eosc | Blue-Cloud2026

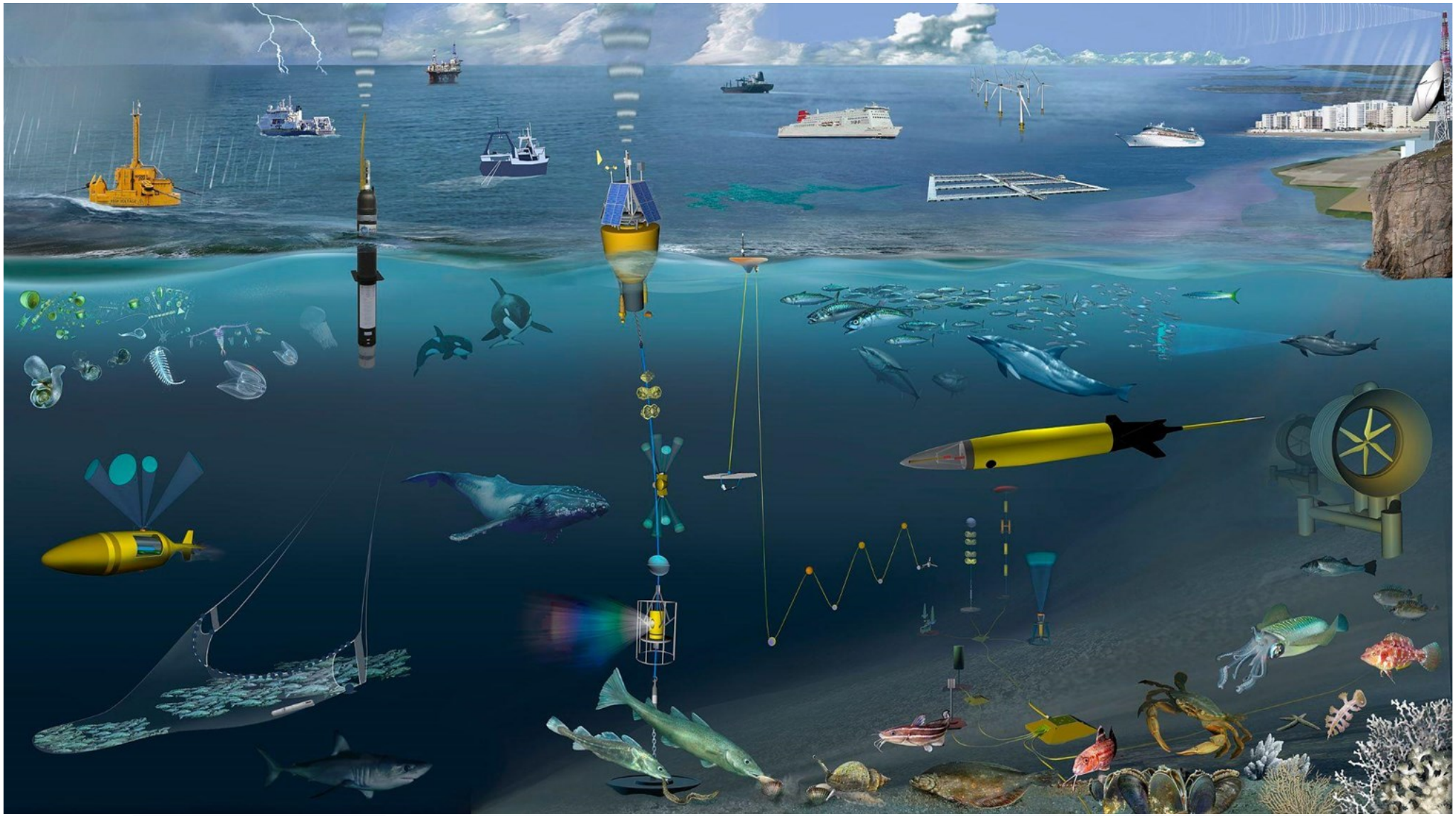
Blue-Cloud interdisciplinary environments

