a related statistical measure such as variance) of a statistical estimate. We present these calculations and implementations of these calculations in R code. For more complicated sampling designs, the use of analytical functions is usually not feasible. In these cases, it is necessary to evaluate precision using resampling techniques such as bootstrapping. This report discusses when bootstrapping is appropriate and gives several worked examples describing how bootstrapping can be applied in different cases.

The evaluation of bias in catch sampling is a difficult subject and most biases are generally hard to quantify. It should be noted that there can be several different types of bias occurring at different points in the data collection and the advice production cycle. This report only considers the type of bias that may occur as a result of sampling—not other biases such as those that may be present in particular estimators or stock assessment models. Our approach to bias builds on the previous work available in the ICES literature to identify and enumerate common sources of bias in catch sampling programs. The information was collated and an evaluation performed as to whether data stored using the RDBES data format can inform about potential biases. Reports are presented that can help member states to identify deviations in their sampling programmes and sampling variability that can potentially lead to bias in catch estimates.

The report is a first step towards providing EU member states with a set of tools that can be used to characterize the precision and bias of their catch sampling data. The aim is to provide a solid foundation that, whilst immediately useful in itself, has greater value as a building-block for future work. To this end, a summary of the further activity that is required to extend the work to other scenarios (such as regional sampling programmes) is presented.
Table of Contents

1 Introduction .......................................................... 1
   1.1 Regional databases supporting the CFP ......................... 1
   1.2 From the Regional Database (RDB) to the new Regional Database & Estimation System (RDBES) .... 2
   1.3 Benefits of the RDBES .................................................. 2
   1.4 Development of RDBES tools to evaluate data accuracy relating to bias and precision ... 6
   1.5 Summary ................................................................. 7

2 Data accuracy ........................................................... 8
   2.1 Introduction ............................................................ 8
   2.2 Scope ........................................................................ 9
   2.3 Approach ................................................................. 9

3 Analytical calculation of variance .................................... 10
   3.1 Sampling without replacement in all three stages .............. 10
   3.1.1 Inclusion probabilities for the general case .................. 10
   3.1.2 Estimation for the general case .................................. 11
   3.1.3 Inclusion probabilities for Simple Random Sampling (SRS) without replacement in each stage .................. 11
   3.1.4 Estimation for SRS without replacement in each stage .... 12
   3.1.5 Simplified variance estimation .................................... 12
   3.2 Sampling with replacement in the first stage .................. 13
   3.2.1 Estimation for the general case .................................. 13
   3.2.2 Estimation for SRS with replacement in the first stage .... 14
   3.2.3 References ............................................................ 14
   3.3 Implementation in R .................................................... 14

4 Bootstrapping ............................................................ 15
   4.1 Introduction ............................................................ 15
   4.2 Types of data needed for input to bootstrapping catch sampling .................................. 17
   4.3 Notation ................................................................. 19
   4.4 Pseudo-code for running bootstrap simulations of example 1 from hierarchy 1 assuming all stages are SRSWR .............................................. 20
   4.4.1 Basic bootstrapping steps for a single simulated sampling effort ...................... 20
   4.5 Example code for bootstrapping a stratified multi-stage sampling design assuming all stages are SRSWR .............................................. 22
   4.6 Example code for bootstrap sampling of single stage by SRSWOR .............................. 23
   4.7 Example code for bootstrapping two-stage sampling by SRSWOR at each stage ........ 24
   4.8 Bibliography ............................................................ 25

5 Bias ........................................................................... 26
   5.1 Method ................................................................. 26
   5.2 Summary ................................................................. 26

6 Future work ............................................................. 28
   6.1 Introduction ............................................................ 28
   6.2 Analytical calculation of variance .................................. 28
   6.3 Bootstrapping .......................................................... 28
   6.3.1 Implementation in R code ......................................... 28
   6.3.2 Post-stratification and domain estimation ....................... 28
   6.3.3 Age-Length Key (ALK) construction .......................... 29
   6.3.4 Estimation based on other sampling designs ................... 29
   6.3.5 Extension to other types of estimation ............................ 29
   6.4 Bias ......................................................................... 29
   6.5 Roadmap ................................................................. 30

7 Summary ..................................................................... 31
<table>
<thead>
<tr>
<th>Annex A3.1:</th>
<th>List of participants</th>
<th>32</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annex A3.2:</td>
<td>R implementation of analytical variance calculation using design-based estimation</td>
<td>33</td>
</tr>
<tr>
<td>Annex A3.3:</td>
<td>Bootstrapping pseudo-code</td>
<td>35</td>
</tr>
<tr>
<td>Annex A3.4:</td>
<td>Potential sources of bias</td>
<td>52</td>
</tr>
<tr>
<td>Annex A3.5:</td>
<td>Example reports to illustrate potential sources of bias</td>
<td>72</td>
</tr>
</tbody>
</table>

**Recommended format for purpose of citation:**

1 Introduction

The management of the Common Fisheries Policy (CFP) should be guided by the principles of good governance. Those principles include decision-making based on the best available scientific advice, which requires harmonized, reliable, and accurate datasets. To achieve this, EU member states (MS) are obliged to collect and manage data in accordance with the recast Data Collection Framework (DCF) (Regulation (EU) 2017/1004) and the Commission Decisions (EU) 2019/909 and (EU) 2019/910. The DCF places a strong emphasis on cooperation between MS and Regional Coordination Groups (RCGs) are established to support this. Furthermore, MS are encouraged to align their data collection in regional work plans.

It is important to realize that the vast majority of datasets that are used in stock assessment or by other end-users are the result of a series of complex data transformations (e.g. data on unwanted catches). Harmonized, reliable and accurate datasets are thereby not only dependent on data of good quality but also on quality assured processes to transform data. It becomes even more complex in a multinational context where different MS use different processes and where these processes are not always fully documented. The result might be that it is not possible to fully assess the quality of the multinational dataset or the impact the quality has on subsequent analyses. One of the reasons for the present situation is that this processing often requires access to detailed data that might be confidential (e.g. commercial fisheries data).

Harmonized, reliable and accurate datasets also require data collection schemes that are built on sound statistical principles. The MS have worked for several years to establish such schemes in the logistically complex environment that fisheries constitute. The work continues and MS are now also focusing on integrating these new designs into data processing/estimation methods and databases. All this work is a prerequisite for future implementation of regional work plans, as integrated data collection also requires integrated data processing/estimation and management.

1.1 Regional databases supporting the CFP

There is an existing commercial fisheries Regional Database (RDB) that is hosted by ICES and currently used to store aggregated effort and landings data, and detailed biological sampling data. The MS in the North Atlantic, North Sea & Eastern Arctic, and Baltic Sea Regional Coordination Groups currently submit data to the RDB annually and use it to support their work. The Long Distance Fisheries RCG also intends to submit data to the RDB in the near future.

Since the RDB was first developed, the requirements of the DCF have become both broader and more complex. Alongside this, there have been changes in wider fisheries management legislation such as the Landings Obligation. During this time, there have also been a number of improvements within scientific data collection practices including the move towards Statistically Sound Sampling Schemes (“4S”), greater regional coordination, and greater transparency in the scientific evidence base used for fisheries advice. The existing RDB is not able to fully support these new requirements. To this end, a new regional database, the Regional Database & Estimation System (RDBES), has been designed and is in the process of being implemented.

One objective of the RDBES is to support the CFP by improving the harmonisation, transparency and quality assurance of datasets used in analyses underpinning scientific advice. The RDBES is also a prerequisite for the implementation of regional sampling plans and production of transparent regional datasets.
1.2 From the Regional Database (RDB) to the new Regional Database & Estimation System (RDBES)

The Commission is generally supportive of the development of compatible regional databases. This is specified in legislation, especially Article 18 of the recast DCF (Regulation (EU) 2017/1004):

“With a view to reducing costs and facilitating access to detailed and aggregated data for end-users of scientific data and other interested parties, Member States, the Commission, scientific advisory bodies and any relevant end-users of scientific data shall cooperate to develop compatible data storage and exchange systems, taking into account the provisions of Directive 2007/2/EC. Those systems shall also facilitate dissemination of information to other interested parties. Such systems may take the form of regional databases. Regional work plans referred to in Article 9(8) of this Regulation may serve as a basis for agreement on such systems.”

The important points about the RDBES development from the Commission’s point of view have been enumerated as:

- To ensure regional database functionality for RCG use is uninterrupted;
- That access to data is provided in line with EU policy (MS ownership of data and agreement before use; RCGs have access to the regional database at all times and can use the data; confidentiality rules);
- The Commission supports any extension of the RDBES to other variables (such as recreational fisheries, large pelagics) and other currently separate databases;
- To be able to use the future RDBES for automatic reporting of DCF deliverables such as Annual Reports or Work Plans–National Correspondents should be able to extract data to create the required tables;
- It is important to ensure compatibility between the ICES RDBES and other similar databases (i.e. the proposed Med&BS regional database).

The RDBES is currently in development and is scheduled to go live during 2022, at which point the existing RDB will become read-only.

1.3 Benefits of the RDBES

The aims of the RDBES are:

1. To make data available for the RCGs;
2. Provide a regional estimation system for ICES stock assessments;
3. To increase the data quality, documentation of data, and the use of approved methods;
4. To facilitate the production of fisheries management advice and reports;
5. To increase the awareness of fisheries data collected and the overall usage of these data.

These aims are fully in line with the DCF and support the Common Fisheries Policy (CFP) aim to “…conserve fish stocks and reduce overfishing in order to provide EU citizens with a long-term stable, secure and healthy food supply.”

---


Aim 1: Make data available for the RCGs

A key aim of the RDBES is to support the DCF work of MS by supporting the RCGs. RCGs are responsible for the coordination of MS sampling activity of commercial fisheries. That sampling is the basis for the estimates of commercial catches used in ICES advice for upcoming years—effective work at RCG level is ultimately needed to fulfil the CFP objectives. The RDBES will allow MS to upload both their commercial detailed biological sample data and aggregated effort and landing data to a new regional database.

The RDBES will support the work of the North Atlantic, North Sea & Eastern Arctic, Baltic Sea, and Long Distance Fisheries RCGs. These RCGs have previously stated their strong support for developing the RDBES. In their 2018 meetings the North Atlantic, North Sea & Eastern Arctic, Baltic Sea RCGs recommended the development and use of the RDBES to store and analyse data. The Long Distance Fisheries RCG also have stated their desire for MS to upload their data to a regional database and require some features of the new RDBES for this to fully occur.

The RDBES could potentially also support the work of the recreational, diadromous, and large pelagic regional data collection, but this will be dependent on whether they wish to pursue this and that the funding is available for any developments agreed upon.

As stated earlier Regulation (EU) 2017/1004 encourages the use of compatible regional databases. Both the RDB and the RDBES fulfil this, however, the RDBES has been designed to better allow RCGs and MS to fulfil their obligations towards documenting and improving data quality, and designing and implementing regional sampling designs. The RDBES is designed to capture information about both biological data and how it was sampled. This allows much more realistic analysis of sampling activity to be performed. Similarly to the existing RDB, the RDBES will allow RCGs to analyse data collected by MS at a regional level but new, regional sampling designs will also be supported by the RDBES. The RDBES has been designed to allow the storage of data from regional sampling schemes, in particular by allowing the specification of different sampling designs for different strata within an overall regional sampling design. The estimation system within the RDBES will also allow for the production of regional estimates. The RDBES will also be used by the RCGs to support regional work plans and sampling schemes (such as referred to in Regulation (EU) 2017/1004).

The RDBES will support MS to implement Article 5 of Commission Implementing Decision (EU) 2016/17 by allowing MS to record their statistically sound sampling designs in a common format.

Unlike the RDB, the RDBES will allow sufficient data about the observation of bycatch and Protected, Endangered and Threatened Species (PETS) to be stored such that the relevant bycatch working groups can use it in their work. As Regulation 812/2004 will be repealed, monitored data on bycatch of PETS in commercial fisheries will be included in ICES RDBES. Regulation (EU) 2019/1241 requires the collection of scientific data on incidental catches of sensitive species and the RDBES provides regional storage for this information.

Regulation (EU) 2017/1004 states “It is important to collect biological data on recreational fisheries where there is a potentially significant impact on the state of the stock…” and the RDBES could provide a regional storage system for this recreational data - this regional storage is currently missing because it cannot be stored in the RDB. On the RDBES development, roadmap the

---

3 For a summary of the recommendations, see 2018 liaison meeting report: https://datacollection.jrc.ec.europa.eu/documents/10213/1239625/2018-10_15th_Liaison_Meeting.pdf/8c3ba088-c9a3-410c-93c-a1c7260b6d6d

possibility of storing recreational data has been included. Whilst it would be good to increase the documentation and transparency of the recreational data used in stock assessments, this is a complicated matter as each of the recreational surveys is conducted in its own way depending on cultural specifications and there is a wide variety of methods being used to sample. The initial plan is therefore to restrict the storage of recreational fisheries data to an aggregated level—with a move to detailed data storage possible in the future.

**Aim 2: Provide a regional estimation system for ICES stock assessments**

The RDBES will be a key part of ICES stock assessment and the way it supports the CFP objective of conserving fish stocks. ICES stock assessment currently depends on many different data calls and many countries and data submitters, including non-EU countries like Norway and Iceland that jointly contribute national estimates of commercial catches to each ICES stock assessment working group. This means there is a duplication of effort and a lack of consistency. The detailed data and processes used by MS to submit data for stock assessment are also not visible outside of the MS and it is hard for ICES to assure the quality of data provided by sometimes dozens of different individuals for a single fish stock. Important data quality indicators such as estimates of precision are often not submitted to stock assessment data calls. The RDBES will resolve these problems by (i) allowing MS to peer-review each other’s estimation algorithms and validate their suitability, (ii) share common functions and tools to obtain those estimates, (iii) strengthening the link between data collectors and stock assessment groups, and (iv) allowing precision of the stock estimates to be correctly calculated and incorporated into the stock assessment models.

Assuring quality is a key element of the ICES advice plan and the RDBES will be an essential tool in the overall quality assurance framework. As a key client of ICES advice and responsible for the success of the CFP, the EU Commission will directly benefit from improvements in ICES stock assessment output.

**Aim 3: To increase the data quality, documentation of data, and the use of approved methods**

Under Article 14 of Regulation (EU) 2017/1004 MS have a responsibility to ensure that primary, detailed, and aggregated data has appropriate quality assurance and control measures applied before transmission to end-users and that these quality assurance measures have been developed in accordance with the procedures adopted by international scientific bodies, STECF and RCGs.

Generally, each MS has a unique format for its national databases and its own procedures for calculating its estimates of commercial catches so it is very difficult to develop, share, or evaluate data quality tools and estimation methods directly between countries. The RDB was a first step in the right direction and allowed the development of a number of common quality tools and a few standard algorithms of estimation. However, the RDB data format makes strong assumptions about the sampling schemes that MS are using that frequently differ from the way national data is actually collected. This has made it difficult for data to be analysed within the RDB. Unlike the current RDB, the new RDBES stores all the important information about how data was collected including all the novel statistically sound sampling variables demanded in Commission Implementing Decision (EU) 2016/1701. These variables include, but are not restricted to, the sampling scheme used, the sampling frames and stratification schemes, the different sampling units (e.g. fishing trips, port-days), and how units were selected for sampling (e.g. simple random sampling, expert judgement) in all sampling levels. This will allow new and better estimates of commercial catches and accompanying quantitative measures of quality to be calculated using

---

5 [https://issuu.com/icesdk/docs/ices_advisory_plan](https://issuu.com/icesdk/docs/ices_advisory_plan)
the RDBES (e.g., the precision of any estimates calculated from the data). It also allows documentation about the different sampling schemes that MS are using to be easily generated.

Using a common regional platform to develop quality checks and having a common, transparent and documented repository of estimation algorithms used in producing the commercial data entering stock assessment also means that MS can develop these procedures in a collaborative peer-reviewed manner which will improve efficiency and effectiveness. For example, it will be possible to encode approved statistical quality checks and estimation procedures that many other MS can review and use. Data quality checks that can be applied consistently on a regional scale will also be necessary for the regional work programmes that are currently being developed.

**Aim 4: To facilitate the production of fisheries management advice and reports**

The RDBES web application will provide certain functionality such as data uploading, and managing permissions but stock estimation and imputation will be performed within the ICES Transparent Assessment Framework (TAF) - this is an open framework for organising stock assessments. All data inputs and outputs are traceable and versioned. The open framework enables stock assessment scientists to easily find, reference, download, and run the assessment from any stage in the process leading to the published ICES advice for a given stock. Anyone will also be able to find, reference, and download the estimation method behind the assessment (but not the underlying data). Basing the stock estimation functions of RDBES on the TAF has a number of advantages: the TAF exists and users are already gaining expertise in it, there is technical and content support available, version control of data and scripts is established, and it provides strong linkages to stock assessment groups.

A key benefit of the RDBES is that it will be used to assure the quality of the DCF process from data collection to stock assessment.

**Aim 5: To increase the awareness of fisheries data collected and the overall usage of these data**

“[Regional databases]...facilitate the work of the EU Member States by reducing the burden of multiple data submissions (for data calls) under different formats. They allow end users to calculate statistical estimates of data tailored to their needs, and help to streamline and ease the reporting of Member States on the EU data collection.”

The aim of the RDBES is that data are available at the highest possible resolution whilst taking into account data ownership, access rights, and confidentiality constraints. This means that it could be possible to use the data for other relevant purposes. For example, currently, MS submit data to both the RDB data call and the Fisheries Dependent Information (FDI) data calls. This involves both a duplication of effort and can produce a lack of consistency. Unlike the existing RDB, the aggregated effort and landings data model in the new RDBES has been designed to be compatible with the FDI data call. Thus it could be possible for MS to use the RDBES to also respond to the FDI data call. In the same way, it should also be possible for MS to use the RDBES to complete part of their DCF annual reports. This would be a benefit for MS since they would not have to duplicate all the data submission work, and a benefit for the STECF since it would remove a possible source of consistency error.

