

Blue-Cloud Piloting innovative services for Marine Research & the Blue Economy

Sara Pittonet Gaiarin, Trust-IT Services,
Blue-Cloud Coordinator

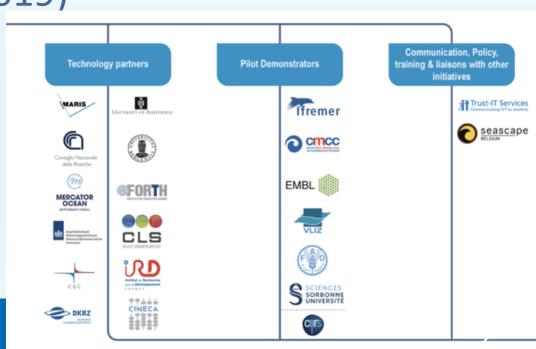






Blue-Cloud in numbers

- Funding: H2020: The 'Future of Seas and Oceans Flagship Initiative' (BG-07-2019-2020) topic: [A] 2019 - Blue Cloud services
- Timing: 36 Months (start October 2019)
- Budget: 5.9 Million Euro
- Partnership: 20 partners
- 13 Blue federated Infrastructures





Overarching project objectives

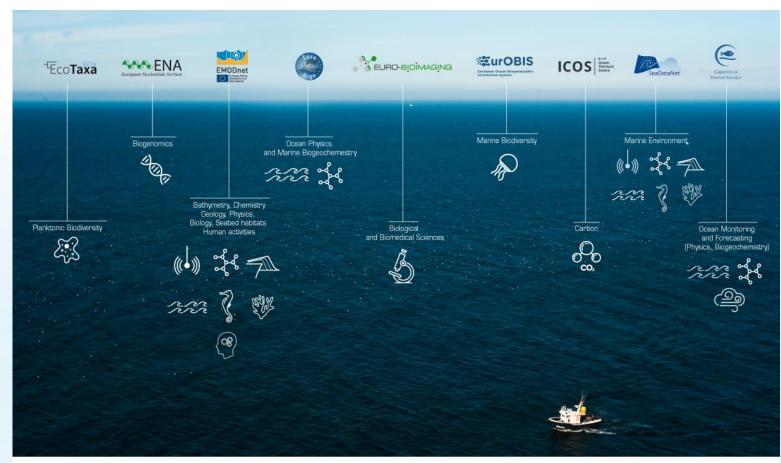
Blue-Cloud aims to become *the* environment for the "Blue community", offering access to **an unprecedented wealth of multi-disciplinary data resources and added-value services** for the benefit of the future marine research and blue economy landscapes



The Blue-Cloud Offer

Blue-Cloud aims at federating leading European marine data management infrastructures with horizontal einfrastructures to create a trusted virtual space - the Blue-Cloud Technical Framework where scientists can access the ocean data, tools, services and research outputs they need to perform research in a more efficient way

https://www.blue-cloud.org/data-infrastructures





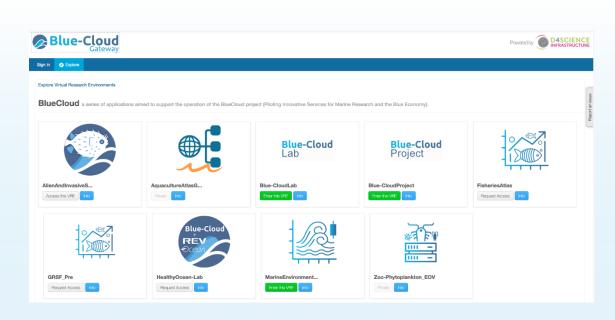








Blue-Cloud Virtual Research Environments



https://blue-cloud.d4science.org/

Each Blue-Cloud VRE includes:

- services that facilitate collaboration between users
- execution of analytical tasks embedded in a distributed computing infrastructure
- services enabling the co-creation of new entire Virtual Laboratories, aimed at realising open science-friendly working environments.

12/02/2021 5



Five real-life demonstrators



Zoo & Phytoplankton EOV products



Plankton Genomics



Marine Environmental Indicators



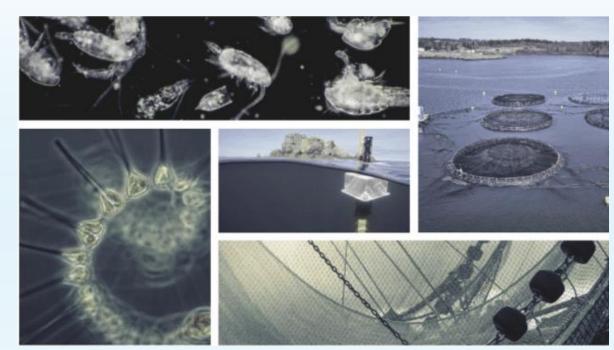
Fish a matter of scales



Aquaculture Monitor

https://www.blue-cloud.org/demonstrators

A number of Blue Cloud Demonstrators are being developed and make use of existing, and drive the further development of additional, Blue Cloud services. They've been originally selected for their importance for marine ecosystems research, conservation, forecasting & innovation in the Blue Economy, not only in Europe but also across the globe





Five real-life demonstrators



Zoo & Phytoplankton EOV products



Plankton Genomics



Marine Environmental Indicators



Fish a matter of scales

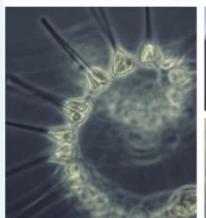


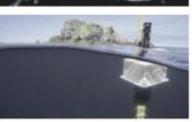
Aquaculture Monitor

https://www.blue-cloud.org/demonstrators

A number of Blue-Cloud Demonstrators are being developed and make use of existing, and drive the further development of additional, Blue Cloud services. They've been originally selected for their importance for marine ecosystems research, conservation, forecasting & innovation in the Blue Economy, not only in Europe but also across the globe













Zoo & Phytoplankton EOV Products



Processing several data resources available under different European marine networks to produce unique zoo- and phytoplankton EOV products.







Thanks to the Blue-Cloud interface, the demonstrator will implement a workflow to apply big data analysis and machine learning (e.g. neural networks) methods on multi-source data sets. The framework will include a ground truth model using near real-time data from LifeWatch data. The resulting workflow will be published as a Blue-Cloud Virtual Lab, and results will be exploited through dissemination in the policy advisory process.



- https://youtu.be/9WwvKT1Gid0
- Beta release in October 2020
- Webinar in February 2021

More info: https://www.blue-cloud.org/news/plankton-climate-change-demonstrator



External Stakeholders Experts Board 14 ESEB Members joined Blue-Cloud

https://www.blue-cloud.org/eseb



Policy Officer at European Commission



Water and Marine group at European Environment Agency



Scientific Officer at European Commission Joint Research Centre



Managing
Director at EGI
Foundation



Executive Director
European Marine Board
and Professor at UHI



Marine Geologist at MARUM at the University Bremen



Head of IT Services at FMBI



Director of IMARDIS



General Secretary of EATiP



Cos4Cloud Coordinator & Chair of the WG "Data, Tools and Technologies" at ECSA



Head of the IOC Project Office for IODE



Managing Partner at Premotec, President of EuroFIR AISBL & member of IMEKO



Chemical oceanographer at GEOMAR



Professor at GEOMAR



The Blue-Cloud Roadmap to 2030

- A co-created, community Vision for the development of the Blue-Cloud beyond project-end and into the future (2030)
- A Strategic Plan towards successfully achieving the Blue-Cloud Vision
- 1st release by March 2021 · 2nd (final) release by July 2022



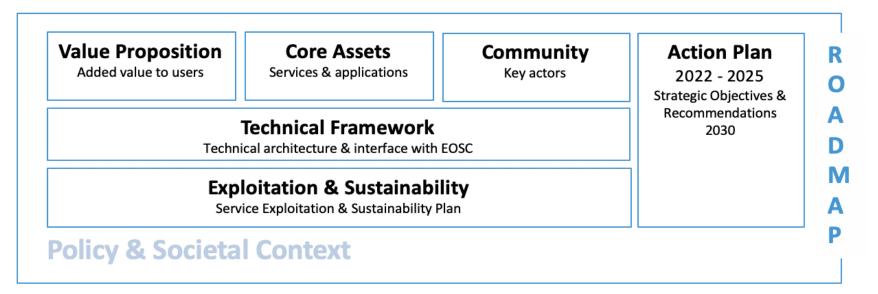
What do we want the Blue-Cloud to achieve?



