Interactive comment on "Global CRM/RM-scaled nutrient gridded dataset GND13" *by* Michio Aoyama

Anonymous Referee #1

Received and published: 5 September 2019

The author outlines a strategy by which he created a gridded nutrient data product from numerous cruise datasets and data products. He presents some of the metrics used to compare various cruises to one another and then outlines methods he used to turn the merged and adjusted data product into a global gridded data product.

This effort is probably worthwhile, but both ahead of its time and behind its time in some ways.

It is ahead of its time because, unfortunately, nutrient CRMs have not yet been used on enough cruises to afford a global reference data set. There are enough measurements in the North Pacific to justify this exercise, but it is not clear that the same is true in, for example, the Atlantic. Quoting from the paper:

"In the Atlantic Ocean, five cruises were also selected as category 1 because RM were used on two of the five cruises, and good tracking standards with excellent quality control were used on the other three cruises."

It is not clear what "good tracking standards with excellent quality control were used" means, and this seems a weak basis on which to base a data product (or at least a basis that is no better than that used by GLODAPv2). Perhaps this could be reworked to be justified based on deep comparisons between the cruises that did have reference materials and those that category 1 cruises that did not? However, one must select some kind of basis for making a merged and internally consistent data product, so arbitrarily selecting a few cruises and calling them reference lines might also be okay but the language used in the descriptive paper should be more clear that this is what was done.

The paper is behind its time because it follows the release of the GLODAPv2 data product (and its recent 2019 update) which does a similar task and gets similar results with essentially the same data. There are some major differences between GLODAPv2 and this data product: 1. The GLODAPv2 data product update has more and more recent data (this might be a mistaken impression on my part). 2. The GLODAPv2 data product process is more meticulous for all properties excepting perhaps nutrients. The following phrase from the paper suggests very little attention was paid to, for example, oxygen, which is critical co-located data for using nutrient distributions.

"For oxygen data, the factors for 30 cruises were assumed to be 1.00 because the high quality control for nutrient analyses on those 30 cruises suggested that the oxygen analyses were also of high quality."

3. GLODAPv2 has more co-located data types. 4. GLODAPv2 does not make adjustments that are smaller than certain threshold values or adjustments in certain variable regions. This paper suggests this is a flaw, but I would point out that this is done deliberately to avoid erasing any potentially interesting signals in the deep ocean. This nu-

trient data product assumes no changes at various depths excepting those measured on category 1 cruises. 5. The gridded GLODAPv2 data product presents more detailed gridding methods and better characterizes gridding uncertainties. 6. The nutrient data in this new data product is traceable to CRMs. So, in most respects, the GLODAPv2 product/gridded product and its presentation simply seems to be better, excepting item 6. Item 6 is a very important idea, however, so this ESSD effort could still be very much worthwhile as an exploration of how large of an impact adding traceability would have on nutrient distributions.

Broadly, I think the best thing the author could do would be to work within the GLODAPv2 data product and concentrate the analysis on proposing adjustments to the cruises therein (and any additional cruises newly added) to bring the GLODAPv2 nutrient data in line with CRM-validated sections. This would also be a much more useful exercise for establishing how this process should be done from future data products when more CRM-validated sections are available, and would mean the new data product could benefit from all of the additional co-located data in the GLODAPv2 product.

If the author is not interested in such a significant revision to make the data product more broadly useful, then a much smaller recommendation would be to spend more time and text motivating and justifying the paper. If the author better explains why traceability is critical (which it is for some applications), why oxygen is included in the analysis and why category 1 oxygen cruises were identified in the way they were, how the gridded fields differ from GLODAPv2 gridded fields, etc., then the paper would be nearly publishable. A critical question is how is this data product better than GLODAPv2 (traceability for nutrients, and perhaps there is more data?), since this is what the readers of this paper are going to be wondering. Alternately, if this data product has already been used for several studies and this ESSD paper is just meant to describe how it was created, then a more thorough presentation of what products have used this data product would be useful.

As the paper is currently written, I have a hard time seeing that there would be a large

user-base for this data product... but I could be wrong and there might be many people interested in using it. An advantage of this being a discussion is that they could write in to correct my error if so.

There are also some bits of unclear language to clean up. This should be done via internal review, so I haven't made an exhaustive list of language suggestions.

