ICES SC–RDB REPORT 2016

Report of the Steering Committee of the Regional Database FishFrame (SCRDB)

29 November – 1 December 2016

ICES Headquarters, Denmark

ICES
CIEM

International Council for
the Exploration of the Sea

Conseil International pour
l'Exploration de la Mer

Contents

Executive summary

Until now the SCRDB has consisted of representatives from the RCM Baltic, RCM North Sea and Eastern Arctic, RCM North Atlantic and the ICES secretariat and meets once annually. Non-EU ICES members (Iceland, Norway, Canada, Russia and USA) and the European Commission were invited to participate in the SCRDB 2016 meeting. This SCRDB 2016 meeting was held at the ICES HQ in Copenhagen 29 November – 1 December 2016.

The RDB is the main prerequisite for development of regional sampling programmes, for standardisation of data, and the tool for ensuring transparency and quality assurance of input data for stock assessment, and for the management of the marine living resources by the EU and non-EU countries in the North Eastern Atlantic area.

The EU member states participating in the RCM Baltic, RCM NS&EA and RCM NA uploaded data in the Regional DataBase (RDB) as a response to a data call launched by the RCM chairs in June 2016. The data call covered landing, effort and sampling data for 2015. Upload of previous years' data (2009-2014) were also requested if these data had changed substantially. The overall response of the data call was satisfactory, even though a few of the MS are still in progress with uploading their complete data sets.

The aim of the annual data calls is to facilitate development of analyses for regional sampling strategies, to solve previous problems related to the upload and also advance in the development of the RDB, in order to build a tool for regional coordination where time series of regional 2009–2015 data would form the base for planning of future data collection.

Since 2012 the RDB has been hosted and maintained by ICES under agreement with the European Commission (MoU). In addition, since 2014, ICES have provided funds for the further development of the RDB. Unfortunately, no funding for further development had been made available through the call for tenders procedures by the EU Commission. As the RDB is an essential tool for the ICES advisory work, ICES has therefore secured 1 Million DKK from the ICES Council to run a two year development project for improving the RDB to a statistical sampling and raising platform, but also to ensure the different end users/uses are accounted for.

The SCRDB membership setup with RCM representatives, the EU Commission, and the ICES secretariat has worked well: however, there are now a number of changes that require that the membership is revisited. The funding of the RDB is split between the EU (via an MoU with DGMARE) and ICES own funding. ICES member countries include non EU-DCF countries that are relevant to the regional scientific stock assessments, and therefore to the development and support of the RDB as the main channel for making commercial fisheries data available for not only the RCM end use, but that of ICES in fisheries management advice. Therefore, it is suggested that non-EU countries should be actively involved with full membership of the SCRDB. In addition, it has been recommended to the long distance and large pelagic RCM's that they should use the RDB, and therefore should be included in the development process.

It has been agreed that a new RDB will be developed over the next two years by the ICES secretariat – this is an ambitious time-scale and is driven by the funding supplied by the ICES Council. At the end of the two years the new RDB should be capa-

ble of both storing data and generating estimates using accepted statistical methods. It is likely that different countries/groups will still need to use InterCatch for a transitional period after the new RDB completion. The long term aim is for the new RDB to completely replace InterCatch. It should be highlighted that the development of the new RDB within such a tight timeframe requires a dedicated support group providing 'instant' support throughout the process. This core RDB development support group is established under the current steering committee and will reflect on suggestions and proposals crucial to the process of developing and building the new RDB.

The RDB Data Policy Document was reviewed at the SCRDB 2016 meeting. The SCRDB does not at this point want to change the present Data Policy. The SCRDB agreed that in the future the data policy should be more open. The new invited non-EU countries will go through the Data Policy to see if they can support the Data Policy as it is, or, if additions are needed in connection with them uploading data to the RDB.

1   Introduction

The Steering Committee for the Regional Database FishFrame (SCRDB) has until now been responsible for strategic planning, technical governance, operational issues and estimates of costs in the overall governance of the regional database (RDB) (Figure 1.1).
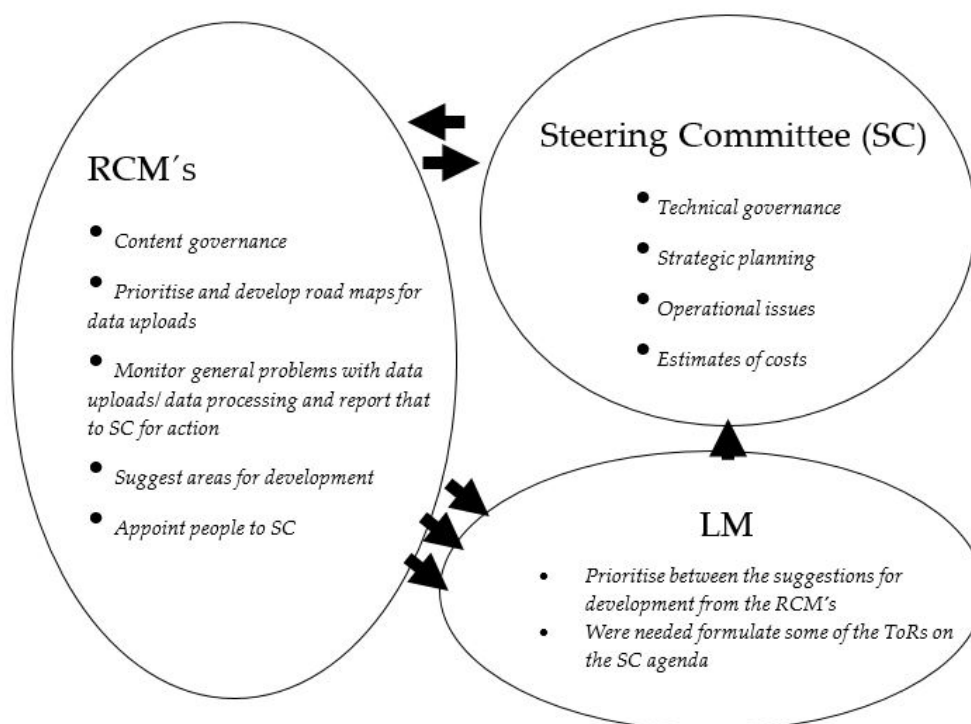
Figure 1.1 The RDB; tasks for and interactions between The Regional Coordination Meetings, the Liaison Meeting and the RDB Steering Committee.

The SCRDB interacts with the Regional Coordination Meetings (RCMs) and Liaison Meeting (LM) on other tasks such as development needs and content governance. Until now the SCRDB has consisted of representatives from the RCM Baltic, RCM North Sea and Eastern Arctic, RCM North Atlantic and the ICES secretariat and meet once annually. Non-EU ICES members (Iceland, Norway, Canada, Russia and USA) and the European Commission were invited to participate in the SCRDB 2016 meeting. This SCRDB 2016 meeting was held at the ICES HQ in Copenhagen, 29 November – 1 December 2016. The Terms of reference are given below. The agenda and list of participants can be found in Annexes 1 and 2.

