



REPORT ON DATA COLLATED & UPLOADED TO VARIOUS DATABASES

Published by:

Helsinki Commission – HELCOM
Katajanokanlaituri 6 B
00160 Helsinki, Finland

www.helcom.fi

Information and views expressed in this publication are the authors' own and might vary from those of the Helsinki Commission or its members.

All rights reserved. Information included in this publication or extracts thereof, with the exception of images and graphic elements that are not HELCOM's own and identified as such, may be reproduced without prior consent on the condition that the complete reference of the publication is given as stated above.

This document is a publication under Work Package 2 of the PROTECT BALTIC project.

For bibliographic purposes this document should be cited as:

"Protect Baltic Report on what data has been collated & uploaded to various databases – HELCOM (2025)"

© Baltic Marine Environment Protection Commission – Helsinki Commission (2025)

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency (CINEA). Neither the European Union nor the granting authority can be held responsible for them.

Layout: Paul Truth

Photo credits: cover photo © Adobe Stock



Funded by
the European Union



EU
MISSIONS

RESTORE OUR OCEAN AND WATERS

D 2.7- Report on what data has been collated and uploaded to the various databases

Deliverable 2.7 of PROTECT BALTIC outlines the data collected and collated so far, specifying the databases and storage systems where this information has been uploaded.

HELCOM and Baltic Sea data

As the lead for Work Package 2 (WP2: Data), HELCOM is a long standing and well-established data hub for regional Baltic Sea data. The HELCOM Secretariat has decades of experience working with national ministries, agencies and data hosts on data reporting and has a strong mandate, granted directly by the countries, to submit data calls, as well as to process and store data on a regional level.

Under this mandate, HELCOM collects and maintains a wealth of data on the Baltic Sea, organizing it into thematic databases that are accessible through online data portals, services, and tools. Many of these will be used in PROTECT BALTIC.

HELCOM Data portals

Online platforms that provide access to datasets and data-related resources for users.

Table 1. List of HELCOM Data portals

Name	Description
HELCOM Map and data service (MADS)	Contains all geospatial data used in HELCOM assessments and reports. Linked with HELCOM Metadata catalogue.
HELCOM Metadata catalogue	Contains metadata records of all geospatial data used in HELCOM assessments and reports as well as in depth documentation of databases. Linked with HELCOM MADS.
BASEMAPS	Basemaps provides access to maritime spatial planning (MSP) data in the Baltic Sea: 'Input data,' as a service from data sources, and 'Output data,' which is MSP plans sea uses.
HELCOM Biodiversity data portal	Contains all macrospecies observations data made available by HELCOM Contracting Parties.
Shipping data platform	Provides access to various shipping-related data collected by HELCOM within the Maritime and Response Working Groups, expert groups and related projects.
Ballast water DST	The HELCOM/OSPAR Ballast Water Exemptions Decision Support Tool contains data on NIS species observations and port sampling.
HELCOM MPA Portal	Access point to relevant information on the Baltic Sea Protected Areas (HELCOM MPAs).

HELCOM Databases

Organized collections of thematic data. Data is typically collected on an annual reporting schedule following regionally agreed reporting formats and quality assurance procedures.

Table 2. List of HELCOM databases

Name	Description
<u>Oceanographic data (nutrients)</u>	Oceanographic monitoring data gathered by HELCOM Contracting Parties.
<u>Hazardous substances in marine environment</u>	Monitoring data on concentrations of contaminants in biota, sediment and seawater.
<u>Biological community data (open sea)</u>	Monitoring data on abundance and biomass of open sea phytoplankton, zooplankton, zoobenthos and phytobenthos.
<u>Biodiversity database</u>	HELCOM Biodiversity database contains macrospecies observations.
<u>Pollution Load Compilation (PLC)</u>	HELCOM PLC Database contains waterborne nutrient and contaminant discharges to the Baltic Sea.
<u>Impulsive noise</u>	Impulsive noise events reported by HELCOM Contracting Parties.
<u>Continuous noise</u>	Continuous low-frequency sound monitored under the HELCOM programme in the Baltic Sea
<u>Shipping accidents</u>	Shipping accidents reported to EMSA/EMCIP and verified by HELCOM Maritime
<u>Illegal oil spills</u>	Location and size of spills observed during aerial surveillance flights by HELCOM Contracting Parties
<u>Disposal sites of dredged material</u>	Disposal and deposit sites of dredged material reported by HELCOM Contracting Parties
<u>Radioactive substance in the environment (MORS Env)</u>	Observations from monitoring of environmental radioactivity in the Baltic Sea
<u>Radioactive discharges (MORS Discharge)</u>	Airborne and waterborne discharges resulting from nuclear facilities in the Baltic Sea and North Sea region

HELCOM data services

Provide standards-compliant methods for accessing environmental data related to the Baltic Sea. These services facilitate machine-to-machine communication, making it possible for GIS professionals, developers, and researchers to integrate HELCOM datasets into their workflows.

