

General comments

This manuscript explains the context of the MAREDAT special issue, the characteristics of the 11 Plankton Functional Types (PFT), and provides details of the quality control and statistical treatment/analysis that was performed systematically across all papers of the MAREDAT special issue. It is therefore an important reference paper for the 12 datasets (11 PFTs + 1 HPLC pigments datasets) that were compiled for the special issue. Additionally, the authors provide a summary of the 11 PFT datasets, showing statistics for biomasses and their vertical and horizontal distributions in oceans. Results from the 12 datasets and from the World Ocean Atlas are compared, addressing the coherence as well as inconsistencies among autotrophic and heterotrophic PFTs, and briefly discussing ecological implications of the results for these two trophic groups. Finally, the authors address the usefulness and limitations of the 12 data compilations, identify specific needs for additional data, and acknowledge recent global sampling initiatives that could help fill these gaps. I find it confusing that the manuscript serves two purposes: 1. Overview of the special issue, and 2. Presentation of the microzooplankton dataset. In my opinion, the microzooplankton dataset should remain the data supplement of Buitenhuis et al. (2010), rather than be included as a supplement to the present manuscript. The revised version of the microzooplankton dataset has been submitted to PANGAEA and should be shown there as the supplement to Buitenhuis et al. (2010). The microzooplankton dataset can still be considered as a contribution to MAREDAT without being re-published in ESSD. I provide specific comments below to modify the manuscript accordingly.

See our answer below.

The manuscript consistently refers to 12 MAREDAT “databases”. This is only semantics, but the term “database” is very dependent on the structure and functions of its management system. I recommend referring to them as 12 “datasets” that are archived and incorporated in the PANGAEA relational database. I realize that the separate papers published in the MAREDAT special issue often use the word database (even in their titles) and that my comment is probably impossible to implement across the special issue, but I suggest to follow my recommendation at least for the present manuscript. Doing so implies many corrections in the text, which I do not report here. The authors should search all occurrences of the word “database” and edit the text as appropriate.

As the reviewer notes, the 12 contributions to MAREDAT are referred to in the contributing papers as databases, and the word datasets is used to refer to subsets, either from individual research cruises, time series stations or individual contributing authors. PANGAEA refers to itself as a data library, it is not a relational database. We therefore feel that it would add confusion rather than clarification to have this paper be the one exception in this special issue from the consistent practice to refer to the 12 contributions as databases.

Overall this is an important contribution to the MAREDAT special issue. It brings an original interpretation that complements the results presented separately in more details in the 11 other papers of the special issue.

Specific comments

Page 1079, line 20. Consistently with my general comment about the microzooplankton dataset, I suggest to replace “Microzooplankton database: doi:10.1594/PANGAEA.779970” by “MAREDAT datasets: <http://www.pangaea.de/search?&q=maredat>”

We have added a reference to the latter webpage at the end of the abstract. See also our reply to the comment on Page 1080, lines 15-16.

Page 1080, lines 4-7. The reference for this sentence is probably Le Quéré and Pesant (2009). That reference can then be removed from the previous sentence.

Reviewer 1 also noticed that the relation between these two papers and two sentences relating to them is unclear. As noted there we have added “the previously defined” to the second sentence to make clear that the 2005 paper defined the key PFTs, while at the 2009 meeting it was decided to make the MAREDAT data synthesis effort.

Page 1080, line 9. I suggest replacing “describing a database of” by “on HPLC- based”. Similarly, on Line 13, I suggest replacing “the HPLC-based phytoplankton pigment database paper” by “HPLC-based phytoplankton pigments”.

Changed as suggested.