---

6 Annex 3 in “Call for proposals MARE/2020/08 Strengthening regional cooperation in the area of fisheries data collection” http://ec.europa.eu/newsroom/document.cfm?doc_id=66541
1.4 Development of RDBES tools to evaluate data accuracy relating to bias and precision

As previously discussed, the annual national work plans and reports of MS are an important record of the data quality processes that are applied at the national level. Specifically, this information is summarized in Table 5A of the EU-MAP. This table typically asks whether documentation on a subject exists and, if so, where that documentation can be found. The subjects covered include sampling design, quality checks at the point of data capture, evaluation of precision and bias, and editing and imputation methods. The contents of these tables have been analysed during inter-sessional work of the RCGs and it has been seen that MS have difficulty answering some of these questions since there is a lack of guidance or tools available on the subject.

In particular, the documentation around data accuracy, bias and precision has been observed to be one of the weaker areas - specifically related to the following questions:

- “Are processes to evaluate data accuracy (bias and precision) documented?”
- “Where can documentation on processes to evaluate accuracy be found?”

When completing this table one MS commented,

“Presently, we do not evaluate bias and precision of our data because we are not aware of routine tools available for such estimates on a national level. As soon as routines are available we will use these. (...)”

It can be seen that tools to evaluate data accuracy relating to bias and precision at a national level are required. Evaluation of this bias and precision at a national level will also be relevant to ICES and the Commission since these data feed into stock assessments and can affect the accuracy of their outputs. To enable this advice to be used by all MS (and ICES member countries if they desire) it should be based on a common data format from which statistical bias and precision can be correctly calculated. The new RDBES data model provides that format since it provides a common structure to describe both the detailed sampling data and, importantly, the sampling design underlying how those data were obtained.

Still, just having data in a sophisticated data structure like the RDBES is not enough: the very estimation of precision and bias for individual programmes is a complex subject frequently found diversely implemented in different countries. For example, there are a number of different estimation techniques that can be used to create inputs for stock assessment from biological data. Broadly these can be categorized as “model-based” and “design-based” estimation methods. (Model-based methods are in common use but involve assumptions on sampling as well as on nature which can be difficult to verify whereas design-based estimators involve only assumptions on sampling which are in principle controllable and easier to scrutinise.)

To resolve this, in the first instance the tools will relate specifically to design-based estimation since substantial further work will be required for it to be applied to other types of estimation. A roadmap has been produced for the work required to extend the tools to these other types of estimation in the future.

This report on evaluating data accuracy (precision and bias) for design-based estimation at a national level covers the following subjects:

---

Definition of the prerequisites that a MS will need to meet to be able to use the tools (e.g. MS data will need to be in the RDBES data format; the MS will need to be carrying out probabilistic sampling and recording certain data);

- Specification of the statistical functions to allow MS to evaluate bias and estimate precision for design-based estimation;
- Identification of further functions that would be required in the future to evaluate data accuracy for other type of estimation, and for regional data estimation;
- Recommendations for further work and a roadmap of how to extend the advice to other types of bias and precision estimation.

### 1.5 Summary

This section has shown how the new RDBES that is currently in development will be better able to support the recast EU Data Collection Framework (Regulation (EU) 2017/1004) than the existing RDB. The RDBES will provide an essential platform for MS and RCGs to fulfil their obligations towards documenting and improving data quality and designing and implementing regional sampling designs. ICES is an important end-user of DCF data and the RDBES will be a key input to ICES stock assessments. In particular, the RDBES will improve data quality and transparency by allowing peer-review of procedures, sharing common functions and tools to obtain those estimates, and allowing precision of stock estimates to be correctly calculated and incorporated into the stock assessment models. As a key client of ICES advice and responsible for the success of the CFP, the EU Commission will directly benefit from improvements in ICES stock assessment output.
2 Data accuracy

2.1 Introduction

This section describes how the general concept of data accuracy is treated within statistical analyses and defines the scope of the types of fisheries sampling that this report covers, and the approach used.

It is useful to first define what is meant by data accuracy in this context:

“Accuracy of data is the closeness of computations or estimates to the exact or true values that the statistics were intended to measure...The concept of accuracy relates a numerical estimate to its true value according to an agreed definition. The closer the estimate is to its true value, the more accurate it is. The difference between the estimate and the true value is called the error of the estimate and error is thus a technical term to represent the degree of lack of accuracy. The error has a random component (variance) as well as a systematic component (bias). It is sometimes better to speak of uncertainty than error, when the term error risks to be confused with a mistake committed, which is a very different matter.” (European Statistical System (ESS) handbook for quality and metadata report, 2020, p.98)

In the context of Table 5A within the DCF National Workplans / Annual Reports the concept of data accuracy is explicitly linked with the terms “precision” and “bias”. In this case, precision can be considered to be inversely related to variance i.e. a higher variance in the random component of the uncertainty means a lower precision.

An informal example, which is often given to illustrate the difference between variance and bias, is that of trying to shoot arrows at a target. Ideally, we would like all our arrows to be in the centre. The diagram below illustrates how the arrows might hit the target in different variance and bias scenarios:

![Diagram showing arrows in different variance and bias scenarios]

Clearly, the desired situation is to have both low variance (high precision) and low bias in our estimates although this may not always be possible in practice.
It should be noted that there can be a number of different types of bias occurring at different points in the data collection and advice production cycle – in this report we only consider bias that may occur as a result of sampling, not other biases such as those that may be present in particular estimators, or stock assessment models.

2.2 Scope

The aim of this work is to produce a first step towards creating general tools that MS can use to evaluate data accuracy but it is not possible to cover all scenarios given the time and resources available. It is thus necessary to restrict the applicability of this work to the following requirements:

1. The data is collected by commercial fisheries sampling programmes performed by a single institute;
2. The sampling programmes considered should be probabilistic;
3. The sampling can be multi-stage, with stratification at any or all levels. Units may be selected with or without replacement;
4. Estimation of the desired parameters should be by design-based estimation;
5. The data is available in the ICES Regional Database & Estimation System (RDBES) format;
6. The sampling data should not require the use of any of the cluster sampling variables defined within the RDBES data model.

2.3 Approach

The evaluation of data precision has been performed using two complementary techniques. For relatively simple sampling designs it is possible to use analytical functions to calculate the precision (or a related statistical measure such as variance) of a statistical estimate. In Section 3 we present these calculations and implementations of them in R code. For more complicated sampling designs, the use of analytical functions is usually not feasible. In these cases, it is necessary to evaluate precision using resampling techniques such as bootstrapping. Section 4 discusses when bootstrapping is appropriate and gives a number of worked examples describing how bootstrapping can be applied in different cases.

The evaluation of bias is a difficult subject and is hard to quantify. The approach presented in Section 5 builds on the previous work available in the ICES literature to identify and enumerate common sources of bias in catch sampling programs. The information is collated and an evaluation is then performed as to whether data stored using the RDBES data format can inform about that bias source. Reports are presented that can help member states to identify deviations in their sampling programmes and sampling variability that can potentially lead to bias in catch estimates.

---

8 https://github.com/ices-tools-dev/RDBES
9 XXselectionMethodCluster, XXnumberTotalClusters, XXnumberSampledClusters, XXselectionProbCluster, XXinclusionProbCluster
3 Analytical calculation of variance

This section presents the variance calculations for design-based estimation using a three-stage sampling design.

3.1 Sampling without replacement in all three stages

Consider the following sampling design in three stages where the primary sampling units are vessels, the secondary sampling units are trips and the tertiary sampling units are hauls.

Stage I: Sampling of vessels
A random sample without replacement of vessels is drawn from all the vessels in the population. The set of vessels in the population is denoted \( U_I \) of size \( N_I \) and the sample of vessels is denoted \( S_I \) of size \( n_I \). Each vessel is looked upon as a cluster of trips.

Stage II: Sampling of trips
For every vessel \( i \) selected in stage I, a random sample without replacement of trips is drawn from all the trips associated with the vessel. The set of trips associated with vessel \( i \) is denoted \( U_{II} \) of size \( N_{II} \) and the sample of trips is denoted \( S_{II} \) of size \( n_{II} \). Each trip is looked upon as a cluster of hauls.

Stage III: Sampling of hauls
For every trip \( q \) selected in stage II, a random sample without replacement of hauls is drawn from all the hauls associated with the trip. The set of hauls associated with trip \( q \) is denoted \( U_{III} \) of size \( N_{III} \) and the sample of hauls is denoted \( S_{III} \) of size \( n_{III} \).

For each haul \( k \) selected in stage III, the weight of discards, \( y_k \), is observed. The problem is to estimate the total weight of discards for all possible hauls, trips and vessels,

\[
t_y = \sum_{u_I} \sum_{u_{II}} \sum_{u_{III}} y_k
\]

and the variance of this estimator. To accomplish this, we need the inclusion probabilities for each stage.

3.1.1 Inclusion probabilities for the general case

For stage I, the first order inclusion probability \( \pi_{II} \) is the probability of vessel \( i \) to be included in the sample \( s_I \). The second-order inclusion probability \( \pi_{IIj} \) is the joint probability of vessel \( i \) and \( j \) to be included in \( s_I \).

For stage II, the first order inclusion probability \( \pi_{IIq|I} \) is the conditional probability of trip \( q \) to be included in the sample \( s_{II} \) (conditional on the stage I sampling). The second-order inclusion probability \( \pi_{IIqr|I} \) is the conditional joint probability of trip \( q \) and \( r \) to be included in \( s_{II} \).

For stage III, the first order inclusion probability \( \pi_{III|IIq} \) is the conditional probability of haul \( k \) to be included in the sample \( s_{III} \) (conditional on stage I and II sampling). The second-order inclusion probability \( \pi_{III|IIql} \) is the conditional joint probability of haul \( k \) and \( l \) to be included in \( s_{III} \).

We summarize these general inclusion probabilities in the table below:
### Inclusion probabilities, general

<table>
<thead>
<tr>
<th>Stage</th>
<th>First-order</th>
<th>Second-order</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>$\pi_{i}$</td>
<td>$\pi_{ij}$</td>
</tr>
<tr>
<td>II</td>
<td>$\pi_{i</td>
<td>q</td>
</tr>
<tr>
<td>III</td>
<td>$\pi_{k</td>
<td>i</td>
</tr>
</tbody>
</table>

(Note that $\pi_{i | i} = \pi_{i}$; $\pi_{i | q | i} = \pi_{i | q | i}$; $\pi_{k | k | q} = \pi_{k | q}$.)

#### 3.1.2 Estimation for the general case

In general, the Horvitz-Thompson (HT) estimator of $t_y$ with respect to all three stages is given by

$$\hat{t}_y = \sum \frac{1}{\pi_{i}} \sum \frac{1}{\pi_{i | q | i}} \sum \frac{y_k}{\pi_{k | i | q}}$$

We can also write $\hat{t}_y$ as

$$\hat{t}_y = \sum \frac{\hat{t}_i}{\pi_{i}}$$

where $\hat{t}_i$ is the HT estimator of the total weight of discards for vessel $i$ with respect to stage II and III:

$$\hat{t}_i = \sum \frac{1}{\pi_{i | q | i}} \sum \frac{y_k}{\pi_{k | i | q}}$$

Similarly, the estimator $\hat{t}_i$ can be written as

$$\hat{t}_i = \sum \frac{\hat{t}_{iq}}{\pi_{i | q | i}}$$

where $\hat{t}_{iq}$ is the HT estimator of the total weight of discards for trip $q$ with respect to stage III:

$$\hat{t}_{iq} = \sum \frac{y_k}{\pi_{k | i | q}}$$

An unbiased estimator of the variance of $\hat{t}_y$ is given by

$$\hat{\sigma}^2(\hat{t}_y) = \sum \left( \frac{1}{\pi_{i}} \sum \frac{1}{\pi_{i | q | i}} \sum \frac{y_k}{\pi_{k | i | q}} \right)^2 + \sum \frac{\hat{\sigma}_i^2}{\pi_{i}}$$

where

$$\hat{\sigma}_i^2 = \sum \left( \frac{1}{\pi_{k | i | q}} \sum \frac{y_k}{\pi_{k | i | q}} - \frac{\hat{t}_{iq}}{\pi_{i | q | i}} \right)^2$$

(Note that for the point estimator we only use the first order inclusion probabilities. For the variance estimator, we also need the second-order inclusion probabilities.)

#### 3.1.3 Inclusion probabilities for Simple Random Sampling (SRS) without replacement in each stage

The inclusion probabilities valid for the case of SRS without replacement in each stage are given in the table below.
3.1.4 Estimation for SRS without replacement in each stage

For SRS without replacement in each stage, the HT estimator of $t_y$ simplifies into

$$\hat{t}_y = \sum_{s_i} \frac{N_i}{n_i} \sum \frac{N_{iIII}}{N_{iIII}} \sum \frac{N_{iq}}{n_{iq}} \sum y_k = \frac{N_i}{n_i} \sum \frac{N_{iIII}}{N_{iIII}} \sum \frac{N_{iq}}{n_{iq}} \sum y_k$$

The estimator can also be written as

$$\hat{t}_y = \frac{N_i}{n_i} \sum s_i \hat{t}_i$$

where

$$\hat{t}_i = \frac{N_{iIII}}{n_{iIII}} \sum \frac{N_{iq}}{n_{iq}} \sum y_k = \frac{N_{iIII}}{n_{iIII}} \sum \frac{N_{iq}}{n_{iq}} \sum \hat{t}_{iq}$$

and

$$\hat{t}_{iq} = \frac{N_{iq}}{n_{iq}} \sum s_{iq} y_k$$

An unbiased estimator of the variance of $\hat{t}_y$ is given by

$$\mathbb{V}(\hat{t}_y) = N_i^2 \left[ 1 - \frac{n_i}{N_i} \right] S_{\hat{t}_y}^2 + \frac{N_i}{n_i} \sum s_i \left[ \left( \frac{1 - n_i}{N_i} \right) S_{\hat{t}_y}^2 + \left( \frac{n_i}{N_i} \right) S_{\hat{t}_{iq}}^2 \right]$$

where

$$S_{\hat{t}_y}^2 = \frac{1}{n_i - 1} \sum s_i \left[ \hat{t}_i - \left( \sum s_i \hat{t}_i / n_i \right) \right]^2;$$

$$S_{\hat{t}_{iq}}^2 = \frac{1}{n_{iIII} - 1} \sum s_{iIII} \left[ \hat{t}_{iq} - \left( \sum s_{iIII} \hat{t}_{iq} / n_{iIII} \right) \right]^2;$$

$$S_{y_{iq}}^2 = \frac{1}{n_{iq} - 1} \sum s_{iq} \left[ y_k - \left( \sum s_{iq} y_k / n_{iq} \right) \right]^2.$$
\[ \hat{y} = \sum_{i} \left( \frac{\pi_{iij} - \pi_{iij} \tilde{t}_{ij}}{\pi_{iij}} \right) \]

Under SRS without replacement in all stages, this would mean using

\[ \hat{y} = N_t^2 \frac{1 - n_i/N_t}{n_i} S_{et}^2 \]

If the sample size in stage I is fixed, an alternative is to use the abridged Yates-Grundy estimator

\[ \hat{y} = -\frac{1}{2} \sum_{i} \sum_{j} \frac{\pi_{iij} - \pi_{iij} \tilde{t}_{ij}}{\pi_{iij}} \left( \frac{\tilde{t}_i - \tilde{t}_j}{\pi_{iij}} \right)^2 \]

In both cases, this would lead to underestimation of the true variance. However, if the variance contributions from stage II and III are small, this underestimation might not be so important.

Another option is to do the variance estimation as if vessels were selected with replacement in stage I. The estimation formula for this situation is given in the next section. This approach might in general lead to both over- and underestimation of the true variance.

### 3.2 Sampling with replacement in the first stage

Consider again a sampling design in three stages where the primary sampling units are vessels, the secondary sampling units are trips and the tertiary sampling units are hauls. The difference from the design in section 1 is that the sampling is done with replacement in the first stage whereas the sampling in subsequent stages is still without replacement.

#### Stage I: Sampling of vessels

A random sample with replacement of vessels is drawn from all the vessels in the population in such a way that, at every draw, \( p_i \) is the probability of selecting vessel \( i \). The set of vessels in the population is denoted \( U_i \) of size \( N_i \). The ordered sample of vessels is denoted \( \{i_1, \ldots, i_{m_i}\} \), where \( i_v \) is the vessel selected in draw number \( v \) and \( m_i \) is the number of draws. Each vessel is looked upon as a cluster of trips.

#### Stage II: Sampling of trips

For every vessel drawing \( i_v \) in stage I, a random sample without replacement of trips is drawn from all the trips associated with the vessel. The set of trips associated with vessel drawing \( i_v \) is denoted \( U_{iit} \) of size \( N_{iit} \) and the sample of trips is denoted \( \{i_{1t}, \ldots, i_{nt}\} \) of size \( n_{i1t} \).

#### Stage III: Sampling of hauls

For every trip \( q \) selected in stage II, a random sample without replacement of hauls is drawn from all the hauls associated with the trip. The set of hauls associated with trip \( q \) is denoted \( U_{iqt} \) of size \( N_{iqt} \) and the sample of hauls is denoted \( \{i_{1qt}, \ldots, i_{ntq}\} \) of size \( n_{i1qt} \).

We assume that the sampling in stage II and III has the properties of invariance and independency.

#### 3.2.1 Estimation for the general case

In general, the Hansen-Hurwitz (HH) estimator of \( t_y \) with respect to all three stages is given by

\[ \hat{t}_y = \frac{1}{m_i} \sum_{v=1}^{m_i} \frac{\tilde{t}_{iq}}{p_{iq}} \]
where $\hat{t}_{iv}$ is the HT estimator of the total weight of discards for vessel drawing $i_v$ with respect to stage II and III:

$$\hat{t}_{iv} = \sum_{s_{iiqv}} \frac{1}{n_{iiqv}} \sum_{s_{ijq}} \frac{y_k}{n_{ijq}}$$

An unbiased estimator of the variance of $\hat{t}_y$ is given by

$$\mathcal{V}(\hat{t}_y) = \frac{1}{m_i(m_i - 1)} \sum_{v=1}^{m_p} \left( \frac{\hat{t}_{iv}}{p_{iv}} - \frac{\hat{t}_y}{N_i} \right)^2$$

(see: Särndal et al., 1992, Result 4.5.1).

