



---

<b>Document title</b>	Outcome of Data reporting workshop (Data reporting WS 1-2019)
<b>Submitted by</b>	Secretariat

---

The online meeting (Skype for business) was held on 26 November at 10-13 CET.

## Agenda Item 1. Opening of the meeting and adoption of the agenda.

1. The meeting was opened, and the agenda was adopted without changes.
2. The meeting was attended by representatives from HELCOM Secretariat, ICES Data Center, Estonia, Germany, Latvia, Poland and Sweden.

## Agenda Item 2. General introduction.

3. ICES presented a general introduction ([Document 2-2](#)) of marine biological community data formats ([Document 2-1](#)), procedures of reporting data to ICES and available applications for extracting data
4. ICES Data Centre suggested to the meeting that the standard data downloads should be based only on the data as submitted to ICES, while the data subsets with omitted records and data calculations should be stored in open repositories, like Github. The proposal was supported by Sweden.
5. The meeting discussed the data summaries in the accessions pages:  
<http://ices.dk/marine-data/tools/Pages/Submission%20status.aspx> and  
<http://data.ices.dk/accessions/AllCommissionSummary.aspx?commission=HELCOM>

The meeting noted the following on the accessions page:

- The page shows data submissions that have been labeled as HELCOM
  - The submission status might not contain the level of information to see whether all planned stations/measurements were completed. For example, label "1C" does not implicate whether the submission is fully reported, and how many submissions are expected. ICES does not have an overview of the planned sampling programs, so checks like these are presently in hands of the national data experts and the respective expert groups.
6. The meeting took note that 'Parameters in the database' summary includes information on number of stations for the combination of (data type, country, RLABO, param, year) which complements the overview on the Accessions page.
  7. The Meeting recalled the procedure that Contracting Parties should label all HELCOM monitoring data with the monitoring programme COMBINE by using attribute MPROG=COMB, to enable agreed extra quality checks applied for HELCOM data and to have the submission to be included in the HELCOM data submissions page.
  8. The Meeting discussed the available formats: ERF 3.2 and Simplified format and what are the differences in these and what implications those might have.
  9. The meeting took note of the information by ICES that the advanced data submitters are recommended to continue using the ERF 3.2 for data submissions and resubmissions, while the simplified format should be used only in cases of the new data providers, or in cases where data is stored in excel files.
  10. The meeting discussed the possible use of other global formats, such as extended Darwin Core format used in EMODnet biology. The meeting took note that in order to be able to decide on introducing new accepted data submission formats, it would have to be assessed against indicator data requirements.
  11. The meeting discussed the issue of re-reporting data and that data harvesting could be a solution to reduce manual workload related to re-reporting of data. The meeting noted that the development of a functional harvesting system requires resources both at national level and ICES side. The meeting was also informed that some Contracting Parties (Finland, Latvia and Sweden), that responded to the

enquiry to apply funding from EU-CEF-TELECOM 2019-2 Public open data call submitted by Secretariat to State & Conservation in September 2019, have applied for funding together with Secretariat and ICES to proceed with this task.

12. The meeting took note of the information that in some cases submission of historical data might not be possible due to the existing quality checks on supporting parameters that are not available or have not been collected at all for the historical data thus omitting these data to be included in the database.
13. The meeting welcomed the information by ICES that when submitting historical data, it is possible to release some quality control procedure and recommended data providers to contact ICES directly (accessions@ices.dk) in cases of reporting historical data with missing supporting parameters.

### Agenda Item 3. Generic issues in reporting biodiversity data to ICES.

14. The meeting discussed the use of reporting taxonomical data and agreed on that the use of WoRMS AphiaID is recommended to be used in reporting.
15. It was clarified by ICES that currently AphiaID is required, but the validity of AphiaID is not checked, meaning that non-valid AphiaIDs can be reported. For legacy data, those can refer to old species lists and there are many species in the current database that do not refer to AphiaID stemming from legacy data. PEG\_BVOL list is mapped to the valid Aphia by the PEG group. Data submissions with RLIST=PEG\_BVOL are validated against species names and validity flags in the PEG list, not AphiaIDs in WoRMS.
16. The meeting welcomed the offer from ICES to compile a table of species and AphiaIDs which is required for carrying out a mapping exercise between existing species IDs and AphiaID. The table is included in the Meeting Documents library: [https://portal.helcom.fi/meetings/Data%20reporting%20OWS%201-2019-689/MeetingDocuments/3-1%20Species\\_DOME\\_2\\_AphiaID.xlsx](https://portal.helcom.fi/meetings/Data%20reporting%20OWS%201-2019-689/MeetingDocuments/3-1%20Species_DOME_2_AphiaID.xlsx)
17. The Meeting concluded that it would be useful to carry out mapping exercise of reported taxon data to AphiaID to make historical data more useful for assessment and time series analysis purposes. The meeting also noted that this process would require taxonomic and substance expertise and would have to happen via the relevant expert groups and national data providers.
18. The meeting noted that in cases, where AphiaID does not exist for species to be reported, it should be added to WoRMS using the established procedure for adding new species.
19. The Meeting pointed out that when reporting data based on PEG list, it is needed to refer to the correct version of the PEG list.
20. The Meeting took note that the PEG group annually maps PEG phytoplankton species to AphiaID and submits the information annually to ICES.
21. The meeting noted that most of the unsuitability of historical data is related to data stemming from 1979-2000 and is due to different reporting formats used then (HELCOM1), which contained limited set of supporting parameters.
22. The meeting took note of the comment on the complexity of ERF 3.2, and that it would be helpful to have an example of a prefilled reporting submission available for data providers.

