

Call MARE/2020/08

Agreement reference: SI2.839444

European Maritime and Fisheries Fund (EMFF)

**Development of the Regional Database for the
Mediterranean & Black Seas**



Work-package 1 - Deliverable 1.1

Compilation of a list of stakeholders and sources used

M.T. Spedicato, G. Tserpes

Partners involved:

COISPA, HCMR, CIBM, IFREMER, NISEA

Table of Contents

Acronyms	1
Executive summary	3
1. Introduction	4
2. Working method	5
3. The Data Collection Framework (DCF) and data quality	7
3.1 National Work Plans and STECF advice	7
3.2 DCF Data Calls and data checks	8
4. The Data Collection Reference Framework (DCRF)	10
5. The main tools in support of DCF	12
6. Consultation with stakeholders and end-users	15
6.1 Consultations and meetings with RCG Med&BS and the Regional Data Base Steering Committee (RDB-SC) ..	15
6.2 Consultations with Member States	15
6.4 Consultation with ICES and RDBES data base Steering Committee	18
7. Mapping the main stakeholders and end users in the Mediterranean and Black Sea	23
8. Current functionalities and data uses	27
Conclusions	31

Acronyms

AER	Annual Economic Report
AR	Annual Report
AS-IS analysis	Analysis of the current state
CFP	Common Fisheries Policy
CPC	GFCM contracting parties
DATRAS	Database of Trawl Surveys
DC	Data Call
DCF	Data Collection Framework
DCRF	Data Collection Reference Framework
DTMT	Data Transmission Monitoring Tool
DVT	Data Validation Tool
ERS	Electronic Reporting System
EU	European Union
EUMAP	European Multi Annual Programme
EWG	Expert Working Group
FDI	Fisheries Dependent Information
GFCM	General Fisheries Commission for the Mediterranean
ICES	International Council for the Exploration of the Sea.
JRC	Joint Research Centre
LDF	Long Distance Fisheries
LFD	Length Frequency Distributions
LM	Liaison Meeting
LP	Large Pelagic
MCDA	Multi Criteria Decision Analysis
MS	Member States
MSFD	Marine Strategy Framework Directive
NA	North Atlantic
NS&EA	North Sea & Eastern Arctic
RCG	Regional Coordination Group
PET	Protected, Endangered and Threatened species
PGDATA	Planning Group on Data Needs for Assessments and Advice
QD	Quality Documentation
RCG Med&BS	Regional Coordination Group of the Mediterranean and Black Sea
RCM Med&BS	Regional Coordination Meeting of the Mediterranean and Black Sea
RDB	Regional database
RDBES	Regional Database and Estimation System
RWP	Regional Work Plan
SAC	Scientific Advisory Committee on Fisheries
SAF	Stock Assessment Form
SC	Steering Committee
SDEF	Standard Data-Exchange format
SS	sampling scheme
STAR	Stock Assessment Results
STECF	Scientific, Technical and Economic Committee for Fisheries
STREAM (project)	STrengtheningREgional cooperation in the Area of fisheries biological data
TAF	Transparent Assessment Framework
VME	Vulnerable Marine Ecosystem
VMS	Vessel Monitoring System
WP	National work plan
WGBYC	Working group on bycatch of protected species
WGFBIT	Fisheries Benthic Impact and Trade-offs
WGCATCH	Working Group on Commercial Catches

WGRDBESGOV	Working Group on Governance of the Regional Database & Estimation System
WKRDB-EST	Workshop on Estimation with the RDBES data model
WKRDB-POP3	Workshop on population of the RDBES data model

Executive summary

This deliverable analyses aspects related to the data quality of the Data Collection Framework (DCF), in relation to the National Work Plans and the Data Calls. Data checks and procedures performed by the Scientific, Technical and Economic Committee for Fisheries (STECF)/Joint Research Centre (JRC) are reviewed, to highlight the need of implementing appropriate quality checks for several data domains. This review also considers the General Fisheries Commission for the Mediterranean (GFCM) work related to the Data Collection Reference Framework (DCRF) Data Call.

An inventory is provided with the main tools in support of the DCF, either from previous grants (e.g. MARE 2014/19 and STREAM projects¹) or as open-source/free tools that can be used with little configuration/upgrade.

Consultations and meetings with the Regional Coordination Group of the Mediterranean and Black Sea (RCG Med&BS) and the Regional Data Base Steering Committee (RDB-SC) are reported. Consultations with Member States—aimed at identifying if a structure comparable to the RDBES with the relevant estimation procedures is in place - are also reported. In most countries, a National database (DB) has been established and is running, though a structure comparable to the RDBES with the relevant estimation procedures is not still in place in most countries.

Consultations with chairs of the Regional Database and Estimation System (RDBES) pointed out that much progress has been done and the RDBES database is expected to be finalised by 2023. Thus, it is pivotal that RDBFIS is progressing towards an RDBES-like structure, starting with a suitable hierarchy as an example, but is also able to store the detailed data in the RCG CS (sampling) and CL (landing) formats for a phase-in period.

An exercise of mapping stakeholders, potential users and their expectations, including 10 main groups of stakeholders and end users, is presented. A number of functionalities/uses of RDBFIS are also outlined.

¹ <https://datacollection.jrc.ec.europa.eu/docs/regional-grants>

1. Introduction

The need of developing a regional database (RDB) for the Mediterranean and Black Sea was atop priority since long time for the Regional Coordination Meeting of the Mediterranean and Black Sea (RCM Med&BS) and subsequently for the Regional Coordination Group of the Mediterranean and Black Sea (RCG Med&BS). Main objectives were to allow an efficient use of the data received from the official RCG data calls and to facilitate a better performance towards efficient management, fast response time of data processing for the Data Calls under the Data Collection Framework (DCF) and more recently under the Data Collection Reference Framework (DCRF).

A Steering Committee of the Med&BS RDB was also established in 2012² to discuss the principles for the governance of a RDB hosting the DCF data. In 2013 this Steering Committee agreed on a Data Policy Document (data confidentiality and data ownership policy). Since then there was no progress on the development of the RDB. The Med&BS RDB Steering Committee (SC) met again in January 2019³, to investigate possible funding sources, structure of the future data base and hosting options, all this until the advent of the current project “Mediterranean & Black Sea regional data base fisheries information system - MED&BS RDBFIS”

RDBFIS aims at providing a tool to support the work of the RCG Med&BS and Member States, improving their performance towards faster data processing and response to data calls, increased quality and robustness of data delivered to end-users, thus ultimately strengthening regional and EU-wide cooperation on data collection.

In particular, RDBFIS is a web-based integrated fisheries information system developed to:

- a. enable reliable scientific advice and support the work of the RCGs;
- b. facilitate the work of the EU MS;
- c. allow end users to calculate statistical estimates of data tailored to their needs;
- d. improve the transparency in the sampling and statistical estimation procedures;
- e. contribute towards/facilitate regional sampling plans.

The RDBFIS foresees to include several data types:

- i) aggregated landings and effort (transversal aggregated data);
- ii) detailed biological data (biological samplings and biological parameters) of demersal and small pelagic species;
- iii) scientific surveys data (MEDITS and MEDIAS at this stage);
- iv) spatial fishing footprints (with main focus on Multi Criteria Decision Analysis - MCDA - for small scale fisheries);
- v) by-catch and PET samplings, recreational fishery, alien species.

Regarding the scientific surveys data, the structure of SOLEMON and Black Sea files is according to the MEDITS format. These data could be thus included in RDBFIS at any further stage.

²Final Report of the 1st Steering Committee Meeting for the Mediterranean & Black Sea Regional DataBase (Med&BS-RDB). Rome, 29-30 November 2012. Also cited in Report of the Regional Coordination Group Meeting for the Mediterranean and Black Sea 2020 (31/08/2020 - 2/09/2020). https://datacollection.jrc.ec.europa.eu/documents/10213/1239599/2020_RCG+MED-BE.pdf/7722d6eaf2ef-4a07-94d1-63d907a307f6?version=1.0

³https://hcmrhellas.sharepoint.com/:b:/r/sites/FISMedBSRDB/Shared%20Documents/Implementation/RDBFIS%20%26%20RDBES/MedBS%20RDB-2019_SC_Jan2019_Report.pdf?csf=1&web=1&e=aAbj7b

The inclusion of the above listed type of data is a particular characteristic of the RDBFIS project, given that the main focus of the Regional Database and Estimation System (RDBES)⁴, under development at ICES, is on biological samplings and related sampling procedures. Still, in the RDBES, the management of by-catch and PET data, as well as of recreational fisheries data, is ongoing. The 2021 RDBES test data call already requested incidental bycatch data⁵.

Work Package 1 (AS-IS analysis) of RDBFIS aims at providing an analysis of the current situation, ongoing studies, tools and developments that are relevant for the RDBFIS. The recommendations from STREAM project, in particular the Deliverable 3.2⁶, regarding the main characteristics of the RDB, are taken into account, considering the assessment of data quality and the emerging configuration of the RDBES. In the deliverable D3.2, besides the revision of the reference lists and data formats, an analysis of ongoing studies and developments (e.g. FishHub; STECF, 2017; ICES, 2018) was done, in order to provide elements for the identification of solutions related to the storage, processing and analysis of the data at regional level also as a basis for the evaluation of options regarding a RDB hosting .

In addition, WP1 of RDBFIS aims at providing an overview of the expectations of various stakeholders (requirement specifications) and potential users, taking into account the RCG Med&BS and RDB Med&BS recommendations.

