



联合国
粮食及
农业组织

Food and Agriculture
Organization of the
United Nations

Organisation des Nations
Unies pour l'alimentation
et l'agriculture

Продовольственная и
сельскохозяйственная организация
Объединенных Наций

Organización de las
Naciones Unidas para la
Alimentación y la Agricultura

منظمة
الأغذية والزراعة
للأمم المتحدة

COORDINATING WORKING PARTY ON FISHERY STATISTICS

Intersessional Meetings of Aquaculture and Fisheries Subject Groups

Twenty-Ninth meeting of the Fisheries Subject (FS)

20-23 June 2022

(Online, Zoom platform + FAO HQ Queen Juliana Room – B323)

Progress report of the CWP ad-hoc Task Group on best practices for streamlining statistical data workflow and confidentiality issues

Author: FAO Secretariat

Document Summary

This document provides the progress report of the work carried out by the ad-hoc Task Group on best practices for streamlining statistical data workflow and confidentiality issues. It includes the rationale for the creation of the Group and some background information. It also illustrates the activities carried out by the Task Group through the creation of a questionnaire circulated to all CWP parties with the aim to collect the different practices currently in place in each organization, covering aspects such as the collection, processing and dissemination of the data as well as confidentiality issues. The key results obtained from the questionnaires are provided in the document. In addition it also illustrates the main outcomes of bilateral meetings held between FAO and the Task Group Members.

CWP members are kindly invited to provide feedback on the results shown and indicate additional ways forward in order to develop the envisaged best practices. In addition, CWP members that have not yet filled the questionnaires are encouraged to do so at their earliest convenience in order to obtain a more comprehensive view of the practices currently in place in the different organizations.

1. Rationale

In response to SDG goals including SDG 14: Life below water, FAO is developing initiatives which promote good governance, participatory decision-making processes and best practices in fisheries. These initiatives include improving the quality and utility of fishery and aquaculture data such as streamlining of arrangements for improving consistency, reducing discrepancies among published global and regional datasets and reducing the reporting burden for countries (e.g. FAO, 2018). This work builds on previous arrangements such as the use of the STATLANT¹ standardized questionnaires (since the 1970s) and formal agreements between FAO and other CWP parties such as EUROSTAT (since the 1980s), tuna RFMOs (since the late 1990s), the Southeast Asian Fisheries Development Center (SEAFDEC) (since 2007) and the Organisation for Economic Co-operation and

¹ STATLANT system of questionnaires is a long-standing standardized statistical inquiry developed by the CWP for the submission of national catch and effort data to international organizations by national statistical offices (<http://www.fao.org/cwp-on-fishery-statistics/handbook/introduction/data-collection-systems/en/>).

Development (OECD) (since 2020 for employment data). FAO is seeking to establish similar agreements with other institutions and other RFBs such as the Regional Fisheries Committee for the Gulf of Guinea (COREP), Fishery Committee for the West Central Gulf of Guinea (FCWC), Regional Commission for Fisheries (RECOFI) and the Western Central Atlantic Fishery Commission (WECAFC). While the work has been mainly focused on capture fishery data, many of the issues and solutions discussed also apply to aquaculture data (e.g. data quality, data processing, capacity building, policy needs).

CWP-26 established five ad-hoc task groups (TGs) to develop its work during the 2019-2022 intersessional period (FAO, 2019), including a TG on best practices for streamlining statistical data workflow with a focus on confidentiality issues (TG-workflow). The membership of the TGs is open to all CWP parties.

2. Background of the ad-hoc Task Group

2.1 Streamlining statistical data workflow

Upon presentation of the above FAO initiatives at the 26th session of the CWP (FAO, 2019), it was noted that these initiatives focus on best practices and guidelines which aim to minimize discrepancies and replications in statistical data and reduce the overall data reporting burden on member countries. Such initiatives may also lead to new or improved data sharing agreements among CWP parties or between member countries and CWP parties. CWP identified various actions which, if implemented, may assist in further streamlining the reporting mechanisms and workflow for capture fishery and aquaculture statistics. These actions include:

- Aligning annual data calls and reporting calendars in order to facilitate data sharing
- Implementing consistent statistical concepts, standards and definitions
- Developing mainstream data provisions which can serve multiple reporting requirements of member countries
- Improving accessibility of data through the use of harmonized and accessible formats
- Identifying and resolving data gaps and discrepancies through collaborative analysis
- Improving transparency through systematic processing and documentation of data sources.