The **Steering Committee for the Regional Database FishFrame (SCRDB)**, chaired by Jørgen Dalskov (Denmark) met 29 November – 1 December 2016 in Copenhagen (ICES HQ), Denmark, to:

a)   Conclude on membership of the steering committee.
b)   Respond to recommendations put forward to the SC-RDB by the Liaison Meeting and ICES expert groups.
c)   Summarize how the RDB has been used in the regional coordination meetings;
d)   Based on requests from the RCG's consider and conclude on possible revision of the data policy document, dealing with access rights, data confidentiality and

data ownership issues, following the consultation process, amend if necessary and adopt the final document.

e)   Based on the findings made during the FishPi project and the requests from the RCGs consider and conclude on a new exchange format and suggest a route forward.

f)   Continue to develop a strategy based on revised DCF and EU-MAP regulations, on development of the RDB, taking requirements from a statistical sound sampling and raising and the landing obligation into account. Report on progress for the short, medium and long term plans developed so far.

2   Background

Regional coordination of the data collection underpinning assessment of marine re-sources and fisheries are primary handled by the five Regional Coordination Meet-ings (RCMs). The RCM setup is initiated in accordance with the EU Data Collection Framework. Non-EU member countries are invited to participate equally at the RCM meeting. These meetings take place every year to review past sampling and to lay down the rules for sampling coordination for the next year in the region. The aim of the meeting is to achieve adequate international sampling coverage, task sharing and cost efficiency.

The work of the RCMs has not been easy, partly because of the complexity of data collection, but also because no central source of data at a regional level has been available to perform the analysis necessary for optimization of the sampling schemes and quality of the data collected. In every case it has been necessary to request data from each country in the region in order to carry out basic analyses, which are neces-sary for coordination. This process is error prone and also time consuming both for the national institutes and the actual meetings of the RCM. This is also reflected in several of the recommendations from the RCMs. This situation has led several RCMs to express a strong need for a Regional Database (RDB) as a data source and tool for their work.

A RDB would also facilitate transmission of data to end-users from a national insti-tute perspective, where work effort can be saved, as well as from an end-user per-spective, where more transparency on the compilation and quality of the data could be achieved. Potential end-users that will benefit from an RDB are thereby all groups which want to make use of tabulations, analyses and graphic presentation of fishery information across countries within a region.

Following a recommendation from the Liaison Meeting in 2009 the European Com-mission organized the workshop "Regional scenarios and Roadmap on Regional Da-tabase" in 2010 (Anon., 2010). A strong need for a regional database (containing biological and transversal data but also VMS data) was expressed by participants from the Baltic and North Sea regions. For the North Atlantic region the opinions were divided. Participants from some Member States saw the possibility to improve the quality of data and data management through a RDB while other considered the present situation with national databases satisfactory and saw a risk with increased workload. The Workshop held in 2010 recommended the development of a roadmap on a regional level to be addressed by the different RCMs giving each region the abil-ity to act on different scenario options. The RCMs (Baltic, North Atlantic and North Sea & Eastern Arctic) responded in their meetings during 2010.

Since then all the three RCMs considered that a database with "disaggregated" (sam-pling data in detailed form and transversal data in a low aggregated form) data would fulfil most of the needs of the RCM. Such database would facilitate analyses on a regional scale and it gives Member States a tool to establish regional data collec-tion programmes and to coordinate their work plans. Also, in order to be able to re-ply to data requests and transfer data routinely to end-users, it would be more cost efficient to use an RDB and it would provide better quality standards compared to the present situation. In the Baltic region MS had already used a regional "disaggre-gated" database for several years. This database, FishFrame, was developed for this purpose. The experiences with FishFrame were positive and the RCM Baltic decided

in 2009 to keep using FishFrame in future. During the RCM North Sea and Eastern Artic 2010 meeting, FishFrame was adopted as platform also in this region.

In 2010, the RCM Baltic and the RCM North Sea and Eastern Artic recommended an interim steering group to be set up with clear terms of references and mandates in order to start the implementation of a RDB including a Steering Committee (SC). The RCM North Atlantic proposed items to be discussed in such a SC. The 7th Liaison meeting endorsed this recommendation. As a consequence, an interim steering group consisting of representatives from the three RCMs, ICES Secretariat and the Commission was put together. This steering group had a meeting in February 2011 in order to elaborate on a governance model for the RDB but also to suggest road maps on how to proceed towards implementation of a RDB from a content point of view as well as from a technical point of view. The outcome of the interim steering group was adopted by the RCMs which also appointed participants to the RDB Steering Committee (SCRDB). The first SCRDB meeting was held in December 2011. Non-EU ICES member countries are invited to participate on both the RCM and in the SCRDB.

3 New financial support for development and maintenance of the RDB

The RDB has been hosted and maintained by ICES since 2012 under agreement with the European Commission (MoU), in addition ICES have provided funds for the further development of the RDB. Based on this funding model the following has been achieved:

Maintenance:

- Operational system, actively used and maintained
- Regional standardization of codes and quality control of input data
- Helpdesk for data providers and users
- Delivery of regional data for the three Regional Coordination Meetings (RCM's) committed to its use
- Agreed data provision (data call) and data access (data policy)
- Main work platform of the three RCM's with all respective countries uploading data to the RDB

Development:

- Further standardisation of codes and quality control of input data, improvements of uploads, and report outputs
- Support for the new landing categories (i.e., landings above and below the minimum conservation size) resulting from the EU landings obligation

3.1 The shared vision for the RDB

The shared vision for the RDB is among other things:

- Reduce the workload for the countries in estimating and providing data, as the RDB would contain (or can utilise from R libraries) all needed methods
- Ensure quality assured standardised statistical methods (expert driven) are used for estimating the data for the stock assessment
- Provide a commercial catch data processing platform for all ICES countries (to avoid an EU and non-EU system for ICES stock assessments)
- Describe and document data quality by using common quality checks across all countries' data
- Reduce the workload for countries submitting data by utilising data, and estimates, from the RDB, as appropriate, to meet end user needs.
- Support the Regional Coordination Groups/Meetings with data and reports for their work
- Data are encapsulated within the RDB (the data is safeguarded in the RDB and the end user understands every change to the data)
- Leverage the body of work already existing in R code projects and developed further by the experts
- Links to other databases e.g. the VMS/Logbook database used by WGSFD, ByCatch regulation, Fisheries independent data (i.e. DATRAS)

3.2 Future development

ICES has secured 1 Million DKK from ICES Council to run a two year development project for improving the RDB to a statistical sampling and raising platform, but also to ensure the different end users/uses are accounted for. It is therefore necessary to

have a RDB Project Expert Group, who has the knowledge of how to generate statistical estimates, and of the kind of outputs that would be needed from the RDB (i.e. for RCM sample design and planning, stock assessments and EU wide analysis). The work would be carried out through workshops (WK funding via the national EC contributions) and intersessionally.

It is anticipated that EC countries, and other ICES member countries, participate.