2. Working method

The working method was based on the review, analysis and inventory of ongoing studies, tools and developments; in particular, the following documentation has been consulted (the list is not exhaustive):

- RCG MED&BS Reports from 2019 to 2021 and the Reports of the Med&BS RDB SC;
- Reports of RCG-NA-NSEA and RCG-Baltic 2019-2021; the RCG ECON Report of 2021; the RCG-LP report of 2020-2021;
- Reports of the Liaison Meetings (LM) 2019-2021;
- National Work Plans of DCF 2020-2021;
- STECF 17-11 - Quality Assurance for DCF data; STECF 20-06 - AER 2020; STECF 20-08 - Evaluation of DCF 2019 reports and DT issues; STECF 20-09 - West Med demersals stock assessments; STECF 20-10 - FDI - Fisheries Dependent Information; STECF 20-15 - Stock assessments in the Mediterranean Sea - (Adriatic, Ionian and Aegean Seas); STECF 20-16 - DCF Evaluation WPs - guidance – templates;
- STECF 21-02 - Methods supporting MED stock assessment; STECF 21-09 - Evaluation of AR and DTi; STECF 21-08 – Annual Economic Report; STECF-21-11- Stock Assessments: demersal stocks in the western Mediterranean Sea; STECF 21-12 - FDI - Fisheries Dependent Information; STECF 21-15 - Stock assessments in the Mediterranean Sea 2021 – Adriatic, Ionian and Aegean Seas; STECF reports 21-17 on the evaluation of Work Plan for 2022-2024 (or beyond);
- ICES 2020. ICES Planning Group on Data Needs for Assessments and Advice (PGDATA); ICES 2020. Report of ICES Workshop on standards and guidelines for fisheries dependent data; ICES 2020 Working Group on Commercial Catches (WGCATCH); ICES 2021. The third workshop on population of the RDBES data model (WKRDB-POP3); ICES 2021. Second workshop on estimation with the RDBES

⁴ ICES. 2021. *The Third Workshop on Population of the RDBES Data Model (WKRDB-POP3)*. ICES Scientific Reports. 3:109. 14 pp. <https://doi.org/10.17895/ices.pub.9375>

⁵ ICES. 2021. *The Third Workshop on Population of the RDBES Data Model (WKRDB-POP3)*. ICES Scientific Reports. 3:109. 14 pp. <https://doi.org/10.17895/ices.pub.9375>

⁶ (<https://datacollection.jrc.ec.europa.eu/docs/regional-grants>)

data model (WKRDB-EST2; outputs from 2020 meeting); RDBES GitHub (<https://github.com/ices-tools-dev/RDBES>) with the relevant documentation;

- FAO. 2019. Monitoring the incidental catch of vulnerable species in Mediterranean and Black Sea fisheries: Methodology for data collection. FAO Fisheries and Aquaculture Technical Paper No. 640. Rome, FAO.
- GFCM, 2018. GFCM Data Collection Reference Framework (DCRF). Version: 21.2
- GFCM. 2021. Forty-fourth session of the Commission, 2–6 November 2021. Post-Adoption Report;
- Reports of the SC of the surveys MEDIAS and MEDITs and their contribution to the RCG.

Furthermore, several RCGs and RDB SC Med&BS meetings were attended and specific/ targeted consultations with representative of RDBES and Member States were held.

An inventory and analysis regarding the following points is also provided:

- (i) the main available tools in support of the DCF carried out in MARE grants (<https://datacollection.jrc.ec.europa.eu/docs/regional-grants>),
- (ii) the available licenced and open-source/free tools that can be used out-of-the-box with little configuration and no development, coupled to
- (iii) information from consultations with the relevant end- users/stakeholders.

An exercise on mapping of the relevant stakeholders and end-users is provided, highlighting their objectives, expectations and needs.

3. The Data Collection Framework (DCF) and data quality

3.1 National Work Plans and STECF advice

Member States prepare and submit National work plan (WP) to the European Commission for approval according to the Article 6 of the Data Collection Framework (DCF) (EU Regulation 1004/2017). Then, STECF evaluates these WPs prior to their approval by the EU Commission.

According to the EU-MAP⁷, new multiannual WPs for data collection have been submitted by Member States on October 2021. The duration of these WPs is three (e.g. 2022-24) or more years. The novelty is that the structure and format of the WP tables and text boxes together with the guidance of WP/Annual Report (AR) templates, have changed⁸. Compared to the previous WPs, the tables have been reshuffled; some have been merged (i.e. tables 1B and 1C into 2.2, 4A and 4C into 2.5, 1G and 1H into 2.6) (STECF 21-17)⁹ and the template tables of the WP are currently conceived with a structure closer to that of a database. Following this new structure of the tables text boxes were revised and merged accordingly.

In addition, the quality assurance framework has been moved from Tables to Quality Annexes 1.1 and 1.2 for biological and economic data, respectively. These Annexes replace the previous text boxes 5A and 5B and introduced the new concept of Quality Documentation (QD) for all sampling schemes referred to in the biological (Tables 2.2, 2.3, 2.4, 2.5 and 2.6), economic and social data (Tables 5.2, 6.1 and 7.1). This Quality Documentation (QD) is considered pivotal to providing a better description of all the sampling schemes referred to in all the sections of the WP. This, in turn, is a step towards the improvement of the statistical support to the data collection under DCF. It also recalls the need that the Member States have to be more explicit about their sampling designs to comply with the requirements of a RDB, like the RDBES, with its estimation procedures and quality assurance aspects¹⁰.

STECF 21-17 noted that some R scripts were developed to ensure consistency between some tables (e.g. Table 2.1 vs. Table 2.2) and between all tables and the Master Code List. These scripts were widely used during the EWG. There is a need for further development of these types of automatic screening for both STECF experts and MS when compiling their WPs and Annual Reports (AR). The availability of such scripts on a public GitHub would greatly improve the editing and quality check burden of populating the set of tables. When publicly available these scripts can be included among the RDBFIS checks.

⁷ Commission Delegated Decision (EU) 2021/1167 of 27 April 2021 establishing the multiannual Union programme for the collection and management of biological, environmental, technical and socioeconomic data in the fisheries and aquaculture sectors from 2022

Commission Implementing Decision (EU) 2021/1168 of 27 April 2021 establishing the list of mandatory research surveys at sea and thresholds as part of the multiannual Union programme for the collection and management of data in the fisheries and aquaculture sectors from 2022.

⁸ Commission Implementing Decision (EU) 2022/39 of 12 January 2022 laying down rules on the format and timetables for the submission of national work plans and annual reports on data collection in the fisheries and aquaculture sectors, and repealing Implementing Decisions (EU) 2016/1701 and (EU) 2018/1283

⁹ Scientific, Technical and Economic Committee for Fisheries (STECF) –Evaluation of work plans for data collection (STECF-21-17). Publications Office of the European Union, Luxembourg, 2021, EUR 28359 EN, ISBN 978-92-76-45479-3, doi:10.2760/744849,JRC127696

¹⁰ RCG NA NS&EA RCG Baltic 2020. Regional Coordination Group North Atlantic, North Sea & Eastern Arctic and Regional Coordination Group Baltic. 2020. Part I Report, 110 pgs. Part II Decisions and Recommendations, 7 pgs. Part III, Intercessional Subgroup (ISSG) 2019-2020 Reports, 154 pgs. (<https://datacollection.jrc.ec.europa.eu/docs/rcg>).

3.2 DCF Data Calls and data checks

Member States have to report the DCF data to different data calls, i.e. RCG Med&BS, DG MARE MED&BS, GFCM/DCRF, DG MARE FDI, that requires a substantial effort in data preparation, given the different formats and codes for data reporting (further elements are provided in Section 5).

This increases the risk of non-compliance with each requested coding and, in turn, challenges the quality of information provided, as the probability of producing misaligned data among the data calls for incompleteness, different codification, different level of aggregation, etc. can be high. For example, STECF 21-02¹¹ noted that the EWG-21-02 also found differences in the landings data between the 2021 Mediterranean and Black Sea, the 2020 FDI and the 2020 AER data calls. When such differences between data call occur, it is not possible to know which data are accurate and which are not, so these issues must be treated by comparing the data procedures for each call. STECF 21-02 also suggested that RCGs should discuss these differences and take necessary steps to explain or resolve them.

The Joint Research Centre of the European Commission (JRC) supports the implementation of the EU Data Collection Framework (DCF) managing five data calls (<https://datacollection.jrc.ec.europa.eu/data-calls>): EU Fishing Fleet (for the Annual Economic Report - AER), EU Aquaculture, Fish Processing Industry, Mediterranean & Black Sea, and the Fisheries Dependent Information (FDI).

The ones particularly relevant for RDBFIS are:

- the Mediterranean & Black Sea data call (Med&BS DC), in which all the biological data are submitted for the stock assessment process providing advice for fisheries management;
- the FDI data call, in which the biological samplings for the Mediterranean and Black Sea are only related to the volume of discards by fleet;
- the EU Fishing Fleet data call, related to fishing fleets, landings, fishing effort and economic data.

In addition to the above, the RCG Med&BS Data Calls also relevant for RDBFIS. In this case the data are at detailed level.

Several checks are made in the upload phase of both Med&BS and FDI Data Calls. A Data Validation Tool (DVT) is available online to check data before the upload for both Data Calls^{12,13}. This tool is currently available with user credentials. During the data upload of both the Data Calls several syntactic checks are performed, in line with the ones implemented in the Data Validation Tools. Such checks are related to codes, in particular to verify that the specifications of the Annex 1^{14,15} of the Data Calls are correctly used in the Data Call tables. Further checks are related to value types (e.g. text or numerical values). In case of errors, submissions are transmitted back to Member States along with an error report. The data are stored at JRC in the DCF database.

At a second step and as regards the Med&BS DC, STECF-EWG working groups on stock assessment evaluate the data quality for the assessed stocks, while specific working groups (e.g. STECF-EWG 21-02 and STECF-

¹¹Scientific, Technical and Economic Committee for Fisheries (STECF) – Plenary Report (PLEN-21-02). EUR 28359 EN, Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-76-40592-4 (online), doi:10.2760/559965 (online), JRC126123.

¹²<https://datacollection.jrc.ec.europa.eu/dc/medbs>

¹³<https://datacollection.jrc.ec.europa.eu/dc/fdi>

¹⁴<https://datacollection.jrc.ec.europa.eu/documents/10213/1446215/ANNEX+1+-+Data+Call+2022.pdf/40afb799-86bc-4b68-a74e-918fcbe72ccd>

¹⁵https://datacollection.jrc.ec.europa.eu/documents/10213/1444011/2022_FDI_Annex1.docx.pdf/17a90766-5480-49d1-93e5-e9fa9153ec83

EWG 22-03¹⁶) make a data quality check also for not assessed stocks. In addition, specific STECF-EWGs assess the data quality of FDI¹⁷

These checks at STECF-EWG level are related to other data features, as coverage, potential anomalies, e.g. the presence of outliers, and data consistency:

- between the data submitted in the different tables of the data call;
- between the data reported at different levels (e.g. segment, national) over time;
- with other data sources.

The upload facilities are implemented (in Java) as Liferay portlets and the DCF data are currently stored in a PostgreSQL DBMS (ICES, PGDATA 2020). The same implementation has the Data Transmission Monitoring Tool (DTMT)¹⁸ that is available at <https://datacollection.jrc.ec.europa.eu/web/dcf/dtmt>. Access to DTMT is restricted and credentials are user-specific. All end-user groups e.g. STECF (and relevant Expert groups), ICES, GFCM etc., are provided with log-on credentials to report data by the JRC. Each Member State has access to their own data, credentials provided to the national correspondents. The European Commission has access to the complete tables.¹⁹ The DTMT has the objective to efficiently monitor and communicate data issues and in the long term to improve the flow and quality of data.

