

## ***Interactive comment on “Copepod species abundance from the Southern Ocean and other regions (1980–2005) – a legacy” by Astrid Cornils et al.***

**A. Atkinson (Referee)**

aat@pml.ac.uk

Received and published: 2 May 2018

This paper describes a valuable data set on copepod abundance, depth distribution and stage structure, mainly from the Southern Ocean. In my opinion this is particularly well suited to the journal Earth System Science Data. This indeed forms a valuable legacy to the enthusiasm of Sigi Schiel, for her cooperative approach to research with the sharing of data. Most of my suggestions are minor but I start with a few that I think are a bit more important.

Section 4. I think that it should be stated very clearly in this section that people using these data should always cite both the doi and the data paper itself. As we move

C1

towards open data and the idea of a big data free for all, I think that the value of the source data should be respected via citation. Citing the data paper as well as the doi is both out of respect for the data themselves and courtesy to their originators, but also allows due traceability of the data. I think this point should be stressed at every opportunity.

Title- Abstract. I would have thought that Sigi's name could perhaps be mentioned in the Abstract so the legacy intimated in the title is made more clear to people who do not know. This could simply be achieved by moving a slightly modified version of the last sentence (lines 361-363) to become the first lines of the Abstract.

Throughout the manuscript: the rationale for this data compilation is firmly around the notion of understanding the effects of climate change. This is indeed one strong reason but I think the authors could be much broader than this on how the data may be valuable. As one example Sigi worked on a series of papers in the early 1990s compiling all available data to look at life cycles of biomass dominants or key copepod species, examples being *Rhincalanus gigas*, *Calanoides acutus* and *Stephos longipes*. Such possibilities could perhaps be mentioned as one potentially valuable use of such depth-resolved, seasonal and stage-resolved data, e.g. Ward et al. (1997).

Section 3.4. I think that a bit more clarification of the internal comparability of this data set is needed. In particular the fact that some of the species perform extensive seasonal vertical migrations could be emphasised a bit more in the text. The implication of this is that nets particularly in surface waters may not always sample the full vertical extent of the populations, particularly those of the biomass dominants. Those working in the Southern Ocean will know this already but nevertheless I think it is important, because many non-specialists may want to use these data.

Following on from this I am not sure about the recommendation to use presence-absence data to account for differences in sampling method (line318). I am sure people will do this anyway... but I do not think this should be a recommendation! Presence

C2

absence data (or in other words the probability of detecting a species) will also depend on a range of factors such as volumes filtered by nets, mesh size, subsample volume analysed and other protocols for looking for rare species).

There are a large number of data compilation, archival and metadata cataloguing initiatives around and it is hard to know the relationship they have with each other and whether the same data sets (or overlapping data sets) are stored in multiple places. For example Southern Ocean zooplankton data are also compiled in COPEPOD and stored BODC and probably a load of other places besides. I think that adding a very short paragraph to highlight other such initiatives and the link between this dataset and the other ones would help the reader. Not a few paragraph, but at least a clarification of whether some of the data are either a) stored or b) catalogued elsewhere would be useful.

The text mentions climate change and ice loss in Antyractica, but this should be qualified to regional ice loss.

Abstract. I think the Abstract needs a bit of work with a sentence to summarise issues of data use (for example vertical coverage, net types which influence internal comparability of the data). People using the data may be in too much of a hurry to read the detail so the really essential stuff should come across in the Abstract. Also I reckon it needs a final sentence outlining the value of such datasets (effects of climate change, biogeographic distribution, life cycle, inputs to various models. . . etc) and the fact that some

#### Minor comments

Line 11 change intermediators to intermediaries Line 39 change Antarctic copepod community to Antarctic epipelagic assemblage. This is because diversity at depth is high. Line 54. The Atkinson et al (2013) paper is 2012. Line 61. Change decreases to decreased Line 70 change to The data sets presented here. Line 92. Better to delete the first sentence Lines 242: for clarity can net times be stipulated as whether GMT or

C3

local (but present offset to GMT) and Julian days stipulated as Jan 1 being either 1 or zero.

Reference Ward P, Atkinson A, Schnack-Schiel SB Murray AWA (1997) Regional variation in the life cycle of *Rhincalanus gigas* (Copepoda: Calanoida) in the Atlantic sector of the Southern Ocean-re-examination of existing data (1928 to 1993). *Marine Ecology Progress Series* 157:261-275

---

Interactive comment on Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2018-36>, 2018.

C4