VLIZ

Patricia Martin Cabrera, VLIZ

Funded by the European Union

The Ocean Observing Landscape



The Blue Data Infrastructures

The Blue Data Infrastructures

Blue-Cloud goal

Blue-Cloud is a federated European FAIR and Open Research Ecosystem for oceans, seas, coastal and inland waters that:

- **Develops a thematic marine extension to EOSC for open web-based science, serving the needs of the EU Blue Economy and the Marine Knowledge agendas**
- **Providing federated access to leading European Marine Data Infrastructures and a Virtual Research Environment with data, models and reusable methodologies that tackle multidisciplinary marine environments**

Blue Cloud services

Discovery Data & Access Service

Easy access to multidisciplinary data sources from a federated search

Blue Cloud VRE

Collaborative research environment: Data sharing, analytics (jupyter, R…)

Blue Data Catalog

Exploitation of results
Increased findability

Blue Cloud VLabs & Workbenches

Access to re-usable Methodologies & data products

Blue Cloud Virtual Labs-VLabs

Coastal Ocean observations along Europe

Coastal currents from observations

Carbon-Plankton Dynamics

Global Fisheries Atlas

Marine Environmental Indicators

Virtual Laboratories

Data sharing

- Workspace
- Dataspace
- Repositories

Data analytics

- High Throughput Computing
- Notebook
- RStudio

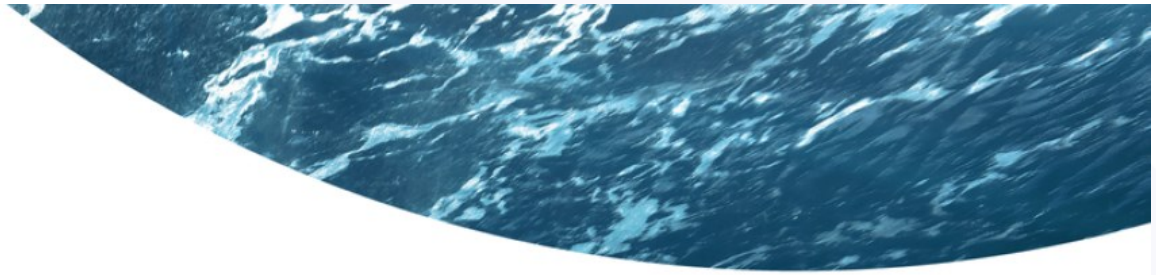
Social networking

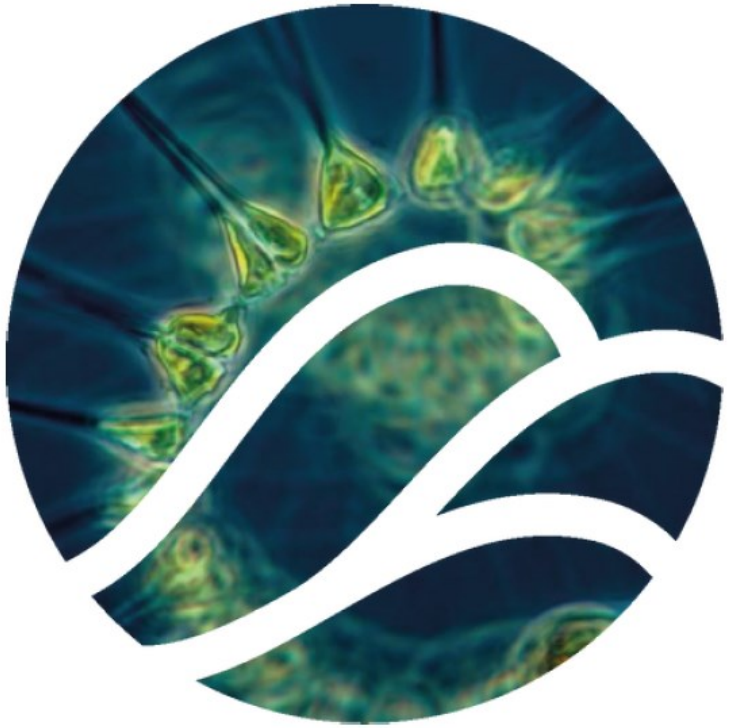
- Messages
- Posts and replies
- User profiling

Research Object Publishing

- Catalogue
- Thredds
- GeoNetwork

Blue-Cloud
Virtual Labs





Zoo- & Phytoplankton Essential Ocean Variable products

This Virtual Lab provides a description of the current state of plankton communities and forecasts their evolution, representing valuable information for the modelling, assessment and management of the marine ecosystems.