One of the STECF EWG 21-02 objectives was to agree on and use data quality checks to resolve issues and to stabilise and freeze time series in view of the stock assessments to be carried out in the following in EWGs of 2021 (EWGs 21-11 and 21-15). Extensive checks were run, identifying issues related to coverage, incompatible estimates between different sources of information, data gaps, and not plausible values related to certain parameters. Based on these outcomes, DG MARE asked all MSs to resubmit historic data series to the Med&BS DC. Issues not resolved with the National Correspondents in the new data call issued on July 2021 were reported in the DTMT. There were also three areas of poor sampling coverage brought to the attention of the RCG Med & BS²⁰ to facilitate the introduction of improvements in the WPs of the involved countries. Subsequently, EWGs 21-11 and 21-15 still identified data issues in the MEDITS survey data, discards and commercial data. Problems were related to raising factors, coverage, reliability of length frequency distributions, and non-updated biological parameters. Mismatches between landings of Med&BS DC and FDI were identified for some species.

Some mismatches were highlighted also within the FDI DC when crosschecking Table A- Catch summary vs. Table H- Landings by rectangle (the latter requiring finer spatial resolution than in Table A). Likewise, inconsistencies were also noted when cross-checking Table G - Effort summary and Table I - Effort by rectangle (the latter requiring finer spatial resolution than in Table I) for the fishing days²¹. Comparison with

¹⁶ <https://stecf.jrc.ec.europa.eu/ewg2203>

¹⁷ Scientific, Technical and Economic Committee for Fisheries (STECF) – Fisheries Dependent -Information – FDI (STECF-21-12). EUR 28359 EN, Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-76-45887-6, doi:10.2760/3742, JRC127727.

¹⁸ Data Transmission Monitoring Tool (DTMT) is a tool that aims at making users aware of issues with the data transmitted by MSs in response to official data calls under the DCF. Any issue entered in the tool requires MS to provide an explanation. In this tool each issue is recorded indicating issue type (coverage, timeliness, or quality) and the severity level (high, medium, or low).

¹⁹ [https://datacollection.jrc.ec.europa.eu/guidelines/dtmt#:~:text=Data%20Transmission%20Monitoring%20Tool%20\(DTMT,flow%20and%20quality%20of%20data](https://datacollection.jrc.ec.europa.eu/guidelines/dtmt#:~:text=Data%20Transmission%20Monitoring%20Tool%20(DTMT,flow%20and%20quality%20of%20data)

²⁰ RCG MED & BS 2021 Annual Meeting (7th-9th September 2021). Recommendation 1. Prioritization of stocks not included in stock assessment to check quality of Med&BS data for the Joint meeting of the RCG Med&BS, DG MARE, JRC and STECF on data quality and availability.

²¹ Scientific, Technical and Economic Committee for Fisheries (STECF) – Fisheries Dependent -Information – FDI (STECF-21-12). EUR 28359 EN, Publications Office of the European Union, Luxembourg, 2021.

EUROSTAT²² data was also carried out with the purpose of cross checking with an external data source, to verify the completeness of the submitted data sets.

The RCG MED & BSheld on September 7-9, 2021, highlighted in Recommendation 1 on Data quality and availability, the need to prioritize stocks not included in stock assessment to check quality of Med&BS data. COM proposed an ad hoc EWG, i.e. the STECF-EWG 22-03, to quality check the Med & BS data not currently scrutinized in STECF stock assessments.

It should be noted that, in principle, some of the data issues highlighted in the STECF-EWG meeting could be avoided with automatic procedures and systematic data quality checks as planned in RDBFIS.

4. The Data Collection Reference Framework (DCRF)

The fisheries Data Collection Reference Framework (DCRF) is an instrument established since 2013 with the aim of achieving efficient data collection in the whole GFCM area. The DCRF²³ is designed to support the implementation of the mid-term strategy (2017–2020) and of the GFCM 2030 Strategy²⁴ towards the sustainability of Mediterranean and Black Sea fisheries. The collection of fisheries-related data is needed to improve the formulation of sound scientific advice by relevant GFCM subsidiary bodies for those fish stocks in the Mediterranean and Black Sea that are shared among EU and non-EU countries.

The core principle behind DCRF design is to reduce data requirements into a simple, reliable and easy to understand format, simplifying the level of data aggregation for the requested variables. According to the GFCM DCRF, data should be reported by fleet segment, defined as a combination of vessel groups by gear (e.g. trawlers, pelagic trawlers, purse seiners, dredges, longliners) and length classes. Data related to passive gears (e.g. trammel nets, gillnets, traps) are reported at the level of '*small-scale vessels with engine using passive gears*' that are more aggregated than the segmentation adopted in the Data Collection Framework.

To support the data transmission, the DCRF online platform²⁵ was built with tools that mirror the DCRF Manual structure and that are in agreement with a set of GFCM recommendations²⁶. An online guide is also available with relevant information. The data-entry and submission are in line with the requirements of the GFCM decisions. The access to the platform is allowed to National Authorities of respective GFCM CPCs through credentials and reporting tools are progressively activated on the DCRF online platform according to the GFCM data submission calendar.

Fisheries quality indicators on the DCRF online platform and related checks workflows were also implemented to enhance the overall quality of fisheries data in the Mediterranean and Black Sea. In the work programme for the period 2021–2023 and implementation of the GFCM 2030 strategy regarding the quality of the fisheries data the Commission decided to:

- i) consolidate the regular application of quality indicators, i.e. timeliness, completeness, conformity, stability and consistency, via the DCRF online platform;

²² Data at <http://ec.europa.eu/eurostat/web/fisheries/data/database>

²³ GFCM, 2018. GFCM Data Collection Reference Framework (DCRF). Version: 21.2

²⁴ FAO. 2021. GFCM 2030 Strategy for sustainable fisheries and aquaculture in the Mediterranean and the Black Sea. Rome. <https://doi.org/10.4060/cb7562en>

²⁵ <https://www.fao.org/gfcm/data/dcrf/platform>

²⁶ Recommendation GFCM/41/2017/6 on the submission of data on fishing activities in the GFCM area of application

- ii) reinforce data collection and analysis at the regional, sub regional and national levels towards regular harmonized monitoring²⁷.

In addition, it was also decided to carry out a technical consultation with national experts for an effective transmission of fisheries data to the GFCM and for the consolidation of data quality assessment.

Late transmissions (timeliness) still represents an issue, mainly in the Mediterranean Sea. Compulsory data (completeness) and lack of adherence to GFCM standards (conformity) had a lower number of issues. Stability issues and several cases of non-coherent values of selected variables across different tasks/subtasks (consistency) were also quite common. There is then the need of raising awareness of the CPC countries (Liaison Meeting Report, 2021)²⁸.

Furthermore, the development of protocols (discards, incidental catch of vulnerable species, scientific surveys)²⁹ increased considerably in recent years, with the aim to provide a harmonized methodological framework for data collection and to promote the comparison of data at the regional and sub regional levels.

Regarding the stock assessment process, input data (e.g. landings, discards, LFDs, biological parameters, survey indices, survey LFDs) for the assessments are validated by the national focal points that upload them on the GFCM platform following a data call process. After validation of an assessment, a Stock Assessment Form (SAF) is compiled including all essential information on data, assessment methodology, estimates of fishing mortality, exploitation rate, spawning stock biomass, recruitment, and short-term projections. The SAFs report, the status of the stock and associated advice are then validated by the Scientific Advisory Committee on Fisheries (SAC). A summary sheet summarizing the main information on the assessment is also compiled by the experts. In 2021 the Stock Assessment Results (STAR) was introduced that is a new framework to organize GFCM stock assessment results. This new tool summarises validated stock assessments so that they could be included in relational databases. It streamlines and automates the information flow from stock assessments to scientific advice, strengthens quality assurance and data dissemination. More information about STAR can be found at <https://github.com/gfcm/star>.

²⁷ GFCM. 2021. *Forty-fourth session of the Commission online, 2–6 November 2021. POST-ADOPTION REPORT*

²⁸ Fisheries Data Collection. 18th Liaison Meeting Report. 5th November 2021

<https://datacollection.jrc.ec.europa.eu/docs/liaison>

²⁹ <https://www.fao.org/qfcm/data/good-practice-guides>; <https://www.fao.org/qfcm/publications/en/>

5. The main tools in support of DCF

An inventory of the main tools in support of DCF is provided herewith. This inventory contains tools developed in the previous regional grants (e.g. MARE 2014/19 and STREAM projects) and open-source/free tools that can be used in RDBFIS with little configuration/upgrade.

Upgrade of the R script developed in STREAM project for data format conversion to answer data calls

To fulfil the Common Fisheries Policy (CFP) objectives, Member States collect fishery data with the financial support of the European Commission under shared management of structural funds (between the Member State and the European Commission through operational programmes). Such data are raised and submitted in the required format to specific data calls for end-users' needs. These are related to the provision of scientific advice for the management of the fisheries resources. Member States have thus developed their internal IT systems for data storage, in some cases adopting the RCG MED&BS Data Call format, in other cases using some internal routines to convert the primary data from their own format to the RCG MED&BS Data Call format.

Some “reshaping” scripts developed in MARE 2014/19 and STREAM projects are already available for the conversion of RCG MED&BS Data Call format to SDEF format.

SDEF (Standard Data-Exchange format) format, adopted in the COST project³⁰, includes 5 tables regarding the biological sampling:

- Table TR (trip): Harbour
- Table HH (Fishing station): Fishing duration, latitude and longitude for the fishing operation, main fishing depth and main water depth.
- Table SL (Species List): Sex, commercial category scale
- Table HL (length): Sex, commercial category scale
- Table CA (sex- maturity- age –weight- length): individual level data (“commercial size category scale”, “single fish number”, “age”, “ageing method”, “sex”, “individual weight”, “maturity stage”, “maturity scale”, “maturity scaling method”)

and one table regarding landing data (CL).

The SDEF and the RCG data formats are adopted in RDBFIS to allow for a transition phase, until RDBES format is implemented.