CWP also agreed that the reporting of national statistics may be further streamlined by:

- Improving collaboration and exchange of data among CWP parties through the use of CWP and other international standard classifications as promoted through the CWP standard for reference harmonization
- Developing comprehensive and accessible data validation, metadata and documentation
- Providing data users with timely and transparent information on data validation rules, metadata and associated documentation
- Encouraging parties to identify and share statistical data and reporting issues such as interpretation and application of standard concepts
- Developing and using comprehensive and accessible Data Collection Reference Frameworks
- Using joint questionnaires, where possible, to reduce the number of different focal points and thus reduce the reporting burden for the national statistical offices

-
- Developing and implementing FAIR data principles for making data findable, accessible, interoperable and reusable.

CWP recognized that the alignment of annual data calls and calendars may potentially result in a synchronous submission of large amounts of data that may result in pressure points within organizations and thus may not necessarily improve specific workflows.

2.2 Data confidentiality

CWP also discussed data confidentiality issues which may hamper the dissemination and exchange of statistical data among CWP parties, e.g. between FAO and EUROSTAT regarding aquaculture statistics where some data (e.g. production by species and FAO area) from EU member countries cannot be published as provided because the aggregated data represent less than three enterprises. In capture fisheries, a specific example is the reporting of gridded (5 x 5 degree) catch and effort statistics in the FAO Atlas of Tuna and Billfish Catches. In some grids, aggregated data (e.g. by flag, gear and month) may represent less than three fishing vessels and may not be able to be released by an RFMO due to its data confidentiality rules.

In general, statistical data confidentiality requirements seek to protect individual identities including, *inter alia*, aquaculture producers. Statistical data confidentiality requirements differ from those of the General Data Protection Regulation (GDPR) and CWP parties use various rules and procedures to deal with statistical data confidentiality (e.g. Annex 1 in Appendix 4, CWP, 2019). The responsibility for decisions on the dissemination of statistical data are generally well defined in RFMOs, and CWP highlighted the following points for consideration in further developing statistical data confidentiality requirements. These requirements should:

- allow metadata from masked data to be published (i.e. flag the existence of masked data in order to facilitate the possible use of such data under specific conditions)
- consider temporal limitations on data confidentiality
- identify aggregation dimensions and levels, whereby individual entities can be no longer identified
- provide clear rules of use for each dataset through published data access and sharing policies
- consider exceptions to data confidentiality rules which may be needed if, for example, the dissemination of data from endangered species is limited by requirements for confidentiality.

3. Objectives of the ad-hoc Task Group

The objectives of TG- workflow were identified as to:

- Review the statistical data workflow of CWP parties and observers and other relevant cases involved in capture fisheries and aquaculture, and identify general issues and constraints which may lead to discrepancies and replications in statistical data and undue data reporting burden on member countries
- Review statistical data confidentiality requirements and rules of CWP parties and other relevant organizations
- Develop best practice guidelines for streamlining the reporting mechanisms and workflow for capture fishery and aquaculture statistics and for reducing the overall data reporting burden on member countries

- Develop best practice guidelines for the implementation of statistical data confidentiality requirements which protect data confidentiality and promote comprehensive, transparent and timely dissemination and exchange of statistical data.

The scope of this work concerns domains covered by the CWP Handbook. Where cases of specific interest to CWP Members are identified, this scope may also include the review of the data workflow of CWP observers and other relevant cases which involve countries' multiple reporting requirements to different organizations. In addition, the review of the statistical data workflow may be split into two streams of work to: (1) Review established streamlining processes (e.g. as described in FAO, 2018, p.94) and identify strengths, weaknesses and opportunities arising from these processes, (2) Review other relevant workflows and identify new requirements for streamlining processes including the use of the CWP Standard for Reference Harmonization.

4. Activities of the ad-hoc Task Group

During the CWP-26 discussion, it was agreed that the work of TG-workflow may be facilitated through the use of questionnaires and specific work cases. For this reason, a short questionnaire (available in Annex 1) was designed to collect the different practices currently at place in each organization, as part of the CWP task group on best practices for streamlining statistical data workflow, with a focus on confidentiality issues. The questionnaire was shared with CWP members on the 8th September 2021.

