



Baltic Marine Environment Protection Commission

Working Group on the State of the Environment and Nature
Conservation

STATE & CONSERVATION
5-2016

Tallinn, Estonia, 7-11 November, 2016

Document title	Information from ICES Data Centre on the reporting of data to COMBINE
Code	3MA-2
Category	INF
Agenda Item	3MA – Environmental monitoring and data
Submission date	25.10.2016
Submitted by	ICES Data Centre

Background

This document includes information on new features related to the reporting of data to HELCOM COMBINE database hosted by ICES and clarifications related to the reporting of data.

Action requested

The Meeting is invited to:

- take note of the new application for the Station Dictionary
- take note that the Simplified Reporting Format is for reporting of new data, but it is not to be used for supplementing existing data submissions,
- take note of the Quality Control Database planned for release in November 2016.

Progress on Data Handling

Station Dictionary

In 2014 ICES reported to the State and Conservation Working Group on progress towards a new monitoring stations concept and management system (part of the discussion on the [BALSAM action plan](#)). The Station Dictionary web-based facility was released in August 2016. ICES regrets that the switch over to the new system and station concept happened as close as it did to the submission deadline. However, both for technical and for content reasons it was felt that we could not support the submission cycle for 2016 based on the old version of the station dictionary. The Station Dictionary was therefore released as a web application for Code Request Management at <http://vocab.ices.dk/Request/>.

Stations can now be added individually, and are based on institute's EDMO codes instead of country governance (which makes it easier for the individual institutes to have control over their stations in the ICES station list, and for ICES to avoid overwriting institute level edits/additions). Bulk uploads of new stations can be made by sending in new stations in the old ACCESS format for the time being. Further developments are planned for bulk changes/uploads of existing station information within the program. In addition:

- All station entries can be queried and viewed via an API using <http://vocab.ices.dk/services/>. This web service also exposes the stations and their attributes in XML.
- All previous submitters of the Station Dictionaries have received passwords.

Suggestions for features to ease working with the facility are welcome.

Simplified Reporting Format

Simplified Reporting Format for biological effects data which was developed under an OSPAR request, was expanded to include the reporting of all aspects of contaminant reporting except the reporting of certified reference material control charts. It is now possible to report contaminant and biological effects of contaminants in biota, sediment and sea water, biological community data including phytoplankton, zooplankton, phytobenthos, and zoobenthos data, as well as marine litter data in the Simplified Format.

The format can be downloaded at

http://ices.dk/marine-data/Documents/ENV/Environment_Formats.zip.

Once the csv file is prepared using the Simplified Reporting Format, data submitters can convert the files at <http://vocab.ices.dk/DataConversion/>. The resulting file is in the Environment Reporting Format version 3.2 (ERF3.2) and can be screened through DATSU <http://dome.ices.dk/datsu> as usual before submission to accessions@ices.dk. Support is available at all stages by contacting accessions@ices.dk.

The Simplified Format has extended reporting to the ICES Database to new users. Many of these new users have no database or advanced computer skills, no guidance concerning data reporting or deadlines, and need extensive instruction in using EXCEL and providing monitoring data. It was found that these new labs were not necessarily analysing the standard parameters usually monitored by the Commissions or were not using standard methods and therefore requested additional codes not previously covered by OSPAR or HELCOM programmes.

The Simplified Reporting Format is for new data not previously reported to ICES. This was interpreted by some to mean that partial submissions of new species or parameters were now accepted. This is not the case so partial submissions must be added to existing data submissions which can be found at <http://dome.ices.dk/browse/>.

Automatic submission facilities

Further adjustments have been made to the automatic resubmission facility and preparations have begun for an automatic submission tool. A consequence of hard deadlines, and the possibility to resubmit data has seen an increase in “unqualified data” being submitted to the database so that Contracting Parties achieve the deadline. This is time consuming to filter out and adds to the workload in making preventative checks in DATSU. It is therefore stressed to the State and Conservation Working Group and data submitters that skeleton files and missing data are not to be submitted to the database. Instead please contact accessions@ices.dk and discuss deadline problems so resubmission arrangements can be made.

Phytoplankton in COMBINE database

Based on feedback from the Phytoplankton experts in relation to developments of phytoplankton indicators, the ICES Data Centre has drafted a phytoplankton view (draft version at <http://dome.ices.dk/views/phytoplankton.aspx>). The view is addressing raised issues in relation to harmonisation of data and integration with the PEG biovolume list (http://www.ices.dk/marine-data/Documents/ENV/PEG_BVOL.zip).

The phytoplankton view was presented at the annual HELCOM PEG meeting (25-29 April 2016). The phytoplankton view can be used by the phytoplankton experts to evaluate the extent that the COMBINE database can support the phytoplankton indicator(s).

In relation to zooplankton and zoobenthos data, the ICES Data Centre has been in contact with experts responsible for these indicators. Initial data extractions have been made available. Work is going on to establish procedures to include oceanographic data (e.g. salinity and temperature) as supporting parameters for these data based on spatial and temporal criteria relevant for the stations.

