WORKSHOP ON POPULATION OF THE RDBES DATA MODEL (WKRDB-POP)

VOLUME 1 | ISSUE 24

ICES SCIENTIFIC REPORTS

RAPPORTS
SCIENTIFIQUES DU CIEM

ICES INTERNATIONAL COUNCIL FOR THE EXPLORATION OF THE SEA
CIEM COUNSEIL INTERNATIONAL POUR L'EXPLORATION DE LA MER

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

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

ISSN number: 2618-1371 I

ICES Scientific Reports

Volume 1 | Issue 24

WORKSHOP ON POPULATION OF THE RDBES DATA MODEL (WKRDB-POP)

Recommended format for purpose of citation:

ICES. 2019. Workshop on Population of the RDBES Data Model (WKRDB-POP).
ICES Scientific Reports. 1:24. 47 pp. http://doi.org/10.17895/ices.pub.5277

Editors

David Currie • Edvin Fuglebakk

Authors

Maciej Adamowicz • Liz Clarke • Laurent Dubroca • Jon Elson • Ana Claudia Fernandes • Karolina Molla Gazi • Kirsten Birch Hakansson • Irina Jakovleva • Ain Lankov • Twan Leijzer • Pedro Lino • Sofie Nimmegeers • Christina Petterson • Jukka Pönni • Nuno Prista • Dália CC. Reis • Ana Ribeiro Santos Jose Rodrigues • Petri Sarvamaa • Ivo Šics • Marijus Špegys • Sven Stötera • Marta Suska • Ioannis Thasitis • Sofie Vandemaele • Julia Wischnewski • Lucia Zarauz

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

Contents

i Executive summary

The aims of this workshop were to explain the new RDBES data model, assist in populating it with real data, and identify any issues or problems.

The data model was explained using a combination of presentations and hands-on sessions. A set of evaluation discussions was held in plenary. During the hands-on sessions, participants worked on mapping their own data to the RDBES data model with guidance from the "RDBES Core Group" members (the group of people developing the RDBES data model). Participants were encouraged to complete the hands-on sessions in the order presented in the agenda and were able to progress at their own pace.

Although issues have been identified and are documented in this report it is not thought that any of them are serious impediments to moving forward with the RDBES data model. The RDBES Core Group will look at the results of this workshop and either respond to individual questions or adapt the data model and documentation as required.

The workshop chairs will keep in contact with the participants to keep them up-to-date with any data model changes and answer questions. This workshop (WKRDB-POP) and the WKRDB-EST workshop to be held later in 2019 should be considered as two halves of the same process so the chairs will also encourage the participants to carry on working on their data transformation routines so that they can attend WKRDB-EST.

ii Expert group information

Expert group name	Workshop on Population of the RDBES Data Model (WKRBD-POP)
Expert group cycle	NA
Year cycle started	2019
Reporting year in cycle	1/1
Chair(s)	David Currie, Ireland
	Edvin Fuglebakk, Norway
Meeting venue and dates	18–22 February 2019, ICES HQ, Denmark, (29 participants)

iii Term of reference

Term of reference	Addressed in this report
a) Describe and explain the RDBES data model to national data submitters using worked examples.	Yes
b) Provide hands-on guidance and assistance in converting data to the RDBES data format for national data submitters.	Yes
c) Identify and document any problems in converting national data formats to the RDBES format.	Yes

1 Introduction

Currently national institutes submit raised commercial fisheries data for use in stock assessment using the InterCatch system. They also submit detailed commercial fisheries data for use by the EU Regional Coordination Groups (RCGs) to the Regional Database (RDB) which is hosted by ICES.

The problems in the current system include:

- Lack of transparency;
- Duplication of effort;
- Lack of consistency;
- Lack of data quality indicators.

The Regional Database and Estimation system (RDBES) is currently in development and will replace both InterCatch and the existing RDB. It will store detailed commercial fisheries sample data alongside improved versions of the CE and CL tables of the old RDB and will allow sample data to be raised for use in stock assessments in a transparent manner.

The aims of the RDBES are:

1. To make data available for the RCGs;
2. To 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.

The RDBES should be seen as part of the movements towards:

1. Statistically Sound Sampling Schemes (4S);
2. Greater regional coordination;
3. Transparent Assessment Framework (TAF);
4. Improved estimates to ICES stock assessments and advice.

The timeline for the transition between InterCatch/RDB to the new RDBES is given in Table 1.1.

In order to fulfil this timeline, it is necessary to explain the model to data submitters and give them support in populating it with their national data. Once sufficient data are converted to the RDBES format the estimation scripts can then start to be developed. To this end the current WKRDB-POP workshop and the later WKRDB-EST workshop should be considered two parts of the same process - the outputs from WKRDB-POP are vital as the inputs to WKRDB-EST.

Table 1.1. Timeline for RDBES development

	RDB System	InterCatch	RDBES
2019	Production Data in/out	Production Data in/out	Development Test data in/out
2020	Production Data in/out	Production Data in/out	Test by selected stocks
2021	Production Data in/out	Production Data in/out	Test by all stocks
2022	Stay alive Data out	Stay alive Data out	Data call for 2021 data
2023	Stay alive Data out	Stay alive Data out	Data call for 2022 and all older years

The RDBES data model for sampled data has been developed over a series of ICES workshops and although some small changes are still being made it is considered to be very close to its final format. The RDBES format for effort and landings data (equivalent to the CE and CL formats in the current RDB) has not been considered in detail yet - this process will begin in March via a Skype meeting of the "Core group". The WKRDB-EST workshop will almost certainly identify extra features will be needed in this format.

It is important to note that the long-term funding of the RDBES is still unresolved. The existing RDB maintenance and hosting is funded by the European Commission under an administrative agreement with ICES. The initial development of the RDBES has been funded by ICES, the European Commission (funding two workshops), and ICES/EU member states (by contributing experts' time, and travel and subsistence for meetings). The funding of the RDBES development beyond October 2019 has not yet been agreed.

2 Describe and explain the RDBES data model to national data submitters using worked examples (ToR a)

Version 1.16 of the RDBES data model was used for the workshop and the details were provided both via the public GitHub repository (https://github.com/ices-tools-dev/RDBES) and the workshop SharePoint.

The data model was explained using a combination of presentations, hands-on sessions, and evaluation discussions in plenary.

During the hands-on sessions, participants worked on mapping their own data to the RDBES data model with guidance from the core-group members. Participants were encouraged to complete the hands-on sessions in the order presented in the agenda. Participants were able to progress at their own pace.

The ICES Data Centre have created a development version of the data uploader and validator which was presented to the participants along with a draft of the data upload format (following the same pattern as the RDB csv upload files). This currently validates against v1.15 of the RDBES data model so it could not be used with the data model version used in the workshop. The ICES Data Centre will update the system to validate against v1.16 of the RDBES model and make it available to the participants as soon as possible.

3 Provide hands-on guidance and assistance in converting data to the RDBES data format for national data submitters. (ToR b)

The workshop participants were asked to create a summary of the work they performed, what was still outstanding, and any issues they encountered. This section contains those reports.

3.1 Thünen-Institute, Germany

Sampling programmes considered

OF (Baltic Sea):

- on-board sampling programme (observer);
- "self-sampling" (sample purchased from fisher);
- harbour-sampling (discard sampling programme for SPF fisheries at the processing plant).

(exemplary datasets of two trips for each type in the sampling programme from 2017)

NS (North Sea):

- on-board sampling (both observer and self-sampling from 2017)

Progress made

OF (Baltic Sea): the two modi of the demersal Sampling programme (observer and self-sampling) both fit into the Hierarchy 1 scheme, the example datasets could be fitted to the respective tables. Some specific questions about some fields remain though.

The harbour-sampling was identified as a Hierarchy 4, upper Hierarchy tables were filled out, but also here, some fields were not yet filled (e.g. LEsequencenumber).

Lower Hierarchies A and B could be filled out. Also Hierarchy D (only bulk measurement of weight, but no length or biological information are gathered) appears regularly in our sampling.

NS (North Sea):

- Hand-made (tested) example for one trip (at sea sampling)
- Upper and lower hierarchies were determined (1 and A/B, respectively).
- Related problems were identified (see below).

Issues identified

OF (Baltic Sea):

The Example dataset did not contain samples from the SPF sampling, which is slightly different from the demersal at-Sea sampling – this might have another Hierarchy and needs to be checked. The BMS sampling programme also need to be evaluated and the Hierarchy identified. Selection methods in the different steps must be checked (with/without replacement? etc.).

NS (North Sea):

The most appropriate upper hierarchy is 1.

Well-suited lower hierarchies are B (length measuring) and A (stratified sampling, 10-12 fish per length class for biological parameters).

Vessel selection (VS): Vessels are selected from telephone list (approach is rather opportunistic, due to the small number of vessels), selection method is UPSWR. But in the case of some companies (vessel owners) vessels are selected for us by a company (selection method NPAH or anything else?)

Lack of information concerning unsampled hauls in the national database: columns FOnoSampReason, FOstartLat, FOstartLon, FOstopLon, FOstopLat, FOdep, FOwaterDep (FO table) can't be filled for unsampled hauls.

No information about sample unit (basket etc.) in the national database: column SAunitType (SA table) can't be filled.

Length data in the national database are given in scm, should they be transformed to the mm-length class (lower boundary) or to the real mm-length for the field BVvalue (table BV)?

The meaning of some fields is not really clear, e.g.:

- SAtotalWeightMeasured and SAtotalWeightLive (SA table), are they the same for the at-sea sampling?
- what is the difference between BiovarID and BVfishID, if they both are unique? Suggestion: first n digits as a serial number of fish (let's assume n fish sampled totally) + 1 digit more for measurement type (0 - length, 1 - age etc.) as a PK column.

Changes that are required either in national database or the RDBES

OF (Baltic Sea):

National database: for the new RDBES we'll need to rename our "self-sampling" as it doesn't fit the definition (technically it's an observer sampling, as we're buying unsorted catches. However, it is also not a real observer trip, e.g. in terms of PETS bycatch observations, gear information, etc.).

The sampling frame per stratification (e.g. VStotal, FTtotal) needs to be added in our data compilation process.

In the RDBES: The catch category "BMS" does not yet appear in the RDBES code list, it should however be added, as we also conduct BMS sampling at landing-sites and it is already reported to Intercatch and RDB.

NS (North Sea):

A field containing information about shipowners has to be added to the national database.

A field containing information about sampling unit has to be added.

Species selection list has to be discussed with colleagues, the corresponding parameter table has to be created and included to the coding process.

