



JRC SCIENCE FOR POLICY REPORT

Scientific, Technical and Economic
Committee for Fisheries (STECF)

–

Compilation of the new DCF
Annual Report template
(STECF-17-17)

Edited by Matthew Elliott, Jon Elson,
John Casey & Hendrik Doerner

This publication is a Science for Policy report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policy-making process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

Contact information

Name: STECF secretariat

Address: Unit D.02 Water and Marine Resources, Via Enrico Fermi 2749, 21027 Ispra VA, Italy

E-mail: stecf-secretariat@jrc.ec.europa.eu

Tel.: +39 0332 789343

JRC Science Hub

<https://ec.europa.eu/jrc>

JRC109375

EUR 28359 EN

PDF	ISBN 978-92-79-67489-1	ISSN 1831-9424	doi:10.2760/471735
-----	------------------------	----------------	--------------------

STECF	ISSN 2467-0715
-------	----------------

Luxembourg: Publications Office of the European Union, 2017

© European Union, 2017

Reuse is authorised provided the source is acknowledged. The reuse policy of European Commission documents is regulated by Decision 2011/833/EU (OJ L 330, 14.12.2011, p. 39).

For any use or reproduction of photos or other material that is not under the EU copyright, permission must be sought directly from the copyright holders.

How to cite: Scientific, Technical and Economic Committee for Fisheries (STECF) – Compilation of the new DCF Annual Report template (STECF-17-17). Publications Office of the European Union, Luxembourg, 2017, ISBN 978-92-79-67489-1, doi:10.2760/471735, JRC109375

All images © European Union 2017

Abstract

Commission Decision of 25 February 2016 setting up a Scientific, Technical and Economic Committee for Fisheries, C(2016) 1084, OJ C 74, 26.2.2016, p. 4–10. The Commission may consult the group on any matter relating to marine and fisheries biology, fishing gear technology, fisheries economics, fisheries governance, ecosystem effects of fisheries, aquaculture or similar disciplines.

Authors:**STECF advice:**

Ulrich, C., Abella, J. A., Andersen, J., Arrizabalaga, H., Bailey, N., Bertignac, M., Borges, L., Cardinale, M., Catchpole, T., Curtis, H., Daskalov, G., Döring, R., Gascuel, D., Knittweis, L., Malvarosa, L., Martin, P., Motova, A., Murua, H., Nord, J., Prellezo, R., Raid, T., Sabatella, E., Sala, A., Scarcella, G., Soldo, A., Somarakis, S., Stransky, C., van Hoof, L., Vanhee, W., Vrgoc, Nedo.

EWG-17-17 (Part 1) report:

Angeliki ADAMIDOU, Angeles ARMESTO, John CASEY, Bram COUPERUS, Henrik DEGEL, Matthew ELLIOTT, Emmet JACKSON, Edvardas KAZLAUSKAS, Estanis MUGERZA, Alastair POUT, Jose RODRIGUEZ, Maria Begoña, SANTOS VAZQUEZ, Alan WALKER.

EWG-17-17 (Part 2) report:

David Currie, Dimitris Damalas, Henrik Degel, Laurent Dubroca, Jon Elson, Maria Teresa Facchini, Jerome Guitton, Jenny Nord

TABLE OF CONTENTS

SCIENTIFIC, TECHNICAL AND ECONOMIC COMMITTEE FOR FISHERIES (STECF) - Compilation of the new DCF Annual Report template (STECF-17-17)	7
1.1 Request to the STECF	7
1.2 STECF response	7
1.3 STECF observations	7
1.4 STECF conclusions	9
1.5 Contact details of STECF members	10
2 Expert Working Group EWG-17-17 report	14
1 Introduction	15
1.1 Terms of Reference for EWG-17-17 (Parts 1 and 2)	15
2 Report of EWG 17-17 Part 1	19
2.1 AR Section 1: Biological data	21
2.1.1 Table 1A, List of required stocks; Table 1B, Planning of sampling for biological variables; Table 1C, Sampling intensity for biological variables	21
2.1.2 Table 1D: Recreational fisheries	22
2.1.3 Pilot Study 1: Relative share of catches of recreational fisheries compared to commercial fisheries	22
2.1.4 Anadromous and catadromous species data collection in fresh water: Table 1E and Text Box 1E	22
2.1.5 Table 1F and Table 1 F(a): Incidental by-catch of birds, mammals, reptiles and fish (Data to assess incidental by-catch and other impacts of Union fisheries on marine ecosystems)	23
2.1.6 Pilot Study 2: Level of fishing and impact of fisheries on biological resources and marine ecosystem	26
2.1.6.1 Data quality evaluation	26
2.1.7 Table 1G: List of research surveys at sea	26
2.1.8 Text Box 1G: List of research surveys at sea	27
2.1.9 Table 1H: Research survey data collection and dissemination	28
2.1.9.1 Data quality issues	28
2.2 AR Section 2: Fishing activity data - Table 2A: Fishing activity variables data collection strategy and Text Box 2A: Fishing activity variables data collection strategy	28

2.3	AR Section 3: Economic and social data (Fleet, aquaculture and processing)	29
2.3.1	Tables and Text Boxes 3A, 3B and 3C: Population segments for collection of economic and social data for fisheries, aquaculture and the processing industry and Pilot Study 3: Data on employment by education level and nationality	29
2.3.1.1	Quality evaluation in Economic and social variables.	29
2.3.1.2	Selection of quality indicators for AR templates.....	30
2.3.2	Pilot Study 4: Environmental data on aquaculture.....	30
2.4	AR Section 4: Sampling strategy for biological data from commercial fisheries - Table 4A: Sampling plan description for biological data and Text Box 4A: Sampling plan description for biological data	31
2.5	Table 4B: Sampling frame description for biological data	32
2.6	Table 4C: Data on the fisheries by Member State.....	32
2.7	Table 4D: Landing locations	32
2.8	AR Section 5: Data quality - Table 5A: Quality assurance framework for biological data and Table 5B: Quality assurance framework for socioeconomic data	33
2.9	AR Section 6: Data availability (Table 6A: Data availability)	33
2.10	Section 7: Coordination	33
2.10.1	Table 7A: Planned regional and international coordination	33
2.10.2	Table 7B: Follow-up of recommendations and agreements.....	33
2.10.3	Table 7C: Bi- and multilateral agreements.....	34
3	Report of EWG 17-17 Part 2	35
3.1	Template scrutiny and automated checks.	36
	Table 1A.....	41
	Table 1C.....	42
	Table 4A.....	43
	Summary	43
3.2	Automation	44
3.2.1	Current examples of international data exchanges using automated checks.....	45
3.2.1.1	Fisheries dependent information (FDI) validation tool	45
3.2.1.2	The regional database FishFrame	46

3.2.1.3	Intercatch.....	46
3.2.1.4	GFCM Data collection Reference Framework online platform	46
3.2.1.5	FishPi tools.....	46
3.2.2	Future automation of the AR and NWP evaluation cycle.....	46
3.2.2.1	Using Excel spreadsheets for validation	47
3.2.2.1.1	Cover sheet.....	48
3.2.2.1.2	Drop-down Menus	48
3.2.2.1.3	Automatic fill of cells	49
3.2.2.1.4	Indications that further information is required.....	49
3.2.2.1.5	Relation to the NWP	50
3.2.2.2	Using R for validation	51
3.2.2.2.1	Framework	51
3.2.2.2.2	Test	52
3.2.2.2.3	Data	52
3.2.2.2.4	Quality checks on table 1A	53
3.2.2.3	Using Web app for validation	56
3.2.2.3.1	Introduction:	57
3.2.2.3.2	Step 1: moving Excel information to a common database.....	57
3.2.2.3.3	Step 2: adding reference tables to the database.....	58
3.2.2.3.4	Step 3: from check list to SQL query	59
3.2.2.3.5	Step 4: from one query to a report.	60
3.2.2.4	Utilising the RDB as a source of metadata for STECF evaluation of National Work plans.....	64
3.3	References	69
4	Contact details of EWG-17-17 participants	70
5	List of Annexes.....	72
6	List of Background Documents.....	72
Appendix 1 - Example of the CheckTemplate_Alltables.xlsx filtered for Table 1C.....		73
Appendix 2 - Using R for validation		74

SCIENTIFIC, TECHNICAL AND ECONOMIC COMMITTEE FOR FISHERIES (STECF) - Compilation of the new DCF Annual Report template (STECF-17-17)

1.1 Request to the STECF

STECF is requested to review the report of the STECF Expert Working Group meeting, evaluate the findings and make any appropriate comments and recommendations

1.2 STECF response

The tasks of the STECF Expert working group (EWG 17-17) were carried out in two consecutive meetings, part 1 and part 2. Part 1 defined the Annual Report templates and a guidance document for Member States on how to fill them and for STECF on how to evaluate them. Part 2 dealt with the testing of procedures for automatic checking of the consistency between the Work Plans (WP) and the Annual Reports (AR). STECF EWG 17-17 part 1 met 16-19 of October and part 2, 23-26 of October 2017. Both meetings were held in Brussels.

Article 11 of the DCF Regulation (EC) 2017/1004 requires Member States to submit annually to the Commission a report on the implementation of their national work plans (WP). A new Annual Report (AR) template is needed from 2018, to allow comparison of implementation against planning of the data collection as described in the WPs.

The Terms of Reference of the meetings were to produce the new Annual Report template for submission by MS, guidelines to be followed by MS and for STECF in their evaluation as well as to define automatic checks of submitted information in the AR. In addition, the EWG was asked to highlight information that may be missing and how these gaps could be addressed through future reporting.

The final reports of the meetings were not available in time for the plenary. The STECF advice is therefore based on the draft versions as of the 6 of November 2017 together with a presentation and discussion with the chair of EWG part 1.

1.3 STECF observations

STECF observes that the outcomes of the meetings consist of a EWG report, Excel tables for the submission of Annual Reports (AR) by MS, a Guidance document to facilitate the submission and evaluation of ARs, as well as a *CheckTemplate* of defined automatic checks for each table of the AR. The EWG suggests three approaches (further explained below) to be used independently or in combination that could be adopted for applying the checks of the *CheckTemplate*.

Draft new Annual Report (AR) template

STECF observes that the draft new Annual Report (AR) template mirrors the Work Program (WP). This enables checking for conformity between the AR and WP, which is an important improvement from previous annual reports (STECF EWG 17-04). STECF notes that the EWG was asked to follow the legally binding structure of the WP but since this structure can potentially be revised in the legislation after 2019, the EWG also suggested changes and modifications to the WP.

STECF notes that both AR and WP development could be seen as a process of continuous improvement. On this end the EWG put forward an additional data table for the AR. The aim of this new Table 1F(a) is to provide an overview of other data collection performed

by Member States that could be used for future determination of the impact of fisheries on the marine ecosystem (excluding incidental by-catch that is reported under 1F). The data asked in this table is collected under the control regulation (EC) 1224/2009 (VMS, logbook information etc.). Additionally, in the case MS carries out stomach sampling under some dedicated sampling programs, this information should also be stated here.

Guidance for submitters and evaluators of the AR

The guidance document for submitters and evaluators of the AR was considered useful by the national experts attending the EWG. STECF notes thus that this document is expected to help Member States filling in each section of the AR. It also contains a specific section for evaluators on what aspects to check for in the evaluation. Furthermore, it provides guidance for MS and evaluators on how to fill the text box related to quality assurance of data.

CheckTemplate and automatic checking

STECF observes that a list of different checks to be applied for the AR (completeness, timeliness, internal consistency etc.) was produced. From this a *CheckTemplate* spreadsheet was produced, listing the different checks for each field of each AR table. The *CheckTemplate* provides the guidelines for programmers to develop a system for data checking of the AR.

STECF notes that three approaches are proposed (to be used independently or in combination) that could be adopted for applying the checks stated in the *CheckTemplate*:

1. *Excel spreadsheets* including examples of different functionalities (drop down menus, automatic fill cells etc.)

2. *R code for validation*

An open-source and transparent tool to end-users was developed.

3. *Web based application*

A database and a web based application was set up as a trial.

The EWG provided worked examples and options of how automated checking can be further developed. It discussed the pros and cons of the different systems but did not advice on what approach the Commission should move towards implementing.

Regional database

STECF notes that detailed biological sampling data from three Regional Coordination Groups (North Atlantic, North Sea & Eastern Arctic and the Baltic), is stored in a common format in the regional database. STECF further notes that for the Mediterranean and Black Sea region, the implementation of the regional database is still under discussion. There is currently no regional database for the RCG for large pelagics. STECF notes that the regional databases provide a very useful tool to facilitate MS producing tables for the Annual Report.

1.4 STECF conclusions

STECF concludes that the STECF EWG 17-17 report referring to both part 1 and part 2 adequately addresses all Terms of References. STECF endorses the outcomes of the EWG. In addition, the STECF discussed the following:

STECF concludes that the draft new Annual report (AR) template allows for assessment of conformity through the mirroring of the WP. Regarding reporting on quality of Economic variables, the outcomes of the data quality subgroup of PGECON should be used as a reference. The guidance document would benefit from a hyperlink to the Eurostat ESS standard for quality report as well as the Quality Guidelines for the DCF ([Moura, 2016](#)). For biological information, the STECF EWG 17-04 has provided a Quality Assurance Framework based on European standards.

STECF concludes that the *CheckTemplate* for each field of all tables provides the basis for programmers to construct automatic checks of the ARs. STECF highlights that it is important to find a balance between flexibility and user-friendliness for the submitters and the necessary consistency between required and submitted information when constructing automatic checks.

STECF concludes that the guidance document provides useful guidance for both MS in their submission of ARs and for the STECF as evaluators of the AR.

STECF reiterates its opinion (from STECF EWG 17-04) and supports the recommendation from the RCG (ToR 6.9) that regional databases should be used to facilitate MS producing tables for the Annual Report. In addition the regional databases could be used for a number of purposes, including:

- Make pre-written RDB data extraction routines available to MS so they can insert the data into the AR template themselves,
- Directly cross-checking data submitted for the Annual Report,
- Providing a complementary data source and reports for Annual Report evaluators to use.
- Providing automatically-generated overviews on e.g. sampling coverage, which provides useful generic information to end-users as part of the quality assessment.

STECF concludes that the two main unresolved issues of the EWG of i) the sampling strategy for biological data from commercial fisheries (tables 4a and d) and ii) data to assess impacts of Union fisheries on marine ecosystems (table 1F) should be further explored by the STECF EWG 17-13.

STECF concludes that the Commission needs to ensure that a workable solution for the automatic checking can come into place in due time before the submission of the AR 2018.

STECF concludes that a database with a web-based application would likely be the preferred option for submission and automatic checking of ARs, as this would provide more flexible functionalities, such as direct comparisons between the Work Plans (WPs) and the Annual Report (AR), consistency checks between years, submission of additional data without the need to resubmit the entire AR etc.

1.5 Contact details of STECF members

¹ - Information on STECF members' affiliations is displayed for information only. In any case, Members of the STECF shall act independently. In the context of the STECF work, the committee members do not represent the institutions/bodies they are affiliated to in their daily jobs. STECF members also declare at each meeting of the STECF and of its Expert Working Groups any specific interest which might be considered prejudicial to their independence in relation to specific items on the agenda. These declarations are displayed on the public meeting's website if experts explicitly authorized the JRC to do so in accordance with EU legislation on the protection of personnel data. For more information: <http://stecf.jrc.ec.europa.eu/adm-declarations>

Name	Address ¹	Tel.	Email
STECF members			
Abella, J. Alvaro	Independent consultant	Tel. 0039-3384989821	aabellafisheries@gmail.com
Andersen, Jesper Levring	Department of Food and Resource Economics (IFRO) Section for Environment and Natural Resources University of Copenhagen Rolighedsvej 25 1958 Frederiksberg Denmark	Tel.dir.: +45 33 33 68 92	jla@ifro.ku.dk
Arrizabalaga, Haritz	AZTI / Unidad de Investigación Marina, Herrera kaia portualdea z/g 20110 Pasaia (Gipuzkoa), Spain	Tel.: +34667174477	harri@azti.es
Bailey, Nicholas	Marine Scotland Science, Marine Laboratory, P.O Box 101 375 Victoria Road, Torry Aberdeen AB11 9DB UK	Tel: +44 (0)1224 876544 Direct: +44 (0)1224 295398 Fax: +44 (0)1224 295511	baileyn@marlab.ac.uk n.bailey@marlab.ac.uk
Bertignac, Michel	Laboratoire de Biologie Halieutique IFREMER Centre de Brest BP 70 - 29280 Plouzane, France	tel : +33 (0)2 98 22 45 25 - fax : +33 (0)2 98 22 46 53	michel.bertignac@ifremer.fr
Borges, Lisa	FishFix, Brussels, Belgium		info@fishfix.eu
Cardinale, Massimiliano (vice-chair)	Föreningsgatan 45, 330 Lysekil, Sweden	Tel: +46 523 18750	massimiliano.cardinale@slu.se
Catchpole, Thomas	CEFAS Lowestoft Laboratory, Pakefield Road, Lowestoft Suffolk, UK NR33 0HT		thomas.catchpole@cefas.co.uk

Name	Address ¹	Tel.	Email
STECF members			
Curtis, Hazel	Sea Fish Industry Authority 18 Logie Mill Logie Green Road Edinburgh EH7 4HS, U.K.	Tel: +44 (0)131 524 8664 Fax: +44 (0)131 558 1442	Hazel.curtis@seafish.co.uk
Daskalov, Georgi	Laboratory of Marine Ecology, Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences	Tel.: +359 52 646892	Georgi.daskalov@gmail.com
Döring, Ralf (vice-chair)	Thünen Bundesforschungsinstitut, für Ländliche Räume, Wald und Fischerei, Institut für Seefischerei - AG Fischereiökonomie, Palmaille 9, D-22767 Hamburg, Germany	Tel.: 040 38905-185 Fax.: 040 38905-263	ralf.doering@thuenen.de
Gascuel, Didier	AGROCAMPUS OUEST 65 Route de Saint Brieuc, CS 84215, F-35042 RENNES Cedex France	Tel: +33(0)2.23.48.55.34 Fax: +33(0)2.23.48.55.35	Didier.Gascuel@agrocampus-ouest.fr
Knittweis, Leyla	Department of Biology University of Malta Msida, MSD 2080 Malta		Leyla.knittweis@um.edu.mt
Malvarosa, Loretta	NISEA S.c.a.r.l.		malvarosa@nisea.eu
Martin, Paloma	CSIC Instituto de Ciencias del Mar Passeig Marítim, 37-49 08003 Barcelona Spain	Tel: 4.93.2309500 Fax: 34.93.2309555	paloma@icm.csic.es
Motova, Arina	Sea Fish Industry Authority 18 Logie Mill Logie Green Road Edinburgh EH7 4HS, U.K	Tel.: +44 131 524 8662	arina.motova@seafish.co.uk
Murua, Hilario	AZTI / Unidad de Investigación Marina, Herrera kaia portualdea z/g 20110 Pasaia (Gipuzkoa), Spain	Tel: 0034 667174433 Fax: 94 6572555	hmurua@azti.es
Nord, Jenny	The Swedish Agency of Marine and Water Management (SwAM)	Tel. 0046 76 140 140 3	Jenny.nord@havochvatten.se

