

## ***Interactive comment on “Dissolved Inorganic Nutrients in the Western Mediterranean Sea (2004–2017)” by Malek Belgacem et al.***

**Toste Tanhua (Referee)**

ttanhua@geomar.de

Received and published: 19 November 2019

This paper takes a systematic approach, using proven methods, to create an internally consistent data product of inorganic nutrient observations in the Western Mediterranean Sea. The work addresses a major issue with marine inorganic nutrient observations, biases in the data. This is a solid work and deserves to be published. However I have a few major issues with the work that I would like to see addressed, and a few minor comments.

The data product is available in a repository and can be easily downloaded. The same goes for the unadjusted data, all collected in a single file. This certainly has a lot of value. It seems to be a two-step procedure with data flagged as questionable (i.e. flag

C1

“3”) removed from the adjusted product. For cruises where the whole cruise was considered as of “poor quality” (as assessed from excessive scatter etc.) are still included in the product, but flagged as questionable. Why include questionable cruises in the product at all? I appreciate having access to the original data (i.e. prior to adjustment), but that does not preclude the need to link to the individual cruise files. These can be in a common format on a dedicated place, or it could be links to the original data file in a repository (NODC, SeaDataNet, or similar). That has value since for instance some of these cruises probably have associated “other” data, such as oxygen etc. that might be of use for the user. I recommend to establish links to the original data files.

This last comment does also go for the meta-data of the cruises. I guess in most cases this would include reference to a cruise report. I could not find any such references, please add links to cruise reports.

For the secondary QC, the authors choose to adjust all data to 5 reference cruises that was considered to have particularly high quality. One of the reasons was the well-known issue with bias in nutrient measurements being introduced by freezing of samples and analyzing them on land post-cruise. However, not all reference cruises had nutrient measured on-board. Why then include them as reference cruises?

Although it seems that the low-nutrient water of the Mediterranean might be less prone to bias due to freezing, the result from this study seem to suggest something different with all three variables being adjusted preferentially upward or downward. That might be an interesting result. Or maybe this is a function of bias in the measurements?

It would be useful to have a directory of crossover plots for all cruises. The method of GLODAP and CARINA could be taken as an example, but a repository on the web where the crossoverplots can be downloaded would go a long way. This would allow users to judge the validity of the adjustments.

Why only discuss a selection of cruises in section 5.4 ? All cruises had adjustments. I recommend to expand this section to cover all cruises.

C2

Minor comments: Line 67: I suggest changing “profiles” for “observations”

Line 116: The CARIED data product is not yet published and available.

Line 115: Please refer to the GO-SHIP nutrient manual.

Line 188: I am not sure if this is a useful metric. The authors discuss the influence having observations in different sub-basins have on this statistic later. Why not create statistic that is for sub-basin by sub-basin?

Line 221: The 2° influence radius is probably fine for the Atlantic Ocean, but mostly not for the Mediterranean Sea. How did the author handle cross-overs that were influenced by observations from nearby other sub-basins where a different nutrient concentration could be expected?

Line 226: If you know that the deep water is (potentially) changing fast, why include it in the crossover analysis? Would it work to have a crossover analysis covering, for instance, 1000 – 2000 meter only? If so, why was that not used, and how did the authors remove temporal natural variability of deep water?

Line 260: Here you decide not to include cruises that could not be adjusted in the product. On the other hand, you do include data that had only questionable data in the product (although flagged as such). Why? An alternative approach could be to include the data with a flag that indicate that the data did not undergo 2nd QC.

Section 6, and possibly elsewhere: GLODAP and CARINA are data products, rather than datasets. The difference being that the products have an additional layer of QC (2nd QC bias adjustment) applied, whereas a data set is a collection of data that are in its original form, possibly with consistent primary QC, unit conversion etc. Not so important perhaps, but a little of semantic difference.

Line 427: Not a complete list of authors for this paper.

Table 2: Why have a different format for this table compared to table 1?

C3

Table 4: It would be useful to include the reference cruises in this table

Table 5: not a big deal, but the “\*” sign in this table is applied in column 2, whereas in other tables (4) it is applied to column 1.

---

Interactive comment on Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2019-136>, 2019.

C4