To allow the conversion of datasets into the relevant formats for the data transmission to the main Data Calls: Med&BS, GFCM/DCRF and FDI, starting from the RCG and SDEF format, two sets of auxiliary scripts (below highlighted in grey) developed in R language in the STREAM project will be updated, to accommodate changes in the Data Calls formats:

- from RCG Med&BS Data Call format to the SDEF (COST);
- from the SDEF (COST) format to the DG MARE MED&BS Data Call;
- from the SDEF (COST) format to the GFCM/DCRF Data Call;
- from the SDEF (COST) format into DG MARE FDI Data Call format (using DG MARE Med&BS Data Call format).

The scripts will be upgraded to R functions to be included in the RDBprocessing package, taking into account the last specifications of DGMARE MED&BS and FDI Data Calls³¹.

Quality checks developed in STREAM projects

³⁰ <https://wwwz.ifremer.fr/cost/Software-Packages>

³¹ <https://datacollection.jrc.ec.europa.eu/dc/medbs> and <https://datacollection.jrc.ec.europa.eu/dc/fdi>

For some quality checks RDBFIS is building on the experience gained within the STREAM project concerning data quality procedures on detailed and aggregated data. The original philosophy behind STREAM quality checks is based on the concept of a 2-step process to verify the consistency of the biological data: 1) *a priori* quality checks (QC), to detect possible inconsistency and inaccuracies in the detailed data; 2) *a posteriori* QC, designed to verify that the data consistency is maintained in the aggregated dataset. Both the *a priori* and the *a posteriori* QC were originally implemented through an Rmd script allowing to produce an automatic report at the end of the procedure, indicating the outcomes of each quality check. This functionality will be upgraded in RDBFIS.

R code developed for STECF EWG 21-02

These scripts have been used during the STECF EWG 21-02 to carry out a number of quality checks on the MED & BS data call tables. The suite of R scripts aimed at supporting the experts in the data preparation for the stock assessment, were developed under R version 3.6.3 (64bit) and tested under R version 4.0.3 (64bit) environment. Some examples follow:

- 1) Checks on landings/discards tables: duplicated records, shift in length distribution, null weight with non-null length-frequency distribution and vice versa, double reporting, mean weight calculation, cumulative function calculation and comparison among the years;
- 2) Checks on catch table: sum of products, landing in weight by gear and metier
- 3) Checks on biological data tables: comparison of maturity/sex ratio at length/age across the years; a similar comparison is made for the growth (von Bertalanffy and age-length key) and for the length-weight relationships.

Two additional scripts (landing and discard) are dedicated to the reconstruction of the LFDs by gear/metier in the years when a gap was present.

All scripts work at stock level so the user must define at the start of the scripts the country, GSA and species for which the analyses have to be run. The scripts can be downloaded through the following link: <https://stecf.jrc.ec.europa.eu/documents/43805/2817637/STECF-21-02+Annexes.zip/6a41ceea-b90b-4de1-bd62-212cd9e58c2e>

RoME

RoME (R code to perform multiple checks on MEDITS Survey data) is an R package aimed at unifying the checks that are made independently over the MEDITS data by the 18 GSAs (Geographical Sub-Area) participating to MEDITS Survey. The package is structured in 55 different functions: each function is related to a specific check and is recalled in a specific order to avoid cascade errors. All the checks have been designed according to INSTRUCTION MANUAL VERSION 9 MEDITS 2017.

There are checks specifically designed for correcting the data contained in TA (haul information), TB (catch by haul), TC (aggregated length, maturity and sex information by haul), TE (individual biological information) and TL (litter by haul) tables. There are also checks developed to verify the consistency of the information across the tables (cross checks). RoME can be run on more than one year at a time.

Since 2012 JRC developed quality checks with routines in the MEDITS database to do cross table consistency tests and conformity to the survey manual checks. These checks shared a similar philosophy to the RoME routines and when this was used before data upload the JRC routines correctly show no error patterns³².

The last version of RoME is 1.4 and it works with R versions 2.15.2.

³² Scientific, Technical and Economic Committee for Fisheries (STECF) – 2013 Assessment of Mediterranean Sea stocks part I (STECF 13-22). 2013. Publications Office of the European Union, Luxembourg, EUR 26329 EN, JRC 86087, 400 pp.

RoME can be downloaded using the following link:

https://www.coispa.it/index.php?option=com_content&view=article&id=34&Itemid=119&lang=it

Regarding other surveys, as the ones in the Black Sea, the tool RoMEBS described below can be used for data checks.

RoMEBS

RoMEBS software was developed in 2019 to support the analysis of Black Sea turbot data. It consists of an adapted version of the RoME that allows multiple checks on MEDITS-like data tables. RoMEBS was modified to be easily run on both demersal trawl survey data (e.g. Black Sea turbot trawl survey) as well as beam trawl survey data (rapa whelk survey).

This is a library of functions developed to work in the R software environment (R version 3.6.0).

The package includes also a function (cood.conv) to support the coordinates conversion from decimal degrees format to the MEDITS one and vice versa, and a wizard function (RoMEBS.wizard), guiding/allowing the user to upload all the useful tables to perform the RoMEBS quality check data analysis.

RoMEBS can be downloaded using the following link:

https://www.coispa.it/index.php?option=com_content&view=article&id=34&Itemid=119&lang=it

FDI data call automatic checks

The data checks on the data collected in the FDI (Fishery Dependent Information, <https://datacollection.jrc.ec.europa.eu/dc/fdi>) data call are carried out automatically during the data uploading phase and after the data are stored in the database (partially through automatic procedures and partially through visual examination).

The majority of the automatic checks concerns the use of valid codes listed in the various appendices of the data call and the type of the data entered (numeric or text). Format and codes used in the data call are available on the website.

The information requested to the Mediterranean and Black Sea Member States is contained in tables A (Catch summary), B (refusal rate), G (Effort summary), H (Landings by rectangle) and I (Effort by rectangle).

In particular, the upload tool verifies the format of the files provided and checks the codes used to specify the following information: country, fishing technique, vessel length, gear type, target assemblage, mesh size range, metier, species, supra-region, sub-region, geographical indicator, EEZ indicator, Nephrops sub-region, deep fisheries, specific conditions related to technical measures (variable name: specon tech).

In addition, in Tables A, G, H and I, the consistency between sub-region codes and EEZ indicator codes are verified; in tables H and I, the format of the c-square and of the geographical coordinates are checked.

A list of checks is available in the STEC 21-12 report; in addition, Ms Zanzi (FDI reference person) provided an extensive list of quality checks after the request of the RDBFIS team.

6. Consultation with stakeholders and end-users

6.1 Consultations and meetings with RCG Med&BS and the Regional Data Base Steering Committee (RDB-SC)

RCG MED&BS 2020 recommended (Recommendation 12) the continuation of the work for a Regional Database which development was considered an urgent priority to allow for the efficient use of the data received from the official RCG data calls and avoid duplication of work.

RCG MED & BS 2021 recommended (Recommendation 5) for the governance of the regional data base RDBFIS to establish a Steering committee. A new chair was elected and Member States had to confirm/nominate SC members (2), DG MARE and GFCM to nominate members (2). Med&BS RDBFIS SC in 2022 should agree on RDBFIS hosting, government policy and cost sharing from 2023 onwards.

RCG MED & BS 2021 recommended (Recommendation 6) to establish a Technical group on the regional data base RDBFIS, following what was agreed with the Steering committee for the regional database during its 4th meeting in July 2021. The technical group has the role to communicate with the regional project Med&BS RDBFIS (MARE/2020/08) regarding technical aspects of RDB development and to further support development of this regional database. In particular, the involvement of national experts is expected in the testing of RDBFIS and the support to RDBFIS regional grant on specificities of Med&BS fisheries data, raising procedures etc. MS should nominate experts on a voluntary basis by the end of 2021.

The Steering Committee for the regional database met twice in 2021, with the aim of agreeing on the structure of the database, data to be included, discuss on hosting and cost-sharing options. The usefulness of the RDBES structure was highlighted.

6.2 Consultations with Member States

The study MARE/2014/19 Med&BS³³ reviewed the characteristics of the main IT systems developed in the Mediterranean Black Sea Region for a proper implementation of the DCF. This review highlighted that not all the MSs had a centralized database for DCF data and that a wide range of data storage systems across MSs existed, which were sometimes incompatible each other. Documentation of the databases was in some cases incomplete. The level of control and methods used for data quality checks varied between MSs and, in several cases, standard quality assurance procedures were not implemented.

In RDBFIS, the project coordinator carried out bilateral consultations with Member States, in order to understand the current state of play of the different data storage and transmission systems. In particular, to evaluate if a structure comparable to the RDBES with the relevant estimation procedures was in place or planned for the near future. Table 1 reports the main features of the National DB, indicating if a structure comparable to the RDBES with the relevant estimation procedures is in place (or planned for the near future), considering that RDBFIS is also implementing an RDBES-like file structure.

In most countries, a National DB is in place and running and, in some cases, an upgrade is foreseen. A structure comparable to the RDBES with the relevant estimation procedures is lacking. France is going in the direction of a RDBES as the sampling scheme (SS) is the same for the north of France and the Mediterranean coast and so the RDBES implementation will be aligned to the RCGs of the North Atlantic

³³ MARE/2014/19 - Strengthening regional cooperation in the area of fisheries data collection in the Mediterranean and Black Sea. AGREEMENT NUMBER – MARE/2014/19 -SI2.705484. D.O.2. FINAL REPORT. <https://datacollection.jrc.ec.europa.eu/mare-2014-2019-strengthening-regional-cooperation>

area once finalised. For Large Pelagic fishery, Cyprus adopted the RDBES hierarchy 5. The concept of sampling hierarchy in RDBES identifies a set of instructions that indicates what sampling levels are included in the multi-stage sampling of the commercial catches and how they are (hierarchically) related to each other. Hierarchy 5 identifies a sampling by selecting from location*time (e.g. harbour-day) as primary sampling unit, then from Landing Events as the secondary sampling unit. Such hierarchy is most used for on-shore sampling, but can be used for at-sea sampling as well.

Table 1 Main features of the National databases (DB) and availability of RDBES formats for biological sampling data.