Name	Address ¹	Tel.	Email
STECF members			
Prellezo, Raúl	AZTI -Unidad de Investigación Marina Txatxarramendi Ugarteaz/g 48395 Sukarrieta (Bizkaia), Spain	Tel.: +34 667174368	rprellezo@azti.es
Raid, Tiit	Estonian Marine Institute, University of Tartu, Mäealuse 14, Tallin, EE-126, Estonia	Tel.: +372 58339340 Fax: +372 6718900	Tiit.raid@gmail.com
Sabatella, Evelina Carmen	NISEA, Via Irno, 11, 84135 Salerno, Italy	TEL.: +39 089795775	e.sabatella@nisea.eu
Sala, Antonello	Italian National Research Council (CNR) Institute of Marine Sciences (ISMAR), Largo Fiera della Pesca, 1 60125 Ancona - Italy	Tel.: +39 071 2078841 Fax: +39 071 55313 Mob.: +39 3283070446	a.sala@ismar.cnr.it
Scarcella, Giuseppe	1) Italian National Research Council (CNR), Institute of Marine Sciences (ISMAR) - Fisheries Section, Largo Fiera della Pesca, 1, 60125 Ancona - Italy 2) AP Marine Environmental Consultancy Ltd, 2, ACROPOLEOS ST. AGLANJIA, P.O.BOX 26728 1647 Nicosia, Cyprus	Tel.: +39 071 2078846 Fax: +39 071 55313 Tel.: +357 99664694	g.scarcella@ismar.cnr.it gscarcella@apmarine.com.cy
Soldo, Alen	Department of Marine Studies, University of Split, Livanjska 5, 21000 Split, Croatia	Tel.: +385914433906	soldo@unist.hr
Somarakis, Stylianos	Institute of Marine Biological Resources and Inland Waters (IMBRIW), Hellenic Centre of Marine Research (HCMR), Thalassocosmos Gournes, P.O. Box 2214, Heraklion 71003, Crete, Greece	Tel.: +30 2810 337832 Fax: +30 6936566764	somarak@hcmr.gr
Stransky, Christoph	Thünen Institute [TI-SF] Federal Research Institute for Rural Areas, Forestry and Fisheries, Institute of Sea Fisheries, Palmallee 9, D-22767 Hamburg, Germany	Tel.: +49 40 38905-228 Fax: +49 40 38905-263	christoph.stransky@thuenen.de

Name	Address ¹	Tel.	Email
STECF members			
Ulrich, Clara (chair)	Technical University of Denmark, National Institute of Aquatic Resources, (DTU Aqua), Charlottenlund Slot, JægersborgAllé 1, 2920 Charlottenlund, Denmark		clu@aqua.dtu.dk
van Hoof, Luc	IMARES, Haringkade 1, IJmuiden, The Netherlands	Tel.: +31 61061991	Luc.vanhoof@wur.nl
Vanhee, Willy	Independent consultant		wvanhee@telenet.be
Vrgoc, Nedo	Institute of Oceanography and Fisheries, Split, Setaliste Ivana Mestrovica 63, 21000 Split, Croatia	Tel.: +385 21408002	vrgoc@izor.hr

REPORT TO THE STECF

EXPERT WORKING GROUP ON Compilation of the new DCF Annual Report template (EWG-17-17)

**Brussels, Belgium
16-19 (Part 1) and 23-27 (Part 2) October 2017**

This report does not necessarily reflect the view of the STECF and the European Commission and in no way anticipates the Commission's future policy in this area

1 INTRODUCTION

The aim of these EWGs is to produce the new Annual Report template, accompanied by guidelines to be followed by MS in their submission and by STECF in its evaluation. The outputs should result in: (i) streamlining of MS reports under the DCF, (ii) allowing for evaluation of both compliance and quality by the STECF, (iii) simplifying procedures through automatization.

Background

A new Annual Report (AR) template on implementation of data collection is needed, as of 2018, to allow comparison against planning, as described in the, newly established, Work Plan (WP) template. The way forward, as proposed by COM and agreed by STECF, is to create a new AR template that will mirror the WP template. This will facilitate checking of certain parts between the two reports (AR and WP), particularly automatic checking between 'planned' and 'achieved'. A two-step evaluation, including both assessment of compliance and quality (the latter is not done so far), is foreseen. To achieve both goals, the new AR template should: (i) be in line with the WP template, (ii) include additional parts to describe quality aspects. The proposed work should take into consideration previous work carried out on the AR template (STECF EWGs 14-07, 16-01, 17-04).

1.1 Terms of Reference for EWG-17-17 (Parts 1 and 2)

Tasks for the EWGs

The work should be carried out in two consecutive EWGs (parts 1 and 2). The EWG 17-17 part 1 should aim at developing the AR template with guidelines and the EWG part 2 should aim at defining automatic checks between AR and WP templates, with appropriate guidance.

EWG 17-17 part 1

In particular, the EWG 17-17 part 1 is requested to:

- (i) develop the new AR template which will follow the WP template format, where appropriate, and include any additional new parts, where needed
- (ii) compile guidelines for MS submission of the AR
- (iii) compile evaluation guidelines for STECF evaluation of the AR

- (iv) highlight information that may be missing and how these gaps can be addressed through future reporting

The new template should allow the assessment of both compliance and quality. In that respect, the EWG should specify in the evaluation guidelines those parts dedicated/relevant to compliance and those parts that are related to quality. The EWG should propose external references and standards that may be relevant for the evaluation. For specific sections of the AR template, time allowing, this EWG can also proceed with the proposal of automatic checks between AR and WP templates.

The EWG is invited to take into consideration the following criteria for the new AR template and guidance:

1. Realisation against the plan (Work Plan)
2. Coverage: national and regional (areas/stocks); fleet; fishing trip; aquaculture/ processing industry enterprises
3. Definition and description of total population
4. Sampling design that respects basic statistical principles: precision (number of PSUs by strata), bias (population not sampled)
5. Existing research survey protocols
6. Quality assurance
7. Data requirements as specified in EU MAP
8. Known end user needs and formats used
9. Regional specificities and recommendations (RCGs, PGECON, end users)
10. Relative importance of data collection per section (fishery; segment; unit; enterprise)
11. Identification of data gaps and MS needs
12. Description of deviations and mitigation measures
13. Useful information for end users (eg. summary statistics, bilateral agreements)

EWG 17-17 part 2

The EWG 17 part 2 will take place after the completion of the EWG 17-17 part 1. The EWG part 2 is requested to:

- (i) select the fields of new AR template that can be automatically checked against the WP template
- (ii) define ranges of values/ naming conventions and outliers, where appropriate
- (iii) run tests of automatic checks using data from past reports, where possible
- (iii) explore possibilities of automated outputs to be used for the AR submission of MS from existing databases at regional/ European level
- (iv) propose further streamlining of the AR template, with proper justification

The EWG is invited to take into consideration the following criteria for the automatic checking between AR and WP:

1. Use of automatic checks already in place by end users (GFCM, ICES, JRC)
2. List of common code lists and naming conventions and/or update where relevant
3. Definition, where possible, of acceptable levels of discrepancy between planning and implementation

Background documents

Both EWGs are invited to use the following list of background documents:

- Recast Basic Regulation (Reg (EU) 2017/1004)¹
- EU MAP (COM Implementing Decision (EU) 2016/1251)²
- Work Plan template (COM Implementing Decision (EU) 2016/1701)³
- Past guidance on AR reporting and evaluation (STECF EWG 14-07⁴, 15-10⁵, 16-08⁶, 17-07⁷) and TOR3 output of STECF EWG 17-07 on new reporting procedures
- Work already carried out on new AR template (STECF EWG 16-01⁸, EWG 17-04⁹)

¹ <http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1505230909712&uri=CELEX:32017R1004>

² <http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1505231032104&uri=CELEX:32016D1251>

³ <http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1505231148141&uri=CELEX:32016D1701>

⁴ https://stecf.jrc.ec.europa.eu/documents/43805/768107/2014-07_STECF+14-13_Evaluation+of+2013+AR+and+Data_JRC91550.pdf

⁵ https://stecf.jrc.ec.europa.eu/documents/43805/1002766/STECF+15-13+-+Evaluation+of+2014+DCF+AR+and+DT_JRC96975.pdf

⁶ https://stecf.jrc.ec.europa.eu/documents/43805/1440437/STECF+16-12+-+Evaluation+of+DCF+AR+and+transmis+issues_JRC10266.pdf

⁷ https://stecf.jrc.ec.europa.eu/documents/43805/1711451/STECF+17-10+-+Evaluation+of+DCF+AR+and+DT_JRC107592.pdf

⁸ https://stecf.jrc.ec.europa.eu/documents/43805/1366312/STECF+16-07+EU+MAP+and+template+National+Work+Plan_JRC101530.pdf

⁹ https://stecf.jrc.ec.europa.eu/documents/43805/1679122/STECF+17-11+-+Quality+assurance+for+DCF+data_JRC107587.pdf

- Guidance on WP template evaluation (STECF EWG 16-01) and experience gained from STECF EWG 16-16¹⁰
- End users reports that make use of/ reference to DCF data and/or reporting (STECF, ICES, GFCM, RFMOs, RCGs, PGECON)
- Final reports of MARE/2014/19¹¹
- Reporting from existing databases and upcoming developments therein
- Data Quality Assurance repositories and documents (ICES PGCCDBS repository¹², WGCATCH¹³, EUROSTAT, JRC, GFCM)
- RCG and PGECON 2017 reports¹⁴
- Any relevant scientific publications

¹⁰ https://stecf.jrc.ec.europa.eu/documents/43805/1561325/STECF+16-25+-+Evaluation+DCF+national+WPs+2017-2019_JRC104918.pdf

¹¹ <https://datacollection.jrc.ec.europa.eu/docs-links/mare-2014-19>

¹² <http://www.ices.dk/community/Pages/PGCCDBS-doc-repository.aspx>

¹³ <http://www.ices.dk/community/groups/SiteAssets/WGCATCH-publications.aspx>

¹⁴ <https://datacollection.jrc.ec.europa.eu/docs/rcm/2017>

2 REPORT OF EWG 17-17 PART 1

As outlined in the ToRs, Part 1 of the EWG developed as key outputs Annual Report templates (Excel tables and Word document) and a guidance document. The AR templates were built on the existing National Work Plan (NWP) templates to allow straightforward comparison between planned (in National Work Plans) and achieved (in AR). Additional elements were added to allow evaluation of both quality and conformity with NWPs and to facilitate automatic cross checking. All of the ToRs were addressed in the meeting and two outputs, Draft Guidance for the submission of Annual Reports, incorporating suggestions for additional text boxes (ANNEX 1) and draft table templates (ANNEX 2). Additionally the EWG also considered whether some information could be auto-filled from regional and transnational data collections, specifically the Regional Database (RDB) and annual economic data call.

The EWG adopted a two layered approach to the meeting's objectives, firstly developing the template followed by construction of quality requirements and guidelines for MSs and STECF evaluation. This was to facilitate reporting of both the conformity with the NWP and the quality of the information provided. The existing NWP templates were augmented by additional text boxes and spreadsheet extensions columns to allow provision of evidence and commentary on what was achieved; any deviations from what was proposed in the NWP; any data gaps and descriptions and indicators of quality. It was highlighted that an evaluation of fitness for purpose should also entail collective evaluation of MS programmes to ensure that EU programme objectives are met. This will involve integration of regionally coordinated elements (evaluation of regional work plans etc.).

In developing the AR structure the EWG needed to keep in mind the constraints imposed by building on the NWP template. It was accepted that the NWP structure was set and changing this was out of the scope of the group. However it was agreed that it should evolve over time and suggestions for change were an important outcome of the meeting. NWP and AR development could be seen as process of continuous improvement with each review cycle building on the last. To a degree this meant a change in the perception of ARs and AWP to considering them as live documents in need of maintenance thereby would ensuring that national programmes and the DCF programme as a whole remained fit for purpose and was able to better adapt to meet the requirements of the recast DCF and any future challenges.

As a principle therefore, the EWG was instructed that they should avoid requesting any additional information that should have ideally appeared in the NWP template. However exceptions were permitted in some instances where inclusion did not impose an undue burden on MS (is reasonable and proportionate) and was in line with legal requirements and most importantly would aid future evaluation. Proposal of entire new sections to the AR which did not follow AWP was also considered necessary in some instances.

On conformity, it was noted that the STECF role was technical one and not one of policing compliance. Key considerations were - did MS do what they set out in AWP? Did they meet the programme objectives as detailed in the EU MAP and did they meet end user needs?

On quality, the question asked was – what is necessary to carry out an effective evaluation of quality? In looking at quality criteria, specific reference was made to the report of EWG 17-04 – Quality assurance for DCF data (STECF -17-11). Data transmission failures were considered to be out of scope the meeting whilst being an integral part of evaluation. Improvement in quality was seen as an incremental process with NWP and ARs providing a snapshot of progress described in two tables (5A for biological variables and 5B for social and economic variables). The role of the Data Quality sub-group of PGECON is elaborated in Section 2.3.1.1 of Part 1 of the report.

Prefilling by the Commission of two tables - Table 7A (Planned regional and international coordination) and Table 7B (Follow-up of recommendations and agreements) was requested to help MSs provide AR information in a uniform way.

The EWG was split into subgroups (see table below), each looking at different sections of the AWP and AR templates as below. The outcomes of the sub-groups are set out in sections 2.1 to 2.7.

EWG Sub-Group Responsibilities

First Name	Last Name	Job title	Section of template	Cooperation between groups
Angeliki	ADAMIDOU	Biologist	commercial fisheries of Section 1	Section 4: sampling design and Section 5A: biological quality assurance
Angeles	ARMESTO	Biologist	commercial fisheries of Section 1	Section 4: sampling design and Section 5A: biological quality assurance
Bram	COUPERUS	Biologist	By catch	environmental impact of fisheries
Henrik	DEGEL	Biologist	Surveys	commercial fisheries of Section 1, Section 4: sampling design and Section 5A: biological quality assurance
Matthew	ELLIOTT	Other	Section 2 - fishing activity and Section 5B on socio-economic quality assurance	Section 3: Socio-economic sections
Emmet	JACKSON	Economist	Section 3: Socio-economic sections	Section 2 - fishing activity and Section 5B on socio-economic quality assurance
Edvardas	KAZLAUSKAS	Economist	Section 3: Socio-economic	Section 2 - fishing activity and

First Name	Last Name	Job title	Section of template	Cooperation between groups
			sections	Section 5B on socio-economic quality assurance
Estanis	MUGERZA	Biologist	recreational fisheries	environmental impact of fisheries, bycatch
Alastair	POUT	Biologist	Section 4: sampling design and Section 5A: biological quality assurance	Automatic checks
Jose	RODRIGUEZ	Biologist	Section 4: sampling design and Section 5A: biological quality assurance	surveys
Maria Begoña	SANTOS VAZQUEZ	Biologist	environmental impact of fisheries	By catch
Alan	WALKER	Biologist	anadromous/catadromous	Section 1 for specific part of commercial fisheries of diadromous species and work on automatic checks

2.1 AR Section 1: Biological data

2.1.1 Table 1A, List of required stocks; Table 1B, Planning of sampling for biological variables; Table 1C, Sampling intensity for biological variables

Tables 1A and 1C have been modified by adding columns to reflect, in the case of 1A, (column *Changes in species landings*) the change in landings of species that can have an impact in the sampling planned, and in case of 1C, to reflect the results of the sampling in terms of numbers of individuals and number of samples.

These new columns (Column "*Achieved number of individuals measured at the national level*" and "*Achieved number of samples*") are needed because an effective sample size together with total number of individual measurements forms the basis of precision indicators. A column to briefly describe the sampling protocol used was also added because the existence of a sampling protocol for each species and variable is considered a primary measure of data quality.

The addition of these columns was suggested by EWG on Quality Assurance of DCF data (EWG-17-04) and agreed by this EWG 17-17.

A text box was proposed with sections to explain "Deviations from the NWP" and the "Actions to avoid deviations". In addition a section to provide evidence of data quality assurance was included in text box. Although the quality evaluation could be only carried out if the information from Table 5A was available. Where this information is not available, some overview can be derived from the AR by giving information in this section on the methodology used to assure the quality of the data collected.

A modification in table 1B was not considered necessary since it relates to long-term planning for the three-year period and affects only the NWP. It is not applicable to the AR.

2.1.2 Table 1D: Recreational fisheries

The principles for the evaluation of the Annual Report must be based on the evaluation of the realisation against the Work Plan. In the case of Recreational Fisheries (Table 1D), this table is useful for compliance/conformity but insufficient for quality evaluation.

The critical requirement for quality evaluation procedures is to have accurate documentation of all components of the programme (design, implementation and analysis). These components are included in NWP Table 5A (Quality Assurance Framework for biological data). However, it seems that Member States (MS), following the guidelines to fill in the tables "this table is intended to identify data to be collected under tables 1A, B and C" have not included any documentation about the surveys carried out on Recreational Fisheries Surveys.

For these reasons, some additional information and clarification in the AR text is requested based on the NWP tables. These would help in the evaluation of the planned surveys and in the quality of the estimates provided. Additional information requested to provide to Member States in the AR text Word document (see draft AR guidance – ANNEX 1).

2.1.3 Pilot Study 1: Relative share of catches of recreational fisheries compared to commercial fisheries

During 2017-2019 many MS will carry out different type of surveys under pilot studies, to obtain catch estimates of recreational fisheries and their impact on different stocks or species. These multi species surveys will allow the evaluation of the impact of these fisheries in the different MS.

A template text box has been provided to guide MS in the reporting of the results of these pilot studies. For example, MS should provide brief descriptions of the results obtained and justifications as to why there have been deviations from the original plans. MS should also report on whether they have achieved the planned outcomes of the pilot study and justification should be provided if this was not the case. Finally MS should report on plans to incorporate the results from the pilot study into regular sampling since this is the overall aim of the pilot study. This format is suggested for all pilot studies in the AR.

2.1.4 Anadromous and catadromous species data collection in fresh water: Table 1E and Text Box 1E

In preparing additional columns and text boxes for details of anadromous and catadromous stock, the following were reviewed: templates and guidance for the previous AR, templates and guidance for the WP, STECF EWG reports and took advantage of the expert opinion available to EWG 17-17.

There are some inconsistencies between the texts in the Legal Text detailing what data should be presented in which tables in the WP, and AR. Therefore, the EWG proposed that the AR template and guidance based on the interpretation that data for anadromous and catadromous stocks should be collected on, and reported in:

- Commercial fisheries in marine waters: should be defined in Tables 1A, 1B, 1C;
- Recreational fisheries in all waters: should be defined in Table 1D;
- Commercial fisheries in freshwaters: data collection of landings and biological data should be defined in Table 1E;
- Data collection for eel recruits, standing stock and silver eels, and for salmon parr, smolts and ascending adults (hereafter called eel and salmon life stages): should be defined in Table 1E, and the salmon requirements should exist for sea trout in the Baltic marine region as well;
- The sampling strategy for biological information from commercial fisheries, and recreational fisheries, and for eel, salmon and sea trout life stages: should be described in Table 4A;
- Information on the Quality of all data collection plans, for fisheries-dependent and independent, should be described in Table 5A.

2.1.5 Table 1F and Table 1 F(a): Incidental by-catch of birds, mammals, reptiles and fish (Data to assess incidental by-catch and other impacts of Union fisheries on marine ecosystems)

The EWG Part 2 proposed splitting the information being requested in this section to two tables – Table 1F "*Incidental by-catch of birds, mammals, reptiles and fish*" and Table 1F (a) "*Data to assess incidental by-catch and other impacts of Union fisheries on marine ecosystems*" because the table now contains more than incidental by-catch data as described below.