### 3.2.2 Estimation for SRS with replacement in the first stage

For SRS with replacement in stage I, the drawing probability $p_{iv}$ is equal to $1/N_i$ for all vessel drawings $i_v$. If SRS without replacement is used in stage II and III, the HH estimator of $t_y$ simplifies into

$$\hat{t}_y = \frac{1}{m_i} \sum_{v=1}^{m_p} \frac{\hat{t}_{iv}}{p_{iv}} = \frac{N_i}{m_i} \sum_{v=1}^{m_p} \hat{t}_{iv}$$

where

$$\hat{t}_{iv} = \frac{N_{iiqv}}{n_{iiqv}} \sum_{s_{iiqv}} \frac{y_k}{n_{ijq}}$$

An unbiased estimator of the variance of $\hat{t}_y$ is given by

$$\mathcal{V}(\hat{t}_y) = \frac{N_i^2}{m_i(m_i - 1)} \sum_{v=1}^{m_p} \left( \hat{t}_{iv} - \frac{\hat{y}_u}{N_i} \right)^2$$

where $\hat{y}_u = \hat{t}_y/N_i$.

### 3.2.3 References


### 3.3 Implementation in R

Annex A3.2 below presents an implementation of these analytical variance calculations using the RDBES data model. It specifically considers Hierarchy 1 but is generalizable to all hierarchies.
4 Bootstrapping

4.1 Introduction

Bootstrapping (Efron and Tibshirani, 1986) in the context of catch sampling is a simulation method for approximating the sampling design and estimation procedures commonly used for providing stock assessors with desired means and variances. In the case of the EU, the designs usually involve multistage sampling, often stratified at one or more stages of sampling, and the estimation procedures are usually design-based.

The general approach when bootstrapping is to,

1. Use the original sample data to obtain the estimates of interest;
2. Repeatedly sample from the original dataset, each time following the original sampling design and estimation procedures as much as possible;
3. Using those “bootstrapped” estimates from (2) to assess bias or to estimate sampling variability of the estimates calculated in step (1).

The basic idea behind bootstrapping is that the original sample (S) is representative of the population (P) from which it was taken, where “representative” means that the sample is sufficiently large to capture the range of values in P, the variability inherent in P, and the frequency distribution of those values in P. Hence, S acts as a pseudo-population from which one can repeatedly sample to observe the behaviour of the design and estimators. The approach works well when the sample sizes are large and the sample data cover the characteristics of the population from which they were taken.

On the other hand, S can be a possibly poor representation of P in several situations, ranging from small sample sizes to poorly chosen strata with mismatched strata sample sizes to poor selection procedures, such as always choosing the fishing operations with the largest landings because one wishes to obtain as many species as possible. The issue of deciding whether a sample size is appropriate or sufficiently large is a difficult one as it depends on 1) the size of the sampling frame, e.g. 10 hauls/trip and 3 hauls are sampled, 2) the choices for stratum sample sizes, e.g. one stratum has a large sampling frame and another a small one but the same number of units were sampled in each, 3) the variability of the sampling frame. For example, the sample of 3 hauls on a fishing trip with 10 hauls can be bootstrapped if it is reasonable to assume that the unsampled trips have similar characteristics and variability displayed in the 3 sampled trips. It would not be appropriate if the 3 sampled hauls were always chosen because of some characteristic that distinguished them from the other hauls on the trip. One must keep in mind that even if the analyst is comfortable with the sample sizes used in the bootstrapping, the entire exercise is based on the observed sample. If that sample is not fully representative, then one would not obtain similar results if a different sample had been used. This is of course true even if the estimates of interest (e.g. standard errors) were analytically calculated since the entire exercise is based on the data available.

In addition to the issue of representation, one cannot estimate variance among sampling units using a design-based estimation approach when only one sampling unit is chosen. Hence, like the analytical method, the variance is underestimated due to the lack of information for one or more stages of the design.

Sampling may not be fully appropriate for bootstrapping when the population is large but a very small sample was taken, e.g. when P contains 1000 vessel × trips but only 8 were sampled. It is unlikely that one can argue successfully that the 8 observations are representative of the entire
population of vessel × trips. Again, this is true whether the analyst is bootstrapping or using an analytical method; the analyst needs to be aware that the bootstrapping cannot overcome the effect the small sample size has on the resulting estimates.

Another instance where bootstrapping will not correct internal flaws in the dataset is when the sampling design cannot be replicated using computer code. An example of this is convenience sampling. Hence, bootstrapping is unlikely to be of use for any non-probabilistic method of obtaining S that requires strict assumptions (e.g. assuming the convenience sample is a reasonable facsimile of SRSWR) unlikely to be true.

Bootstrapping works best when the sample sizes are sufficiently large that one is comfortable that S is representative of P, when the sample data were collected according to a design strategy that can be replicated in computer code, and when the estimation procedures can be reproduced without error.

In the case of probabilistic (simple random or unequal probability) with replacement sampling at all stages of the design, the original sample dataset is the pseudo-population. Multistage sampling can be done directly and repeatedly on the original sample following the original design and original sample sizes at each stage and in each stratum. For each new “bootstrap” multistage sample dataset, the desired estimates can be calculated and stored. After a large number of repetitions of the sampling from the pseudo-population, the bias and precision of the estimators used on the original data can be assessed by calculating the mean and variance of the bootstrapped estimates. Pseudo-code for performing this bootstrapping procedure using the sampling design in Example 1 of Hierarchy 1 in the RDBES Data Model documentation is given in the next section and some R code (Box 2) for a slightly different multi-stage sampling design is shown in Annex A3.3 below.

In the case of probabilistic (simple random or unequal probability) where sampling is without replacement at one or more stages of the design, the original sample dataset is used to construct a new pseudo-population. The pseudo-population for the stage sampled without replacement contains multiple copies of the original dataset S from that level. We need a bit of notation here. Let the population have N sampling units and an SRSWOR of n vessels is selected. In the simplest case, suppose \( \frac{n}{N} = k \) is an integer (e.g. \( N = 100, n = 20 \), then \( k = 5 \)). Then, the pseudo-population is simply the set of \( k \) replications of the sample. So, in the case of SRSWOR of PSUs but SRSWR at all lower stages, the pseudo-population would be the \( k \) copies of the sample of PSUs and all child sample data associated with those PSUs.

In the case where sampling is SRS but \( k \) is not an integer (e.g. \( N = 100, n = 15 \), then \( k = 6.667 \)), there are three approaches possible. First, if the sample size is small relative to the population size, then the SRSWOR is treated as an SRSWR and the pseudo-population is the sample. A commonly used cutoff is to treat an SRSWOR as being with replacement when \( \left( 100 \times \frac{a}{N} \right) \% < 5 \). Second, is to construct a pseudo-population that contains \( k \) copies of S (\( \lfloor a \rfloor \) is the largest integer less than \( a \)) plus \( N - n \times \lfloor k \rfloor \) randomly chosen units from S. In this approach, the pseudo-population has two sources of variability: the original sampling variability plus a variability due to the selection of the additional units to fill out the population to its full size. Hence, any bootstrapping based on this method usually includes a loop to repeatedly create new pseudo-populations and to perform bootstrapping on each. We do not recommend this approach because the pseudo-population characteristics such as its mean and variance vary among the different random realizations of the pseudo-population. The third method is a simpler and more useful approach to creating a pseudo-population from which bootstrap SRSWOR can be taken. In this approach (Bickel and Freedman, 1984; Chao and Lo, 1985; Sitter, 1992) there is randomization between two different pseudo-populations made up of either \( k \) or \( \lfloor k \rfloor + 1 \) copies of the sample S so that in either case, the mean of each of the pseudo-populations matches the mean of S.
We recommend the Sitter (1992) method for constructing the pseudo-population since it has some nice behaviours relative to the bootstrapping that is planned. Some example R code is provided in Boxes 3 and 4 in Annex A3.3 below to demonstrate the method for single-stage and multi-stage designs.

When constructing pseudo-populations for multi-stage sampling designs where SRSWOR occurs at more than one level, the recommended approach (Sitter, 1992) is to construct the pseudo-population for each stage with WOR sampling before and during bootstrapping. For example, suppose both the PSUs and SSUs are selected using SRSWOR and the TSUs are selected by SRSWR. The PSUs might be vessel × trips, the SSUs, fishing operations nested within each vessel × trip, and the TSUs are individual fish (or possibly samples of X buckets of unsorted fish) that are treated as having been selected by SRSWR. Construction of the pseudo-population would start by replicating the sampled SSUs the required number of times to “fill out” all operations for each sampled PSU following the Sitter procedure. Once each sampled PSU has been fully recreated and the desired quantities obtained for that SSU, the pseudo-population of PSU would be constructed by again following the method described by Sitter (1992) and then re-sampling from the pseudo-population.

Once the pseudo-population is created for a multi-stage sampling design, bootstrapping proceeds similar to that for SRSWR. Multistage sampling is done directly and repeatedly on the pseudo-population following the original design and appropriate sample sizes at each stage and in each stratum. For each new “bootstrap” multistage sample dataset, the desired estimates can be calculated and stored. After many repetitions from the pseudo-population, the bias and precision of the estimators used on the original data can be assessed by calculating the mean and variance of the bootstrapped estimates. R code for performing the SRSWOR bootstrapping procedure using the sampling design in Example 1 of Hierarchy 1 is given in Box 4 in Annex A3.3 below.

### 4.2 Types of data needed for input to bootstrapping catch sampling

For the examples given in this document, there are generally two or three datasets needed for input to the bootstrapping simulations. Tables 4.1–4.3 list the more common variables needed for calculating numbers at length, numbers at age and total discard weight for a given species within a given fishery within a stratum. The tables assume the stratum is quarter; modifications can be made if there are other strata at different stages of the sampling. The lists assume that fish are sampled for length from a fishing operation from a trip on a vessel within a stratum.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vessel ID</td>
<td>Unique identifier of the sampled vessel from the reference fleet</td>
</tr>
<tr>
<td>Quarter</td>
<td>1, 2, 3, 4</td>
</tr>
<tr>
<td>Trip ID</td>
<td>Unique identifier of the sampled trips from the vessel in the quarter</td>
</tr>
<tr>
<td>Fishing Operation ID</td>
<td>Unique identifier of the sampled FO within the sampled trip</td>
</tr>
<tr>
<td>Species ID</td>
<td>Unique identifier of the species for which total discard weight is of interest</td>
</tr>
<tr>
<td>Discard Weight</td>
<td>Observed discard weight for the species in the sampled FO</td>
</tr>
</tbody>
</table>
### Table 4.2. Variables needed to estimate numbers at length (NAL) by stratum. The list is based on providing an estimate for a single year. Additional information could be required if there is a need for one or more conversions (such as converting fork length to total length).

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landed Weight</td>
<td>Observed landed weight for the species in the sampled FO</td>
</tr>
<tr>
<td>Total Landed Weight</td>
<td>Reported total landed weight for the entire fishery from which vessels, trips and FOs were sampled in the listed quarter</td>
</tr>
</tbody>
</table>

### Table 4.3. Variables needed to estimate age-length-key (ALK) by stratum. The list is based on providing an estimate for a single year and assumes that numbers at length are separately calculated before creating the numbers at age (NAA) for the stratum.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vessel ID</td>
<td>Unique identifier of the sampled vessel from the reference fleet</td>
</tr>
<tr>
<td>Quarter</td>
<td>1, 2, 3, 4</td>
</tr>
<tr>
<td>Trip ID</td>
<td>Unique identifier of the sampled trips from the vessel in the quarter</td>
</tr>
<tr>
<td>Fishing Operation ID</td>
<td>Unique identifier of the sampled FO within the sampled trip</td>
</tr>
<tr>
<td>Catch Category</td>
<td>Catch category (DIS, LAN, BMS, ALL, .....)</td>
</tr>
<tr>
<td>Species ID</td>
<td>Unique identifier of the species for which total discard weight is of interest</td>
</tr>
<tr>
<td>Length Class</td>
<td>Number indicating an observed length class for the species in this FO</td>
</tr>
<tr>
<td>Length Unit</td>
<td>mm, cm, scm, ...</td>
</tr>
<tr>
<td>Length Type</td>
<td>Type of measurement (total, fork, ...)</td>
</tr>
<tr>
<td>Number at Length</td>
<td>Number of fish in the length class in the FO sample</td>
</tr>
<tr>
<td>Sample Weight of FO</td>
<td>Weight of the sample in the listed catch category in the FO</td>
</tr>
<tr>
<td>Total Weight of FO</td>
<td>Total weight of the species in the listed catch category in the FO</td>
</tr>
<tr>
<td>Total Weight of Trip</td>
<td>Total weight of the species in the listed catch category in the sampled trip over all FOs, sampled or not</td>
</tr>
<tr>
<td>Total Weight of Stratum</td>
<td>Total weight of the species in the listed catch category over all trips, sampled or not, within the quarter</td>
</tr>
<tr>
<td>Catch Category</td>
<td>Catch category (DIS, LAN, BMS, ALL, .....)</td>
</tr>
</tbody>
</table>
### Variable Name | Definition
--- | ---
Species ID | Unique identifier of the species for which total discard weight is of interest
Fish ID | Unique identifier for an individual fish in the sampled species in this sampled OF in the sampled trip
Length Class | Number indicating an observed length class for the species in this FO
Length Unit | mm, cm, scm, ....
Length Type | Type of measurement (total, fork, ...)
Age Class | Number indicating the estimated age of the fish sampled in the FO on this trip
Age Unit | Year, month, ....

#### 4.3 Notation

\(x(y)\) indicates \(x\) is nested within \(y\), with Area/Quarter Strata (note that simpler stratification would remove some of the subscripts)

- \(N_a = \text{total number of vessels in reference fleet in the area stratum } a, a = 1, ..., A\) (\(A\) is the total number of areal reference fleets)
- \(n_a = \text{number of vessels selected from the reference fleet in the area stratum } a\)
- \(T_{v(a,q)} = \text{total number of trips taken by vessel } v \text{ within area } a, v(a,q) = 1, 2, ..., n_a, \text{ and quarter } q, q = 1, 2, 3, 4\) (note that this could be 0 for some combinations of quarters, areas and vessels)
- \(t_{v(a,q)} = \text{number of trips observed for vessel } v \text{ within area } a \text{ and quarter } q\)
- \(H_{t(v,a,q)} = \text{total number of fishing operations performed on trip } t \text{ for vessel } v \text{ within area } a \text{ and quarter } q\)
- \(LW_{h(t,v,a,q)} = \text{total landings weight for the } h^{th} \text{ fishing operation performed on trip } t \text{ for vessel } v \text{ within area } a \text{ and quarter } q \text{ (reported)}, h = 1, ..., H_{t(v,a,q)}\)
- \(DW_{h(t,v,a,q)} = \text{total discard weight for the } h^{th} \text{ fishing operation performed on trip } t \text{ for vessel } v \text{ within area } a \text{ and quarter } q \text{ (usually estimated)}, h = 1, ..., H_{t(v,a,q)}\)
- \(h_{t(v,a,q)} = \text{number of fishing operations observed on trip } t \text{ for vessel } v \text{ within area } a \text{ and quarter } q\)
- \(Lw_{h(t,v,a,q)} = \text{sampled landings weight for the } h^{th} \text{ fishing operation sampled on trip } t \text{ for vessel } v \text{ within area } a \text{ and quarter } q \text{ (reported)}, h = 1, ..., h_{t(v,a,q)}\)
- \(Dw_{h(t,v,a,q)} = \text{sampled discard weight for the } h^{th} \text{ fishing operation sampled on trip } t \text{ for vessel } v \text{ within area } a \text{ and quarter } q \text{ (usually estimated)}, h = 1, ..., h_{t(v,a,q)}\)
- \(Fl_{h(t,v,a,q)} = \text{total number of fish in the sampled landings for the } h^{th} \text{ fishing operation sampled on trip } t \text{ for vessel } v \text{ within area } a \text{ and quarter } q \text{ (usually estimated)}, h = 1, ..., h_{t(v,a,q)}\)
- \(f{L}_{h(t,v,a,q)} = \text{number of fish sampled for length in the sampled landings for the } h^{th} \text{ fishing operation sampled on trip } t \text{ for vessel } v \text{ within area } a \text{ and quarter } q, h = 1, ..., h_{t(v,a,q)}\)
- \(FD_{h(t,v,a,q)} = \text{total number of fish in the sampled discards for the } h^{th} \text{ fishing operation sampled on trip } t \text{ for vessel } v \text{ within area } a \text{ and quarter } q \text{ (usually estimated)}, h = 1, ..., h_{t(v,a,q)}\)
$f_D_{h(t,v,a,q)}$ = number of fish sampled for length in the sampled discards for the $h^{th}$ fishing operation sampled on trip $t$ for vessel $v$ within area $a$ and quarter $q$, $h = 1, \ldots, h_{t(v,a,q)}$

$f_{aL_{l,h(t,v,a,q)}}$ = number of fish sampled for age within length class $l$, $l = 1, \ldots, \min(X, f_{L_{h(t,v,a,q)}})$, (where $X$ is a species-specific upper limit of the number of fish ideally sampled from each length class) in the sampled landings for the $h^{th}$ fishing operation sampled on trip $t$ for vessel $v$ within area $a$ and quarter $q$, $h = 1, \ldots, h_{t(v,a,q)}$

$f_{aD_{l,h(t,v,a,q)}}$ = number of fish sampled for age within length class $l$, $l = 1, \ldots, \min(X, f_{D_{h(t,v,a,q)}})$, (where $X$ is a species-specific number of fish to be sampled from each length class) in the sampled discards for the $h^{th}$ fishing operation sampled on trip $t$ for vessel $v$ within area $a$ and quarter $q$, $h = 1, \ldots, h_{t(v,a,q)}$

4.4 Pseudo-code for running bootstrap simulations of example 1 from hierarchy 1 assuming all stages are SRSWR

A simple example of how bootstrapping is performed: Assume vessels are randomly selected with replacement (WR) from a reference fleet. Trips from these vessels are then randomly sampled within each quarter (stratum). Fishing operations within each selected trip are randomly selected with replacement from multiple fishing operations that occurred on the trip. Within a fishing operation, fish are randomly selected for length measurements and assigned to length classes once measured. For each observed length class, a subset of fish is randomly selected for ageing. Note that this code is easily modified to allow for the case where fish are selected for ageing from a self-sample that is separate from the sample of fish measured for length.

In reality, no stage is actually sampled WR, but this bootstrapping approach could be used for SRSWOR at every stage if one desires a conservative estimate of the variance of the estimates, i.e. a variance estimate larger than the WOR sampling variance.