3.2 Natural Resources Institute Finland

Sampling programmes considered

- Commercial pelagic trawl fisheries (herring, sprat, vendace, 2019 onwards)
- Legacy Commercial data (same as above, up to 2018)

Progress made

Both lower and upper hierarchy identified.

- Lower hierarchy is A or both B and C
- Upper hierarchy is 1

Initial export code for data from 2019 onwards using hierarchies 1, B and C. RDBES 1.17 should allow for simplifying to hierarchy 1 and A. Legacy data differs only in ad-hoc vessel selection.

Tables populated during the workshop:

- Sampling Design (DE)
- Sampling Details (SD), currently omitted from export
- Vessel Selection (VS)
- Fishing Trip (TR)
- Fishing Operation (FO)
- Species Selection (SS)
- Species List Details (SL), currently omitted from export
- Sample (SA)
- Frequency Measure (FM)
- Biological Variable (BV)

Issues identified

- Hierarchy 1 is at-sea, but sampling is onshore.
- Count of hauls is mandatory in trip aggregated data, this might not be available.
- Stock assessment does not care about this since samples are trip aggregated and from whole catch. Better make it optional in RDBES.
- Amount of waste in catch was considered, but that will not be included in RDBES.

Changes that are required either in national database or the RDBES

- RDBES: Need onshore variant of hierarchy 1 or some kind of harbour-observer.
- RDBES: Count of hauls should be optional for trip aggregated data
- RDBES: Add all fish species option for species details
- RDBES: Add weight field in FM record (or equivalent)

3.3 Wageningen Marine Research, Netherlands

Sampling programmes considered

- Pelagic at-sea sampling scheme (herring)
- Demersal onshore sampling (plaice)
    - Stratification by Quarter

Both for year 2017.

Progress made

- Pelagic
    - Hierarchies identified:

    - Upper Hierarchy : 1
    - Lower Hierarchy : Depends on the corresponding species (A, B or D). For herring the lower hierarchy is A.
    - Tables populated:
    - None
- Demersal
    - Hierarchies identified:
- Upper Hierarchy : 7
    - Lower Hierarchy : Depends on the corresponding species (A, B or D). For plaice the lower hierarchy is A.
    - **Tables populated:**
    - All upper hierarchy tables (auxiliary as well). The Species List (SL) need to be updated.

Preliminary code developed.

Issues identified

- We did not have time to populate the tables for the pelagic sampling scheme.
- Issues understanding some field definitions from the excel file descriptions (for example the difference between SAtotalWtLive and SAtotalWtMes).
- We need to develop code to adapt to the needs of the RDBES.

Changes that are required either in national database or the RDBES

Database:

- We do not store information in our database about the sampling unit total (SAtotal).

3.4 Swedish University of Agricultural Sciences, Sweden

Sampling programmes considered

All the commercial sampling schemes currently developed by the Institute of Marine Science of SLU Aqua under NWP 2018–2019 were considered for purposes of the identification of upper hierarchies (see Table 2.1). With regards to the population of the RDBES data model with real data, effort was put previous to the meeting into the population of hierarchy 2 with data from sampling schemes “SWE - SkaKat (at-sea) - Act - 20 - MixTrawl” and “SWE - SkaKat (at-sea) - Act - 20 - NepTrawlGrid” and, during the meeting into the population of hierarchy 8 with data from sampling scheme “SWE - Balt (self) - Act - 22/32 - DemTrawl”.

Table 3.1. Correspondence between sampling schemes (as reported in NWP 2018/19) and RDBES v1.16 upper hierarchies

Scheme	Scheme	Proposed hierarchies	Probable Future developments
Baltic at-sea	SWE - Balt (at-sea) - Act - 24/25 - DemTrawl	2 (1,3)	new1
Baltic self-sampling	SWE - Balt (self) - Act - 22/32 - DemTrawl	8	---
Baltic at-sea / self-sampling	SWE - Balt (at-sea/self) - Pass – 23, 24, 25, 27/29 - DemNets	3,8	new1
Baltic at-sea / self-sampling	SWE - Balt (at-sea/self) - Pass – 24, 25 - DemLonglines	3, 8	---
Skagerrak/Kattegat at-sea	SWE - SkaKat (at-sea) - Act (several)	2	1,3
Skagerrak/Kattegat at-sea	SWE - SkaKat (at-sea) - Pass - 20/21 - NepPots	2	1,3
Baltic other (market, stock specific)	SWE - Balt (stock spec) - Act - 24/29 - HerSpr	?	new2
Skagerrak/Kattegat other (market stock specific)	SWE - SkaKat (stock spec) - Act – 20, 21 - Cod	?	new2
Skagerrak/Kattegat other (market stock specific)	SWE - SkaKat (stock spec) - Act – 20, 21 – HerSpr	?	new2

Progress made

Identification of upper hierarchies:

Upper hierarchies were identified for the different sampling schemes (Table 3.1). In some cases, the sampling schemes could be unequivocally assigned to a hierarchy, but not all.

Population of the format:

- SWE - SkaKat (at-sea) - Act - 20 – MixTrawl
  SWE - SkaKat (at-sea) - Act - 20 - NepTrawlGrid
  - Population of RDBES format with 2016 data
  - All RDBES tables were populated
- SWE - Balt (self) - Act - 22/32 - DemTrawl
  - population of RDBES with 2017 data
  - All RDBES tables were populated

Issues identified

Identification of upper hierarchies:

Some difficulties were experienced in univocally assigning sampling schemes to RDBES hierarchies. The main causes for these difficulties were:

a) shortcomings in the 4S specification of the present national sampling protocols (e.g. scheme "SWE - Skagerrak/Kattegat (stock spec)", "SWE - Balt (stock spec)")[1]

b) departures in sampling brought about by constraints felt in the practical implementation of the sampling schemes (WKPICS). E.g. sampling scheme "SWE - Balt (at-sea) - Act - 24/25 - DemTrawl" is traditionally considered under upper hierarchy 2 which has trip as the PSU. Trip was the established sampling unit and hierarchy 2 corresponds to sampling class A of ICES WKPICS 2 (ICES 2013) having well defined estimation methods. However, in practice a fishing trip can rarely be selected randomly because lists of trips do not exist at the beginning of the quarters that allow for controlled probability of selection. Furthermore, the need to plan staff time leads to the planning of trips on a weekly basis depending on observer availability a situation that also causes departures from the originally intended random trip selection. Accordingly, it is not infrequent that alternative hierarchies like hierarchy 1 (that involves selection of vessel at the first stage of sampling) or hierarchy 3 (that involves selection of time/week at the first stage of sampling) provide suitable alternative hierarchies to fit the data collected under this sampling scheme.

c) need for new hierarchies in RDBES. Two new upper hierarchies were identified that need to be included in the RDBES: new1 (vessel (UPSWOR) > time ->trip (within time) > haul > spp select > sample) and new2 (location > time > spp select > sample).

Population

The RDBES data model was successfully populated but significant difficulties were felt during that exercise. These mostly related to non-availability of some information in national database (e.g. estimates of total weights of discards are stored in national database after some raising - the total number of baskets available/sampled is only available in paper protocols; species lists and sampling probabilities are not declared in the national database; etc.) and a need to combine information from different databases before populating (e.g. to obtain total number of trips logbook data must be fetched from another database).

Changes required

The following changes are examples of adaptations needed in national databases:

- Explicit information on the sampling hierarchies used needs to be added;
- Details on species selection, namely the species lists used (if any), need to added;
- Sampling variables that define the sampling fraction, i.e. no. of baskets/boxes in catch fraction and number of baskets/boxes sampled, need to be added;
- A link needs to be specified between length distribution and specimens so that lower hierarchy A can be properly represented;
- New report formats are needed;
- Other minor changes (e.g. incorporation of harbour of departure and arrival; date-time of landing and date-time of sampling may need to be distinguished).

The following changes are examples of adaptations needed in national sampling designs and protocols:

- Clarification of sample selection procedures at upper hierarchy level (see above)
- Definition of best practice when dealing with practical constraints (see above).

[1] The Institute of Marine Research of SLU Aqua is developing efforts to review, update and/or clarify some of its sampling schemes, adapting them to EU-MAP requirements of statistically sound sampling.

- Clarification of the recording of species lists, in particular in what concerns the recording of incidental bycatch (which hauls is it registered? which parts of the fishing operation where effectively observed?);
- Separate field recording of some sampling unit (e.g. species and length composition by box or basket);
- Recording of marine litter (note: not currently required under EU-MAP nor handled by the RDBES but likely of future interest. Changes are already made in national database for entering litter.).

The following changes are examples of adaptations needed in the RDBES:

- Two new upper hierarchy are needed:
  - **new1 (vessel > time > trip (within time) > haul > spp select > sample)** (Figure 3.1) is needed to accommodate at-sea observer and self-sampling programmes where for vessel selection (e.g. at start of the quarter, with UPSWOR) is followed by time selection (e.g. weeks in quarter when observers are available) and selection of trips to sample (within weeks with observers).
  - **new2 (location > time > spp select > sample)** (Figure 3.2) is needed to accommodate stock specific onshore programmes where sampling takes place by size category within preselected combinations of locations (e.g. harbours) and time (e.g. days).

Schemes for these new hierarchies were discussed and developed during the workshop in a format similar to the one used in annexes of the documentation of the RDBES.

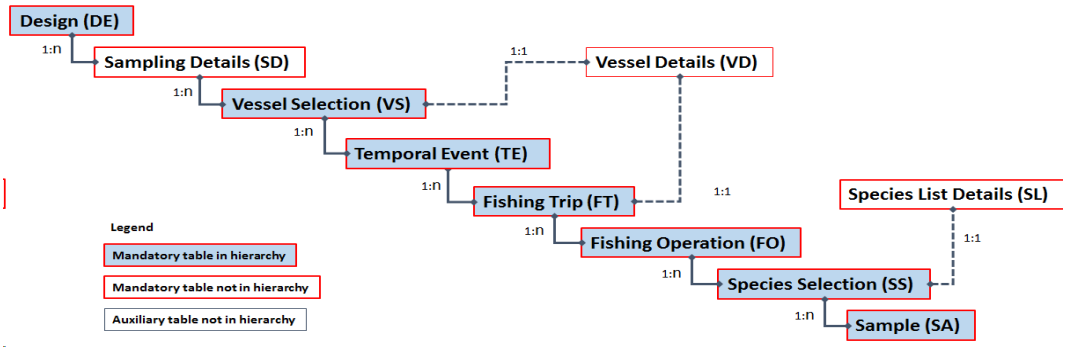


Figure 3.1. new1 hierarchy.