Data services are documented in the [HELCOM Metadata catalogue](#) and are accessible from [HELCOM ArcGIS REST Services Directory](#) as follows:

1. Select a thematic service from the service directory.
2. Select the interface type from the top of the page (WFS / WMS / Rest).

Available service types are:

- **OGC Web Feature Service (WFS)** enables users to request and manipulate vector feature information from the HELCOM databases in a machine-readable format (for example, GeoJSON or GML).
- **OGC Web Map Service (WMS)** provides pre-rendered map images of spatial datasets directly from HELCOM's servers.
- **ESRI ArcGIS Rest interface** allows direct access to HELCOM's spatial data through ArcGIS-compatible REST endpoints.

HELCOM data tools

The software tools developed for use in HELCOM assessments are listed below and are available in the HELCOM GitHub repositories.

Table 3. List of HELCOM data tools

Name	Description
SPIA desktop tool (GitHub)	ArcGIS Pro toolbox to run the spatial pressure and impact assessment (SPIA) with any pressure layer and ecosystem component datasets that meet the specified requirements.
SPIA online tool	Online tool for conducting the spatial pressure and impact assessment (SPIA). The newest input datasets are available, with links to corresponding metadata. Users can explore results and trace them back to contributions from human activities, pressure layers and ecosystem components. Downloadable outputs, including raster files and statistical information.
Sufficiency of Measures (SOM)	A Python model developed by HELCOM to assess the potential improvements in environmental state and pressures achievable with existing measures.
HEAT	HELCOM Eutrophication Assessment Tool (HEAT) is used to carry out HELCOM Eutrophication assessment and indicator result calculation from monitoring data.
CHASE	HELCOM Hazardous Substances Status Assessment Tool (CHASE). The CHASE tool integrates monitoring data on hazardous substances in water, sediments and biota as well as bio-effect indicators and is based on a substance- or bio-effect-specific calculation of a 'contamination ratio'.

<u>BEAT</u>	HELCOM Biodiversity Assessment Tool (BEAT) is used for creating integrated assessments of biodiversity in the Baltic Sea based on biodiversity indicator results.
-------------	---

List of used abbreviations

BioDiv db	HELCOM Biodiversity Database
EMODnet	European Marine Observation and Data network
ESA	European Space Agency
ESAS	European Seabirds at Sea
DATRAS	Database of trawl surveys
DOI	Digital Object Identifier
ICES	International Council of for the Exploration of the sea
ICES DOME	Data repository hosted by ICES
MADS	HELCOM Map and data service
MPA	Marine Protected Area
Laji.fi	Finnish species database
LUKE	Natural Resources Institute Finland
OBIS	Ocean Biodiversity Information System
POHJE	Finnish benthic species data system
SYKE	Finnish Environment Institute

Data calls

Environmental data was collected through a data call issued on 25 March 2024 (Data call 1), while data related to MPA management was gathered through two data calls, issued on 2 November 2023 (Data call 2) and 19 December 2024 (Data call 3). Additionally, a large number of datasets on environmental data has been collated directly from open sources.

The geodatabases utilized for PROTECT BALTIC data include:

- [HELCOM Map and data service](#) (MADS)
- [HELCOM Biodiversity Database](#) (BioDiv DB)
- MPA Portal Database (MPA Portal DB) – a new version is currently under development and will be published under Work Package 8 (WP8: MPA Portal) of PROTECT BALTIC.

Original datasets received from Contracting Parties are stored on HELCOM's restricted drive, accessible only to designated Secretariat data personnel. General data storage for non-restricted work files and harmonized spreadsheet data is maintained on PROTECT BALTIC's Work Package 2 (WP2: Data) SharePoint site, with access limited to relevant project partners.

Data call 1: Summary

For Data call 1, the collected data can be categorized into:

1. Species observation data
 - A significant portion of species observation data was collated and uploaded in cooperation with the HELCOM RED LIST II project. Data quality checking, processing and uploading concluded in autumn 2024.
 - Most of the remaining data was obtained from open sources, including international data repositories such as ICES DOME, which also hosts several HELCOM databases.
2. Other environmental data (e.g. supporting parameters for modelling referred to as predictor variables)
 - Much of this data was sourced from open-access platforms.
 - All collected data was formatted as spreadsheets, shapefiles, or raster files to facilitate further processing by other work packages.
 - Selected final products will be integrated into HELCOM MADS and BioDiv DB, following these principles:
 - Map layers, coverage data, and raster files will be included in MADS.
 - Species observations conforming to the BioDiv DB data model (a Darwin Core variant) will be added to BioDiv DB.

A total of 112 datasets were requested from eight countries: Germany, Denmark, Estonia, Finland, Lithuania, Latvia, Poland and Sweden. The detailed requests can be found in the original data call document “PROTECT BALTIC Data call 1” (DOI: [10.5281/zenodo.14858548](https://doi.org/10.5281/zenodo.14858548)).