Line by line comments: Line 10: nutrient → "nutrients" or "nutrient concentrations" Line 12: what is meant by "comparability between stations was ensured" Line 12: collected from which source(s)? Line 14: Suggested rephrasing: Cruises that used certified reference materials (CRMs) for seawater nutrient concentration measurements were used as reference sections to... Line 15: What is meant by similar protocols? Are there O2 reference materials? Line 23: suggestion: "upper and lower" or "shallow and deep" Line 24: suggestion: delete "and from geographically similar ocean waters"... also delete "reliably"... the word reliably is covered by "with complete confidence." Line 25: delete "accepted" since it is redundant with "certified" and the idea, later in the sentence, that people are expected to use the CRMs. P2/L3: "earth" -> "Earth" P2:L9: biases... among -> consistent disagreements... between P2:L15-20: This text seems to imply that deep ocean nutrient changes would be expected if we had more reproducible measurements. However, this paper has not yet presented any literature suggesting that we would expect there to be these changes. I'd recommend adding that literature to the first paragraph of the introduction if any, and being cautious about applying small adjustments based on deep ocean differences over time. P3/L5: ->factors P3/L2 through P3/L6: suggested shortening: The implication is that...among the laboratories did not improve between 2008 to 2018 to the same degree that it did for nitrate/phosphate, and the correction factors for silicate were indeed more variable and uncertain than the correction factors for nitrate and phosphate." P3/L6: nutrients->nutrient P3/L10: suggested delete: "by reducing the magnitude of those standard deviations."... this is unnecessary and it is unclear whether it is referring to the inter-lab deviations or the deviations of the RM homogeneity P3/L15: This is unclear. Perhaps:

Disagreements between cruises at depth tend to be smaller when reference materials are used (then quantify this statement or refer to the section where this information is presented). P3/L18: suggested change to either ->provided a synthesis of... or provided synthesis results of P3/L19: This needs a reference to GLODAPv2. P4/L3: I don't follow the logic... why is the quality of oxygen data high just because nutrient reference materials were used? P4/L11: What is meant by "good tracking standards with excellent quality control"? P4/L12: dataset ->data product P5/L10: What are the median filter parameters? Table 2: if there are 30 cruises in category 1, how are there 112 cruises for the category 1 row for nitrate in table 2? I suspect number of cruises should be number of cruise-intersections or number of profiles used for comparisons. P6/L20: it does not imply that. Also, table 2 perhaps implies that category 2 had more consistently-measured O2 than category 1 despite it coming from a much larger pool of research groups. This suggests the reproducibility of the category 1 oxygen data may be low. P7/L5: vertically integrated? Combined uncertainty of measurement uncertainty? Is this combined measurement uncertainty? Broader question: would it make sense to use density interpolated values or multiple linear regression estimated values to limit the impacts of heaving and shoaling further? P7/L12: This logic doesn't make sense since CRMs were also used for silicate (I think. If I am wrong and they weren't used for silicate then you have the related problem with this logic that the ratio between the category 1 silicate value and the other category silicate values is similar between silicate, nitrate, and phosphate). You explain this later on line 16, but by then the reader is already confused. Just omit the silicate information from line 11. P7/L20 could be assumed to be of the same what? P8/L1: variabilities is defined long after it is used. It is also defined again on lines 6 and 8. P8/L17: GLODAPv2 requires a reference here Fig 7: axis labels are not in English and the figure is low resolution. The figure titles are all the same and confusing. The figure legends are not explained. Category 6 looks very good, yes? Why is this? Fig8: phosphate is miss-spelled. Why are the axes reversed? Since they are reversed, why is a 1:1 line still plotted? P9/L3: It is not that the synthesis could not detect differences when the differences were small, it is that

C5

the people putting the data product together chose not to apply adjustments when the differences were small or potentially real differences. This is an important point, because the approach used by these authors assumes there are no changes in the deep ocean. This means this new data product would eliminate and miss the deep ocean changes that they said motivated their work. P9/L10: Are these gridded uncertainties or uncertainties in the measurements? They seem much too small to account for potential gridding errors. P9/L13: what is meant by "chose profiles of factors determined from the global dataset?" Step 2 is also inadequately explained. P10/L8: multiplied by the volume and the density, yes? It is unclear what is meant by "volume corresponding to the density" P11/L9: how were these uncertainties calculated? P11/L11-12: what is meant by nitrate silicate and oxygen being "small," and phosphate similar? Especially if it is large in the next sentence? P12/L4: suggested deletion: ", which is the basic dataset used to more accurately characterize the spatial distribution of nutrients in the global ocean,"

Side note, there only seems to be 4 cruises in the Atlantic in Fig. 1, is the 5th category 1 cruise the Arctic cruise?

C6