Partners:

VLIZ LOV LIÈGE université

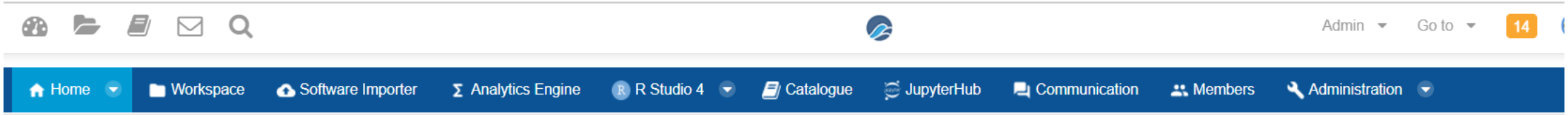
Data sources through Blue-Cloud:

EurOBIS, EMODnet Biology, LifeWatch, GEBCO, SeaDataNet, World Ocean Atlas, NOAA, Copernicus Marine Service, Argo GDAC, GlobColour

UN SDGs addressed

Services

- Zooplankton EOVs
- Phytoplankton EOVs
- Scientific validation



Phytoplankton EOV docs

Phytoplankton EOV products

The phytoplankton Essential Ocean Variables (EOV) demonstrator aims to provide a methodology to generate **global open ocean three-dimensional (3D) gridded products of** (1) **chlorophyll a concentration (Chla)**, which is a proxy of the total phytoplankton biomass, and (2) **Phytoplankton Functional Types (PFT)**, as a proxy for phytoplankton diversity, based on vertically-resolved *in situ* data of ocean physical properties (temperature and salinity) matched up with satellite products of ocean color and sea level anomaly.

The Machine Learning method

The methods have been developed following the method of Sauzede et al. (2016), which relies on machine learning, specifically on an artificial neural network (Multi-Layer Perceptron, MLP), and retrieves the vertical distribution of biogeochemical properties from merged ocean colour and hydrological data. The MLPs consist of several layers: one input layer, one output layer and one or several hidden layers. Each layer is composed of neurons, which are elementary transfer functions that provide outputs when inputs are applied.

Here, following the same philosophy as the method developed by Sauzede et al. (2016), two different MLP-based algorithms are developed for the independent retrieval of the Chla and of the PFT EOV products:

1. The first MLP retrieves the depth-resolved Chla product and is trained using in-situ depth-resolved measurements of Chla, temperature and salinity (T/S), from the global BioGeoChemical-Argo (BGC-Argo) observation network (coriolis Global Data Center), matched-up with global satellite-derived products. The MLP input layer is composed of three main components:

- Surface satellite-based inputs from the Copernicus Marine Environment Monitoring Service (CMEMS) and GlobColour, such as the ocean colour remote sensing reflectance (Rrs) at five wavelengths, the photosynthetically available radiation (PAR), and the sea level anomaly (SLA),
- Depth-resolved ocean physical properties such as components derived from a principal component analysis (PCA) of the T/S vertical profiles and the mixed layer depth ,

Zoo & Phytoplankton EOV products

Guidelines on how to run the scripts with the data provided or your own data

Data sources

VARIABLES	DATA SOURCES	DATA ACCESS
Satellite-derived reflectance	OCEANCOLOUR_GLO_OPTICS_L3_REP_OBSERVATIONS_009_086	Blue Cloud
Satellite Sea Level Anomaly	SEALEVEL_GLO_PHY_L4_REP_OBSERVATIONS_008_047 product	Blue Cloud
Physical data: T, S, MLDGlobal ARMOR 3D products	CMEMS MULTIOBS_GLO_PHY_REP_015_002	Blue Cloud
BGC-Argo Float NetCDF files (S-files)	ftp.ifremer.fr/ifremer/argo/; http://www.argo.ucsd.edu	Blue Cloud
Satellite-derived Photosynthetically Available Radiation	ftp://ftp.hermes.acri.fr	Blue Cloud Vlab
High-performance liquid chromatography (HPLC) data	http://www.obs-vlfr.fr/proof/cruises.php	Blue Cloud Vlab
Bathymetry	https://www.gebco.net/data_and_products/gridded_bathymetry_data/	Blue Cloud Vlab