Response rate and data availability

Some of the requested datasets were solicited for the first time. A key objective of Data call 1 was to assess the availability of the requested data. However, the response rate was low, which is typical for ad hoc data calls that are not part of an established reporting framework with dedicated resources at data-providing institutions (for more information see D 2.5 - Updated gap analysis of data to get a clear overview of further data needs and infrastructure development).

At the time of writing, several datasets remain pending or under preparation (i.e. undergoing harmonization and external validation by experts). Nonetheless, the proportion of country-specific data is anticipated to remain between 10–30%.

Furthermore, none of the country-specific datasets were submitted in the exact required format. Instead, Contracting Parties provided or linked to datasets that could be utilized for project purposes but did not fully meet the requested format or include the required metadata on observations. The HELCOM Secretariat subsequently compiled and harmonized the received data.

Data sources and harmonization

Most of the datasets for Data call 1 were sourced from open-access data repositories, including:

- European Union institutions.
- International organizations (e.g. ICES).
- National open-data platforms (e.g. laji.fi in Finland).

The open-data sources used by PROTECT BALTIC for data collation and harmonization are listed in Table 4. The data collated under Data call 1 is included in the overview of species and environmental data, respectively.

Data call 2

On 2 November 2023, Task 2.4 under WP2 issued a data call to Contracting Parties to identify a contact person (or multiple contacts, if necessary) within each country who is familiar with and available to provide advice and/or data on:

1. The existing types of Marine Protected Areas (MPAs), including both nationally established and currently planned MPAs.
2. Information regarding MPA Management Plans.

The detailed requests can be found in the original data call document “Combined Data Request to BioDiv - WP2 of PROTECT BALTIC” (DOI: [10.5281/zenodo.14858548](https://doi.org/10.5281/zenodo.14858548)). Further details regarding this data call are provided in Deliverable 2.5 Updated Gap Analysis of Data to Obtain a Clear Overview of Further Data Needs and Infrastructure Development.

All collected data is stored in HELCOM’s restricted SharePoint site for use by the respective work packages and for MPA Portal development. Once finalized, the data will be published in the Baltic Sea MPA Portal.

Data call 3

On 19 December 2024, Task 2.4 issued a data call to Contracting Parties regarding measures outside of MPA Management Plans.

The objective of the third data call was to:

1. Review the previously submitted information and provide additional data on MPAs in their respective areas, where relevant.
2. Supply shapefiles/polygons for all MPAs, including any zoning information associated with them.
3. Provide information on national/domestic measures or sectoral instruments that affect MPAs but are not included in official MPA Management Plans.

The detailed requests can be found in the original data call document “Second Data Request to BioDiv - WP2 of the PROTECT BALTIC (Measures)” (DOI: [10.5281/zenodo.14858548](https://doi.org/10.5281/zenodo.14858548)). Further details regarding this data call are provided in Deliverable 2.5 Updated Gap Analysis of Data to Obtain a Clear Overview of Further Data Needs and Infrastructure Development.

All collected data is stored in HELCOM’s restricted SharePoint site for use by the respective work packages and for MPA Portal development. Once finalized, the data will be published in the MPA Portal.

Data collation overview

Much of the data collated for use in PROTECT BALTIC has been collated from open sources and actively collected by the project partners. This data is used concomitantly with data reported directly to the HELCOM Secretariat by the countries. A list of open sources used can be found in Table 4.

Table 4. Used open sources for collected datasets

Dataset	Open source
Benthic species distribution models	SYKE (Fin)
Benthic species observations	ICES (int), POHJE (Fin), laji.fi (Fin)
Bird species distribution models	ICES/ESAS (int)
Bird species observations	OBIS (int)
Bottom organic matter content concentration	
Chlorophyll a coverage	ICES (int)
Coloured organic dissolved matter (cdom)	ESA, Copernicus Publications (EU)
Fish and decapod crustacean species distribution models	ICES (int), LUKE (Fin)
Fish and decapod crustacean species observations	ICES DATRAS (int), LUKE (Fin)
Hypoxia	ICES (int)
Sediment maps	EMODnet (EU)
Sediment maps point	
Sedimentation accumulation rate	EMODnet (EU)
Spawning and nursery areas of fish	LUKE (Fin)
Spawning and nursery areas of fish point	LUKE (Fin)
Turbidity	EMODnet (EU)
Sea water potential temperature at sea floor - bottomT [°C]	Copernicus
Ocean mixed layer thickness defined by sigma theta - mlotst [m]	Copernicus
Sea ice area fraction - siconc	Copernicus
Sea ice thickness - sithick [m]	Copernicus
Sea surface height above sea level - sla [m]	Copernicus
Sea water salinity - so [0.001]	Copernicus
Sea water salinity at sea floor - sob [0.001]	Copernicus
Sea water potential temperature - thetao [°C]	Copernicus
Eastward sea water velocity - uo [m/s]	Copernicus