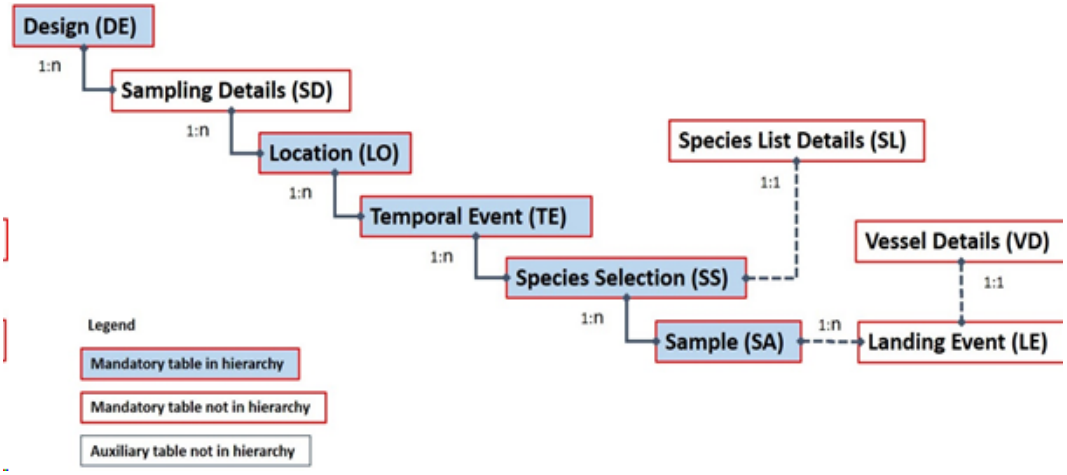


Figure 3.2. new2 hierarchy.

3.5 National Marine Fisheries Research Institute, Poland

Sampling programmes considered

At sea and on shore sampling programme for commercial fisheries in the Baltic Sea.

The programme consists of four sampling schemes: Demersal, Pelagic, Freshwater, Salmonids with a separate lists of PSUs for each quarter. Each sampling scheme is stratified by targeted stocks. All species and catch categories are covered by sampling.

During the workshop data from at-sea sampling from 2017 will be used.

The identified hierarchy for at-sea sampling is hierarchy 1. The vessels are selected with UPSWR.

Progress made

For at-sea sampling, hierarchy no. 1 seems to be the most appropriate one. All upper hierarchy tables have been successfully populated with 2017 test data.

List of tables populated during the workshop:

- Design;
- Sampling details;
- Vessel selection;
- Vessel details;
- Fishing trip;
- Fishing operation;
- Species selection;
- Species list details;
- Sample.

In the case of onshore sampling it was confirmed during the discussion that Poland may apply hierarchy no. 1 which was originally designed for at-sea sampling. Aggregation level "trip" has to be entered in the Fishing Operation table, as the total number of hauls might not be known at the time of collecting the sample.

Length frequency and biological data tables have not been populated yet, but no major issues are anticipated. Lower hierarchies A (length stratified biological samples), B (only length frequency data) and D (no length or biological data) were identified in the data collected in the sampling programme. In the national database records containing biological data are linked directly to length measurements.

Issues identified

Although hierarchy no.1 was identified as the most appropriate, hierarchy no. 2 is still taken into consideration because the true primary sampling unit is a trip of a vessel. The list of the upcoming trips is not available in advance but it might be assumed that all trips of a particular vessel have equal probability of being selected for sampling.

During at-sea observer trips, the information on all fish species is collected in Poland. However, it is unclear how the Species List Details table should be populated. At least three options are considered:

- all marine species (without listing of species);
- all marine species listed in the National Workplan for Data Collection;
- all marine species that occurred in the samples in the previous years.

Currently, one common list of vessels selected for sampling is used for both at-sea and onshore selection of PSUs. Once the vessel owner is contacted, he is first asked about a possibility to take

observer on-board. If he refuses, then he is asked for a possibility to take a sample from the landings.

Some vessels are excluded from a sampling frame.

Changes that are required either in national database or the RDBES

If hierarchy no. 1 is going to be applied, the stratification of PSU's in the national programme needs to be updated in a way that a vessel belongs to only one strata. Currently, vessels can belong to more than one strata because it was assumed that the PSU type is a trip of a vessel.

3.6 DTU Aqua, Denmark

Sampling programmes considered - all commercial sampling programs since 2000, see Table 3.2.

Table 2.2. Overview of Danish sampling schemes

No	Year_data_collection	Sampling_Scheme_ID	Sampling_Scheme_type	Comments	RDBES_hierarchy
1	1995–1999	DNK-observer	At-sea	This period is from the start of the Danish at-sea observer program. Need to find documentation of the program, but assuming same set up as in 2002–2009	1
2	2000–2001	DNK-observer	At-sea	Same as for the period 1997–1999	1
3	2002–2009	DNK-observer	At-sea	Able to find EU national program documenting some of the sampling design. Vessels are selected for sampling based on expert knowledge, so these are representative for the fisheries. Experts are people from DTU Aqua and the industry.	1
4	2010–2013	DNK-observer	At-sea	Changed to a formal design with probability based selection of the PSU. The sampling frame is a vessel list, but each vessel is weighted accordingly to number of trip the year before.	2
5	2014–2018	DNK-observer	At-sea	Still a formal design with probabilistic selection of the PSU, but each vessels is only present in one strata. The sampling frame is a vessel list without any weighting.	1 | 2
6	2019	DNK-observer	At-sea	Same as before, but vessels without logbooks are included	1 | 2
7	2009–2015	Gillnets	At-Sea	Reference fleet for sampling gillnetters. (This hierarchy is a bit strange. We have a reference fleet of vessels. Then we select a day and sample all of them (if fishing). On one of the vessels an observer will be present on the others it will be self-	1

No	Year_data_collection	Sampling_Scheme_ID	Sampling_Scheme_type	Comments	RDBES_hierarchy
				sampling. Try to fit this into one of the existing hierarchies)	
8	2016–2018	Gillnets	At-Sea	Ad-hoc sampling scheme. Kind of continuing the reference fleets for gillnetters (2011–2016)	1
9	2009–2019	LLD_ANA_0_0_0	At-sea	Ad-hoc sampling scheme	2
10	1979–2001	DNK-auctions-HUC	Onshore	We have harbour samples in our national database back to 1979. Unsure about the design	?
11	2002–2014	DNK-auctions-HUC	Onshore	This sampling scheme do not really have a design - it is ad-hoc quota sampling, where the purpose it to get a specific number of individual from a stock each year. Each year we sampled a specific auction and the main group of auction are remarkably similar throughout the period.	new hierarchy needed - suggested hierarchy 9, Species/size category focused
12	2015–2018	DNK-auctions-HUC	Onshore	Changed to a formal design and probability based selection of the PSU. All selected auctions are sampled (census) and the selection of the SSU (harbour day) is systematic. The group of auctions are selected for each quarter based on the same quarter the year before. A stock is only sampled in an auction if the auction belongs to the group accounting for 80% of the landings, trips and value of that stock, so if the stock is encountered in another auction, then it is not sampled	New hierarchy needed - suggested hierarchy 9, Species/size category focused
13	1979–2018	DNK-factories	Onshore	Sampling of unsorted landings for reduction and human consumption by DTU Aqua. The PSU is the factories and then the staff samples a fishing trip for us.	New hierarchy needed (very similar to hierarchy 6, but it goes location and the fishing trip)
14	1979–2018	DNK-factories	Onshore	Sampling of unsorted landings for reduction by the Danish control.	2
15	2011–2018	DNK-sprat	At-sea	Self-sampling of sprat at-sea. The fishers sample on a voluntary basis.	1
16	2001–2018	DNK-sandeel (TBM)	At-sea	Self-sampling of sandeel at-sea. The fishers are obliged to sample hauls on all trips targeting sandeel. DTU Aqua subsamples afterwards.	2

No	Year_data_collection	Sampling_Scheme_ID	Sampling_Scheme_type	Comments	RDBES_hierarchy
17	2004–2014	DNK-sandeel (FAST)	Onshore	Real-time monitoring of the sandeel fishery. Samples taken by the Danish control	2

Progress made

Identification of upper hierarchies:

- Most of the Danish sampling schemes 2000–2018 fit one of the hierarchies.
- Two new hierarchies are needed.
- Some of the old one are a bit difficult to fit, since it is unclear how the samples have been selected.

Population of the format:

Full population of the data model with data from at-sea sampling schemes 2016 (see Table 3.2 no 5, 8 and 9). Sampling scheme no 5 has been put in both hierarchy 1 and 2 to test the difference.

Issues identified

Vessel details (VD):

Why is VDtype mandatory?

Fishing Operation (FO):

All Fishing Operations within an observed Fishing Trip are recorded in the Danish at-sea sampling schemes, but with different sampling protocols.

FOtotal: All valid Fishing Operations. The same as FTfoNum

FOsampled: Here we have a problem, since the Fishing Operations are sampled with different sampling protocols. On some FO's both discard and landings are sampled, FOcatchReg = "All", on others only landings are sampled, FOcatchReg = "Lan", and on others again none is sampled, FOcatchReg = "Non". So FOsampled depends on what you want to estimate e.g. if discard then only FOcatchReg = "All" would be included. If landings are to be estimated then both "All" and "Lan" would be included. Here I have only filled FOsampled, when everything is worked up.

FOselectMeth: The observers should as a minimum have one FO fully worked up per day (FOcatchReg = "All"). The number depends on trip length. This is considered systematic (SYSS). FOcatchReg = "Lan" will at the moment be considered ad-hoc (NPAH).

Species selection (SS) and Species List Details (SL):

In the Danish at-sea sampling schemes 1-3 the observers should record all species observed. This is not the case, since e.g. the phylum Echinodermata and the class Gastropoda are never or very rarely recorded in the database. Some species are mostly recorded at the class level e.g. Cephalopoda.

This needs to be discussed with the observers, so only species normally recorded and at the level normally recorded are in the species list.

Ideas:

1. A species list based on what we normally observe (last 10 years). One for the following regions 27.4, 27.3.a, 27.3.22-24 and 27.3.2532. Not sure where to set the limits, so that will require some scrutinizing and expert help.
2. Start conservative - only include the certain e.g. commercial species. Work on the rest. Be careful when handling species, where some observers identify on class and some on species level. The latter need to be taken up to the general level.