Page 1080, lines 15-16. Consistently with my general comment about the microzooplankton dataset, I suggest to replace “(here, we briefly published by Buitenhuis et al., 2010)” by (Buitenhuis et al., 2010)”

Based on a comment by reviewer 1, we have clarified that the MAREDAT microzooplankton database has 682 differences relative to the 2010 paper. To maintain a transparent documentation of the development of a database over time (which was also emphasized as important by reviewer 1) we felt that it would not be appropriate to use the same reference for these two different datasets. However, neither has there been enough new work on the microzooplankton database that we felt it was justified to include an individual paper in the special issue on microzooplankton. While we agree that the solution we’ve come up with here is not ideal, we wanted to make sure that the available microzooplankton biomass data were included in MAREDAT, so that they can be used and updated along with the other databases.

Page 1081, line 20. Consider replacing “cycling of alkalinity” by “cycling of calcium carbonate, and thus alkalinity and atmospheric CO₂”.

Changed as suggested.

Page 1081, line 22. Consider replacing “are useful to evaluate” by “are also useful to parameterize carbon and nitrogen cycling in ecological models”

We have added a reference to Luo et al. (subm.), who evaluate global marine nitrogen fixation directly from the MAREDAT data without using an ecological model.

Luo, Y.-W., Lima, I. D., Karl, D. M., and Doney, S. C.: Data-based assessment of environmental controls on global marine nitrogen fixation, Biogeosciences Discussion, subm.

We have mentioned the usefulness of the MAREDAT data to biogeochemical models in the summary section.

Page 1082, line 2. Instead of saying “like MAREDAT”, please provide the name and references of these initiatives. This would be a useful information for the reader.

We have added “(e.g. the CARINA special issue in ESSD http://www.earth-syst-sci-data.net/special_issue2.html)”.

Page 1082, line 8. “We accept that datasets. Hence, we” is a very tortuous sentence. I suggest to simply say “Because data are often too scarce to provide global coverage, we”

Reviewer 1 also made this point and we changed this to “At this point, there is”

Page 1082, line 15-18. I suggest adding after the first sentence of the paragraph: “Each of the 12 datasets can be cited with the DOIs reported in Table 2.”

Changed as suggested.

Page 1082, line 16. Please replace “from the PANGAEA World Data Centre” by “at PANGAEA Data Publisher for Earth and Environmental Science”

Changed as suggested.

Page 1083, lines 11-15. Consistently with my general comment about the microzoo- plankton dataset, I suggest to move these lines to Section 3.2.4 Heterotrophs.

This paragraph deals with the database construction and gridding, while 3.2.4 compares biomass concentrations of the different heterotrophs, in this instance we think it is best to keep these separate. See also our reply to the comment on Page 1080, lines 15-16.

Page 1083, line 26 to Page 1084, line 8. It is unclear to me how Zero values were collected and then transformed. The authors should clarify if all zero values in the raw datasets are absence and NOT “lack of measurement”. I do not understand why zero values “are usually underrepresented, especially in the deep sea”? Do the authors wish to say that true zero abundances are sometimes substituted by no value (e.g. NaN, n/a, nil, or just a blank), so that the information about a true absence is lost.

Yes. We have added this: “(i.e. represented by a blank entry rather than a 0.)”

Or do they wish to say that information about a true absence is lost due to log transformation. Finally, I gather that the authors have generated two types of datasets:

1. “Non-zero observations” were log transformed and QC’d with a 2-sided Chauvenet’s criterion;
2. “Total observations” were not log transformed and were QC’d with a 1-sided Chauvenet’s criterion. In any case, this should be explained more clearly.

We think this confusion was caused by an illogical order of presenting the information. We have moved the information about the gridding of non-zero observations to 2.2 gridding (after 2.1 on the

QC). This should make it clear that, as we state in 2.1, we performed a single one-sided Chauvenet test. On all but the diatom datasets, the data was not normally distributed and was log-transformed. The exclusion of zero values for the QC on the log-transformed data is explicitly mentioned. After the QC we gridded the data, which included (statistical) information on non-zero observations that passed the QC.

Page 1084, line 1. Replace “abundance/biomass” by “abundance or biomass”

Changed as suggested.

Page 1085, lines 1-14. Consistently with my general comment about the microzooplankton dataset, I suggest to move these lines to Section 3.2.4 Heterotrophs.

