Review ESSD-2017-58, BGC Argo

The authors / data providers have prepared and presented a potentially useful data set. A user gets easy access and very good descriptions. Researchers have good options to use the data in depth profile formats or to take advantage of the euphotic zone summaries. The authors have performed a valuable service by compiling these particular bio-optical data from the larger ARGO data set and by specifying careful but consistent quality control procedures. The data seem like a very good fit to this journal. I applaud the effort and recommend publication. (Having Roesler et al. 2017 in open access proved very important for evaluation of this data set.)

I recommend changes to the presentation of this data that should make it much more useful. I make three over-arching suggestions followed by a sequence of line-specific technical suggestions or questions.

1) Global impact

The data set derives from 105 profilers operating over a time period of 40 months. Although one would have to dig deeply through the data to confirm, I believe that no single profiler operated for the entire 40 months. We have instead a compilation based on a series of relatively short (20 to 30 month) deployments each in a relatively restricted region. Although the map presented here as Figure 1 and the number of "stations" approaching 10'000 seem impressive, on the scale of an evolving global ocean they give us only a snapshot. A useful, unique, challenging snapshot, hard-won in the face of funding and operational constraints, but a relatively short Atlantic-focussed glimpse none-the-less. This data set basically misses most of the Indian and almost the entire Pacific Ocean and, not surprisingly, stays well away from ice-influenced regions. We should feel very well served to have these data! But we should not pretend that they provide us an encompassing view of an evolving global ocean. These authors hint at these limitations in their conclusions (page 11, line 15), where they mention " new regions", "improved vertical and temporal frequency", etc. Against these cautions, I feel that the title which includes the phrase "biogeochemical and bio-optical applications at the global scale" greatly overstates the impact.

The authors might consider several other efforts to compile global biologically-relevant or carbon cycle-relevant data sets for the oceans, including Peloquin et al. (HPLC chlorophyll) and Valente et al. (near-surface bio-optical properties) - both of which they do cite - or Sauzéde et al. (in situ fluorescence) or Bakker et al. (SOCAT, surface ocean CO2) which they do not cite. Taking only those four examples (all from ESSD), we typically see data collected over 2 to 6 decades presenting several 10s of thousands (up to millions, for SOCAT) of stations, profiles and measurements. In this paper we see nearly 10'000 observation in only 40 months (closer to 5'000 for the first optical depth) - the promising impact of ARGO technology which the authors could highlight more clearly - but also clear temporal and spatial limitations compared to other data compilations. In presenting their very careful quality control discussion, these authors have failed to show us clearly what their bio-ARGO profilers have achieved for ocean observations and also how these data contribute to, fit with, supplement, or surpass prior and on-going pan-oceanic data compilations. For this reader, these authors have missed an opportunity to quantify how "The BGC-Argo sampling approach can therefore help the scientific community accumulate observations on biological and biogeochemical properties of the ocean" (page 2, line 10). We want to do more than merely "accumulate"?

2) Representativeness

From the understandable view of these biological and bio-optical oceanographers, the ideal ocean situation occurs when a profiler reaches the ocean surface near local noon with small waves and a cloud-free sky. In this paper we encounter a series of qualitative statements about variances from those ideal conditions:

Page 7, line 33 "unstable" meteorological conditions.
Page 8, line 4 - Again a focus on stability (now of the water column!) and "deteriorated sky and sea conditions".
Page 8, line 24 "worsening meteorological conditions and deepening mixed layer depths".

But, the global ocean represents a windy, cloudy, stormy place. Large regions have persistent coverage of stratus clouds at certain times of the year. For some parts of the ocean we have almost no cloud-free images despite nearly 40 years of daily satellite observations. From a biological view, we should appreciate disturbed conditions and vigorous mixing processes. I believe the standard (non-biological) ARGO profilers rise to the surface without regard to sky or wind? If, in this data set, the authors, consciously or sub-consciously, focus on and lead the user toward mid-day profiles under calm seas and clear sky conditions, we together suffer the risk of developing a serious bias?

3) Accuracy

Having overcome many of the serious technical, operational and funding challenges of gaining useful bio-optic data from autonomous ocean profilers, and having applied a consistent set of quality control procedures intended primarily to remove spikes and outliers (and secondarily to identify instrument drift) the authors then present all data as uniformly certain. Although the authors show data distributions in several plots, we see no error bars and no uncertainty shading. For intercomparisons, these authors give us only global ranges (min-max values).

The description needs to thoroughly address uncertainty in a specific section. After all processing, what remaining uncertainties apply to what data? What changes in operation, instrumentation or data processing could or could not address those uncertainties? For the derived properties Z_{EU} and Z_{PD} we get statistical (standard deviation) uncertainties (e.g. Table 1 on page 17) but those derive from the data processing and do not address uncertainties in the underlying measurements? For many real-world measurements uncertainty remains very difficult to quantify but in this case the reader needs from the authors at least a sense of confidence and uncertainty for each product. If, as the authors clearly hope, these data prove useful for global models, quantitative uncertainty information will prove an absolute requirement.

These authors need to explicitly address issues raised by Roesler et al. In its present form, this manuscript appears to have added those issues, and applied a uniform 2x correction, after the fact or at least late in the preparation process. Roesler et al attempted to exclude from their analysis exactly those mid-day, clear sky, high irradiation conditions that this compilation seems to favour - how does that mismatch affect the values presented here? Roesler et al - using many of these same data! - showed a very strong regional dependence of correction factors, indicating that a global uniform application of a 2x correction factor would in fact prove seriously wrong in almost all cases for almost all regions. Yet here we read about a simple 2x correction for chlorophyll values? Roesler et al. provided explicit regional- or biome-based correction factors that this paper should have considered?