Frequency Measure

In Denmark we often weigh the fish per length class. These weights are used to estimate mean weight at length | age when estimating catch (number) at length/age. At the moment we squeeze these weights into the CA table by making a row per fish. This is not correct, since it is not single fish measurements, so it is suggested to add the following fields to Frequency Measure (FM): Additional measurement, Additional measurement type and Additional measurement unit

Subsampling for biological variables - not length stratified

Problem:

In Denmark we often sample small pelagic the following way; we have a sample of herring from a fishing trip (here 1150 kg) where we take a subsample of (here 29.957 kg, equal to 100 fish). All these fish are length measured. Only 25 of the fish are aged. The subsampling is not length stratified but an unstratified subsample. We sample the fish in a systematic way, so the fish are order by length and then each 4th fish is selected in this case. This won't fit into hierarchy A, since there is no length stratification involved in the subsampling

Solution:

Utilizing the possibility to add a subsample in the sample table. The first sample (for the representative length measurements) will have Lower Hierarchy = B. The subsample thereof (for biological measurements) will link to the sample via ParentSampleID and have Lower Hierarchy = C.

Changes that are required either in national database or the RDBES

National database

Currently data are not stored by sampling scheme, but under national projects relating to labs. This will be changed in future, since it will make reporting and documentation much easier. This is a minor change.

We may include the sampling frames and non-responses in future, these are at the moment stored in excel sheets outside. Species Lists may also be included in future.

Inclusion of an additional measurement in Frequency Measure (FM)

In Denmark we often weight the fish per length class. These weights are used to estimate mean weight at length | age when estimating catch (number) at length/age. At the moment we squeeze these weights into the CA table by making a row per fish. This is not correct, since it is not single fish measurements, so it is suggested to add the following fields to Frequency Measure (FM), Additional measurement, Additional measurement type and Additional measurement unit.

New Hierarchies in the RDBES

A new hierarchy is needed for the Danish sampling at auctions, see suggested hierarchy 9, called Species/size category focused elsewhere in this report. The new hierarchy will cover the Danish auction program back to at least 2002, probably earlier.

A new hierarchy is needed for the Danish sampling of unsorted landings at factories. It is very similar to hierarchy 6, but Onshore Event should be Location.

3.7 Institute for Agricultural and Fisheries Research (ILVO), Belgium

Sampling programmes considered

*Observer at sea:

- TBB_DEF_>221 kW fleet segment
  - <2015 and >2017 non-probability based sampling (quota sampling)
  - <2018 and >2014 probability based sampling (SRSWR)
- TBB_DEF_<=221 kW fleet segment
  - 2004 – now: non-probability based sampling (quota sampling)

*Onshore sampling: <2011

During the WKRDB-POP, we only focused on the observer at sea data

Progress made

Both lower and upper hierarchy were identified for the at sea sampling programmes.

- Upper hierarchy is 1 for probability based sampling (SRSWR) and 2 for non-probability based sampling (quota sampling)
- Lower hierarchy is A or B or D when the non-probability based sampling (quota sampling) as well as when the probability based sampling (SRSWR) took place.

During the workshop we focused on hierarchy 2 (non-probability based sampling - quota sampling) as the most recent at sea sampling data (2018) are also obtained through quota sampling (for the TBB_DEF_>221 kW fleet segment as well as for the TBB_DEF_<=221 kW fleet segment). The tables mentioned below were populated (using an R-script) during the workshop for one observed trip from the TBB_DEF_>221 kW fleet segment:

- Sampling Design (DE)
- Sampling Details (SD)
- Vessel Selection (VS)
- Fishing Trip (TR)
- Fishing Operation (FO)

Issues identified

In the past hierarchy 2 (design class A) was identified for the probability based observer at sea sampling (SRSWR) programme as the most appropriate. But after a short description of the sampling design (see below) and a group discussion during the WKRDB-POP, it was decided that hierarchy 1 (design class B) is considered to be more appropriate than hierarchy 2 for cases like ours.

Short description of the probability based observer at sea sampling (SRSWR) programme

The TBB_DEF_>221 kW fleet segment comprises beam trawl vessels with a capacity of more than 221 kW, operating in North Sea, the English Channel, the Irish Sea, the Celtic Sea, South of Ireland and the inner part of the Bay of Biscay. The TBB_DEF_>221 kW trip duration is on average 8-10 days and one trip can cover several areas. Selecting a vessel from the TBB_DEF_>221 kW fleet segment for sampling is done by a random draw from a vessel list (with replacement) whenever an observer was available. Only the vessels that are willing to take observers onboard and those that are suited, from a logistic point of view, to have an observer onboard are included in the vessel list (sampling frame). Throughout the entire year more or less the same vessel list is used. The vessel owner of the selected vessel is contacted and is asked whether or not an observer was allowed to be onboard during the first planned trip. Non–responses and refusals are documented. An observer onboard will sample every second haul (systematic sampling) so sampling takes place around the clock to reflect typical working conditions.

Changes that are required either in national database or the RDBES

So far, no changes are needed either in the national database or in the RDBES in order to be able to populate the RDBES tables. Of course it needs to be noted that during the RDB-POP, not all tables were populated for hierarchy 2 (quota sampling) and that we didn't focus on the hierarchy 1 case (probability based sampling - SRSWR) or on the old onshore sampling programme. So, potentially, in future, new issues might be encountered.

3.8 Institute of food safety, animal health and environment (BIOR), Latvia

Sampling programmes considered

For Baltic Sea five sampling programmes are selected:

- On board sampling for demersal fish trawl fishery;
- On board sampling and self-sampling for Pelagic trawl fishery in the open sea;
- Self-sampling for Pelagic trawl fishery in the Gulf of Riga;
- Self-sampling for trapnets in the Gulf of Riga coastal fishery;
- Onboard and self-sampling in the coastal fishery.

Progress made

During the workshop precise hierarchies for all sampling programmes were identified. For demersal and pelagic trawl fishery and coastal fishery Hierarchy no. 1 was selected. For coastal fishery, for special case then additional information about salmon and trout are collected, Hierarchy no. 2 were selected.

Three hierarchy tables were populated during the meeting (onboard sampling for demersal fish trawl fishery, coastal fishery data with Hierarchy no. 1 and Hierarchy no. 2. No special need to prepare test file for pelagic fishery data, the principles of how the data are collected and prepared is similar to demersal fishery data.

During the workshop understanding about lover hierarchy data tables were obtained. Method with samples numbering should be introduced.

Method of data conversation to the new lover hierarchy data tables format were tested.

For the coastal fishery is necessary pay more attention to data collection to be able to fill upper hierarchy tables.

Issues identified

None.

Changes that are required either in national database or the RDBES

At this moment national database are not used for RDB file preparation. It is not possible to extract whole trip info from the national database. RDB data files are prepared using MS Excel and Access.

3.9 Department of Fisheries and Marine Research Cyprus

Sampling programmes considered

Cyprus data collection framework includes Trawlers, Polyvalent vessels and Small-scale fishery. The following scenarios were considered:

- On-board and onshore sampling for Trawlers (OTB). Contrary to the designed scheme the reality falls into an ad hoc application. For onshore equal probability simple random sampling without replacement is taking place. PSU is the vessel with a systematic selection after the first random vessel pick and then an alteration between the two fishing vessels (that the fleet consists of) is taking place. Following that a systematic selection of sampling day at port is selected.
- Onshore sampling of SSF. Equal probability simple random sampling without replacement with PSU being the landing site. Then a random sampling for the day occurs. During the visit a census sampling occurs of all vessels from a priority species list. Most of the times though vessels and catch are not so many and a census is taking place on all species.
- On-board and onshore sampling for polyvalent vessels (operate mostly as LLS for Large pelagics but they can also use all SSF gears).

Progress made

Identify the hierarchy schemes that each of the sampling designs belongs to.

OTB -> For On board sampling the Hierarchy 1 will be selected were an ad hoc selection method occurs. On shore falls in Hierarchy 5 as a random systematic approach of Vessel selection and port day.

SSF-> Currently only on shore sampling is taking place for this fleet as the vessels are generally small and they cannot easily accommodate observers. A new Hierarchy Aka No 9 will be needed based on Hierarchy 5 to account for the second stage of random day selection.

LLS -> Hierarchy 1 for on board sampling is proposed as a starting point for the current implementation. In future once it is possible to apply the designed scheme on the field the Hierarchy 3 will be more appropriate. For on shore sampling an attempt with hierarchy 5 is proposed.

During the meeting priority was given to SSF case for which it was recognized from the beginning the need for developing a new Hierarchy and it was necessary to conclude and justify this case with the help of RDBES Core Team members. During the course of the meeting it was possible to configure the tables design and populate them with actual samples.

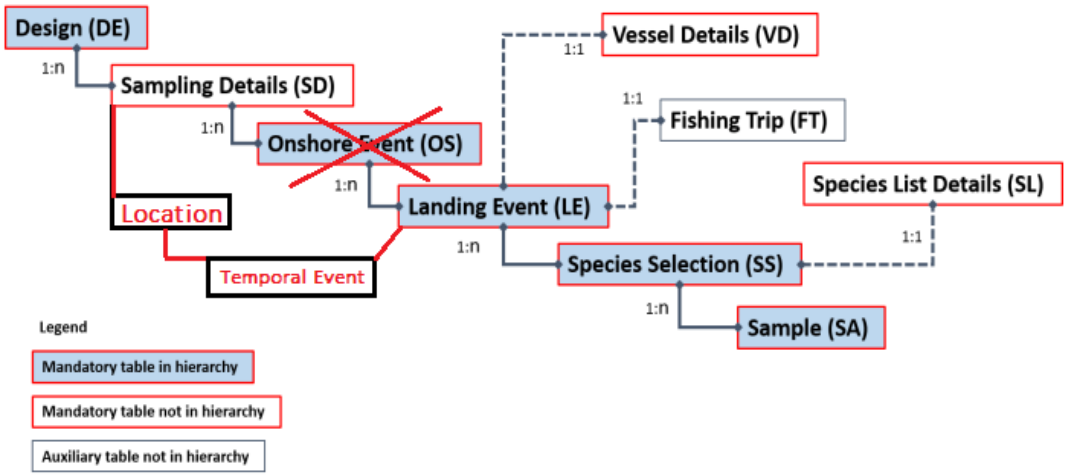
Issues identified

Regarding SSF the Hierarchy 5 looked to be the ideal combination but it appears that it cannot be adequately descriptive of the sampling activity. In Cyprus case a random selection of the ports (without replacement) occurs every two weeks. Six ports (current staff can accommodate three sampling days per week) are selected out of 16. In next stage a random selection (without replacement) of the sampling days occurs (work days and work hours only) for each of the weeks. Alternative ports and days are also recorded (from the draw) in case weather, logistical and/or technical issues are not allowing the programmed activity to happen.

Regarding the lower Hierarchies and biological sampling an issue occurs in Hierarchy A which implies stratified selection of the biological sample. In CYP case samples are not always abundant and most of the time a census approach is applied in order to reach species measurement quotas. The method proposed during the workshop (Kirsten, Denmark) looks that it fixes this issue.

