**Earth System
Science Data
Discussions**

Interactive comment on "Global marine plankton functional type biomass distributions: coccolithophores" *by* C. J. O'Brien et al.

Anonymous Referee #1

Received and published: 5 September 2012

The authors present a data compilation for coccolithophores. The task is worthwhile – such a data compilation is much needed and would be widely used. Unfortunately the data compilation as presented is not worthy of publication. Hardly any of the prominent researchers in this area have been involved. In addition, there are serious deficiencies in how the dataset has been put together.

A major point is that no effort appears to have been made to engage the specialist scientists who have spent months or years at sea collecting such data. This will inevitably greatly reduce the amount of data included in the dataset and therefore its value (this compilation only scratches the surface of what is out there; there is vastly more data than has been assembled here). The recommended approach for these compilation exercises, and the one most usually followed (for instance in similar compilations that

have been carried out for diatoms and diazotrophs), is to write to a wide range of the foremost scientists in the field to request them to send data. This appears not to have been done in this case. I wrote out a list of 15 experts that I would have expected to have been involved; only one is actually involved. Of the two multi-author edited books titled "Coccolithophores" that have been published in the last 20 years (the most recent was published in 2004), not a single chapter author is included in the author list of this paper. This is in stark contrast to Luo et al 2012, for instance, which appears to involve all of the acknowledged most active data collectors and has therefore been able to pull in the bulk of the data on nitrogen-fixers and nitrogen fixation.

Some ill-advised decisions have been made, including to calculate cell biomass even when the species was not reported and so the conversion from cell number to cell mass is unknown. It is also unfortunate and counter-productive to calculate biomass from the coccosphere size, i.e. the size including the coccoliths. Biomass is conventionally used to refer to the organic carbon only. The term is not used to refer to the sum of POC + PIC, and to do so in the dataset is misleading and would lead to confusion and even scientific errors in subsequent studies if they used the 'biomass' data without realising that in fact it represented both POC and PIC. It does seem sensible to me to include in the dataset estimates of both POC and PIC (as well as of cell concentrations), but not to lump the two of them together. The issue of morphotypes, and whether to include this information if available, is not mentioned. The most widely used taxonomy guides, those produced by Jeremy Young and colleagues, should be mentioned.

Most seriously, there is no mention of the different techniques used to quantify coccolithophore abundance. A key paper on the topic (Bollmann et al, 2002) is not even cited. This is important because there is quite a wide variety of different techniques with important differences in the information they yield. Some of the techniques are: (1) visual counting from scanning electron microscope images, (2) visual counting under a light microscope, (3) visual counting under a light microscope using cross-polarised light, (4) flow cytometry, and (5) automated identification using the SYRACO software. De-

pending on the technique used there is a great variation in the quality of the resulting information. Some methods yield robust estimates, others are not proven. Some can count small coccolithophores but not large. Other techniques can count large but not small coccolithophores. Some are more laborious tending to lead to smaller numbers of cells being counted and hence, potentially, errors associated with concentration estimates derived from small numbers. A cell or biomass concentration estimate that is derived from seeing just one cell on a microscope image and then scaling up has much greater uncertainty compared to caluculations from counting 100 cells in a sample.

These issues are not even mentioned in the paper, much less dealt with. For instance, it is arguably better to exclude information from some sources in order to maintain a high overall standard to the data quality. At the very least a column should be added to the dataset to indicate the quantification method, so that subsequent users of the dataset can make their own decisions about which sorts of data to include and which not. However, another issue, how to deal with zero counts, is correctly identified and discussed. Uncertainties associated with scaling up from counts of scarce species should be calculated and reported in the dataset.

The authors are to be congratulated on having attempted this important service for the wider community. However, they would have been better advised to have made a much wider trawl for available data, and to have made their compilation in conjunction with, rather than separate from, the main people collecting the data.

References

Bollmann J et al. 2002. Techniques for quantitative analyses of calcareous marine phytoplankton. Marine Micropaleontology, 44, 63–185.

Luo Y-W et al. 2012. Database of diazotrophs in global ocean: abundance, biomass and nitrogen fixation rates. Earth Syst. Sci. Data, 4, 47-73, doi:10.5194/essd-4-47-2012.

C173

Interactive comment on Earth Syst. Sci. Data Discuss., 5, 491, 2012.