



Baltic Marine Environment Protection Commission

Working Group on the State of the Environment and Nature
Conservation

STATE & CONSERVATION
15-2021

Online, 4-8 October 2021

| | |
|------------------------|---|
| Document title | Baltic Data Flows project – report of progress |
| Code | 3J-11 |
| Category | INF |
| Agenda Item | 3J – Progress of relevant HELCOM expert groups and projects |
| Submission date | 13.9.2021 |
| Submitted by | Secretariat |

Background

The HELCOM Secretariat was informed in fall 2019 about a potential project call by INEA (CEF-TC-2019-2: Public Open Data) which could be utilized to strengthen current HELCOM data collation and creation of public regional data products such as indicator and assessment datasets. The Secretariat requested State & Conservation contacts to explore the possibility of relevant national institutes/organizations that are responsible of collecting and making available environmental monitoring data to HELCOM to signal their interest in drafting a project proposal for this call.

The project proposal, coordinated by HELCOM Secretariat, gathered 6 partners from 4 HELCOM CPs (Finland, Latvia, Sweden, Denmark (ICES)) was approved for applying by HELCOM HODs and was successfully reviewed and Grant Agreement was signed in August 2020. The project period is 1.10.2020-30.9.2023. The project kick-off meeting was held on 12 October and [project website](#) established in April 2021.

The project provides concrete support for the HOLAS III process e.g., through the addition of new datatypes to existing data flows for eutrophication, through the implementation of DCAT-AP/INSPIRE in regional metadata catalogues, the further development of HELCOM Map and Data services and upgrade ICES Oceanographic data portal. It also works to review and develop methodology for hazardous substances indicator calculation and the further development of the CHASE tool. With regards to biodiversity the project reviews and updates indicator data requirements for harmonized indicator data product, methodology for indicator calculation, specifically for each biodiversity indicator and further develops the methodology for integrated biodiversity assessment and assessment tool (BEAT).

Action requested

The Meeting is invited to take note of the HOLAS III relevant work planned under the project and progress of work under the Baltic Data Flow project.



Baltic Data Flows 2020-2023

The Baltic Data Flows project, co-financed by the Connecting Europe Facility of the European Union, seeks to enhance the sharing and harmonisation of data on marine environment originating from existing sea monitoring programmes, and to move towards service-based data sharing.

In particular, open datasets will be made available by HELCOM to a wider community, such as European open data ecosystem, researchers, NGOs and private sector, in order to benefit from the availability of harmonised environmental data. Wider dissemination is achieved by sharing HELCOM metadata records to European Data Portal.

Baltic Data Flows will improve the capacity building of the national environmental data hosting organisations and providers of the consortium, in terms of quality control and solutions to make harmonised environmental data available. Members of the consortium will build and enhance their ICT infrastructure to support better the data sharing process.

Furthermore, data harvesting systems based on Application Programming Interfaces (APIs) will be developed with the aim to automatically integrate national datasets into a combined and harmonised regional dataset for partner institutes SMHI and SYKE for the environmental monitoring data that is reported to ICES.

Furthermore, Baltic Data Flows will further develop tools and indicator data flows for eutrophication, hazardous substances and biodiversity assessment. This will be carried out by further developing existing data flows, data view and aggregation tools at data host ICES and by further defining the indicator data needs. The focus of the work is indicators related to eutrophication, hazardous substances and biological community data reported to ICES (Seasonal succession of dominating phytoplankton groups, Zooplankton mean size and total stock (MSTS), State of the soft-bottom macrofauna community).

The work will include also further development of assessment tools and tool outputs in close cooperation with relevant expert groups / indicator leads and ultimately State & Conservation. For the biodiversity related parts, synergy with HELCOM BLUES¹ project outputs on Biodiversity assessment Tool (BEAT) will be explored.

The project will run from October 2020 to October 2023. Project partners include HELCOM, ICES, LHEI, SMHI, Spataneo inc., Stockholm University, and SYKE.