Member State	National DB main features	RDBES format for biological samplings
Bulgaria (BGR)	The national database is under construction. The database will host records from logbook as well as scientific data collected by three Bulgarian Institutes. The system will store fishery Bulgarian data and will be set up in January 2022.	Even if data are not organized in RDBES format, efforts will be made to provide them according to the formats required by the RDBFIS.
Croatia (HRV)	A centralized database system established, hosted by the Ministry. The Research Institutes are connected to the DB to enter data, running the quality checks, analysis and Data Calls' outputs. This system will be re-developed using state of the art programming techniques.	Neither structure nor hierarchies of RDBES are included in the system for the raising of biological samples to the catches. Relevant discussions will take place to decide whether they will be included in the new system.
Cyprus (CYP)	A centralized database system has not been established in CYP yet.	The RDBES hierarchy 5 was adopted for Large Pelagic fishery. The adoption of other RDBES hierarchies is under investigation.
Greece (GRC)	Up to 2020, the National Database was hosted in the Hellenic Centre for Marine Research (HCMR) and the data storing, analysis and data call obligations were supported by the IMAS-Fish (an Integrated Fisheries Information System) After 2020 a National Centralize Database (NCD) established for storing all the data collected in the framework of the Data Collection Framework (DCF).	No provision has been made for the adoption of the RDBES structure as well as the sampling scheme model followed by ICES.
France (FRA)	A centralized database has been established supporting the Data Calls and the other obligations.	The sampling scheme (SS) is the same for the north of the France and the Mediterranean coast. France is involved in data submission to RDBES and Med&BS RDB and the burden of multiple data delivery needs to be taken into consideration.
Italy (ITA)	A centralized database has been established supporting the Data Calls and other obligations. FishDataNet	Neither structure nor hierarchies of RDBES are included in the system.
Malta (MLT)	National database of Malta (i.e. FIS) is in a transition phase. Currently, FIS is used to host logbook, fishing vessels register and survey data. Economic survey data are not stored in FIS. The	RDBES-like schemes will be adopted in accordance with the regional developments.

	DCF consortium starts a tender for developing a new system. Biological data from observers will be incorporated in this new system.	
Romania (ROU)	A centralized database has been established and hosted in Constantia for supporting the Data Calls and other obligations of the country. The plan is to develop a new state of the art information system.	Neither structure nor hierarchies of RDBES are included in the system. However, ROU could respond positively to the adoption of one or more RDBES hierarchies if this is deemed useful.
Slovenia (SVN)	A centralized database has been established in SVN supporting the Data Calls and other obligations.	Neither structure nor hierarchies of RDBES are included in the SVN system.
Spain (ESP)	SIRENO is the system, in ORACLE, implemented to keep and manage the fisheries information and for feeding DORI (national Database). SIRENO is the database which contains disaggregated data and hosts also transversal data from INFOBASE. SIRENO stores fleet, landings, discards, lengths, other biological information (individual weight, sex, maturity, age...) and surveys. Interface of biological data connected to INBIO (R tool) for estimating biological parameters. DORI, is a web-based application with two functions: - Data Repository (data on economic, social, biological variables and fishing activity); - Data Calls (generate reports and queries). Data on fishing activities are processed and INFOBASE is created (a series of linked tables).	Spain has delivered data for ICES RDBES in 2020 and 2021: - In 2020, it was for testing and development of scripts. - In 2021, it was for testing the RDBES data model and usage.

Remarks on the consultation

Bulgaria highlighted the added value of the RDBFIS to provide common facilities to be used by MSs by assuring homogenous process to assess quality and to deliver data to end users of the region.

Rumania and Slovenia as well underlined the importance of the system under development.

Also Italy observed that a tool with a regional dimension to support data calls and quality checks is useful and underlined the importance of the system under development to spare duplication of national efforts and support optimization. Italy engages in visiting the project web site and in getting more acquainted with the issues at hand, though it was underlined the importance of not duplicating efforts and considering what is already available.

France found RDBFIS usefulness in providing data to GFCM and easing analysis at regional level for specific requests. It could be useful to provide raw data and then use the RDBFIS to perform analysis. The link among LP (Large Pelagic) RCG (Regional Coordination Group), RDBES and Med&BS RDB is not clear; this is a crucial point for France, which has both Atlantic and Mediterranean coasts that are subject to fishery monitoring and sampling activities.

An RDBES approach was developed for the samplings of large pelagic in Cyprus.

The Spanish DCF team commented on the usefulness of the compatibility between RDBFIS and the Spanish national database because every change requires big work. Also, they expressed the hope that the Med&BS RDBFIS can help to reduce the burden to respond to multiple data calls. The team from Spain also commented on the test carried out for RDBES: in general, the processing of the data was carried out smoothly, but two difficulties were highlighted:

- Species code: WoRMS Aphia IDs codes is mandatory. ASFIS codes is optional. ASFIS codes are used in the rest of Data Calls.
- RDBES only allows level 6 metier codes. This creates problems with tuna metiers (level 7 by agreement of RCG LP), tuna metiers with level 6 cannot be distinguished.

The topic related to the VMS data storage is under discussion at the meeting of the RCG Steering Committee.

6.4 Consultation with ICES and RDBES data base Steering Committee

The RDBES chair and the responsible of ICES Data Centre have been consulted on the development of the RDBES, to allow for synergies. This is an important step, taking into consideration the need of a further evolution of standards and data formats to adopt at Mediterranean and Black Sea level (focus of tasks 3, 4 and 5 of WP4 in RDBFIS).

ICES expert groups (PGDATA, WGCATCH, WKRDB) and the wider ICES end user community strongly recommend to raise the data collected by EU MSs to the total production using statistically sound methods (4S sampling approach), in order to provide a documented, quality assured and accurate description of the catches to be used in scientific advice for fisheries management.

The existing InterCatch and Regional Database (RDB) database systems were respectively used by national institutes to submit raised commercial fisheries data for use in ICES stock assessments and to submit detailed commercial fisheries data to the Regional Coordination Groups (RCGs) (e.g. North Atlantic, North Sea & Eastern Arctic, Baltic and Long Distance Fisheries).

The RDB structure was based on ICES exchange format where people can have a common format for sharing their sample data and effort/landings aggregated data. This was predominantly used by the RCGs. This was fine to a certain point; many assumptions had to be made to follow a data structure, translating whatever was done nationally to that structure. Thus, information was lost on how the actual sampling was taking place; e.g., picking vessels or ports or trips from a list, with limitation on the data use and interpretation.

Figure 1 illustrates the detailed data flow from national institutes to RCGs (NA, NS&EA, Baltic and LDF) and ICES Expert Groups in the system based on InterCatch and RDB. The data is collected through two data calls to the countries. The way countries are raising/estimating the data for upload to InterCatch for ICES Expert Groups and advice is not known.

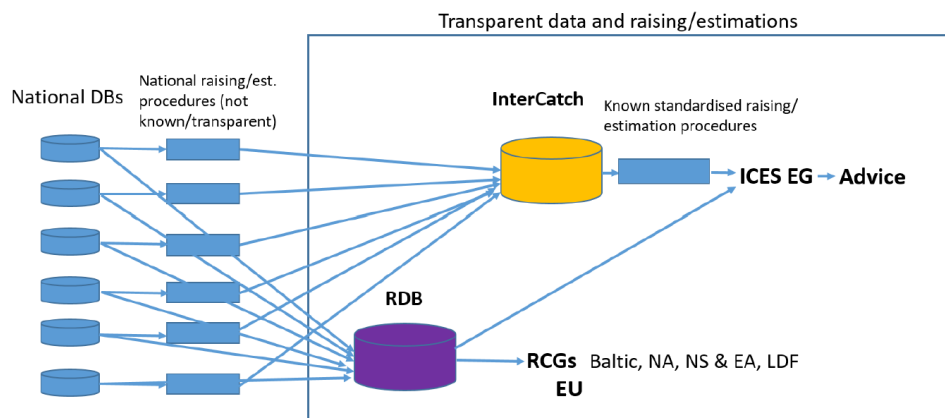


Figure 1. Scheme of the data flow from National DB to InterCatch and RDB data base.

The significant difference in the new RDBES, developed in collaboration with the RCGs, is that it provides a common structure to describe both the disaggregated sampling data and, most importantly, the statistical approach adopted. RDBES has, hence, a crucial role to play in increasing transparency and improving the quality of stock assessment within ICES (ICES, PGDATA, 2020). With the new system, there will be a standardized way on how samplings were taken to be representative of a sampling scheme and data collecting methods are now embedded in the system.

Figure 2 schematises the detailed data flow from national institutes to RCGs (NA NSEA, Baltic and LDF) and ICES Expert Groups in the new RDBES system. The data are collected through one data call to the countries. The raising/estimating of data is documented for the ICES Expert Groups and advice is done.

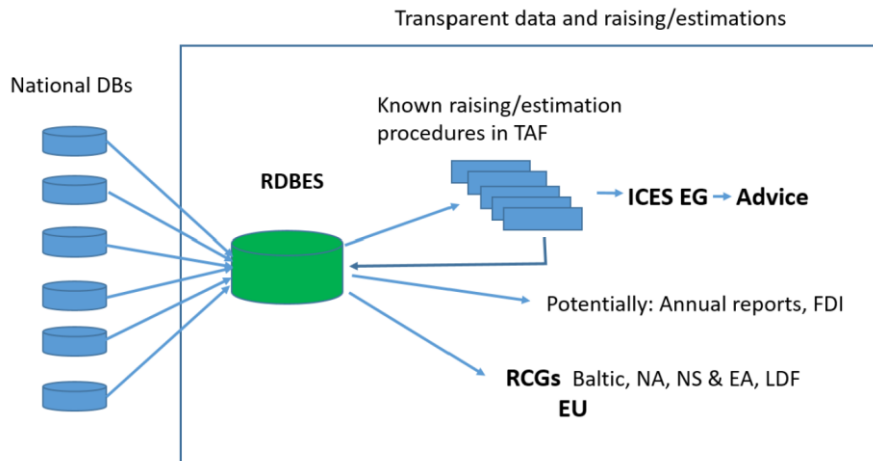


Figure 2. Scheme of the data flow when RDBES will be used in the future for commercial fisheries.

The RDBES is designed to ensure that data can be made available for the coordination of regional fisheries data sampling plans for the North Atlantic (NA), North Sea & Eastern Arctic (NS&EA), Baltic and Long Distance Fisheries (LDF) RCGs, and to provide ICES with a regional estimation system that will allow statistical estimates for stock assessment from detailed sample data in a transparent manner.

The centralised RDBES will allow the uploading of detailed data for biological data with the scripts running the raising procedures; a transparent way for sharing functions. The use of the appropriate approach for the raising procedures is under the national responsibility. Effort and landings data are at aggregated level. Whilst a lot of effort in the ICES community has been put towards developing statistically sound

sampling schemes (4S), it is recognised that design-based estimation is not currently in widespread use. There is therefore a need to help countries migrate their current raising procedures to data in the new RDBES format and using the ICES Transparent Assessment Framework (TAF)³⁴.