4.4.1 Basic bootstrapping steps for a single simulated sampling effort

1. Take an SRSWR of $n$ vessels from the $N$ vessels in the reference fleet.

   Within each quarter:

2. For each vessel selected in step (1), take an SRSWR of $t_{v(q)}$ trips from the full set of $T_{v(q)}$ trips by the vessel in that quarter.

3. For each trip selected in step (2), take an SRSWR of $h_{t(v,a,q)}$ fishing operations from the full set of $H_{t(v,a,q)}$ fishing operations on that trip by that vessel:

   a) If estimates of discard weights for the quarter are desired, then perform required calculations.

   If biological parameter estimates are desired:

4. (Figure 2) For each fishing operation selected in step (3), take a

   a) SRSWR of $f_{L_{h(t,v,a,q)}}$ individual fish for length measurements in the landings and
   b) SRSWR of $f_{D_{h(t,v,a,q)}}$ individual fish for length measurements in the discards

5. If Numbers at Length (NAL) are required, perform required fishing operation-level (or trip-level) calculations (depending on aggregation level).

   a) Store the $NAL_q^b$ calculated in step (5) in a temporary file. The notation indicates that it is the estimated NAL for quarter $q$ obtained from bootstrap simulation $b$
6. For each length class observed in step (4), take a
   a) SRSWR of \( f_aL_{i,h(t,v,q)} \) individual fish for age/weight measurements in the landings
   b) SRSWR of \( f_aD_{i,h(t,v,q)} \) individual fish for age/weight measurements in the discards
7. For stratum-level (quarterly) Age-Length Key (ALK\(_b^q\)), use all data from step (5) combined
   over all trips within the stratum
8. For calculating the quarterly Numbers at Age (NAA) or mean length at age (\( \overline{LAA} \)), use
   the NAL estimate from step (5) raised to the appropriate level and the ALK estimate from
   step (7)
   a) Store the \( NAA_b^q \) calculated in step (8) in a temporary file. The notation indicates that
      it is the estimated NAA for quarter \( q \) obtained from bootstrap simulation \( b \)
   b) Store the mean length at age \( \overline{LAA}_b^q \) from step (8) in a temporary file.
9. Repeat steps (2) to (8) for each quarter to obtain the estimated \( NAA_b^q \), \( NAL_b^q \), and \( \overline{LAA}_b^q \) for
    every quarter.
   a) Calculate and store the bootstrap estimates of the annual \( NAA_b^p \), \( NAL_b^p \), and \( \overline{LAA}_b^p \) using
      the usual stratified weighted mean of the quarterly estimates
10. Repeat steps (1) through (9), \( B \) times to obtain \( B \) estimates of \( NAA_b^p \), \( NAL_b^p \), and \( \overline{LAA}_b^p \) and
    if desired the quarterly estimates \( NAA_q^b \), \( NAL_q^b \), and \( \overline{LAA}_q^b \).
11. The \( B \) estimates from step (10) can be used to
   a) Assess bias within each quarter by calculating the mean of the bootstrapped \( NAA_q^b \),
      \( \overline{LAA}_q^b \), and \( NAL_q^b \) values from the \( B \) simulated samplings and comparing the means
      to the \( NAA_q \), \( \overline{LAA}_q \), and \( NAL_q \) calculated using the original dataset
   b) Assess bias of the annual estimates by calculating the mean of the bootstrapped
      \( NAA_b^p \), \( NAL_b^p \), and \( \overline{LAA}_b^p \) values from the \( B \) simulated samplings and comparing
      the means to the \( NAA \), \( \overline{LAA} \), and \( NAL \) calculated using the original dataset
   c) Assess the standard errors (or variances or relative standard errors) of the original
      quarterly estimates of \( NAA_q \), \( \overline{LAA}_q \), and \( NAL_q \) by calculating the standard errors (or
      variance or relative standard errors) of the \( B \) estimates of \( NAA_q^b \), \( NAL_q^b \), and \( \overline{LAA}_q^b \).
   d) Assess the standard errors (or variances or relative standard errors) of the original
      annual estimates of \( NAA \), \( \overline{LAA} \), and \( NAL \) by calculating the standard errors (or
      variance or relative standard errors) of the \( B \) estimates of \( NAA_b^p \), \( NAL_b^p \), and \( \overline{LAA}_b^p \).
   e) Assess the correlation among the NAA or NAL by calculating the variance-covariance
      matrix for the desired vector. This is of use for determining the effect of the
      sampling strategy on the distribution of the vector of numbers at age or length (Note
      that a typical approach by a stock assessor is to consider the values in a NAA or NAL
      vector to be independent and distributed as a multinomial. The reality is they are
      not!).
4.5 Example code for bootstrapping a stratified multi-stage sampling design assuming all stages are SRSWR

In this example, there is a single fish species for which biological parameters, numbers at length and numbers at age, are desired by stratum. Strata are combinations of quarter, area and gear (indicated by “qX geartypeY areaZ”) and all together there are five strata. There are three stages of sampling:

- vessel × trip (PSU) within strata;
- individual fish for length measurements on a trip (SSU) where only a single fishing operation was sampled;
- and length-stratified sub-sampling of fish (TSU) for aging and weight measurements.

The catch category (DIS or LAN) is also available and could have been included as a stratum at the SSU level but was not used in this example.

The bootstrapping R code in Box 2 in Annex A3.3 below assumes that there are two data files (Box 1) for this example: the first file (“fish”) contains a record for each length class observed on each sampled trip of the number of fish observed in that length class on that trip and the second file (“indfish”) contains individual fish records for age, length class, and weight for each trip where fish were sampled. The sampling for age was stratified by length class and the age sub-sample size was the maximum of 10 fish or the number of fish observed in that length class at
the second stage. This section is easily modified to allow for a separate self-sample of fish from a trip to be used for developing the ALK.

The example R code was developed to assess the bias and precision of the quarterly estimated NAL and NAA from the original datasets. In the code, these estimates are not expanded to the entire population or even to the sampling frame since data on total discard or landing weight by stratum was not available; instead, the results from this analysis provide the estimated NAL and NAA for the sampled trips only. Hence, many steps that are needed to expand the bootstrap estimates to the overall totals for the entire fishery are not shown. Note though that the expansion factors are constants and so it should not be difficult to include the additional computational steps. One need only have the values stored and accessible to the R program environment.

Box 1. A) Example of the trip level dataset “fish”. B) Example of the subsample data for individual fish ages and weights “indfish”.

<table>
<thead>
<tr>
<th>NewStratum</th>
<th>CatchCat</th>
<th>UniqTrip</th>
<th>Length</th>
<th>SumStationWt</th>
<th>SumSampleWt</th>
<th>SumNumAtLen</th>
</tr>
</thead>
<tbody>
<tr>
<td>q12Active22-24</td>
<td>DIS</td>
<td>201</td>
<td>22</td>
<td>0.312</td>
<td>0.097</td>
<td>1</td>
</tr>
<tr>
<td>q12Active22-24</td>
<td>DIS</td>
<td>201</td>
<td>120</td>
<td>19.570</td>
<td>4.541</td>
<td>1</td>
</tr>
<tr>
<td>q12Active22-24</td>
<td>DIS</td>
<td>201</td>
<td>140</td>
<td>19.570</td>
<td>4.541</td>
<td>2</td>
</tr>
<tr>
<td>q12Active22-24</td>
<td>DIS</td>
<td>201</td>
<td>160</td>
<td>19.570</td>
<td>4.541</td>
<td>3</td>
</tr>
<tr>
<td>q12Active22-24</td>
<td>DIS</td>
<td>201</td>
<td>170</td>
<td>19.570</td>
<td>4.541</td>
<td>3</td>
</tr>
<tr>
<td>q12Active22-24</td>
<td>DIS</td>
<td>201</td>
<td>180</td>
<td>19.570</td>
<td>4.541</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RatioStaWtSampWt</th>
<th>TrpNumAtLen</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.216495</td>
<td>4</td>
</tr>
<tr>
<td>4.309623</td>
<td>5</td>
</tr>
<tr>
<td>4.309623</td>
<td>9</td>
</tr>
<tr>
<td>4.309623</td>
<td>13</td>
</tr>
<tr>
<td>4.309623</td>
<td>13</td>
</tr>
<tr>
<td>4.309623</td>
<td>22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UniqTrip</th>
<th>CatchCat</th>
<th>Length</th>
<th>Age</th>
<th>Weight</th>
<th>NewStratum</th>
</tr>
</thead>
<tbody>
<tr>
<td>201820301</td>
<td>DIS</td>
<td>210</td>
<td>2</td>
<td>83</td>
<td>q12Active22-24</td>
</tr>
<tr>
<td>201820301</td>
<td>LAN</td>
<td>360</td>
<td>3</td>
<td>461</td>
<td>q12Active22-24</td>
</tr>
<tr>
<td>201820301</td>
<td>LAN</td>
<td>370</td>
<td>2</td>
<td>510</td>
<td>q12Active22-24</td>
</tr>
<tr>
<td>201820301</td>
<td>LAN</td>
<td>370</td>
<td>2</td>
<td>471</td>
<td>q12Active22-24</td>
</tr>
<tr>
<td>201820301</td>
<td>LAN</td>
<td>380</td>
<td>2</td>
<td>565</td>
<td>q12Active22-24</td>
</tr>
<tr>
<td>201820301</td>
<td>LAN</td>
<td>380</td>
<td>2</td>
<td>579</td>
<td>q12Active22-24</td>
</tr>
</tbody>
</table>

4.6 Example code for bootstrap sampling of single-stage by SRSWOR

Suppose instead that at one or more stages in the last example, an SRSWOR is required. There are several different approaches to bootstrapping without replacement sampling designs (cf. Mashreghi et al., 2016). One possible approach is to construct a pseudo-population at the appropriate level based on the method described by Sitter (1992). Some example R code is provided Box 3 in Annex A3.3 below.
In the example in Box 3, a population of \(N=350\) values is created and a sample of \(n = 100\) is taken by SRSWOR. The code includes how to create the pseudo-population and do bootstrapping to obtain means and variances. It also shows that using the Horvitz-Thompson estimator of the total inside of the bootstrapping loop leads to a biased estimate whereas calculating an estimate for the population total after bootstrapping is unbiased. This set of code also shows that one could simply take an SRSWR and adjust the variance estimate with the finite population correction factor after bootstrapping to obtain the same results as those based on the pseudo-population.

### 4.7 Example code for bootstrapping two-stage sampling by SRSWOR at each stage

In the example in Box 4 in Annex A3.4, a population of \(N = 350\) PSUs is created and a sample of \(n = 30\) is taken by SRSWOR. For each PSU, there are between 10 and 20 SSUs of which a sample of 5 is taken for every PSU. For our example, suppose a PSU is a vessel \(\times\) trip and each PSU has between 10 and 20 hauls. On each trip, the observer samples 5 hauls for the weight of discards \((d_{L\ell}(t,v))\) and the weight of landings \((l_{L\ell}(t,v))\) of a single species of fish but has the total weight of the species for every haul, sampled or not \((W_{h(t,v)})\). Of interest is estimating the total discard weight of the species for the entire population of PSUs. The code shows how to create the pseudo-population and do bootstrapping to obtain the estimated total discard weight and its variance.

For this example, we used the following method for estimating discard weight:

1. For each sampled haul, estimate the total discard weight:
   
   \[
   \overline{dW}_{h(t,v)} = \frac{d_{W_{h(t,v)}}}{l_{W_{h(t,v)}}} W_{h(t,v)}
   \]
   
   The assumption here is that a haul is sampled by the observer before the discards and landings are separated during the processing of the haul by the crew. Hence, the proportion of the observer’s sample that is discards is considered equivalent to the proportion of the total landed weight of the haul that is discards.

2. For each sampled vessel \(\times\) trip the haul data are aggregated as follows. Total landings and discards for sampled hauls on vessel \(\times\) trip \(t(v)\) are defined as the sums of all sampled hauls

   \[
   LW_{t(v)} = \sum_{h(t,v)} LW_{h(t,v)}
   \]

   and

   \[
   DW_{t(v)} = \sum_{h(t,v)} \overline{dW}_{h(t,v)}.
   \]

   The unsampled hauls are excluded from the raising procedure. Note that the total landings weight for a haul, \(LW_{h(t,v)}\), is assumed to be known (e.g. provided by captain) or, if not, it can be calculated using \(\overline{W}_{h(t,v)} = W_{h(t,v)} - \overline{dW}_{h(t,v)}\) in place of \(LW_{h(t,v)}\) in the above equation.

3. To obtain estimated total discards for the entire population of PSUs, vessel \(\times\) trip-level estimates from step (2) are aggregated as follows. The estimated total discard weight, \(\overline{DW}\), is given by

\[
\overline{DW} = \frac{\sum_{t(v)\in S} DW_{t(v)}}{\sum_{t(v)\in S} LW_{t(v)}} \times LW
\]
where the sums in the ratio are over all observed trips in S and $LW = \sum_{t(v)\in P} LW_{t(v)}$ is the total landed weight reported for all vessel × trips in P, observed or not.

4.8 Bibliography


5 Bias

5.1 Method

The main ICES reports that considered the topic of bias in relation to catch sampling programs are:


The potential sources of bias that these reports identify were collated and an analysis was performed to see whether the RDBES could provide information that helped to evaluate that source of bias. The full analysis is presented in Annex A3.4 below. For specific cases where the RDBES can already provide insights into bias, some example reports have been developed and these are provided in Annex A3.5 below. In some cases where the RDBES is currently not capable of informing about the bias, some changes are suggested to the WGRDBESGOV10 core group that could enhance the capacity of the system in bias analyses.

5.2 Summary

In the table below, there were 41 issues identified from the ICES literature referenced above, classified into six categories. It was found that the RDBES can already provide comprehensive information about 12 of these issues, and partial information about a further 18 issues. This is a great improvement over systems such as the RDB or InterCatch, which can provide little or no information about these potential sources of bias.

The RDBES could not provide information about 11 issues. Typically, this required information that would not normally be stored in a commercial fisheries database, such as detailed protocols or training records.

Table 5.1. Table 1 Can the RDBES inform about the bias issues identified?

<table>
<thead>
<tr>
<th></th>
<th>1) Yes</th>
<th>2) Partially</th>
<th>3) No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Design</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>2) Protocol</td>
<td>4</td>
<td>4</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>3) Quality / precision</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>4) Analysis</td>
<td>1</td>
<td>8</td>
<td>2</td>
<td>11</td>
</tr>
</tbody>
</table>

10 https://www.ices.dk/community/groups/Pages/WGRDBESGOV.aspx
### Table

<table>
<thead>
<tr>
<th>Issue Category</th>
<th>1) Yes</th>
<th>2) Partially</th>
<th>3) No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>5) Analysis / estimation</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>6) Estimation</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>18</td>
<td>11</td>
<td>41</td>
</tr>
</tbody>
</table>

### Graph

The graph shows the number of responses for each issue category with options for 1) Yes, 2) Partially, and 3) No. The categories include:

- Design
- Protocol
- Quality / precision
- Analysis
- Analysis / estimation
- Estimation

The legend indicates how the responses are categorized.
6 Future work

6.1 Introduction

The RDBES aims to tackle long-standing needs of commercial catch sampling and estimation within ICES in terms of both precision and bias. It is a collaborative effort that involves experts and many countries and that will take several years to develop (the RDBES development roadmap is reviewed and updated annually\(^1\)). The tools to evaluate data precision and bias described within this report are necessarily limited in scope to ensure the work was feasible to complete with the time and resources available. They should not be considered as a complete solution but as a first step towards creating tools that can be applied more widely.

This section describes some ways in which this work could be extended. A roadmap is presented which shows how this could be approached. It will of course be necessary to gain feedback from users about the tools specified in this report and implement any identified improvements.

6.2 Analytical calculation of variance

The current work on implementing the analytical functions in R code was only done for totals so could be extended to include other estimates, such as numbers at length. It could also be extended to additional estimators (such as ratio estimators and age-length keys) within simple sampling designs (simple random sampling with or without replacement at each stage). Since the analytical calculation of variance is only feasible in relatively simple sampling programmes it is unlikely to be of much use when confronted with real-world data. For this reason it is not recommended to continue the development of these functions beyond the work done in this report.

6.3 Bootstrapping

6.3.1 Implementation in R code

Implement the bootstrapping pseudo-code as an R function which uses the correct RDBES field names. As far as possible the code should be modularized and split into functions to allow it to be more flexible.

6.3.2 Post-stratification and domain estimation

For this work, the bootstrapping approach is based on the estimation of parameters for the strata defined in the sampling design and reported in the data. However, often we want to produce estimates for domains that do not match the sampling strata. The bootstrapping procedure and the input datasets can be amended do this correctly. If the number of samples in strata are proportional to the size of the population (e.g. number of fishing trips) then the calculation of post-stratification variance based on the sampling strata will be reasonable, but in most cases, this will

---

not be the case and the post-stratification variance will actually be higher than that calculated based on the sampling strata.

### 6.3.3 Age-Length Key (ALK) construction

In the examples provided, the ALK is compiled across all samples to give an unweighted stratum-level ALK. Since the bootstrapping approach must honour the sampling design, the estimate of the accuracies associated with the numbers at age using this estimator are correctly captured. But the question of whether the current estimation procedure of simply collating all data collected within a stratum to construct the ALK provides the best estimator (more accurate as defined here) is unknown. The bootstrapping method could be used to investigate whether this is the most appropriate way to estimate the ALK or should a weighted approach be used instead.\(^{12}\)

### 6.3.4 Estimation based on other sampling designs

The procedures provided in this report are based on sampling designs that assume that the units in the sampling frame are selected based on simple random sampling, that is, every unit in the frame is equally likely to be the one selected. There are sampling efforts in MS that rely on unequal probability selection, usually based on a probability proportional to size, for example, assigning the likelihood of selecting a vessel to have an observer on board based on the vessel’s total landings in the prior year. Hence, more active or larger vessels are more likely to be sampled than the “smaller” vessels. The bootstrapping approach described here can be modified to allow for unequal probability sampling.

### 6.3.5 Extension to other types of estimation

This work only considered national design-based estimation but there are other types to be considered:

- National design-based estimation using the RDBES clustering variables
- Regional design-based estimation without/with the RDBES clustering variables
- National model-based estimation without/with the RDBES clustering variables
- Regional model-based estimation without/with the RDBES clustering variables

Of these, the regional design-based estimation should be the priority since it is likely not difficult to extend the work to accommodate a regional approach, and this extension is relevant to the move towards designing regional sampling programmes.