D 2.7 - Report on what data has been collated and uploaded to the various databases

Northward sea water velocity - vo [m/s]	Copernicus
Mass concentration of chlorophyll a in sea water (CHL)	Copernicus
Mole concentration of nitrate in sea water (NO3)	Copernicus
Mole concentration of phosphate in sea water (PO4)	Copernicus
Mole concentration of ammonium in sea water (NH4)	Copernicus
Mole concentration of dissolved molecular oxygen in sea water (O2)	Copernicus
Sea water ph reported on total scale (pH)	Copernicus
Surface partial pressure of carbon dioxide in sea water (spCO2)	Copernicus
Secchi depth of sea water (ZSD)	Copernicus
Mole concentration of dissolved molecular oxygen in sea water (at bottom) (O2)	Copernicus
Net primary production of biomass expressed as carbon per unit volume in sea water (daily accumulated) (PP)	Copernicus
Sea surface wave significant height (SWH)	Copernicus
Sea surface wave mean period from variance spectral density second frequency moment (MWT)	Copernicus
Bathymetry/depth	Baltic Sea Hydrographic Commission
Wave exposure index	AquaBiota
Chlorophyll-a	ICES
Chlorophyll-a	EMODnet
Water body phosphate	EMODnet
Dissolved oxygen concentration	EMODnet
Dissolved oxygen concentration	ICES
Water body silicate	EMODnet
Inorganic nitrogen	EMODnet
CEFAS sedimentation accumulation rates	CEFAS
Shipping density	HELCOM
Potential cumulative impacts on benthic habitats	HELCOM
Baltic Sea Impact Index BSII	HELCOM
Baltic Sea Pressure Index	HELCOM
Baltic Sea Impact Index BSII	HELCOM
Potential cumulative impacts on benthic habitats	HELCOM
Total Suspended Matter Baltic Sea	EMODnet
Volume absorption coefficient of radiative flux in sea water due to dissolved organic matter and non algal particles ADG443 [m-1]	Copernicus
Volume absorption coefficient of radiative flux in sea water due to phytoplankton APH443 [m-1]	Copernicus
Volume backwards scattering coefficient of radiative flux in sea water due to particles BBP443 [m-1]	Copernicus
Mass concentration of cryptophytes expressed as chlorophyll in sea water CRYPTO [mg/m3]	Copernicus
Mass concentration of diatoms expressed as chlorophyll in sea water DIATO [mg/m3]	Copernicus
Mass concentration of dinophytes expressed as chlorophyll in sea water DINO [mg/m3]	Copernicus
Mass concentration of greenalgae and prochlorophytes expressed as chlorophyll in sea water GREEN [mg/m3]	Copernicus
Mass concentration of microphytoplankton expressed as chlorophyll in sea water MICRO [mg/m3]	Copernicus
Mass concentration of nanophytoplankton expressed as chlorophyll in sea water NANO [mg/m3]	Copernicus

Mass concentration of picophytoplankton expressed as chlorophyll in sea water PICO [mg/m ³]	Copernicus
Mass concentration of prokaryotes expressed as chlorophyll in sea water PROKAR [mg/m ³]	Copernicus
Surface ratio of upwelling radiance emerging from sea water to downwelling radiative flux in air	Copernicus
Volume attenuation coefficient of downwelling radiative flux in sea water KD490 [m ⁻¹]	Copernicus
Sea surface wave maximum height VCMX [m]	Copernicus
Sea surface wave significant height VHM0 [m]	Copernicus
Sea surface primary swell wave significant height VHM0_SW1 [m]	Copernicus
Sea surface secondary swell wave significant height VHM0_SW2 [m]	Copernicus
Sea surface wind wave significant height VHM0_WW [m]	Copernicus
Sea surface wave from direction VMDR [°]	Copernicus
Sea surface primary swell wave from direction VMDR_SW1 [°]	Copernicus
Sea surface secondary swell wave from direction VMDR_SW2 [°]	Copernicus
Sea surface wind wave from direction VMDR_WW [°]	Copernicus
Sea surface wave maximum crest height VMXL [m]	Copernicus
Sea surface wave from direction at variance spectral density maximum VPED [°]	Copernicus
Sea surface wave stokes drift x velocity VSDX [m/s]	Copernicus
Sea surface wave stokes drift y velocity VSDY [m/s]	Copernicus
Sea surface primary swell wave mean period VTM01_SW1 [s]	Copernicus
Sea surface secondary swell wave mean period VTM01_SW2 [s]	Copernicus
Sea surface wind wave mean period VTM01_WW [s]	Copernicus
Sea surface wave mean period from variance spectral density second frequency moment VTM02 [s]	Copernicus
Sea surface wave mean period from variance spectral density inverse frequency moment VTM10 [s]	Copernicus
Sea surface wave period at variance spectral density maximum VTPK [s]	Copernicus
Air density [kg/m ³]	Copernicus
Backscatter distance to model function - bs_distance	Copernicus
Surface downward eastward stress - eastward_model_stress [N/m ²]	Copernicus
Surface downward eastward stress - eastward_stress [N/m ²]	Copernicus
Eastward wind - eastward_wind [m/s]	Copernicus
Vertical component of surface downward stress curl - model_stress_curl [N/m ³]	Copernicus
Divergence of surface downward stress - model_stress_divergence [N/m ³]	Copernicus
Magnitude of surface downward stress - model_stress_magnitude [N/m ²]	Copernicus
Wind to direction - model_wind_to_dir [°]	Copernicus
Surface downward northward stress - northward_model_stress [N/m ²]	Copernicus
Surface downward northward stress - northward_stress [N/m ²]	Copernicus
Northward wind - northward_wind [m/s]	Copernicus
Eastward wind - se_eastward_model_wind [m/s]	Copernicus
Wind speed - se_model_speed [m/s]	Copernicus
Atmosphere relative vorticity - se_model_wind_curl [s ⁻¹]	Copernicus
Divergence of wind - se_model_wind_divergence [s ⁻¹]	Copernicus
Northward wind - se_northward_model_wind [m/s]	Copernicus
Vertical component of surface downward stress curl - stress_curl [N/m ³]	Copernicus