There has been a drive to implement the EAFM for several decades now, a management framework that should also take into account the impact of fisheries on the wider marine ecosystem (on target species, on by-catch species, on trophic relationships and on habitats). The Commission Implementing Decision (EU) 2016/1251 of 12 July 2016 adopting a multiannual Union programme for the collection, management and use of data in the fisheries and aquaculture sectors for the period 2017-2019 requires that data to assess the impact of European Union fisheries on marine ecosystems in Union waters and outside Union waters should be collected. Decision (EU) 2016/1251 specifies the requirement for documenting incidental by-catch of all birds, mammals and reptiles and fish protected under Union legislation and international agreements (including species listed in Table 1D) in all types of fisheries. However, Decision (EU) 2016/1251 is quite vague at specifying other types of data to be collected for evaluating additional fishing impacts.

For determining the impact of fishing on habitats (physical loss and physical disturbance), the level of fishing activity gathered from VMS/logbook data or other sources is needed (specifically for bottom-contacting gears) to provide an indication of potential impact on habitat quality. VMS data are collected as part

of Regulation (EC) No 1224/2009 and although in principle available, a significant issue could be the existence of time lags in the availability of the VMS data.

For other information, 2016/1251 mentions "data for estimating the level of fishing and the impact of fishing activities on marine biological resources and on marine ecosystems, such as effects on non-commercial species, predator - prey relationships and natural mortality of fish species in each marine region". Because by-catch (and discards) are already covered in separate tables we have concentrated on data that would provide information on trophic relationships which could be obtained from the analyses of stomach contents.

In the light of these requirements Table 1F with 1F (a) have been modified to:

- Evaluate incidental by-catch programmes submitted by the MS
- Provide an overview of other data collection for future determination of the impact of fisheries on the marine ecosystem (excluding incidental by-catch). Because MS have not indicated figures/activities/targets, etc. in the NWP, Table 1F is purely informative and can not be used for checking conformity.

Changes and additions are summarised and justified below:

- The purpose of the proposed data collection by MS is to obtain reliable incidental by-catch data. For this the observation effort needs to be provided and data need to be appropriately collected. As with other types of data, data need to be stored and transmitted to end users.
- The WP Table 1F was modified by adding columns to reflect, in the case of incidental by-catch, MS inputs on all of these aspects.
- In relation to coverage, new columns have been included: **Average number of PSUs during the reference years** (this column is already available in Table 4A of the WP), **Planned number of PSUs** (also available in Table 4A of the WP) and **Number of PSU sampled for by-catch** to obtain an image of the realised sampling effort. The columns from Table 4A are repeated here, in revised Table 1F) to obtain the information at a glance for each Member State.
- In relation to sampling design, it is important to stress that incidental by-catch observation effort is NOT the same as general observation effort, since it requires a specific and different kind of observation to that normally used. Incidental by-catch rate can only be estimated if the observation effort "at haul level" has been recorded. An incidental by-catch is by definition a rare event and often the individual(s) animals involved are relatively large. By taking a small subsample (e.g. a fish basket) from a large catch, incidental by-catch will almost always be missed. The whole catch or at least a substantial part of the catch needs to be scanned to record incidental by-catch events. New columns have been added to obtain this information from MS, these new columns are **% Trawls codend observed** and **% sorting observed**. Explanation on the guidelines have been provided to specify how these columns should be filled.

- In relation to data, a column has been added, **Is there any mitigation device?**, because mitigation devices (i.e. acoustic deterrent devices, escape devices, etc.) could affecting the incidence of by-catch. New columns have been added, **#Fish, #Mammals, #Birds, #Reptiles** to obtain information on number of individuals incidentally by-caught by group. Another column, **Additional data/samples available?**, has been added to list if additional information is been collected.
- In relation to data storage, two columns (are also available in Table 5A of the WP) have been added **Are data stored in a national database?** and **Are data stored in international database(s)?**. The rationale behind is that MS report where the data are stored.
- New columns have been added to obtain information on MS activities in relation to stomach content sampling and VMS/logbook data (and other data on vessel location and activity recorded under Regulation (EC) No 1224/2009). These new columns have been added to the table to provide an overview of the data collection on aspects that could be useful to assess the effect of fisheries on trophic interactions and habitats. This part of the Table 1F is purely informative and is not for checking conformity, because MS have not indicated figures/activities/targets, etc. in the WP for these data. These new columns are **Any stomach content data collected?, Group of species sampled, #species sampled, #stomach sampled**.
- Two columns for incidental by-catch data (also available in Table 5A of the WP) have been added in relation to data storage, **Are data stored in a national database?, Are data stored in international database(s)?**. As there could be a time lag between collection and data availability, a new column has been added to obtain information on the extend of this time lag, **Effective time lag for availability**.
- In relation to VMS data and after some discussion in the group, it was decided to add the following columns, **are VMS/logbook/EM data collected?**, to list whether information is available that can be used to determine the impact of fishing on habitats, if the answer is "yes", in the next column added, **Type of data**, MS are asked to provide which type of data are available (i.e. VMS, logbook, etc.). Another column has been added to ask in which database de data are stored and the contact, **Database where the data are stored**, and another column on **Effective time lag for availability**.
- For the text of the AR additional information requested from MSs includes, (for example):
 - for the incidental by-catch section we ask that species or families to be listed (if the identification is available);
 - number of samples collected and the state of the animals incidentally by-caught (i.e. were they released alive, dead, or collected for sampling).

- For the stomach content section, to provide additional information by listing the species sampled, the number of stomach per species sampled, details on methodology, etc.
- A new text box, **Text Box 1F: Data to assess incidental by-catch and other impacts of Union fisheries on marine ecosystems**, has been added to complement the information provided by MS in the Tables .

2.1.6 Pilot Study 2: Level of fishing and impact of fisheries on biological resources and marine ecosystem

Most MS have mentioned they will carry out pilot studies to obtain incidental by-catch data and/or other information to assess the impact of fisheries on the ecosystem (very few mention other components different from incidental by-catch).

A template text box has been provided to guide MS in the reporting of the results of these pilot studies. For example, MS should provide brief descriptions of the results obtained and justifications as to why there have been deviations from the original plans. MS should also report on whether they have achieved the planned outcomes of the pilot study and justification should be provided if this was not the case. Finally, and equally importantly, MS should report on plans to incorporate the results from the pilot study into regular sampling since this is the overall aim of the pilot study.

Similarly, guidance has been provided to evaluators, where particular emphasis should be given to the overall quality of the Pilot Study Report, the provision of sampling protocols, whether sampling design and protocols follow internationally agreed protocols, soundness of conclusions drawn from the Pilot Study and the follow-up suggested by MS, among others.

2.1.6.1 Data quality evaluation

A section on Data quality was discussed and has been added to the AR guidelines both for MS and for evaluators. Important aspects are related to the provision of details on sampling protocol and sampling design for incidental by-catch data collection and stomach content collection. A series of questions have been included to guide MS in their responses. This is particularly important because for example, in the case of incidental by-catch, very few MS have mentioned at sea observers and in this case almost all will be observers carrying out biological and other sampling the ones collecting also by-catch data. In this case it would still be useful to know the duties of the observer in order to judge their likely dedication to observing by-catch. Additional specific questions are whether data quality issues are taken into account and on how data and samples are stored.

2.1.7 Table 1G: List of research surveys at sea

The EWG proposed the addition of a number of columns to table 1G:

“Deviation from fixed temporal range” and “Deviation from fixed spatial range” is a simple indication if the spatial and temporal coverage differs for the information given in the NWP.

“Indication if AR comments by MS are required concerning effort achieved” indicates if the Member State is requested to provide a comment. The request is based on the discrepancy between the “Days at sea planned” and the “Target planned” given in the NWP and the “Days at sea achieved” and “Target achieved” in the AR. If one of the measures exceeds a certain margin (say 10 percent), the column shows an “x” implying that a comment, which explaining this non-conformity is mandatory. The reason for introducing this interactive functionality is to support that explanations for non-conformities are provided where necessary already in connection with the Member States initial submission of the AR. A missing “x” should not prevent the Member State to provide a comment if the Member State finds it relevant. If the comments are to extend to fit into the table format then the comments should be put in the text box section and only a reference to this should be given in the comment field.

“Indication if AR comments by MS are required concerning temporal and spatial coverage” has the same functionality as the one above but concerning discrepancies between survey area and period. If these parameter values are not the same in the NWP and the AR this releases an “x” in the column “Indication if AR comments by MS are required concerning temporal and spatial coverage”.

“Type of MS participation” is associated with shared cost procedure and indicates how the Member State contributes to the survey. This information is already given in the NWP (covering a three-year period) but any changes from that in AR year must be indicated in order to resolve any financial implications.

“Other data assimilations” is added in order to provide information if auxiliary data are collected (e.g. CTD data and stomach data) and submitted to other data holders than the database given in Table 1H.

2.1.8 Text Box 1G: List of research surveys at sea

The following points have been included in the text box related to the AR:

- Graphical representation (map) showing the positions (locations) of the realized samples. Here the Member state should include a map showing the result of the survey carried out. The map should provide the possibility to compare the survey results in terms of stations and location with the information stated in the AWP.
- For internationally coordinated surveys, provide a link to the latest meeting report of the coordination group - Most international surveys are coordinated and quality assured by a working group and the work of these groups are very relevant for the review if issues that are more specific are discovered and needs review.
- List the main use of the results of the survey (e.g. indices, abundance estimates, and environmental indicators) - In order to provide an

overview of the use of the data collected during scientific surveys the Member State should list the various uses of the data.

2.1.9 Table 1H: Research survey data collection and dissemination

The EWG proposed the addition of a number of columns to table 1G:

"Is the sampling carried out?" indicates if the Member State actually has carried out the specified type of sampling listed in the NWP. If not, an explanation should be given in the "AR comment" indicated by "x" in the "Indication if AR comments are required by MS" field. The reason for the interactive functionality is the same as for Table 1G.

"Relevant International database" indicate if there exists an international database, which holds all or part of the survey results obtained and gives the name of the database. The following column, "Is data uploaded to the relevant database (previous column)", inform if the data obtained are uploaded to the database and hereby makes it available for further use.

"Other data assimilations" is a possibility to inform on a lower level if the data type listed in the NWP is submitted to other data holders than the relevant international database.

2.1.9.1 Data quality issues

The majority of surveys where more than one Member State are involved are coordinated by a working group which coordinate the work between the participants and perform a continuous quality assurance covering most aspects of the survey. Therefore, the role of the AR in connection with scientific surveys is mostly to provide the basis for an evaluation of the degree of conformity with the NWP.

2.2 AR Section 2: Fishing activity data - Table 2A: Fishing activity variables data collection strategy and Text Box 2A: Fishing activity variables data collection strategy

This section should include only those transversal data that are not available through administrative sources (logbooks, landing declarations and sales notes) and so collected through bespoke surveys. It was noted that many member States had provided information on their entire transversal data collection and much of this was not needed. It was clarified that information on metier was optional for MSs to provide as an aid to describing complementary data collection.

Additional information requested for the AR includes the following:

- Data source for complementary data collection;
- Achieved coverage of data collected under complementary data collection;
- Response Rate, %.

2.3 AR Section 3: Economic and social data (Fleet, aquaculture and processing)

2.3.1 Tables and Text Boxes 3A, 3B and 3C: Population segments for collection of economic and social data for fisheries, aquaculture and the processing industry and Pilot Study 3: Data on employment by education level and nationality

The STECF Expert Working Group (STECF EWG 17-07) observed that the assessment of data transmission issues and the evaluation of AR should be better aligned. To date for instance, the EWG 17-07 evaluated AR on data collection activities performed in 2016, but assessed the data transmission issues of the data call from the previous year, which means e.g. 2015 biological data and 2014 economic data. This alignment would be needed to link data transmission failures with the corresponding annual reports, and to link those directly with any subsequent necessary amendment required in the Work Plans for the following period. This is essential for the new AR.

STECF EWG 17-07 noted that, as in previous years, the online compliance platform provided by the JRC on the DCF website facilitated the work of the experts to evaluate the annual report. It is the suggestion of EWG 17-17 that this platform is also utilised when MS are completing their annual reports. This will result in the actual achieved data transmission being presented in the annual report (as Achieved Sample Rate % and Response Rate). Using reports from the JRC database will not only ensure consistency in naming conventions but give the real data achievements (except in instances where data has been omitted for confidentiality reasons). To that end the new columns suggested in the revised AR tables should be populated from the summary of data transmission to the JRC for the reference year.

2.3.1.1 Quality evaluation in Economic and social variables.

The quality assurance task for RCG was indicated in STECF 17-11 on Quality Assurance for DCF Data saying that after RCMs were developed into Regional Coordination Groups following the recast of the Data Collection Framework regulation (2017/1004) it extended the scope of the groups which also have clear aims to develop and implement procedures, methods, quality assurance and quality control in particular regional groups. The role of the RCGs has a clear objective to improve quality assurance and quality control.

To address this PGECON established a subgroup on Quality Assurance Framework and drafted the first recommendations relevant to the reporting of quality. These quality measures have been extended into the AR text boxes for 3A, 3B and 3C and now request information to be reported on selected quality assurance principles, namely 'sound methodologies', 'sampling strategies', 'state of data accuracy and clarity'.

In line with additional quality reporting requirements, guidelines for filling AR template were complemented for the relevant sections. These checkpoints were recommended to be used by PGECON 2017 Subgroup on quality assurance and can reflect part of quality assurance process in MS. Further quality assurance principles, and their implementation, should be continuously tackled in annual PGECON meetings, particularly in Subgroup of QAF.

2.3.1.2 Selection of quality indicators for AR templates

STECF EWG 17-11 reports that the quality of data collection has generally been quantified through figures like sample rate, response rate, coverage rate and CV. It was suggested that these figures could also further be used as quality indicators. However, the evaluation of “acceptable data quality” should be performed with caution. Fleet segments are often small populations which cannot necessarily be regarded homogeneous. Therefore CV can easily be high even though the coverage is high, too. Therefore CV is not selected as quality indicator for new AR templates. Based on The SGECA 09-02 report dealing on quality aspects of the collection of economic data-methods of calculation of the indicators and sampling strategies, for evaluation of quality for economic and social variables in fleet, aquaculture and fish processing sectors, indicators were unchanged from previous AR.

Following previous AR practice, for economic and social variables, data quality will be evaluated by response rate, achieved sample rate and achieved coverage rate for fishing activity variables. However, as AR template had to be developed according to the WP template, some WP amendments (which are not introduced yet) have been taken into account and was included in AR tables. Changes were made based on the summarized remarks and proposed amendments listed in STECF 17-11 ANNEX 4.

2.3.2 Pilot Study 4: Environmental data on aquaculture

For the aquaculture environmental Pilot Study this relates to the following EU MAP requirement:

- Environmental data may be collected on the basis of pilot studies and extrapolated to indicate totals relevant to the total volume of fish produced in the Member State.
- Environmental data shall be collected every two years.

The data requirements for aquaculture were considered in some detail at PGECON last year (Subgroup of Fisheries Data Collection Experts Group 6th Planning Group on Economic Issues – PGECON, 15-19 May 2017, Vilnius, Lithuania). PGECON noted that the variables set out in the EU MAP were not specified in sufficient detail and further work was needed to make sure that the data collection was useful. PGECON recommended that this should be done through a workshop meeting in 2018. Following the workshop the next time data would be collected then would be 2020. This would be done on the basis of any guidelines agreed there. There seems little merit in pre-empting the outcomes

of pilot studies by asking for figures from those studies in the AR. Similarly there seems no value in trying to anticipate what the figures might look like in any future annual report since reporting would fall out of the scope of the current EUMAP. The draft AR guidelines require only an update on any work carried out on pilot studies (if any). This information would be mostly of interest to carry forward to PGECON and the proposed workshop in 2018. As for other pilot studies, a text box requesting information on progress with any pilot studies is included in the draft AR template.

2.4 AR Section 4: Sampling strategy for biological data from commercial fisheries - Table 4A: Sampling plan description for biological data and Text Box 4A: Sampling plan description for biological data

The proposed structure of the annual report template for table 4A serves multiple purposes. Firstly it records the conformity of the MS data collection to the WP, secondly it provides a snapshot of the data collected and thirdly, it provides some insight into some quality issues".

The conformity section of the template is based on the archived number of PSU in the reporting year. This can be compared directly with the planned number of PSU set out in the work plan. This is the simple and best measure of the conformity of the data collection. This is achieved with the addition of columns that record the total number of PSU available to be selected in the reporting year and the achieved number of PSU totals that were actually sampled.

The data collection part of the template provides a snapshot of the data collected and is quantified in terms of the on-shore locations vessels, trips, species length measures biological species and biological measures (age sex weight and maturity). These follow the schemes and stratum rows set out in Work Plan table 4A. and thus provide this information on a scheme and stratum basis.

The EWG was asked to introduce additional fields allowing for some quality assessment in Table 4A. EWG advised, on the one hand, to calculate the total coverage in terms of PSUs and, on the other hand, to provide total number of unique vessels with activity within the strata and total number of trip. These two columns should allow, in comparison with unique vessels sampled and unique trips sampled, an objective first approach to understand how selection of sampling units (PSUs, SSUs, etc) is implemented along the different steps of the sampling scheme.

It is agreed by the by the EWG plenary that providing a breakdown of the data serves the purpose of quantifying the data collected. However it was recognised that the data collection community is evolving, MS are at different stages in that evolutionary process. It may not be possible for all MS to provide data in that form for data collected during 2017. Such summaries are an aspiration that could be achieved over time. It was stressed that the role of the STECF AR evaluation meeting was not to penalise MS for failing to deliver data, rather to provide positive feedback to MS in order to aid the improvement of data collection schemes across MS.

The MS that can utilise a regional data base for the storage of achieved sampling data (which is nearly all MS in the Baltic, North Sea Eastern Arctic and North Atlantic region), can utilise the functionality of such a databases to automate the production of the AR table 4A. The AR table 4A template is therefore seen as a major driving force in collating sampling at the regional level with all the advantages of harmonisation, cost efficiency, transparency and data delivery that can thus be achieved. These are all facets of “data quality”, as set out in the Quality Assurance Framework of the European Statistical System¹⁵.

2.5 Table 4B: Sampling frame description for biological data

The EWG agreed that this table should be presented as it appears in the National Work Plan without any additional information. No extra fields are requested to be completed in this table.

2.6 Table 4C: Data on the fisheries by Member State

Table 4C describes the distribution of vessels, fishing effort and landings across the fleets.

According to the criteria allowed in the regulation (EU 2016/1701) MS can choose different criteria under the field “Fleet segment / Metier” to complete the information in their WPs. While this information is useful for getting a picture of the national fisheries (e.g. fleet segment according to LOA, metier level 6 as an extended domain provided to end-users, etc.), the utility of this table to get a better understanding of the fisheries dynamics and the sampling issues (e.g. coverage, impact in sampling results, etc) is not fully exploited.

The comprehension of the AR results could be improved if information provided in this table is given following the same segmentation that is used for at-sea sampling programmes in Table 4A-4B. The information in Table 4C can then be compared directly to the population described in 4A.

It was suggested that national stratification in Table 4A may or may not provide sufficient information for end-users or third parties to get an overview of the fisheries (e.g. fleet segment according to LOA or metier level 6). However it was argued that this was not the primary purpose of either the NWP or the AR.

2.7 Table 4D: Landing locations

The EWG plenary agreed that this table should be presented as it appears in the National Work Plan without any additional information. However it was subsequently suggested that there is no fundamental difference between tables 4D and 4C in that both should set out a logical grouping of populations and therefore both should be updated with the current year’s values. It is noted that Table 4D has the ability to demonstrate if a MS has a high proportion of foreign landings, and thus an obligation to sample those landings. The absence of this

¹⁵ http://ec.europa.eu/eurostat/documents/64157/4392716/qaf_2012-en.pdf/8bcff303-68da-43d9-aa7d-325a5bf7fb42

information may risk leaving gaps in any pan-European data collection scheme. Further expert deliberation on this issue is warranted and a version of an augmented Table 4 D is provided in the draft AR guidance

2.8 AR Section 5: Data quality - Table 5A: Quality assurance framework for biological data and Table 5B: Quality assurance framework for socioeconomic data

The table in the annual report serves to provide an update snapshot of the state of play with regard to the development of quality assurance frameworks by MSs. Evaluation is as for the National Work Plans and no additional columns have been added.