See our reply to the comment on Page 1080, lines 15-16.

Page 1086, line 2. I suggest rewording “(Fig. 4), though (). This increase in patchiness” to “(Fig. 4). Although (), the increased patchiness in both groups”.

We have changed “though” to “although”, but the contrast that “although” refers to is to the “clear increase” in the first part of the sentence. There is no reason to connect the different relationships for autotrophs and heterotrophs to the following sentence, which provides a potential mechanism for the increase.

Page 1086, line 8. Replace “zooplankton” by “macrozooplankton and pteropods”?

But mesozooplankton also show coordinated movement, in particular vertical migration of copepods, which becomes more prominent across their size range, and this could be seen as intermediate behaviour to the more prominent swarming in some taxa of macrozooplankton. We have changed this to “show coordinated vertical migration and/or swarming”.

Page 1086, line 20. I am puzzled by the sampling bias proposed by the authors for a peak in diatom biomass at 125m (Fig. 2A). This is also reflected in Fig. 2D for the sum of phytoplankton and it seems to match an absence of observations OR true zero abundances of forams and pteropods at 125m I’m sure the authors have looked at the actual availability of observations at 125m. Is the bias towards diatoms due to only one of a few observations? It is probably impractical to run the QC for high values “per depth layer”, but this might have removed this bias. In any case, I suggest to explain in more details this possible bias because this is “the feature” that strikes the eye in Figure 2.

We’ve added a clarification of this diatom peak by rewriting the second paragraph of 3.2.2 and adding: “The diatom biomass profile (Fig2a) has a distinct peak at 125 m. This peak is caused by a single observation of 7210 µg C/L, which was measured in a massive accumulation of mat forming Rhizosolenia, a feature which has been regularly observed in various oligotrophic environments (Shipe et al. 1999).” Although it is due to a single observation, the Shipe reference argues that diatom mats could make a significant contribution to the total silica production in the Sargasso Sea. Furthermore, there are similarly high concentrations at shallower depth, so that this point is not rejected by the QC. All the PFT databases have been treated with the same QC procedure, therefore, as the reviewer notes, it would be impractical to rerun this.

Page 1087, line 20. It is unclear why the presence of zeaxanthin is consistent with the lack of diazotrophs.

We have changed “exhibited maximum concentration around the equatorial region” to “exhibited a maximum concentration around the equatorial region, and is an order of magnitude lower poleward of 50°”.

Page 1087. The comparison of the Total phytoplankton from MAREDAT (with and without Phaeocystis) with those from HPLC and WOA should be discussed in more details in Section 3.2.5, notably the surface (or coastal) maxima shown in MAREDAT (with and without Phaeocystis) and HPLC that is not shown in WOA. I suggest showing the four total phytoplankton profiles in a separate pane, i.e. Fig. 2E, which would also point to differences in the magnitude of the average profiles. It would be useful for the discussion to compute average profiles (totals only) for coastal and open ocean separately.

The subsurface maximum at 10 m depth is reproduced in the WOA2005 data at all latitude bands if average depth profiles are calculated for every 30° of latitude. For the sum of picophytoplankton, diazotrophs (outside the polar oceans), coccolithophores and diatoms the depth profiles are variable, with maxima anywhere between 0 and 30 m, and no consistent pattern with latitude. We therefore think it is too early to draw any conclusions from it, and on the other hand don't think the difference is large enough (relative to the variability of the concentrations that can be seen in Figure 3) to require an explanation.

We have added a paragraph on the changes in biomass between coastal and open ocean at the end of section 3.2.6, and coastal and open ocean panels to Fig. 5. This includes a cautionary note about the large standard deviations which are visualised in Fig. 5, but not Fig. 2. We have therefore chosen to present these as epipelagic averages (0-200m) rather than depth profiles.