ToRs:

1) Follow, and advise on the development of the project
2) Provide substantial input to the user requirement specifications, including:
3)
   a) The drafting of a requirement specification document.
   b) Specify data exchange format,
   c) Define user roles, processing of data, data checks, methods for estimation, output.
4) Be responsive to the project team in providing input to issues in the implementation of the RDB.
5) Testing and approval of developments

If the user requirement specification document is not finished by end Feb. 2017 then there should be a workshop finalising it.

Review in March 2017 the progress in RDB participant countries in populating the data exchange formats.

1st Workshop (end March start April 2017, but might be postponed due to ICES assessment WG commitments): Write the user requirement specifications of the new RDB regarding: Data exchange format, checks, user roles, processing of data, methods for raising and output.

The core support group of the RDB Project Expert Group is suggested to consist of:

Kirsten Birch Håkansson, DTU Aqua, Denmark   kibi@aqua.dtu.dk

Nuno Prista, SLU Aqua, Sweden   nuno.prista@slu.se

Alastair Pout, Marine Scotland Science, Scotland   A.Pout@MARLAB.AC.UK

Norbert Billet, IFREMER, France   norbert.billet@ird.fr

Peter van der Kamp, Wageningen Marine Research, Netherlands
                                                    peter.vanderkamp@wur.nl

David Currie, Marine Institute, Ireland   David.Currie@Marine.ie

Sub-group members: (10-14). People with primary good knowledge of statistical raising but also knowledge of national data is good.

4   Membership of the SCRDB steering Committee

4.1   Background

Currently the membership of the Steering Committee for the RDB (SCRDB) is comprised:

- Nominated experts from the RCMs/RCG's using the RDB
- The host of the RDB (ICES)
- The EU Commission

This has worked well, however there are now a number of changes that require that the membership is reconsidered. Firstly, the funding of the RDB is split between the EU (via an MoU with DGMARE) and ICES own funding (3 development grants in 2014, 2015, 2016). ICES member countries include non EU-DCF countries that are relevant to the regional scientific stock assessments, and therefore to the development and support of the RDB as the main channel for making commercial fisheries data available for not only the RCM end use, but that of ICES in fisheries management advice. Secondly, to avoid parallel systems for the provision of data and estimation procedures for countries within and outside the EU but sharing stocks/stock assessment advice. In both cases, non-EU countries should be actively involved with full membership. Thirdly, the long distance and large pelagic RCG's have been recommended to use the RDB, and therefore should be included in the development process.

4.2   New configuration

From 2017, to ensure a broader participation and applicability of the RDB, but also recognizing the obligation EU countries have under the DCF, the current SCRDB recommends the following setup for approval by the RCG's.

- The RCG-RDB-Steering Committee will also include RCG's that are not actively submitting data, but interested to follow the development.
- The RCG-RDB-Steering Committee will invite non-EU countries as observers (primarily Norway, Iceland, Faroe islands, Greenland and Russia).
- A new group under the ICES system, the ICES RDB Steering Group, will be created with similar ToR's and the same chair as the RCG-RDB-Steering Committee. Non-EU countries that are part of ICES will have full membership of this group.
- The two groups will meet sequentially (2 days for RCG-RDB-Steering Committee and 0.5 days for ICES RDB Steering Group) in order that all can follow the discussion and proposed decisions, but allow the EU and non-EU countries to have a separate reporting line for the outcomes.
- This will allow some flexibility in addressing the needs of the RCG's as end users, and ICES as an end user.
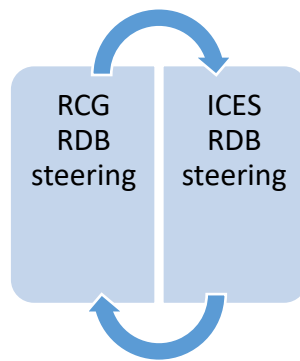
4.3 Future outlook

The structure will be reviewed and the possibility to merge these 2 groups under one ICES group would be considered.

5   Responds to recommendations put forward to the SCRDB by the Liaison Meeting and ICES expert groups

In 2016 the Regional Coordination meetings recommended actions to be taken in relations to the RDB. These recommendations have been reviewed by the Liaison meeting 2016 (LM 2016). The following recommendations have been endorsed by LM 2016.

LM 10. RCM-LDF 2016 -2 and RCM LP 2016: Inclusion of RCM LDF data into Regional Database	
RCM-LDF 2016 RCM LP 2016 Recommendation	RCMs LDF and LP, in line with the expressed interest from the Commission to set up a data base system holding all relevant fisheries data in the RCMs LDF and LP regions, recommends to set up a database system serving the RCMs LDF and LP needs. These needs are to be combined with the need to improve data transmission to end-users of data collected through the DCF and coordinated by those RCMs. A strong preference of both RCMs is to build upon currently established systems to ensure short term progress and minimize development needs, thus being cost-efficient.
	The time spent before the meetings and during the meeting correcting data, merely highlight the importance of automating and standardizing the data transmission and its previous checking. Tools for data format conformity checks, such as those developed under project DGMARE 2014/19, can be considered a first step towards this standardization.
	However the RCM LP agrees that a Regional Database would facilitate even more the work of the Data Provider and would be a crucial step in the process of standardization and data exchange. Thus, RCM LP recommends expanding the scope of the Regional Data Base FISHFRAME, hosted by ICES, to include EU Large Pelagic fisheries data. This would imply the inclusion of specific fields and codes of interest for the LP fisheries.
	RCMs LDF and LP recommend the inclusion of their areas of competence into the Regional Database Fishframe as currently hosted by ICES and governed by the RCMs through a Steering Committee. This would imply the inclusion of areas outside the EU waters as well as the inclusion of the relevant species and metiers involved.
Follow-up actions needed	RCMs LDF's and LP's chairs to contact the Steering Committee for the Regional Database Fishframe. Upon acceptance, both RCMs to provide representation and support to the Steering Committee.
	ICES to facilitate the inclusion of data from the RCMs LDF

	and LP areas.
Responsible persons for follow-up actions	Chair of RCMs LDF and LP, SC-RDB, ICES, Commission
Time frame (Deadline)	Prior to the 2017 data call
LM Comments	LM combined this with a similar recommendation from the RCM LP and endorses it.
Follow up from SCRDB	ICES secretariat will make it possible for countries to upload data on species and from areas under the remit of RCM LP and RCM LDF. This will require that the chairs of the RCM LP and RCM LDF check the reference tables in the regional database under tools "Lookup Manager", then "View/Edit Tables" and press the button "Download [table name]" and report additional needs in accordance with existing types and possibility (harbours, metiers, areas and species that are missing in the lists) to the secretariat.
	The RDB-SC was informed that RCM LP and RCM LDF primarily need a place to store data and that resources for estimation procedures presently are not need.
	The RDB-SC advice countries and institutes involved in the data collection under the remit of RCM LP and RCM LDF to participate in the testing of the new data model / formats (see SharePoint RDB-SC 2016) that will be undertaken in January-February 2016 and further in the pan-regional data group that were initiated by RCM NA 2016.