Eutrophication assessment tool development and support

This is covered under separate reports in agenda item 4J.

Hazardous substances

Through the HELCOM BalticBOOST project, ICES has supported the work of the HELCOM Expert Network on Hazardous Substances (EN-HZ) in considering the application of the OSPAR assessment system (cf ‘MIME R-script’) for HELCOM core indicators (see proposal in meeting document 4J-15). The aim is to integrate the scripts into the data arrangements for the core indicators. The data-flow and national checking procedure to ensure that the relevant data is made available for the indicator calculations through HELCOM COMBINE, and that the indicator calculations are correct will be implemented through a HELCOM workspace, where the indicator R-scripts would form the first step (see document 4-26).

Outlier Proposal

At OSPAR MIME 2015, ICES proposed developing a facility for examining the whole database and time-series to identify outliers and to generate a summary report for MIME contact points. It would be in the online DOME view. Contracting Parties could then scrutinise and address national data issues ahead of assessment steps.

ICES are currently investigating previous outlier checks used in the past environment checks (SAS code). ICES will provide a short informational document on this prior to OSPAR MIME, which will also be made available to State and Conservation contacts and HELCOM EN-HZ.

Quality Control Database

The database of all Quality Control checks performed on all data managed at ICES is planned for release online in November 2016. The database is for all data types, and as such has proved challenging to present as a coherent list. The draft online tool for searching contents of the quality control database will be made available to State and Conservation 5-2016, and feedback would be welcome from HELCOM Contracting Parties.

The principle of the tool is to present the basic information for the quality checks performed for data submitted to ICES, allowing filtering by the three proposed headings (dataset type, programme – COMBINE for example, and nature of check). These can be changed if other criteria are more relevant. The title will have a ‘rollover’ information icon where the full meta record of the check is displayed. This includes information such as references to the decision basis for the inclusion of checks, date the check was established, SQL encoding of the check etc. The records will be fully downloadable in a standard format.

The screenshot shows the 'QC Database Query Tool' interface. It features three dropdown menus for filtering: 'Select dataset ...', 'Select program ...', and 'Select check nature ...'. The 'Select check nature ...' menu is open, showing options: 'Assessment requirement', 'Conversion issues', 'Cross-check', 'Outlier (statistical)', 'Position', and 'Range (Accepted values)'. Below the filters is a table of results with 24 rows. The table has two columns: a description of the check and a count. The third row is highlighted in blue.

Check Description	Count
Valid value - RECO check	33
PARAM Bile metabolites must be reported in matrix BI (bile) (WGBEC2013)	39
PARAM ACHL should be reported in MATRIX GI (gill) in shellfish	40
PARAM EROD and CYP1A should be reported in MATRIX US9 or UMIC (WGBEC2013)	41
Report DRYWT% instead of MOCON% (OSPAR2009)	42
Suspicious value in VALUE/PERCR/DETL or LMQNT	212
Effects parameters VDSI, INTS, IMPS, PCI for mixed sex populations require NOINP > 30	213
QA - PARAMs liver histopathology should report PARAM AGE	217
Survival time SURVT should be reported in MATRIX WO for <= 30 days (MUNIT=d)	220
ENV Hardcode - Value check: Out of range	226
MUNIT options for fish disease parameters are limited	232
MUNIT options for CYP1A are restricted	233
EROD should be reported with MUNIT pmol/minute/mg protein	234
MUNIT options for TBTINs are restricted	235
MUNIT options for SN ions are restricted	236
Effects parameter INTS, TREM SPCY and INTCLPRT should have integer range 0-4 (WGBEC2010/SGIMC2011)	238
Effects parameters VDSI and IMPS should have integer range 0-6 with MUNIT SI (WGBEC2010)	239
Effects parameters VDSI and IMPSI should have index range 0-5 with MUNIT IDX (WGBEC2010)	240
Effects parameter INTSI should have index range 0-4 with MUNIT IDX (WGBEC2010)	241
Effects parameter PCI should have index range 0-3.5 with MUNIT IDX (WGBEC2010)	242

At the bottom of the interface, there is a pagination bar showing '1 2 3 4 5 ...' and '1-20 of 25'.

Figure 1 QC checks online tool

Information on the Environmental Reporting Format (ERF)

The ICES Data Centre can accommodate reporting of both microlitter and macrolitter. No data have been reported using ERF3.2 or the Simplified Format. It should be noted that litter can be reported as a separate litter data type, or it can be added via extra litter records to any of the existing contaminant data types for biota, sediment or water and to the biological community data types for benthos and plankton.

Data Submission summaries are available on the ICES web:
www.ices.dk/marine-data/tools/Pages/Submission%20status.aspx.

See HELCOM Data:

- “All” for status of all data submissions to HELCOM
- “Last” six year status
- “Parameters in database”