In its present form, without an explicit discussion of uncertainty, this manuscript will fail to meet the needs and expectations of many potential users.

Specific comments as follows:

Data considerations

Why, at http://www.seanoe.org/data/00360/47142/ (corresponds to doi http://doi.org/10.17882/47142) do we find two versions, version 1 and version 2? Readme text in version 2 explains the differences, but the deletion of these 20-some profiles received no mention or justification in the text. Properly, a doi should point exclusively to a single version of any data. Second version should carry a second doi.

We get profile data in separate parameter-based files (e.g. CHLA, CDOM), space delimited. We get the derived optical depth data in one file, properly comma delimited. (.csv). Why do we not get all files in .csv format?

Floats to breach surface around local noon. 100 floats times 100 profiles per each gives roughly 10000 profiles. At once per 10 days, 100 profiles would take 1000 days, not quite 3 years. Average profile interval must therefore approach more like 13-15 days? No profiler operates for the full 40 months, but wrong to specify an average of 10 days over that full time period. More properly an average of 10 days while operating?

Page 2, Lines 16, 17: “nitrate concentrations” Has any ARGO biogeo float solved the NO3 challenge? Not sure why the authors mention this?

Page 4, correcting the FDOM. Removed spikes outside 25 and 75, and then additionally remove spikes greater than 4 x the mean value? Identification and quantification of temporal deterioration assumes consistency of deep water masses?

We need to know the duration of each profiler’s operation. Can we determine the average lifetime of each float? E.g float 7900591 operated from 2013-12-20 to 2015-07-05, e.g. 30 months, over which time it took roughly 70 profiles (average interval of 13 days but even greater because for the first month after deployment it apparently profiled once per day). We also need number of profiles per profiler? To understand anything about contamination or other performance deterioration as a function of cumulative time in the water, we need to know average deployment duration for the profilers as well as number of profiles by each. Not hard to extract and perhaps graph this information? This information would also prove helpful in making the points about large effort and great resources needed to achieve even this level of spatial and temporal coverage.

Page 5, line 2: “ E_d (0_-)” and Page 5, line 9: “ E_d (0_+)” typos? How related to E_d (lambda)?

Page 5, line 17, tracking possible biofouling. Interruption of the time series occurred in real-time or during post-processing?

Page 6, line 2 - simple vertical average of CHLa, CDOM, etc., for optical depth? How did these optical depths compare to other data, globally or regionally?

Page 6, comparison with satellite data.

The authors have missed an important opportunity to connect and compare these data to the bio-optic data reported by Valente et al. These authors cite that data set (page 9, line 20), but only once for the purpose of confirming the bio-optic properties reported here (and buried in a summary sentence in which the reader can’t determine which external paper connected to which parameter reported here). In fact the Valente et al data represent an important partner for these data. Those data stop at 2012, these data extend through 2015. This data includes Labrador Sea and Southern Ocean locations missing in the Valente et al. data. That data has many more values, e.g. for Chl A, that could give a much better regional comparison (North Atlantic, for example) for these data. These authors do not need to do or show the work of attempting to merge this data into a Valente et al. framework but they should explicitly outline the connection points and opportunities. These authors could also very much learn from the graphic approaches (e.g. Valente et al. figure 3 and figure 10) and users of both data sets need some convergence of variable names and units (e.g. for FDOM treated very differently in the two data sets).

Page 7, line 22 “open-ocean environments” If by “open-ocean” the authors mean ‘collected in areas with depths greater than 1000 meters, as opposed to shallower continental shelf regions, then we can perhaps accept their definition so long as they describe what they mean more carefully. If, however, “open-ocean" should imply a broad spatial coverage of large ocean regions, then the absence of measurements from the Pacific stands out. The authors can correctly say `within limitations of project-driven resources, deployments focused on some of the important carbon-export regions of the Atlantic Ocean, in all cases in regions with depths greater than 1000 meters’?

Page 8, line 31. And North Pacific equals 0 %.

Page 9, line 16. I don't believe, from this data set, the authors can make any statistically-valid statement with respect to any part of the Pacific.

Page 9, final paragraph. Again, a good description of percentages within regions of deployments, but must include recognition of very large areas (e.g. Pacific) with no deployments.

Page 19, legend to Figure 1. Dot and diamonds hard to distinguish, should specify red diamonds and blue dots.

Page 19, Figure 2. Reader to assume that data from float 6901439 were truncated at some point, with data after that time point discarded?

Page 20, Figure 3. Panel A - Why assume only non-photochemical quenching? Does NPQ include by definition active surface avoidance? Panel C includes the 2x factor?

Page 20, Figure 4. Why the blue circles indicate spikes in corrected (red) curves? Also, in this case, we assume corrected greater than raw due to a sensor performance issue tracked from the deep FDOM? But at what point would biofouling or sensor deterioration have disqualified the measurements?

Page 21, Figure 5. My old eyes see blurred blue dots, no blue open circles.

Page 21, Figure 6. Ascent speed, profile time? Variability of cloud shading over that time period?

Page 23, Figure 8 - Not a useful way to show geographic data, confusing. Use maps instead, as in Figure 9 (or Figure 10 of Valente et al.)

Supplement consists of a single table defining the acronym codes for specific ocean regions. If, as I suspect, it applies to or derives from a larger ocean regional description scheme, the authors should cite the external references. If it represents a product custom to ARGO generally or bio-ARGO specifically, the authors should include the table as an appendix in this ESSD manuscript. Otherwise, according to Copernicus archive procedures as this reviewer understand them, the archive process could preserve the manuscript but not the supplement.