Divergence of surface downward stress - stress_divergence [N/m3]	Copernicus
Atmosphere relative vorticity - wind_curl [s-1]	Copernicus
Divergence of wind - wind_divergence [s-1]	Copernicus
Wind speed - wind_speed [m/s]	Copernicus
Magnitude of surface downward stress - wind_stress_magnitude [N/m2]	Copernicus
Wind speed (WIND)	Copernicus
Northward wind (WIND)	Copernicus
Eastward wind (WIND)	Copernicus
Quality flag wind speed (WIND)	Copernicus
Wind to direction (WIND)	Copernicus
Surface downward eastward stress - eastward_stress [N/m2]	Copernicus
Surface downward eastward stress bias - eastward_stress_bias [N/m2]	Copernicus
Surface downward eastward stress standard deviation of differences - eastward_stress_sdd [N/m2]	Copernicus
Eastward wind - eastward_wind [m/s]	Copernicus
Eastward wind bias - eastward_wind_bias [m/s]	Copernicus
Eastward wind standard deviation of differences - eastward_wind_sdd [m/s]	Copernicus
Surface downward northward stress - northward_stress [N/m2]	Copernicus
Surface downward northward stress bias - northward_stress_bias [N/m2]	Copernicus
Surface downward northward stress standard deviation of differences - northward_stress_sdd [N/m2]	Copernicus
Northward wind - northward_wind [m/s]	Copernicus
Northward wind bias - northward_wind_bias [m/s]	Copernicus
Northward wind standard deviation of differences - northward_wind_sdd [m/s]	Copernicus
Sea water turbidity TUR [FNU]	Copernicus
Mass concentration of suspended matter in sea water - SPM [g/m3]	Copernicus
Volume backwards scattering coefficient of radiative flux in sea water due to particles (BBP)	Copernicus
Surface ratio of upwelling radiance emerging from sea water to downwelling radiative flux in air	Copernicus
Mole concentration of dissolved inorganic carbon in sea water - dissic [mole/m3]	Copernicus
Sinking mole flux of particulate organic matter expressed as carbon in sea water - expc [mol/m2/d]	Copernicus
Sea floor depth below sea level - model_depth [m]	Copernicus
Net primary production of biomass expressed as carbon per unit volume in sea water - nppv [mg/m3/day]	Copernicus
Mole concentration of phytoplankton expressed as carbon in sea water - phyc [mmol/m3]	Copernicus
Mole concentration of silicate in sea water - si [mmol/m3]	Copernicus
Surface partial pressure of carbon dioxide in sea water - spco2 [Pa]	Copernicus
Mole concentration of zooplankton expressed as carbon in sea water - zooc [mmol/m3]	Copernicus
Eastward sea water velocity assuming no tide - uos_detided [m/s]	Copernicus
Northward sea water velocity assuming no tide - vos_detided [m/s]	Copernicus
Sea surface height above geoid assuming no tide - zos_detided [m]	Copernicus
Sea surface height above sea level - sla [m]	Copernicus
Upward sea water velocity - wo [m/s]	Copernicus
Sea ice thickness - ice_thickness [cm]	Copernicus
Ice thickness range in cm - thickness_range [cm]	Copernicus