The RDBES database and web application/system is now implemented on a test server³⁵. The countries of the ICES community can upload data for all sampling schemes, according to 13 specified upper hierarchies. The data are allocated to 15 types of tables. The checks in the RDBES currently ensure that data is in a valid format and that valid codes have been used. The data can be exported in the same format as the uploaded RDBES format (PGDATA, 2020).

The ICES WKRDB-EST working group aims at developing R script estimations for each hierarchy. That is done by making R estimation scripts for each table, then the total estimation for a hierarchy will be a combination of the estimations belonging to the tables in the hierarchy.

Recently (May-June 2021) a workshop was held to develop improved ratio estimation methods that will be used as estimation routines in RDBES. Currently, most countries use ratio estimators for their national estimation of commercial catch data. ICES WGCATCH will evaluate the outcomes of this workshop prior to their incorporation as regular estimation routines in RDBES.

According to Roadmap for the RDBES (Report of the Liaison Meeting, 2021) it is currently planned that this new system will replace both InterCatch and RDB database systems by 2023. Milestones of the roadmap are:

- a fully operational RDBES such that statistical estimates for stock assessment can be produced from detailed sample data in a transparent manner by 2022.
- support the work of the relevant RCGs by 2022 (in parallel with existing RDB for first year)
- incorporate detailed data on Bycatch and PETS AND/OR Recreational data in the RDBES by 2023.

The RDB and InterCatch will be terminated by 2024.

Following an agreement with ICES, the existing RDBES data model will include PETS and recreational fisheries information.

The population of RDBES will require that historical data are resubmitted. A test call was launched by ICES on June 2021 for testing, after that RCGs/WGRDBESGOV/Core group had identified 19 stocks and 2 bycatch species.

The storage of VMS/ERS is not contemplated in the RDBES, ICES manages a spatial fishery DB with specific data calls but no raw VMS data are available.

Regarding data access, according to the current policy, EU Member States grant access to the RCGs for detailed and aggregated data and may also grant access to specific ICES WK groups for aggregated data. Anyone can request data and the DCF rules are followed with Member States deciding what is to be shared. Commercial fishing data are not given full access and only the data inventory is public. There is a data policy in place, which defines the level of data aggregation prior to granting access to any group. A form has been introduced that should be signed by groups requiring data, in order to ensure understanding of terms of use.

An example of an RDBES hierarchy that can be considered closer to some demersal and small pelagic biological samplings in the Mediterranean, though some other hierarchies can also work, is schematised

³⁴ Fisheries Data Collection Report 18th Liaison Meeting between the Chairs of RCGs, key end users, European Commission DG Mare and JRC. 5th November 2021.

³⁵ <https://github.com/ices-tools-dev/RDBES>

below. It is taken from the *"Documentation of the Regional Database and Estimation System. Data Model. RDBES Data Model doc. v. 1.19.2; 23 September 2021"* that is possible to download from the RDBES GitHub. This hierarchy can be applied both in case of at-sea and on-shore samplings.

Hierarchy 2: Sampling from a list of trips

Hierarchy 2 (scheme in figure 3) is most used for at-sea sampling, but can be used for on-shore sampling as well. Different sampling schemes which all fit the hierarchy 2 are:

- at-sea sampling from available fishing trips; some or all fishing operations (hauls/sets) are sampled;
- at-sea sampling from available fishing trips; aggregated fishing operations (hauls/sets) are sampled;
- on-shore sampling from available fishing trips; aggregated fishing operations (hauls/sets) are sampled when landed;
- on-shore sampling from available fishing trips; the last fishing operation (haul/set) is sampled.

A list of trips to sample from is not always available, but a vessel list is used as a proxy for selecting trips. In reality we often try to adapt sampling design where we sample from a vessel list into this hierarchy, when it may be more appropriated to use hierarchy 1. Such hierarchy however is more complex than hierarchy 2 as it includes an additional table into the upper hierarchy tables (i.e. the vessel selection table, with associated the auxiliary –mandatory- table of vessel details)³⁶.

Upper hierarchy tables:

- Design (DE)
- Fishing Trip (FT)
- Fishing Operation (FO)
- Species Selection (SS)
- Sample (SA)

Auxiliary tables:

- Sampling Details (SD) (mandatory)
- Vessel Details (VD) (mandatory)
- Species List Details (SL) (mandatory)

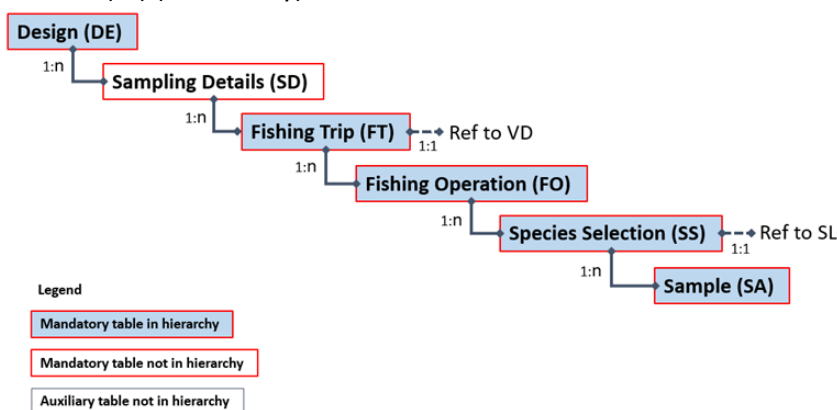


Figure 3. Example of the Hierarchy 2 of the RDBES.

³⁶ "At-sea sampling with trips as primary sampling units. When trips can be selected randomly from a fleet of vessels, at least approximately, it is often reasonable to treat vessel-trips as the primary sampling units. In such cases, the list of all trips (obtained at the end of the year) makes up the sampling frame. This is a virtual frame that cannot be used in stage 1 to select the trips. The actual selection is typically based on a frame with a vessel list crossed with time. For a fleet with day-trips this can easily be achieved by randomizing the selection of days and vessels. For fleets with varying trip-length it is more difficult to selected vessels and trips with approximately equal inclusion probabilities. It can be helpful to create strata where vessels with a similar trip length are grouped", (WKPICS, 2013 p. 41)

Finally, as regards the data from scientific surveys, it is important to recall that at ICES level, a common online database exists since long time, i.e. the Database of Trawl Surveys (DATRAS)³⁷ that has been developed to collate and document the survey data. It is a public tool that assures data quality, standardised data formats and calculations, data handling and availability. With the possibility for instant remote access, the data from DATRAS are used for stock assessments and fish community studies by the ICES community and public users.

Survey data are continuously updated by national institutions. DATRAS has an integrated quality check utility. Data products, as the ones useful for the MSFD, and raw data, can be freely downloaded according to the ICES Data policy.

In addition to DATRAS and Eggs&Larvae database, an Acoustic database is created at ICES for the acoustic surveys³⁸. This database hosts information on fisheries observations collected from various pelagic surveys coordinated by ICES. Visualization of sampling stations of Eggs&Larvae and acoustic surveys are available on the ICES website.

³⁷ <https://www.ices.dk/data/dataset-collections/Pages/Fish-trawl-survey.aspx>

³⁸ <https://www.ices.dk/data/dataset-collections/Pages/Fish-trawl-survey.aspx>

7. Mapping the main stakeholders and end users in the Mediterranean and Black Sea

An exercise for mapping of stakeholders, the potential users and their main expectations was built on the basis of different information sources and evaluations. This mapping is reported in table 2.

Table 2. Key stakeholders and end users with an outline of their main expectations.

Stakeholders&end-users	Expectations
<p>Directorate-General Maritime Affairs and Fisheries-MARE</p> <p>stakeholder and end user</p>	<p>The objective of DG MARE is to ensure availability of data for reliable advice in the Mediterranean and Black Sea that will allow the implementation of fisheries management related to the European Union Common Fisheries Policy (CFP).</p> <p>The CFP is a science-based policy. Reliable advice is based on the availability of robust data and estimates.</p> <p>Allowing end users to compute statistical estimates tailored to their needs is a further objective.</p> <p>An additional aim is to coordinate the Data Collection Framework (DCF) and facilitate the work of the EU Member States in harmonising and streamlining (and ease) the reporting on the EU data collection.</p> <p>Expectations are that RDBFIS will meet these objectives as a tool to improve data quality and accessibility and to facilitate Member States in the compliance to their obligations.</p> <p>In addition, according to the Article 18 on Compatible data storage and exchange systems of the EUMAP (Regulation (EU) 2017/1004 (recast)), Member States, the Commission, scientific advisory bodies and any relevant end-users of scientific data shall cooperate to develop compatible data storage and exchange systems, taking into account the provisions of Directive 2007/2/EC. This with a view to reducing costs and facilitating access to detailed and aggregated data for end-users of scientific data and other interested parties. Those systems shall also facilitate dissemination of information to other interested parties. Such systems may take the form of regional databases. Regional work plans referred to in Article 9(8) of the EUMAP may serve as a basis for agreement on such systems.</p>
<p>European Commission (Directorate-General Environment-DG-ENV)</p> <p>Stakeholder and end-user</p>	<p>DG-Environment coordinates the implementation of the Marine Strategy Framework Directive (Directive 2008/56/EC) and the Birds and Habitats Directives.</p> <p>The implementation of legal requirements by Member States requires fisheries data (fishery dependent data and scientific surveys data) on fish stocks, relevant ecosystem components (e.g. on PETS, macrobenthic fauna, etc.) collected in the Data Collection Framework (DCF), in particular for reaching the Good Environmental Status under the MSFD for Descriptors 1, 3, 4, 6, 10 in the region RDBFIS covers, as well as for the strict protection of species under the Birds and Habitats Directives and the management of Natura 2000 sites and other marine protected areas.</p> <p>RDBFIS can facilitate the access to checked and standardised data.</p>
<p>Member States</p>	<p>Member States (MSs) as authorities in charge of (coordinated) implementation</p>