Quality and conformity indicators are provided in the tables for biological and socio-economic data and information relating to these is in the relevant sections of this report.

As noted in Section 3, PGECON have established a Quality Assurance Framework which can be utilised to provide further feedback to the Commission and end users on the methodologies and quality of data in Member States.

2.9 AR Section 6: Data availability (Table 6A: Data availability)

An additional column has been added for MSs to show when data actually became available in the reporting year compared with the times shown in the NWP.

2.10 Section 7: Coordination

2.10.1 Table 7A: Planned regional and international coordination

An additional column has been added to the NWP template to indicate numbers of staff sent to each meeting by Member States and a further column to comment on where MSs were absent from significant meetings. The EWG requested that the template be pre-filled by the Commission with relevant meetings, although this would not preclude MSs from adding additional meetings where appropriate.

2.10.2 Table 7B: Follow-up of recommendations and agreements

The EWG recommended that the template be prefilled by the Commission with relevant recommendations to help ensure consistency of responses from MSs. An additional column is provided in the template for MSs to indicate how recommendations have been followed up by MSs.

2.10.3 Table 7C: Bi- and multilateral agreements

The table in the annual report serves to provide an update snapshot of existing agreements. Evaluation is as for the National Work Plans and no additional columns have been added.

3 REPORT OF EWG 17-17 PART 2

As outlined in the ToRs, Part 2 of the EWG reviewed the draft AR templates and guidance in detail. The EWG in consultation with Part 1 experts, added to the guidance and structure of the tables where necessary and considered what automation would help with the screening and evaluation of the submitted reports. Whether the process was within the Excel spreadsheets, externally using R or a Web based approach all the methods are dependent on the values and the formats data entered or submitted in Excel. As a first step - each of the fields in each of the tables were defined: formats, range limits, thresholds, reference lists, the checks that would be needed and the consequences of failure. These are catalogued in the spreadsheet reproduced in ANNEX 3.

The EWG considered the different stages of the evaluation process and where these checks might apply and who would benefit. Automating the pre-evaluation would serve both Evaluators and Submitters. As the NWP forms part of the AR then the same automation or checks could also be used in any pre-evaluation of the NWPs. Templates could be designed in Excel (or compatible) to limit the data provided and flag when further information is required based on what has been entered. MS could use shared R code for pre-screening their own ARs before submission. The code would provide summaries of errors or improvements and indicate when additional information might be required. Similarly, a Web application would give a pre-submission report after upload allowing MS the time to correct any omissions before formally submitting their final report. The whole process could be further enhanced by referring to MS transversal and biological data held on Regional Databases (see Section 3.2.1.2). Standard reports from these databases could summarise sampling achievements against population data. Tying these reports to any pre-screening exercise would allow MS to review and answer any flagged deviations from what was planned. The EWG presents working examples of each of these scenarios in this report. The EWG also summarises the current use of automatic checking in data exchanges by ICES, GFCM and JRC.

Key outputs:

- Review and draft edits of the AR templates and guidance with reference to compliance, conformity and quality and possible automation.
- A reference list of the checks types that could be applied to the fields and reports for consistency, conformity and quality.
- A catalogue of the AR tables and fields with the checks that could be applied to each as a key reference for any coding.
- Working examples:
 - of a table from the of AR template in Excel offering drop down lists and threshold checks
 - R-code for screening and reporting NWPs
 - Web App for screening and reporting on NWPs.

To cover the detailed review the EWG was split into 3 subgroups (see table below), each looking at different tables within the AR templates. This process was comprehensive and covered two tasks:

1. to review the draft templates and guidance with reference to the rationale provided by 17-17(1) and the regulation. Editing the draft, commenting and documenting particular concerns and issues.
2. to identify what checks could be applied to what fields within each table.

EWG Sub-Group Responsibilities

First Name	Last Name	Section of template
David	CURRIE	Group 1: 1A-1C, 1E, 1F, 4A, 5A, 6A; Lead - Links to RDB
Dimitrios	DAMALAS	Group 1: 1A-1C, 1E, 1F, 4A, 5A, 6A
Henrik	DEGEL	Group 2: 1D, 1G, 1H; Lead - Automation Excel
Laurent	DUBROCA	Group 1: 1A-1C, 1E, 1F, 4A, 5A, 6A; Lead - Automation with R
Jon	ELSON	Group 2: 1D, 1G, 1H
Maria Teresa	FACCHINI	Group 1: 1A-1C, 1E, 1F, 4A, 5A, 6A
Jerome	GUITTON	Group 3: 2A, 3A-3C, 5B; Lead - Automation with Web app.
Jenny	NORD	Group 3: 2A, 3A-3C, 5B

3.1 Template scrutiny and automated checks.

The EWG (2) had not seen the draft templates nor guidelines before the meeting so the work carried out in the first group 17-17(1) was presented by Venetia Kostopoulou and Henrik Degel. The aims and rationale behind each table was discussed in plenary. Laurent Dubroca gave a presentation of the data quality validation in the GFCMs Data Collection Reference Framework where four operational quality indicators have been defined for their data exchanges (GFCM 2016). These indicators are: (a) **Conformity**: checks if a value conforms to the syntax of its definition (format, type, range). (b) **Stability**: checks if values vary at an acceptable level based on values of the recent past. (c) **Coherency**: checks if reported values are equal across different data tables. (d) **Accuracy** (precision and bias): checks the degree to which values vary from a true or expected value. The first three have been tentatively implemented. Although in this instance the GFCM are dealing with disaggregated sample data and our AR templates are details that describe the data collection (Meta-data) these Quality Indicators could form the basis of an automated process here.

The EWG discussed the different types of checks, the general integrity and consistency expected for these tables with reference to the scheduled evaluation process (Figure 1).

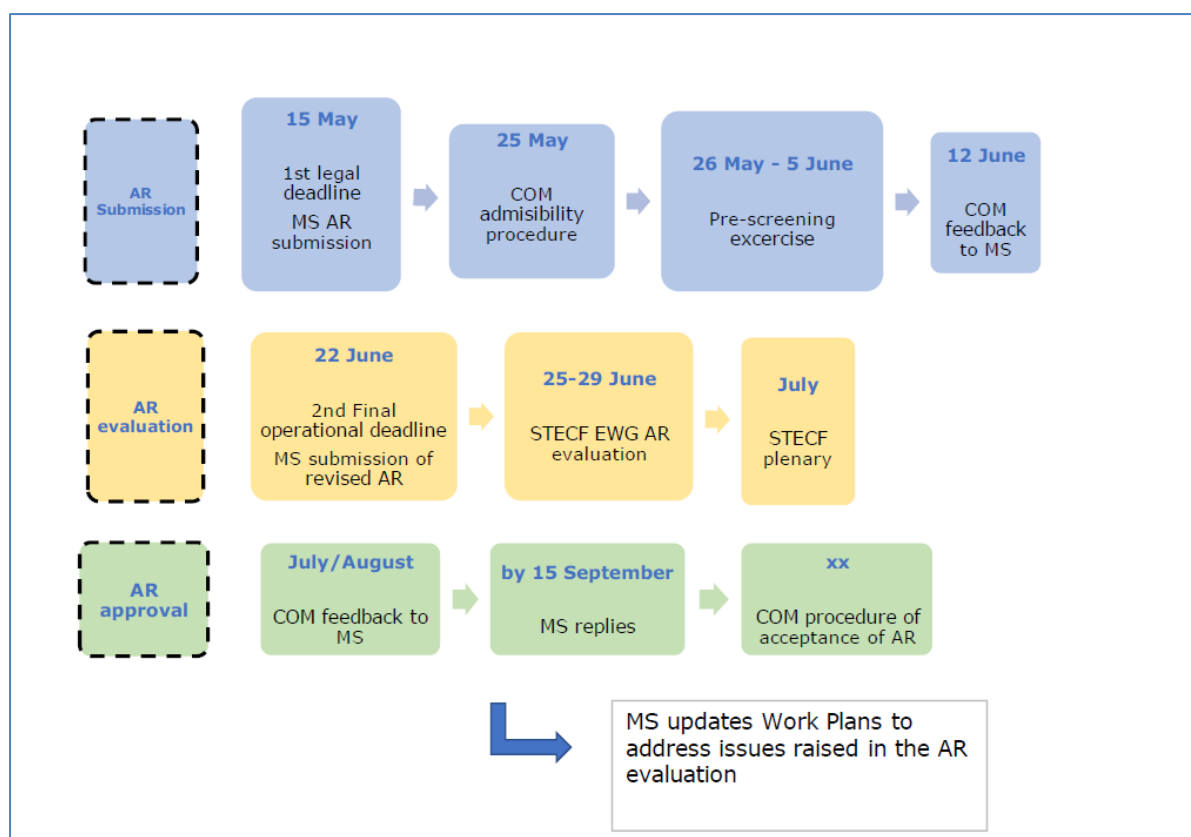


Figure 1 - AR annual cycle for evaluation dates - dates refer to 2018 for the AR2017 (STECF XXX)

The schedule is made up of three stages: the Submission, Evaluation and Approval stage. There are four points within those stages where automated checks could apply (i) pre-submission by the MS, (ii) pre-screening by STECF, (iii) pre-resubmission by the MS and (iv), the AR evaluation by STECF. MS will have their own screening exercises before the initial submission and before any re-submission. The STECF pre-screening captures any departures from what is expected, omissions, errors and insufficient explanations for departures from what is expected. It also screens for insufficient explanations for non-compliance. In the proposed schedule (Figure 1) MS are given the opportunity to correct or improve on the AR before the evaluation phase. Providing MS with a standard automated procedure where they can see what they need to do to improve their AR pre-submission, could make the pre-screening stage superfluous.

The choice of automation and the form that it would take whether within excel, R or a database/Web application was briefly discussed. Dealing with the present situation and establishing something relatively quickly is challenging, and on a practical level, including the NWP in the submission of the AR adds to the complication. There is the potential for transcription errors. MS could inadvertently provide the wrong NWP or reference the wrong values. Whatever checks were to be applied to the AR, they would need to be effective for pre-screening the NWP component as well. The NWP fields need to be checked before the AR fields can be compared.

Henrik Degel demonstrated a modified template developed in EWG1 which would help with manual data entry, cells are populated based on values in adjacent cells and or other tables, values were limited to lists and dropdown menus referring to external reference lists would help with completing them - simple flags that highlight a departure from what was expected. The spreadsheet is the standard format for the AR. The discussion about whether the automation should be carried out within Excel, or using external coding (R) or a database application was deferred until the group had looked in detail at what automation or checks would be required for these tables. Section 1.2.2

To be able to record each of the checks or consider what might be automated for each fields in each of the tables in the AR, the EWG first categorised the Check Types. Table 1 lists the check types that could be applied to any value in any combinations of fields and the consequences of a fail that would help MS pre-submission, pre-screeners and or evaluators. For ease the acronym of the Check Type was used as a reference ID in the detailed review.

Table 1- Types of checks that can be applied to each field, combination of fields in the

RefID	Check Type	Check description	Result	When are checks applied?	Needed
IVC	Individual value check	Accepted ranges for numerical values, code lists/drop down lists for others	Not allowed to enter unaccepted data. MS would need to contact administrators if code lists need to be expanded or ranges are incorrect.	Pre-submission	Reference lists (from WP and external sources)
ICC	Internal Consistency check	Do aggregations of a value from 1 AR table equal the aggregation of that value from a different AR table?	Mandatory comments	Pre-submission	
CC	Completeness check	Are all required data provided?	Not allowed to submit incomplete data	Pre-submission	

PIC	Problem Indicator check	Some fields are suggested that allow MS to indicate problems or incompleteness (e.g. "Deviation from fixed temporal range" in table 1G)	Mandatory comments	Pre-submission , pre-screening	
TC	Temporal check	Are the values for this year significantly different from the values for last year?	Mandatory comments	Pre-submission and pre-screening	Thresholds
ERC	External reference check	Do the values for submitted data match an external reference list? E.g. Official landings figures	Mandatory comments	Pre-submission and pre-screening	
WP	WP check	Compare the AR values to the WP values e.g. Are the achieved numbers significantly lower or higher than the planned numbers?	Mandatory comments	Pre-submission and pre-screening	Thresholds
BQC	Basic Quality Check	If response rates fall below a certain level then a comment is required	Mandatory comments	Pre-submission and pre-screening	Thresholds

Most of the checks types in Table 1 focus on the integrity of the data in the reports, primarily ensuring the data in the report reflects what MS have done sufficiently and without any errors. In terms of quality, the evaluator would be assured of the quality of the metadata describing the sampling schemes and the

achievements and the links to any external information or assessment of the sampling schemes.

With reference to quality and conformity, these checks cover the quality of how the templates are completed – a necessary first step in ensuring the underlying meta-data is correct and sufficient for evaluating the quality of a sampling scheme. Conformity was considered in terms of what was required to meet the regulation referencing the lists and thresholds provided in the draft Guidance and Decision 2016/1701 and the multi-annual EU plan.

For the detailed review a template was designed to record which of the Check Types listed in Table 1 are applicable to the fields in each table. The template was to provide a description of the checks; what the source for any check might be; when the check would best be applied; what the consequence of a fail might be; and where appropriate, an example of a pass and a fail. The template headers and descriptions are presented in Table 2 with two examples of how the table was completed.

Table 2 – Header descriptions for the template used by the EWG to describe the checks required for each field in each table of the Annual Report.

Template headers	Description	Examples	
WP or AR	Does the field in the AR template refer to historic NWP data (WP) or new AR data (AR)?	WP	AR
Table name(s)	What is the name of the table or tables for which check applies?	1A	4A
Name of the field(s)	What is the name of the field in the AR table?	MS	Achieved number of PSU in the reporting year
Link	What other tables are linked to this table by this field?	All tables	4A
Checktype	Reference ID for the Check Type?	IVC	WPC/BQC
Check Description	Describe what the actual check would be?	Limited to closed list	Achieved Nb PSUs.4A.AR ~ Planned Nb PSUs 4A.WP
DataSource/Thresholds	What is the data source for this value?	ISO 3166-1 alpha-3 code e.g. 'DEU'	difference of more than 25 %(?) between

			Achieved (AR) and Planned (WP)
DS exists? Link ?	Is there an electronic source for this and if so where?	https://unstats.un.org/unsd/methodology/m49/	
When applied	When would this check apply?	NWP pre-submission or pre-screening	Pre-screening
Consequence	What would happen if this did not pass the check?	Reject submission	Warnings
Warnings	What would be the warning or response?	Country list is not in the M49 standard list.	values differ by xx %
Example ok	What would pass?	FRA	350 in AR, 360 in WP
Example ko	What would fail?	FRO	50 in AR, 360 in WP

The EWG went into subgroups to complete tasks:

1. to review in detail the draft templates and guidance with reference to the rationale provided by 17-17(1) and the regulation.
2. to identify what checks could be applied to what fields within each table.

The intention was for the work to be carried out in subgroups but because of the size of the task the EWG members tended to work in pairs. Each table was considered independently but the links and the impacts on other tables was also recorded. Progress meetings were held once or twice each day in plenary to help ensure consistency.

The completed table is reproduced in ANNEX 3.

The completed CheckTemplate is too large to review easily so some of the considerations for Table 1A, 1C and 4A are documented below to provide examples of what was considered for all of the tables.

Table 1A

An automated approach for populating table 1A of the NWPs has already been developed by France. It was presented to the Regional Coordination Groups in 2016 (RCGNA 2016) and was used by a number of member states to populate their NWP submissions. It uses the EUROSTAT database for deriving the share of the landings and the MARE FIDES file, <https://webgate.ec.europa.eu/fides/index.cfm> for deriving the TAC share at EU level. Using the thresholds defined in the Decision 2016/1701 the flags on

whether a MS should sample can also automatically be completed. Using EUROSTAT limits its use for MS who do not provide their data but the code can be adapted to source other information. MS have had to adjust the figures to be more realistic to account for TACs and landings being shared across a number of species (e.g. TAC and landings are reported for Megrin but two species of Megrin are listed in Table 1A).

To check the NWP component of the AR, Evaluators could run the code for the same time period and compare one against the other but the evaluator would need an understanding of the data source and how the MS resolved the limits to the code when the Eurostat/TAC and DCF "stock" definitions are not aligned. This would be an issue for quality evaluation.

The AR only asks MS to flag any stocks when there has been a significant change in the annual landings from the reference 3-year average which could have affected achievements. An additional field in the report asking for the landings for each of the species/area for the report period and a field calculating the difference could answer that concern – however that would rely on the source data used for NWP being up to date. Eurostat is unlikely to have been updated by the submission date for the AR but the MS may be able to provide comparable if provisional figures. Further fields providing the source for the landings and TAC share would be informative. A threshold for the difference could be set to flag if exceeded and MS could comment if it affected sampling or not.

Table 1C

The review of Table 1C is provided in ANNEX 1. The automatic checks listed range from verifying compliance against the format, against reference lists, the coherence against the agreed NWP, the completeness of the sampling against the NWP and the deviation from basic quality standards (e.g. minimum numbers of individual to be measured).

Basic Individual Values Checks (IVC) for each column of the NWP of the table should be performed to verify the conformity of the table against given reference code lists or numeric ranges. Considering that values for MS, Species, Region, RFMO, Area/Stock could be reported in different ways, these checks are preparatory for the evaluation checks and assist the submitters warning them when codes used in filling tables are not syntactically correct or are not allowed.

As the NWP has effectively been resubmitted as part of the AR Individual Values Checks (IVC) on multiple fields between the NWP fields in Table 1C the evaluated NWP (on the JRC website) should be performed to verify the coherence between the two versions. The automation would invite the submitters to:

- correct values in case of copy-paste errors
- provide an explanation in AR Comments in case of missing/additional information in respect to the submitted WP
- provide an explanation in AR Comments in case of deviation from the submitted WP.

A Completeness Check (CC) would be required in Table 1C to verify if the information about any bi-and multilateral agreements related to MS participating

in sampling is accounted for in the "MS participating in sampling" field and invite MS to complete information before the submission.

Once the compliance and the completeness has been verified, checks on Work Plan (WPC) are foreseen both in pre-submission and pre-screening phase in order to give the submitters the opportunity to explain deviations from WP, if any, and the allow the evaluators to identify deviations or gaps in the WP. Specifically, in table 1C the checks verify:

- coverage of the species selected in the submitted WP for the sampling in the given MS-Species-Region-RFMO-Area/Stock (COMMISSION IMPLEMENTING DECISION (EU) 2016/1251, Table 1a-b-c, Stocks in Union waters).
- coverage of the species for which sampling of biological variable is agreed (COMMISSION IMPLEMENTING DECISION (EU) 2016/1251, chapter III)
- deviation in number of samples and number of measured individuals respect to the WP

Other automatic checks for the submitters and evaluators would be in reference to the past ARs in order to identify significant changes in numbers of individuals measured and allow the submitters to include comments in the AR and the evaluators to consider possible gaps in sampling.

Table 4A

Setting an exact deviation threshold over which conformity is an issue and "achieved" values are far from "planned", was not decided during the EWG. This would need a sensitivity analysis to be conducted on the historical WPs and ARs submitted. That would allow an "optimal" level of deviation to be identified over which DCF/EU-MAP implementation could be considered a failure or a success. For the time being, an arbitrary threshold of 25% deviation from "planned" has been suggested in the CheckTemplate spreadsheet to generate a warning message.