Sea ice classification - sea_ice_classification	Copernicus
Ice concentration range - concentration_range [%]	Copernicus
Sea ice area fraction - ice_concentration [%]	Copernicus
Sea ice extent - sea_ice_extent [1e6km2]	Copernicus
Sea water alkalinity expressed as mole equivalent - talk [mol/m3]	Copernicus
Mole concentration of dissolved iron in sea water - fe [mmol/m3]	Copernicus
Bathymetry	GEBCO
Baltic Sea Pressure Index	HELCOM
Baltic Sea Pressure Index BSPI	HELCOM
Potential effect of continuous noise to mobile species	HELCOM
Potential cumulative impacts of eutrophication and hazardous substances	HELCOM
AIS vessel tracks: Drifting longlines, Fixed gear, Pole and line, Purse seines, Trawlers, Trollers, Unknown	Global Fishing Watch
Distance from shore in meters	Global Fishing Watch
Distance from port in meters	Global Fishing Watch
Anchorage (global database of anchorage locations where vessels congregate)	Global Fishing Watch
Global Datasets of AIS-based Fishing Effort and Vessel Presence	Global Fishing Watch
Industrial vessels detected with satellite imagery and classified with deep learning	Global Fishing Watch
Offshore infrastructure detected with satellite imagery and classified with deep learning	Global Fishing Watch

Species observation data

A significant portion of species observation data was collated and uploaded in cooperation with the HELCOM RED LIST II project (see D 2.3 - Updated Biodiversity Database with new data and information). In total, not previously reported data was provided for 3891 species, representing close to 6 million new species observations for use in PROTECT BALTIC's modelling and analyses. This data represents a combination of data reported to HELCOM as part of official data calls, and data harvested from open sources (see Table 5 for an overview of data sources for the species observation data).

As much of the collected data was newly acquired and had not been previously obtained, it did not fully align with the data models defined in the data calls, and in the case of PROTECT BALTIC Data call 1, the call specified extensive metadata on monitoring (e.g. device type, sediment characteristics) which have not previously been requested.

All received species data is harmonized into the format specified in the data call template spreadsheets and integrated into the BioDiv DB, where applicable and in compliance with its data model for mandatory parameters (Darwin Core). The newly processed data is now available for other project work packages in spreadsheet format for further processing (see Table 6 for an overview of the current status of recently reported observational species data).

Table 5. Sources of data for the species observations included in the HELCOM Biodiversity database.

Contracting Parties - DE (Institution unknown)
Contracting Parties - DK (Institution unknown)
Contracting Parties - EE (Institution unknown)
Contracting Parties - FI (Institution unknown)
Contracting Parties - LT (Institution unknown)
Contracting Parties - LV (Institution unknown)
Contracting Parties - PL (Institution unknown)
Contracting Parties - RU (Institution unknown)
Contracting Parties - SE (Institution unknown)
Association Fish and Environment Mecklenburg-Vorpommern e.V.
Department of Aquatic Resources, Swedish University of Agricultural Sciences
Estonian Marine Institute
Institute of Food Safety, Animal Health and Environment
National Institute of Aquatic Resources, Technical University of Denmark
National Marine Fisheries Research Institute
Natural Resources Institute
Nature Research Center
Provincial Government of Åland Islands
University of Rostock
University of Tartu
Lithuanian Sea Museum
Zoological Museum University of Copenhagen (now Natural History Museum of Denmark) and Fisheries and Maritime Museum Esbjerg Denmark
DMM - German Oceanographic Museum, Stralsund
European Seabirds at Sea Database
Finnish Ministry of the Environment
FTZ - Research and Technology Centre Westcoast
GSM - Society for the Conservation of Marine Mammals
University of Gdansk
Museum of the World Ocean
ITAW - Institute for Terrestrial and Aquatic Wildlife Research
Pro Mare
The Fishery and Maritime Museum, The National Forest and Nature Agency
NRM-Swedish Museum of Natural History
Gauja National Park
Aarhus University
LUNG - State Agency for Environment, Nature Conservation and Geology Mecklenburg-Vorpommern
MSC - Marine Science Center
GIOŚ - Chief Inspectorate for Environmental Protection
SMHI - Swedish Meteorological and Hydrological Institute

IOW - Leibniz Institute for Baltic Sea Research
Environmental Protection Agency
The Danish Environmental Protection Agency
OCEANA
Finnish Institute of Marine Research
ICES - International Council for the Exploration of the Sea
MADS - Danish national database for marine data
The Swedish Species Information Centre
SYKE - Finnish Environment Institute
LUOMUS - Finnish Museum of Natural History
MariLim
The Danish Biodiversity Information Facility, Botanical Museum
Klaipeda University, Marine research institute
Landesamt für Natur und Umwelt des Landes Schleswig-Holstein
Uusimaa ELY Centre
Estonian Environment Agency
Parks and Wildlife Finland
Agency for Environment, Nature Conservation, and Geology
Lund University
Alleco
BfN - Federal Agency for Nature Conservation
Blinova and Tolstikova 1972
Bluemel et al. 2002
Fuerhaupter et al. 2008
Kylin 1947
Meyer et al. 2005
Mäensivu 2006
Nystroem 2009
Scheinin and Soederstroem 2005
Schubert and Blindow 2003
Schubert et al. 2003
Wærn M 1945
Wærn M 1952
TraFi - Finnish Transport Safety Agency
Swedish Transport Agency
Maritime and Coastguard Agency, Environmental Policy Branch
ILT - Human Environment and Transport Inspectorate
Ministry of Ports: Official Contact Points for BWMC A-4 Exemptions
Parks and Wildlife Finland and Finnish Environment Institute
AquaNIS
Marine Monitoring AB
Johanna Bergkvist