Results of this exercise aimed at:

- reviewing the statistical data workflow of CWP parties and identifying general issues and constraints which may lead to discrepancies and replications in statistical data and undue data reporting burden on member countries
- reviewing statistical data confidentiality requirements and rules of CWP parties and other relevant organizations

4.1 Results of the Questionnaire on Approach to Data Workflow

Ten respondents originally submitted the completed questionnaire, with results presented at the CWP Intersessional Meetings in November 2021, and six more respondents completed the questionnaire after the CWP Intersessional Meeting. For the two main sections, the main updated results are the following:

4.1.1 Approach to data workflow

- Almost all respondents (90 percent) are responsible for data collection and collation.
- In terms of frequency, the majority of respondents receive data yearly or biannually (although depending on the data it can be ad hoc or more frequent submissions), and 70 percent of respondents make formal data calls (mostly thorough emails), and the majority provide a questionnaire to the data providers.
- Consumption data is not collected by any of the respondents, while aquaculture data, socio-economic data and trade data are only collected and collated by 30 percent. The most widely collected data is that of nominal catch (70 percent).
- CWP standard and definitions are widely endorsed, while just over half of the organizations use other international standards. The majority of organizations use ASFIS three-alpha codes to categorize species.

-
- All respondents but one indicated that data is shared outside their organization, mostly through online portals accessible to the public, and data goes through a process of validation and standardization before being published
 - Most organizations (75 percent) publish either data validation practices or metadata, but very few respondents publish both.
 - Just under half of respondents do not use (or not explicitly use) Data Collection Reference Frameworks.
 - Most organizations collaborate and/or exchange data with other organizations, but a smaller amount (53 percent) uses a joint questionnaire and harmonizes data.

In addition, it is important to highlight how organizations are not aligned on other aspects of data workflow. For example, although of the majority of respondents stated that the deadline for data submission is aligned with internal workflows and deadlines, the deadlines for the data submissions for each organization are spread throughout the year. Similarly, when using estimations to fill data gaps, respondents are almost equally split between positive and negative answers, and even within the positive answers, there are differences dependent on data type.

4.1.2 Approach to confidentiality

- All organizations receive confidential data from data providers, with the original data (55 percent) or with masked original data (35 percent).
- The large majority of respondents do not share data as confidential even in the case the providers did not flag it as confidential.
- The data considered most confidential is licensing and compliance data, observer data, and catch and effort data. In particular, vessel-level data was given as a common example of data variables considered as being confidential. Aquaculture data, socio-economic data and trade data, on the other hand, were not significantly flagged as confidential.
- When disseminating data, half of the respondents indicated confidential data is masked at the most detailed level and at all the subsequent aggregates in which it is included, while the other half mask the confidential data but disseminate the aggregates.
- Overall, no temporal limitations are applied to confidentiality rules.

4.2 Bilateral meeting results

Following the Intersessional Meeting (November 2021) agreed next steps, bilateral meetings between FAO and the Task Group Members were carried out to further analyze and discuss the practices highlighted in the questionnaire. Detailed results are shown in Annex 2. For the two main sections, the main outcomes of these bilateral meetings are explained below.

4.2.1 Approach to data workflow

Overall, there is consensus that improved data sharing, reduced burden on data providers and harmonized outputs are important objectives, and there is willingness to work towards their even partial achievement. However, a number of obstacles were highlighted both in the questionnaire as well as in the bilateral meetings, as below:

- Legal ties to Commissions and subsidiary bodies/member countries.
- Differences in the detail requested by the organizations to data providers.
- Constraints in harmonized deadlines due to national calendars.

There was general agreement that the CWP concepts and standards should be adhered to as much as possible, and that CWP Members should communicate and coordinate concerning new data collection programmes/plans, or for overlapping datasets. CWP Members should actively strive for coherence and comparability between data collection standards.

4.2.2 Approach to confidentiality

The discussions carried out highlighted how confidentiality regulations are tightly linked to the data providers, meaning CWP Members have little scope for enforcing new standards. However, overall there is interest in exploring ways to disclose confidential data (examples include aggregating all confidential data in one group, or disclosing data at a higher level of granularity), and to develop common approaches between CWP Members.