¹ The "HELCOM Biodiversity, Litter, Underwater noise and Effective regional measures for the Baltic Sea" (HELCOM BLUES) project is led by HELCOM and co-funded by the European Union. More information at <https://blues.helcom.fi>

Summary and timeline of project activities

The project activity summary table can be read alongside the GANTT Timeline, available [here](#). Below is a summary on progress of each project activity.

| Activity # | Activity title |
|------------|---|
| 1 | Increase of capacity at national data host institutes |
| 2 | Further development of existing data sharing platforms |
| 3 | Development and implementation of data harvesting |
| 4 | Addition of new datatypes to existing data flows (Eutrophication) |
| 5 | Further development of data processing and software used in hazardous substances assessment |
| 6 | Development of data processing and software to be used for biodiversity assessment |
| 7 | Dissemination and impact assessment |
| 8 | Project management |

Activity 1: Increase of capacity at national data host institutes

This activity is led by HELCOM in the role of project coordinator with the participation of LHEI and SMHI. It includes the following tasks:

- **Task 1.1.** Establishment of database platform for storing monitoring data (LHEI)
- **Task 1.2.** Establishment of metadata catalogue (LHEI)
- **Task 1.3.** Processing and harmonising biological community data according to a regionally agreed format (SMHI)
- **Task 1.4.** Synchronization of monitoring station visit-IDs between data types (SMHI)
- **Task 1.5.** Adding new export format to SHARK data (harvest API) (SMHI)

Task 1.1 is ongoing with the database platform established and currently undergoing testing by LHEI. Data migration and cleaning has been ongoing since August. In early September, stations, cruises, hydrological and nutrients data was uploaded and tested. The workshop for administrators and supervisors of the database administration is scheduled for 20 September. The first workshop for operational users within LHEI is scheduled for the last week of September. Data migration will continue in October and November 2021. **Task 1.2** is linked to task 1.1 with the database platform including the metadata catalogue. The remaining tasks are scheduled to start in quarter 3 of 2021.

Activity 2: Further development of existing data sharing platforms

This activity is led by HELCOM with the participation of ICES and Spatineo. It includes the following tasks:

- **Task 2.1.** Implementation of DCAT-AP/INSPIRE in regional metadata catalogues
- **Task 2.2.** Further development of HELCOM Map and Data service
- **Task 2.3.** Development and implementation of KPIs to monitor the access and use of datasets
- **Task 2.4.** Implementation of end user survey
- **Task 2.5.** Upgrade ICES Oceanographic data portal

Task 2.1 has started; the HELCOM Secretariat have established a Windows server 2016 (16GB RAM, 50GB disc space) for the new HELCOM metadata catalog. The latest stable GeoNetwork opensource and the required software will be installed on the server and new HELCOM metadata catalog supporting INSPIRE will

be established. The HELCOM Secretariat have developed Python scripts to fix errors in the current metadata records, to update HELCOM specific information in the current metadata records, and to validate metadata records locally. The validation includes INSPIRE technical guidance requirements and HELCOM specific information requirements. **Task 2.2** will extend across the duration of the project and HELCOM are already exploring ways to include additional data analysis and visualization tools alongside the HELCOM Map and Data Service (MADS). **Task 2.3** is almost complete, KPIs have been developed with baseline data used from findings of the MADS End-user Survey. KPIs will be measured against findings from a repeat of the End-user survey, along with the project impact assessment survey (Activity 7.4), at the end of the project. An interactive [Power BI dashboard](#) has been developed to provide partners with a dynamic analysis of the KPIs. The KPIs will be finalized at the next Activity 2 meeting scheduled on 29 September. **Task 2.4** the Map and Data Service (MADS) End-user survey was developed and deployed online using the Survey Monkey platform. It was made available from 30 March until 16 May 2021 (7 weeks). The survey was composed of 23 questions, [found here](#). The survey was disseminated widely across partner networks, relevant HELCOM Expert Working Groups, and the HELCOM Secretariat. The survey was also promoted on HELCOM social media channels ([LinkedIn](#), [Facebook](#), [Twitter](#)). A total of 158 respondents completed the survey. The survey report can be [found here](#) along with data summarized in a [Power Bi dashboard](#). **Task 2.5** has been started by ICES with the automation and standardization of data scripts.