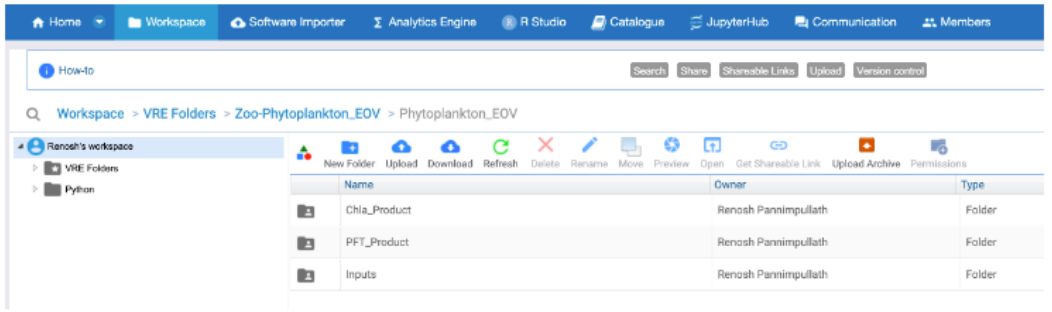
Data access= Blue Cloud, Data is accessible via the Blue Cloud Data Access service. It is also uploaded in the Vlab.

Data access= Blue Cloud Vlab, Data is not accessible via the Blue Cloud Data Access service. It has been uploaded in the Vlab.

Step by step guideline to use the service

All necessary files to generate monthly global 3D Chla product and the Phytoplankton functional types (PFT) for the year 2018 are located in `Workspace/VRE Folders/Zoo-Phyto`

The input files are located in the `Inputs` folder in the `Phytoplankton_EOV` folder. The `Inputs` folder contains all the necessary input data to derive the Chla and PFT products f