### Agenda Item 4. Specific issues by biodiversity data type

#### Zoobenthos:

23. The Meeting took note of the issues observed with zoobenthos data extractions carried out by indicator experts ([Document 4-1](#)). The issues were the following:
  1. Data from all countries are not available
  2. Inconsistency in taxonomy used in reported data
  3. Differences in reporting of units and replicate samples (Examples in [Attachment 1](#))
24. The meeting took note of the inconsistency of reported units for zoobenthos and that for indicator calculation, actual counts/weight in the sample should be reported together with the sampled area (grab size).

25. The meeting took note that in the COMBINE manual for phyto- and zooplankton the preferred reporting unit is per volume, while for zoobenthos, there is no preferred/required way, which probably results in the current situation of inconsistent reporting of units.
26. The meeting agreed that these issues need to be documented and changed in the reporting guidance for the reported data to be useful for assessments.
27. The meeting discussed on reporting of lengths of zoobenthos and it could be potentially useful for assessment purposes.
28. The meeting welcomed the information by ICES that reporting of lengths of zoobenthos will be possible (not mandatory), and will be implemented in cooperation with the EN BENTHIC.

#### Phytoplankton:

29. The Meeting took note of the importance and thanked PEG group for continuous work on refining the existing biovolume file, related intercalibration exercise and mapping of species names to AphiaIDs.
30. The meeting discussed that it is required to thoroughly map the reporting requirements against the indicator needs for each of the data types. The process should be started well in advance of HOLAS III and included in the work plan. ICES Data Centre should be involved in the process.
31. The meeting took note of the information from German experts that working with legacy data (1979 onwards) has been problematic due to various issues listed in [Document 4-3](#) (provided after the meeting).
32. The meeting took note of the information from ICES that the majority of issues in historical data is related to data stemming from 1979-2000 and is resulting from different reporting formats used back then (HELCOM1), which contained a limited set of supporting parameters.
33. The meeting took note of the various gaps observed in the database for phytoplankton (listed in [Document 4-3](#), provided after the meeting).
34. The meeting welcomed the information by ICES that historical data can be resubmitted using simplified format or ERF3.2 to be able to tackle some of the issues stemming from use of old formats (HELCOM1).
35. The meeting agreed that it would be recommended to report any historical data that are stored in national databases but not included in COMBINE using ERF 3.2 format.

#### Zooplankton

36. The meeting took note of the information from German experts that working with legacy zooplankton data has been problematic due to two reasons:
  1. Legacy data for 1979-1996 is lacking information necessary to process and analyze zooplankton stock composition for the zooplankton mean size and total stock indicator.
  2. Normally biomass and abundance is reported using separate sample IDs, although the sampling event is the same, which has caused confusion in data extractions. The issue probably stems from the old reporting formats, and has to be examined further.
37. The meeting discussed on the use of SMVOL and FLVOL attributes in reporting of sample volume for zooplankton data.
38. The meeting took note of the information that presently, if the two fields are reported correctly, the values in them match. Information on whether the sample volume is based on the flowmeter readings is stored in the supporting field CLMET.
39. The meeting concluded that one field for reporting of the volume could be sufficient and to forward the question to the HELCOM ZEN.
40. The meeting discussed about the suitable range of volumes for better quality check for sample volume.
41. The meeting agreed that a proper maximum volume would be 50 m<sup>3</sup> and quality check for that can be implemented.

## Agenda Item 5. Conclusions and next steps

42. The meeting agreed that the conclusions and recommended actions resulting from the workshop should be presented and reviewed by thematic expert groups (PEG, ZEN, EN BENTHIC) and should be submitted to State & Conservation for commenting.
43. The meeting agreed on the following recommended actions, which should be taken forward:

**Action 1:** Data submission summaries at ICES accessions page on HELCOM (<http://data.ices.dk/accessions/AllCommissionSummary.aspx?commission=HELCOM>) should be amended to explicitly inform the user that it displays submissions that have been labeled as HELCOM.

