Earth System
Open Access Science
Data Discussions

ESSDD

Interactive
comment

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

Melissa M. Zweng et al.

melissa.zweng@noaa.gov

Received and published: 4 January 2018

Dear Referee #2, thank you for taking the time to review this paper. We appreciate your careful look at the manuscript, particularly the requests for clarification of terms and information, and will revise the manuscript based on your suggestions.

P1, Abstract: I will include the time period of the data.

P1, L8: The definition of "cast" is one that often causes confusion, and I was remiss in not providing more detail. A cast is a single profile taken concurrently, and the definition of a 'profile' is a set of measurements for a variable taken at discrete depths in the water column. For an instrument like a moored or drifting buoy that does not move up or down in the water column, a 'profile' is a time snapshot of the measurements

taken at different depths by the instruments on the mooring chain. I will include this information in the updated manuscript.

P2, L3-4: I sought to introduce the WOD concept but deferred the detailed explanation of it until section 5. I will add more detail to this paragraph.

P3, L3: Thank you for pointing out that I mentioned WOD13 before explaining what it was. I will correct this.

P5, L22-28: I will add detail about WOD download methods and formats here.

P6, L18: I will include a reference to the data documentation here.

P7, L10: The acronym is spelled out on P2, L21, but I will write it out again here for clarity.

P7, L13: I would say 1950-1960 is the first decade with geographically broader sampling, but the data in Baffin Bay and the Chukchi Sea in 1940-1950 do make a strong case.

P8, L23: GODAR is spelled out on P4, L26, but I will do it again here. I will also include a citation with more information about the project.

P9, L3: Thank you! I will correct this.

P9, L6-7: I will include the dates of these project data.

P9, L12-17: This is a logical suggestion and I will move the text as described.

P9, L13: The parentheses are correct here.

P11, L2: I will make changes as described.

Figure suggestions are noted and agreed with; I will make changes accordingly.

Finally, we very much appreciate your close look at the WOD netCDF format. Tim Boyer (tim.boyer@noaa.gov) maintains WOD and its netCDF format and provided a

response that addresses many of your concerns, and has made several corrections to the format based on your comments. He is also happy to discuss further questions and encourages you to contact him directly. Please see the following from Tim:

Note regarding all netCDF comments - the World Ocean Database (WOD) netCDF contiguous ragged array format has been revamped over the last six months. The new format is expected to be available by the end of January, 2018. An example file (all CTD data for 2017 - through Sep. - in WOD) is available at ftp://ftp.nodc.noaa.gov/pub/WOD/SELECT/wod_ctd_2017.nc

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

- Thank you. Following the reviewer's recommendation and using the CF compliance checker reveals the use of type SHORT for arrays [variable]_row_size instead of type INTEGER. This has been corrected. After this correction, the format is compliant with CF contiguous ragged array conventions.

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

- In the updated version of the format, the following is given as reference: references = "World Ocean Database 2013. URL:http://data.nodc.noaa.gov/woa/WOD/DOC/wod_intro.pdf

>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.

- Good suggestions. We have discussed 'keywords'. The problem with keywords is that not all files (especially when ragged array netCDF is implemented in our request software - WODselect) will contain each type of data. There are some files without temperature profiles for instance. We have discussed more general keywords from GCMD but at this time have decided not to include any in our global attributes. This may change in the future.

As for license, this has been another source of discussion. Since the WOD is a U.S. Government data set, it is by law available without restriction - even the restriction of requiring citation. For this reason, we have not included the license keyword.

Finally, as part of the format update, units has been added as an attribute to most variables. Some few, such as 'originators_cruise' still do not have units. But all of the measured variables do have units - with the exception of salinity which is unitless.

>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.

- long_name and/or standard_name are now included for each variable.

>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".

- The problem with doing this is that there is no way a priori to know which files will have all flag values of 0. And if there were, to have one file with flag as metadata since all values are 0 and other files with flag as a variable since there are non-zero flags

C4

would be inconsistent and could create problems for reading software.

>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" ; _____Salinity_WODflag:flag_meanings = "accepted range_out inversion gradient anomaly gradient+inversion range+inversion range+gradient range+anomaly range+inversion+gradient" ;

- This has been done in the updated format.

>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.

- We discussed using groups - and in fact we do use groups for plankton data which are much more complex than ocean profile data. The problem with groups is that many applications don't recognize them. For instance, our own THREDDS server simply ignores plankton data in WOD since it cannot interpret the group information. Many, but not all applications have moved from netCDF3 to netCDF4, but even some of those which have made the switch are not equipped to handle groups.

————————————————

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