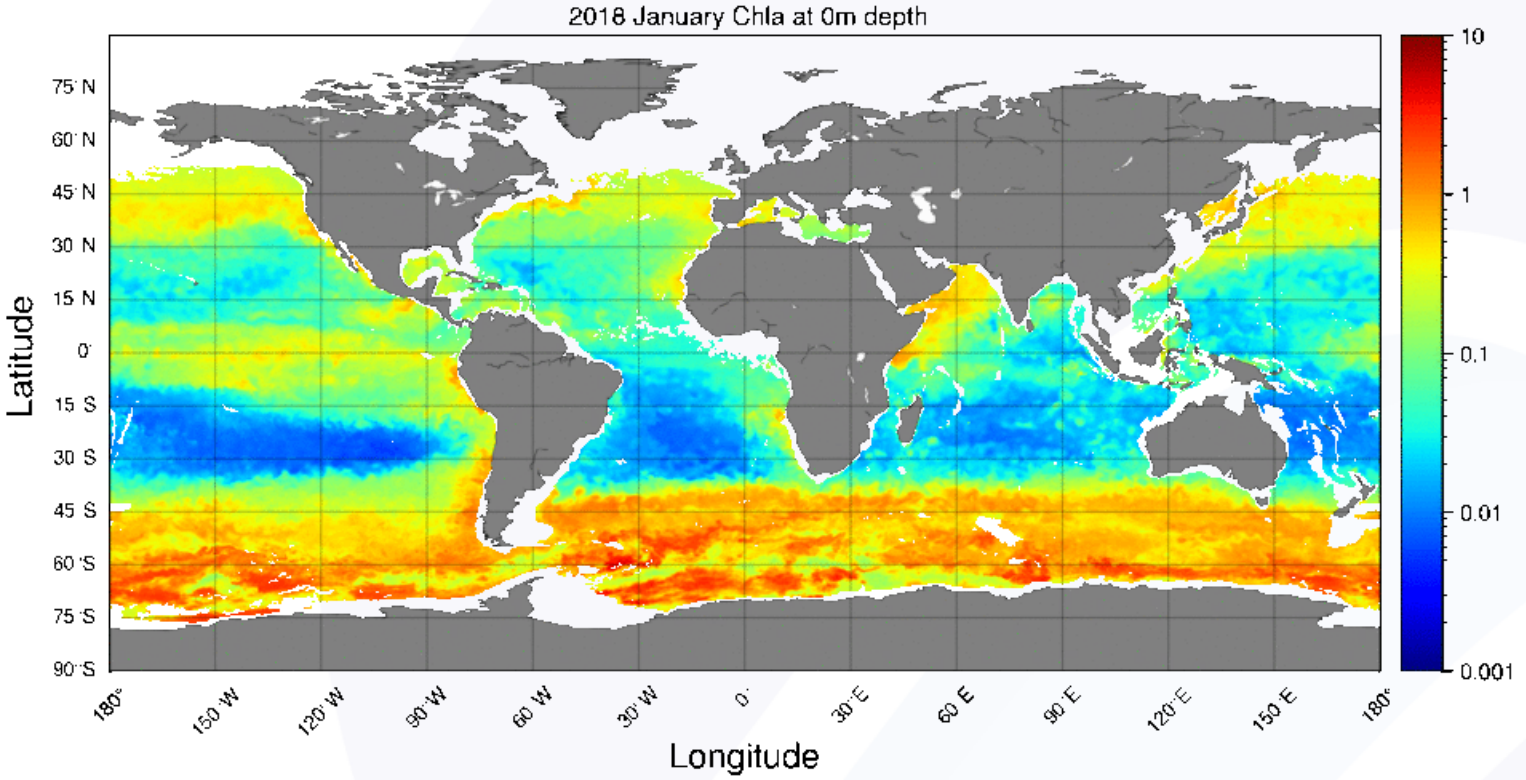
The `Chla_Product` and `PFT_Product` folders include three subfolders: `Programs`, `Outputs`, and `Plots`.

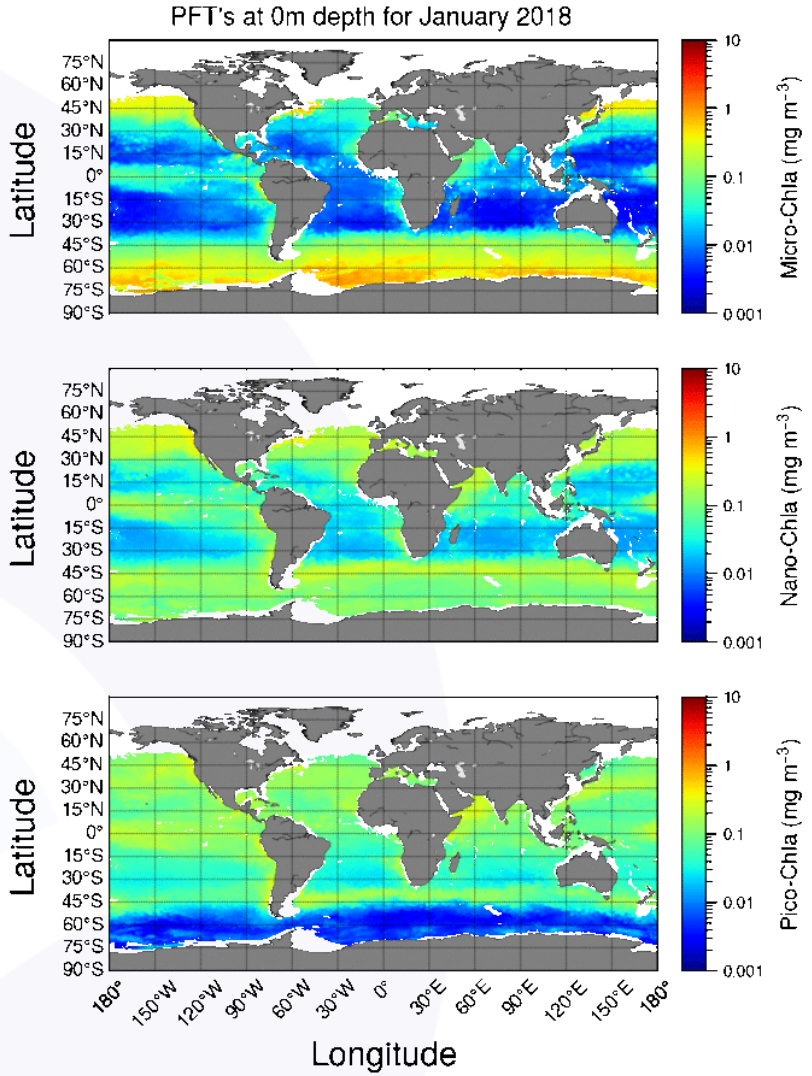
- The `Programs` folder contains 2 Jupyter notebooks and two folders (`Functions` and `Models`).
  - The `Functions` folder contains all the necessary functions required to generate the 3D global products, and
  - The folder `Models` contains the trained MLP models and PCA models.
- The `Outputs` folder contains the output global 3D products generated for each month of the year 2018 (.nc files).
- The `Plots` folder contains the visualization of the output products as , .png files (2D spatial plots for 36 depths).