Changes that are required either in national database or the RDBES

Proposed diagrammatic representation of what would Hierarchy 9 need to be is given based on Hierarchy 5 diagram. Attempt to populate an example with this configuration was positive at the end.



In order to achieve the above mentioned diagrammatic flow it was necessary to add a LOid column in Temporal Event table to connect it with the port selection. A TEid column was then needed in Landing Event table to connect the sequence.

3.10 Estonian Marine Institute

Sampling programmes considered

Baltic Sea:

- Onshore sampling from trawlers (herring/sprat)
- Onshore sampling from coastal fishery
- Onshore sampling from processing industry

Progress made

During the workshop Hierarchies identified:

- for onshore sampling from trawls and coastal fishery Upper hierarchy 8 and Lower hierarchy A
- for onshore sampling from processing industry Upper hierarchy 7 and Lower hierarchy A

Hand-made example table for Hierarchy 8 was filled with data from SD28 (Q1).

Also partly filled example table Lower hierarchy A.

Some problems were encountered in filling Lower Hierarchy A with table BV but the problems were solved during the meeting

Upper hierarchy 7 was not examined during the meeting

Issues identified

None.

Changes that are required either in national database or the RDBES

Vessels and landing data from national database are not accessible to Estonian Marine Institute (EMI) currently because the database is stationed in Ministry of Rural Affairs and database currently not available for our institute. The information on vessels and landings will be available only to the beginning of the next year.

3.11 Klaipedos University, Lithuania

Sampling programmes considered

The programme consists of three sampling schemes:

"Self-sampling "on shore for pelagic fisheries. The fisher simply randomly take a half box of fish. Sampling program fit to Hierarchy 4. PSU is a landing day. SSU is a trip. The selected fishing trip is stratifying by quarter and subdivision. The lower hierarchy was indicated as A (based on data 2016–2017)

Observers at sea for discard sampling in demersal fishery vessels particularly for the Baltic Sea region. Some or all fishing operations are sampled where some could be aggregate from haul and some from the whole trip. PSU is a fishing trip. The selected fishing trip is stratifying by quarter and subdivision. Sampling program is fit to Hierarchy 1. Lower hierarchies A or B (based on data 2016–2017)

Observers on shore for demersal fisheries. Sampling program is fit to Hierarchy 4. PSU is a landing day. The selected fishing trip is stratifying by quarter and subdivision. The lower hierarchies A, B and C (based on data 2016–2017).

Progress made

Tested RDBES data model for onshore sampling data. In the case of onshore sampling it was confirmed to be applied Hierarchy 4. All datasets were successfully completed for that Hierarchy:

- Design;
- Sampling details;
- Onshore;
- Fishing trip;
- Species selection;
- Species list details;
- Sample.

Established the 40 species list for sampling discards in Baltic Sea fisheries, mostly freshwater species in coastal zone. There are only six main species for landings.

Almost have finished R script to convert data from national database format to RDBES format.

Issues identified

Only some minor issues were found out, which clarified at the Workshop.

However, historical data will not be able to upload properly: national data stratification until 2010 is not in line with Hierarchies and 2011–2014 data does not contain vessel's details.
Changes that are required either in national database or the RDBES

Lack of information concerning unsampled hauls in the national database. Also need to add for SA table presentation; the condition in which the sample was presented

3.12 AZTI, Spain

Sampling programmes considered

Three different types of sampling schemes will be tested:

- on shore sampling scheme;
- on board sampling scheme;
- Revising the draft document made in the project RECOLAPE about sampling of large pelagic and how to accommodate collected data to the RDBES.

Progress made

- Identify hierarchy for on shore sampling. All our ICES onshore sampling have hierarchy 5;
- Identify hierarchy for on board sampling. All our ICES onboard sampling have hierarchy 2. This could be discussed as we select vessel*week in a matrix, and then the trips within that week and vessel. But I think it is a reasonable approach;
- Go through the variables and solve doubts about the meaning of some variables and how to populate the tables;
- Revise the document on tropical tuna sampling and how to accommodate it to the RDBES
- Start populating mandatory tables for hierarchy 5 (all onshore sampling schemes). populated tables: DE, SD, LE, SS.
- To do: finish populating the tables.

Issues identified

- *Tropical tuna sampling:* Several issues were identified that will need further discussion by the core group and also with the people directly involved in this fishery. The main issue

was to decide whether a new hierarchy is needed, or the sampling design can be fit in the existing ones assuming that sampling wells in a port*day may be as sampling boxes of mixed trips (similar to Danish port sampling).

- *Mixed species*: two options were discussed to include the data in the Sample Table when species are landed together with the same commercial name (i.e. megrims, monkfish):
  - a) Record the individual species as SAspeciesCode (i.e. MON and ANK) using SAcommercialSpecies to indicate that they were landed together (i.e. as MNZ). At the moment it is not clear whether the variable SAspeciesCode correspond to individual species or to mixed species, this could be clarified in the documentation
  - b) Record in one line the data for the mixed species (i.e. MNX) and then use Parent id to record the data of the subsample, indicating the species composition (ANK and MON)
- *Sampling subsamples*: In the cases where in on board sampling, a subsample is taken from the basket to measure for length (i.e. if there is a large amount of small pelagics), then the subsample can be recorded in the same table, using the parent Id.
- *Lower hierarchies*: we sample biological variables independently from length. Usually for one landing event we have either length or biological sampling, but it could happen that we have both. The solution proposed was to register two lines in the sample table, one line for each hierarchy (B &C)
- *Species list for concurrent sampling*:
  - *onshore sampling*: our protocol states that we should sample everything that is landed. A solution would be to use broad categories in the species list. It may be useful to agree on broad groups that can be used for countries in the same situation.
  - *onboard sampling*: Need to check how to do it for onboard sampling: how do we deal with invertebrates, benthic fauna, litter… More categories may be needed
- *Selection of SU:*
  - selection of port*days: systematic selection with a random start, in matrix of port*days. Two issues:
    - In practice we have deviations from the selection in the matrix. What shall we put in the selection method? SRSWOR/ NPEJ/other?
    - how does this systematic selection fit in the proposed new hierarchies where you first select a port/vessel and then a day/week?
- *selection of vessels (SSF*): there is a protocol to randomize (first vessel every X minutes), but in practice we know it is not random. What shall we put in the selection method? SRSWOR/ NPEJ/other?

Changes that are required either in national database or the RDBES

1. A solution to accommodate tropical tuna data is needed and it may require changes in the RDBES
2. During the WK we have discussed many issues and solved many doubts. I think that some of this doubts could be used to improve the documentation:

   - *Use of parent Id*: At the moment we were told not to fill in the KeyId (shaded in the data model tables). But then it is not clear to me how this ParentId be recorded in the tables to make it clear how it relates to the rest of the tables
   - *Auxiliary tables:* Tables that can be both mandatory and auxiliary tables depending on the hierarchy: is it needed to fill in the stratification and total/sampled units when they work as auxiliary tables? It was recommended not to fill in that information, but it is not clear in the documentation
   - *Samples out of your sampling frame*: It can happen that you end up having some samples that were not initially in your sampling frame. It was suggested to use an extra

strata "Strata offdraw". Guidelines would be needed to explain how to fill in total units and selection procedures in these cases

- *Sample table:*
  - variable SAtotal: does it refer to the number of commercial species (which may be mixed in some cases: megrims, lophius.) or to the number of individual species already split
  - strata should be defined for each species, or for each combination of species and commercial category? It is not clear because apart from the strata, there is a column for commercial category (SAcommCat). I think this is important that this is clear to fill in the SAtotal and SAsampled variables
- *Landing event table:*
  - LEhaulNumber: refers to the total number of hauls?
  - LEfullTripAvailable: does this variable refers to the trip or to the landing event? It is confusing.
  - LEsequenceNumber: would this be a kind of LandingEventId?

3.13 IEO, Spain

Sampling programmes considered

Onshore and on-board length sampling programs for landings and discards in the ICES area considered. This includes all length sampling schemes under current DCF/EUMAP National Work Plan.

1) Onshore schemes:

PSU: port/day. SSU type: Landing event. PSU selection is quasi systematic. SSU selection is random.

Species/Stocks covered: Concurrent sampling.

Temporal strata: Quarter with systematic monthly allocation.

Unique exception: Scheme IEO_M_VR_GC. Small longline fishery targeting blackspot sea bream in the Gulf of Cadiz selected for sampling because of national interest. Fishery concentrated in just one species and one port. PSU: port. SSU: day. TSU: species category.

Species/Stocks covered: *Pagellus bogaraveo*.

Temporal strata: Quarter with systematic monthly allocation.

Table 3.3. IEO Onshore schemes:

IEO_M_BC_AC	Bottom otter trawl targeting demersal fish in Divisions 8a, 8.b, and 8.d (OTB_DEF_>=70_0_0)
IEO_M_BC_PN	Bottom otter trawl targeting demersal fish in 9.a North, Portuguese waters (OTB_DEF_>=55_0_0)
IEO_M_BJ_CN	Bottom otter trawl targeting demersal and pelagic fish in North Iberian waters (Divison 8.c and North of 9.a) (OTB_DEF_>=55_0_0 and OTB_MPD_>=55_0_0)
IEO_M_BT_CN	Minor-gear fleet operating with gillnets in North Iberian waters (Division 8.c and North of 9.a) (GNS_DEF_60-79_0_0)
IEO_M_CR_CN	Purse-seine targeting small pelagic fish in North Iberian waters (Divison 8.c and North of 9.a) with access to Bay of Biscay waters (Divisions 8.a, 8.b, and 8.d) (PS_SPF_0_0_0)
IEO_M_NM_AC	Gillnet targeting hake in Subarea VII and Divisions 8.a, 8.b, and 8.d (GNS_DEF_120-219_0_0 and GNS_DEF_>=100_0_0)
IEO_M_LN_CN	Handline targeting mackerel in North Iberian waters (Division 8.c) (LHM_ SPF _0_0_0)
IEO_M_MR_AC	Bottom otter trawl targeting hake in Subarea 7 (OTB_DEF_100-119_0_0)
IEO_M_NS_CN	Minor-gear fleet operating with pots targeting Octopus vulgaris in North Iberian waters (Divison 8.c and North of 9.a) (FPO_MOL_0_0_0)
IEO_M_PL_AC	Bottom longline targeting demersal fish in Western Scotland, Western Ireland and Bay of Biscay fishing grounds (Subarea 7 and 8.a, 8.b, and 8.d) (LLS_DEF_0_0_0)
IEO_M_PL_CN	Bottom longline targeting demersal fish in North Iberian waters (Divison 8.c and North of 9.a) (LLS_DEF_0_0_0)
IEO_M_PR_CN	Bottom pair trawl targeting both pelagic and demersal fish in North Iberian waters (Divison 8.c and North of 9.a) (PTB_MPD_>=55_0_0)
IEO_M_RP_AC	Bottom otter trawl targeting megrims in Subarea 7 (OTB_DEF_70-99_0_0)
IEO_M_RS_CN	Gillnet targeting anglerfish in North Iberian waters (Divison 8.c and North of 9.a) (GNS_DEF_>=100_0_0)
IEO_M_SB_GC	Bottom longline targeting silver scabbard fish in the Gulf of Cadiz (Division 9.a South) (LLS_ DWS _0_0_0)
IEO_M_VL_CN	Gillnet targeting hake in North Iberian waters (Divison 8.c and North of 9.a) (GNS_DEF_80-99_0_0)
IEO_M_VR_GC	Longline targeting blackspot sea bream in the Gulf of Cadiz (Division 9.a South) (LHM_ DWS _0_0_0)

2) On-board schemes:

PSU: Vessel/Trip. SSU: Haul.