Besides internal consistency checks among the AR and NWP fields in Table 4A, numerous fields in Table 4A will also have to be cross checked against corresponding fields in Tables 1A, 1B, 1C, 1E, 1F, 4B, 5A, and 7C. Some of these cross-checks are introduced in the CheckTemplate spreadsheet.

The example of the achievements versus planned is given in Table 2. Achieved number of PSUs in the AR is checked against the planned value submitted in the NWP and if a difference of more than 25% is identified then a warning message is generated.

Summary

One field can have multiple Check types and others may not require any checks at all. There are a few inconsistencies in the overall table as some check types that could be applied may have been omitted. Further checks could also be included. Because some of the fields and contents are going to be consistent between tables there is some repetition. However, as a first draft this table **does provide a tool that any programmer or pre-screener or evaluator could use to see what might need to be applied to quality assure the data in the AR and the thereby the information you can take from it.**

The text table below summarises the checks identified for all the templates from Table 1A to 6A (excluding Table 4B-D and the extension to 1F). Overall for 246 fields/field groups (e.g. Species+Area) 399 checks were identified.

NWP or AR	Pre-screening	Pre-submission	AR evaluation	Unallocated	Grand Total
NWP	4	252	3	13	272
NWP and AR		6			6
AR	16	73	7	25	121
Grand Total	20	331	10	38	399

This CheckTemplate is preliminary and does need more work but the summary above highlights that most of the checks identified so far can be applied at a pre-submission stage and mainly refer to the fields associated with the pre-evaluated NWP.

This does not imply that there is little AR evaluation that can be automated it does suggest however that, at this stage, there are limited automated checks that would solely form part of any AR evaluation process.

Any evaluation checks could form the basis of any pre-submission checks or pre-screening checks. The pre-submission check would highlight to the submitter what action they need to take if a value in that field fails a check.

The checks for some fields would require a review of all NWP submissions to compile limited lists or reference ranges.

The same check could provide the prompt that the submitter needs to provide more information and once submitted provide the prompt for the evaluator to check that that information has been provided and it is sufficient.

Reference list for Thresholds: For automation a single source for the thresholds that might be used in an evaluation or for flagging an action from the MS would be useful. If they were either all in one place or visible on the spreadsheet it would be a useful reference for coding, data submitters as well as evaluators. Some of them may be arbitrary to flag an action or based on compliance thresholds quoted in the regulation.

3.2 Automation

To provide a balanced pragmatic and consistent approach to checking for conformity there needs to be consistency in the data submitted by MS. For conformity, evaluators need to see whether what the MS has done is what was agreed that they would do and to be able to review the quality of what MS are doing - assuming these tables allow you to do it. MS need to be able to submit sufficient information to describe what they have done relative to what was expected of them.

For any assessment of quality, the Meta-data must be correct so that there is a link running through each of these tables from the descriptions and summaries of the target populations for the data that is required, to the regulation and thresholds that may be applied to the surveys and sampling programmes collecting the data, to the documentation and quality framework applied to the sampling schemes and processes.

If the tables are too rigid and entry is limited to a list of variables or entries that do not account for a particular process or variable there is the danger that for completeness a MS might feel obliged to enter something "similar" to what they have done rather than what they have "actually" done. Better to have data that is correct rather than data that is forced to be wrong by the constraints on the submission. This would help in developing best practice on how to complete NWP templates and thereby answer some of the concerns that the current NWP templates do not capture what or how MS have designed or are managing their programmes.

Any pre-submission checks should give a MS sufficient warning of potential issues with their submission. If these issues persist into the evaluation then the MS would have had the warning and opportunity to provide sufficient explanation for any departure from what was expected.

The exercise of reviewing every field in the table and cataloguing what checks might be applied to those fields was extensive but not exhaustive and can be added to as NWP and AR templates develop and other evaluation criteria are adopted. For now, for each field, we have a preliminary list of the basic checks that would need to be applied to each field with an example of the result. Links to other tables, dependency on other fields, reference lists and external sources are also given. The AR fields that can be compared against the NWP fields have been identified once the NWP fields have been checked.

The different processes and current automated procedures used in data exchanges and some worked examples of what might be adopted in the future are presented in the following sections:

3.2.1 Current examples of international data exchanges using automated checks

Some tools related to fishery data are used in different RFMOs and IOs. A short description of them is given in this section.

3.2.1.1 Fisheries dependent information (FDI) validation tool

FDI data call asks for excel spreadsheet. A validation tool is provided by the JRC: the DVTool (<https://datacollection.jrc.ec.europa.eu/dc/effort>). This tool is a set of macros developed in Visual Basic Applications (VBA) and embedded in a specifically designed Excel Workbook. The DVTool checks codification and duplication problems, and operates cross-checks between tables where necessary. The excel template highlights lines containing errors using colour coding. The same validation tool applies also to various other data calls (e.g.: Mediterranean & Black Sea, Fleet-Economic).

3.2.1.2 The regional database FishFrame

The regional database FishFrame (<https://www.rdb-fishframe.org/>) provides quality check during the upload of new data. The data format and codification problems are tested and precludes the upload of the data without it being corrected.

3.2.1.3 Intercatch

Intercatch is a web-based system (<https://intercatch.ices.dk/>) where biological and catch data relating to fisheries and stocks are uploaded to feed stock assessment analyses in ICES. During the upload of national data, the data format, and codification of the variables are tested. A cross check between catch tables and length or age distribution is made using the sum-of-product procedure: if the weights of catches in different tables diverge by more than 20% the user has to correct the data submitted. Detected errors ask for correction and preclude the transmission of the data.

3.2.1.4 GFCM Data collection Reference Framework online platform

The DCRF online platform of the GFCM (<https://gfcmlsharepoint.com>) provides the submitter with a spreadsheet interface to upload the data - based on excel. The Excel templates include quality checks and quality indicators that describe the number of reported data rows, the number of cells completed correctly, the number of cells with errors or missing values and the number of duplicate values. Detected errors warn the user but do not preclude data transmission.

3.2.1.5 FishPi tools

During the fishPi project (MARE/2014/19), one of the deliverables developed guidelines to evaluate the quality of data at national and regional levels using shared tools (<https://github.com/ldbk/fishPifct>). From the format and check definitions given in a spreadsheet file and the data to be checked, the R script will test data integrity and perform the quality checks specified by the user and deliver an automatic report.

3.2.2 Future automation of the AR and NWP evaluation cycle

The completed CheckTemplate provides the guidelines for data checkers or programmers if we want to apply automated procedures in our AR evaluation or screening.

There are three relatively simple approaches that could be adopted for applying checks:

1. Excel (or compatible) spreadsheets
2. Shared code which, based on stored and shared reference lists, criteria and data exchange formats, can report and check NWPs and ARs
3. Web based application with a store of submitted NWPs and ARs

They are considered here in isolation but they may all form part of an adopted process. How each may be used is illustrated in the sections below.

JRC has the facility to host a database but not necessarily the resources or 'appetite' to do so.

Database construction is an option but a simple approach may not stop at being simple it needs to be maintained and users need support. Something apparently simple can turn into a monster and can take too long to get started.

FDI datacalls, and uploads to intercatch stop the uploads if there is an error or issue with one data item. Assessing the Meta-data for these reports needs to be less restrictive than FDI datacalls.

We do not necessarily need to be too restrictive with ARs in the first instance. We would apply the checks to the Meta-Data pre-submission. The system should not block a pre-submission but would report back on it – a feedback process where the report provides an indication of where the errors might be in what they have submitted.

Ideally, we would have an interactive system where MS can review their uploads and reports and re-submit based on those responses before finally submitting for evaluation.

R code offers a swift option where everyone's submissions, data reference lists and thresholds can be stored independently and linked to when required.

The Web application is just a tool as is the Markdown in R. In the examples below all the processes start with a submitted report in excel format. Both the R example and the Web App examples refer to existing but edited NWP for demonstration.

The whole process would be enhanced by tying regional submissions to a RDB which should hold or have records of all the biological data and transversal data collected by a MS under the DCF in the reported year. RCMNA (2014) reviewed differences between the two reference datasets, the AR submissions and the data held on the RDB. Although the regional DB was under development and there were upload issues, there were marked differences in the comparisons which informed the RDB developers and the RCGs on how to improve on their data calls and improve the upload process. If the data on the RDB is quality assured then it would be a source for AR checks and conformity as well as quality evaluation and could also be used as a source for the template data itself (Section 3.2.2.4).

3.2.2.1 Using Excel spreadsheets for validation

Drop down lists or limit to lists only work in a standard Excel spreadsheet if the data is being entered a cell at a time. For a number of these tables the data entered will be based on exports or other compiled data and these will be copied to the AR sheet a block at a time. The FDI and GFCM (see previous section) offer a more complex but stricter approach before the data is uploaded where coded excel spreadsheets or web based spreadsheet interface provides direct indicators if there is an issue with the cell entry or comparisons with other data.

The existing spreadsheets constituting the present National Work Plan and the new Annual Report can be modified so that to become more user-friendly and support a more efficient evaluation by introducing a number of functionalities to the spreadsheets. Table 1G and Table 1H are used as examples of how the various spreadsheets in the Annual Report could be automated. A working

example is available on in Annex 4 – *Extended functionality of AR template .xism*.

The functionalities suggested in the sections below are embedded in the spreadsheets which is a demonstration of how the various functions would be experienced by the user and a suggestion of how the spreadsheet might be automated.

In EXCEL, it is possible to lock selected cells in order to prevent the user modifying the essential parts of the spreadsheet and hereby make the Workbook inconsistent with the format used for general evaluation and crosschecking. This allows the user to enter data and add parameter values to fixed reference tables without corrupting the necessary consistency (see below under the section “Drop-down Menus”).

3.2.2.1.1 Cover sheet

A cover spreadsheet has the function that users are able to state basic parameter values such as “Member State”, “AR Data Year” and “Date of submission” in one sheet. The parameter values are then automatically allocated to all relevant cells. This assures that the values are consistent throughout the whole workbook. In addition, information about contact persons can be given in the cover sheet.

3.2.2.1.2 Drop-down Menus

If automatic checks are to be introduced, it is of vital importance that all values (except numbers) comply with some naming rules of the parameters. Building on the present spreadsheet solution (as an intermediate situation on our way to a more permanent database solution for example) the use of “Drop Down menus” are a possibility. A drop down menu presents the list of possible values as shown in Figure 2 and the user then selects a value among this closed list of valid values. This assures that values are consistent within a spreadsheet and across Member States.

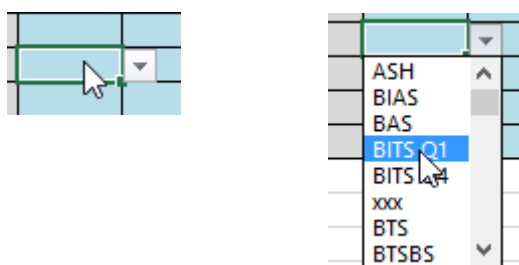


Figure 2 - Drop-down Menu listing the possible valid survey acronyms. The Drop-down possibility is activated by selecting the field (left). The value is selected by selecting the wanted value in the list. (Right) The list can be scrolled down by the bar to the right.

It is possible as well to write a value in the cell associated with a Drop-down Menu as long as the value is included in the Drop-down list. In order for the user to save time and effort, it is possible as well to copy a given valid value to a number of entry lines below.

To prevent the user to be trapped in a situation where no predefined value is available to describe the data, it should also be possible for the user to add a value to the list of valid values if the user keeps book of the updates in a designated sheet (e.g. in the Cover sheet). Based on that sheet, the "official" menu can be updated in future templates.

The Drop-down tables should be defined in a special spreadsheet for convenient administration.

3.2.2.1.3 Automatic fill of cells

The work of filling in the template can be significantly reduced, if the value in a given cell can be automatically derived from one or more parameter values already filled in e.g. if the survey acronyms is given, then the name of the survey, List of MS participating, the total area covered by the survey, the relevant international planning group - RFMO/RFO/IO and the relevant international database can be automatically filled in saving time and preventing errors. This functionality is based on a lookup table defining the relationship between the different parameter values. The look-up tables should be defined in a special spreadsheet together with the Drop-down tables

3.2.2.1.4 Indications that further information is required

Time constraints and efficiency are very important factors in meeting a tight schedule from the Member State submitting the Annual Report to the final approval. A typical table in the AR often contains so many lines it is not easy to see where AR-comments are required or not based on discrepancies between measures in the NWP and the AR. Therefore, **an automatic indication in the spreadsheet, which implies the Member State is required to provide an AR-comment, will make it easier for the Member State to complete all justified comments before the submission and easier for the evaluator to see where comments are needed saving time and effort in communicating with MS and resubmissions.** The criteria for requesting an

AR-comment could be if the discrepancy in selected measures in the NWP and the AR exceeds plus/minus a certain percentage.

MS	Name of survey	Acronym	Mandatory (Y/N)	Threshold (Y/N)	Agreed at RCG level	List of MS participating	Type of MS participation	Total area covered by the survey	Total area covered by the survey (Additional)	Month start	Month finish	Frequency	Days at sea planned	Type of main sampling activity	Planned target	Map	Relevant international planning group - RFMO/RFMO/O	Relevant international database	WP Comments	Days at sea achieved	Achieved target	Other data assimilations	Deviation from fixed temporal range	Deviation from fixed spatial range	Indication if AR comments by MS are required concerning effort	Indication if comments by MS are required concerning percentage
DNK	Baltic International Trawl Survey	BITS Q1	Y	N	Y	DNK, DEU, POL	Physical	SD 27.25-32	SD 27.22	March	March	Annual	25	Fish Hauls	50	Fig 7.1	ICES WGBIFS	DATRAS		25	53	Y	N	N		
DNK	International Bottom Trawl	IBTS Q1	Y	N	Y	DNK, UK, SCO	Physical	27.IVa	27.IIIa	Feb.	Feb.	Annual	17	Fish Hauls	48	Fig 7.2	ICES IBTSWG	DATRAS		17	10	Y	N	N	x	

Figure 3 - Supporting information. An indication (Table G1) is given (x) because the discrepancy between "Planned target" and "Achieved target" exceeds 5 pct.

These thresholds, which induce comments, could be more easily handled and adjusted if they were all held together on a reference sheet within the template.

3.2.2.1.5 Relation to the NWP

It is a requisite for most of the functionalities in the AR that the NWP as well being consistent follows the same naming rules as is suggested for the AR. It is therefore suggested that the NWP is subject to the same naming rules and that similar functionalities are introduced at the earliest opportunity. In general, the same type of functionalities can be embedded in the NWP as is suggested for the AR.

Some parameters in the NWP are a mix of several basic variables. For instance, can the parameter: "Type of data collected" be considered as a mix of "Data Category", "Data Object" and "Area Covered by the MS"? (Figure 4). This will make it possible to compare the data collected with the need for assessment data. Furthermore, the split will make comparisons across Member States easier and link to calculations by species which can be used in the Regional context.

AR date of submission				
n	Type of data collected			
	Data Category	Data object	Area covered by the MS	Additional area covered by the MS (if needed)
1	Biological data	Cod	SD 27.25-32	SD 27.22-24
1	Environmental data	Litter	SD 27.25-32	SD 27.22-24
1	Oceanographic	CTD	IVa	IIIaN

Figure 4 - The original parameter "Type of data collected" can be split into three basic variables.

3.2.2.2 Using R for validation

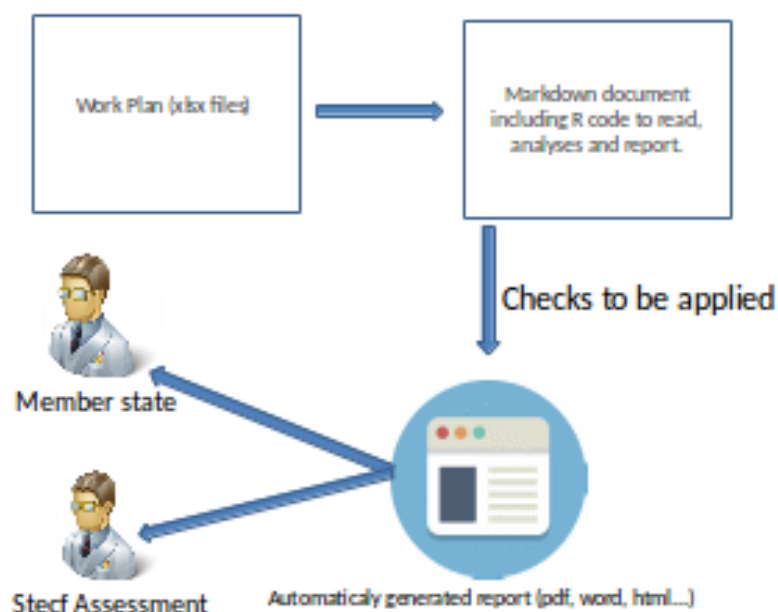


Figure 5 - information flux diagram between work plans

3.2.2.2.1 Framework

A simple validation procedure was developed in order to provide a simple, open-source and transparent tool to end-users (assessors and member-states, Figure 5). The validation procedure is embedded in a markdown document including R code. This framework is aimed for reproducibility and transparency, following the recommendation of the reproducible research statement (Gentleman and Lang 2004). Consequently, the document is self-consistent: the code used to process and to analyse the annual report are embedded in the document itself.

Analyses are carried out using the R environment (R Core Team 2017). R is a free software environment for statistical computing and graphics. The reproducibility of the results presented in the report relies on the use of a dialect of the Markdown language called Pandoc for word processing using the Knitr R package. Markdown is a plain text formatting syntax designed so that it can optionally be converted to HTML using a tool by the same name. Pandoc is a Markdown dialect which extends the conversion capability to word processing file (docx, doc and odt) and pdf, among other formats. Pandoc understands a number of markdown syntax extensions, including document metadata (title, author, date), footnotes, tables, figures and references. Knitr is an R package (a set of functions extending the R capabilities). The R code used to process and analyze the data is included directly in the report. Results are then produced dynamically. This framework has demonstrated the capacity to improve the conduct and the presentation of data analysis in a way that another person can understand and replicate (Baumer et al. 2014).

For example, if the calculus of $1+1$ is needed, the code to compute it is written in the report using special hooks, as in this simple example:

```

'''{r                                     test00,warn=FALSE,cache=TRUE,echo=TRUE}
#comment:                               addition                               example.
1+1
'''

```

This code is evaluated during the compilation of the report by the knitr command and it prints the following result:

```

#comment:                               addition                               example.
1+1

## [1] 2

```

The result is 2. All the numerical values related to the quality checks of the annual report, including tables and figures are generated following these steps.

3.2.2.2.2 Test

The procedure was tested on the national work plans available in <https://datacollection.jrc.ec.europa.eu/wps> for 4 variables of the table 1A.

3.2.2.2.3 Data

Work plan in excel files of 24 countries were read and saved in an R object of class named "wp", encompassing all the tables available in the work plans. This class was created on purpose for the work plan.