RCM NA 6: UPLOAD LOGS	
RCM NA 2016 Recommendation	The RCM NA strongly recommends that: 1. Upload Logs continue to form part of the data call for the RDB and require submission by all MS whether they are able to upload data or not. 2. *An Intersessional Group is given sufficient time to review Upload Logs to qualify the data, report and distribute actions or advice to MS, RCMs and the RDB administrator to ensure the quality of data on the RDB is maintained and improved and end-users are aware of the limits to the data.*
Justification	The development of the RDB is an ongoing process which has to be enhanced based on user's feedback. There are still inconsistencies and errors in the data on the RDB that have been caused by the IT system design itself, by non-restrictive reference lists or due to insufficient data checks by MS. Data gaps and limit the potential for data analysis and delays RDB use in the regional coordination process. The completion of an Upload log was included in the data calls for the RCM 2015 and 2016 and was to be completed so that users can assess the limitations of the data and therefore what interpretations or analysis can be done with it. The RDB will be developed to record the status of the data within it, but until this feature is available a standard log submitted at the time of each data call can provide RCGs and data users with a simple reference to what data is not on the system as well as what is. Intersessional work is required to review the issues listed in the upload logs, qualify them and identify actions and the responsibly body (RCM, RDB administrator and SC-RDB, or MS) that can deal with them. This intersessional work is pan regional.

Follow-up actions needed	1. RCM Chairs to appoint pan regional contributors for intersessional work.
	 2. Intersessional group to report on 2016 upload logs
	      a. Identify a clear communications strategy
	      b. Identify actions and responsible bodies
	      c. Communicate these actions
	      d. Collate responses
	 3. RCM chairs to include an updated upload log in data call 2017 and, when relevant ask MS to consider reload their data.

Responsible persons for follow-	RCM Chairs, Intersessional group, MS, RDB administrator

up actions	
Time frame (Deadline)	RCMs 2017
Follow up from SCRDB	RDB-SC recognizes that this presently is a task for the RCM chairs. Feedback on the RDB from users through summaries from upload logs is highly appreciated and will be dealt with, if summaries are provided, in the 2017 meeting. Requirements or actions might however need to be prioritized given the available amount of resources (incl. human resources)

RCM NA 8: Data analysis subgroup	
RCM NA 2016 *Recommendation*	RCM NA recommends that a data analysis subgroup is established to facilitate intersessional work of the RCMs
Justification	Increasingly the work of the RCM RCG can be and needs to be supported by work that utilises regional data. This can often be achieved very efficiently using open source software such as R, operating on common data formats as has been demonstrated in the COST project and the fishPi project.
Follow-up actions needed	RCG Chairs to identify a task leader. Task leader to establish mailing list of any and all interested individuals. RCG chairs to consider role, operation, remit, time commitments, if necessary funding, of the work of such a group, and to begin a dialogue with interested parties.
Responsible persons for follow-up actions	RCG Chairs, Task leader
Time frame (Deadline)	Intersessionally 2016-2017
SCRDB Comments	SCRDB appreciate the formation of a data analysis group that potentially could support the forthcoming development of the RDB. SCRDB realises that this group, as a RCG subgroup, need to have access to detailed data. A specific SharePoint, with a more restricted access (only active members in the data analysis group) than the normal RCM SharePoint, will be established by ICES to meet their need.

LM 12. RCM NA 10: Foreign Landings	
RCM NA 2016 Recommendation	RCM NA recommends that: 1. MS upload their samples of non-flag vessels (foreign landings) to the RDB. 2. The WKCATCH investigate suitable methods for estimating non Flag landing fractions.
Justification	27% of landed tonnages are by foreign vessels landing into countries other than their own flag country. Follow up work from previous recommendations on foreign landings have demonstrated that the RDB has the facility to hold non flag vessel sampling data. MS should therefore upload any and all foreign landings they collect.
Follow-up actions needed	RCG Chairs to put out a data call. MS to upload their samples of non-flag vessels (foreign landings) to the RDB. WKCATCH investigate suitable methods for estimating non Flag landing fractions.
Responsible persons for follow-up actions	RCG Chairs, MS and WGCATCH
Time frame (Deadline)	2017
LM Comments	LM endorses this recommendation and this needs to refer to ICES to be included in the ToRs
SCRDB Comments	SCRDB underline that the RDB has the facility to hold data on landings from foreign vessels. It needs to be highlighted in the data call that MS are supposed to upload this data as well. To be considered by the RCG chairs prior to the data call 2017.

LM 20. RCM NA 2: ASSIGNING TRIPS TO MÉTIER	
RCM NA 2016 Recommendation	The RCM NA strongly recommends a data compilation workshop to: 1. Standardise the processes that use trip based transversal data to determine the métier. 2. Consider criteria for aggregating data for different end-users (JRC, ICES and other RFMOs). *3.* Investigate a framework for managing métier and fleet descriptions when needed.
Justification	All catch, effort and sample data is uploaded to the RDB by métier and limited to a reference list of RCM agreed métiers defined by ICES area. MS submit their data to ICES on Intercatch disaggregated by métier and the JRC for FDI datacalls and other RFMOs also request the data by métier without reference to the RCM agreed list of métiers. MS have independently developed their own code and processes for calculating species assemblage and rules for merging métiers. This is initially based on the transversal data for a trip which may or may not cover more than one métier. These methods are not necessarily consistent between countries and even between agencies within countries and could be based on, for example, a foreshortened list of species assemblages; rules for particular gears; catch by weight and or value and relative ratios. The impact of the different methods is not known but as data is being compared more readily at a regional or international level, so as to improve on confidence in current assumptions simple rules and standards need to apply. Since 2009 RCMNA has regularly recommended MS provide, maintain and update fleet descriptions for all sampled métiers to better define and compare similar fleets and the sampling of them between nations. This has been only moderately successful with no clear repository for them. ICES Expert Working Groups are now also requesting fishery and métier descriptions as part of their data calls. A repository and better strategy is required for collating and maintaining MS descriptions.
Follow-up actions needed	JRC or ICES to setup a transversal data workshop: 1. to review current algorithms and processes for allocating a trip to a métier based on catch data. 2. to provide standard guidelines for determining or allocating a trip to a métier or multiple métiers and how to aggregate the data.

	3. to define a strategy for storing and maintaining national fishery descriptions relative to the defined métiers.
Responsible persons for follow-up actions	ICES or JRC to set up workshop NC and MS to provide contributors and to implement guidelines
Time frame (Deadline)	3 years
LM Comments	Liaison endorses this recommendation but the RCGs should consider how this is best progressed. RCGs should provide ToRs and a formal proposal for a joint JRC ICES series of workshops.
SCRDB Comments	SCRDB support the initiative, as consistent assignment of trips to métiers between MS are of relevance for the quality of transversal (landings and effort) data in the RDB when compiling data on a regional scale.

6   The use of the Regional Database in the Regional Coordination Meetings 2016

EU member states (MS) participating in the RCM Baltic, RCM NS&EA and RCM NA uploaded data in the RDB as a response of a data call launched by the RCM chairs in June 2016. The data call covered landing, effort and sampling data for 2015. Upload of previous years' data (2009-2014) were also requested if these data had changed substantially. The overall response of the data call was satisfactory, even though a few of the MS still are in progress with uploading their complete data sets.