100% of vessels within the scheme are considered; refusals recorded.

Selection procedure of the PSU: Equal probability simple random selection with replacement from the official vessel list.

Unique exception: Scheme IEO_S_RP_AC. Ad-hoc selection of PSU. This scheme accounts for 10 PSU by year, the 4.5% of the total on-board sampling effort allocation.

Table 3.4. IEO On-board schemes

IEO_S_TB_CN	**Bottom otter and pair trawl in North Iberian waters (Divison 8.c and North of 9.a) (OTB_DEF_>=55_0_0, OTB_MPD_>=55_0_0 and PTB_MPD_>=55_0_0)**
IEO_S_BC_GC	Bottom otter trawl targeting both crustaceans and demersal fish in the Gulf of Cadiz (Division 9.a South) (OTB_MCD_>=55_0_0)
IEO_S_CR_GC	Purse-seine targeting small pelagic fish in the Gulf of Cadiz (Division 9.a South) (PS_SPF_0_0_0)
IEO_S_GN_CN	Gillnets targeting anglerfish and hake in North Iberian waters (Divison 8.c and North of 9.a) (GNS_DEF_>=100_0_0 and GNS_DEF_80-99_0_0)
IEO_S_RP_AC	Bottom otter trawl targeting megrims in Subarea 7 (OTB_DEF_70-99_0_0)

Progress made

Identification of the hierarchies for the sampling schemes:

Hierarchy proposed:

- Onshore schemes: **Hierarchy 5**
  - exception: IEO_M_VR_GC **Hierarchy 7**
- On-board schemes: **Hierarchy 1**

For the market sampling Hierarchy 5 a complete review of the tables were done. This include: Table Design, Table Sampling Details, Table Onshore Event, Table Landing Event, Table FishingTrip, Table Species Selection, Table Sample, Table FrequencyMeasure and auxiliary tables for VesselDetails and SpeciesListDetatails. Review of all fields done during the week thus allowing a clarification of the required data.

Issues identified

Completion of "Table Sample" and the way mixed species have to be uploaded to the database subject to more than one criteria. This affects the use of SAstratum, SAspeciesCode and SAcommercialSpecies. May be this could be further clarified with examples.

Changes that are required either in national database or the RDBES

Evaluation of new required fields needed. Preparation of the RDBES format in a first stage will be done based on current database output (RDB format currently generated by IEO database and required extra information). Evaluation of permanent changes to include new required fields in the database to be done after generation of new format completely clarified. The reason is to avoid unnecessary and potentially conflicting changes in the IEO database, which is a multipurpose database used by the all the Fisheries Programs run by the IEO (ICES area, Mediterranean, CECAF, Long distance, etc.).

3.14 Cefas, UK (England and Wales)

Sampling programmes considered

Onshore and offshore catch sampling programmes were tested on the RDBES data model.

Progress made

1. Offshore programme

Hierarchy 1 was identified to be the appropriate to this sampling programme. All Upper hierarchy tables:

1. Design table;
2. Sampling details;
3. Vessel selection;
4. Vessel details:
5. Fishing trip;
6. Fishing operation;
7. Species selection;
8. Sample.

The lower hierarchies for the offshore programme were: Hierarchy A for sample where species are measured and stratified by length for otolith collection the species; Hierarchy B for samples where species are measured, but no biological information is collected and Hierarchy D for species that were observed. However, the tables were not populated

The list of species still needs to be defined.

2. Onshore programme

The onshore programme is more complex. Regional strata are sampled differently – PSU is port day but how that is selected varies between regions.

Hierarchy 5 may be sufficient to cover all onshore programmes, but we need to use clusters and unequal probability. One region for the demersal programme was covered in the week. Sampling in the Southwest was used as a case study. Days of the week in one stratum are sampled every week with unequal probability. Two gear groups are identified and on a selected day a landing is selected from one gear group every time and one randomly 50% of the time. Each gear group has 3 species lists with one or 2 selections allocated to each. For some landings of a selected species the box sampled may need to be subsampled to save time using a throw rate (e.g. measure 1 throw 1).

TABLE: Design, TABLE: SamplingDetails, TABLE: OnshoreEvent, TABLE: LandingEvent, TABLE: SpeciesSelection, TABLE: Sample, TABLE: BiologicalVariable, TABLE: Frequency Measure were completed with real data for one sampling event, one trip and one species length sample with age data.

This still needs to be reviewed by one of the core group.

The lower hierarchies for the onshore programme were: Hierarchy A for samples where species are measured and stratified by length for otolith collection; Hierarchy B for samples where species are measured, but no biological information is collected.

Issues identified

1. Onshore

Strata as above; Species selection where multiple species groups are sampled at different rates

Multiple programmes (Demersal; Crustacean; Pelagic) Within these programmes stratifications are sampled following different sampling hierarchies or Programmes are sampled together e.g. Random day within a week for a single port for Demersal programme in one strata vs. bus route within a selected week.

- First port - Demersal
- Second port - Demersal and Crustacean
- Third port - same
- Fourth port - Crustacean

Could be considered as different sampling programmes but need to maintain a link for reporting.

2. Offshore

Offdraw selections;

Adhoc selection of age data throughout trip;

Drawlists based on previous years activity - how do you account for immigration and emigration?

Sample table (SA) - Weight Live field is mandatory. We estimate weight. There is no reference if this is estimated or collected weight.

Species lists – is it necessary to have a reference list for species for each stratum within a sampling frame when the protocol is for example to sample or observer ALL finfish and commercial shellfish? Commercial shellfish might need to be defined. It could be an administrative burden keeping these lists up to date and relevant to each sampling scheme and year.

Changes that are required either in national database or the RDBES

Currently none of our sampling databases have a reference to the sampling design. Sampling probabilities are collected for onshore and offshore vessel/trip selections but these are stored remotely from the sample data. They can be linked. An observerDB re-write; production of an observer activity management DB; OCR upload of onshore sample selection data will help improve the connectivity.

3.15 Ifremer, France

Sampling programmes considered

Length sampling program for landings and discards will be considered. Two sampling programs are active.

The French at-sea sampling program (named OBSMER) seems to follow a hierarchy 3, where month is the primary sampling unit, and vessel the second sampling unit. In this case a set of vessels is sampled every month. The pool of vessels is selected according to the port, the métier, the group of ICES area and the species or group of species they are targeting.

The French onshore sampling program (names OBSVENTE) seems to follow hierarchy 8, where month is the primary sampling unit, and vessel the second sampling unit. In this case a set of vessels is sampled every month. The pool of vessel is selected according to the port, the métier, the group of ICES area and the species or group of species they are targeting.

In the national data compilation, this sampling program becomes a hierarchy 3 thanks to the use of national landings to identify the trips related to each onshore event.

For the biological parameters, the sampling plan is *ad-hoc*: it includes onshore and scientific campaign (in separated sampling "scheme"), with quarter as psu (sometimes), and target individuals weights, size, maturity and age by species in subarea or group of subarea, without any consideration regarding the *métier*.

The selection of vessel is random with equal probability in theory, month are sampled sequentially. ⅓ of the trips are sampled in sequence for the at-set sampling program. Difference in species selection for landings and discards. Sampling for the size category onshore. To be completed and clarified.

Test : few trips from OBSMER and OBSVENTE (from different places) will be used to test the hierarchy 3 and 8 (and others if needed...). A test to assess to utility of having the onshore sampling program converted in hierarchy 3 will be done.

Progress made

Import from RDB to RDBES for the hierarchy 3 seems to be possible using extra data (the landings national population data to build the vessel list among other). Improvement of the related messy R script.

Issues identified

None for now.

Changes that are required either in national database or the RDBES

None for now.

3.16 IPMA, Portugal

Sampling programmes considered

The national data collection (PNAB/DCF), includes the following sampling schemes for the ICES area 27.9.a:

1. At-sea sampling (concurrent) - A vessel list is used to select a vessel and then a fishing trip to sample in that vessel. Three different sampling protocols (static gears, mobile gears and purse-seiners) are used for the onboard sampling but the same procedure is adopted until getting to the trip level. In the at-sea sampling, lengths are recorded at haul level and biological variables may also be collected for a group of species, if present in the discarded fraction.
2. Onshore sampling (concurrent) – Auction*day is the PSU and trip the SSU. Concurrent sampling is performed at trip level where all the species and size categories are to be sampled for lengths.
3. Onshore sampling (species focus) – Auction*day is the PSU and Size Categories the SSU. A box (sample) is taken from each SC to do the length measurements.
4. Biological sampling (species focus) – Auction*day is the PSU and Size Categories the SSU. A commercial sample is obtained specifically for the biological sampling. Vessel information is recorded and landing event may or may not be recorded.

Progress made

- Identification of the hierarchies for each of the sampling programmes.

1. At-sea sampling: Upper Hierarchy 1 and Lower Hierarchies A or B

2. Onshore sampling, concurrent: Upper Hierarchy 6 and Lower Hierarchy B
3. Onshore sampling, species focus: New upper hierarchy 9? Lower Hierarchy A or B
4. Biological sampling, species focus: New upper hierarchy 9 or 7? Lower hierarchy A

- Production of the R code for population Hierarchy 1 tables was performed during and completed after the WK. The at-sea sampling data from mobile gear (bottom otter trawl for demersal fish - OTB_DEF) in 2017 was used.
- List of tables populated during the workshop:
  - Design,
  - Sampling details,
  - Vessel selection,
  - Vessel details,
  - Fishing trip,
  - Fishing operation,
  - Species selection,
  - Species list details,
  - Sample.
- List of tables populated after the workshop:
  - Frequency measure,
  - Biological variable

Description of the population procedure and assumptions

DE table

- Considered the at-sea sampling for OTB_DEF fleet in 2017.