**Action 2:** National data providers should be reminded to always report the HELCOM Monitoring data flagged as HELCOM COMBINE data (by defining MPROG=COMB), so the data are passing the HELCOM-specific quality checks, and are visible on the respective submission summary pages.

**Action 3:** It would be useful to carry out analysis on the usability of other global formats (e.g. extended Darwin Core used in EMODnet biology) to be used as acceptable format for data reporting and whether that would fulfill all indicator data requirements.

**Action 4:** It is required to thoroughly map reporting requirements against indicator needs for each data types. The process should be started in cooperation with ICES well in advance of HOLAS III and should be included in the work plan and parameters required by HELCOM would be listed in data reporting section of the relevant monitoring guidance document.

**Action 5:** Carry out mapping of species codes used in DOME without AphiaID to WoRMS/AphiaID utilizing the taxonomic expertise of relevant thematic groups (PEG, ZEN, EN BENTHIC) and indicator leads. This can be done using the 'Confirm or suggest mappings' tab in the file provided by ICES: [https://portal.helcom.fi/meetings/Data%20reporting%20WS%201-2019-689/MeetingDocuments/3-1%20Species\\_DOME\\_2\\_AphiaID.xlsx](https://portal.helcom.fi/meetings/Data%20reporting%20WS%201-2019-689/MeetingDocuments/3-1%20Species_DOME_2_AphiaID.xlsx).

**Action 6:** The COMBINE manual / reporting guideline should be updated regarding data reporting requirements for zoobenthos data: The actual counts/weight in the sample should be reported together with the sampled area (grab size), not the aggregated values.

**Action 7:** ICES is requested to implement quality check for zoobenthos data: When MUNIT is "nr", values in VALUE field should be whole numbers and not contain decimals.

**Action 8:** Implementing a quality check for zoobenthos data for area of the sampling device is needed, exact details should be defined with EN BENTHIC.

**Action 9:** HELCOM ZEN is requested to discuss the need to use both reporting fields in record 34: SMVOL and FLVOL. And to recommend reporting of only SMVOL in combination with CLMET as the indicator of whether the sample volume is calculated based on the flowmeter readings, haul length, or other methods.

**Action 10:** ICES is requested to implement quality check for zooplankton data: SMVOL field should have values < 50 m<sup>3</sup>.

**Action 11:** ICES Data Centre is requested to provide examples of prefilled reporting submissions in ERF 3.2 as a part of help documentation for data providers.

**Action 12:** Resubmission of historical data that has been reported in the past using old format (HELCOM1) to be re-reported using simplified format or ERF 3.2. Where legacy data are not found in national databases and cannot be resubmitted, ICES will reconvert the original HELCOM1 format to add previously unmapped species. These data will still be considered incomplete but can be used as 'presence' data.

**Action 13:** It is recommended for Contracting Parties to report any historical data that are stored in national databases but not included in COMBINE using ERF 3.2 format (or the related Simplified format)

and carry this out in coordination with ICES Data Center staff. This could be achieved by specific data call to report historical data.

**Action 14:** There is a need to further develop data outputs and dataview, to be further defined with relevant expert groups and indicator leads.

**Action 15:** Sweden is requested to provide the details of the data access issues for phytoplankton data

**Action 16:** ICES is requested to check and answer the findings reported in the document 4-3 and involve the workshop participants and the related expert groups (PEG, ZEN, EN BENTHIC) to the discussion, as necessary.

44. The Meeting was closed. The list of participants is included in Annex 1.

45. The draft meeting outcome document was circulated to the participants in 4 December for commenting by 9 December. Outcome was revised based on received comments.

#### Annex 1. List of participants.

Name	Contracting Party	Organisation
Anastasiia Kovtun-Kante	Estonia	Estonian Environment Agency
Anna Osypchuk		ICES
Astra Labuce	Latvia	Latvian Institute of Aquatic Ecology
Hans Mose Jensen		ICES
Iveta Jurgensone	Latvia	Latvian Institute of Aquatic Ecology
Janina Kownacka	Poland	National Marine Fisheries Research Institute
Joni Kaitaranta		HELCOM Secretariat
Lars Johan Hansson	Sweden	Swedish Agency for Marine and Water Management
Marilynn Sorensen		ICES
Michał Iwaniak	Poland	Institute of Meteorology and Water Management
Neil Holdsworth		ICES
Piotr Margonski	Poland	National Marine Fisheries Research Institute
Sven-Henrik Kleber	Germany	Federal Institut of Hydrology
Wera Leujak	Germany	Federal Environment Agency
Violetta Koszuta	Poland	IMGW-PIB