

## ***Interactive comment on “An inventory of Arctic Ocean data in the World Ocean Database” by Melissa M. Zweng et al.***

### **Anonymous Referee #2**

Received and published: 27 November 2017

P1, Abstract: Time period of the collected data should be clarified in abstract.

P1, L8: definition of ‘casts’ is unclear for me.

P2, L3-4: In the paper, data distributions are discussed from different perspectives. “.. WOD represents a collection of ...” should be described more details.

P3, L3: It is better to add short description about “WOD13”.

P5, L22-28: WOD provides data by three different format types (WOD native ASCII format, CSV format and netCDF format), and user selects one of them. To read WOD native data format, the user requires ODV software package or use source codes in FORTRAN, C, Matlab etc.... The sample codes are provided. These information should be given in the sentence.

C1

P6, L18: “... in WOD13.” should refer to “URL:[http://data.nodc.noaa.gov/woa/WOD/DOC/wod\\_intro.pdf](http://data.nodc.noaa.gov/woa/WOD/DOC/wod_intro.pdf)”

P7, L10: “... (due to the International Geophysical Year (IGY))”

P7, L13: Is it “Until 1940-1950, ...”?

P8, L23: “The International Global Ocean Data (GODAR) project ...”

P8, L23: When is the GODAR project period? What types of data were mainly provided by the project?

P9, L3: “... Korabev et al. (2014), Yashayaev and Seidov (2015))”

P9, 6-7: when is the period of the data covered by ASOF, SBI, NPEO and International Arctic Buoy Program?

P9, L12-17: this sentence describes about variable types. I suggest to move the sentence in 5.1 and change the subtitle as “Data distributions by time, space and instrument and variable types”.

P9, L13: “... instrument type. However...”

P11, L2: Aagaard, K. and Carmack, E.C.: The role of sea ice and fresh water in the Arctic circulation, J. Geophys. Res., 94, 14, 14485-14498, 1989.

The style of other references should be followed as above.

Figure 2: 80N should be marked in the figure.

Figure 4: 80N should be marked in the figure.

Figure 5: It is better to show what the abbreviations in the figure mean .

Figure 5: Since some points are overlapped, they are unclear.

CSV data file: References of NODC code, WOD code and OCL code should be provided in the data file.

C2

NetCDF data file: >The data format does not comply with CF convention. I suggest to use CF checker <http://puma.nerc.ac.uk/cgi-bin/cf-checker.pl>

><https://www.nodc.noaa.gov/OC5/WOD13/> should be added in the global metadata as "reference".

>NetCDF file should include minimum global and variable metadata for the data in the file. "keywords" is highly recommended in global metadata. For "keywords", you may choose from GCMD science keywords or something else. "Licence" is also recommended. An example of "licence" is "This data is made freely available by NODC. User must display the this citation in any publication as: < Boyer, T.P., J. I. Antonov, O. K. Baranova, C. Coleman, H. E. Garcia, A. Grodsky, D. R. Johnson, R. A. Locarnini, A. V. Mishonov, T.D. O'Brien, C.R. Paver, J.R. Reagan, D. Seidov, I. V. Smolyar, and M. M. Zweng, 2013: World Ocean Database 2013, NOAA Atlas NESDIS 72, S. Levitus, Ed., A. Mishonov, Technical Ed.; Silver Spring, MD, 209 pp., <http://doi.org/10.7289/V5NZ85MT>>". Variable metadata "units" is highly recommended for each variable. For instance, "salinity" variable does not have "unit" metadata.

>It is sometimes difficult to guess what variable is given from variable name. "long\_name" and "standard\_name" are highly recommended for all variables as variable metadata.

>If all values of a variable are same, it is not necessary to add it as a variable. For instance, all flag values for a variable 'X' take 0, I suggest to include it as "flag" metadata in variable "X".

>Data structure should be as simple as possible. For instance, the following two variables (exampleA) are able to summarize into one variable (example B). :

exampleA )

```
short Salinity_WODflag(z) ;
_____ Salinity_WODflag:flag_definitions = "WODf" ;
C3
```

short WODf ;

```
_____ WODf:long_name = "WOD_observation_flag" ;
_____ WODf:flag_values = 0s, 1s, 2s, 3s, 4s, 5s, 6s, 7s, 8s, 9s ;
_____ WODf:flag_meanings = "accepted range_out inversion gradient
anomaly gradient+inversion range+inversion range+gradient range+anomaly
range+inversion+gradient" ;
```

Example B)

```
short Salinity_WODflag(z) ;
_____ Salinity_WODflag:long_name = "WOD_observation_flag" ;
_____ Salinity_WODflag:flag_values = 0s, 1s, 2s, 3s, 4s, 5s, 6s, 7s, 8s, 9s ;
_____ Salinity_WODflag:flag_meanings = "accepted range_out inversion gra-
dient anomaly gradient+inversion range+inversion range+gradient range+anomaly
range+inversion+gradient" ;
```

>Data structure is complex since several variables are included in one file. To make groups using netCDF4 format might be easier to understand for users.

Interactive comment on Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2017-63>, 2017.