Federal Maritime and Hydrographic Agency, BSH
Contracting Parties - SE (other, described on notes)
HOLAS 3 data call

The status of complementary species observation data provided through Data call 1 for each country is presented in Table 6, which is updated daily. This is data that has been collated to complement the species observational data already processed and included in the HELCOM Biodiversity Database. A snapshot of the dataset statuses was taken on 1 February 2025. The coding system is explained below:

- A Data is under approval by the respective expert
- C There is country-specific data available
- I There is no information about the data
- N Country has informed that the data is not available
- O There is open data available and possibility for national data (still under investigation)
- P Data availability is postponed due to national data collection schedule
- S Data is only partly available. There are big gaps comparing to data call template.
- U Data is under harmonization process

Table 6. Status of different species observation datasets per country

Dataset	DE	DK	EE	FI	LT	LV	SE
Benthic species observations	O	C	C	O	O	O	O
Bird species observations	U	U	A	U	U	U	U
Fish and decapod crustacean species observations	O	O	O	C	O	O	C

Environmental data

In this context, other environmental data refers to all collected datasets that do not pertain to species observations. This data is primarily utilized by work packages involved in modelling and analysis and many of the datasets represent predictor variables used to drive the models.

Direct collating of predictor variable data

A predictor variable is a dataset used in predictive modelling such as species distribution models as an explanatory (or independent variable or feature). For instance, depth and water temperature could predict the distribution of a given fish species if we use a predictive model within a spatial context, e.g. with fish presence/absence as the response (or independent) variable and depth and temperature as the explanatory variables (fish p/a ~ depth + temperature). The predictor variables are key datasets for PROTECT BALTIC as they will be used to predict the distribution of a large set of species and habitats in the whole Baltic Sea through species distribution models. The results of these models will then be used as inputs by other work packages in their activities in PROTECT BALTIC.

PROTECT BALTIC has collected 148 predictor datasets (included in Table 4) which were all reviewed and evaluated for their suitability for the planned modelling within the project.

Out of these 148 datasets, 22 were evaluated as suitable (see Table 7), 62 kept for further review, and 64 were not suitable. Most of predictors were collected from Copernicus (78 predictors), followed by HELCOM with the second most (10 predictors). The collated predictors are available at a pan-Baltic scale and are the product mostly of satellite observations that are processed and extrapolated to the whole Baltic Sea as raster files at a given resolution (e.g. Copernicus: 1NM). The temporal span is highly variable and ranges from 1 to 61 years depending on the predictor.

The data is being maintained on the PROTECT BALTIC SharePoint site, with access restricted to relevant project partners.

All predictor variables data which are included in the final modelling will be made available also within HELCOM MADS once the first round of modelling concludes, to ensure that all the input data can be accessed and reviewed in the same place.

Table 7. Datasets evaluated as suitable for modelling in PROTECT BALTIC.

Sea water potential temperature at sea floor - bottomT [°C]	Baltic Sea Physics Reanalysis	Copernicus
Ocean mixed layer thickness defined by sigma theta - mlotst [m]	Baltic Sea Physics Reanalysis	Copernicus
Sea ice area fraction - siconc	Baltic Sea Physics Reanalysis	Copernicus
Sea water salinity - so [0.001]	Baltic Sea Physics Reanalysis	Copernicus
Sea water salinity at sea floor - sob [0.001]	Baltic Sea Physics Reanalysis	Copernicus
Sea water potential temperature - thetao [°C]	Baltic Sea Physics Reanalysis	Copernicus
Mass concentration of chlorophyll a in sea water (CHL)	Baltic Sea Biogeochemistry Reanalysis	Copernicus

Mole concentration of nitrate in sea water (NO ₃)	Baltic Sea Biogeochemistry Reanalysis	Copernicus
Mole concentration of phosphate in sea water (PO ₄)	Baltic Sea Biogeochemistry Reanalysis	Copernicus
Mole concentration of ammonium in sea water (NH ₄)	Baltic Sea Biogeochemistry Reanalysis	Copernicus
Mole concentration of dissolved molecular oxygen in sea water (O ₂)	Baltic Sea Biogeochemistry Reanalysis	Copernicus
Sea water pH reported on total scale (pH)	Baltic Sea Biogeochemistry Reanalysis	Copernicus
Surface partial pressure of carbon dioxide in sea water (spCO ₂)	Baltic Sea Biogeochemistry Reanalysis	Copernicus
Secchi depth of sea water (ZSD)	Baltic Sea Biogeochemistry Reanalysis	Copernicus
Mole concentration of dissolved molecular oxygen in sea water (at bottom) (O ₂)	Baltic Sea Biogeochemistry Reanalysis	Copernicus
Net primary production of biomass expressed as carbon per unit volume in sea water (daily accumulated) (PP)	Baltic Sea Biogeochemistry Reanalysis	Copernicus

