GOBAI-O2 regionally-tuned predictors and more diverse ensembles of ML algorithms should lead to increased confidence in estimates of ocean interior [O2]

Authors: Jonathan D. Sharp, Andrea J. Fassbender, Brendan R. Carter, Gregory C. Johnson, Cristina Schultz, John P. Dunne

General comments:

The objective of the manuscript is to present the GOBAI-O2 tool, a 4D gridded product of O2 concentrations in the global ocean. It is based on machine learning algorithms trained on observations from BGC-ARGO and GO-SHIP in 7 regions and applied to temperature and salinity fields constructed from the Argo network. This product allows a fairly fine prediction of O2 concentrations from 2004-2021 on 58 vertical levels with a spatial resolution of 1°x1° allowing an analysis of spatial variability, seasonal cycles and decadal trends in O2.

The article is well constructed and written. The authors clearly present the methodology, and the prediction uncertainties. The authors indicate that GOBAI-O2 provides homogeneous O2 coverage improving O2 observations where spatial and temporal gaps are present in some regions.

The authors mention at the end the limitations of the product but they do not specify the added value of GOBAI-O2 compared to the existing observation networks. For example, it would be interesting to compare the GOBAI-O2 contribution vs. the ARGO-O2 network (with and without GO-SHIP). What is the real contribution of GOBAI-O2 ?

In BGC-ARGO, few O2 data have been qualified properly and adjusted in delayed mode even if a strong global efforts is and will done by the different GDAC. In this context, the authors do not precise how many O2 profiles from Argo network exists and how many have been used for the training ? What is the ratio total vs. qualified ? Probably the efforts will lead to more usable ARGO O2 profiles and thus contribute significantly to the overall O2 content coverage. In this case it will be interesting to know the added value of GOBAI-O2 predictions (metric comparison of the two approaches)

Another use of GOBAI-O2 not mentioned by the authors would be the use of GOBAI-O2 predictions to generate quality time series in areas poorly covered by reference data (long time series) which would allow for a finer qualification of O2 measurements from different platforms and often sensitive to drift over time. This product would be much better than the fields from WOA2018.

Also GOBAI-O2 has been trained from the Winkler O2 data of GO-SHIP but it would have been interesting to start from the O2 profiles from the ship's CTD and adjusted via the Winkler data. The vertical resolution would then be significantly improved. What are the limitations? Access to adjusted O2 profiles? If so, the document should mention and alert to this crucial point. It is now becoming essential to follow the FAIR data principles for all platforms.

The authors also mention the lack of other platforms to improve predictions, but this concerns in particular fixed moorings, which would be a plus in certain regions to increase the temporal resolution of observations (from minutes to months) over the entire water column, but only if a mooring array is available, otherwise a fixed point will not be significant and will not bring much. Also, the contribution of gliders sections will be relevant if we are interested in coast-open sea exchanges because most of the gliders are deployed in these specific sub-regions and their integration in the learning methods will not necessarily bring much.

Specific comments:

- a diagram explaining the principle of FNN and RFR would help readers understand the different algorithms used in this paper
- Table 2: Units of O_2 is missing
- Figure 7: The O_2 anomaly over depth (panel D) is close to zero between 2010-2015. Why? This is because GOBAI-O2 is centered on the year 2012? In this case, explain why it is centered on 2012.
- Figure 8: Units of O_2 is missing. O_2 uncertainties are higher near the equator and subtropical zones. Explain why