

Comment on essd-2022-61

Anonymous Referee #1

Referee comment on "A repository of measured soil freezing characteristic curves: 1921 to 2021" by Élise G. Devoie et al., Earth Syst. Sci. Data Discuss.,
<https://doi.org/10.5194/essd-2022-61-RC1>, 2022

The manuscript «A Repository of 100+ Years of Measured Soil Freezing Characteristic Curves» by Devoie et al. describes a new data base pooling published and unpublished measurements of soil freezing characteristic curves, while also providing a repository to report such measurements in the future. I very much like and support this effort and recommend publication after a number of comments and weaknesses are addressed.

Major comments:

- The critical point on data uncertainty is left to an accompanying paper, Devoie et al., 2022b, in prep., which is not available at this point? As a potential user of the database, having some quality assessment of the different data sets in the data base (or even data uncertainty in the best of worlds) is critical. As an example, if one wants to look into the freezing point depression and a measurement at -0.09 degree C is reported, it is important to know if the temperature sensor has an accuracy of 0.01 K or 0.1 K. I assume that such information is extremely hard to come by in most cases, but outsourcing this uncertainty part to an external manuscript (which is not even available yet) is not a good solution. So ideally, the authors should include uncertainties or quality assessments in the database wherever possible, and discuss this briefly in the manuscript. If such data quality assessment cannot be provided for any of the data, this should at least be mentioned, as it is a highly important limitation for some applications. On a general note, the ESSD paper should contain all aspects related to data use, and uncertainty is a rather important aspect. This still offers the possibility for a companion paper on the measurement techniques and related uncertainties.
- The csv file for metadata contains doi's and author + year, but not the full reference. Are all published studies also referenced in the main paper, so that the citation can be

found there? If yes, it would be good to mention this somewhere, so that the authors of the original studies can receive due credit if their data are downloaded from the database and used. It could be mentioned e.g. under "data availability" that you encourage to additionally refer to the original publication whenever possible. If no, it would be good to provide a list of the references of all published studies somewhere so that it can be linked to the data in the database. Potentially related to this, in SFCCMetadata.csv I am not sure about the last column "file". Are these pdf's available as a collection somewhere, or is this only an internal file path? In case of the latter, consider replacing it with the full citation!?

- Add a section on how new entries to the database will be quality-controlled in the future. Otherwise, measurements of questionable quality may be added in the future and weaken the usability of the database. See Minor comments.
- I.184: I can't access <https://github.com/egdevoie/SFCCRepository>, this gives an error for me. It may be a problem on my end, but worth checking.

Minor:

I am unsure about the title. Is the fact that 100 year old measurements are included really that important? As a user, I would mostly care about having the best possible measurements, no matter how old they are. A measurement conducted 100 years ago should have yielded the same result for the same soil as a measurement conducted today, so unlike quantities that do change over time (e.g. climate-related data), the "length of the time series" does not play a role here. But I leave this point up to the preference of the authors.

I. 3: very minor point: "the soil ice content" is also determined by the water content before onset of freezing (i.e. water plus ice content). So something like "partitioning of soil water and ice" would be more precise.

I. 54: it would be good to add a sentence on dissolved ions and other molecules in the soil water which depress the freezing point and can concentrate in the brine upon freezing, leading to further depression of the freezing point of the brine.

I.61: add year in reference

I.150: I do not agree with the causality. If the last water to freeze is also the last water in drying soils, this only means that the low-temperature end of the SFCC can be approximated by looking at drying soils. For the rest of the SFCC range, this is not necessarily true, and it is in fact contested by some studies (e.g. overview in Karra and Painter, 2014).

I.170: are you sure about 1, and not rather "can exceed the porosity of the dry soil"?

Appendix B: it would be good to add a field on "known errors/uncertainties", e.g. if authors are not sure whether the position of the temperature and soil water sensors may have shifted during measurements. This could be essential to quality-control new entries to the database, see Major comments.

References

Painter, S. L., & Karra, S. (2014). Constitutive model for unfrozen water content in subfreezing unsaturated soils. *Vadose Zone Journal*, 13(4).