For whom? By when?



How do we get there?





The Blue-Cloud Roadmap: Process

Public consultation on Blue Cloud Roadmap v1.0 (July-Oct 2021)

Public, open consultation, including targeted dissemination, promotion and dialogue with large pool of identified stakeholders. Tailored dialogue with key communities to ensure awareness, uptake and interaction.

Targeted survey on Blue **Cloud Roadmap Concept** (Sept – Nov 2020)

Targeted survey & dialogue (tailored to specific communities): EU projects (e.g. EuroSea) EOSC, Food thematic cloud; EU policy & decision makers (EC) & National and International government; European initiatives (Marine data infrastructures and e-infrastructures); Global initiatives STEP 5: Consultation analysis, Workshop #2, Review

Roadmap v2.0 & final Roadmap delivered At Final Conference

STEP 4: Roadmap v1.0 & consultation

STEP 3: Survey analysis, Workshop #1

Roadmap v1.0 based on targeted stakeholder feedback

Roadmap v1.0 release based on targeted stakeholder feedback from 6 key communities, Blue-Cloud partners and ESEB.

STEP 2: Concept tailored surveys

STEP 1: Concept Note

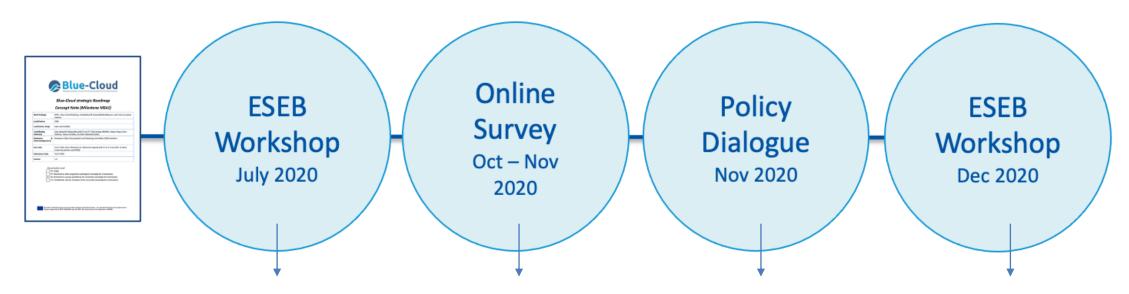
Blue Cloud Roadmap Concept Note (Spring-Summer 2020)

Develop based on review of policy context, EOSC developments & wider landscape, Input from Steering Committee, Blue Cloud Demonstrators, External Science Advisory Board (ESEB) (M6.1; July 2020)

Engaging & consulting the B-C community Phase 1: July-December 2020



BLUE-CLOUD ROADMAP CONSULTATION PROCESS SO FAR



Feedback collection about technical concept and demonstrators Collection of inputs from the marine community, to be used in the development of the "Roadmap to 2030" Start a dialogue on the Blue-Cloud can meet EU policy priorities in marine research, environment, economy, policy making, and decisioning, to feed into the Blue-Cloud Roadmap 2030

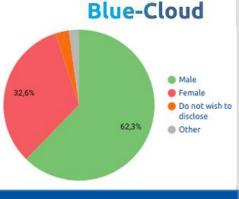
Collect inputs for the initial blueprint for the Blue-Cloud Roadmap to 2030, built on results from consultation

Engaging & consulting the B-C community Online survey Blue-Cloud Roadmap

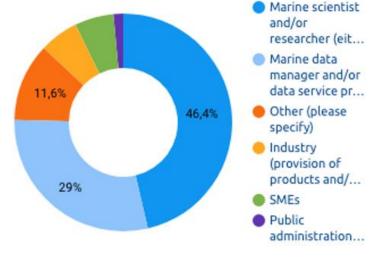
October - November 2020

Profile of respondents

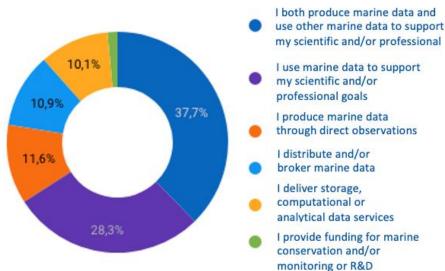
Respondents October 12th to December 7th, 2020



Respondent types



Connection with marine data



Sectors

(please note this was a multiple-choice question)

+ Sector	Quantity	
Academia	51	
Blue economy	8	
Business	12	
Government	27	
Non-governmental organisation	17	
Other	11	





Synergies Overview

Horizontal Contribution to the 2030 Roadmap





Demonstrators

EuroSea

ODYSSEA

Cos4Cloud

FNS - Cloud

Part Author Business

SEE OPEN

Blue-Cloud has built synergies with European and international stakeholders











Examples of concrete synergies



BeOpen is contributing to the exploitation of data from their observatory (maritime and water transports) into the Blue-Cloud demonstrator on marine environmental indicators, thus interlinking different sets of data for several thematic clouds.



EuroSea is interested in the Blue-Cloud demonstrators on Biodiversity (Plankton Genomics), Fisheries and Aquaculture to further exploit FAIR data efforts on the network level for forecasts and observations using data from Copernicus, EOSS and GOOS.



FNS-Cloud and Blue-Cloud are addressing the engagement with users from the Blue-Cloud Fisheries Demonstrator for integrating datasets from different disciplines, such as nutrition.



Odyssea and Blue-Cloud are developing synergies for feeding each other's demonstrators with relevant data, especially for the demonstrators on environmental indicators, fisheries and aquaculture.

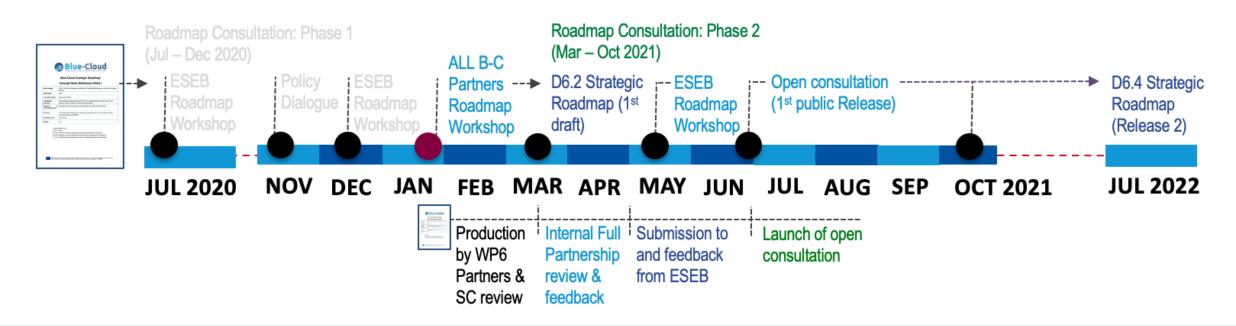


The synergy is aiming to integrate citizen science data and observations of COS4CLOUD into the Blue-Cloud through the Biodiversity and Environmental indicators demonstrators.