Page 1088, line 25. Although I do not object to the idea that a peak at 20m might be “real”, the authors should explain how vertical migrations could result in this peak. It could be argued that organisms aggregate near the surface to feed at night (peak) and disperse in the water column during the day. This behavior would however mean that the vertical integration of biomass for these organisms would be overestimated by up to 2x

Reviewer 1 also thought vertical migration was not a convincing explanation for this peak, and we have changed this to “Macrozooplankton include some species that swarm, which could explain this sharp biomass peak, but it's also possible that some sampling bias such as proposed for the high diatom concentrations at 125 m depth occurred at a biomass concentration that wasn't quite high enough to be excluded by Chauvenet's criterion.”

Page 1091, line 19. I fail to see an inverted pyramid in Figure 5. Perhaps it would be worth doing this figure for coastal (0-100m), open ocean (0-100) and open ocean (>100m) separately.

Again, reviewer 1 made a very similar objection, and we have changed this to: “The global biomass data for each plankton group (Fig. 5) do not show a blunt food pyramid as is typically found in terrestrial ecosystems, but instead confirm the high H:A (heterotroph:autotroph) ratio of around 1 (see also Fig. 2D) that was previously found by Gasol et al. (1997).”

Page 1093, line 4. The Tara-Oceans expedition collected samples during 2009-2012 and will complete its sampling of the world ocean with a circumpolar expedition in 2013. Please correct the text accordingly.

Changed as suggested.

Tables and Figures

Table 1. Two decimal places are probably not necessary for most PFT biomasses reported in Table 1

We have removed two decimal places from the max and 1 from mean and stdev.

Table 2. Please write the full doi reference, e.g. doi:10.1594/PANGAEA.777385, for each line of the table. Please add characteristics of the HPLC dataset as well. The DOI of that dataset must be easy to find along the others.

A table in horizontal format is much easier to read, so this was a trade-off between keeping a horizontal format and having a reference that can be copied and pasted in one go. We have added the HPLC dataset to the table.

Note from authors to ESSD: The dot between PANGAEA and the database number got lost in the table conversion. Please make sure this is present after PANGAEA.

Figure 1. Consistently with my general comment about the microzooplankton dataset, I suggest to remove Figure 1 from this overview paper. It is unfortunate that the revised microzooplankton dataset is not presented as a separate paper in the MAREDAT special issue, but the original dataset is nevertheless described in Buitenhuis et al. (2010). The reader should be referred to that publication and to the dataset at PANGAEA if he wished to examine it in more details. Moreover, the vertical distribution of microzooplankton is already shown in Figure 2.

See our reply to the comment on Page 1080, lines 15-16.

Figure 2. I suggest showing the four total phytoplankton profiles in a separate pane, i.e. Fig. 2E, which would also point to differences in the magnitude of the average profiles. It might be useful for the discussion to compute average profiles (totals only) for coastal and open ocean separately. Make sure that each pane is much larger.

We have included a coastal and open ocean panel in Fig.5, which is our figure which is our nearest figure to figure 4 in Gasol et al. (1997). See also our reply to the comment on Page 1087.

Note from authors to ESSD: Reviewer 1 also asked for Fig. 2 to be larger.

Figure 3. The authors should reorganize and add panes on that figure. A stack of five panes for autotrophs on the left, a stack of five heterotrophs in the middle and a stack of five panes on the right showing: zonal averages from WOA, HPLC, sum of autotrophs (Fig. 3A-E), sum of heterotrophs (Fig. 3G-L) and zonal values of H:A ratios. The latter could help the discussion in Section 3.2.6. Make sure that each pane is large enough to read.

The first changes were made as also requested by reviewer 1. There is a contradiction between adding lots of panels and keeping the panels large enough to read. The sum and ratios of the zonal averages would mostly depend on data coverage of the individual PFTs rather than show the geographical distribution of the profiles shown in Fig. 2.

Figure 4. needs substantial improvement of the format

The lines were made thicker, to allow for the loss of definition during conversion.

Figure 5. Perhaps it would be worth doing this figure for coastal (0-100m), open ocean (0-100) and open ocean (>100m) separately.

We've added coastal and open ocean graphs to Fig. 5, and discuss these in 3.2.6.