### 6.4 Bias

There are a number of actions described in the table listing the potential sources of bias in Annex A3.4 below. When these actions are considered, priority should be given to those actions that have the most benefit but that do not require a change in the RDBES data model. Some of the most important issues raised are inherently informed by the existing RDBES data model and will not require any further development work. For example, the way that issue numbers 1 (sampling

---

design), 4 (spatial and temporal coverage), 5 (sample allocation schemes), and 8 (PSU selection) are considered in the future will be improved by sampling data being uploaded to the RDBES.

6.5   Roadmap

A suggested roadmap for further development of the topics raised above is presented below. This is presented as a guide, assuming both funding and expert time will be available. Still, it will be necessary to further quantify the resources required and identify if those resources can be made available. Since the development of estimation methods for the RDBES is ongoing any further work should be considered within the overall RDBES roadmap and the work plan of related groups such as the EU Regional Coordination Groups (RCGs).

<table>
<thead>
<tr>
<th></th>
<th>2021 Q1</th>
<th>2021 Q2</th>
<th>2021 Q3</th>
<th>2021 Q4</th>
<th>2022 Q1</th>
<th>2022 Q2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implement bootstrapping as R function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extend bootstrapping to regional sampling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bootstrapping for post-stratification and domain estimation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Development of further example reports for potential bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WGRDBESGOV Core Group to consider the identified actions that are related to bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
7 Summary

This report shows how the new RDBES that is currently in development will be better able to support the recast EU Data Collection Framework (Regulation (EU) 2017/1004) than the existing RDB. The RDBES is an essential platform for MS and RCGs to fulfil their obligations towards documenting and improving data quality and designing and implementing regional sampling designs.

The evaluation of data precision was performed using two complementary techniques. For relatively simple sampling designs it is possible to use analytical functions to calculate the precision (or a related statistical measure such as variance) of a statistical estimate. These calculations and implementations of them in R code are presented in this report. For more complicated sampling designs, the use of analytical functions is usually not feasible. In these cases, it is necessary to evaluate precision using numerical techniques, the main one of which is bootstrapping. This report discussed when bootstrapping is appropriate and gives several worked examples describing how bootstrapping can be applied in different cases.

The evaluation of bias is a difficult subject and is hard to quantify. The approach followed in this report was to build on the previous work available in the ICES literature and identify and enumerate the main common sources of bias in catch sampling programs they describe. The information was collated and an evaluation performed as to whether data stored using the RDBES data format and reports issues from them can inform about the potential for bias in catch estimates. A set of example reports was coded that demonstrates the utility of the RDBES in relation to bias issues and can already help member states to identify how deviations in their sampling programmes and sampling variability may potentially lead to bias in their catch estimates.
## Annex A3.1  List of participants

<table>
<thead>
<tr>
<th>Name</th>
<th>Institute</th>
<th>Country</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary Christman</td>
<td>USA</td>
<td><a href="mailto:marycchristman@gmail.com">marycchristman@gmail.com</a></td>
<td></td>
</tr>
<tr>
<td>Liz Clarke</td>
<td>Marine Scotland Science</td>
<td>United Kingdom</td>
<td><a href="mailto:liz.clarke@gov.scot">liz.clarke@gov.scot</a></td>
</tr>
<tr>
<td>David Currie</td>
<td>Marine Institute</td>
<td>Ireland</td>
<td><a href="mailto:david.currie@marine.ie">david.currie@marine.ie</a></td>
</tr>
<tr>
<td>Annica Isaksson de Groote</td>
<td>SLU</td>
<td>Sweden</td>
<td><a href="mailto:annica.isaksson.de.groote@slu.se">annica.isaksson.de.groote@slu.se</a></td>
</tr>
<tr>
<td>Kirsten Birch Håkansson</td>
<td>DTU Aqua</td>
<td>Denmark</td>
<td><a href="mailto:kih@aqua.dtu.dk">kih@aqua.dtu.dk</a></td>
</tr>
<tr>
<td>Nuno Prista</td>
<td>SLU</td>
<td>Sweden</td>
<td><a href="mailto:nuno.prista@slu.se">nuno.prista@slu.se</a></td>
</tr>
<tr>
<td>Jose Rodriguez</td>
<td>IEO</td>
<td>Spain</td>
<td><a href="mailto:jose.rodriguez@ieo.es">jose.rodriguez@ieo.es</a></td>
</tr>
</tbody>
</table>
Annex A3.2  

R implementation of analytical variance calculation using design-based estimation

The following code creates a function that acts on an R object containing sampling data, and calculates a Horvitz-Thompson estimate of the population total of a univariate variable (for example landing weight) and the associated variance for that estimate, assuming simple random sampling without replacement (SRSWOR) was used at each stage. The code has been tested for a 3-stage sampling design and can be found on Github\(^{13}\).

```
estimateHTtotalMultiStageSRSWOR <- function(RDBobj, stages=stages, varOfInterest="SAtotalWtLive"){
  # this function calculates, and outputs, the Horvitz-Thompson estimate
  # of the population total of a single univariate variable
  # and the resulting variance of the estimate, assuming
  # SRSWOR (simple random sampling without replacement) is used at each stage
  # The function has 3 arguments:
  # RDBobj - the object containing the data, in RDBES format
  # stages - the sampling stages in the data
  # this avoids the need to specify hierarchies used in RDBES
  # varOfInterest - a character string specifying the name of the variable
  # for which we are estimating the population total (the "y variable")

  # set up some objects
  nStage <- length(stages)
  idPrev <- idList <- piList <- list()
  y <- nTotal <- nSamp <- meanStage <- ymean <- list()
  ssqTerm <- ssqStage <- tStage <- nTotStage <- nSampStage <- list()
  y[[nStage]] <- RDBobj[[stages[[nStage]]]][,varOfInterest]
  estVarTot <- vv <- list()

  # create lists of the key variables required in the calculations
  # nTotal - total number of units in each stage
  # nSamp - total number of samples in each stage
  # idList - the unique identifier for each unit in each stage
  # idPrev - a unique identifier for the units in the previous stage
  # at each stage (except stage 1)
  for (i in 1:nStage) {
    dat <- RDBobj[[stages[[i]]]]
    nTotal[[i]] <- dat[,paste(stages[[i]],"numTotal",sep="")]
    nSamp[[i]] <- dat[,paste(stages[[i]],"numSamp",sep="")]
    idList[[i]] <- dat[,paste(stages[[i]],"id",sep="")]
    if (i>1) {
      idPrev[[i]] <- dat[,paste(stages[[i-1]],"id",sep="")]
    } # end if
  } # end for

  # calculate terms in the variance at each stage
  # this needs to be done from the lowest hierarchy upwards
```

\(^{13}\) https://github.com/ices-eg/WK_RDBES/tree/master/Special_Request_20_05
# meanStage is the mean of of the y variable at each stage
# nTotStage & nSampStage are the numbers of units in each stage
# ssqStage is the sum of squares term in the variance estimate
# tStage the estimate of the "population" total of the y variable
# at each stage
for (i in nStage:1) {
  if (i==1) {
    # for the first stage (which is calculated last) we use sum not tapply
    meanStage[[i]] <- mean(y[[i]])
    ymean[[i]] <- rep(meanStage[[i]],length(idList[[i]]))
    nTotStage[[i]] <- rep(mean(nTotal[[i]]),length(idList[[i]]))
    nSampStage[[i]] <- rep(mean(nSamp[[i]]),length(idList[[i]]))
    ssqStage[[i]] <- sum(((y[[i]]-ymean[[i]])^2)/(nSamp[[i]]-1))
    tStage[[i]] <- sum(nTotal[[i]]/nSamp[[i]]*y[[i]])
  } else {
    # as there are several units of upper hierarchies in each stage,
    # we use tapply
    meanStage[[i]] <- tapply(y[[i]],idPrev[[i]],mean)
    ymean[[i]] <- meanStage[[i]][match(idPrev[[i]],names(meanStage[[i]]))]
    nTotStage[[i]] <- tapply(nTotal[[i]],idPrev[[i]],mean)
    nSampStage[[i]] <- tapply(nSamp[[i]],idPrev[[i]],mean)
    ssqStage[[i]] <- tapply(((y[[i]]-ymean[[i]])^2)/(nSamp[[i]]-1),idPrev[[i]],sum)
    tStage[[i]] <- tapply(nTotal[[i]]/nSamp[[i]]*y[[i]],idPrev[[i]],sum)
  }
  # add in cases where variance is zero because the whole population
  # at that stage was sampled
  ssqStage[[i]][is.infinite(ssqStage[[i]]) & nTotStage[[i]]==nSampStage[[i]]] <- 0
  ssqTerm[[i]] <- nTotStage[[i]]^2*(1-nSampStage[[i]]/
    nTotStage[[i]])/nSampStage[[i]]*ssqStage[[i]]
  if (i>1) y[[i-1]] <- tStage[[i]][match(idList[[i-1]],names(tStage[[i]]))]
} #end for
estTot <- tStage

# now calculate the sum of each term sequentially from the lowest hierarchy
# to the first heirarchy
vv <- estVarTot <- list()
vv[[nStage]] <- 0
for (i in nStage:2) {
  idPrevVec <- idList[[i-1]][match(names(ssqTerm[[i]]),idList[[i-1]])]
  estVarTot[[i]] <- tapply(ssqTerm[[i]]+vv[[i]],idPrevVec,sum)
  vv[[i-1]] <- nTotStage[[i]]/nSampStage[[i]]*estVarTot[[i]]
}

# output the point estimate of the population total, and
# the associated variance estimate for the point estimate
output <- list(estTot=estTot,estVarTot=estVarTot)
return(output)
}
Annex A3.3  Bootstrapping pseudo-code

Box 2. R code for bootstrapping two stage sampling of trips and then fish within trips. It includes stratified SRSWR of fish for age where the strata are length classes observed in a trip.

<table>
<thead>
<tr>
<th>Box 2. R code for SRSWR of several stages</th>
</tr>
</thead>
<tbody>
<tr>
<td># PREPARATORY STEPS</td>
</tr>
<tr>
<td># Number of bootstraps desired for simulating results</td>
</tr>
<tr>
<td>TB &lt;- 2500</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td># List of strata names</td>
</tr>
<tr>
<td>stratalist &lt;- unique(fish$NewStratum)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td># matrices to store stratum-level NAL results (means and variances) from the bootstraps</td>
</tr>
<tr>
<td>stratbootNAL &lt;- matrix(0, nrow=length(unique(fish$Length)), ncol=length(stratalist))</td>
</tr>
<tr>
<td>rownames(stratbootNAL) &lt;- sort(unique(fish$Length))</td>
</tr>
<tr>
<td>colnames(stratbootNAL) &lt;- stratalist</td>
</tr>
<tr>
<td>stratbootrtvarNAL &lt;- matrix(0, nrow=length(unique(fish$Length)), ncol=length(stratalist))</td>
</tr>
<tr>
<td>rownames(stratbootrtvarNAL) &lt;- sort(unique(fish$Length))</td>
</tr>
<tr>
<td>colnames(stratbootrtvarNAL) &lt;- stratalist</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td># matrices to store stratum-level NAA results (means and variances) from the bootstraps</td>
</tr>
<tr>
<td>stratbootNAA &lt;- matrix(0, nrow=length(unique(indfish$Age)), ncol=length(stratalist))</td>
</tr>
<tr>
<td>rownames(stratbootNAA) &lt;- sort(unique(indfish$Age))</td>
</tr>
<tr>
<td>colnames(stratbootNAA) &lt;- stratalist</td>
</tr>
<tr>
<td>stratbootrtvarNAA &lt;- matrix(0, nrow=length(unique(indfish$Age)), ncol=length(stratalist))</td>
</tr>
<tr>
<td>rownames(stratbootrtvarNAA) &lt;- sort(unique(indfish$Age))</td>
</tr>
<tr>
<td>colnames(stratbootrtvarNAA) &lt;- stratalist</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td># need column indicator for stratum ID</td>
</tr>
<tr>
<td>colid &lt;- 0</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>#### OUTSIDE LOOP FOR STRATUM ANALYSES</td>
</tr>
<tr>
<td>for (strata in stratalist)</td>
</tr>
<tr>
<td>{</td>
</tr>
<tr>
<td># matrix to store NALs from each bootstrap sample within a stratum</td>
</tr>
<tr>
<td>}</td>
</tr>
</tbody>
</table>
Box 2. R code for SRSWR of several stages

```r
bootNAL <- matrix(0, nrow=length(unique(fish$Length)), ncol=TB)
rownames(bootNAL) <- sort(unique(fish$Length))
bootNAA <- matrix(0, nrow=length(unique(indfish$Age)), ncol=TB)
rownames(bootNAA) <- sort(unique(indfish$Age))

# create a temporary dataset of lengths for the stratum (bootstrap samples taken from this)
tempdata <- fish[fish$NewStratum == strata, ]

# need a list of trips in tempdata
trips <- unique(tempdata$UniqTrip)
# need total number of trips in tempdata
lentrips <- length(trips)

##### BOOTSTRAPPING LOOP (within a stratum)
for (boots in 1:TB)
{

# for each bootstrap need 2 datasets to store the data from the bootstrapped trips
# one for the length sampling and one for the subsampling of ages by length

# matrix to store bootstrap samples of trips for length measurements
# first row will be removed later – it contains all 0s
bootfish <- matrix(0, nrow=1, ncol=dim(fish)[2])
colnames(bootfish) <- colnames(fish)

# matrix to store bootstrap subsamples of age data from bootstrapped trips
# first row will be removed later – it contains all 0s
bootind <- matrix(0, nrow=1, ncol=dim(tempind1)[2])
colnames(bootind) <- c("UniqTrip", "CatchCat", "Length", "Age", "Weight", "NewStratum")

# take a SRSWR of trips within the stratum using SRSWR
boottrip <- sample(trips, lentrips, replace = T)

####### TRIPS LOOP
for (tripid in boottrip)
```

}
Box 2. R code for SRSWR of several stages

```
{
  # fish lengths from a sampled tripid
  bootdata <- tempdata[tempdata$UniqTrip==tripid,]
  rownames(bootdata) <- bootdata$Length

  # total number of fish originally sampled on the trip
  numlens <- sum(bootdata$SumNumAtLen)

  # relative frequencies of each length in dataset
  Sprobs <- bootdata$SumNumAtLen/sum(bootdata$SumNumAtLen)

  # TAKE A BOOTSTRAP SRSWR OF FISH WITHIN THE TRIP FOR LENGTH MEASUREMENTS
  fishsamp <- sample(bootdata$Length, numlens, replace=T, prob=Sprobs)

  # frequencies of fish lengths in the bootstrap sample of fish
  numtimes <- table(fishsamp)

  # create new fish dataset (contains only 1 record for each length) for the trip
  tempfish <- bootdata[which(!is.na(match(rownames(bootdata),fishsamp))],]
  # order it by lengths
  tempfish <- tempfish[order(tempfish$Length,decreasing=FALSE),]

  # update the frequencies of fish sampled in bootstrapped trip by length
  tempfish$SumNumAtLen <- numtimes

  # calculate new estimated numbers per trip (using the ratio of sample to total weight
  # within the catch category
  tempfish$TrpNumAtLen<-ceiling(tempfish$SumNumAtLen*tempfish$RatioStaWtSampWt)

  # save bootstrapped trip of fish length data to temporary file (used for bootstrapped NAL)
  bootfish <- rbind(bootfish, tempfish)

  # obtain the data from indfish for the same tripid
}
```
### Box 2. R code for SRSWR of several stages

```r
tempind <- indfish[indfish$UniqTrip==tripid,]

# list of lengths in the bootstrap sample of fish in trip = tripid
indlens <- unique(tempfish$Length)

# SRSWR STRATIFIED BY LENGTH CLASS OF FISH AGES USING LENGTHS FROM TEMPFISH
for (j in indlens)
{
  # number of times a length appears in fish dataset
  size1 <- tempfish$SumNumAtLen[tempfish$Length == j]
  # take a SRSWR of size1 or 10, which ever is smaller from indfish
  samp1 <- sample(min(10, size1), replace=T)
  subsample1 <- tempind[tempind$Length == j,]
  # store results for that length class
  bootind <- rbind(bootind, subsample1[samp1,])
}
# close j loop

# close tripid loop

# ANALYZE the “boots” BOOTSTRAP SAMPLE AND STORE RESULTS
# the two datasets bootfish and bootind contain “new” trips that have the same numbers of
# fish for length measurements but may contain different numbers of fish for the age
# measurements

# remove the first row of the bootfish and bootind datasets (all zeroes)
totfishrecs <- dim(bootfish)[1]
if (totfishrecs>1) bootfish <- bootfish[2:totfishrecs,]
totindrecs <- dim(bootind)[1]
if (totindrecs>1) bootind <- bootind[2:totindrecs,]

# store NAL for the “boots” bootstrap
numlen <- tapply(bootfish$TrpNumAtLen, bootfish$Length, sum)
bootNAL[which(!is.na(match(rownames(bootNAL),names(numlen)))), boots] <- numlen

# calculate NAA for the “boots” bootstrap
```
Box 2. R code for SRSWR of several stages

```r
# construct ALK for the "boots" bootstrap
bootALK <- table(bootind$Length, bootind$Age, useNA="no")
bootALK <- bootALK/apply(bootALK, 1, sum)

# multiply ALK by NAL for the "boots" bootstrap
numagelen <- sweep(as.matrix(bootALK[which(!is.na(match(rownames(as.matrix(bootALK)),rownames(as.matrix(bootNAL))))),]), MARGIN=1, as.matrix(bootNAL[which(!is.na(match(rownames(as.matrix(bootNAL[,1])),rownames(as.matrix(bootALK))))),boots]), "*")

# collapse to get NAA and store
bootNAA[which(!is.na(match(rownames(bootNAA),rownames(t(numagelen))))),boots] <- apply(numagelen, 2, sum)
}

# close boots loop

# store the means and variances of the bootstrapped NAA and NAL for stratum = "strata"
colid <- colid + 1
stratboottrvarNAL[,colid] <- sqrt(apply(bootNAL, 1, var))
stratbootNAL[,colid] <- apply(bootNAL, 1, mean)

templist <- apply(bootNAA,2, sum)

tempbootNAA <- bootNAA[,templist>0]
stratboottrvarNAA[,colid] <- sqrt(apply(tempbootNAA, 1, var))
stratbootNAA[,colid] <- apply(tempbootNAA, 1, mean)
}
# close strata loop
```