Code

```

#read the original wp and build an object from it
fichwp<-"./data/EWG_16-01_tables.xlsm"
sheet<-getSheetNames(fichwp)
sheet<-sheet[grepl("Table",sheet)]
sheetname<-gsub(" ","_",substr(sheet,1,7))
#build a wp class
slots0<-paste0("slots=list(",paste0(sheetname,"='data.frame'","collapse="))
slots0<-paste0(substr(slots0,1,nchar(slots0)-1),",")
proto0<-
paste0("prototype=list(",paste0(sheetname,"=data.frame(),"collapse="))
proto0<-paste0(substr(proto0,1,nchar(proto0)-2),"))")
eval(parse(text= paste0("setClass(Class='wp',"slots0,proto)))
#read the main object
wp<-new("wp")
for(i in 1:length(sheet)){
  tmp<- read.xlsx("./data/EWG_16-
01_tables.xlsm",sheet=sheet[i],startRow=4)
  eval(parse(text=paste0("wp@",sheetname[i],"<-tmp")))
}

wpall<-new("wp")
listwp<-dir("../wp",patt="xls",full=T)
listwp<-listwp[grepl("WP_",listwp)]
#listwp<-listwp[!grepl("Belgium",listwp)]
for(j in listwp){

```

```

print(j)
for(i in 1:length(sheet)){
  try(tmp<- read.xlsx(j,sheet=i,startRow=4),silent=T)
  try(tmp<-tmp[!is.na(tmp[,1]),],silent=T)
  try(names(tmp)<-
gsub("?", "", gsub(")", "", gsub("(", "", gsub(",", "", gsub("%", "", gsub("/", "", name
s(tmp)))))),silent=T)
  try(eval(parse(text=paste0("wpall@",sheetname[i],"<-
rbind(wpall@",sheetname[i],"tmp)"))),silent=T)
  try(rm(tmp),silent=T)
}
}
## [1] "../wp/WP_Austria_2017-2019.xlsx"
## [1] "../wp/WP_Belgium_2017-2019.xlsx"
## [1] "../wp/WP_Bulgaria_2017-2019.xlsm"
## [1] "../wp/WP_Croatia_2017-2019.xlsm"
## [1] "../wp/WP_Cyprus_2017-2019.xlsm"
## [1] "../wp/WP_Denmark_2017.xlsx"
## [1] "../wp/WP_Estonia_2017-2019.xlsx"
## [1] "../wp/WP_Finland_2017-2019.xlsm"
## [1] "../wp/WP_France_2017-2019.xlsx"
## [1] "../wp/WP_Germany_2017-2019.xlsx"
## [1] "../wp/WP_Greece_2017-2019.xlsm"
## [1] "../wp/WP_Hungary_2017-2019.xlsm"
## [1] "../wp/WP_Italy_2017-2019.xlsx"
## [1] "../wp/WP_Latvia_2017-2019.xlsx"
## [1] "../wp/WP_Lithuania_2017-2019.xls"
## [1] "../wp/WP_Malta_2017-2019.xlsm"
## [1] "../wp/WP_Netherlands_2017-2019.xlsm"
## [1] "../wp/WP_Poland_2017-2019.xlsx"
## [1] "../wp/WP_Portugal_2017-2019.xlsx"
## [1] "../wp/WP_Romania_2017-2019.xls"
## [1] "../wp/WP_Slovenia_2017-2019.xlsx"
## [1] "../wp/WP_Spain_2017-2019.xlsm"
## [1] "../wp/WP_Sweden_2017-2019.xlsm"
## [1] "../wp/WP_UnitedKingdom_2017-2019.xlsx"

```

3.2.2.2.4 Quality checks on table 1A

For the variable MS (Member state), the script tests if the MS is in the iso-9989 list. If not, it gives to the user the number and the type of errors.

For the variable Reference years, only the value "2017-2019" is considered as correct. If not, it gives to the user the number and the type of errors.

For the variable Species, only the value included in a reference list built using the species reported in the work plan and manually corrected are considered to be correct. If not, the script gives to the user the number and the type of errors.

For the variable Selected for sampling, only Y (for yes) or N (for no) are expected. If not, it gives to the user the number and the type of errors.

The results are summarized in a table for each country. The code and example of the output is given below.

Code:

```
# `` ` {r check1A, include=T,cache=T,echo=F,results="asis"}
fct<-function(wpull){
#MS
ms<-read.csv("ms.csv")
checkms<-all(wpull@Table1A$MS%in%ms$code)
whichwrong<-which(!wpull@Table1A$MS%in%ms$code)
idmswrong<-
unique(wpull@Table1A$MS[which(!wpull@Table1A$MS%in%ms$code)])
checkmswrong<-length(whichwrong)
#refyear
checkrefyear<-all(as.character(wpull@Table1A$Reference.year)%in%c("2017-
2019"))
whichwrong<-which(!wpull@Table1A$MS%in%c("2013-2015"))
idrefyearwrong<-
unique(wpull@Table1A$Reference.year[which(!wpull@Table1A$Reference.year
%in%c("2017-2019"))])
checkrefyearwrong<-length(whichwrong)
#Species
spp<-read.csv("Species .csv")
spp1<-spp%>%filter(!grepl("spp",code))
spp1<-spp1%>%filter(!grepl("\\(",code))
checkspp<-all(wpull@Table1A$MS%in%spp$code)
whichwrong<-which(!wpull@Table1A$Species%in%spp$code)
idsppwrong<-
unique(wpull@Table1A$Species[which(!wpull@Table1A$Species%in%spp$code
)])
checksppwrong<-length(whichwrong)
#region
region<-read.csv("Region.csv")
checkregion<-all(wpull@Table1A$Region%in%spp$code)
whichwrong<-which(!wpull@Table1A$Region%in%region$code)
idregionwrong<-
unique(wpull@Table1A$Region[which(!wpull@Table1A$Region%in%region$cod
e)])
checksppwrong<-length(whichwrong)
#selected for
checkyesno<-all(wpull@Table1A[,7]%in%c("Y","N"))
whichwrong<-which(!wpull@Table1A[,7]%in%c("Y","N"))
idyenowrong<-
unique(wpull@Table1A[which(!wpull@Table1A[,7]%in%c("Y","N")),7])
checkyesnowrong<-length(whichwrong)
#refyear
rez<-data.frame(variable=c("MS","Reference year","Species","Selected for
sampling"),
  test=c(checkms,checkrefyear,checkspp,checkyesno),
```

```

nbwrong=c(checkmswrong,checkrefyearwrong,checkspwrong,checkyesno),
  idwrong=c(paste(idmswrong,collapse=","),
    paste(idrefyearwrong,collapse=","),
    paste(idsppwrong,collapse=","),
    paste(idyesnowrong,collapse=","))
  )
  )
  return(rez)
}

wp<-new("wp")
rez<-data.frame()
listwp<-dir("../wp",patt="xls",full=T)
listwp<-listwp[grepl("WP_",listwp)]

for(j in listwp){
  #print(j)
  for(i in 1:length(sheet)){
    try(tmp<- read.xlsx(j,sheet=i,startRow=4),silent=T)
    try(tmp<-tmp[!is.na(tmp[,1]),],silent=T)
    try(names(tmp)<-
gsub("?", "", gsub(".", "", gsub("(", "", gsub(")", "", gsub("%", "", gsub("/", "", name
s(tmp)))))),silent=T)
    try(eval(parse(text=paste0("wp@",sheetname[i],"<-
rbind(wp@",sheetname[i],"",tmp)"))),silent=T)
    try(rm(tmp),silent=T)
  }
  cat("\n")

cat(paste0("# #",gsub("_2017", "", gsub("xlsm", "", gsub("xlsx", "", gsub("_2017
-2019", "", gsub("../wp/WP_", "", j))))))
cat("\n")
reztmp<-fct(wp)
#reztmp$id<-j
pander(reztmp,split.tables=Inf)
cat("\n\\newpage")
}

```

Example of output for some of the NWPs (full output is provided in Appendix 2)

Austria.

variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	TRUE	1	
Species	FALSE	1	
Selected for sampling	FALSE	0	NA

Belgium.

variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	TRUE	1	
Species	FALSE	1	
Selected for sampling	FALSE	0	NA

Bulgaria.

variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	11	2013-2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA

Croatia.

variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	11	2013-2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA

3.2.2.3 *Using Web app for validation*

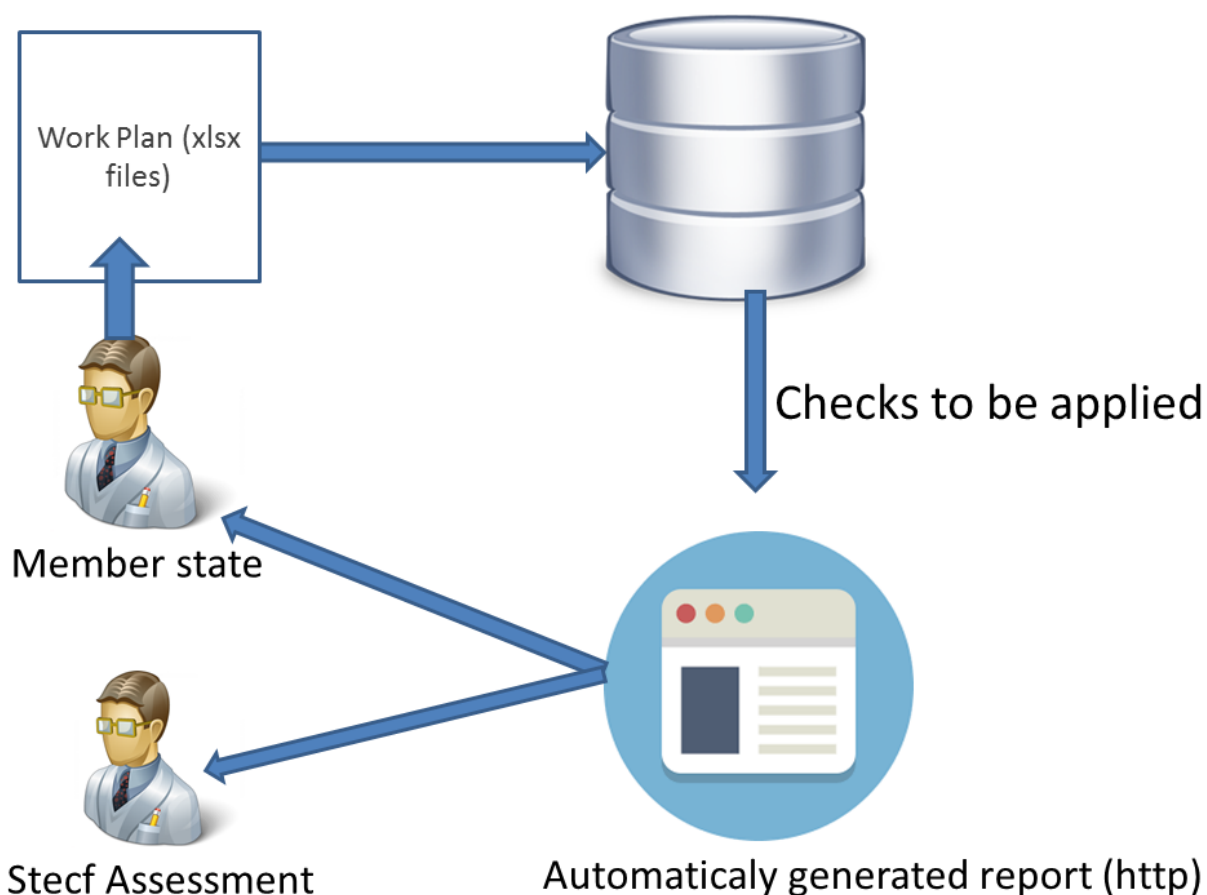


Figure 6 - Proposed information flow diagram

3.2.2.3.1 Introduction:

To set up a partial automated check on the Annual Reports we first have to store information in a common database. Storing the submitted Reports will allow queries to be performed on them based on the checks listed by this expert group in the completed CheckTemplate. Sql can be used and is a simple language which is relatively easy to interpret and more transparent to the uninitiated.

In this exercise we used Postgres opensource software to create a database and a set of php script to develop a web based application. Table 3A from the French and German NWP's were used and edited to capture discrepancies.

3.2.2.3.2 Step 1: moving Excel information to a common database

We used R script (to be finalised) to read all the NWP's provided by STECF and to store the information of each script to a common database.

```

setwd("Where the files are")

library(xlsx)
library(RPostgreSQL)
-- connexion to a local database performed by Postgresql software

```

```

con<- dbConnect(PostgreSQL(),host="localhost", port=5432,user= "x",
password="x", dbname="ar_report")

--List of the available Report
liste<-dir('.', pattern='xlsx')

for (fichier in liste)
{
pays<-strsplit(fichier,'_')[[1]][2]
tout<-loadWorkbook(fichier)
liste_sheet<-names(getSheets(tout))

for (sheet in (liste_sheet))
{

tmp<-read.xlsx(fichier,sheet,startRow=4,header=TRUE)
tmp<-tmp[!is.na(tmp$MS),]
--we add a unique number of each lines, the country included in the file name
and the date of submission
donnees<-cbind(seq(1:dim(tmp)[1]),pays,tmp[,c(-13,-12)],Sys.time())

sheet_new=tolower(gsub(" ","",substr(liste_sheet[sheet],1,8))) -- no blank and
uppercase in names of table
column_name<-names(dbGetQuery(con,paste( "select * from ",sheet_new,'
limit 1',sep=")))
names(donnees)<-column_name -- column names of data are the same as the
table
dbWriteTable(con,sheet_new,donnees,append = TRUE, row.names = FALSE) --
send the data frame in the table
}
}

```

3.2.2.3.3 Step 2: adding reference tables to the database.

For each field a reference list would be required. A table would be created and populated with allowed values (These are defined in the check list based on the guidance and EU MAP). The example below shows the limited number of values allowed for the Variable field in table 3A of the NWP and AR.

	variable [PK] character varying(100)
1	Average price per species
2	Consumption of fixed capital
3	Days at sea
4	Energy consumption
5	Energy costs
5	Engaged crew
7	Gross value of landings
3	Investments in tangible assets, net
9	Long/short Debt
0	Mean age of vessels
1	Mean LOA of vessels
2	Non-variable costs
3	Number of fishing enterprises/units
4	Number of vessels
5	Operating subsidies
6	Other income
7	Personnel costs

Figure 7 - view of ref_economic_var table. It include all the variable required and defined in EU Decision(EU) 2016/1251 table 5A

3.2.2.3.4 Step 3: from check list to SQL query

In reference to the CheckTemplate table (ANNEX 3) a sql query could be created to carry out those checks.

Check done	Table Name	Fields involved	Checktype	Check description	Data sources/threshold	Link if available	When applied	Consequence
1 3A		MS	IVC	content of the fields has to be included in reference list	ISO 3 letters code	http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016D1251&qid=1508852008005&from=FR	Pre-submission	Not allowed

The first check in the list for table3A (our case study) is to look at the MS column to see if values are included in the reference list of member state (ISO3 country code). We deliberately edited one incorrect record so that to check if the query captured the error. The SQL for the check and the result are below

```
--Last submission for France
with preselection as (
select * from table3a
where submission_data=(select max(submission_data) from table3a where pays
like 'France')
and pays like 'France' )
, --extract lines from preselection for where the MS value is not included in
ref_ms table
```

```

check_list as (
select distinct 'MS code list problem' as Checks,row_names
from preselection A left join ref_ms B on (A.ms=B.ms)
where B.ms is null
)
-count the number of lines with warning about MS values and details provide
line number of each

select distinct checks as ABSCISSE,count(*) as TOTAL ,string_agg(row_names,'-
') as details from check_list group by checks

```

	abscisse text	total bigint	details text
1	MS code list problem	1	1

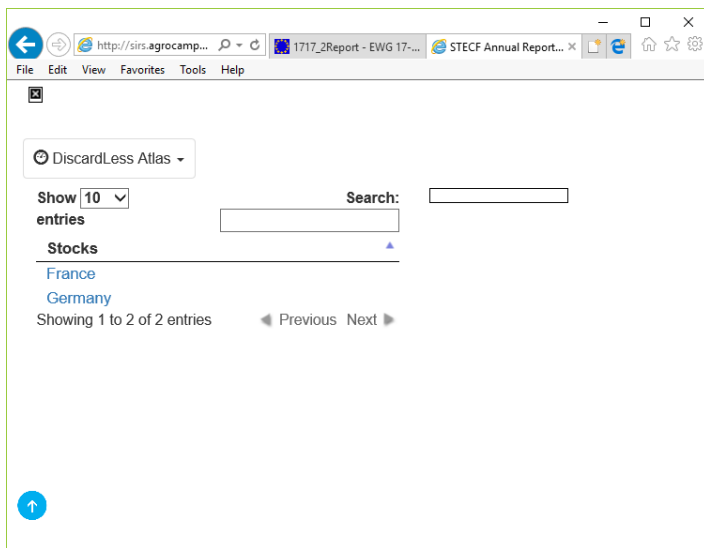
Figure 8 - result of the query

3.2.2.3.5 Step 4: from one query to a report.

Once the check list has been translated into SQL, we used php script to produce the report. Here for this example we have used an on line atlas application already used for the *STECF Balance between fleet capacity and fishing opportunity working group* - http://sirs.agrocampus-ouest.fr/stecf_balance_2017/.

This is just a sample implementation that could give some ideas on the way to produce reports and not all the indicators refer to real data. The example as a proof of concept is available here - http://sirs.agrocampus-ouest.fr/sirs_ar/

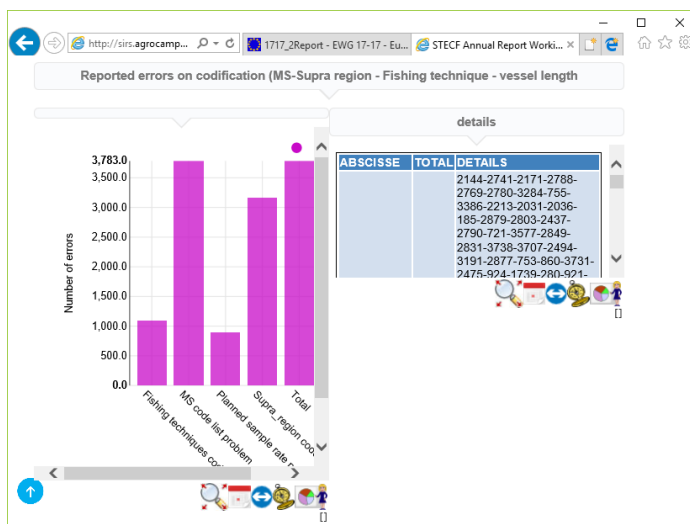
Screenshots from the application are shown below and provide examples of how errors or issues may be displayed:



[Choose the country...](#)

View the report....





Fleets without required Economic variable

FLEET	VARIABLE
DEU-Baltic Sea, North Sea and Eastern Arctic, and North Atlantic-Beam trawlers*-10m-<12m	Average price per species
DEU-Baltic Sea, North Sea and Eastern Arctic, and North Atlantic-All-All	Investments in tangible assets, net
DEU-Baltic Sea, North Sea and Eastern Arctic, and North Atlantic-Inactive vessels-40 m or larger	Engaged crew
DEU-Baltic Sea, North Sea and Eastern Arctic, and North Atlantic-Inactive vessels-40 m or larger	Engaged crew

Fleets without required Social variable

FLEET	VARIABLE
FRA-Baltic Sea, North Sea and Eastern Arctic, and North Atlantic-Demersal trawlers and/or demersal seiners *-0-< 10 m	FTE National
FRA-Baltic Sea, North Sea and Eastern Arctic, and North Atlantic-Purse seiners *-18-< 24 m	Employment by education level
FRA-Baltic Sea, North Sea and Eastern Arctic, and North Atlantic-Vessels using active and passive gears *-10-< 12 m	FTE by gender
FRA-Baltic Sea, North Sea and Eastern Arctic, and North Atlantic-Vessels using active and passive gears *-10-< 12 m	Employment by

3.2.2.4 Utilising the RDB as a source of metadata for STECF evaluation of National Work plans

Detailed sampling data from 3 RCGs (North Atlantic, North Sea and Baltic) is stored in a common format within 3 instances of the Regional Database (RDB). This data can form the basis of metadata reports relating to some aspects of both the execution of the national work plan and metrics of data quality.

