SCOR WG #168: 4D-BGC
Year 1, Meeting 1 (M1.1)
February 18, 2024 - New Orleans, LA
Meeting Minutes

A. Roundtable introductions (1:00 - 1:22)

In person:
J. Sharp, R. Sauzède, L. Delaigue (guest), A. Poteau (guest), P.R. Renosh (guest), H. Bittig, T. Fujiki, S. Lauvset, H. Garcia (guest), T. Ehmen (guest), N. Mackay (guest), K. Fennel, L. Keppler, J. Muller, K. Turner, T. Ito, A. Watson, E. Twigg, P. Chamberlain (guest)

Virtual:
P. Oke, R. Kerr, W. Cheah, L. Djeutchouang, A. Rochner

B. SCOR introduction from Emily Twigg (1:22 - 1:37)

Emily presented slides about SCOR as an organization, activities, working group coordination, and capacity development activities

C. Introduction on the SCOR working group from Jon Sharp (1:37 - 1:52)

Jon presented slides about how the group came together, its motivation, terms of reference, and planned timeline of activities

Discussion (1:52 - 1:58):
Katja: this WG should think about how to ensure that products that are developed are transitioned to operational production and updated regularly, so they don't go defunct after their first release (+ update)
Raphaëlle: Copernicus is an option for this type of operational production
Jens: are the products we're talking about exclusively BGC-Argo-focused?
Jon: the group was motivated by BGC-Argo, however we are interested in considering all observational datasets that are available

D. Introduction to existing observation-based products from Raphaëlle Sauzède (1:58 - 2:10)

Raphaëlle presented slides reviewing different types of data products by spatial and temporal resolution, and provided examples of existing data products

Discussion (2:10 - 2:19):
Siv: How can we ensure that 3D products that are developed are coherent with 2D products?
Jon: could be a recommendation for interior carbon products that they are compared against surface products during development
Lydi, Henry, and Katherine: discussed
Raphaëlle: bio-optical 3D products can be compared with satellite products at surface, and can improve upon those surface products. Uncertainties have to be coherent also

E. Lightning presentations (2:20 - 3:34):

Carbon:
Louise Delaigue: SOCA-CO2 - a 4D product of ocean carbon, built with GLODAP, floats, and neural networks
Tobias Ehmen: Unicorns project - uses deep learning trained on GLODAP to reconstruct ocean DIC and C*
Lydia Keppler: MOBO-DIC - adaptation of SOM-FFN to reconstruct ocean interior carbon from GLODAP data
Jens Müller: Comparing decadal trends in the anthropogenic carbon storage from 1994-2004 and 2004-2014 and inferring interior acidification rates
Katherine Turner: using covariance fields from CMIP6 to reconstruct ocean carbon from simulated Argo temperature and salinity observations within model world

Oxygen:
Jon Sharp: GOBAI-O2 - a 4D dissolved oxygen product from machine learning and Argo and GLODAP observations

- Hernan Garcia: (during lightning presentation of Jon): How do we combine data with different uncertainties?
- Jon: quantify uncertainty with ensemble maps

Takamitsu Ito: using ensemble maps of reconstructed oxygen to quantify uncertainty
Hernan Garcia: New release of World Ocean Atlas includes oxygen climatology including Argo data for the first time

Optics:
Xiaogang Xing: (1) work in data QC: slope factor, non-photochemical quenching, etc. (2) work in merging dissimilar datasets, in particular for float oxygen from different DACs and against discrete matchups
Raphaëlle Sauzède: SOCA-PFT - a reconstruction of phytoplankton community composition using machine learning on fluorescence profiles to assign community composition and SOCA-type neural networks to map
P.R. Renosh: SOCA-light - a machine learning model to predict and map vertical light profiles from satellite observations

Modelling:

Andrea Rochner: CO2 flux drivers from ESMs, and the effect of the assimilation of different datasets on CO2 flux. 4D-BGC benefits: (1) boundary conditions, (2) validation, (3) new knowledge about links between ocean physics and biogeochemistry

F. Break (3:34 - 3:56)

G. Discussion (3:56 - 5:00)

Brief introductions of WG members who were not present

Why all converge to neural networks (question from Katherine)
Historical (NN used for pCO2 surface products from Peter and also a lot used in satellite ocean color), ensemble of ML is more and more popular
RF more used in land (high frequency data)

Consistency in methods would help?
Would it be possible to make recommendations on standard interpolation techniques to use
Different techniques give in some case different results

It would really help to give recommendations among the different variables on how to validate the method (e.g. train/ test data).

It would help if everybody use all the same test data to validate their methods → validation intercomparable)

Delete all an area can be complicated as we don't want to remove data that contain information (point made by Katja)

Make available synthetic datasets for intercomparison

The small scale variability is not easy to have with coarse models → use eddy resolved hindcast?

If it fails it is ok and it responds to the question but if works is necessary it is not sufficient
Synthetic datasets are used in the framework of the Global Carbon Budget

ESM has been used from Lyddi + Jon + Tobias
It would be interesting to share some codes to be coherent in the way of interpolating data in vertical (it is different if it is low resolution in situ data or high resolution data from Argo)

Siv made the point that Glodap is a really tiny dataset and that people should think about using and gathering other datasets for variables such as nutrients.

→ recommendation of gathering such databases?

Which is the impact of the temperature products that we use as inputs of our models? We should check and answer this question.

We should be very transparent on what products we use