The aim of the data-call was to facilitate analyses for regional sampling strategies, to solve previous problems related to the upload and also advance in the development of the RDB in order to build a tool for regional coordination where time series of regional 2009-2015 data would form the base for planning of future data collection.

The RCMs have, since the implementation of the RDB in 2010, systematically been working towards better harmonization and completeness of the data in the RDB. Harmonization has in recent years been sufficiently improved through harmonized codes for harbours (LOCODE), restricted lists of regionally agreed métiers and adoption of the World Register of Marine Species (WoRMS) AlphaID for species. Upload logs have been facilitated and used to identify (and were possible solve) obstacles for MS to upload data to increase completeness.

The RCMs have since 2010 been using RDB data for

- Overviews and ranking of métiers within the different regions
- Overviews of regional landings and effort to inform revised sampling designs.
- Analysis of landings abroad
- Overviews of sampling intensities for different métiers
- Overviews of biological sampling of stocks
- Basic quality control of regional data

Since 2013 the RCMs have been provided with an R package "RCMfunctions" utilising some of the tools developed during the COST project (COST 2006), and others developed independently since. The functions within this package support the data analysis carried out before or during the RCMs and provide tools for:

- Basic data manipulation tasks for regional data sets.
- Quantifying landings and effort data by different grouping variables (such as harbours, fleet segments, metiers etc.)
- Plotting functions for sampled length frequencies, and length and age scatterplots.
- Mapping functions.

In 2016 a sampling sampling summary function was developed that:

1. Provides the institute uploading the data an overview of what has been successfully uploaded.
2. Provides the potential end user with a summary of what data has been collected, and therefore what is potentially available The caveat being that this is raw sampling data, not estimates of population parameters, and therefore of limited use.

3. Provides a concise summary which can form the basis of the proposed annual and link with the revised sampling tables 4A and 4B (and 4C and 4D) in the national work plans. As such these summaries give the appropriate scrutiny group a clear quantitate table of the achieved sampling in relation to the proposed national programmes. In this way the assessment of the statistical basis of the assessment of the sampling schemes would be dramatically improved.

The RCMs were testing the scripts. Examples of the outputs are shown below

	Number of unique harbours visited	Number of unique vessels	Number of unique voyages	Vessel days	Number of unique species	Number of lengths recorded	Number of species with biological information	Number of ages (with lengths)	Number of individual fish weights	Number of fish of determined sex	Number of maturity records
BEL	10	20	33	233	48	535005	7	6090	0	0	0
DEU	38	70	166	381	137	211920	16	23583	22775	18138	20765
DNK	23	143	418	976	106	264212	51	27065	57928	8781	0
ENG	38	123	194	342	127	272978	10	2102	0	2174	0
ESP	20	47	171	646	159	208140	8	0	13652	13518	0
EST	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
FIN	1	1	179	107	32	12532	5	1624	5462	5017	4293
GBR	4	1	164	162	57	181922	NA	NA	NA	NA	NA
IRL	13	25	72	378	119	128918	6	3024	3024	1373	0
LTU	4	2	21	120	15	15286	3	1063	4111	3011	1063
LVA	2	1	133	147	33	62404	5	13704	13704	11100	9514
NLD	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
POL	12	23	91	229	69	72752	16	8867	8867	8747	8357
PRT	12	1	95	85	204	23334	3	0	883	591	0
SCT	21	116	219	785	129	500646	4	8500	0	0	0
SWE	36	65	119	144	139	72602	7	3794	7348	22623	328

Example 1 showing the output table of sea-sampling events by country from the sampling summary function.
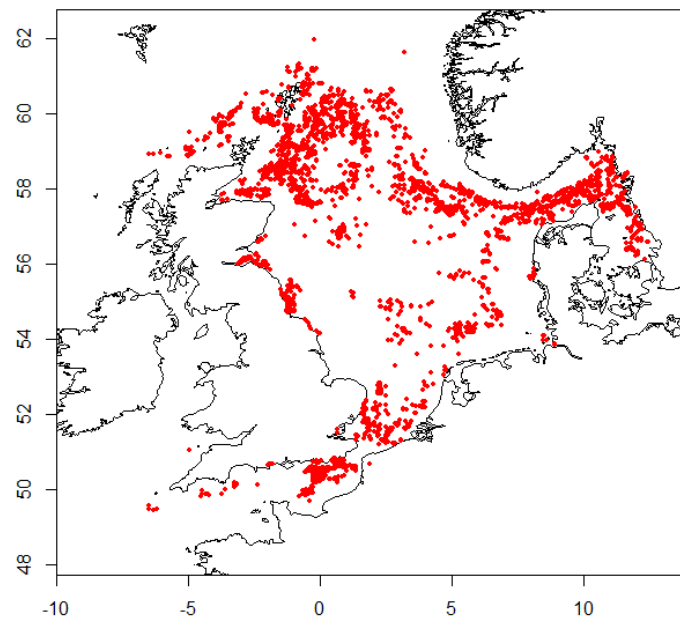
Example 2 Output of the plotting functions showing positions from sea-sampling trips, in this instance North Sea cod.