The RDB data could be used for a number of purposes including:

1. Pre-filling elements of the Annual Report by STECF,
2. Make pre-written RDB data extraction routines available to MS so they can insert the data into the AR template themselves,
3. Directly cross-checking data submitted for the Annual Report,
4. Providing a complementary data source and reports for Annual Report evaluators to use.

Both the requirements to a) evaluate the execution of the national Work Plan and b) the quality of the data collected by member states, fall within the remit of the STECF. The STECF 17-11 report notes that regional data bases should be used to facilitate MS producing tables for the Annual Report and to inform STECF during evaluation of the quality of the data collected by MS.

The process of achieving this aim will involve a number of steps:

- If RDB data is to be used for the Annual Report compilation or evaluation processes then the timing of the RDB data call must be early enough so that it is completed before the data is required for the Annual Report.
- Data uploaded to the RDB needs to be identifiable to the National Workplan commitments set out in table 4A, dealing with sampling designs, and potentially 1C dealing with the biological data collection and for the required species. This will involve changes being made to the current RDB data format.
- If the AR and RDB data did not match then thought must be given to how corrections should be made e.g. should the RDB data be re-submitted and used to re-populate the AR or is it sufficient to submit correct AR data and allow the RDB data to be inconsistent?
- The drafting of an Annual Report template which is compatible with being populated in part or in full, through RDB derived data. This template should however not be such that the formats preclude MS which do not use the RDB.
- The automated checking of the elements of National Work plan templates with the corresponding elements of Annual Report templates. This could be done on a detailed level if the RDB format matched that used in the AR, or at an aggregated level if the formats don't align (e.g. the total number of samples per species could be checked rather than checking the number of samples by strata).
- Guidelines for the assessment of the elements of any automated procedure as to their relevance, how they should be interpreted, and the insights that they can, or cannot, provide about the execution and quality of the MS data collection.

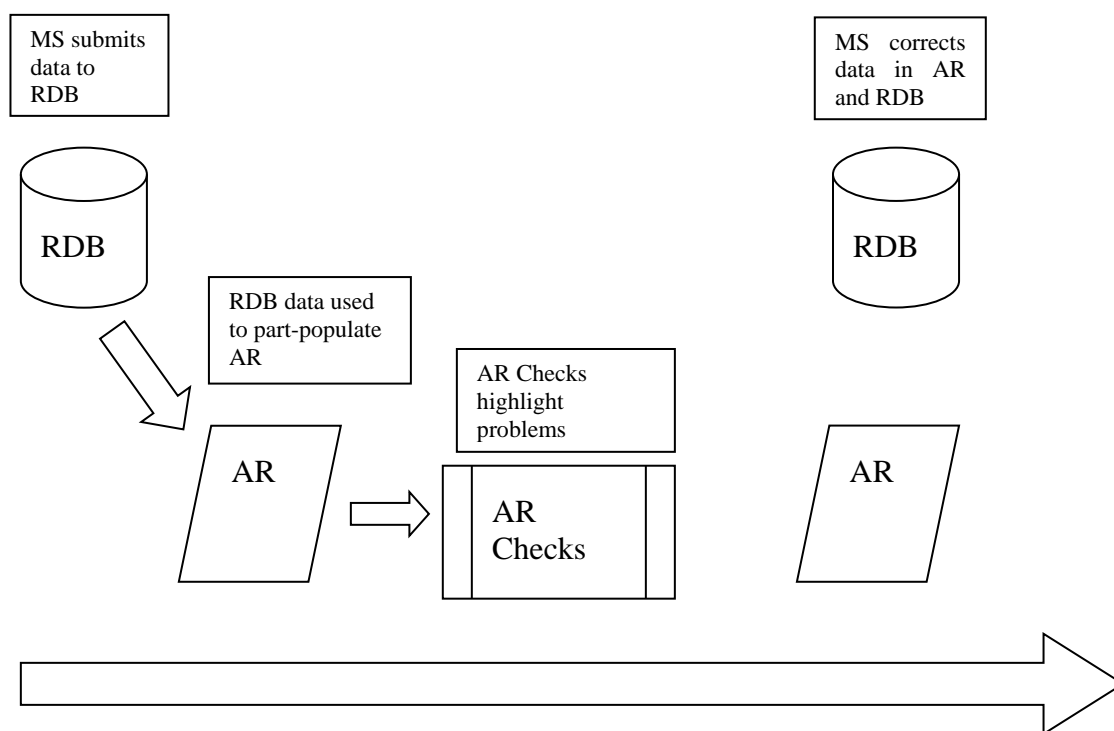


Figure 9 - Possible RDB/AR submission process

It is recognised that an automated process can have huge efficiency and quality benefits, but it is one that cannot, and should not, be imposed on MS. Rather it is a process that can be developed and refined over time.

The RDB also has its own data quality checks and validations so this could improve the quality of the data in the AR without needing to create new AR-specific data quality checks

An important point to make is that the Annual Report is of a much higher importance to MS than the RDB, therefore it is to be expected that the data submitted in the AR is more accurate than that currently in the RDB. Therefore if discrepancies are detected between the AR and RDB initially it is likely that the AR is correct – it would be expected that this would change in future years so that using the RDB in the AR process would have the side-effect of increasing the quality of RDB data.

Example

The data for 2016 in the RDB cannot be directly related to National Work plans, but a simple audit of the sampling data by sampling country can provide an indication of the potential of the automation process. These tables are generated using the R function `sampInv` (which was also used to summarise the sample data). Figure 10 quantifies the sampling by sampling country, and sample type; Figure 11 summarises the sample numbers by species for a particular country.

sampCtry	sampType	days	sites	vessels	trips	nspp	lengthFreq	bioSites	bioTrips	bioSpp	wght	ages	lengths	sex	maturity
BEL	S	157	9	16	34	15	403677	9	32	7	0	4215	4215	0	0
BEL	D	NA	NA	NA	NA	NA	NA	9	33	7	9106	9106	9106	8593	6903
DEU	S	220	13	10	33	107	235321	32	128	15	23194	27088	29973	18429	21985
DNK	S	242	28	128	341	96	210771	27	336	59	53786	17830	53786	12424	88
DNK	M	231	36	258	811	32	79557	36	812	31	52561	23570	52561	5243	3733
DNK	D	14	11	1	17	9	1164	11	17	9	752	343	752	56	25
ENG	S	230	42	134	223	128	308090	28	150	10	0	3037	3143	3022	0
ENG	M	311	67	560	1972	52	271891	42	871	18	1905	21318	24745	15593	3427
ENG	D	2	2	2	2	1	199	NA	NA	NA	NA	NA	NA	NA	NA
ESP	S	331	32	128	281	182	302766	4	12	32	34189	4981	34239	33558	32468
ESP	M	273	22	447	1313	146	408772	12	136	12	7317	5645	8357	6953	7047
ESP	V	NA	NA	NA	NA	NA	NA	19	272	18	14615	11660	17632	15066	8946
EST	M	95	14	23	137	6	22215	14	144	6	15755	15078	15755	14756	14849
FIN	M	94	1	1	124	26	46087	1	116	5	4021	3647	4021	3845	3564
FIN	S	115	1	1	186	27	10955	1	185	8	5307	2597	5308	5048	3918
IRL	D	222	1	174	636	26	156744	1	529	18	31696	25811	31721	11372	0
IRL	S	172	14	26	50	103	66526	10	30	6	1844	1844	1844	539	0
LVA	S	128	2	1	137	20	54458	1	117	6	15274	14585	15274	12769	12787
LVA	V	NA	NA	NA	NA	NA	NA	1	178	5	1458	1439	1466	185	0
NIR	S	156	7	1	139	57	150239	4	34	4	0	1861	1861	0	0
NIR	D	37	3	1	91	19	8277	2	27	3	1673	1737	1750	1600	1600
NLD	M	275	11	118	659	41	172448	7	368	18	18336	17435	18336	0	16843
POL	S	117	14	25	79	65	52826	14	77	14	7670	7670	7670	7657	7420
POL	M	101	15	40	131	23	17393	15	130	15	6697	6697	6697	6697	6551
PRT	M	246	21	1	3039	160	4944820	NA	NA	NA	NA	NA	NA	NA	NA
PRT	S	73	10	1	73	137	20931	3	11	2	490	0	491	363	0
PRT	V	NA	NA	NA	NA	NA	NA	9	350	23	16123	4816	20807	20124	17299
SCT	S	305	21	118	228	123	507931	19	178	4	0	7837	7837	0	0
SCT	M	228	18	293	844	44	234158	10	529	11	0	18053	18053	3322	3322
SWE	S	114	39	92	154	104	108179	38	152	7	7743	7981	34935	26121	155
SWE	M	194	41	5	356	13	45689	41	366	5	24449	19811	24682	15882	15877
SWE	D	12	8	11	12	20	4709	8	12	1	0	0	4881	4881	0
WLS	M	24	13	37	48	10	3278	5	17	4	0	449	449	158	0

Figure 10 Sampling by country

sampCtr	spp	days	sites	vessels	trips	nspp	lengthFreq	bioSites	bioTrips	bioSpp	wght	ages	lengths	sex	maturity	name	AphiaID
SCT	Gadus morhua	302	19	160	291	1	34063	18	268	1	0	5153	5153	0	0	Atlantic cod	126436
SCT	Melanogrammus aeglefinus	283	20	156	312	1	137670	20	278	1	0	4955	4955	0	0	Haddock	126437
SCT	Merlangius merlangus	310	20	158	312	1	92138	20	299	1	0	4423	4423	0	0	Whiting	126438
SCT	Pollachius virens	290	10	121	244	1	33191	10	233	1	0	4234	4234	0	0	Saithe(=Pollock)	126441
SCT	Lophius piscatorius	262	17	135	233	1	8323	5	86	1	0	2178	2178	0	0	Angler(=Monk)	126555
SCT	Scomber scombrus	231	15	98	184	1	12479	3	63	1	0	1813	1813	1813	1813	Atlantic mackerel	127023
SCT	Clupea harengus	219	16	101	165	1	11347	2	33	1	0	1509	1509	1509	1509	Atlantic herring	126417
SCT	Lepidorhombus whiffiagonis	174	12	90	149	1	13365	7	57	1	0	1230	1230	0	0	Megrim	127146
SCT	Lophius budegassa	69	5	58	76	1	522	5	57	1	0	271	271	0	0	Blackbellied angler	126554
SCT	Glyptocephalus cynoglossus	226	17	111	176	1	13008	2	6	1	0	78	78	0	0	Witch flounder	127136
SCT	Micromesistius poutassou	109	9	33	52	1	5676	1	1	1	0	46	46	0	0	Blue whiting(=Poutassou)	126439
SCT	Nephrops norvegicus	129	15	148	247	1	198857	NA	NA	NA	NA	NA	NA	NA	NA	Norway lobster	107254
SCT	Eutrigla gurnardus	265	17	100	169	1	28222	NA	NA	NA	NA	NA	NA	NA	NA	Grey gurnard	150637
SCT	Hippoglossoides platessoides	211	16	92	138	1	20923	NA	NA	NA	NA	NA	NA	NA	NA	Amer. plaice(=Long rough dab)	127137
SCT	Pleuronectes platessa	240	20	140	242	1	17483	NA	NA	NA	NA	NA	NA	NA	NA	European plaice	127143
SCT	Microstomus kitt	243	18	136	245	1	15772	NA	NA	NA	NA	NA	NA	NA	NA	Lemon sole	127140
SCT	Trisopterus esmarkii	167	17	64	94	1	15160	NA	NA	NA	NA	NA	NA	NA	NA	Norway pout	126444
SCT	Limanda limanda	215	19	99	150	1	13087	NA	NA	NA	NA	NA	NA	NA	NA	Common dab	127139
SCT	Trisopterus minutus	153	14	75	103	1	10441	NA	NA	NA	NA	NA	NA	NA	NA	Poor cod	126446
SCT	Merluccius merluccius	267	14	123	220	1	9609	NA	NA	NA	NA	NA	NA	NA	NA	European hake	126484
SCT	Amblyraja radiata	203	8	55	79	1	5191	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SCT	Argentina sphyraena	177	11	60	81	1	5145	NA	NA	NA	NA	NA	NA	NA	NA	Argentine	126716
SCT	Scyllorhinus canicula	206	16	91	135	1	4404	NA	NA	NA	NA	NA	NA	NA	NA	Small-spotted catshark	105814
SCT	Molva molva	224	14	106	181	1	3824	NA	NA	NA	NA	NA	NA	NA	NA	Ling	126461
SCT	Chelidonichthys cuculus	120	13	64	91	1	3754	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SCT	Sebastes viviparus	72	8	24	37	1	3580	NA	NA	NA	NA	NA	NA	NA	NA	Norway redfish	127255
SCT	Leucoraja naevus	168	18	89	136	1	2541	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SCT	Trachurus trachurus	136	14	62	88	1	2435	NA	NA	NA	NA	NA	NA	NA	NA	Atlantic horse mackerel	126822
SCT	Argentina silus	37	7	11	18	1	2318	NA	NA	NA	NA	NA	NA	NA	NA	Greater argentine	126715
SCT	Mullus surmuletus	37	10	26	30	1	1319	NA	NA	NA	NA	NA	NA	NA	NA	Surmullet	126986
SCT	Reinhardtius hippoglossoides	18	1	3	5	1	1226	NA	NA	NA	NA	NA	NA	NA	NA	Greenland halibut	127144
SCT	Raja montagui	84	9	44	63	1	1126	NA	NA	NA	NA	NA	NA	NA	NA	Spotted ray	105887
SCT	Phycis blennoides	93	8	34	58	1	1092	NA	NA	NA	NA	NA	NA	NA	NA	Greater forkbeard	126501
SCT	Anarhichas lupus	82	4	64	97	1	1086	NA	NA	NA	NA	NA	NA	NA	NA	Atlantic wolffish	126758
SCT	Helicolenus dactylopterus	44	6	12	16	1	982	NA	NA	NA	NA	NA	NA	NA	NA	Blackbelly rosefish	127251
SCT	Hippoglossus hippoglossus	79	5	73	101	1	947	NA	NA	NA	NA	NA	NA	NA	NA	Atlantic halibut	127138
SCT	Pollachius pollachius	58	5	46	64	1	799	NA	NA	NA	NA	NA	NA	NA	NA	Pollock	126440
SCT	Zeus faber	38	8	30	46	1	724	NA	NA	NA	NA	NA	NA	NA	NA	John dory	127427
SCT	Raja clavata	83	12	46	61	1	702	NA	NA	NA	NA	NA	NA	NA	NA	Thornback ray	105883
SCT	Enchelyopus cimbrius	105	13	37	59	1	585	NA	NA	NA	NA	NA	NA	NA	NA	Fourbeard rockling	126450
SCT	Scophthalmus maximus	78	7	69	97	1	541	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SCT	Sebastes mentella	8	1	2	3	1	490	NA	NA	NA	NA	NA	NA	NA	NA	Beaked redfish	127254
SCT	Callionymus lyra	107	13	49	70	1	481	NA	NA	NA	NA	NA	NA	NA	NA	Dragonet	126792
SCT	Brosme brosme	69	6	24	36	1	394	NA	NA	NA	NA	NA	NA	NA	NA	Tusk(=Cusk)	126447
SCT	Dipturus flossada	57	10	25	32	1	335	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SCT	Leucoraja fullonica	47	6	19	25	1	324	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SCT	Agonus cataphractus	33	10	20	30	1	323	NA	NA	NA	NA	NA	NA	NA	NA	Hooknose	127190
SCT	Rajidae	16	2	2	4	1	320	NA	NA	NA	NA	NA	NA	NA	NA	Rays and skates nei	105711
SCT	Platichthys flesus	39	11	18	24	1	264	NA	NA	NA	NA	NA	NA	NA	NA	European flounder	127141
SCT	Sebastes norvegicus	13	1	2	4	1	238	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SCT	Myoxocephalus scorpius	49	11	27	36	1	236	NA	NA	NA	NA	NA	NA	NA	NA	Shorthorn sculpin	127203

Figure 11 Sampling by species

3.3 References

- GFCM (2016): Meeting on GFCM fisheries data requirements. Scientific Advisory Committee on Fisheries.
<http://www.fao.org/gfcm/reports/technical-meetings/detail/en/c/462331/>
- Baumer, Ben, Mine Cetinkaya-Rundel, Andrew Bray, Linda Loi, and Nicholas J. Horton. 2014. "R Markdown: Integrating a Reproducible Analysis Tool into Introductory Statistics." *Technology Innovations in Statistics Education* 8 (1).
- Gentleman, Robert, and Duncan Temple Lang. 2004. "Statistical Analyses and Reproducible Research." *Bioconductor Project Working Papers*. Working Paper 2.
- R Core Team, R. 2017. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.<http://www.R-project.org/>.