stakeholders and end-users	<p>of the EU fisheries and environmental legislation and as key users of data would be interested in:</p> <ul style="list-style-type: none"> - sharing data, reducing the workload and costs needed for data preparation in different Data Calls; - improving the data quality in a standardised manner for the region would significantly reduce the data transmission issues; - applying commonly agreed validation rules to the datasets before data submission; - applying common rules for allowing access to end users; - reducing the workload of data preparation for the Data Calls by automatically aggregating the data in Data Call formats; - streamlining the data delivery processes/exchange between MSs, as the DCs are increasing; - ensuring that an RDB does not duplicate current tools; - reporting facilities to facilitate the compilation of tables needed for the preparation of the Work Plans and the submission of the Annual Reports for internal checks and compliance to the formats and guidelines; - applying procedures RDBES-like, to facilitate WP implementation, for those Member States operating both in the ICES area and in the Mediterranean one.
<p>Regional Coordination Group for the Mediterranean and Black Sea (RCG MED&BS)</p> <p>stakeholders and end-users</p>	<p>The RCG Med&BS considered the development of a regional database as a priority to allow:</p> <ul style="list-style-type: none"> • efficient use of the data received from the RCG data call; • appropriate management of the data used by the RCG. <p>RCGs have a central role in implementing the standards for the national sampling plans and are legitimate to provide support and advice on what is needed for common tools such as Regional Data Bases, R scripts, routines and packages.</p> <p>In addition, RCGs could be tasked to perform data analysis and produce standard and quality reports at regional level.</p> <p>Further RCGs' objectives are:</p> <ul style="list-style-type: none"> ✓ organizing dedicated workshops, ✓ establishing guidelines for quality assurance and control, ✓ identifying best practices and tools that MSs should follow. <p>All these actions can be supported by a RDB.</p> <p>RCG MED&BS interest for a regional database is also to:</p> <ul style="list-style-type: none"> - promoting the use of data at regional level supported by a tool as RDBFIS for several purposes, including the improvement of statistical samplings of biological data; - increasing the data quality to reduce the Data Transmission Monitoring Tool (DTMT) impact on MSs; - aligning the sampling procedures and estimations to the ones in RDBES, given that some countries are working both in the ICES area and in the Mediterranean; - supporting the MSs, for the preparation of the WPs and ARs according to the recent guidelines and schemes; - designing Regional Work Plans (RWP) at RCG level.

<p>Scientific, Technical and Economic Committee for Fisheries (STECF)</p> <p>stakeholders and end-users</p>	<p>STECF is considered a stakeholder and an end-user. As an independent scientific body assisting the EU Commission, it is legitimate to evaluate the performance and implementation of the DCF at MSs level.</p> <p>STECF EWGs evaluate the data, which quality is pivotal for robust and reliable stock assessments and advice and evaluation of management plans, including landing obligation. STECF EWGs evaluate Work Plans (WPs), Annual Reports (ARs) and Annual Economic Report (AER).</p> <p>STECF can expect that web-based RDBFIS application to support the preparation, management and assessment of the DCF data would ensure efficiency and transparency in the data transmission process, also decreasing the number of data transmission issues.</p> <p>RDBFIS can provide support in terms of standardised data checks, and transparency of the sampling and raising procedures, when an RDBES-like approach will be implemented, at least for some hierarchies as defined in the RDBES.</p>
<p>Joint Research Center (JRC)</p> <p>end-users</p>	<p>In case JRC will continue analysing the quality and coverage of data submitted by MSs under the DG MARE Med&BS and FDI Data Calls and making data available to the STECF working groups, it would be interested into a tool as RDBFIS that can facilitate this process with Member States.</p> <p>JRC would be interested in improving methodology to assess data quality e.g. in terms of data completeness (temporal and spatial coverage), but also in terms of transparency as regards the raising procedures from samplings.</p> <p>If the further development of the RDBFIS will go in the direction of a RDBES like tool the advantage would be to get available basic information on the methodology applied for raising procedures that, in turn, will be also an advantage for EWGs.</p>
<p>General Fishery Commission for the Mediterranean (GFCM) and Scientific Advisory Committee on Fisheries (SAC)</p> <p>stakeholders and end-users</p>	<p>GFCM provides binding Recommendations on several aspects for fishery management and conservation of fisheries resources at Mediterranean and Black Sea levels. The mandate of SAC is to provide independent advice on the technical and scientific basis for decisions related to fisheries conservation and management, including biological, social and economic aspects.</p> <p>Thus, GFCM seeks for high quality information on catches and fishing effort as well as on other relevant data for the conservation and management of fishery resources.</p> <p>SAC regularly reviews the assessments carried out by the expert WGs, then it would be interested that checked datasets and data quality reports are available to support WGs running stock assessment models and advices.</p> <p>In addition, there is a wide interest in GFCM for conservation aspects related to the mitigation of fishing on incidental catch of vulnerable species, PETs and VMEs.</p>
<p>International Council for the Exploration of the Sea (ICES)</p> <p>stakeholders and end-users</p>	<p>ICES can be interested in making tools available aimed at the RDBES data, that can be applied to the detailed RDBES data to perform analysis of the uploaded detailed data. ICES is also working on developing an estimation package written in R aimed at the fish stock assessment working groups and the RDBES data, where the catch and age and length distributions are estimated using statistical</p>

	<p>estimation methods. ICES can also be interested in making such estimation package available.</p> <p>Several WGs are organised in ICES on a large number of topics involving assessments that support the implementation of MSFD, as the working group on Fisheries Benthic Impact and Trade-offs (WGFBIT) that develops methods and performs assessments to evaluate benthic impact of fisheries at regional scale, using data from different sources, including trawl surveys and fishing footprints. Similar considerations hold for the working group on bycatch of protected species (WGBYC) that launched data calls requesting data on fishing effort, monitoring effort and PETS.</p> <p>Tools like RDBES have the role of supporting such uses.</p>
<p>Expert Working Groups (WGs) of STECF, GFCM, ICES, RCGs and International Conventions</p> <p>end-users</p>	<p>Experts attending WGs on stock assessment may be interested in a direct access to all the information needed to run any stock assessment models. For example, based on the length/age structures of the catches and total production by métier/gear, biological parameters, catch per unit effort, spatial allocation of fishing effort and origin of the catches. All this can be facilitated by the support of a web-based tool as RDBFIS.</p> <p>GFCM is also developing systems to share outcomes of validated stock assessments online. Clear systems and mechanisms would have to be put in place to ensure the agreement of non-EU countries covered by the Data Collection Reference Framework (DCRF) of the GFCM and not the DCF.</p> <p>Experts attending WGs on stock assessment may be interested in reporting facilities to automatize and standardize the compilation of data quality report including information related to e.g., the coverage of the sampling, the methodology used for estimations of the parameters, related CVs and the sampling procedures. Experts are also interested in reducing the burden of such reporting, thus relying upon the support of a tool as RDBFIS. It would be advisable to interact with the data submission requirements of the GFCM through the DCRF.</p> <p>Experts attending WGs on stock assessment, in particular STECF EWGs are in charge to identify the Data Transmission issues that have an impact on the assessments. A tool such as RDBFIS would significantly reduce the occurrence of errors. This would simplify and speed up the process of verification of MSs compliance with their obligations to collect and make available DCF data.</p> <p>Experts are also interested in better understanding the effects of sampling procedures and estimations on the assessment results. This would improve their advice, besides improving the transparency of the evaluation process.</p> <p>Institutional Bodies working for the objectives of the Marine Strategy Framework Directives and the achievement of Good Environmental Status can be interested on data of ecosystem components (e.g. on PETS).</p>
<p>Non-Governmental Organizations (NGOs)</p> <p>stakeholders</p>	<p>A number of NGOs and public initiatives are monitoring fisheries closely around the world and are raising awareness among the general public. Hence, they would be interested in having access to datasets and data quality reports for producing documents/outputs informing the general public on fisheries status, also making their own assessments/evaluations.</p>

8. Current functionalities and data uses

Based on the listed stakeholders and potential users from the previous sections, the current functionalities and data uses can be schematised as follows.

The information to be stored in a regional database, such as RDBFIS, can be grouped in three main components, each one reflecting a different level of aggregation/confidentiality according to the definitions in the Regulation (EU) 2017/1004 ³⁹:

- i) **Meta data**, i.e. data giving qualitative and quantitative information on the collected primary data. Metadata are, for example, those reported in the tables of the National Work Plan and Annual Report of the DCF;
- ii) **Detailed data**, i.e. data based on primary data in a form that does not allow natural persons or legal entities to be identified directly or indirectly. Data at this aggregation level are, for example, the ones delivered to the RCG Med&BS data call
- iii) **Aggregated data**, i.e. the output resulting from summarizing the primary or detailed data for specific analytical purposes. Data at this aggregation level are, for example, the ones delivered to the DG MARE Med&BS, FDI, GFCM data calls.

In Table 3 the main functionalities that a Med&BS RDB should foresee are reported with indication of the potential user, the data aggregation level, the category (among *data check*, *data import*, *data transformation*, *data analyses* and *data export*) and the functionality description. If IT tools (R script/package, function implemented on web platform, etc.) are already in use in Mediterranean and Black Sea context to support the implementation of the listed functionalities, a note on the name/source of the tool is reported in the last column.

Table 3– List of the main functionalities of a Med&BS-RDB. Links with the Data Calls tables or an IT system are highlighted

RDB Users	Data aggregation level	Functionality	Description of functionality and link with the Data Calls tables or IT system	Available in existing IT systems in Mediterranean context
National data submitters/owners	Detailed data	Data check	Perform quality checks on collected detailed data before submission RCG format: CS table	YES. External tool for data quality checks: fishPifct R package based on functions from SDEFQuality R package (developed under MARE/2014/19 Med&BS project) and on functions developed under FishPi project; COST functions for data exploration. Scripts on a priori quality checks developed under Task 6.1 of STREAM project. Additional data quality functions developed in RDBFIS project: visual check of haul and trip position, summary on the number of individual by trip for which biological data have been collected (length, sex, maturity, weight and

³⁹Regulation (EU) 2017/1004 of the European Parliament and of the Council of 17 May 2017.