Activity 3: Development and implementation of data harvesting

This activity is led by ICES with the participation of SMHI, SYKE and HELCOM. It includes the following tasks:

- **Task 3.1.** Definition and agreement on a data flow process and content to be harvested
- **Task 3.2.** Definition of requirements for API to be harvested
- **Task 3.3.** Implementation of the API at data providing institutes
- **Task 3.4.** Implementation of harvesting solution to retrieve data

Tasks 3.1 and 3.2 have started, all partners have presented on the types of marine data currently harvested and hosted by different database platforms. Partners are currently in the process of defining the harvesting process and formats (Darwin core vs. ERF 3.2). The APIs used have also been presented. **Tasks 3.3 and 3.4** are scheduled to start in quarter 4 2021 following completion of 3.1 and 3.2.

Activity 4: Addition of new datatypes to existing data flows (Eutrophication)

This activity is led by SYKE with the participation of SMHI, ICES and HELCOM. It includes the following tasks:

- **Task 4.1.** Definition of ferrybox data products for assessment
- **Task 4.2.** Definition of earth observation data products for assessment
- **Task 4.3.** Implementing tools for creating indicator data products
- **Task 4.4.** Making assessment data products FAIR

Task 4.1 has started with a review of available national and international ferrybox data from SYKE and SMHI. SYKE have provided information on their internal database 'Algabase' and status of ferrybox and bottle data. SMHI have mapped where ferrybox data is submitted and available. ICES have presented the plan to receive ferrybox data to assessment data flow. A data flow model has been developed and available [here](#). **Task 4.2** has started with a review of available national earth observation data by each activity partner. SYKE have presented [Tarkka](#) as a resource for open satellite data, based on a 20x20km grid from Sentinel II data. SMHI have presented the availability of Cyanobacterial data products based on Sentinel III and MERIS data. A data flow model has been developed and approved, available [here](#). Updates on tasks 4.1 and 4.2 have been presented to the HELCOM Intersessional Network on Eutrophication (IN-Eutrophication) who last met on 1-2 September 2021. **Tasks 4.3 and 4.4** are scheduled to start in quarter 3 2021.

Activity 5: Further development of data processing and software used in hazardous substances assessment

This activity is led by HELCOM with the participation of SYKE, LHEI and ICES. It includes the following tasks:

- **Task 5.1.** Reviewing and developing methodology for hazardous substances indicator calculation
- **Task 5.2.** Developing methodology for integrated assessment.
- **Task 5.3.** Further developing hazardous substances online assessment tool
- **Task 5.4.** Making assessment data products FAIR

Task 5.1 has started with a process to review and harmonise the HELCOM indicator needs, HELCOM Monitoring and Assessment Guidelines, and ICES quality checks (DATSU) to ensure all are fully aligned. Work has started to improve the visualisation of data and assessment outputs and plans to adapt the confidence assessment to correspond to each assessment unit have been approved by participants of the recent EN-HZ meeting on 2 September. The proposed approach and structure will ensure greater harmonisation with the BEAT and HEAT integrated assessment tools applied under the biodiversity and eutrophication assessments, respectively. **Task 5.2 and Task 5.3** will be carried out by NIVA Denmark Water Research, the sub-contractor who was selected following an open procurement process in May 2021. The 'notice for procurement' was published on the EU TED (Tenders Electronic Daily), an online version of the 'EU Supplement to the Official Journal', dedicated to European public procurement. The procurement notice is available [here](#). This work will improve the assessment confidence levels within the CHASE integrated assessment tool. This will ensure that if other/new substance concentration indicators are endorsed as core then the tool will have the capacity to include these by HOLAS III. Other developments aim to improve the efficiency of the data flows (i.e., connecting them to the COMBINE/DOME system under ICES) to work towards greater automation. **Task 5.4** is scheduled to start in quarter 3 2021.

Activity 6: Development of data processing and software to be used for biodiversity assessment

This activity is led by SYKE with the participation of SMHI, LHEI, SU, ICES and HELCOM. It includes the following tasks:

- **Task 6.1.** Reviewing and updating indicator data requirements for harmonized indicator data products
- **Task 6.2.** Reviewing and developing methodology for indicator calculation, specifically for each biodiversity indicator
- **Task 6.3.** Developing methodology for integrated biodiversity assessment and assessment tool
- **Task 6.4.** Making assessment data products FAIR