Box 3. R code for Sitter (1992)

```r
#### Sitter 1992 approach

# Pseudo-Population when N/n is not an integer

# Random size of Pseudo-Pop is not a problem for

# variance estimation, i.e. can be ignored, but could affect estimation of

# population total if pseudo-pop size is used

# create a population of N = 350 values
N <- 350
pop <- rnorm(N,500,75)
POPmean <- mean(pop)
POPtotal <- sum(pop)

# take a SRSWOR from the population
n <- 100
SRSWOR <- sample(pop, n, replace=F)

# estimate of the population mean and its variance
SRSWORmean <- mean(SRSWOR)
SRSWORvar <- (1-(n/N))*var(SRSWOR)/n

# estimate of the population total and its variance
SRSWORht <- N*SRSWORmean
SRSWORhtvar = N^2 * SRSWORse^2

# calculations needed for creating the pseudo-population
f <- n/N
k <- (N/n)*(1-((1-f)/n))
k1 <- floor(k)
k2 <- ceiling(k)
```


Box 3. R code for Sitter (1992)

```r
n1prime <- n-1
n2prime <- n
a1 <- (k1*(1-(n1prime/(n*k1))))/(n1prime*(n*k1-1))
a2 <- (k2*(1-(n2prime/(n*k2))))/(n2prime*(n*k2-1))
qs <- (((1-f)/(n*(n-1)))-a2)/(a1-a2)
PPB1 <- rep(SRSWOR, k1)
PPB2 <- rep(SRSWOR, k2)
len1 <- length(PPB1)
len2 <- length(PPB2)

B <- 50000
# place to store bootstrapping results
BootRes1 <- matrix(0, nrow=B, ncol=6)
colnames(BootRes1) <- c("SRSWORmean", "SRSWORvar", "HTBtotal", "HTBbias",
                  "SRSWRmean", "SRSWRvar*(1-f)")

# bootstrapping loop
for (b in 1:B)
{
  # take a sample from the appropriate pseudo-population
  ranumber <- runif(1)
  if (ranumber < qs)
  {
    sampWOR <- sample(PPB1, n1prime, replace=F)
    poplen <- len1
  }
  if (ranumber >= qs)
  {
    sampWOR <- sample(PPB2, n2prime, replace=F)
    poplen <- len2
  }
  BootRes1[b,1] <- mean(sampWOR)
}
```
Box 3. R code for Sitter (1992)

```r
# Original Population
POPmean
POPtotal

# Original Sample
SRSWORmean
SRSWORvar
SRSWORht

# Means of bootstrap quantities
apply(BootRes1, 2, mean)

# variances of bootstrap quantities
apply(BootRes1, 2, var)

# SOME RESULTS
> # Population
> POPmean = 500.6
> POPtotal = 175224.7

> # Original Sample
> SRSWORmean = 499.18
> SRSWORvar = 35.28
> SRSWORht = 174712.50
```

```r
BootRes1[b,2] <- (1-(n/N))*var(sampWOR)/n
BootRes1[b,3] <- poplen*BootRes1[b,1]

# compare to SRSWR
tempsamp <- sample(SRSWOR, n, replace=T)
BootRes1[b,5] <- mean(tempsamp)
BootRes1[b,6] <- var(tempsamp)*(1-f)/n
```

```r
#Original Population
POPmean
POPtotal

#Original Sample
SRSWORmean
SRSWORvar
SRSWORht

# Means of bootstrap quantities
apply(BootRes1, 2, mean)

# variances of bootstrap quantities
apply(BootRes1, 2, var)

#SOME RESULTS
> # Population
> POPmean = 500.6
> POPtotal = 175224.7

> # Original Sample
> SRSWORmean = 499.18
> SRSWORvar = 35.28
> SRSWORht = 174712.50
```
Box 3. R code for Sitter (1992)

```
> # Means of bootstrap quantities (should reproduce the values calculated from the original sample)
> apply(BootRes1, 2, mean)

         SRSWORmean     SRSWORvar   HTBtotal   HTBbias   SRSWRmean
SRSWRvar*(1-f)  499.169   35.278      178925.23  4216.19      499.228
         34.894
```

Box 4. R code for two stage with SRSWOR at both stages. Estimating total discards in the population.

```r
# Population of N = 350 PSUs
Npsu <- 350

# Each PSU has between 10 and 20 SSUs
SSUs <- sample(10:20, Npsu, replace=T)

# Total Number of SSUs for the Npsu PSUs
totnumSSU <- sum(SSUs)

# Create data for each SSU
# True Total Haul Weight
SSUtotwts <- sample(200:700, totnumSSU, replace=T)

# True proportion of total haul weight that is discards
SSUdisprops <- runif(totnumSSU, min=0.05, max=0.20)

# Create the population
pop <- data.frame(matrix(0, nrow=SSUs[1], ncol = 11))
colnames(pop) <- c("PSUid", "SSUid", "totSSU", "SSUtotwt", "SSUdisprop",
               "trueSSUdiswt", "trueSSUlanwt", "SSUsampwt", "sampSSUdiswt",
               "sampSSUlanwt",
               "ratioSampDisLan")

# Fill in matrix for the first PSU
pop[,1] <- 1
pop[,2] <- c(1:SSUs[1])
pop[,3] <- SSUs[1]
pop[,4] <- SSUtotwts[1:SSUs[1]]
pop[,5] <- SSUdisprops[1:SSUs[1]]

# Do the same for the remaining PSUs
for (i in 2:Npsu)
{
  tpop <- data.frame(matrix(0, nrow=SSUs[i], ncol = 11))
```

Box 4. R code for two stage with SRSWOR at both stages. Estimating total discards in the population.

```r
# Define the dataset
pop <- data.frame(PSUid = 1:totnumSSU, SSUid = 1:totnumSSU, SSUsampwt = runif(totnumSSU, min=0.1, max=0.15)*pop[,4], sampSSUdiswt = runif(totnumSSU, min=0.05, max=0.20), sampSSulanwt = pop[,8]-pop[,9])

# Calculate weights for discards and landings for each SSU
pop[,6] <- pop[,4]*pop[,5]
pop[,7] <- pop[,4]-pop[,6]

# Create sample weights of discards and landings
pop[,10] <- pop[,8]-pop[,9]
pop[,11] <- pop[,9]/pop[,10]

# Parameter Values
TrueTotDisWt <- sum(pop[,6])
TrueTotLanWt <- sum(pop[,7])
TrueTotWt <- sum(pop[,4])

# Population Estimate of Total Discards (can never be known)
# Might be useful for estimating bias due to using a ratio estimator
EstTripDis <- matrix(0, nrow=Npsu, ncol=2)
colnames(EstTripDis) <- c("EstDisWt", "SumLanWt")
for (i in 1:Npsu) {
  # R code for two stage with SRSWOR at both stages
  # Estimating total discards in the population.
```

Box 4. R code for two stage with SRSWOR at both stages. Estimating total discards in the population.

```r
PSUdata <- pop[pop$PSUid==i,]
EstTripDis[i,1] <- sum(PSUdata$ratioSampDisLan*PSUdata$SSUtotwt)
EstTripDis[i,2] <- sum(PSUdata$trueSSUlanwt)
}
EstPopTotDisWt <- TrueTotLanWt*sum(EstTripDis[,1])/sum(EstTripDis[,2])

# take a SRSWOR of PSUs from the population
npsu <- 30
SRSWORpsu <- sample(1:Npsu, npsu, replace=F)
PSUsample <- pop[which(!is.na(match(pop$PSUid,SRSWORpsu))),]

# for each sampled PSU, take a SRSWOR of nssu SSUs
nssu <- 5
temppop <- pop[pop$PSUid == SRSWORpsu[1],]
SRSWORssu <- sample(1:temppop$totSSU[1], nssu, replace=F)
PSUSSUsamp <- temppop[which(!is.na(match(temppop$SSUid,SRSWORssu))),]

for (i in SRSWORpsu[2:npsu])
{
  temppsu <- pop[pop$PSUid == i,]
  SRSWORssu <- sample(1:temppsu$totSSU[1], nssu, replace=F)
tempsu <- temppsu[which(!is.na(match(temppsu$SSUid,SRSWORssu))),]
  PSUSSUsamp <- rbind(PSUSSUsamp, tempsu)
}

# Original Sample Estimate of Total Discards
EstTripDis <- matrix(0, nrow=Npsu, ncol=2)
colnames(EstTripDis) <- c("EstDisWt", "SumLanWt")
rowid <- 0
for (i in SRSWORpsu)
{
  EstTripDis[rowid+1,1] <- sum(PSUdata$ratioSampDisLan*PSUdata$SSUtotwt)
  EstTripDis[rowid+1,2] <- sum(PSUdata$trueSSUlanwt)
  rowid <- rowid+1
}
```

# Original Sample Estimate of Total Discards
EstTripDis <- matrix(0, nrow=Npsu, ncol=2)
colnames(EstTripDis) <- c("EstDisWt", "SumLanWt")
rowid <- 0
for (i in SRSWORpsu)
Box 4. R code for two stage with SRSWOR at both stages. Estimating total discards in the population.

```r
{ 
  PSUSSUdata <- PSUSSUsamp[PSUSSUsamp$PSUid==i,]
  rowid <- rowid + 1
  EstTripDis[rowid,1] <- sum(PSUSSUdata$ratioSampDisLan*PSUSSUdata$SSUtotwt)
  # Using true landings weight for the hauls not an estimated weight
  EstTripDis[rowid,2] <- sum(PSUSSUdata$trueSSUlanwt)
} # Using the reported true total landings weight for the entire population
EstSampTotDisWt <- TrueTotLanWt*sum(EstTripDis[,1])/sum(EstTripDis[,2])

# Calculations needed for creating the pseudo-population of PSUs used in bootstrapping
fp <- npsu/Npsu
kp <- (Npsu/npsu)^((1-(1-fp))/npsu))
k1p <- floor(kp)
k2p <- ceiling(kp)
n1primep <- npsu-1
n2primep <- npsu
a1p <- (k1p*(1-(n1primep/(npsu*k1p))))/(n1primep*(npsu*k1p-1))
a2p <- (k2p*(1-(n2primep/(npsu*k2p))))/(n2primep*(npsu*k2p-1))
qsp <- (((1-fp)/(npsu*(npsu-1)))-a2p)/(a1p-a2p)
PPB1p <- rep(SRSWORpsu, k1p)
PPB2p <- rep(SRSWORpsu, k2p)
len1p <- npsu*k1p
len2p <- npsu*k2p

# Number of bootstrap samples to use
B <- 50000
# place to store bootstrapping results
BootRes1 <- matrix(0, nrow=B, ncol=1)
colnames(BootRes1) <- c("EstTotDis")
}```
Box 4. R code for two stage with SRSWOR at both stages. Estimating total discards in the population.

```r
# bootstrapping loop
for (b in 1:B)
{
  # sample PSUs from the pseudo-population
  rannumber <- runif(1)
  if (rannumber < qsp)
  {
    sampWORp <- sample(PPB1p, n1primep, replace=F)
    psulen <- len1p
  }
  if (rannumber >= qsp)
  {
    sampWORp <- sample(PPB2p, n2primep, replace=F)
    psulen <- len2p
  }
  # Calculations needed to estimate total discards in population
  EstTripDis <- matrix(0, nrow=length(sampWORp), ncol=2)
  colnames(EstTripDis) <- c("EstDisWt", "SumLanWt")
  rowid <- 0
  for (i in sampWORp)
  {
    Nssu <- unique(pop$totSSU[pop$PSUid==i])
    fs <- nssu/Nssu
    ks <- (Nssu/nssu)*(1-((1-fs)/nssu))
    k1s <- floor(ks)
    k2s <- ceiling(ks)
    n1primes <- nssu-1
    n2primes <- nssu
    # sample SSUs from sampled PSUs
    Nssu <- unique(pop$totSSU[pop$PSUid==i])
    fs <- nssu/Nssu
    ks <- (Nssu/nssu)*(1-((1-fs)/nssu))
    k1s <- floor(ks)
    k2s <- ceiling(ks)
    n1primes <- nssu-1
    n2primes <- nssu
  }
}
```
Box 4. R code for two stage with SRSWOR at both stages. Estimating total discards in the population.

```r
# EstTripDisc now contains the data for the trips in this bootstrap sample
da1s <- (k1s*(1-(n1primes/(nssu*k1s))))/(n1primes*(nssu*k1s-1))
da2s <- (k2s*(1-(n2primes/(nssu*k2s))))/(n2primes*(nssu*k2s-1))
qss <- (((1-fs)/(nssu*(nssu-1)))-a2s)/(a1s-a2s)
SRSWORssu <- unique(PSUSSUsamp$SSUid[PSUSSUsamp$PSUid==i])
PPB1s <- rep(SRSWORssu, k1s)
PPB2s <- rep(SRSWORssu, k2s)
len1s <- length(PPB1s)
len2s <- length(PPB2s)
rannumbers <- runif(1)
if (rannumbers < qss)
{
  sampWORs <- sample(PPB1s, n1primes, replace=F)
poplens <- len1s
}
if (rannumbers >= qss)
{
  sampWORs <- sample(PPB2s, n2primes, replace=F)
poplen <- len2s
}
rowid <- rowid + 1
if (i == sampWORp[1])
{
  tempboot <- PSUSSUsamp[PSUSSUsamp$PSUid==i,]
tempboot1 <- tempboot[match(sampWORs, tempboot$SSUid),]
}
if (i != sampWORp[1])
{
  tempboot <- PSUSSUsamp[PSUSSUsamp$PSUid==i,]
tempboot1 <- tempboot[match(sampWORs, tempboot$SSUid),]
}
```
Box 4. R code for two stage with SRSWOR at both stages. Estimating total discards in the population.

```r
EstTripDis[rowid,1] <- sum(tempboot1$ratioSampDisLan*tempboot1$SSUtotwt)
EstTripDis[rowid,2] <- sum(tempboot1$trueSSUlanwt)
}
# store estimate of total discards from the bootstrap sample
BootRes1[b,1] <- TrueTotLanWt*sum(EstTripDis[,1])/sum(EstTripDis[,2])
}

#Original Population Discard Weight
TrueTotDisWt
# Estimated Discard Weight Based on All Population Data
EstPopTotDisWt
# Estimated Discard Weight Based on Sample Data
EstSampTotDisWt
# Mean of the Bootstrap Estimates based on SRSWOR
mean(BootRes1[,1])
# Range of values of the Bootstrap Estimates
quantile(BootRes1[,1])
# Standard Error of the Bootstrap Estimates
sqrt(var(BootRes1[,1]))
# Relative Bias of mean of Bootstrap estimates compared to the sample used to create the pseudo-population
100*EstSampTotDisWt/mean(BootRes1)
# Confidence Interval Endpoints (90 and 95% CIs)
quantile(BootRes1[,1], probs=c(0.025, 0.05, 0.95, 0.975))

#### SOME RESULTS
```

> #Original Population Discard Weight
> TrueTotDisWt
[1] 312805.9
> # Estimated Discard Weight Based on All Population Data (had a sample been taken from every PSU)
Box 4. R code for two stage with SRSWOR at both stages. Estimating total discards in the population.

