

Earth Syst. Sci. Data Discuss., referee comment RC2
<https://doi.org/10.5194/essd-2022-384-RC2>, 2023
© Author(s) 2023. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on essd-2022-384

Anonymous Referee #2

Referee comment on "The Ant-Iso dataset: a compilation of Antarctic surface snow and ice isotopic observations" by Jiajia Wang et al., Earth Syst. Sci. Data Discuss.,
<https://doi.org/10.5194/essd-2022-384-RC2>, 2023

General comments

The presented dataset of stable isotope composition in precipitation, snow and ice in Antarctica is useful and novel but require better description in the article and careful editing of the dataset. There is high potential of data being useful in the future. The description of methods and materials should be improved. The dataset is available via the link and seems to be complete but does not follow all common standards.

Such important characteristics like time coverage, elevation range and types of sample points should be provided in the abstract and dataset description.

Specific comments, questions and suggestions are listed below.

Specific comments:

Why authors of the Antarctic surface snow isotopic dataset available for download at <https://doi.org/10.5281/zenodo.7294183> are different from the authors of the submitted ESSD article?

It is not clear why the title of ESSD paper is "The Ant-Iso dataset: a compilation of Antarctic **surface snow** isotopic observations" while there are 80 points of ice cores and 235 points of firn cores. You should either revise the title or exclude ice cores from the dataset.

The dataset needs to be revised and polished. You should thoroughly check the dataset to be sure that it follows the common requirements of a dataset.

At least 25 samples do not have coordinates, elevation and year of sampling. 117 samples do not have coordinates and year of sampling. I doubt that such values without any spatial and temporal references could be useful. If it is not possible to obtain the metadata, they should be excluded from the dataset. By the way column "Sample label" suggests location for some of the samples without coordinates. For example, 753 Molodezhnaya, 754 Amery-G1, 755 GM7, 756 GM10, 757 GM13, 758 Dome C, 759 Mirny, 760 Pioneerskaya, 761 Vostok 1, etc. Probably you can use it after careful check.

You need to define parameters in the dataset. It is not clear what the difference between

published and calculated distance is and why did you need to calculate it? The same relates to elevation.

It is also not clear how the quality flag was assessed.

Does "Firn temperature or surface air temperature" relate to exact location and time of sampling or is it somehow averaged? How was it calculated or assessed? What is "Accumulation of snow/ice per year" and how was it estimated?

Data in columns should be formatted in a single style. For example, column "Averaging length (years or depth)" contains different data in very different style that prevent easy processing and analysis of the dataset. I suggest splitting the column into two different ones ("Averaging years" and "Averaging depths") and putting only numerical values in each of them. If needed additional explanation you could add another "Comments" column with text.

Column "Sampling date" contains not dates but years in different formats (both numeric and text) that prevent processing and filtering. You may consider splitting the column into two different ones – "sampling year start" and "sampling year finish".

Column "Sample type" have errors in writing that prevent grouping samples by types. You should carefully check every type and provide exact number of points of every type in the article.

Lines 68-74. Are you talking about the dataset by Masson-Delmotte et al. (2008), mentioned earlier, or your dataset? Clarify it and if it relates to the dataset by Masson-Delmotte et al. (2008), provide more details about its actual content rather than "it provides an observational basis".

Lines 78-79 Provide sufficient references of "numerous new samples and measurements that have been acquired by different researchers"

Lines 79-81 Add quantitative estimation of the "additional observations" and described in numbers the difference between MD08 and your dataset.

Lines 90-91 How many data points did you make publicly available for the first time? It is one of the most important things to show the value of your dataset.

Lines 119-121 Add numbers of points to the figure caption

Lines 118-123 Consider merging figures 1 and 2. You can add information from Fig.2 to Fig.1 (b)

Line 177-178 Include the same figure for δD , %o

Line 214 Since you have several files in Supplement you should reference more precisely.

Line 252 Content of the Word file with Supplement differs from the supplement described here. You should provide detailed description of the supplement files.

Technical corrections:

Affiliations and even country names have different formats

Line 174 Rewrite "...we do not quantitatively calculate the quantitative relationship..."