VD table

- The vessel length segmentation used at national level is the same one adopted for the DCF (vessel lengths: [0-10[, [10-12[, [12-18[, [18-24[, [24-40[).

VS table

- Populated with the number of vessels sampled and the total vessels in fleet for the same vessel length category. Information on fleet is not derived from national sampling database. External data has to be prepared for this purpose.
- The vessel list used for making the contacts includes mainly the cooperative vessels, so the selection method considered for this exercise is the 'UPSWR', but further discussion on this is needed.

FT table

- Populated with the number of trips sampled and the total trips in fleet for the same vessel length category. This information is not derived from national sampling database. External data has to be prepared for this purpose.

FO table

- All hauls from each trip were recorded in this table because there is also important information collected for the unsampled hauls (e.g. fishing duration, species landed weight)
- Sampled hauls have 'All' assigned for the 'FOcatReg' and unsampled hauls have only 'Lan'.

SL table

- The list includes all species that can be caught in Portuguese ICES area 27.9.a: if any new other species is caught, there will be the need to validation and possibility of updating the list
  - Some species/groups present in the sample but with no aphia code were removed from the list to populate RDBES
  - Information on litter, biological waste (e.g. damaged fish not identified) and sediment (e.g. rocks, mud, shells) that are always weighted and recorded onboard were removed from the list used to populate the RDBES.

SS table

- The SLid not filled because in our point of view it only makes sense if groups of species are populated in the Species List table. In the case of discriminating all species this SLid should not be mandatory in this table.

SA table

- Sex: some species are sexed during measuring but they may not be weighted by sex. So, information provided in the table does not discriminate information by sex. The alternative would be to calculate the weights per sex, by ignoring the recorded weight for the total, or by redistributing it? Still need to see how to do.

FM table

- Type of length is species-specific and is now starting to be recorded along with the other biological data in the database. For this exercise all recordings were considered 'total length' but in future there will be the possibility of discriminating it better.

Issues identified during the population of the tables

In general there were difficulties in understanding some of the field definitions that resulted in more time spent to prepare the code

VS table

- Each row should give information on the number of vessels total and sampled by sampling scheme. I don't understand why the VDid links in this table. VDid present in this table means that each row will contain vessels that are present in the following VD table? If we put information by VDid here, the total and sampled vessels will be the same in all rows.

FO table

- Put all hauls performed in the trip, as we collect general information (e.g. spatio temporal and landings composition) in the unsampled hauls. Some of that information will be needed for the estimation procedure and is not present in the trip levels (e.g. fishing duration in the trip = sum fishing duration from each haul).
- If only sampled hauls are to be in this table, then there is the need to have a new column with the fishing duration in the FT table.

SL table

- Missing codes for litter, sediment (e.g. stones, rocks), and biological waste (e.g. fish bones, damaged species not identified). They are discriminated and weighted when present in the sample, which should be representative of the total catch.

- RDBES should accommodate all species codes considered by MS: missing codes in RDBES can lead to some assumptions for the 're-classification' of that species, that will not reflect what was really sampled and will not match the national databases...

SS table

- Each row should give information on the number of species total and sampled in each catch category per fishing operation. The SLid only makes sense if groups of species are populated in the Species List table. In the case of discriminating all species this SLid should not be mandatory in this table. In our case, SLid present in this table would means that each row would contain the same species that are present in the SL table.

SA table

- In FO table refers to 'All' in 'FOcatReg' for sampled hauls. When in SA table that 'All' catch is separated into 'Lan' and 'Dis'. Some doubts on whether these fractions should be considered as strata or not.
- Sometimes species are recorded as species at sample level and as genus at haul level because fisher mix more than one species in boxes recorded as landings (e.g. measured ANK and landings refer to ANF). A decision has to be made for the population of this information in the table due to possible duplication of data. Maybe those species should not be raised at haul level when they don't match between what was sampled and recorded as landed? Or should we do it anyway? But then it should be highlighted somehow.

Changes that are required either in national database or the RDBES

RDBES – Need to include fishing duration at trip level (FT table) to be used as auxiliary variable in discards estimation, in case the unsampled hauls are not present in the FO table.

Other issues identified prior and during the workshop:

- Onshore sampling for species focus – national sampling doesn't fit in any of the hierarchies previously proposed in the RDBES data model. A new hierarchy proposal during the workshop (H9) seem to accommodate this type of sampling scheme. But still it needs to be tested.
- Biological sampling for species focus – during the workshop and after a deeper analysis some doubts between Hierarchy 9 and Hierarchy 7: need to see what makes more sense when trying to populate the tables.

3.17 Azores, Portugal

Sampling programmes considered

Sampling programmes for landings and discards will be considered: at-sea sampling programme for discards and onshore landings sampling programme.

Progress made

As no historical (and present) data from Azores are available, a sample set of at-sea sampling data for handlines targeting cephalopods was used. Hierarchy 1 was identified as the proper one and hierarchy B as lower hierarchy was considered.

Issues identified

None.

Changes that are required either in national database or the RDBES

None yet identified.

3.18 Marine Institute, Ireland

Sampling programmes considered

From 2017 Q3 the MI's demersal at-sea catch sampling programme was changed to a 4S programme. Vessels that fall within the sampling frame (Irish registered vessels >10m length using the gear types OTB, SSC, GNS and TBB and with target assemblages DEF and CRU) are stratified temporally (year and quarter) and spatially (based on which ICES areas the majority of their fishing activity occurred in the same quarter in the previous year) – there are then 3 vessel lists per quarter (vessels mostly fishing in 6, mostly fishing in 7a and mostly fishing in 7b-k). Random selections are then made from these lists and sampling coordinators try and contact the selected vessels to arrange trips for at-sea observers to sample. Vessels are selected with unequal probability, based on their length and the number of trips they have previously made. No clustering or subsampling is used.

From 2017 the Marine Institute's port sampling programme operated by defining a list of ports to be sampled and then setting semi-random quarterly targets for the number of sampling events at each port. Non-random quarterly sampling targets are also defined for the stocks of interest. The sampling frame consists of landings by Irish registered vessels at Irish ports. The samplers use expert judgment to fulfil the targets. No clustering or subsampling is used.

Progress made

Hierarchy 1 has been identified as the correct hierarchy for the at-sea catch sampling programme. Initially it was thought that Hierarchy 3 would be the correct one to use and a full set of data were populated for that. However after discussion it was agreed that Hierarchy 1 would be most appropriate - the extract code will need to be adapted to account for this change of hierarchy.

Hierarchy 5 has been identified for the port sampling programme – this is because the hierarchical sampling involves first sampling port-days within a quarter, then the landing events on those days, then taking length and biological sample from those landings. The code to extract the data has not yet been developed for this.

Issues identified

A careful consideration of where to place the variables indicating quarterly stratification is required to avoid the situation where a sampling unit is placed into 2 different strata (this is not allowed).

Changes that are required either in national database or the RDBES

None identified so far.

4 Identify and document any problems in converting national data formats to the RDBES format. (ToR c)

In skype meetings held prior to WKRDBES the core development group of RDBES has continued to develop the data model towards v1.17. Some of the issues addressed in those skype meetings concern new developments that were not examined during WKRDB-POP (e.g. specific needs from WGBYC); others overlap with the issues experienced and reported by the WK participants, having found possibly distinct solutions. An effort is now needed to compile the suggestions and different needs towards a definitive (and public) v1.17 of the RDBES data model.

Upper hierarchy summary

Participants were asked to identify which hierarchies they would use for their sampling programmes - this is summarized below.

Table 4.1. Upper hierarchy summary.

Upper Hierarchy	Short description	Institute count
1	Select vessel, then trips	14
2	Select trips	6
3	Select time, then vessels	2
4	Onshore, then trip, then landing event	2
5	Onshore, then landing event	7
6	Onshore, then trip	3
7	Onshore	4
8	Select time, then vessel, then landing	3

It can be seen that there is a big difference in the popularity of the different hierarchies, with the selection of vessels being a very common design pattern. These differences largely reflect national differences in sampling but can also result from ambiguities in the definition of sampling schemes.

New hierarchies proposed

The following new hierarchies were proposed during the workshop.

Table 4.2. New hierarchies proposed

Short description	Institute count
Tropical tuna	1
Species/size category focused	3
Select location, then time, then landing event	1
Select vessel, then time, then trip	1

The RDBES Core Group will evaluate these new hierarchies and include them in the data model if appropriate.

New fields proposed

FM - add a weight field to allow people to record the total live weight in grammes of the length class

SA - add a field with the subsample level e.g. the initial sample would be level 0, the first level of subsampling would be level 1etc.

Documentation

Some fields were repeatedly misunderstood (e.g. the difference between the SAtotalWeight-Measured and SAtotalWeightLive fields in the SA table) which indicates that the explanation for these fields should be clarified in the documentation.

Species selection

Considerable time was spent discussing the species selection table and its implications both for data recording in the RDBES and estimation from its data. A very diverse use of concurrent sampling and species lists has been reported by e.g. WKISCON2. In brief, some sampling programmes do not use species lists at all, sampling all species in the catch/catchCategory; but it is not infrequent sampling programmes only sample a subset of species or even carry out species-focused sampling (the same single species is sampled in each occasion).

The SS table allows the declaration of species selection lists and selection methods, being fundamental to the generation of the true 0s and NAs during estimation and hence to the correct estimation of all species-related RDBES outputs. A preliminary R-script testing the capabilities of the SS table to generate correct NAs and 0s generation had been carried prior to the workshop and was shared with participants for further validation and testing. The fundamentals of this script lay the basis for some of the work that will be carried by WKRDB-EST in what concerns this particular table.

Additional discussions were held on the need (or not) to record all species present in the markets when complex species selection schemes, involving some kind of selection of species are employed. The answer to this question probably differs depending on the variable that is being estimated (e.g. frequency of occurrence, total volume/numbers, lengths and other biological information). A preliminary R-script was done and shared among participants with code for simulations of different species selection scenarios.

The overall preliminary conclusion is that the variables currently present in the RDBES SS table appear to fit the main needs of the vast majority of sampling schemes (that either sample concurrently or carry out a census of the species from a species list). This result will likely be further evaluated during upcoming WKRDB-EST.

With regards to more complex and rare situations such as those involving random selection of species from those species present in, for example, a market*day, significantly more complex estimators will be involved, with results being currently inconclusive with regards to the sufficiency of variables incorporated in v1.16 of the RDBES.