Contribution of COS4CLOUD to the Blue-Cloud 2030 Roadmap.



Next steps



- Blue-Cloud Open Workshop (March 2021)
- Blue-Cloud Roadmap Workshop (June 2021)



Thank you! Questions?

Follow Blue-Cloud

Website: www.blue-cloud.org

E-mail: info@blue-cloud.org

Twitter: @BlueCloudEU

LinkedIn: Blue-Cloud Org

Join our newsletter: https://bit.ly/2VdZD95







Big data and machine learning methods to enhance biodiversity data

Renosh PR, Sorbonne Université (SU) – CNRS/Laboratoire d'Océanographie de Villefranche (LOV) **Alexander Barth**, University of Liege – GHER

Gert Everaert & Patricia Cabrera, Flanders Marine Institute (VLIZ)

















Agenda

10:00: Introduction, Blue-Cloud Vision and Mission, by Sarah Pittonet, Trust-IT Services

10:05: Zoo-Phytoplankton EOV demonstrator insights, by Patricia Cabrera, VLIZ

10:10: Deriving global ocean 3D Chlorophyll-a concentrations using machine learning techniques, by Renosh Pannimpullath Remanan, LOV

10:20: Deriving gridded data products for zooplankton distribution using variational analysis and neural networks, by Alex Barth, ULiège

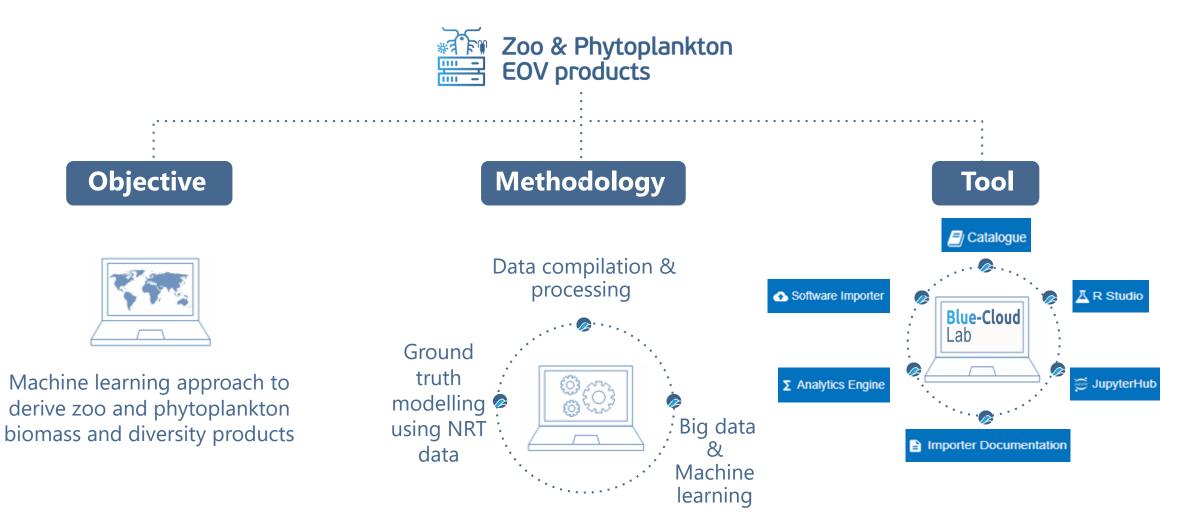
10:30: Modelling zoo- and phytoplankton interactions, by Gert Everaert, VLIZ

10:40: Next steps, by Patricia Cabrera, VLIZ

10:45: Q&A, moderated by Sarah Pittonet, Trust-IT Services

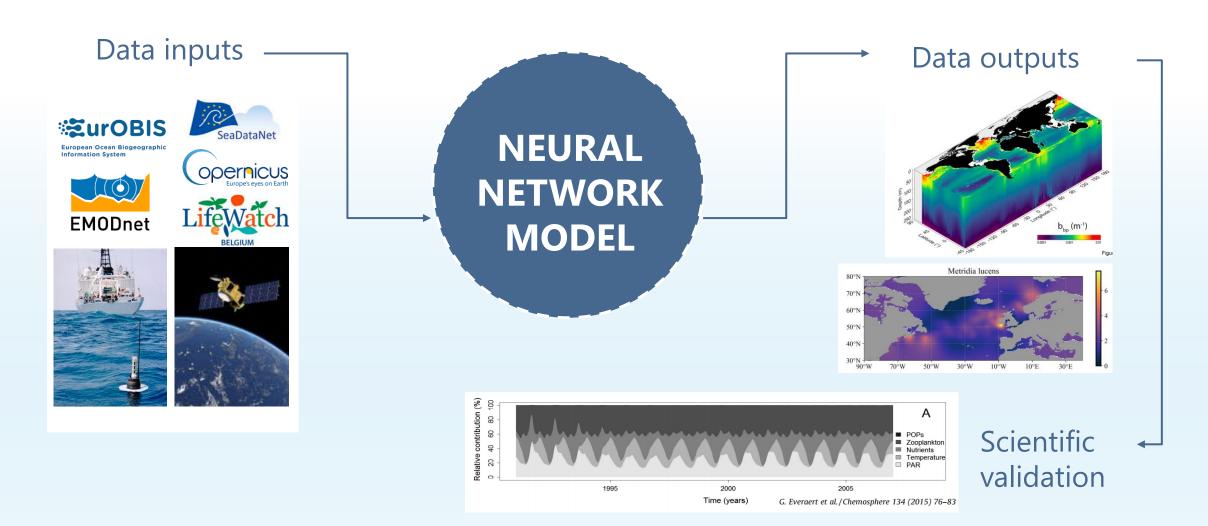


Demonstrator insights





Demonstrator workflow





Phytoplankton EOV products

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



Renosh P. Remanan, renosh.pr@obs-vlfr.fr Raphaëlle Sauzède, raphaelle.sauzede@imev-mer.fr Julia Uitz, julia.uitz@imev-mer.fr Hervé Claustre, claustre@obs-vlfr.fr









Outline

- Introduction
- Objectives
- Methodology
- Data
- Validation
- How to use the demonstrator?
- Conclusions and perspectives



Introduction







- Through their photosynthetic capacity, **phytoplankton organisms form the basis of the** marine food web on which all life in the ocean depends.
- These microalgae are also key to the biological carbon pump, by which the oceans sequester carbon and hence **mitigate the impact of climate change** induced by CO2 anthropogenic emissions.
- Chlorophyll-a (Chla) is the key pigment associated with phytoplankton photosynthesis and is thus widely used as a proxy for the phytoplankton biomass in the ocean.
- Knowledge of global ocean 3D Chla will contribute to improve the proficiency and vastly reduce the uncertainty regarding the present state of marine ecosystems and their response to ongoing and future climate change.