				age), summary on the number of trips/hauls monitored by year by port, metier, sampling method.
National data submitters/owners	Detailed data	Data import	Submit detailed data uploading the tables on the IT System	NO. Datasets in Excel format are sent by email to the RCGs chairs When RDBFIS will be completed and populated, the data could be extracted directly from the database.
National data submitters/owners	Aggregated data - DG MARE MED&BS Data Call	Data check	Perform quality checks on tables required by DG MARE MED&BS Data Call before submission Med&BS Data Call tables: Catch, Landings, Discards, ALK, GP, SRL, SRA, ML, MA.	YES. When the National Correspondent uploads the data on the JRC website, automatic checks are carried out and a logfile is produced; STECF EWG 21-02 scripts on data quality. External tool for data quality checks: - scripts on a posteriori quality checks developed under Task 6.1 of STREAM project. These scripts have been integrated and included in an R package for quality checks that will be embedded in RDBFIS. - Additional functions developed in RDBFIS project: checks concerning the number of sampled trips (for landing and discard), number of lengths measured, number of age measurements.
National data submitters/owners	Aggregated data - DG MARE MED&BS Data Call	Data import	Upload the tables of DG MARE MED&BS Data Call on the IT System	YES. Facility already available on JRC Web site to the MSs delivering datasets.
National data submitters/owners	Aggregated data - DG MARE FDI Data Call	Data analyses and transformation	Convert the length and age structures included in the tables delivered to the DG MARE Med&BS Data Call into DG MARE FDI Data Call format FDI Data Call: Table A. Catch summary	YES. Auxiliary script developed under the task 3.2 of STREAM project for the conversion of the DG MARE Med&BS Data Call to the DG MARE FDI Data Call format This script will be updated according to the new data call specifications and included in an R package that will be embedded in RDBFIS. Currently, only Tables A, B, G, H, I, J are required for Mediterranean Sea.
National data submitters/owners	Aggregated data - DG MARE FDI Data Call	Data check	Perform quality checks on tables required by DG MARE FDI Data Call before submission FDI Data Call: Tables A, B, G, H, I, J	YES. When the National Correspondent uploads the data on the JRC website, automatic checks are carried out and a logfile is produced. Specific functions developed in RDBFIS project: e.g. temporal/spatial coverage, consistency of effort between G

				and I, consistency of landings between A and H. The report of the EWG 21-12 is taken into account for the list of checks. Currently, only Tables A, B, G, H, I, J are required for Mediterranean Sea. In case other FDI tables will be required in the future, other check functions will be needed.
National data submitters/owners	Aggregated data - DG MARE FDI Data Call	Data import	Upload the tables of DG MARE FDI Data Call on the IT System	YES. Facility already available on JRC Web site to the MSs delivering datasets.
National data submitters/owners	Aggregated data - GFCM/DCRF Data Call	Data analyses and transformation	Convert the length and age structures included in the tables delivered to the DG MARE Med&BS Data Call into GFCM Data Call format	YES. Auxiliary scripts developed under the task 3.2 of STREAM project for the conversion from the DG MARE Med&BS Data Call to the GFCM/DCRF Data Call format. This script will be updated according to the new Data Call specifications and included in an R package that will be embedded in RDBFIS.
National data submitters/owners	Aggregated data - GFCM/DCRF Data Call	Data check	Perform quality checks on tables required by GFCM Data Call before submission GFCM DCRF table: Task II.2, Task III, Task VII.2, Task VII.3.1, Task VII.3.2.	YES. Excel files to perform the quality checks on GFCM/DCRF Data Call are available on the GFCM platform to the MSs delivering the datasets. Specific functions developed in RDBFIS project: empty fields, duplicated records, consistency maturity, consistency length data, check of missing combination GSA/Fleet segment per year, check mismatching species/Catfau and sex per maturity stages, consistency length-weight data.
National data submitters/owners	Aggregated data - GFCM/DCRF Data Call	Data import	Upload the tables of GFCM Data Call on the IT System	YES. Facility implemented into the DCRF online platform, available within the GFCM extranet
Data check and data analyses by experts	Detailed data	Data check	Verify the coverage of the sampling; calculate data quality indicators (precision, accuracy, representativeness, completeness and comparability) RCG format: CS table	YES. Some functions implemented in COST package allows to investigate data e.g. checking completeness, calculating precision, verifying the representativeness of the sampling against the total production. Some (temporal, spatial, technical) coverage functions developed under RDBFIS project: number of trips monitored by metier, year, port and sampling method; number of individual measurements collected for length, weight, maturity, sex and age.
Stock assessment	Aggregated	Data check	Verify the coverage	YES. Facility implemented into the

experts	data - DG MARE and GFCM/ DCRF Data Calls		and timeliness of the database Med&BS Data Call tables: Catch, Landings, Discards, ALK, GP, SRL, SRA, ML, MA.	DCRF online platform, available within the GFCM extranet. Also on the JRC web site reports on coverage in relation to the data calls under its responsibility are available (STECF EWG 20-08; 21-02; 22-03),. An additional JRC report was available as result of specific checks carried during at data preparation meeting for the stock assessments (STECF EWG 21-02).
National data submitters/owners	Scientific survey data - DG MARE MED&BS Data Call (MEDITS and Black Sea surveys)	Data check	Perform quality checks on TA, TB and TC tables required by DG MARE MED&BS Data Call before submission	YES. When the National Correspondent uploads the data on the JRC website, automatic checks are carried out and a logfile is produced. RoME package is available for the systematic data quality checks. A new version of RoME to be included in RDBFIS is under development. RoMEBS package will be updated for the check of Black Sea surveys and embedded in RDBFIS.
National data submitters/owners	Scientific survey data - DG MARE MED&BS Data Call	Data import	Upload the tables of DG MARE MED&BS Data Call on the IT System	YES. Facility already available on JRC Web site to the MSs delivering datasets
National data submitters/owners	Scientific survey data - DG MARE	Data analyses and transformation	Estimation of abundance indices from scientific survey	YES. Auxiliary scripts developed by JRC for the stock assessment purpose and available during the STECF EWG 21-02.
National data submitters/owners	Aggregated data	Data check	Perform quality checks on collected detailed data before submission RCG format: CL table	New data quality functions developed in RDBFIS project: Spatial and temporal coverage of landing and landing values.
National data submitters/owners	Aggregated data	Data import	Submit aggregated data uploading the tables on the IT System	NO. Datasets in Excel format are sent by email to the RCGs chairs. When RDBFIS will be completed and populated, the data could be extracted directly from the database.

Conclusions

In the RDBFIS development, the main existing tools in support to the DCF, such as those developed in MARE/2014/19 Med&BS and STREAM regional grants, JRC tools and FDI checking tools, R routines for checking trawl surveys data, will be considered and incorporated, where appropriate, with suitable adaptations.

Several R tools have been designed so far to develop quality checks (e.g., for trawl survey data; in STREAM project and new tools under development in STREAMLINE project) to be applied both on detailed and aggregated data. In the RDBFIS project, these tools will be generalized to facilitating the adaptation to changes in data call formats, for using algorithms already implemented in STREAM (e.g. raising of sampling data), and for implementing RDBES-like procedures. In addition, new tools for quality checks are in development.

RDBFIS will be a web-based centralized platform where data can be uploaded and data quality checks performed using common procedures.

The consultation with MS highlighted that:

1. all Member States, except Cyprus, have established databases to store and analyze the data collected in the framework of DCF and report to several data calls;
2. the MSs underlined the importance of the RDBFIS grant;
3. a key aspect is that RDBFIS will ensure common procedures for data quality checks;
4. the needs of the MS have to be considered in the RDBFIS (the case of Cyprus to use SDEF format and RDBES hierarchies are considered useful);
5. RDBES structure, hierarchies or algorithms are not yet included or contemplated in most of the Med&BS MSs established information systems.

Regarding detailed biological data (biological samplings and biological parameters) of demersal and small pelagic species, RDBFIS aims at allowing the storage of detailed information on the sampling scheme, sampling frame strata and hierarchy followed in the collection of the detailed biological data in line with RDBES initiatives. Thus, a RDBES-like structure will be incorporated in the RDBFIS, also taking into consideration the outcomes of previous RDB SC meetings. This will allow in the future to describe both the disaggregated sampling data and, most importantly, elements of the sampling design.

Envisage in RDBFIS also a RDBES-like structure would increase the compatibility of the results among the European sea basins and a more synchronized approach in data storing and processing. This would be particularly useful for countries in which both the ICES and Mediterranean communities coexist, as in Spain and France.

However, it is pivotal that RDBFIS is able to store the detailed data also in the RCG CS (sampling) and CL (landing) formats, used so far as exchange format, and including the COST data models and functions (that make use of Age-Length keys, ratio estimators, and unbiased designed based estimations). These COST functions are also used in the RDBES. This will allow an easy use of the information already available and a certain flexibility in the meanwhile statistical concepts behind the RDBES data format, such as the implementation of 4S sampling design, are fully developed by all the Member States of the Mediterranean and Black Sea. As foreseen by the RDBFIS project, this will allow to properly take into consideration the more appropriate “hierarchies” and estimation methods according to the approaches developed in the RDBES.

Regarding data from scientific surveys, a common online database exists at ICES level, i.e. the Database of Trawl Surveys (DATRAS). This database has been developed to collate and document the survey data. It is a public tool that assures data quality, standardised data formats and calculations, data handling and availability.

The mapping exercise included 10 main stakeholders and end users. The more common expectations can be summarised as follows:

- interest in sharing data and reducing the workload and costs needed for data preparation for the different Data Calls;
- improve data quality in a standard manner for the region, significantly reducing the data transmission issues;
- implement the standards for the national sampling plans;
- perform data analysis, establish guidelines for quality assurance and control, identify best practices and tools;
- align the sampling procedures and estimations to the ones in RDBES;
- improve statistical sampling of biological data;
- support the preparation of WPs and ARs, accounting for the recent guidelines and schemes, and preparation of Regional Work Plans (RWP);
- support the evaluations of relevant ecosystem components (e.g. PETS) collected in the Data Collection Framework (DCF) but also linked to the Marine Strategy Framework Directive;
- robust and reliable stock assessments and advice, evaluation of management plans, including the landing obligation, for a sound implementation of the CFP.

Seventeen main functionalities were identified indicating the potential user, the data aggregation level, the category (among data check, data import, data transformation, data analyses and data export) and the functionality description.

The expected tangible outcomes of building Med&BS RDBFIS are⁴⁰:

- database structure, referential integrity;
- common codification system, common exchange format;
- upload procedures, advanced data mining subsystem;
- validation procedures;
- data processing tools to support specific data calls (including VMS analysis procedures, estimation of fishing pressure from small scale fisheries using a Multi-Criteria Decision Analysis);
- automatic reporting tools (linked to the DCF processes);
- a state-of-the-art web-based user interface to interact with all processes (upload data, validate data, report data, search, compare, compile, aggregate, plot, visualize data etc);
- compatibility with RDBES data model.

In conclusion, the added value of RDBFIS implementation will consist of providing a tool to support the work of the RCG Med&BS and Member States, improving their performance towards faster data processing and response to data calls, increased quality and robustness of data delivered to end-users, thus strengthening regional and EU-wide cooperation on data collection.

⁴⁰ *Report of the 17th Liaison Meeting. Meeting between the Chairs of RCGs, PGECON, key end users, JRC and the Commission. Webex24 - 25 September 2020*