Most importantly, a strong need to document historical/current species lists used in the different sampling programmes was identified. If for main target species their inclusion/exclusion from historical lists is relatively easy to determine, a much more complex situation occurs when bycatch (e.g. cephalopods, marine birds and mammals, invertebrates) and taxonomic resolution starts being considered. The brief discussion among participants confirmed that criteria for sampling/recording such cases may have changed over-time and most likely differed among observers; also, that methods suggested like frequency of occurrence of samples are not always useful to reconstruct previous sampling practices, particularly if species are rare. Given the significance of this topic for the quality of estimates used by some ICES EGs (WGCEPH, WGBYC etc.) and the probable need to combine different sources of data (commercial sampling, survey data) in the identification of which species can effectively be assumed to have been consistently recorded across the time-series, it is envisioned that a specific workshop will be the most promising means to devise common guidelines and secure coordination of species list prior to data provision of bycatches.

Commercial Species

Clearer guidance needs to be given on how to record samples of mixed commercial species.

Litter data

Thought needs to be given as to how to include litter data that is used in estimation for assessments. ICES has a EG dealing directly with marine litter (WGML) and litter data in general could also be included in the RDBES although this is beyond its current scope.

Conclusion

Participants from 20 different institutes based in 17 countries participated in the hands-on workshop at ICES. It was to be expected that this workshop would generate some new requirements for the RDBES data model since it was the first opportunity for many institutes to actually try populating the data model with their national data. Although issues have been identified it is not thought that any of them are serious impediments to moving forward with the RDBES data model. The RDBES Core Group will look at the results of this workshop and either respond to individual questions or adapt the data model and documentation as required.

The workshop chairs will keep in contact with the participants to keep them up-to-date with any changes and answer questions. As mentioned earlier in this report this workshop (WKRDB-POP) and WKRDB-EST should be considered as two halves of the same process so the chairs will also encourage the participants to carry on working with their data so that institutes can use their own data in WKRDB-EST.

Annex 1: List of participants

Name	Country	email
David Currie (Chair)	Ireland	David.Currie@Marine.ie
Edvin Fuglebakk (Chair)	Norway	Edvin.Fuglebakk@hi.no
Ain Lankov	Estonia	ainparnu@gmail.com
Ana Claudia Fernandes	Portugal	acfernandes@ipma.pt
Ana Ribeiro Santos	UK (England and Wales)	ana.ribeirosantos@cefas.co.uk
Christina Petterson	Sweden	christina.e.pettersson@slu.se
Dália CC. Reis	Portugal	Dalia.CC.Reis@azores.gov.pt
Ioannis Thasitis	Cyprus	ithasitis@dfmr.moa.gov.cy
Irina Jakovleva	Lithuania	irina.jakovleva@zuv.lt
Ivo Šics	Latvia	ivo.sics@bior.lv
Jon Elson	UK (England and Wales)	jon.elson@cefas.co.uk
Jose Rodrigues	Spain	jose.rodriguez@ieo.es
Jukka Pönni	Finland	Jukka.Ponni@luke.fi
Julia Wischnewski	Germany	julia.wischnewski@thuenen.de
Karolina Molla Gazi	Netherlands	karolina.mollagazi@wur.nl
Kirsten Birch Hakansson	Denmark	kih@aqua.dtu.dk
Laurent Dubroca	France	laurent.dubroca@ifremer.fr
Liz Clarke *	UK (Scotland)	L.Clarke@MARLAB.AC.UK
Lucia Zarauz	Spain	lzarauz@azti.es
Maciej Adamowicz	Poland	madamowicz@mir.gdynia.pl
Marijus Špegys	Lithuania	marijus.spegys@apc.ku.lt
Marta Suska	Poland	msuska@mir.gdynia.pl
Nuno Prista *	Sweden	nuno.prista@slu.se
Pedro Lino *	Portugal	plino@ipma.pt
Petri Sarvamaa	Finland	Petri.Sarvamaa@luke.fi
Sofie Nimmegeers	Belgium	Sofie.Nimmegeers@ilvo.vlaanderen.be
Sofie Vandemaele	Belgium	sofie.vandemaele@ilvo.vlaanderen.be

Sven Stötera	Germany	sven.stoetera@thuenen.de
Twan Leijzer	Netherlands	twan.leijzer@wur.nl

*** Part time via WebEx**

Annex 2: Resolution

2018/2/ACOM41 The **Workshop on Populating the RDBES data model (WKRDB-POP)** co-chaired by David Currie, Ireland and Edvin Fuglebakk, Norway, will meet in ICES HQ, Copenhagen, 18–22 February 2019 to:

a) Describe and explain the RDBES data model to national data submitters using worked examples.

b) Provide hands-on guidance and assistance in converting data to the RDBES data format for national data submitters.

c) Identify and document any problems in converting national data formats to the RDBES format.

WKRDB-POP will report by 30 April 2019 for the attention of ACOM

Supporting information

Priority	The activities of this workshop will promote the development of the Regional Database and Estimation System, RDBES. This workshop will help countries to correctly convert their national data formats to the RDBES format. The RDBES when it is implemented works as a database for the Baltic Sea, North Sea and Eastern Arctic, and North Atlantic Regional Coordination Groups (RCGs). The RDBES will also function as a database and estimation system for ICES Fisheries Advice. The development will concentrate on harmonization, quality assuring, documentation, approved estimation methods and transparency. Consequently, these activities are considered to have a very high priority.
Scientific justification	The RDBES will be extensively used by the RCGs and ICES both to store detailed fisheries sample data and use it for estimation - therefore it is essential that national data submitters are familiar with the RDBES format and confident in correctly converting their national data to this format. **ToR a) – Describe and Explain the Data Model** The RDBES data format will be explained using its documentation, and a number of worked examples. These worked examples will play an important role in illustrating the types of decisions that data submitters will need to make. **ToR b) – Provide hands-on guidance and assistance** This is the most important part of the workshop and will occupy the majority of the workshop's time - it will entail the RDBES Core Group providing practical assistance to the attendees. The workshop attendees must be familiar with their own national sampling programme designs, and must have made preparations necessary to provide real datasets of their national samples to the workshop. The Core Group will then help them to convert their data to the new RDBES format. The more work that attendees have done in trying to populate the RDBES format with their own data before the workshop the more value they will gain from this work. When new questions are identified and resolved they can be added to the RDBES "Frequently Asked Questions" so that other people can benefit from the answers. **ToR c) – Identify problems**

	If it is not clear how particular data should be converted to the RDBES format then this will be recorded for future discussion and resolution.
Resource requirements	The two co-chairs, and potentially the rest of the 5 active members of the RDBES Development Support Core Group will be requested to participate as hands-on instructors/demonstrators.
Participants	~20 people excluding the Core Group.
Secretariat facilities	ICES HQ meeting room and facilities
Financial	No financial implications.
Linkages to advisory committees	There are no direct linkages with the advisory committees, but there is a link to WGCATCH and PGDATA and most of the stock assessment Working Groups will benefit from the development of the RDBES.
Linkages to other committees or groups	The group and the development of the RDBES will support and benefit from the fishPi2 project.
Linkages to other organizations	The RDBES will support the work done by the RCGs under the European Commission, EC. The aim is also make the RDBES support the countries in providing data for the data calls under the EC.

Annex 3: Agenda

During *Hands-on* sessions, participants will work with mapping their own data to the RDBES data model with guidance from the core-group members.

During *Evaluation* sessions, we will have plenary discussions on issues identified.

Participants are encouraged to complete the hands-on sessions in the order presented in the agenda. We will facilitate that participants can progress at different pace, and will reschedule the evaluation sessions if needed.

Monday 18 February

14:00	Agenda and presentation of participants	chairs	
14:30	Introduction to the RDBES Questions and discussion	Henrik Kjems-Niel-sen	ToR a)
15:00	Upper hierarchies in the RDBES	David Currie	ToR a)
15:30	Short case-study: IRL At-Sea Sampling	David Currie	ToR a)
16:00	Coffee break		
16:30	Compile overview of datasets: sampling programs, countries, stocks		ToR c)
17:30	Break for the day		

Tuesday 19 February

09:00	Recap of Day 1 Questions and discussion	Chairs	
10:00	*Hands-on:* identifying sampling hierarchies and RDBES tables		ToR b)
11:00	Coffee break		
11:30	*Hands-on* continued *Evaluation*: RDBES hierarchies		ToR c)
13:00	Lunch		
14:00	Design based estimation in the RDBES Questions and discussion	Edvin Fuglebakk	ToR a)

15:00	Coffee break		
15:30	*Hands-on:* Fill fields in upper hierarchies		ToR b)
17:00	Break for the day		

Wednesday 20 February

09:00	Summary of progress by participants		
10:00	*Hands-on* continued		ToR b)
11:00	Coffee break		
11:30	*Hands-on* continued *Evaluation:* RDBES fields in upper hierarchies		ToR c)
13:00	Lunch		
14:00	True zeroes in the RDBES Questions and discussion	Nuno Prista	ToR a)
14:45	RDBES Upload/Validation tool Demo	Henrik Kjems-Nielsen	
15:30	Coffee break		
16:00	*Hands-on:* specify species selection tables		ToR b)
17:30	Break for the day		

19:30 Social dinner: RizRaz - Kompagnistræde 20 (https://goo.gl/maps/zp6UmpqZTQy)

Thursday 21 February

09:00	*Hands-on:* fill in species selection and sample table		ToR b)
10:30	Coffee break		
11:00	*Hands-on:* continued		ToR b)
12:00	*Plenary:* species selection and sample table		ToR c)

13:00	Lunch		
14:00	Biological parameters in the RDBES Questions and discussion	Kirsten Birch Håkansson	ToR a)
15:00	Coffee break		
15:30	*Hands-on:* fill in biological parameters		ToR c)
16:30	*Evaluation:* lower hierarchies and biological parameters		
17:00	Break for the day		

Friday 22 February

09:00	Compile *dataset report*	ToR c)
10:30	Coffee break	
10:45	Presentation of some dataset reports: CYP, FIN, NLD	ToR c)
12:30	Discussion to summarize common issues, identified changes needed in national sampling programs or databases	ToR c)
13:30	End of Workshop	

Dataset reports: At the end of the workshop we would like each participant to contribute a short report for inclusion in the workshop report. This should contain:

- a description of how far the mapping exercise progressed
- a list of all mapping-issues identified
- a description of any changes that needs to be considered for national sampling programs or databases in order to populate the RDBES model.

We hope that some participants will be willing to present their reports at the workshop before the plenary discussion on Friday.

In addition, all participants are encouraged to document the mapping between their national database and the current version of the RDBES, in such a manner that it can be developed into automatic data conversion routines at their institutions. Participants that would like to prototype data conversion scripts as part of the mapping exercise, are encouraged to do so.