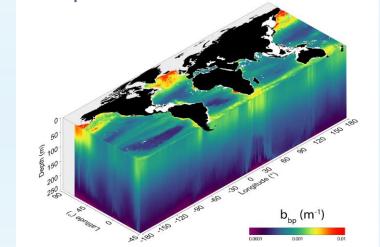
Objectives







- Generate global ocean 3D fields of Chlorophyll-a & Phytoplankton size classes using a machine learning-based method, following the approach of Sauzède et al. (2016).
- This approach takes advantage of both the fine vertical resolution of Argo T/S profiling floats and the synoptic coverage of ocean color satellite imagery in order to extend a surface biogeochemical variable to depth and create a 3D product.
- The beta version of the Phytoplankton-EOV generates 3D fields of monthly average Chla for the year 2018.





Methodology



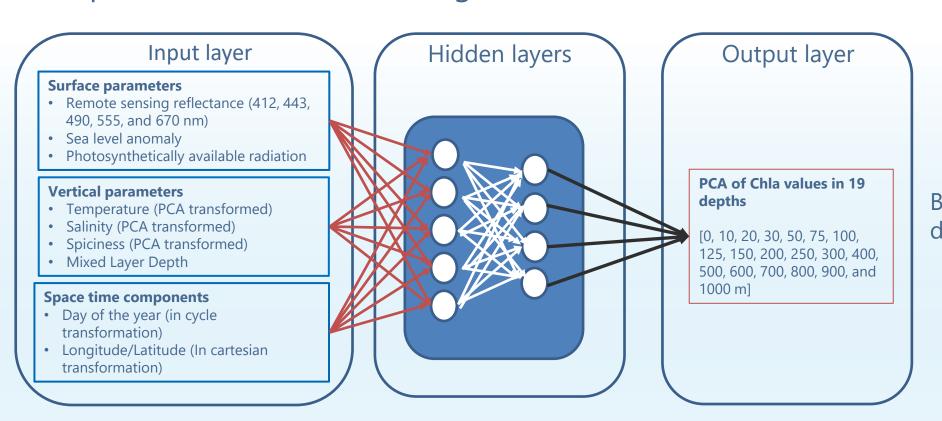




Multi Layer Perceptron (MLP) model training

Satellite data

BGC- Argo data



BGC- Argo data









How to derive global 3D Chla?

Monthly mean global satellite data (CMEMS & Globcolor)

Monthly mean global Physical data (CMEMS)

Input

Surface parameters

- Remote sensing reflectance (412, 443, 490, 555, and 670 nm)
- ❖ Sea level anomaly
- Photosynthetically available radiation

Vertical parameters

- Temperature (PCA transformed)
- Salinity (PCA transformed)
- Spiciness (PCA transformed)
- Mixed Layer Depth

Space time components

- Day of the year (in cycle transformation)
- Longitude/Latitude (In cartesian transformation)

MLP model

Global 3D monthly mean Chla fields (19 depths)

[0, 10, 20, 30, 50, 75, 100, 125, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, and 1000 m]



Data







Variable	Data used for	Source	Data type	Data measured from
BGC-Argo data of Chlorophyll-a, Temperature and Salinity	MLP training & Validation	Agro-GDAC (Coriolis data center)	Freely accessible	Depth-resolved
Satellite derived Remote sensing reflectance (Rrs)		CMEMS	Freely accessible	Surface
Sea Level Anomaly (SLA)	MLP training and generation of	CMEMS (DUACS)	Freely accessible	Surface
Satellite derived Photosynthetically Available Radiation (PAR)	output 3D Chla product	CMEMS / Globcolour	Freely accessible	Surface
Physical data (Salinity and Temperature)	Generation of ouput 3D Chla product	CMEMS (ARMOR3D)	Freely accessible	Depth-resolved
HPLC pigment database	Validation	LOV	Accessible upon request	Depth-resolved



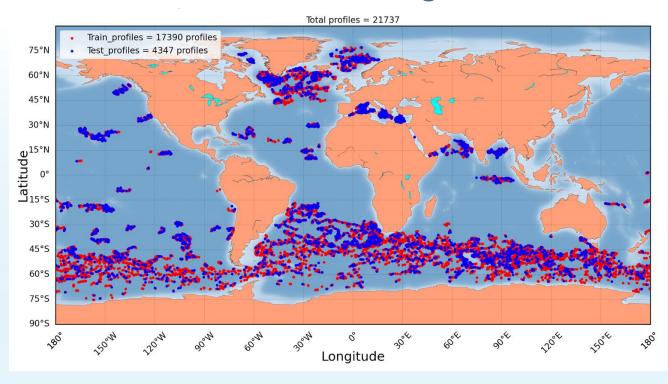
Validation



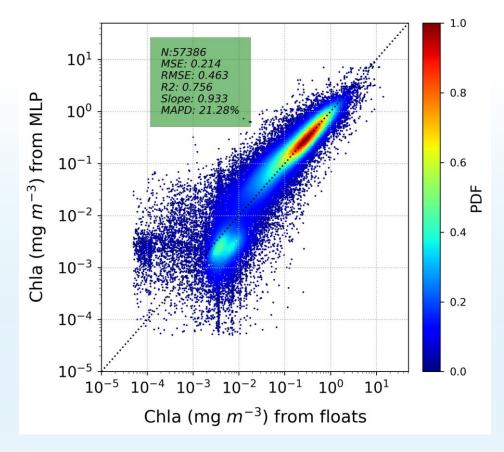




Distribution of the Argo Chla



MLP model validation



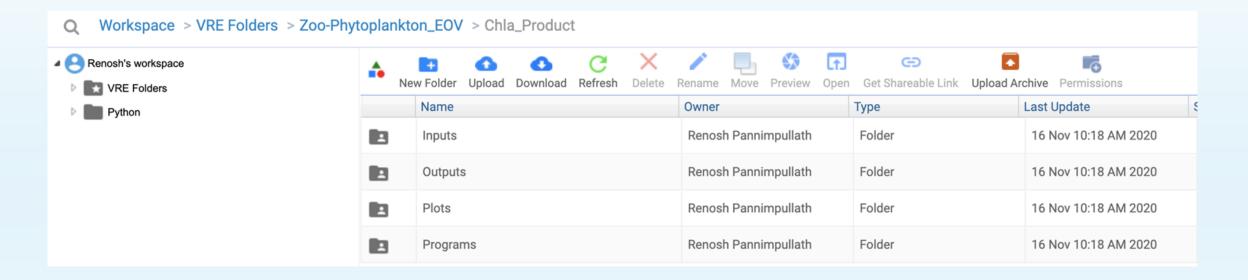


How to use the Demonstrator?