4 CONTACT DETAILS OF EWG-17-17 PARTICIPANTS

¹ - Information on EWG participant's affiliations is displayed for information only. In any case, Members of the STECF, invited experts, and JRC experts shall act independently. In the context of the STECF work, the committee members and other experts do not represent the institutions/bodies they are affiliated to in their daily jobs. STECF members and experts also declare at each meeting of the STECF and of its Expert Working Groups any specific interest which might be considered prejudicial to their independence in relation to specific items on the agenda. These declarations are displayed on the public meeting's website if experts explicitly authorized the JRC to do so in accordance with EU legislation on the protection of personnel data. For more information: <http://stecf.jrc.ec.europa.eu/adm-declarations>

EWG 17-17 (Part 1) participants

Name	Address	Email
Angeliki ADAMIDOU	NAGREF-Fisheries Research Institute Nea Peramos 64007 KAVALA Greece	adamidou@inale.gr
Angeles ARMESTO	Instituto Español de Oceanografía Subida a Radio Faro- 50 36390 VIGO Spain	angeles.armesto@vi.ieo.es
Bram COUPERUS	Wageningen Marine Research PO Box 68 1970AB IJMUIDEN Netherlands	bram.couperus@wur.nl
Henrik DEGEL	DTU-Aqua, Øde Hastrup Vej 85, 4000 ROSKILDE, Denmark	hd@aqua.dtu.dk
Matthew ELLIOTT	Marine management Organisation Foss House- Kings Pool-1-2 Peasholme Green- YO17PX YORK	matt.elliott@marinemanagement.org.uk
Emmet JACKSON	Bord Iascaigh Mhara Crofton Road N/A DUBLIN	jackson@bim.ie
Edvardas KAZLAUSKAS	Agricultural Information and Rural Business Center Lithuania	edvardas.kazlauskas@vic.lt
Estanis MUGERZA	AZTI Fundation Txatxarramendi Irla z/g 48395 SUKARRIETA Spain	emugerza@azti.es
Alastair POUT	Marine Scotland Science 375 Victoria Road AB119DB ABERDEEN United Kingdom	a.pout@marlab.ac.uk
Jose RODRIGUEZ	Spanish Oceanographic Institute IEO San Martin s/n 39005 SANTANDER Spain	jose.rodriquez@st.ieo.es
Maria Begoña SANTOS VAZQUEZ	Instituto Español de Oceanografía Centro Oceanografico de Vigo- Subida a Radio Faro 50 36390 VIGO Spain	m.b.santos@ieo.es
Alan WALKER	CEFAS Pakefield Road NR330HT LOWESTOFT	alan.walker@cefas.co.uk

European Commission		
Name	Address	Email
John Casey	STECF secretariat	JRC-stecf-secretariat@ec.europa.eu

Venetia Kostopoulou	DG MARE	venetia.kostoupoulou@ec.europa.eu
---------------------	---------	--

JRC expert		
Name	Address	Email
John Casey	DG JRC.	john.casey@ec.europa.eu

EWG 17-17 (Part 2) participants

Name	Address	Email
Jenny Nord	Swedish Agency for Marine and Water Management, Gullbergs Strandgata 15, 40439 GOTHENBURG, Sweden	jenny.nord@havochvatten.se

Name	Address	Email
David Currie	Marine Institute Rinville, ORANMORE, CO. GALWAY, Ireland	david.currie@marine.ie
Dimitris Damalas	Hellenic Centre for Marine Research, Agios Kosmas- Hellinikon, 16777 ATHENS, Greece	shark@hcmr.gr
Henrik Degel	DTU-Aqua, Øde Hastrup Vej 85, 4000 ROSKILDE, Denmark	hd@aqua.dtu.dk
Laurent Dubroca	IFREMER	laurent.dubroca@gmail.com
Jon Elson	CEFAS	jon.elson@cefes.co.uk
Maria Teresa Facchin	COISPA, Via dei Trulli- 18/20, 70126 BARI, - TORRE A MARE, Italy	facchini@coispa.it
Jerome Guitton	Agrocampus oust, 65 rue de saint brieuc, 35700 RENNES, France	jerome.guitton@agrocampus-ouest.fr

European Commission		
Name	Address	Email
Hendrik Doerner	European Commission, Joint Research Centre, Unit D.02, Via e. Fermi 2749, 21027 Ispra VA, Italy.	hendrik.doerner@ec.europa.eu
Venetia Kostopoulou	DG MARE	venetia.kostoupoulou@ec.europa.eu

5 LIST OF ANNEXES

Electronic annexes are published on the DCF dedicated report section of the STECF website on:
<https://stecf.jrc.ec.europa.eu/reports/dcf-dcr>

List electronic annexes documents:

EWG-17-17 – Annex 1 – DRAFT Guidance for the submission of Annual Reports under Regulation (EU) 2017/1004, Commission Implementing Decision (EU) 2016/1701 and Commission Implementing Decision (EU) 2016/1251

EWG-17-17 – Annex 2 - EWG 17_17_AR TABLES WITH BASIC CHECKS FINAL.XLSM

EWG-17-17 – Annex 3 – CheckTemplate.XLSM

EWG-17-17 – Annex 4 Extended functionality of AR template.xlsm

6 LIST OF BACKGROUND DOCUMENTS

Background documents are published on the meeting's web site on:
<https://stecf.jrc.ec.europa.eu/ewg1717ii>

List of background documents:

EWG-17-17 – Doc 1 - Declarations of invited and JRC experts (see also section four of this report – List of participants)

Appendix 1 - Example of the CheckTemplate_Alltables.xlsx filtered for Table 1C

WP or AR	Table name(s)	Name of the field(s)	Link	Check type	Check Description	DataSource/Thresholds	DS exists? Incl. Link if available	When applied	Consequence	Warnings	Example
WP	1C	Key-fields in WP, Key-fields in AR	1A, 1B	WPC	The number of entries for each key MS/Species-Region-RFMO-Area/Stock in the submitted WP must be the same in the new AWP.	Submitted workplan	https://datacolle.ction.ec.europa.eu/wpl-main/wpl	Pre-submission and pre-screening	Mandatory comment	Missing/additional rows should be included in the new AWP with an explanation for the deviations in AR Comments (it could depend on some deviation in table 1A)	e.g. adding a new line and specify in the AR comment the following reason: "add new line to include a species due to changing in landings, not considered in the submitted WP"; e.g. "missing row for Length variable for given key-fields in the submitted plan."
WP	1C	MS	1A, 1B	IVC	Limited to reference list	ISO 3166-1 alpha-3 code	-	Pre-submission	Submission not allowed	Invalid value for Country code. Value must be an integer.	e.g. for Croatia the correct code is "HRV" and not "CRO"
WP	1C	MS participating in sampling	7C	IVC	Limited to reference list. More than one value can be inserted separated by a "-" without spaces	ISO 3166-1 alpha-3 code	-	Pre-submission	Submission not allowed	Invalid value for Country code. Value must be an integer.	e.g. for Croatia the correct code is "HRV" and not "CRO"; "ITA-SLO"
WP	1C	MS participating in sampling, MSs (7C)	7C	CC	In case the value in this field include more than one MS, check if in Table 7C is reported a row with a Bi-and multilateral agreements related to MS participating in sampling.	Agreements listed in table 7C	-	Pre-submission	Submission not allowed	No planned regional and international coordination or bi- and multilateral agreements found in Table 7C. Insert it in Table 7C	e.g. for Italy in table 7C is reported a row related to "Data collection for scientific surveys in the Adriatic area (GSA 17); Medion, Modis and Solomon (see Table 1C)" but it is no present in Table 1C
WP	1C	Sampling year	1B	IVC	Integer value between 1900 and current year	-	-	Pre-submission	Submission not allowed	Invalid value for sampling year. Value must in the reference range.	e.g. 1890
WP	1C	Species	1A, 1B	IVC	Limited to reference list	Scientific name in the "ASFIS List of Species for Fishery Statistics Purposes"	http://www.fao.org/fishery/collection/asfis/en	Pre-submission	Submission not allowed	Invalid value for Species. Value must be included in the reference list.	e.g. "Boop boops" and not "Boop boop"
WP	1C	Region	1A, 1B	IVC	Limited to reference list	COMMISSION IMPLEMENTING DECISION (EU) 2016/1251, Geographical stratification by region, "all region" is allowed.	http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:32016D1251&id=149999057084	Pre-submission	Submission not allowed	Invalid value for Region. Value must be included in the reference list.	e.g. "Mediterranean Sea" (level II) and not "Mediterranean Sea and Black Sea" (level III)
WP	1C	RFMO/RFO/IO	1A, 1B	IVC	Limited to reference list. If not applicable NONE is allowed	Regional fisheries management organisations (RFMOs)	https://ec.europa.eu/fisheries/fishinfo/data/infoc/en	Pre-submission	Submission not allowed	Invalid value for RFMO/RFO/IO. Value must be included in the reference list.	e.g. "CFOM" and not "CFOM"
WP	1C	Area / Stock	1A, 1B	IVC	Limited to reference list	Names of Sub-areas and Divisions of FAO fishing areas 27 and 37	https://ec.europa.eu/fisheries/fishinfo/data/infoc/en	Pre-submission	Submission not allowed	Invalid value for Area/Stock. Value must be included in the reference list.	e.g. "GSA 10" and not "GSA 100"
WP	1C	Variables	1B	IVC	Limited to the reference list	It can be one of the following values: "Length", "Weight", "Age", "Sexratio",	-	Pre-submission	Submission not allowed	Invalid value for variable name. Value must be included in the reference list.	e.g. "Maturity" and not "Sexual maturity".
WP	1C	Data sources	-	IVC	Limited to the reference list	It can be: surveys, "Commercial samples", "Market samples", "Discard samples"	-	Pre-submission	Submission not allowed	Invalid value for Data source. Value must be included in the reference list.	e.g. "Commercial samples" and not "Commercial" or "commercial samples"
WP	1C	Planned minimum no of individuals to be measured at the national level	-	IVC	Integer value between 0 and 1000000	-	-	Pre-submission	Submission not allowed	Invalid value for Planned minimum no of individuals to be measured at the national level. Value must be an integer.	e.g. 1500 and not 1500.3
WP	1C	Planned at the national level in WP, Planned at the national level in AWP	-	IVC	The numbers of individuals to be measured at national level in the submitted WP must be the same in the new AWP.	Submitted workplan	https://datacolle.ction.ec.europa.eu/wpl-main/wpl	Pre-submission	Mandatory comment	Differences in the numbers submitted in the WP and the numbers reported in the new AWP must be justified.	ex change in number of planned individuals due to the application of sampling optimization procedures are justified in the AR comment
WP	1C	Planned minimum no of individuals to be measured at the regional level	-	IVC	Integer value between -1 (used if not applicable) and 1000000.	-	-	Pre-submission	Submission not allowed	Invalid value for Planned minimum no of individuals to be measured at the regional level. Value must be an integer.	e.g. 1500 and not 1500.3
WP	1C	Planned at the regional level in WP, Planned at the regional level in AWP	-	WPC	The numbers of individuals to be measured at regional level in the submitted WP must be the same in the new AWP.	Submitted workplan	https://datacolle.ction.ec.europa.eu/wpl-main/wpl	Pre-submission and pre-screening	Mandatory comment	Differences in the numbers submitted in the WP and the numbers reported in the new AWP must be justified.	e.g. change in number of planned individuals due to the application of sampling optimization procedures are justified in the AR comment
AR	1C	Achieved number of individuals measured at the national level	-	IVC	Integer value between 0 and 1000000	-	-	Pre-submission	Submission not allowed	Invalid value for Achieved number of individuals measured at the national level. Value must be an integer.	e.g. 1000 and not 1000.5
AR	1C	Achieved number of individuals measured at the national level.	-	WPC	Check if an explanation is given if the number of achieved is equal to 0.	-	-	Pre-submission and pre-screening	Mandatory comment	Text in AR Comments not found for 0 planned.	e.g. 0 individuals achieved for a rare species.
AR	1C	Achieved number of individuals measured at the national level	-	TC	Check if a significant change (%) occurs in the achieved number of individuals measured compared to the past years. A comment must be present when the change in % is lower than 50% or greater than 150%.	Past Annual Report	-	Pre-submission and pre-screening	Mandatory comment	Text in AR Comments not found for significant change in values respect to the past years.	
AR	1C	% of achievement (100%/J)	-	IVC	Integer positive value. It is automatic filled in so no error should be occurred.	-	-	Pre-submission	Submission not allowed	Invalid value for % of achievement. Value must be an integer	e.g. 90% and not 90.5%
AR	1C	% of achievement (100%/J)	-	WPC	Integer between 50 and 150. A comment must be present when the percentage of % of achievement (100%/J) is out of the range.	-	-	Pre-submission and pre-screening	Mandatory comment	Insert a brief explanation for values lower than 50% or greater than 150% for the given species/area and add the general explanation for the deviation in "Text Box 1 C" (paragraph Deviations from the AWP).	e.g. missing comment for a number under the threshold.
AR	1C	Achieved number of samples	-	IVC	Positive integer or equal to 0	-	-	Pre-submission	Submission not allowed	Invalid value for Achieved number of samples. Value must be an integer	-
AR	1C	Achieved number of samples	-	WPC	Check if an explanation is given if the number of achieved is equal to 0.	-	-	Pre-submission and pre-screening	Mandatory comment	Insert in comment explanation for 0 value.	e.g. Rare species not found at all.
AR	1C	Achieved number of samples	-	TC	Check if a significant change (%) occurs in the achieved number of samples compared to the past years. A comment must be present when the change in % is lower than 50% or greater than 150%.	Past Annual Report	-	Pre-submission and pre-screening	Mandatory comment	Text in AR Comments not found for significant change in values respect to the past years.	
AR	1C	Sampling protocol	-	IVC	Mandatory free text	-	-	Pre-submission	Submission not allowed	A text to explain the sampling protocol adopted must be reported in AR comment column	e.g. e.g. 10 ind/haul, max 50 ind/box (comm. Cat), 5 ind/cm/quarter
WP	1C	Species, Area/Stock	1A, 1B	WPC	For all the row (identified by MS-Reference years-Species-Region-RFMO-Area) in Table 1A selected for the sampling (Y) and for all the years and the variables selected for collecting data in table 1B a row must be present in Table 1C.	COMMISSION IMPLEMENTING DECISION (EU) 2016/1251, Stocks in Union waters	-	Pre-submission and pre-screening	Submission not allowed	Add missing rows related to MS-Reference years-Species-Region-RFMO-Area and variable XX in Table 1B selected for the sampling (in Table 1A)	e.g. in Table 1A A. foliaceae is selected for sampling in and in table 1B the Length variable is selected for collecting data, but in Table 1C the entry related to Length variable is missing
WP	1C	Region, RFMO/RFO/IO	1A, 1B	IVC	The value of RFMO/RFO/IO is limited to the reference list for the given Region	Regional fisheries management organisations (RFMOs)	https://ec.europa.eu/fisheries/fishinfo/data/infoc/en	Pre-submission	Submission not allowed	Invalid RFMO for the given region.	eg. ICES cannot be in Mediterranean Sea
WP	1C	Area / Stock, Region	1A, 1B	IVC	The value of Area/Stock is limited to the reference list for the given Region.	COMMISSION IMPLEMENTING DECISION (EU) 2016/1251, Stocks in Union waters	http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:32016D1251&id=149999057084&...	Pre-submission	Mandatory comment	Specify a reason in the AR comment column why a species not included in the reference list has been reported.	e.g. <i>Alpha minuta</i> is not included in the Mediterranean Sea but it is relevant in the Italian CRA 9, where a Management Plan is in place.

Appendix 2 - Using R for validation

Austria.

variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	TRUE	1	
Species	FALSE	1	
Selected for sampling	FALSE	0	NA

Belgium.

variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	TRUE	1	
Species	FALSE	1	
Selected for sampling	FALSE	0	NA

Bulgaria.

variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	11	2013-2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA

Croatia.

variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	11	2013-2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA

Cyprus.

variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	165	2013-2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA

Denmark.

variable	test	nbwrong	idwrong
MS	TRUE	0	

Reference year	FALSE	165	2013-2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA
Estonia.			
variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	165	2013-2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA
Finland.			
variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	187	2013-2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA
France.			
variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	624	2013-2015,2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA
Germany.			
variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	900	2013-2015,2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA
Greece.			
variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	900	2013-2015,2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA
Hungary.			
variable	test	nbwrong	idwrong
MS	TRUE	0	

Reference year	FALSE	900	2013-2015,2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA
Italy.			
variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	900	2013-2015,2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA
Latvia.			
variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	979	2013-2015,2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA
Lithuania.xls			
variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	979	2013-2015,2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA
Malta.			
variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	1102	2013-2015,2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA,N*
Netherlands.			
variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	1187	2013-2015,2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA,N*
Poland.			
variable	test	nbwrong	idwrong
MS	TRUE	0	

Reference year	FALSE	1229	2013-2015,2015
Species	FALSE	1	
Selected for sampling	FALSE	0	NA,N*
Portugal.			
variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	1463	2013-2015,2015,2012-2014,2011-2013
Species	FALSE	78	
Selected for sampling	FALSE	0	NA,N*
Romania.xls			
variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	1463	2013-2015,2015,2012-2014,2011-2013
Species	FALSE	78	
Selected for sampling	FALSE	0	NA,N*
Slovenia.			
variable	test	nbwrong	idwrong
MS	TRUE	0	
Reference year	FALSE	1547	2013-2015,2015,2012-2014,2011-2013
Species	FALSE	78	
Selected for sampling	FALSE	0	NA,N*
Spain.			
variable	test	nbwrong	idwrong
MS	FALSE	6	DATA SOURCE ICES areas: EU landings. The data for this table came from FIDES (Fishery Data Exchange System) catch reporting,DATA SOURCE NAFO areas: landings of the EU. The data for this table are from the STATLANT21A database,DATA SOURCE MED areas: EU landings according to RCM Med,DATA SOURCE CECAF areas: Data source of landings from Spain and from other EU fleets : RCM-LDF and Joint Scientific Committees of Sustainable Fisheries Partnership Agreements EU-coastal States,DATA SOURCE ICCAT, IOTC, IATTC areas: National data submitted to tuna RFMOs, published on the corresponding websites,DATA SOURCE WCPFC areas:the thresholds was calculated (Column Share%) using the catch data from the WCPFC web that are listed as catches from Portugal, Spanish catches come from data base of Control regulation of Spain. The percentage captured by Spain taking into account that 100% would be the catches of Spain + Portugal.
Reference year	FALSE	1898	2013-2015,2015,2012-2014,2011-2013,NA

Species	FALSE	91	
Selected for sampling	FALSE	0	NA,N*
Sweden.			
variable	test	nbwrong	idwrong
MS	FALSE	6	DATA SOURCE ICES areas: EU landings. The data for this table came from FIDES (Fishery Data Exchange System) catch reporting,DATA SOURCE NAFO areas: landings of the EU. The data for this table are from the STATLANT21A database,DATA SOURCE MED areas: EU landings according to RCM Med,DATA SOURCE CECAF areas: Data source of landings from Spain and from other EU fleets : RCM-LDF and Joint Scientific Committees of Sustainable Fisheries Partnership Agreements EU-coastal States,DATA SOURCE ICCAT, IOTC, IATTC areas: National data submitted to tuna RFMOs, published on the corresponding websites,DATA SOURCE WCPFC areas:the thresholds was calculated (Column Share%) using the catch data from the WCPFC web that are listed as catches from Portugal, Spanish catches come from data base of Control regulation of Spain. The percentage captured by Spain taking into account that 100% would be the catches of Spain + Portugal.
Reference year	FALSE	2163	2013-2015,2015,2012-2014,2011-2013,NA
Species	FALSE	91	
Selected for sampling	FALSE	0	NA,N*
UnitedKingdom.			
variable	test	nbwrong	idwrong
MS	FALSE	7	DATA SOURCE ICES areas: EU landings. The data for this table came from FIDES (Fishery Data Exchange System) catch reporting,DATA SOURCE NAFO areas: landings of the EU. The data for this table are from the STATLANT21A database,DATA SOURCE MED areas: EU landings according to RCM Med,DATA SOURCE CECAF areas: Data source of landings from Spain and from other EU fleets : RCM-LDF and Joint Scientific Committees of Sustainable Fisheries Partnership Agreements EU-coastal States,DATA SOURCE ICCAT, IOTC, IATTC areas: National data submitted to tuna RFMOs, published on the corresponding websites,DATA SOURCE WCPFC areas:the thresholds was calculated (Column Share%) using the catch data from the WCPFC web that are listed as catches from Portugal, Spanish catches come from data base of Control regulation of Spain. The percentage captured by Spain taking into account that 100% would be the catches of Spain + Portugal.,*EC Noted in evaluation of 2013 programme (ARES 2016 4766107) The UK argues that all fish were landed abroad into South Africa and Mauritania. Considering the fact that the quantities landed by

the UK vessels were less than 3% of the total EU landings in the previous years, the UK was granted derogation in its National Programme 2011-2013 for all the species from long-distance fisheries. This derogation will have to be reviewed in the light of the most recent specifications under the IOTC agreement - Review has yet to occur but just one UK vessel now operates in this area.

Reference year	FALSE	2465	2013-2015,2015,2012-2014,2011-2013,NA,2013-2014
Species	FALSE	92	
Selected for sampling	FALSE	0	NA,N*

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: <http://europa.eu/contact>

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696, or
- by electronic mail via: <http://europa.eu/contact>

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: <http://europa.eu>

EU publications

You can download or order free and priced EU publications from EU Bookshop at: <http://bookshop.europa.eu>. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see <http://europa.eu/contact>).

STECF

The Scientific, Technical and Economic Committee for Fisheries (STECF) has been established by the European Commission. The STECF is being consulted at regular intervals on matters pertaining to the conservation and management of living aquatic resources, including biological, economic, environmental, social and technical considerations.

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent, evidence throughout the whole policy cycle.



EU Science Hub
ec.europa.eu/jrc



@EU_ScienceHub



EU Science Hub - Joint Research Centre



Joint Research Centre



EU Science Hub