Length distribution for Atlantic cod by sampCtry
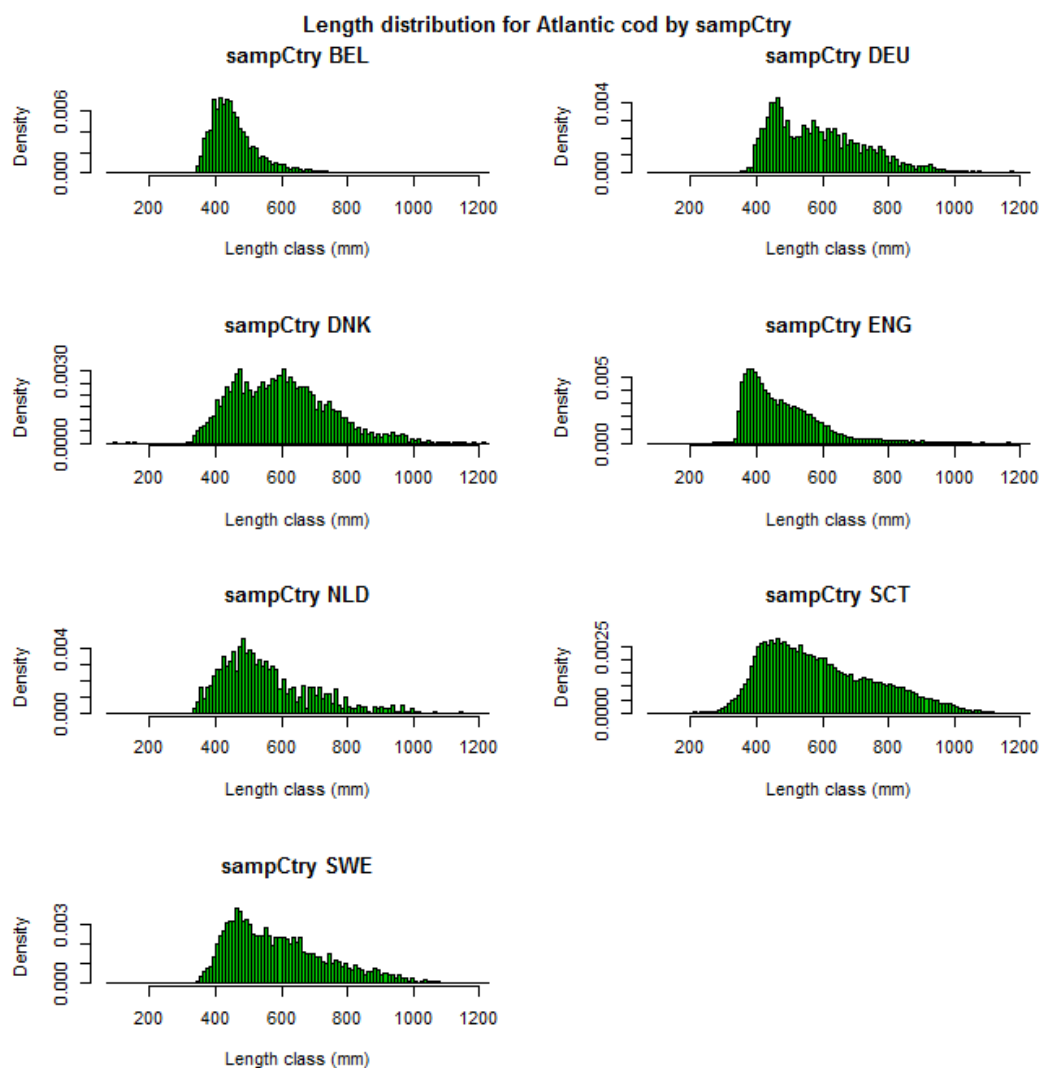
Example 3 Length Frequency plots showing the length distributions of the measured landed fraction of cod by sampling country.

Reference

COST. 2006. Studies and Pilot Projects for Carrying Out the Common Fisheries Policy Call for Proposal Ref : FISH/2006/15 – Lot 2 Project No :SI2.467814 Project Acronym :COST. EUROPEAN COMMISSION.

7   Development strategy for the Regional Data Base (ToR e and f)

The RDB will be developed over the next 2 years by the ICES secretariat – this is an ambitious time-scale and is driven by the funding supplied by ICES Council. At the end of the 2 years the new RDB should be capable of both storing data and utilising appropriate statistical estimation methods. It is likely that different countries/groups will still need to use InterCatch for a transitional period after the new RDB completion. The long term aim is for the new RDB to completely replace InterCatch. It was highlighted that the development of the new RDB within such a tight timeframe would require a dedicated support group providing 'instant' support throughout the process. This group will be established under the current steering committee and will reflect on suggestions and proposals crucial to the process of developing and building the new RDB.

Development of InterCatch and the existing RDB will be minimised during this 2 year period with ICES prioritising the new RDB instead. The RDB development is intended to be an agile, iterative process and will require extensive input from scientists. The new RDB should not have any "black boxes" inside it – all processing should be transparent and subject to review and verification by experts. ICES secretariat (project manager Henrik Kjems-Nielsen, ICES) will be responsible for the project management of the process and for asking for help on specific topics. This input will be managed via the use of groups of experts – the exact groups required were not decided but they are likely to include: a "Core" group, a Data Group, an "R Peer Review" group, a "User Design" group and a "User Test" group. The experts will be invited from groups such as WGCATCH, RCG's, other institutional bodies and the RDB steering group. The steering group will need to decide on the input of different groups, they have to act adaptable and flexible. Some questions need to be answered sooner than others, for example, the security model that should be used needs to be decided early in the process.

The current RDB will serve as a starting point for the development of the new RDB. Procedures that can be reused efficiently will be used again. Due to the new development framework this will be limited. Other procedures need to be redeveloped matching the new RDB approach as structural differences between RDB Fishframe and the new RDB might require this. These differences emerge e.g. from the fundamentally different approach to have various analysis outside the RDB rather than an integral part of the RDB.

The figure 1 below gives an overview of the flow of data in the new RDB from data uploads and the interaction with the national experts to the two main end users RCM/RCG and ICES, but data could also be downloaded for other relevant end users e.g. STECF.
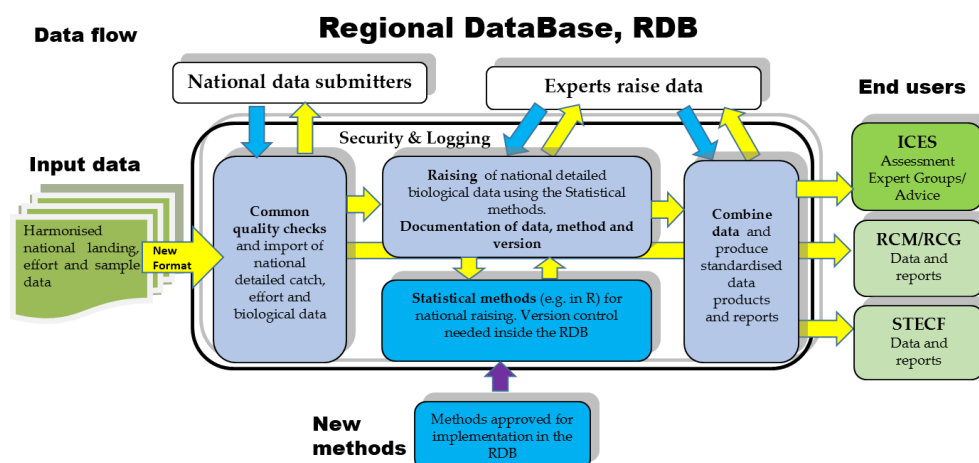
Figure 7.1. The new RDB system structure

A high-level design document (around 25 pages long and following a similar format to the Acoustic design document) will be written by the SCRDB and act as an overall guide to the development. This will not attempt to specify all the implementation details – these details will need to be fleshed out as part of the agile development process followed by ICES.

As part of the discussion the data model as developed by the WKRDB2014, WKRDB2015 (initiated by SCRDB) and the FishPi project, was presented. The idea was that this could be a starting point for the development of the new data model. The fact that this model is focused on commercial sampling (CS) was not considered a problem as this fits into the principle 'get our feet wet'. The CS fishPi model is considered to be well developed. The Landings and Effort data model is not as far progressed as the sampling data model and will require further work.

To provide a workable prototype of the data CS element of the model it is desirable to test the suitability of the data format and in particular that all the required fields can be accommodated in a logical fashion. To that end this format will be circulated to MS and members of the data group with a view to it being populated with examples of real data from as many schemes as possible. Progress will be reviewed by SCRDB in March 2017.