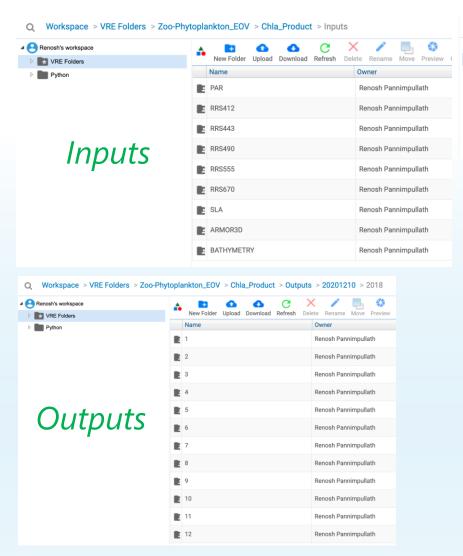


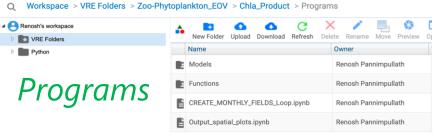


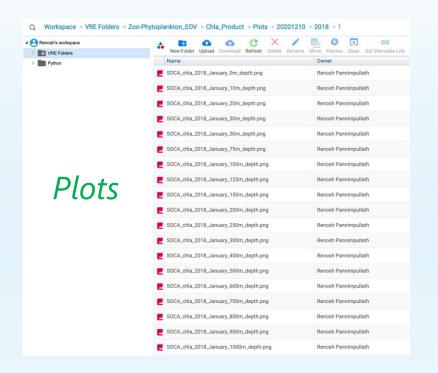














2 Jupyter notebooks









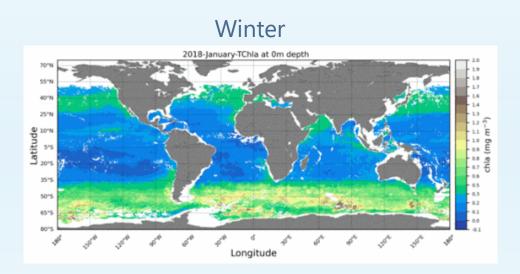


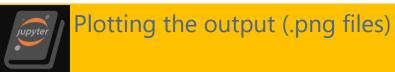
Making output files (netcdf)

Location
Chla_Product -> Programs

Name of the Notebook

CREATE_MONTHLY_FIELDS_Loop.ipynb



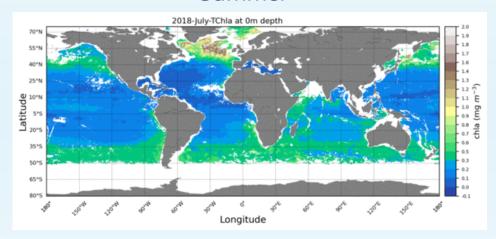


Location
Chla_Product -> Programs

Name of the Notebook

Output_spatial_plots.ipynb

Summer











Conclusions and perspectives

- Users can generate the global 3D monthly mean Chla concentrations for the year 2018 and corresponding plots, using the 2 jupyter-notebooks provided in the "Programs" folder.
- The output ".nc" files will be stored in the "Output" folder and corresponding plots (.png) will be stored in the "Plots" folder.
- If users wish to save these data and plots in different folders, they can edit the path in the notebook.
- In the short term, the MLP model will be refined in order to improve the quality of the retrieved products.
- The Chla concentrations derived using the MLP model will be validated against HPLC reference pigment data.





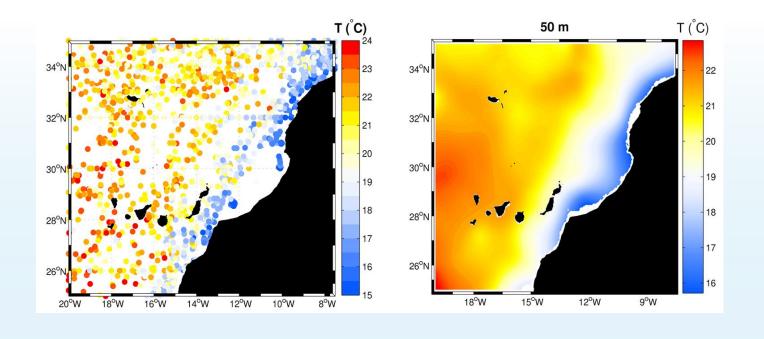
Interpolation and data analysis



Difficulties in oceanography:

- Data distribution/scarcity
- Confidence in measurements
- Representativity error

Possible usage: climatologies, reference state, model validation

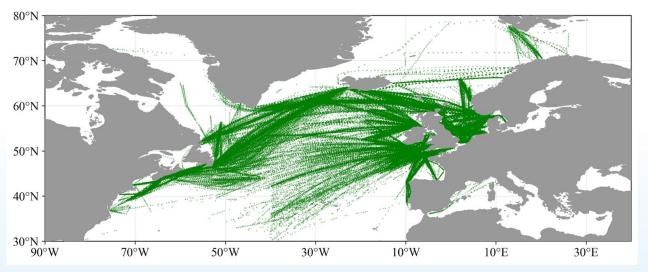


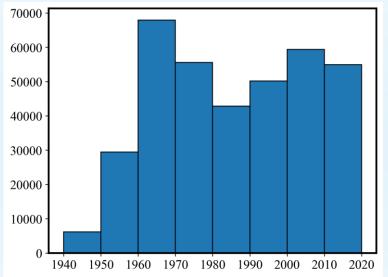


The dataset



- Data from the <u>continuous plankton</u> <u>recorder operated by the Marine</u> Biological Association, UK
- Heterogeneous distribution
- Data since 1940
- Data | position | number of ind/m³







The method



DIVAnd: Data Interpolating Variational Analysis in n dimensions

https://github.com/gher-ulg/DIVAnd.jl

Code written in Julia (https://julialang.org/)

DIVAnd



DIVAND (Data-Interpolating Variational Analysis in n dimensions) performs an n-dimensional variational analysis/gridding of arbitrarily located observations. Observations will be interpolated/analyzed on a curvilinear grid in 1, 2, 3 or more dimensions. In this sense it is a generalization of the original two-dimensional DIVA version (still available here https://github.com/gher-ulg/DIVA but not further developed anymore).

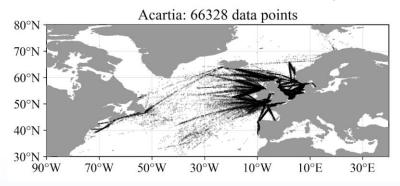


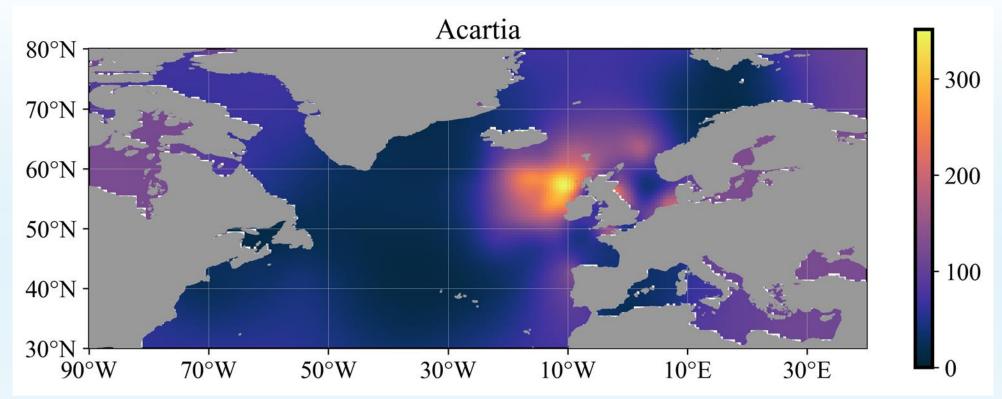
DIVAnd Application





Spatial interpolation







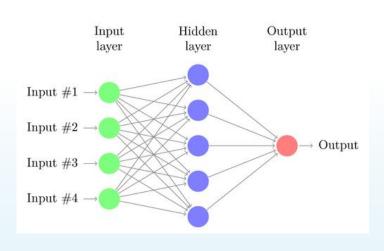




DIVAnd with a Neural Network

$$g(x) = f(v_1, v_2, ..., v_n) + x'$$

- The beta version of the Phytoplankton-EOV generates 3D fields of monthly average Chla for the year 2018.
- $v_1, v_2, ... v_n$ are environmental variables related to the probability
- $f(v_1, v_2, ..., v_n)$ is a neural network
- x' is a residue varying of a specified length-scale (DIVAnd)
- Also applicable to occurrence probability with a slightly different formulation