Bathymetry/depth	Baltic Sea Bathymetry Database	Baltic Sea Hydrographic Commission
Wave exposure index		AquaBiota
Total Suspended Matter Baltic Sea	EMODnet Physics	EMODnet
Volume attenuation coefficient of downwelling radiative flux in sea water KD490 [m-1]	Baltic Sea Multiyear Ocean Colour Plankton, Reflectances and Transparency	Copernicus
Sea ice thickness - ice_thickness [cm]	Baltic Sea ice concentration, extent, and classification time series	Copernicus
Sea ice area fraction - ice_concentration [%]	Baltic Sea ice concentration, extent, and classification time series	Copernicus

Collating environmental data via data call

After collation, the data is structured into raster, harmonized spreadsheet, or shapefile formats. The final outputs from analysis and modelling will be added to the MADS database where relevant (as outlined in D 2.4 - Updated Map and data service for the Baltic Sea region). Additionally, some outputs may be accessed or utilized in computational analyses within the Baltic Sea MPA Portal. The intended destinations of the final data products are detailed in Table 8.

Table 8. Current data formats and locations of end products after processing. Possible usage in MPA portal not yet known.

Dataset	Current data format	End products
Benthic species distribution models	Shapefile	MADS
Benthic species observations	Spreadsheet	BioDiv db, spreadsheet
Bird species distribution models	Rasterfile	MADS
Bird species observations	Spreadsheet	BioDiv db, spreadsheet
Bottom organic matter content concentration	Shapefile, spreadsheet	MADS
Chlorophyll a coverage	Shapefile	MADS
Coloured organic dissolved matter (cdom)	Rasterfile	MADS
Fish and decapod crustacean species distribution models	Rasterfile	MADS
Fish and decapod crustacean species observations	Shapefile, spreadsheet	BioDiv db, spreadsheet
Hypoxia	Shapefile, spreadsheet	MADS
Sediment maps	Geodatabase	MADS
Sediment maps point	No data available	
Sedimentation accumulation rate	Geodatabase, rasterfile	MADS
Spawning and nursery areas of fish	Geodatabase, shapefile, rasterfile	MADS
Spawning and nursery areas of fish point	No data	
Turbidity	Multiple formats	MADS

At present, three datasets contain partially restricted data. The terms of use for these datasets must be confirmed with the respective data providers. Furthermore, PROTECT BALTIC utilizes data from previous and ongoing HELCOM projects, such as RED LIST II, for which separate permission from data providers is required if data beyond what is already published in MADS and BioDiv DB is to be used.

For several datasets, Contracting Parties have indicated that the available data had already been submitted in the context of previous HELCOM projects, such as RED LIST II and HOLAS 3. This data is accessible via MADS and BioDiv DB and is actively utilized by PROTECT BALTIC.

The status of collected data for each country is presented in Table 9, which is updated daily. A snapshot of the dataset statuses was taken on 1 February 2025. The coding system is explained below:

- A Data is under approval by the respective expert
- C There is country-specific data available
- I There is no information about the data
- N Country has informed that the data is not available
- O There is open data available and possibility for national data (still under investigation)
- P Data availability is postponed due to national data collection schedule
- S Data is only partly available. There are big gaps comparing to data call template.
- U Data is under harmonization process

Table 9. Status of different datasets per country.

Dataset	DE	DK	EE	FI	LT	LV	SE
Benthic species distribution models	N	U	N	U	A	P	I
Bird species distribution models	N	U	A	N	U	U	N
Bottom organic matter content concentration	C	I	C	C	P	N	I
Chlorophyll a coverage	O	O	O	O	C	O	O
Coloured organic dissolved matter (CDOM)	O	O	O	O	O	O	O
Fish and decapod crustacean species distribution models	P	N	N	U	N	P	C
Hypoxia	C	O	O	O	C	O	O
Sediment maps	O	O	O	O	O	O	O
Sediment maps point	N	N	N	N	P	N	N
Sedimentation accumulation rate	S	S	S	O	S	S	O
Spawning and nursery areas of fish	P	U	C	U	I	P	C
Spawning and nursery areas of fish point	P	I	N	U	I	P	I
Turbidity	O	O	O	O	O	O	O



The PROTECT BALTIC project is funded by the European Union under Grant agreement ID 101112866. This publication was funded by the European Union. Its contents are the sole responsibility of the author(s) and do not necessarily reflect the views of the European Union.



Funded by
the European Union



**EU
MISSIONS**

RESTORE OUR OCEAN AND WATERS