Zoo & Phytoplankton EOV products

OUTPUTS

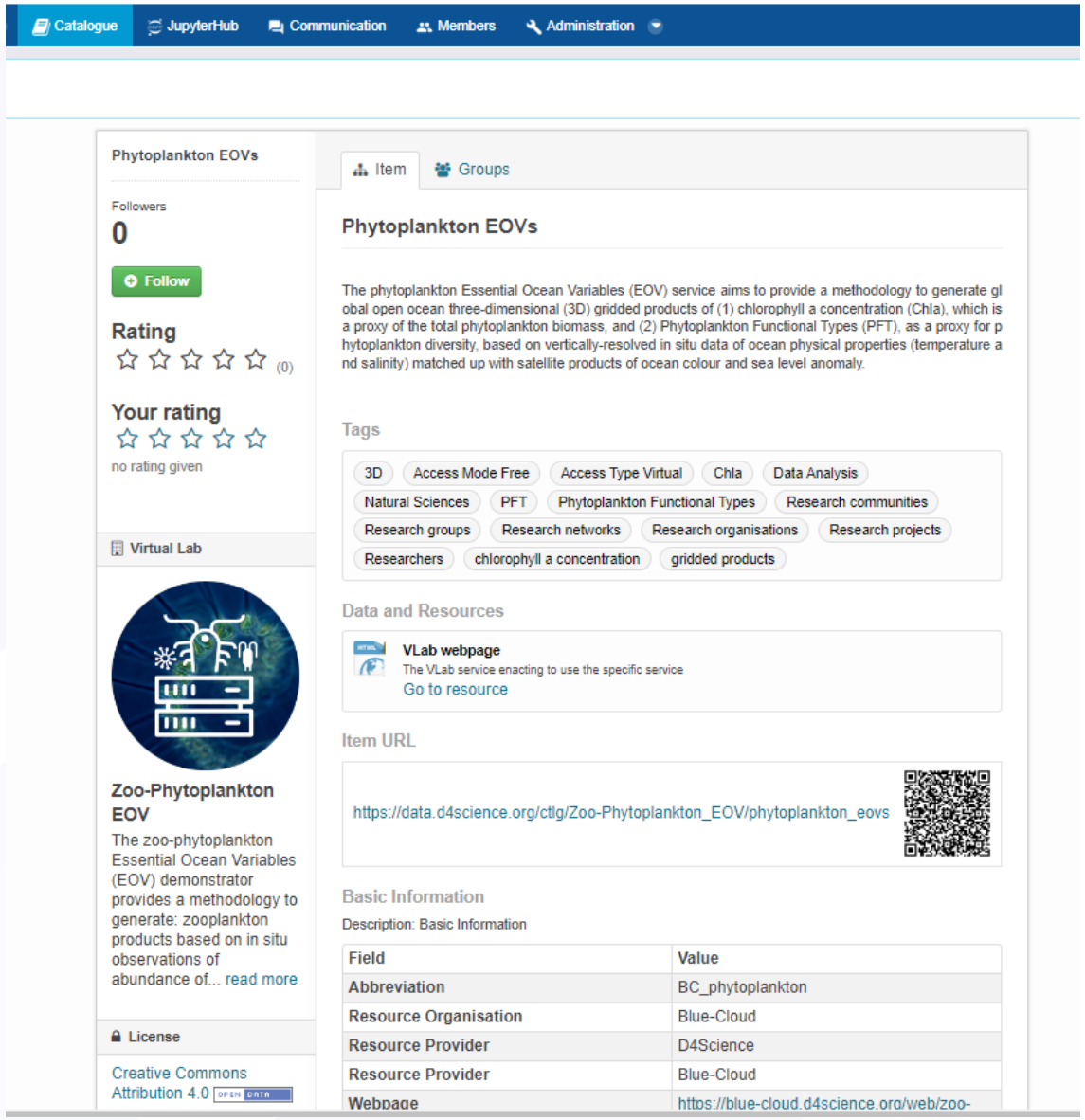
Deriving global ocean 3D Chlorophyll-a concentrations using machine learning techniques