Actions:

- Jørgen Dalskov to draft the outline of the high level design document, including authors for the different sections
- SCRDB to write the sections of the high level design document, done by the first of March 2017
- Alistair Pout to provide through the RCG Data Group the means whereby countries have the opportunity to populate the new format; progress to be reviewed through on-line meetings by 1/3/17.
- Jørgen Dalskov/ICES secretariat to arrange a SCRDB WebEx to discuss the next steps, done before first of April.
- WGCATCH is recommended to add "Consider estimation and raising" to their ToRs

- High-level design document to be sent to national institutes responsible for data collection and processing

8   Test of data model on sampling data

The Commercial Sampling sub Schema for the RDB data model

Part of the envisaged data model for a revised RDB relating to the commercial sampling is shown in Figure 8.1. The nested CS data part of the model consists of the SE-CA structure tables. These differ from the TR-CA structure in the existing RDB Fish-Frame data model in that there is a SE table to record the sampling events, above the level of trip. Also the CA table (which records individual age, length weight measures etc. on individual fish/shellfish) is moved to be below the HL table, where length frequency are recorded. Many of the tables are also revised with additional fields, and many of the code lists used are those aligned at the regional level. One to many relationships (1:n) would exist in the SE-TR-HH-SL-HL-CA hierarchy between these tables.

The primary keys for the SE – CA tables will match the number of rows in each table and the exact form of the primary key has yet to be defined, as an appropriate key may differ between data sets. This differs from the approach in the original Fish-Frame data model where the primary key was derived from a specified number of fields and the population of those filed was therefore mandatory. One approach to defining this is that the variables for the primary key is defined explicitly from the pertinent fields for the sampling situation as recorded in the table, so the "xxxKey" field records those pertinent fields and an appropriate "xxxxID" field is generated from a concatenation of those fields. For example a primary key for the fishing operation (HH table) could be defined as trip code station number and landing fraction e.g. foKey = "trpCode_staNum_landFrac". The actual value of the key would then have the form "M20150123_3_A" for example depending on the exact code lists being used.

The calculation of an unbiased estimate in a probability based sampling design requires the calculation of the correct sample weight $\omega_{s,v,t,u,f,a}$ at each level S, V, T, U, F of a hierarchical multistage sample. The combined sample weight which is the inverse of the product of the sampling probabilities for all the levels of the multistage sampling:

$$\omega_{s,v,t,u,,f,a} = \frac{1}{\pi} = \frac{S}{s} \times \frac{V}{v} \times \frac{T}{t} \times \frac{U}{u} \times \frac{F}{f} \times n_{s,v,t,u,f,a}$$

where ω is the sample weight and π the inclusion probability. To this end each of the SE – CA tables has a field were the total "xxxTotal" and the number sampled can be recorded "xxxxSampled" can be recorded. These are used to generate the "xxxSampProbability" field for each table.

Where the total are not available an auxiliary variable (such as weight or a time period), can be used to generate a ratio estimate. The exact form this may take may be one of the areas that need to be determined as different auxiliary variables may be appropriate in different sampling situations. The auxiliary variable obviously also has to be recorded, and thus recorded in the appropriate table, for the calculation of the sample weight.

To the right of the csPi model (figure 8.1) are (some) of the existing or envisaged tables relating to regional data. These tables relate to for example, species (based on WoRMS), location, metier, areas. The envisaged primary key for each table is in bold, e.g. WoRMS AlphiaIDs and LOCODE for the location codes. A vessel table would be

particularly useful for regional data, and the EU vessel register is the clear candidate for this table.

To the left csPi model (figure 8.1) is an envisaged voyage table. This would consist of trip level data derived from logbook and sales note data. This would not be part of the RDB but would be a table used to populate parts of the CS structure at the national level as required. This table would remain at the national level.

It is envisaged that the SE-CA format can be utilised for the recording of sampling situations for commercial fisheries designs at the level of schemes and stratum from the primary sampling unit (PSU) down to the individual fish or shellfish. The SE table will record the scheme and the stratum, and the PSU totals of the respective sampling frame. It should be applicable to on-shore sampling involving visits to specific locations on particular dates, with the sampling from known vessels, market categories, as well as when vessels are not know and fish area sampled from market categories or purchased from vendors. For the at-sea sampling the situations where the PSU is considered to be the vessel or the vessel x trip can hopefully be recorded.

The proposed format has been developed at the WKRDB 2014 in Aberdeen, WKRDB 2015 in Séte and refined and converted into an R object during the fishPi project (MARE2014/19). The prototype is defined as an R object in both the "fishPiFunctions" and the "fishPifct" packages made in the fishPi project. The current list of field names are in the Annex 9 of the fishPi report.

To provide a workable prototype of the data CS element of the model it is desirable to test the suitability of the data format and in particular that all the required fields can be accommodated in a logical fashion. To that end this format will be circulated to MS and members of the data group with a view to it being populated with examples of real data from as many schemes as possible. Progress will be reviewed by SCRDB in March 2017.
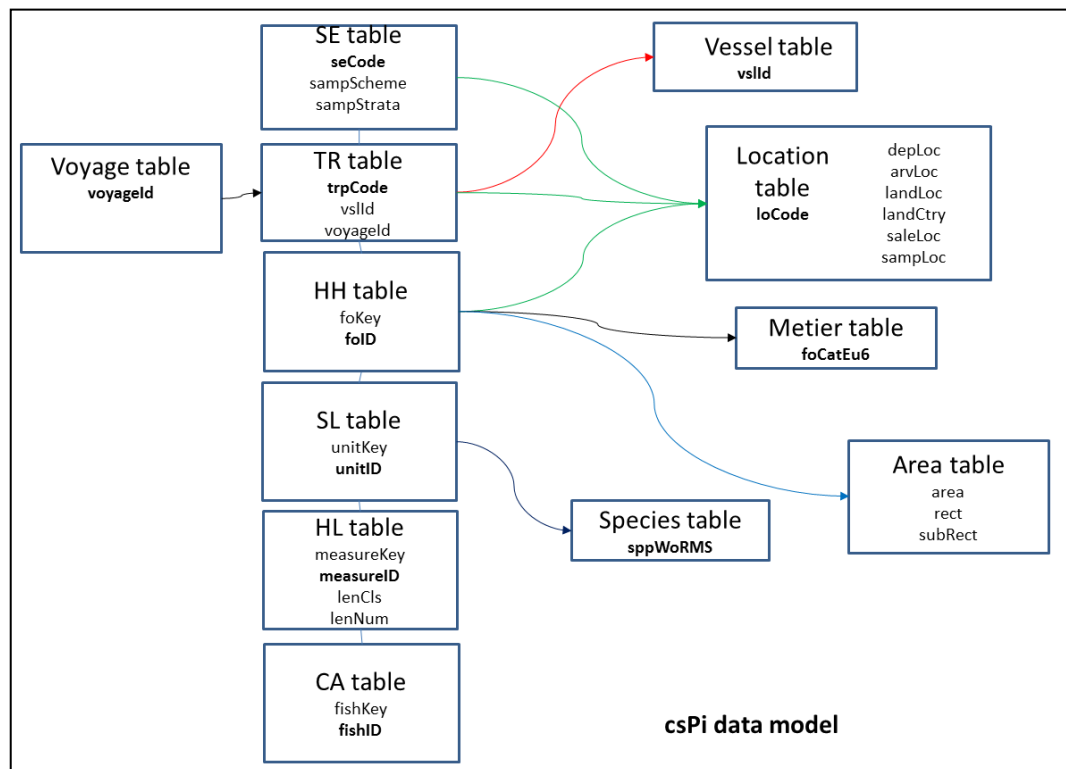
Figure 8.1. Part of the revised data model for the RDB showing the csPi tables for recording of commercial sampling data and related tables.

