

## ***Interactive comment on “UDASH – Unified Database for Arctic and Subarctic Hydrography” by Axel Behrendt et al.***

**T. Tsubouchi (Referee)**

takamasa.tsubouchi@uib.no

Received and published: 22 January 2018

### **1. General comments**

This paper presents new hydrographic dataset (temperature and salinity) in the Arctic Ocean and Subarctic sea for the period of 1980 - 2015. The dataset is prepared by combining publically available hydrographic data from various data sources. Proper quality control check and data flagging have been done to secure the quality of dataset. The dataset is available on online, easily downloadable and well documented.

Any field campaigns in the Arctic Ocean require significant logistic preparations and often diplomatic negotiations. After years of preparation, a bad weather, unfavorable sea ice condition, change in political situation can prevent implementation of the scheduled

C1

field campaigns. This may be one of reasons why there is no widely recognized publically available hydrographic dataset in the Arctic Ocean. As nicely illustrated by figures 1-3, the quality of WOD 2013 is simply not good enough to conduct any scientific studies. In reality, individual researchers or even organizations own their quality controlled dataset, which is not publically available. One of consequences of this is that more than 10 years old climatological temperature and salinity fields based on WOA 1998, EWG Arctic Ocean Atlas and other sources, as known as PHC climatology (Steele et al., 2001), is still widely used to validate climate model outputs (e.g. Il'ĭnskĭ et al., 2016). The latest version of PHC 3.0 was updated 12 years ago on March 2005. Given the rapid and drastic ocean condition change in the Arctic Ocean from late 1990's (e.g. Proshutinsky et al., 2009; Polyakov et al., 2017), the PHC 3.0 dataset does not represent the ocean condition in present days. Therefore, this UDASH dataset is a milestone to foster various scientific studies in the Arctic Ocean and Subarctic sea.

Having said that, there are three points needs to be improved before it could be published. Once these issues are properly addressed, I would certainly recommend the editor to publish this manuscript in the Earth System Science data journal. First, the dataset is somehow limited for wide exploitation in the climate science community because of the way it is prepared. Second, on-going and future plan needs to be refined. Third, figures and tables need to be improved.

### **2. Major concerns**

#### **2.1. Limitation of dataset exploitation**

The dataset is somehow limited for wide exploitation in the climate science community because of the way it is prepared. There are two issues here. First, it is unfortunate that authors choose to prepare the dataset north of 65N. The choice of 65N somehow limits the potential exploitation of the dataset in future. The physical boundary for the Arctic Mediterranean is not 65N, but Greenland-Scotland Ridge (Aagaard et al., 1985; Mauritzen 1996; Dickson et al., 2008). A recent GRACE satellite and data assimilation

C2

based study shows that uniform sea level variability both in the Arctic Ocean and the Nordic Seas together on the time scale from weekly to inter-annually (Fukumori et al., 2015). As it is, the dataset chops off the southern part of the Nordic Seas. This will hamper future scientific attempts to address physical interactions between the Arctic Ocean and the Nordic Seas. Thus, I strongly suggest to widen the boundary down to 60°N. Since quality control procedure and necessary scripts are already prepared, I presume this can be done relatively easily.

The other issue is that the UDASH dataset does not provide gridded temperature and salinity fields. Since the gridding procedure requires decent time and effort, in reality, this dataset would not be widely used by numerical modelers. As a result, it is likely that the more than 10 years old PHC 3.0 dataset would still remain as a standard observation based dataset to validate the climate models. I understand that preparation of gridded dataset is out of scope of this paper. However, authors could still try to fill the gap. There are several efforts going on to prepare in situ observation based world gridded temperature and salinity datasets. For instance, EN4 by met office in UK (Good et al., 2013: <https://www.metoffice.gov.uk/hadobs/en4/>), MIMOC by NOAA in USA (Schmidt et al., 2013: <https://www.pmel.noaa.gov/mimoc/>). Other possibilities could be found here ([http://www.argo.ucsd.edu/Gridded\\_fields.html](http://www.argo.ucsd.edu/Gridded_fields.html)). Authors could contact them directly. There is a chance that they would agree to incorporate the UDASH dataset for their next generation of dataset. In this case, authors could mention this in section 4, on-going and future plan.

## 2.2. On-going and future plan

On-going and future plan needs to be refined. As it is, it is a wish list rather than a feasible plan. Authors need to be more realistic to define what can be done and what can not be done by identifying amount of work involved for each task. Do authors have funding to update the dataset? If yes, how many months / years it would be? Is updating every year realistic in practice? I believe updating the UDASH database every 3-5 years with same quality control scheme to extend the time series towards present

C3

days would be very beneficial to the climate scientific community. Going backwards before 1980 may require additional quality control check.

Again, it is important to create the gridded dataset either directly (i.e. by authors) or indirectly (i.e. by someone else) in order to replace the PHC 3.0 dataset to validate the climate models. Authors should explore any opportunities to make this happen. It is also important that authors promote the UDASH dataset to data assimilation community so that UDASH data will be assimilated to next generation of the data assimilation products. For example, ECCO (<http://www.ecco-group.org>), ORAS4 by ECMWF (<https://www.ecmwf.int/en/research/climate-reanalysis/ocean-reanalysis>), and name more (<https://climatedataguide.ucar.edu/data-type/oceanic-reanalysis>).

Concrete and feasible future plan which guides a future direction of the next generation of the UDASH dataset and exploitation plans in the climate science community would be worth to see.

## 2.3. Figures and tables

Figures and tables need to be improved. Generally speaking, figures and tables describing about the quality control check (figure 1-7 and table 1) are well prepared. However, figures and tables describing about the UDASH dataset itself (figure 8-16 and table 2-5) need to be presented in more systematic way. What readers want to see is when, where, what kind of, and how much data are available in the UDASH dataset. I think figure 11 and figure 12 is a good way to provide such information. In practice, I would suggest following improvements.

(a). Prepare another figure between figure 10 and figure 11 about number of profiles available in different months as same format of figure 11 and figure 12. Six distribution maps can be prepared by splitting months into six categories: Jan.-Feb., Mar.-Apr., ... Nov.-Dec.. Below the six distribution maps, histogram of number of profiles as a function of year (same as figure 11 and figure 12) with the six different categories in different colors can be shown.

C4

(b). Remove figure 13 and figure 14.

(c). For figure 15, prepare it as a same format of figure 11 and figure 12. Below the six distribution maps with different depth ranges, histogram of number of profiles as a function of year with the six different profile depth range categories in different colors can be shown. Double check the fraction of the full dataset. When they are summed up, it should become 100%.

(d). For figure 16, add additional four distribution panels. Temperature and salinity distribution in the Halocline layer (e.g.  $50\pm 20$  m) and in intermediate layer (e.g.  $1,250\pm 250$  m) would be useful to illustrate the UDASH dataset.

(e). Remove table 4 and 5. Put the fraction of full archive information in figure 11 and figure 12 as done in figure 15.

(f). For figure 8, prepare T-S plots for both Amerasian Basin and Eurasian basin. Plot both before QC and after QC in different colors. The depth range does not have to start from -500m. Choose more appropriate depth range to zoom in the profiles.

(g). Figure 9c. Why this looks different from figure 8a (bottom right, after QC in Amerasian Basin)? They should be identical.

### 3. Minor concerns

Captions for figures and tables are too short, and often not well explained. This needs to be improved.

For example, Figure 3, 4, 5: Specify cruise number, ID, year, month so that the figures can be reproduced. Table 1, 2, 3: captions are too short. Table 1: Most of data source acronym should be in capital letter.

---

Interactive comment on Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2017-92>, 2017.