Zoo & Phytoplankton EOV products

Data, methodologies and outputs available through the Blue Data Catalog



Zoo & Phytoplankton EOV products

The phytoplankton EOV product published as Copernicus CMEMS product

Zoo & Phytoplankton EOV products

With the support of the Blue-Cloud Hackathon:

The [Wildlife Tracker for Oceans team] used the output data from the Phytoplankton EOV product to update their application that visualizes the movements of marine megafauna in real-time. Chl-a data was used to include hotspot of Phytoplankton concentration.

Blue-Cloud
Virtual Labs

The Wildlife Tracker for Oceans

The "Wildlife Tracker" is a cloud geo-framework dedicated to Marine Protected Areas (MPAs) management based on biologging and ocean satellite data. The platform offers a unique opportunity to overlay and enrich the movement tracks of wildlife over ecogeographical data layers such as Phytoplankton hot spots to observe in near real-time what may be influencing the animal activities and to spatially assess their meaningful habitats as MPAs.

Partners:

Wildlife Tracker
Real-time movement analytics

whaleshark PROJECT

Data sources through Blue-Cloud:
The Global ocean three-dimensional (3D) key phytoplankton product of chlorophyll-a (Chla) concentration, as a proxy for total phytoplankton biomass from Vlabs.

Main target users:
Marine scientists dedicated to MPAs and conservation efforts

Services introduction:
The Wildlife Tracker has diversified its activities and it can offer real-time wildlife monitoring based on alert system, web map gallery creation, and biologging data enrichment with satellite data. As a cloud-based platform the service is done by monthly/yearly subscription and we offer cooperation to non-profit foundations based on blue economy model.

UN SDGs addressed

14 LIFE BELOW WATER

Conclusions

- Blue-Cloud provides a Virtual Research Environment that **boost scientific collaboration and Big Data Analysis**.
- Integrating **multidisciplinary data** allows to create innovative data products that can respond complex research questions.
- Federating **Blue Data Infrastructures** builds the bridges needed to develop **Data Lakes**.
- Providing data, products & reusable methodologies feeds the **Digital Twin Ocean**.
- Blue-Cloud data, products and VRE enable our one main objective: **Ocean integration** towards our **One Ocean** #UNDecade

GEBCO Compilation Group (2022) GEBCO 2022 Grid (doi:10.5285/e0f0bb80-ab44-2739-e053-6c86abc0289c).

eosc | Blue-Cloud2026

blue-cloud.org

@bluecloudeu

blue-cloud org

Funded by
the European Union