5 Next steps

The results of the questionnaire and the bilateral meetings will provide a basis for the TG's discussion on the development of best practice guidelines, in particular for confidentiality standards, as there seems to be limited scope for workflow guidelines. In this respect, and as emerged from the conversations with Task Group Members, further discussion is needed to identify potential ways to facilitate and encourage the dissemination of confidential data. In addition, only three CWP have not filled the questionnaire on data workflow and confidentiality and they are kindly invited to fill it in order to have a comprehensive overview of all Members.

Questionnaire on Approach to data workflow

In the context of the CWP Task Group on "Best practices for streamlining statistical data workflow with a focus on confidentiality issues", we would like to get information on your organization's approach to dealing with data flows.

***Required**

1. What is the role of your organization in terms of statistical information? *

Mark only one oval.

- Data collection and collation
- Data harmonization
- Data analysis Data
- dissemination
- Other:

2. Please list the agencies and/or countries which your organization receive data from? *

3. How often is data received/collected? *

Mark only one oval.

- Weekly
 - Monthly
 - Quarterly
 - Biannually
 - Yearly
 - Other:
-

4. Does your organization make formal calls for data? *

Mark only one oval.

Yes

No

5. If yes, how are these data calls carried out?

6. What are your organization's deadlines, if any, for the submission of the data? *

7. Are these deadlines aligned with your organization's workflow and internal deadlines (e.g. working groups, Commission meetings)? If so, how? *

8. How is data submitted? * *

Mark only one oval.

My organization provides the questionnaires to the data providers.

The data providers have access to an online database.

Data providers share their own questionnaires.

Other:

9. Which data is collected/collated? *

Tick all that apply.

- Nominal catch
- Landed catch
- Aquaculture
- Socio-economic (fleet, employment)
- Trade
- Consumption
- Catch and effort
- Licensing and compliance
- Biological
- Observer data

Other: _____

10. Are CWP endorsed concepts, standards and definitions applied to the collated data (ex. live weight vs landed weight) ? * *Mark only one oval.*

- Yes
- No

11. If yes, how and when are these concepts, standards and definitions applied?

12. Are other international classifications used in the collection and dissemination of data? *

Mark only one oval.

- Yes
- No

13. If yes, how and when are these concepts, standards and definitions applied?

14. Are species categorized using ASFIS three-alpha codes? *

Mark only one oval.

Yes

No

15. Is data shared outside your organization, and if so how? *

16. Does data go through a process of validation and standardization before being published?

Mark only one oval.

Yes

No, it is published as it is received.

Other:

17. Are data gaps filled through estimations?

Mark only one oval.

Yes

No

Other:

18. Do you publish comprehensive data validation practices and metadata? If so, please specify. *

1. 19. Do you use Data collection reference frameworks? If so, please specify. *

-
2. 20. Does your organization collaborate and/or exchange data with other organizations? *

Mark only one oval.

- Yes
- No

3. 21. If yes, is a joint questionnaire used and data harmonized?

Mark only one oval.

- Yes
- No
- A joint questionnaire is used but data is not harmonized.

Questionnaire on Approach to Confidentiality

In the context of the CWP Task Group on "Best practices for streamlining statistical data workflow with a focus on confidentiality issues", we would like to get information on your organization's approach to dealing with confidential data.

1. Does your organization receive confidential data from its data providers? *

Tick all that apply.

- Yes, the data is confidential and the original data is submitted.
- Yes, confidential data is flagged and original data is masked.
- No, confidential data is not submitted.

Other: _____

2. Do you disseminate data as confidential while it was not considered as confidential by your data providers? * *Mark only one oval.*

- Yes
- No
- Other:

3. Which data does your organization consider as being confidential?

Tick all that apply.

- Capture data (quantities or price)
- Aquaculture data (either quantities or price)
- Socio-economic data (fleet and employment)
- Trade data
- Consumption data
- Catch and effort data
- Licensing and compliance data
- Biological data
- Observer data
- Not relevant

Other: _____

4. If you selected any of the above categories, please provide examples of data variables which are considered as being confidential.