9   Status of the data policy document dealing with access rights, data confidentiality and data ownership issues (ToR d)

The SCRDB does at this point not want to change the present Data Policy[1]. The SCRDB agreed that in the long run the data policy should be more open. The new invited non-EU countries will go through the Data Policy to see if they can support the Data Policy as it is or if additions are needed, in connections with them uploading data to the RDB.

Under the RCMs/RCGs a data group will be working on the data and statistical estimations. Since the group is a subgroup working under the RCMs/RCGs, the persons do not need to be present at the RCMs/RCGs to be able to have access to the detailed data across the RCMs/RCGs in the RDB. Non-EU countries will not be a part of this data group.

WGCATCH is working on the statistical estimations of data, they can get access to the detailed data in the RDB through a data call to countries.

---

[1] http://ices.dk/marine-data/Documents/Data_Policy_RDB.pdf

10   RCG data group

The RCG data group is a loose grouping of 24 individuals from 13 countries with skills in the statistical computing language R (R Core Team, 2015[2]) which was assembled following initiatives in the RCMs in 2016. The aim of this grouping was to facilitating intersessional and pan regional work with regional data in the data formats used by the RDB (CS CL and CE data formats). Members of the group have been involved in RCM data work, the fishPi project and data uploads to the RDB, but not all are regular attendees of the RCM's. Such a group broadens the pool of skilled people able to develop the RDB and the provision key statistical advice.

The SCRDB agreed to the establishment of a secure pan-regional section of the RCM share point site for holding RDB data sets as R objects with a user membership limited to those in the data group. Use of the data is in accordance to the data policy document.

---

[2] R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

11 RDB modification: Area Codes

Area codes used in the RDB were updated in December 2016 from the Arabic numeral codes used presently to the FAO Fishing Areas for Statistical Purposes Codes http://www.fao.org/fishery/cwp/handbook/h/en. Hence, ICES area fall into major fishing area 27 and so, for example, ICES sub area 4a will become 27.4.a and area 25 in the Baltic will become 3.d.25. FAO major fishing area 21 covers the NAFO convention area and FAO major fishing area 37 covers the Mediterranean and Black Sea. These codes will bring the RDB into line with InterCatch, the EU and those used in the COST project.

Annex 1: SCRDB Agenda

Tuesday, 29th November 2016

<u>13.00 - 13.15</u>

Welcome, introduction of the participants, organization & house rules, adoption of the agenda.

<u>13.15 – 14:30</u>

<u>Presentations</u>

The aim of the Regional Data Base (RDB) (Jørgen)

What is the status of the RDB and which data types and data are at present hold in the database and ideas for the future? (Henrik)

Examples of analysis made using the data in the RDB. (Alastair)

Outcome of the Council and the Bureau meeting discussions on the RDB. (Neil)

<u>14.30 – 15:00</u>

ToR a): Conclude on membership of the steering committee.

<u>15.00 – 15:30</u>

ToR b): Respond to recommendations put forward to the SC-RDB by the Liaison Meeting and ICES expert groups. (Presentation Jørgen – rapporteur Katja)

<u>15.30 – 16:00 Coffee break</u>

<u>16.00 – 17:00</u>

Webex with USA and Canada

<u>17.00 – 17:15</u>

Tor c): Summarize how the RDB has been used in the regional coordination meetings. (Presentation and rapporteur Katja).

<u>17.15 – 18:00</u>

ToR d): Based on requests from the RCG's consider and conclude on possible revision of the data policy document, dealing with access rights, data confidentiality and data ownership issues, following the consultation process, amend if necessary and adopt the final document. Any data policy issues related to EU and non-EU countries should be considered and taken into account. (Presentation & rapporteur Neil and Henrik)

18.00 End of the day

Wednesday, 30th November 2016

<u>09:00 - 10.30</u>

Continue ToR d)

<u>10.30 – 11:00 Coffee break</u>

11.00 – 12:30

ToR e): Based on the findings made during the FishPi project and the requests from the RCGs consider and conclude on a new exchange format and suggest a route forward. (Presentation Katja - rapporteur Alastair)

12:30 – 14:00 Lunch

14:00 – 15:30

ToR f): Continue to develop a strategy based on revised DCF and EU-MAP regulations, on development of the RDB, taking requirements from a statistical sound sampling and raising and the landing obligation into account. Report on progress for the short, medium and long term plans developed so far. (Rapporteur David, Sieto and Peter)

- Establishment of RDB Project Expert Group subgroups under the SCRDB

15.30 – 16:00 Coffee break

16.00 – 17:00

Webex with USA and Canada

17.00 – 18:00

Continue ToR e):

18.00 End of the day

Thursday, 1st December 2016

09.00 - 10.30:

Writing text for the report and "stand alone document" like the policy document.

10.30 - 11.00: Coffee break

11.00 - 13.00:

Agree on data policies, development strategies etc. and finalising the report.

13:00 Closure of the meeting

Annex 2: SCRDB 2016 meeting list of Participants

For the SCRDB meeting 2016 it was decided also to invite all non-EU ICES member countries plus representatives from the EU Commission. In addition to representatives from the Regional Coordination meetings for The Baltic, The North Sea & Eastern Arctic, The North Atlantic, the Large Pelagic and the Long Distance Fisheries, representatives from Norway and Iceland participated in the SCRDB meeting. A representative from Canada joined the meeting via Webex for two hours.

Name	Country	E-mail
Jørgen Dalskov, chair	Denmark	jd@aqua.dtu.dk
Neil Holdsworth	ICES	NeilH@ices.dk
Henrik Kjems-Nielsen	ICES	henrikkn@ices.dk
Rui Catarino	ICES	rui.catarino@ices.dk
Eskild Kirkegaard	ICES	eskild.Kirkegaard@ices.dk
Asta Gudmundsdottir	Iceland	asta@havogvatn.is
Katja Ringdahl	Sweden	Katja.Ringdahl@slu.se
Peter van der Kamp	Netherlands	peter.vanderkamp@wur.nl
Sieto Verver	Netherlands	Sieto.Verver@wur.nl
Alastair Pout	Scotland	a.pout@marlab.ac.uk
David Currie	Ireland	David.Currie@Marine.ie
Edvin Fuglebakk	Norway	edvin.fuglebakk@imr.no
Norbert Billet	France	norbert.billet@ird.fr
Mike McMahon *	Canada	Mike.McMahon@dfo-mpo.gc.ca

* Parttime via Webex.