Neural network (Multilayer perceptron)

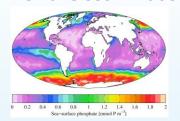


Blue-cloud Neural network technique

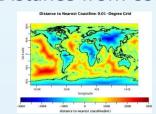


SeaDataCloud
T/S climatology

World Ocean Atlas



Distance from coast





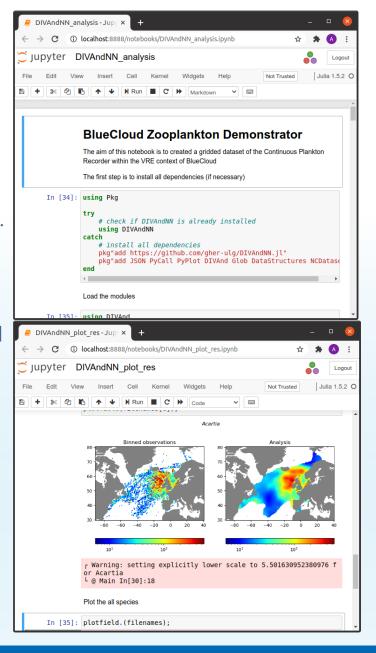
Using other sources of information to improve the interpolation. Used "Co-variables":

- Sea water temperature (SeaDataCloud)
- Salinity (SeaDataCloud)
- Distance from coast (NASA Goddard Space Flight Center)
- Bathymetry (GEBCO)
- Nitrate, Silicate and Phosphate (World Ocean Atlas 2018)



Approach

- Reproducible approach:
 - Version control using GitLab (at ULiège).
 - Continuous integration (testing that the code still runs after every pushed commit).
 - Declare all necessary direct software dependencies.
 - Full snapshot of the dependency tree (direct and indirect dependencies).
 - Projects can be easily instantiated (i.e. the exact Julia environment can be replicated with a single command).
- Using two Jupyter notebooks
 - Analysis
 - Preparation of covariables.
 - Split data into training and validation dataset.
 - DIVAnd + neural network analysis.
 - Visualization of the results.

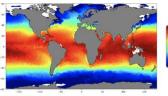




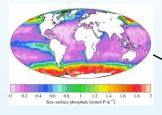
Overview



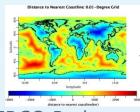




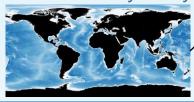
World Ocean Atlas (nutrients)

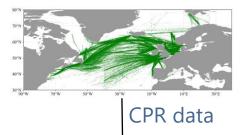


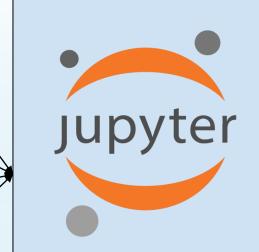
Distance from coast



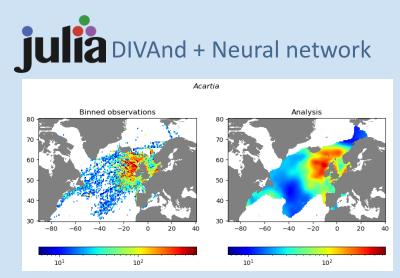
GEBCO Bathymetry









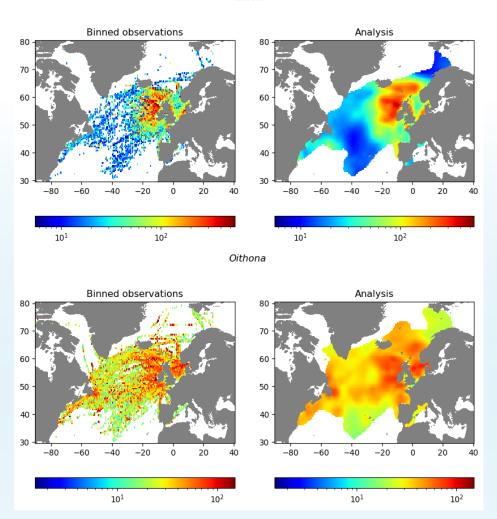




Some example results



Acartia



- Inhomogeneous sampling -> inhomogeneous error spatially
- A relative error map is derived allowing to mask the value far away from the observations







- Collaborative virtual research environment have a large potential to boost scientific productivity
- Capture the relationship between specie distribution and environment parameters using a neural network
- Spatial (and temporal) coherence: constraints used in variational inverse methods
- Possibility to use irregularly sampled observations
- Perspective: explore the potential of convolutional networks (e.g. <u>DINCAE</u>, <u>Data INterpolating Convolutional Auto-Encoder</u>)



Scientific validation

Modelling zoo- & phytoplankton interactions



Gert Everaert, gert.everaert@vliz.be Viviana Otero, viviana.otero@vliz.be







- Marine phytoplankton fuels marine food webs
- Drivers that limit phytoplankton abundance:
 - Temperature
 - Nutrients
 - Light
 - Zooplankton grazing
- Changes in space and time