```r
> EstPopTotDisWt
[1] 351515.4
> # Estimated Discard Weight Based on Sample Data of 30 PSUs and 5 SSUs/PSU
> EstSampTotDisWt
[1] 353891.9
> # Mean of the Bootstrap Estimates based on SRSWOR
> mean(BootRes1[,1])
[1] 355803.7
> # Range of values of the Bootstrap Estimates
> quantile(BootRes1[,1])
0% 25% 50% 75% 100%
284957.9 344167.4 355730.1 367227.4 423707.4
> # Standard Error of the Bootstrap Estimates
> sqrt(var(BootRes1[,1]))
[1] 17013.32
> # Relative Bias of mean of Bootstrap estimates compared to the sample used to create the pseudo-population
> 100*EstSampTotDisWt/mean(BootRes1)
[1] 99.46268
> # Confidence Interval Endpoints (90 and 95% CIs)
> quantile(BootRes1[,1], probs=c(0.025, 0.05, 0.95, 0.975))
2.5% 5% 95% 97.5%
322741.0 328053.6 384104.4 389356.9
```
Annex A3.4  Potential sources of bias

The table below summarizes a number of potential sources of bias identified in the ICES litera-
tures and describes if and how the RDBES can provide information about them. The columns
have the following meanings:

- **Issue Category**
  - The broad category the issue falls within
- **Id**
  - A numerical id only used within this report
- **Issue description**
  - Text describing the potential source of bias and its effects. In the majority of cases
descriptions were copied directly from the source.
- **Source**
  - The report where the issue was identified. The reports considered were:
- **Affects**
  - Which variables does this source of bias affect (e.g. age)? Some issues will affect all
variables
- **Can the RDBES tackle it at the moment?**
  - Yes/No/Partially. Includes a description of how the RDBES can provide information
about the issue
- **Could the RDBES potentially tackle this in the future?**
  - Could the RDBES potentially give information about this issue in the future? This
might require a change in the RDBES data format.
- **Comment**
  - General comments related to the issue
- **Actions**
  - Actions required by specific groups to make progress on the issue
<table>
<thead>
<tr>
<th>Issue Category</th>
<th>Id</th>
<th>Issue description</th>
<th>Source</th>
<th>Affects</th>
<th>Can the RDBES tackle it at the moment?</th>
<th>Could the RDBES potentially tackle this in the future?</th>
<th>Comment</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design</td>
<td>1</td>
<td>Sampling design. Minimization of bias through sampling design, or at least an ability to identify and quantify biases, is more critical than minimization of variance (SGPIDS).</td>
<td>SGPIDS, WKPI CS, 2011</td>
<td>All variables</td>
<td>Yes</td>
<td>In the RDBES sampling design is documented alongside the data collected. It is then possible to evaluate whether the design might be a source of bias (e.g. non-probabilistic sampling, systematic non-responses)</td>
<td>-</td>
<td>Groups such as WGCATCH promote good practice in sampling design. Groups like WGRDBESGOV promote good practices in populating the RDBES format. Once fully populated the RDBES will be a valuable tool to analyse practical implementation of sampling designs.</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Coverage, design. If only part of the population is covered, the frame has under-coverage that will lead to bias unless the variables of interest (e.g. discard rates; species or size compositions) are the same in the parts of the population covered or not covered, or if only a very small part of the population is not covered.</td>
<td>WKPI CS, 2011</td>
<td>All variables</td>
<td>Partially</td>
<td>We can see what we have. The RDBES stores population data and sampling data – these can then be compared to determine coverage. The RDBES does not store sampling frame data.</td>
<td>-</td>
<td>The overall sampling frames are described in the national work plans, and this is needed to evaluate the overall design. The RDBES will be able to support evaluation of national work-plans in the future.</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Coverage, country. An example of under-coverage would be the non-sampling of vessels of a national fleet that land in another country. This fraction may vary from year to year leading to a variable bias if activities, gears etc. differ from vessels landing in the home country.</td>
<td>WKPI CS, 2011</td>
<td>All variables</td>
<td>Partially</td>
<td>Using data in the RDBES it will be possible to quantify the magnitude of non-national vessel activity, but it is not possible to evaluate if it is included in a sampling frame.</td>
<td>-</td>
<td>National workplans provide detailed information about the sampling frames used. The recommended way of handling this situation within the RDBES needs reviewing. Further analysis and discussion is possible at the Regional Coordination Groups and in the future when designing regional sampling plans. This issue will also be discussed and prioritized by the WGRDBESGOV “Core group”.</td>
</tr>
<tr>
<td>Issue Category</td>
<td>Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>----------------</td>
<td>----</td>
<td>--------------------</td>
<td>--------</td>
<td>---------</td>
<td>----------------------------------------</td>
<td>--------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>Spatial and Temporal coverage: it has been discussed during the workshop that any discrepancy between the sampling and fishing effort coverage do not lead to a bias when the sampling is done randomly following a well-designed protocol. In other cases, the temporal coverage in terms of mean discrepancy between proportion by units of time plus existence of non-sampled strata must be evaluated.</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>Yes</td>
<td>All RDBES CS tables representing sampling stages includes the variable “XXselection-Method” that allows the distinction between probabilistic and non-probabilistic sampling and methods therein. It is also possible, within the RDBES to compare the temporal distribution of CS data with that of CL and CE data and detect systematic departures from the sampling rates expected for specific time domains.</td>
<td>-</td>
<td>WKACCU seems to confound bias and sampling error, both of which affect estimates and may lead to departures of estimates from the true totals they try to quantify. We have focused on the bias which is frequently considered more severe from a data quality perspective.</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>Appropriate time period and spatial coverage: Biological variables change through time and space and in some time periods and areas their determination may be less accurate than in other. There are recommendations from ices groups (e.g., ICES WKMAT) that orient countries during data collection and reduce biases in analyses.</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>Yes</td>
<td>Information on timing and space of samples is available in RDBES data model that allows investigation into this sources of bias.</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Issue Category</td>
<td>Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>----------------</td>
<td>----</td>
<td>------------------</td>
<td>--------</td>
<td>---------</td>
<td>----------------------------------------</td>
<td>------------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>Sampling allocation scheme: estimation of the randomness of the sampling. Is sampling pure random with a sampling protocol well followed, or is sampling allocation made on ad hoc or opportunistic observations? A non random sampling is clearly a source of bias which needs to be reported. In the case of length sampling: Random sampling of boxes/trips: This bias, linked to the follow-up of a sampling protocol focuses more on the randomness of both the choice of boxes to sample (always the top box, vs. real random,) and the choice of trips (opportunistic, real random).</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>Yes</td>
<td>All RDBES CS tables representing sampling stages include variables of type “selection-Method” that allow the distinction between probabilistic and non-probabilistic sampling and methods therein.</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>Non sampled strata: Usually, imputation rules exist for non sampled strata, thus this bias will be an evaluation of the appropriateness of the imputation rules used. E.g., Population of vessels: are all vessels included in the population that forms the sampling frame?</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>No</td>
<td>Partially</td>
<td>A suggestion for declaration of out-of-frame non sampled fractions has been put forward by WKRDB-EST2 to the WGRDBESGOV “Core Group”. Inclusion of such a feature will allow the identification of (known) parts of the population that are not included in the sampling frame.</td>
<td>WGRDBESGOV “Core Group to discuss the recommendation.</td>
</tr>
<tr>
<td>Issue Category</td>
<td>Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>----------------</td>
<td>----</td>
<td>-------------------</td>
<td>--------</td>
<td>---------</td>
<td>-------------------------------------</td>
<td>--------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>In general, the PSU is the first level in hierarchy of sampling units, each representing a cluster of fishing trips, hauls within trips, boxes of fish within hauls etc. For the overall raising procedure to be unbiased, the selection of samples at each stage should be random, and the raising factors are derived from the sampling fraction at that stage.</td>
<td>WKPICS, 2011</td>
<td>All variables</td>
<td>Yes</td>
<td>In contrast with the previous RDB system, the RDBES allows clear identification of sampling levels used by countries in their multi-stage sampling programmes. Furthermore data-submitters can specify for each level of sampling the selection method they used when selecting samples (see “Sample selection methods” in the RDBES documentation)</td>
<td>-</td>
<td>The RDBES makes substantial progress towards identification and documentation of this bias</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>Source of information: it is unlikely that one source of information encompasses the statistics of all fisheries, including the temporal, spatial and fishing activity stratification. In all cases, the advantages and limitations of the sources used should provide a clear view on the related bias.</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>Partially</td>
<td>RDBES documents the source of information of CL and CE data but not the source of information used in defining the sampling frames of CS data.</td>
<td>No</td>
<td>Variables that document the sources used to derive the sampling frames of CS data could be added to RDBES but are unlikely add much value to the data.</td>
</tr>
<tr>
<td>Issue Category</td>
<td>Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>----------------</td>
<td>----</td>
<td>--------------------</td>
<td>--------</td>
<td>---------</td>
<td>--------------------------------------</td>
<td>------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>Quality assurance protocol: Existence and follow-up of a sampling protocol.</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>No</td>
<td>No</td>
<td>The RDBES does not contain detailed information on all protocols used in everything that is sampled (e.g., age reading, some specifics of onboard or onshore sampling, etc). Supplementary information will be needed to quality assure those aspects. To facilitate finding this information the name of the sampling scheme used in the RDBES should be the same as the name used in other sources (e.g. national workplans).</td>
<td>Information on quality assurance protocols should be included in national workplans</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>Sampling protocol: Existence and adherence to a sampling protocol that yields representative selection of fish for length measurements.</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>No</td>
<td>No</td>
<td>See comments in issue 10</td>
<td>Information on sampling protocols should be included in national workplans</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>Non response rate: the percentage of refusal is one of the most important sources of bias for on-board observers. This case discussed in general in Cochran, 1977 has also been addressed by the recent workshop on discards (Anon, 2003) in the frame of the DCR.</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>Yes</td>
<td>-</td>
<td>Non-responses are a source of potential bias in the calculation of all parameters</td>
<td>An example report was developed that demonstrates how plots build from RDBES data can highlight the presence of non-responses in the data. See Annex A3.5.</td>
</tr>
<tr>
<td>Issue Category</td>
<td>Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>----------------</td>
<td>----</td>
<td>-------------------</td>
<td>--------</td>
<td>---------</td>
<td>---------------------------------------</td>
<td>-------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>Responses. WKPICS highlighted the importance of recording non-events, such as documenting failed sampling attempts where procedures were followed but fishermen or merchants barred access to landings or a trip. These events could create bias so need to be accounted for in raised estimates.</td>
<td>WKPICS, 2011</td>
<td>All variables</td>
<td>Yes</td>
<td>The RDBES allows the recording of non-responses (see “Non-responses and missing values due to quota sampling” in the RDBES documentation)</td>
<td>-</td>
<td>The RDBES will allow this potential source of bias to be analysed.</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>Who collects the samples, the staff responsible for sampling, by the crew or by the port staff? There are potential conflicts of interest in some of these players that may induce bias in sampling selection that then propagates to final estimates.</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>Yes</td>
<td>All RDBES CS tables representing sampling stages includes variable “sampler” that allows the distinction between different sources of data, including self-sampling, control, observer, etc. That information can be used to evaluate this source of bias.</td>
<td>-</td>
<td>The RDBES will allow this potential source of bias to be documented.</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>Species replacement: species thrown away (discarded) because replaced by another. This behaviour, linked to the carriage capacity, must be evaluated if it occurs, either by a well-designed sampling programme (no bias) or by external source (risk of bias).</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>No</td>
<td>No</td>
<td>Unclear description. The RDBES does not seem to be able to analyse this behavior.</td>
<td>-</td>
</tr>
<tr>
<td>Issue Category</td>
<td>Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>----------------</td>
<td>----</td>
<td>--------------------</td>
<td>--------</td>
<td>---------</td>
<td>--------------------------------------</td>
<td>-------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>Damaged fish landed: some cases were reported of fishers proposing for sale incomplete, i.e., fish partially cut for any reason, such as bite by a shark.</td>
<td>WKACCU, 2008</td>
<td>Landing and discard weight</td>
<td>No</td>
<td>RDBES CL table includes variable “CLcatchCategory” that under code “RegDis” allows the reporting of logbook registered discards, exemptions to the landing obligation and damaged fish. The current code list does not yet allow for the separation between these categories.</td>
<td>Yes</td>
<td>This feature could be added in an upcoming update to the data model. CS data can be reported in different levels of processing. Analyses of these records may provide for additional indications of potential bias in official estimates.</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>Slipping behaviour: In general, this behaviour is linked to specific fisheries such as pelagic trawling. The more or less rare occurrence of rejecting all the catch before it comes on the vessel deck needs to be evaluated.</td>
<td>WKACCU, 2008</td>
<td>Slipping weight</td>
<td>Yes</td>
<td>RDBES CS data model allows the recording of slipping events as samples associated to variable “SSobservationActivityType”</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>Working conditions: evaluating the sampled weight with a scale needs proper conditions, which are not always possible. Sampling for discards needs also good conditions for taking the sample and enough time and space for carrying the scientific work. Any constraint on working conditions may lead to a bias in the final estimates.</td>
<td>WKACCU, 2008</td>
<td>Landing and discard weight</td>
<td>Partially</td>
<td>The RDBES BV table allows the declaration of different types of measurement equipment. That information can provide insight into the accuracy of some measurements.</td>
<td>Yes</td>
<td>The WGRDBEDGOV “Core group” have been discussing the implementation of a quality scale for all biological measurements.</td>
</tr>
<tr>
<td>Issue Category</td>
<td>Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>---------------</td>
<td>----</td>
<td>-------------------</td>
<td>--------</td>
<td>---------</td>
<td>--------------------------------------</td>
<td>-------------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>Staff trained for age reading: information such as the time since the last training or information on the experience of the staff are the elements to determine the risk of bias on age reading. Some international calibration workshop evaluate the competence of age readers for estimating age structure for stock assessment purpose, Age readers formally approved by such a forum, would lead to an absence of bias for this parameter; experience of the staff is an element to determine the risk of bias on estimating the sex of certain species (e.g., Pandalus).</td>
<td>WKACCU, 2008</td>
<td>Age structure, sex-ratio</td>
<td>No</td>
<td>The RDBES is not capable of holding information of this kind.</td>
<td>No</td>
<td>Supplementary information will be needed that can be included in national protocols and/or workplans. This is part of the overall aspects of training and possibilities to increase accuracy of biological determinations</td>
</tr>
</tbody>
</table>
| 20 | Quality documentation on biological variables: Existence of a validity control for the appropriateness of the reading to evaluate the true age (check with tagging or in vivo growing programs); Existence of a recent age reading calibration workshop; Existence of a recent international exchange in order to compare the results of age reading by several readers from different countries on the same material. Usually, the exchange is carried out in preparation of an age reading workshop or at regular interval to assess the need of convening such a workshop.; International reference set: Existence and routinely use of an agreed international reference set. ; The risk of bias is inherent to the species/stock itself, depending on the difficulty of reading the age. The international calibration workshops use software able to evaluate such a bias; Existence of a routine calibration validation of the equipment used.; How are immature issues being addressed? Is the method used well described and approved? ; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007, survey protocols); existence of a protocol for dealing with immature; Existence and follow-up of an international sampling protocol (ICES WKMAT 2007,
<table>
<thead>
<tr>
<th>Issue Category</th>
<th>Id</th>
<th>Issue description</th>
<th>Source</th>
<th>Affects</th>
<th>Can the RDBES tackle it at the moment?</th>
<th>Could the RDBES potentially tackle this in the future?</th>
<th>Comment</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>21</td>
<td>Processing and evaluation methods for age, sex and maturity: Some reading methods are known to be biased for estimating some or all ages.</td>
<td>WKACCU, 2008</td>
<td>Age structure</td>
<td>Partially</td>
<td>The RDBES BV table includes information on processing method for all biological measurements. Presently only codes for age structures exist, but codes for other relevant methods can be added.</td>
<td>Yes</td>
<td>The present RDBES code list would need to be expended.</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>Staff trained for species identification: information such as the time since the last training or information on the experience at sea are the elements to determine the risk of bias on species identification at the end of a sampling. This source of bias must be combined to the previous one as on one hand a species easy to identify do not present major risk of bias even for a novice, and on the other hand a species difficult to identify is not a source of bias if sampled by a taxonomist.</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>No</td>
<td>The RDBES is not capable of holding information of this kind. Supplementary information will be needed that can be included in national protocols and/or workplans.</td>
<td>No</td>
<td>This is part of the overall aspects of training and possibilities to increase accuracy of taxonomic identifications</td>
</tr>
<tr>
<td>Issue Category</td>
<td>Issue Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>----------------</td>
<td>----------</td>
<td>-------------------</td>
<td>--------</td>
<td>---------</td>
<td>--------------------------------------</td>
<td>-------------------------------------------------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>Existence of an identification key: photographs or sketches of species of relevance in a given fishing area are very useful tools to ensure correct species identification. The absence of such identification keys, however, is not to be considered a source of bias when the staff that conduct the species identification is trained and experienced in taxonomy.</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>No</td>
<td>RDBES is not capable of holding information of this kind. Supplementary information will be needed that can be included in national protocols and/or NWP.</td>
<td>No</td>
<td>This is part of the overall aspects of training and possibilities to increase accuracy of taxonomic identifications</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>Species subject to confusion: The risk of bias is inherent to the species itself, depending on the difficulty of its identification. A way of evaluating the bias could be through a reference table of species to be agreed by an inter-national forum. The setting of such a table, specific to fishing areas/regions, should be addressed by the ICES PGCCDBS.</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>Partially</td>
<td>The RDBES is not capable of knowing whether a recorded species has been mis-identified. The RDBES SL table contains the taxa recorded in each sampling event. The table can be compared with regional standards for completeness. If MS provide that level of detail, differences between SL tables of different sampling events may indicate lack of consistency between observers</td>
<td></td>
<td>Reference tables of species are now defined in EU legislation on data collection. However, not all sampling programmes in each country sample all the species in that list (some are a subset of species or are species focused)</td>
</tr>
<tr>
<td>Issue Category</td>
<td>Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>----------------</td>
<td>----</td>
<td>--------------------</td>
<td>--------</td>