Task 6.1 has been started with HELCOM providing a suggestion to amend biological community data (phytoplankton, zooplankton, zoobenthos) monitoring guidelines to cover required data reporting aspects. The HELCOM Phytoplankton Expert Group (PEG) and Zooplankton Expert Network (ZEN) was updated on the suggested changes to indicator data product creation and data reporting practicalities, i.e. what information is required to be reported and on what level to be useful for indicators. The monitoring guideline amendments were presented and discussed also in [HELCOM EN BENTHIC 5-2021 meeting](#) and were submitted to HELCOM State & Conservation 14-2021 meeting. After collecting comments from CPs, the modifications were agreed and are implemented to both COMBINE manual (zoobenthos) and guidelines (phytoplankton and zooplankton) by State & Conservation 15-2021 meeting. **Task 6.2** has started with a review of the biodiversity taxonomy linked to the World Register of Marine Species (WoRMS). A cross-check of missing brackish/freshwater species has been conducted resulting in the addition of one benthic freshwater species. Flow charts for indicator calculation of [phytoplankton](#) and [zooplankton](#) have been developed, each showing how data is combined from a variety of sources and fed into the BEAT assessment tool. **Tasks 6.3. and 6.4** are underway for indicator calculations and results, the HELCOM Secretariat GitHub

repository will be used for data collection for BEAT including the tool itself. This task shares synergies with a parallel project ([BLUES](#)) that is working on a pelagic habitats assessment.

Activity 7: Dissemination and impact assessment

This activity is led by HELCOM with the participation of ICES and Spatineo. It includes the following tasks:

- **Task 7.1.** Create visual identity for the project
- **Task 7.2.** Create project website
- **Task 7.3.** Outreach for dissemination by project partners
- **Task 7.4.** Impact assessment

Task 7.1 has been completed and the results can be viewed on the [Baltic Data Flows project website](#). They include unique project logos and presentation templates. **Task 7.2** has been completed with the website launch on 31 March 2021. The website URL is: <https://balticdataflows.helcom.fi/>. The website currently features blog posts, as featured in the links below:

- [Monitoring HELCOM data services using the 'Spatineo Monitor'](#)
- [Three questions to... Rita Poikane and Astra Labuce, LHEI](#)
- [Three questions to Matthew Richard, the BDF Project Coordinator – HELCOM](#)

Task 7.3 is ongoing. A [project video](#) has been produced with inputs from project partners (HELCOM, LHEI and SMHI) who recorded a video message in response to the following two questions: 1) *Why do we need harmonised data on the marine environment?*, 2) *What will be the concrete benefits of the Baltic Data Flows project?* The video was shared widely across social media platforms ([Twitter](#), [Facebook](#), [LinkedIn](#)), and the dedicated BDF website and page of the HELCOM website. **Task 7.4** is scheduled to start towards the end of the project in quarter 1 2023, however with help from Spatineo the survey tool to be used for the impact assessment has been developed, available [here](#). The project impact assessment will be partly informed by a repeat of the MADS End-user survey along with a key informant interview.

Activity 8: Project Management

This activity is led by HELCOM as the project coordinator. The kick-off workshop of the Baltic Data Flows was organized on 12 October 2020. The initial plan during the Grant Agreement preparation phase was to have a face-to-face meeting in Helsinki, which is the location with the highest number of project participants. Due to Covid-19 related travel restrictions the workshop was organized as an online meeting using the Zoom platform provided by the Project Coordinator. All partners were able to attend and all together 27 persons attended the kick-off meeting. An internal project workspace was established on SharePoint accessible only by the project participants for preparing, sharing, and making available content to project partners. The project workspace is a closed workspace under [HELCOM Meeting Portal](#) hosted and managed by HELCOM Secretariat. All project partners signed up to HELCOM Meeting portal as external users and access rights to the closed project workspace was given by the project coordinator. A Project contact list has been established for each activity and is available in the workspace. The Consortium Agreement has been completed by the Project Coordinator and signed by all partners. The first project Steering Group meeting was held on 12 March 2021, 6-months after the kick-off meeting. All partners of the project attended, represented by designated Steering Group focal points. Each lead partner prepared an update on each activity, focusing on main tasks, deliverables, milestones, timelines as well as interlinkages with other project activities. The minutes of meeting are available [here](#). Activity specific group meetings are organized by lead partners in agreed timelines (typically monthly meetings). Steering Group Meetings will be held every 6 months with the next one scheduled for 23 September 2021. The project data management plan has also been developed.