5. Please describe the confidentiality rules that you apply when validating, processing and storing data. If possible, please include a link, or the text, of your organization's confidentiality policy. *

6. When disseminating data, how does your organization handle confidential data? *

Tick all that apply.

- You replace confidential data by estimates.
- You mask the confidential data at the most detailed level and at all the subsequent aggregates in which it is included
- You mask the confidential data at the most detailed level but disseminate the aggregates.

Other: _____

7. If your organization follows confidentiality rules, are temporal limitations applicable (e.g. data is confidential only for a set amount of time)?

Contact person

4. Name *

5. Organisation *

6. Email *

CWP - Task group on best practices for streamlining statistical data workflow and confidentiality issues

Main outcomes of bilateral meetings

WORKFLOW

- **HARMONIZATION OF DATA CALLS AND SUBMISSIONS**
 - To facilitate workflow, reduce overburdening for data providers, and avoid inconsistencies, RFMOs and other agencies could act as “intermediary” between data providers and FAO/other CWP parties who require access to the same data. This could apply to specific units such as species of interest (for example in the case of tuna RFMOs).
 - When planning deadlines, consider the availability of national data, which means deadlines can be constrained by the data providers.
 - Learn from current examples, where joint questionnaires have been deployed, a good degree of success has been achieved (e.g. FAO-OECD employment questionnaires, which resulted in a higher response rate).
 - Overall, it can be a difficult objective as many organizations are tied to Commissions and subsidiary bodies.
- **LOW LEVEL OF SUBMISSION FOR CERTAIN DATA**
 - Analyse the reasons behind low levels of submissions (for example low success of country level programmes and initiatives, time consuming/non user friendly questionnaire, overburdened data providers, politically sensitive data, format/level of aggregation not applicable).
- **STANDARDIZATION OF CWP STANDARDS AND CONCEPTS**
 - Overcome the loss of granularity of the original data that stems from mapping local concepts (for example SSF gear and vessel types).
 - Consider simplification of CWP definitions, for better understanding by the data providers.
 - Where new frameworks are being developed, include as many CWP concepts and standards as possible. Organizations should reach out to CWP for support and any further discussions if needed.
- **INTRODUCE NEW CWP CONCEPTS AND STANDARDS**
 - Standard list of effort units by gear for small-scale fisheries (proposal already submitted to CWP).
 - Inclusion of standards and definitions for recreational fishing, discards and slipping, non-commercial catches.
 - CWP Glossaries, terms and concepts should keep into account new CWP partner specificities.
 - Employment concepts (e.g. distinctions between deep-sea and marine coastal fishers as ILO ISCO 08).
- **VALIDATION AND HARMONIZATION OF ESTIMATES**
 - Sharing of estimate methodologies between CWP parties and investigate potential of adopting a common approach.
 - Attempt to identify which approaches each partner uses on gap completion estimates.
- **JOINT QUESTIONNAIRES FOR SHARED DATA**
 - Although a useful tool, consider the granularity of the data received, and the need to reduce the resolution to meet the requirements of some CWP parties, therefore potentially losing detail (for example if data is submitted by gear, but only required by some organizations by species).

CONFIDENTIALITY

- **COMMON CONFIDENTIALITY RULES**

- Most CWP members receive confidential data from their data providers => at least for the TG members, there is little scope for enforcing confidentiality rules to the data providers.
- Some data is considered as more sensitive, such as price data, vessel locations, etc.
- Actions and changes are dependant on the individual confidentiality policies.
- COMMON DATA ESTIMATION APPROACH
 - TG members were more interested in finding ways allowing to disclose confidential data than in developing common approaches to estimate data gaps related to confidential data points.
- DISSEMINATION OF AGGREGATES WITH CONFIDENTIAL DATA
 - Sharing of approaches between CWP parties and investigate potential of adopting a common set of approaches. Approaches include:
 - Disclosing data at a higher level of granularity
 - Aggregating all confidential data into one group
 - Exclude confidential data when negligible
 - Adding statistical noise in the data to bypass the confidentiality issue
 - Relax confidentiality rules after a set amount of time (for example five years)
 - The European Fishery Statistics framework is under review and should be released by 2024. ESTAT welcomes collaboration with FAO to develop common approaches to best way to deal with confidential data.
 - If in agreement with the organizations policy, an agreement to disseminate aggregates with confidential data is feasible.