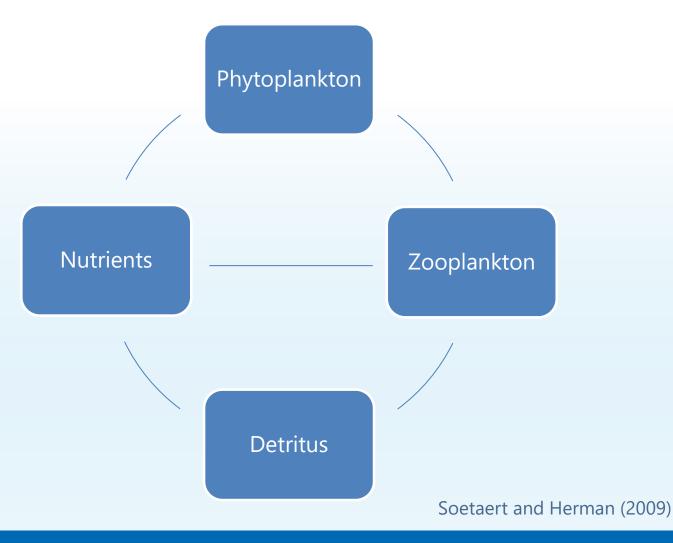
© ESA



Ecosystem model



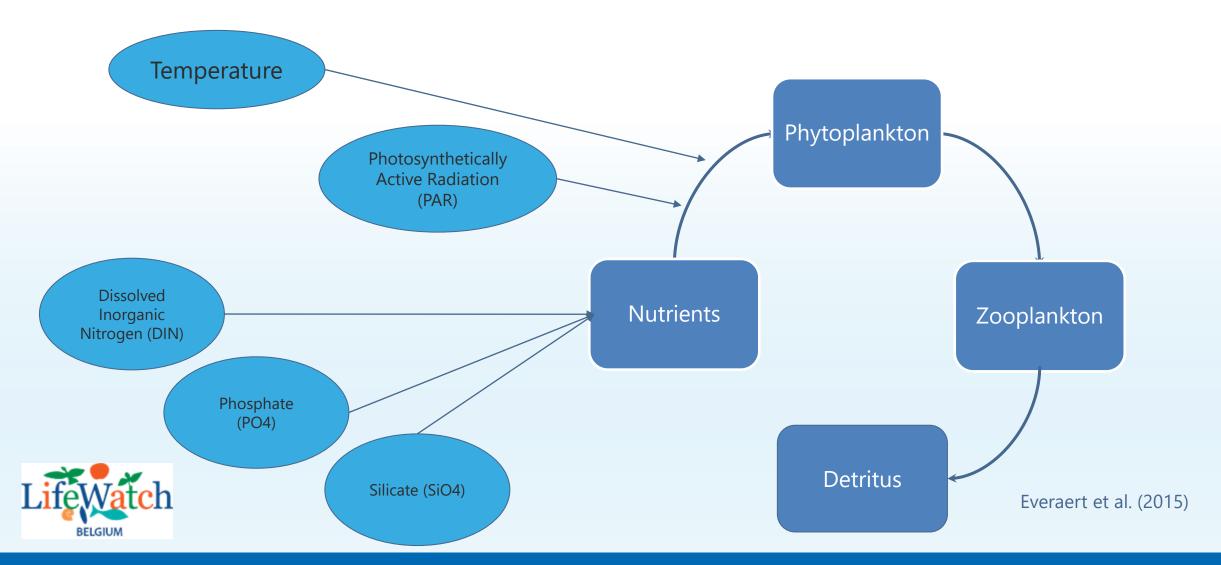
Nutrients, Phytoplankton, Zooplankton and Detritus model for aquatic environments NPZD Model







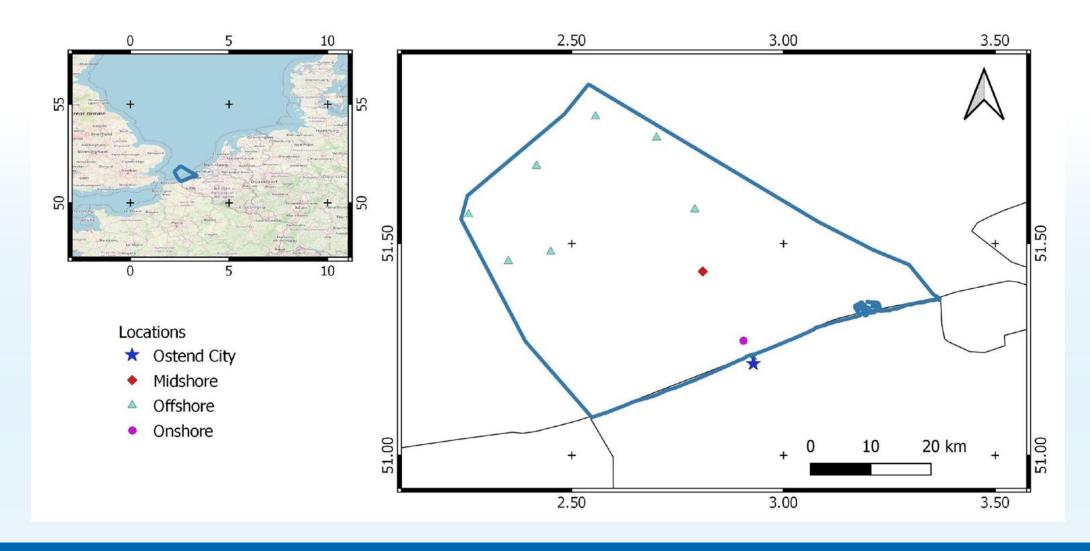






Study area

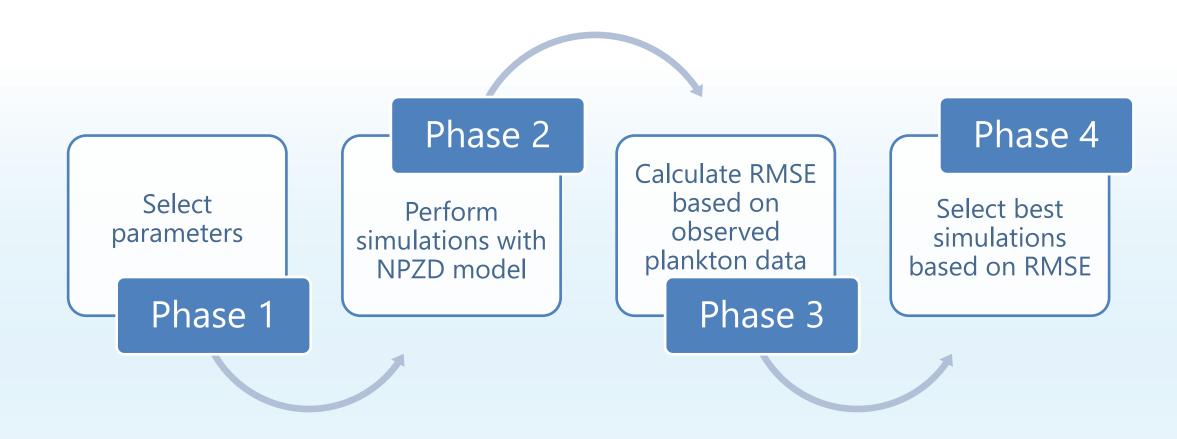






Model calibration and validation

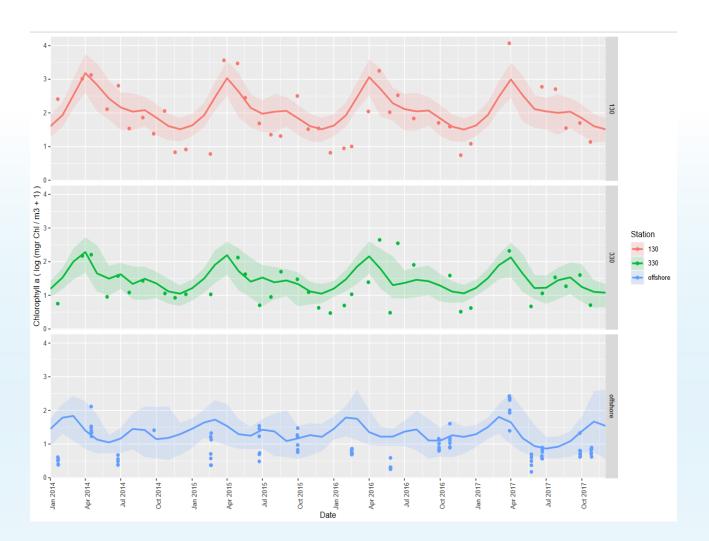


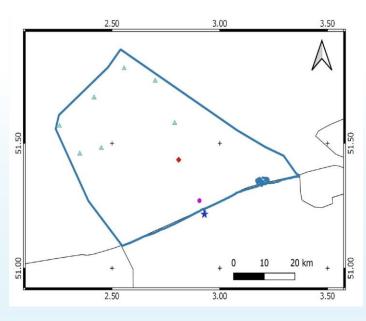




Phytoplankton abundances







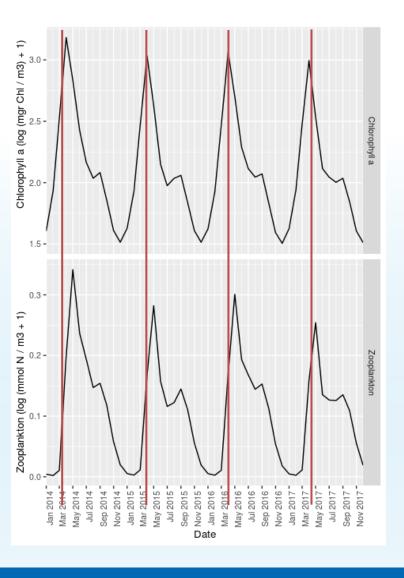


Otero et al., in prep.