<td>---------</td>
<td>----------------------------------------</td>
<td>-------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td>Unit definition: Existence and follow-up of an international agreed definition and specifications. Effort statistics obtained through a census or a sampling programme.</td>
<td>WKACCU, 2008</td>
<td>Yes</td>
<td>RDBES CE table includes variables “CEdataTypeForScientificEffort” and “CEdataSourceForScientificEffort” that provide that information. RDBES CE also requests effort estimation to be carried out according to STECF guidelines (WKTRANSVERSAL II report)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>26</td>
<td></td>
<td>Size of the catch effect: When catches are big and only a guestimated fraction has been sampled, the bias is more likely than when a significant fraction of the catch (say more than 10%) is taken for sampling. In general this information is absent even from the raw samples.</td>
<td>WKACCU, 2008</td>
<td>Partially</td>
<td>The RDBES allows the quantification of sampling fractions in all CS tables representing sampling stages. The RDBES can store the number of units sampled and the total number of units at each stage – it is then possible to identify small samples taken from large hauls. It is not currently possible to record how, for example, the size of a haul was calculated (e.g. measured or estimated).</td>
<td>-</td>
<td>-</td>
<td>WGRDBESGOV “Core group” to discuss and evaluate possibilities of detecting this source bias</td>
</tr>
<tr>
<td>27</td>
<td></td>
<td>Changes in fishers’ behaviour when observed, on board sampling.</td>
<td>WKPICS, 2011</td>
<td>No</td>
<td>The RDBES can’t identify this source of bias.</td>
<td>No</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Issue Cate-</td>
<td>Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>28</td>
<td>Taxonomic changes: changes in species nomenclature over time, e.g. the splitting of sandeel species in the face of new knowledge, may impact the consistency of a time-series.</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>Partially</td>
<td>New uploads to the RDBES will be required to use the latest, valid aphia id for a species. An aphia id may become invalid – in this case MS can be asked to re-upload data using the new code, or the RDBES host can update the aphia id. In RDBES consistency of species recording in time can be checked by comparing SL records across time periods.</td>
<td>-</td>
<td>-</td>
<td>Information on the species targeted by each sampling programme should be requested in national workplans. Those target species should be recorded in the SL table of RDBES</td>
</tr>
<tr>
<td>29</td>
<td>Area misreporting: like for the species misreporting, there may be a sudden increase of a species reported in an uncommon neighboring area. This type of bias may be assessed by checking the consistency between different sources e.g. logbooks, VMS, sales notes, questionair surveys, cpue trends of commercial vs. surveys, ...</td>
<td>WKACCU, 2008</td>
<td>Landings weight, effort</td>
<td>Partially</td>
<td>Not all different sources of information will be incorporated in the RDBES. But RDBES CS data can be used to estimate landings of a species in specific spatial domains that can then be evaluated against similar estimates in RDBES CL data and/or compared with other types of external data like VMS, survey data, etc.</td>
<td>-</td>
<td>-</td>
<td>WGRDBESGOV is promoting workshops (e.g., WKRDB-EST) with the aim of developing code for the RDBES / TAF. Some of that code will allow the identification of situations where this type of bias may impact final estimates.</td>
</tr>
<tr>
<td>Issue Category</td>
<td>Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>----------------</td>
<td>----</td>
<td>-------------------</td>
<td>--------</td>
<td>---------</td>
<td>---------------------------------------</td>
<td>-------------------------------------------------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>Missing part: ratio between the retained fractions estimated on-board by observers and the landings of a species. A statistical test can be performed to evaluate if the slope is significantly different from one.</td>
<td>WKACCU, 2008</td>
<td>Landing weight</td>
<td>Partially</td>
<td>RDBES CS data can be used to estimate total landings of a species that can then be compared with RDBES CL data</td>
<td>-</td>
<td>See the comment on issue 29</td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>Quantity misreporting: known as the most current bias in fisheries statistics, this bias may be assessed together with area misreporting and with the addition of sources like economic surveys and EU control database.</td>
<td>WKACCU, 2008</td>
<td>Landings weight, effort</td>
<td>Partially</td>
<td>RDBES CS data can be used to estimate landings of a species in specific spatial domains that can then be evaluated against similar estimates in RDBES CL data and/or compared with other official data</td>
<td>-</td>
<td>See the comment on issue 29</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>Species misreporting: A sudden increase of an unexpected species may occur in the statistics, thus pointing out a potential risk of species misreporting. This case is generally linked to quota consumption. Another way of detecting such a bias is dissimilarities between on-board observers reporting for the same fishing activity, or dissimilarities between on-board observers and landing statistics</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>-</td>
<td>Estimates derived from RDBES CS data may be compared with the official estimates provided in CL table to identify this type of bias.</td>
<td>-</td>
<td>See the comment on issue 29</td>
</tr>
<tr>
<td>Issue Category Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------------------------------------------------------------------</td>
<td>----------------</td>
<td>----------</td>
<td>----------------------------------------</td>
<td>--------------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Management measures leading to discarding behavior: the specification of the measure and the date of entry into force are indications of potential bias, if not monitored through a well-designed sampling program.</td>
<td>WKACCU, 2008</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Information on impacts of management measures can be pursued by independent projects looking into this issue.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>High grading; selecting a given size range for landing a species depending on the market demand or to reduce the quota consumption automatically change the discarding ogive. High grading behaviour may be evaluated by interviews and/or on-board observers.</td>
<td>WKACCU, 2008</td>
<td>Partially</td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Change in selectivity: bias linked to the characteristics of the gear and evaluation whether the length structure sampled is representative of the exact characteristics of the gears used at the population level.</td>
<td>WKACCU, 2008</td>
<td>Yes</td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Issue Category</td>
<td>Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>----------------</td>
<td>----</td>
<td>-------------------</td>
<td>--------</td>
<td>---------</td>
<td>----------------------------------</td>
<td>---------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Statistical processing: when direct biological measurements (e.g., age reading, mean weight) are impossible, statistical methods may be used to estimate those variables that may introduce bias in analysis (e.g., length-weight relationships, Von Bertalanffy models). The time between the references used for modellings and the actual time strata is an indication on the potential induced bias.</td>
<td>36</td>
<td>WKACCU, 2008</td>
<td>Age structure, mean-weight</td>
<td>Partially</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Incomplete ALK: Appropriateness of the imputation rules used, e.g., for filling length classes without age information.</td>
<td>37</td>
<td>WKACCU, 2008</td>
<td>Age structure</td>
<td>Partially</td>
<td>Estimates carried out under RDBES will be registered in TAF allowing future discussion on appropriateness of methods and calculations used.</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Issue Category</td>
<td>Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>----------------</td>
<td>----</td>
<td>-------------------</td>
<td>--------</td>
<td>---------</td>
<td>----------------------------------------</td>
<td>------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>Specific handling of some biological variables to reduce bias (e.g., Plus group: bias linked to the setting of the plus group, and the existence or not of international agreement; Skipped spawning: following ICES WKMAT recommendation, is skipped spawning known to happen and taken into account?; Catchability effect: for some species the catchability by sex vary over time. If such behaviour related change in catchability occurs, do the estimates take this into account following an agreed protocol?)</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>Partially</td>
<td>Estimates carried out under RDBES will be registered in TAF allowing future discussion on appropriateness of methods and calculations used.</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>Raising variable: For raising to the population, different raising procedures must be compared and also raising the retained fraction to be compared with the landing statistics is a solution to assess the relevance of the variable used for raising (WKDRP, 2007).</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>Partially</td>
<td>RDBES is designed in a way that makes it possible to estimate catches by a variety of probabilistic methods including ratio estimation. The specifics of ratio estimation within the RDBES have not yet been developed.</td>
<td>Yes</td>
<td>There are initiatives are planned that make these analyses possible (e.g. WKRATIO scheduled for 2021)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WKRATIO to provide indications on analyses comparing different methods. WGRDBESGOV to include implementation of those analysis in the system.</td>
<td></td>
</tr>
<tr>
<td>Issue Category</td>
<td>Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>---------------</td>
<td>----</td>
<td>-------------------</td>
<td>--------</td>
<td>---------</td>
<td>--------------------------------------</td>
<td>------------------------------------------------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>40</td>
<td>40</td>
<td>Conversion factor: information such as the age and the methodology used for the conversion factor, are indications on the potential induced bias. The magnitude of the conversion factor used is also an indication, e.g. estimating landing weight from fillet or from gutted fish will lead to different amplification of a bias.</td>
<td>WKACCU, 2008</td>
<td>landings and discrad weight, length frequencies, mean weights</td>
<td>Partially</td>
<td>Yes</td>
<td>Inclusion of similar conversion factors in BV may be needed</td>
<td>WGRDBESGOV “Core group” to evaluate possibilities of fully detecting this source of bias and help to improve conversion factors</td>
</tr>
<tr>
<td>Issue Category</td>
<td>Id</td>
<td>Issue description</td>
<td>Source</td>
<td>Affects</td>
<td>Can the RDBES tackle it at the moment?</td>
<td>Could the RDBES potentially tackle this in the future?</td>
<td>Comment</td>
<td>Actions</td>
</tr>
<tr>
<td>----------------</td>
<td>----</td>
<td>-------------------</td>
<td>--------</td>
<td>---------</td>
<td>--------------------------------------</td>
<td>-------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Estimation</td>
<td>41</td>
<td>Grouping statistics: some commercial naming include groups of several species, e.g. lophius, megrims. It may also be the case that a commercial naming includes incidentally other species, as often encountered with the elasmobranchs (e.g. mixture of ray species in a box of Raja clavata). Scientific sampling surveys are generally used to quantify the percentage of species within the relevant commercial names, and if it is the case, there is no major risk of bias.</td>
<td>WKACCU, 2008</td>
<td>All variables</td>
<td>Yes</td>
<td>The RDBES SL table includes both commercial and scientific denominations allowing for the recording of sampled species composition within commercial denominations. Those recordings can be used to split national statistics of very aggregated categories (e.g., split national landings of monkfishes into landings of angler and black-bellied angler)</td>
<td>-</td>
<td>This feature has been implemented but not fully tested with real data yet.</td>
</tr>
</tbody>
</table>
Annex A3.5  Example reports to illustrate potential sources of bias

Comparison of sample data to population data report

The following plots illustrate graphical reports made using the RDBES data format that can be used by MS to inspect their data and help them identify errors in their datasets and sources of potential bias in their catch estimates. The R code for production of these plots can be found on Github\(^ {14} \) and had as a starting block earlier developments achieved during 2019 RCG intersessional subgroups on fisheries and sampling overviews.

Note: These type of plots should be used with caution since the patterns they may highlight can only be evaluated with full knowledge of the original sampling plan and its implementation (i.e. by combining RDBES, national sampling plan information, and implementation notes) and some observed patterns might be explained by either biases in sampling or by the natural variation occurring in a probabilistic sampling scheme.

The first figure displays landings per month from a certain fleet (in tonnes, bars, left y-axis) alongside numbers of trips sampled including and excluding non-response (a potential source of bias). By analysing plots like this, data submitters will be able to scrutinize aspects of quality such as errors in data (e.g. is the large number of samples in month 10 real or an error in the data submitted?). In parallel, the graph provides data estimators with a quick overview of the data available along a temporal dimension, facilitating the consideration of potential sources of bias such as non-responses (e.g. could the non-responses observed impact the estimates?) and departures from the sampling plan during its implementation that may need to be accounted for during estimation to avoid biases (e.g. if the goal of this plan was to sample 2 trips per month, clear departures in the implementation occurred in many months). Additionally, the graph may also allow insight into potential precision issues such as those brought about by lack of coverage in months where landings could be known to be important and highly variable.

The second figure displays similar data (population and sampled totals) but focuses on the spatial coverage. The figure displays landings (right map, in tonnes) and effort (left map, as days-at-sea) per ICES statistical rectangle alongside the number of hauls samples and their position. These type of plot can be used to visually check for potential biases in the sampling data such as those caused by “observer effects”, but is also useful to identify biases potentially caused by errors in the data (e.g., a sample collected where no effort took place) or to visually check for abnormally high concentration of data in some regions and not others (either probabilistic, or more importantly, non-probabilistic).

Different variations of both these type of plots can be coded, e.g., in the case of the bar plot one might be interested in plotting effort instead of landings or seeing the data at different time resolutions (e.g. quarters instead of months). It most likely that several of these graphs will need to be combined during analyses of potential biases (e.g. if sampling goals are established at quarterly level, data estimators may still be interested in monthly bar plots to check if implementation issues such as all trips being sampled in the first month of each quarter could potentially impact the estimates).

\(^ {14} \) https://github.com/ices-eg/WK_RDBES/tree/master/Special_Request_20_05
CLofficialWeight1000 by CLmonth

- target trips including non-response
- trips effectively sampled

Month

Number of trips

sum of CLofficialWeight1000

y:100% x:100% all_data
Summary of selection methods report

This report provides a summary of the different selection methods used in a sampling scheme – in particular it highlights the number of units that were selected using probabilistic and non-probabilistic methods at each stage. In the case where the data contains information on un-sampled units these are also summarized. The R code can be on Github\textsuperscript{15}.

\textsuperscript{15} https://github.com/ices-eg/WK_RDBES/tree/master/Special_Request_20_05
**Bias - Summary of selection methods and declaration of none sampled, ESP-AZTI_DCF_Onboard_Sampling 1966**

**Selection methods**

The selection methods are grouped the following way:

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Selection method type</th>
</tr>
</thead>
<tbody>
<tr>
<td>CENSUS</td>
<td>Census (CENSUS)</td>
<td>Probabilistic</td>
</tr>
<tr>
<td>NotApp</td>
<td>NotApp</td>
<td>NA</td>
</tr>
<tr>
<td>NotSam</td>
<td>NotSam</td>
<td>NA</td>
</tr>
<tr>
<td>NPHA</td>
<td>Ad Hoc Sampling (NPHA)</td>
<td>None-probabilistic</td>
</tr>
<tr>
<td>NPEJ</td>
<td>Expert Judgement (NPEJ)</td>
<td>None-probabilistic</td>
</tr>
<tr>
<td>NOPG</td>
<td>Quota Sampling (NOPG)</td>
<td>None-probabilistic</td>
</tr>
<tr>
<td>SRSWOR</td>
<td>Simple Random Sampling Without Replacement (SRSWOR)</td>
<td>Probabilistic</td>
</tr>
<tr>
<td>SRSWR</td>
<td>Simple Random Sampling With Replacement (SRSWR)</td>
<td>Probabilistic</td>
</tr>
<tr>
<td>SYSS</td>
<td>Systematic Sampling (SYSS)</td>
<td>Probabilistic</td>
</tr>
<tr>
<td>UPSWOR</td>
<td>Unequal Probability Sampling Without Replacement (UPSWOR)</td>
<td>Probabilistic</td>
</tr>
<tr>
<td>UPSWR</td>
<td>Unequal Probability Sampling With Replacement (UPSWR)</td>
<td>Probabilistic</td>
</tr>
</tbody>
</table>

NotApp and NotSam are not considered in these overviews.

**Overall - summary**

Number of samples selected in a probabilistic and none-probabilistic way per sampling level.

The use of probabilistic selection is highlighted in green and number of none-probabilistic selections are highlighted in red.

<table>
<thead>
<tr>
<th>Sampling scheme</th>
<th>Year</th>
<th>Country</th>
<th>Sampling unit level</th>
<th>Sampling unit type</th>
<th>Number of probabilistic selections</th>
<th>Number of none-probabilistic selections</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESP-AZTI_DCF_Onboard_Sampling</td>
<td>1966</td>
<td>DK</td>
<td>1</td>
<td>Vessel</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>ESP-AZTI_DCF_Onboard_Sampling</td>
<td>1966</td>
<td>DK</td>
<td>2</td>
<td>Fishing trip</td>
<td>72</td>
<td>3</td>
</tr>
<tr>
<td>ESP-AZTI_DCF_Onboard_Sampling</td>
<td>1966</td>
<td>DK</td>
<td>3</td>
<td>Fishing operation</td>
<td>69</td>
<td>3</td>
</tr>
<tr>
<td>ESP-AZTI_DCF_Onboard_Sampling</td>
<td>1966</td>
<td>DK</td>
<td>4</td>
<td>Species selection</td>
<td>69</td>
<td>3</td>
</tr>
<tr>
<td>ESP-AZTI_DCF_Onboard_Sampling</td>
<td>1966</td>
<td>DK</td>
<td>5</td>
<td>Sample</td>
<td>720</td>
<td>3</td>
</tr>
</tbody>
</table>
## Detailed summary per sampling level

Number per selection methods used per sampling level.

In each table the stratum from the level above is stated to make identification easier.

### 1. Sampling unit

<table>
<thead>
<tr>
<th>Sampling unit type</th>
<th>Sampling scheme</th>
<th>Year</th>
<th>Country</th>
<th>Stratum above</th>
<th>Stratum here</th>
<th>Selection method</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vessel</td>
<td>ESP-ADIF Onboard_Sampling</td>
<td>1988</td>
<td>DK</td>
<td>DE_stratum1</td>
<td>VS_stratum1</td>
<td>RPSU</td>
<td>5</td>
</tr>
<tr>
<td>Vessel</td>
<td>ESP-ADIF Onboard_Sampling</td>
<td>1988</td>
<td>DK</td>
<td>DE_stratum1</td>
<td>VS_stratum2</td>
<td>SRSWR</td>
<td>5</td>
</tr>
<tr>
<td>Vessel</td>
<td>ESP-ADIF Onboard_Sampling</td>
<td>1988</td>
<td>DK</td>
<td>DE_stratum2</td>
<td>VS_stratum1</td>
<td>RPSU</td>
<td>5</td>
</tr>
<tr>
<td>Vessel</td>
<td>ESP-ADIF Onboard_Sampling</td>
<td>1988</td>
<td>DK</td>
<td>DE_stratum2</td>
<td>VS_stratum2</td>
<td>SRSWR</td>
<td>5</td>
</tr>
</tbody>
</table>

### 2. Sampling unit

<table>
<thead>
<tr>
<th>Sampling unit type</th>
<th>Stratum above</th>
<th>Stratum here</th>
<th>Selection method</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fishing trip</td>
<td>VS_stratum1</td>
<td>U</td>
<td>RPSU</td>
<td>3</td>
</tr>
<tr>
<td>Fishing trip</td>
<td>VS_stratum1</td>
<td>U</td>
<td>SRSWR</td>
<td>7</td>
</tr>
<tr>
<td>Fishing trip</td>
<td>VS_stratum2</td>
<td>U</td>
<td>RPSU</td>
<td>5</td>
</tr>
<tr>
<td>Fishing trip</td>
<td>VS_stratum2</td>
<td>U</td>
<td>SRSWR</td>
<td>5</td>
</tr>
</tbody>
</table>

### 3. Sampling unit

<table>
<thead>
<tr>
<th>Sampling unit type</th>
<th>Stratum above</th>
<th>Stratum here</th>
<th>Selection method</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fishing operation</td>
<td>U</td>
<td>U</td>
<td>SRSWR</td>
<td>60</td>
</tr>
</tbody>
</table>

### 4. Sampling unit

<table>
<thead>
<tr>
<th>Sampling unit type</th>
<th>Stratum above</th>
<th>Stratum here</th>
<th>Selection method</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Species selection</td>
<td>U</td>
<td>U</td>
<td>CENSUS</td>
<td>60</td>
</tr>
</tbody>
</table>

### 5. Sampling unit

<table>
<thead>
<tr>
<th>Sampling unit type</th>
<th>Stratum above</th>
<th>Stratum here</th>
<th>Selection method</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>U</td>
<td>U</td>
<td>SRSWR</td>
<td>120</td>
</tr>
</tbody>
</table>
## None sampled

### Overall - summary

Only sampled data declared

<table>
<thead>
<tr>
<th>Sampling scheme</th>
<th>Year</th>
<th>Country</th>
<th>Sampling unit level</th>
<th>Sampling unit type</th>
<th>Sampled or not</th>
<th>Industrial Decline</th>
<th>No Contact</th>
<th>Details Available</th>
<th>Observer Decline</th>
<th>Other</th>
<th>Quota Reached</th>
<th>No reason given</th>
</tr>
</thead>
</table>

### Detailed summary per sampling level

Number of none sampled per reason for not sampling.

In each table the stratum from the level above is stated to make identification easier.

#### 1. Sampling unit

Only sampled data declared

<table>
<thead>
<tr>
<th>Sampling unit type</th>
<th>Sampling scheme</th>
<th>Year</th>
<th>Country</th>
<th>Stratum above</th>
<th>Stratum here</th>
<th>Sampled or not</th>
<th>Reason for not sampling</th>
<th>Number of samples</th>
</tr>
</thead>
</table>

#### 2. Sampling unit

Only sampled data declared

<table>
<thead>
<tr>
<th>Sampling unit type</th>
<th>Stratum above</th>
<th>Stratum here</th>
<th>Sampled or not</th>
<th>Reason for not sampling</th>
<th>Number of samples</th>
</tr>
</thead>
</table>

#### 3. Sampling unit

Only sampled data declared

<table>
<thead>
<tr>
<th>Sampling unit type</th>
<th>Stratum above</th>
<th>Stratum here</th>
<th>Sampled or not</th>
<th>Reason for not sampling</th>
<th>Number of samples</th>
</tr>
</thead>
</table>

#### 4. Sampling unit

Only sampled data declared

<table>
<thead>
<tr>
<th>Sampling unit type</th>
<th>Stratum above</th>
<th>Stratum here</th>
<th>Sampled or not</th>
<th>Reason for not sampling</th>
<th>Number of samples</th>
</tr>
</thead>
</table>