Phyto and zooplankton interactions





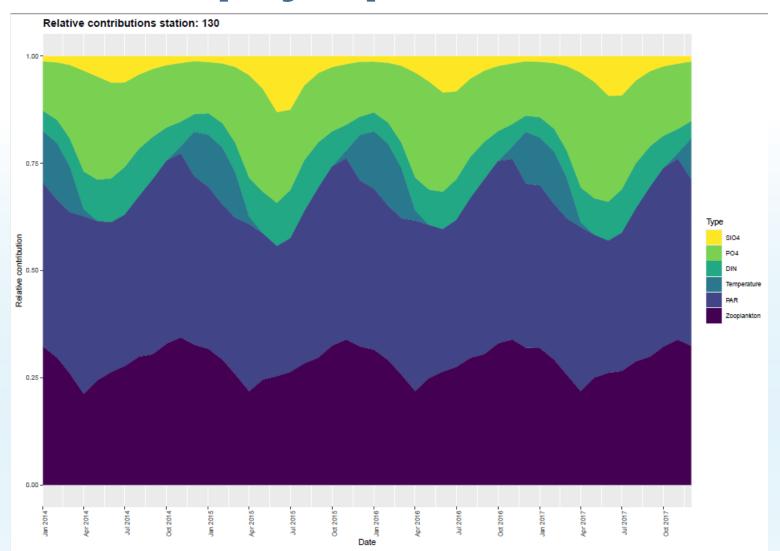


Otero et al., in prep.









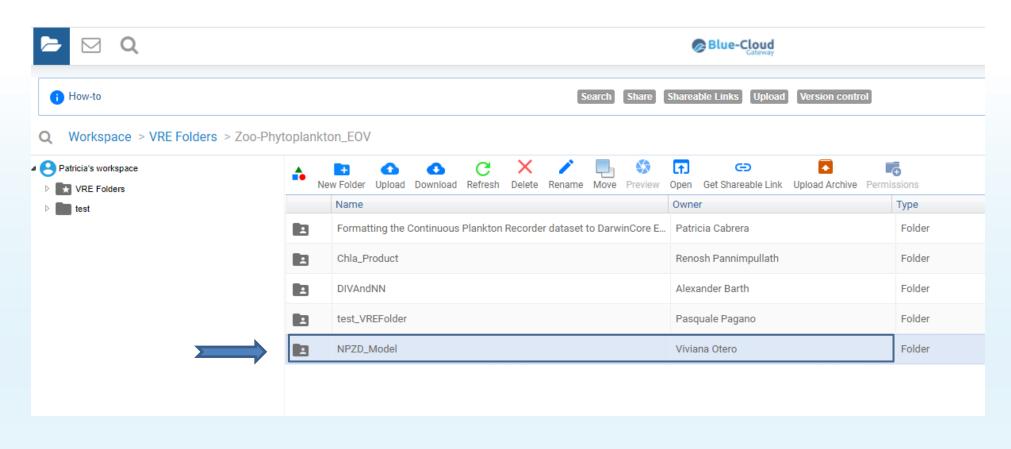


Otero et al., in prep.



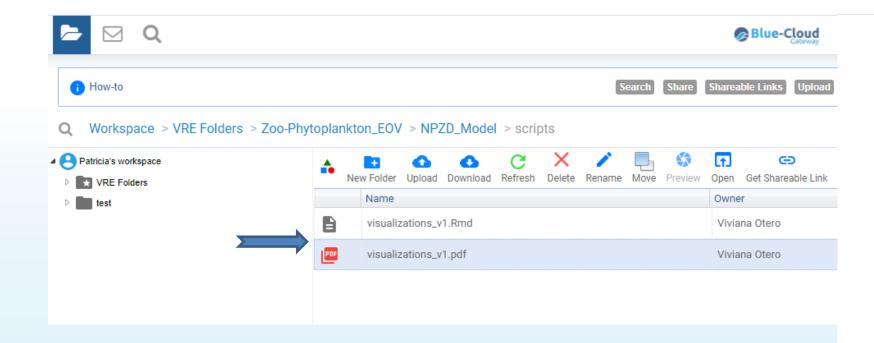












Modelling phyto- and zoo-plantkon interactions

Viviana Otero

2021-01-18

Contents

Co	ntext																					
	Modelling approach							 					 								 	
	Input and validation																					
	Input data																					
	Visualizing the data																					

Context

Marine phytoplankton primary production is the base of the marine food web and regulates functions in coastal ecosystems. Understanding how primary production changes through time and space is of key importance to better quantify the effects of human impact on the ocean.

With the methodology presented in this document, it is possible to analyse which factors drive the phytoplankton abundance and how these factors change in space and time. In this document, we are focused on the Belgian part of the North Sea, therefore the parametrization and visualizations shown here correspond to this particular area. For other areas, this document can be used as a guideline to adjust the relevant variables and data sources to obtain similar analyses.

Modelling approach

The ecosystem model for Nutrient, Phytoplankton and Zooplankton (NPZ) was used to simulate changes in plankton density from 2014 to 2017 (Soetaert and Herman, 2009). This model describes daily changes in phyto and zoo-plankton density based on abiotic parameters (Figure 1). The variables in the model are expressed in mmol N m-3 for nutrient, phytoplankton and zooplankton densities. Daily changes in these variables are expressed in mmol N m-3 d-1.

This model is useful to describe marine and freshwater systems. The state variables are nutrients, phytoplankton and zooplankton. In this document, we focused on a marine system, as in Everaert et al. (2015). Nutrients are defined as the total density of Dissolved Inorganic Nitrogen (DIN), Phosphate (PO4) and Silicate (SiO4). DIN is defined as the sum of NH4, NO3 and NO2.



Conclusions



- Integration of different EOV variables, not only to see data-driven trend, but also to understand interactions in a mechanistic way.
- The myriad of inputs are streamlined stepwise towards one common denominator.
- Data and scripts are available and annotated in a RMarkdown document.
- Marine systems are under multiple pressures, but which one is most important and how do these change over time and in space?
- The novelty relates to the quantification of the relative importance of each driver, and this is of interest for the Blue Community.



NEXT STEPS

Phytoplankton EOV

- Refinement of Chla product
- Development of global 3D phytoplankton community product

Zooplankton EOV

Final test on Blue-Cloud VRE of DIVAnd + Neural Network

Scientific modelling

- Application of the model at other regional seas
- Validation of the model with zoo and phytoplankton products



Survey







Thank you



Renosh P. Remanan, <u>renosh.pr@obs-vlfr.fr</u>
Julia Uitz, <u>julia.uitz@imev-mer.fr</u>
Raphaëlle Sauzède, <u>raphaelle.sauzede@imev-mer.fr</u>
Hervé Claustre, <u>claustre@obs-vlfr.fr</u>



Alexander Barth, <u>a.barth@uliege.be</u> Charles Troupin, <u>ctroupin@uliege.be</u>



Viviana Otero, <u>viviana.otero@vliz.be</u>
Gert Everaert, <u>gert.everaert@vliz.be</u>
Patricia Cabrera, <u>patricia.cabrera@vliz.be</u>
Lennert Schepers, <u>lennert.schepers@vliz.be</u>