

Earth Syst. Sci. Data Discuss., referee comment RC3
<https://doi.org/10.5194/essd-2021-268-RC3>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on **essd-2021-268**

Anonymous Referee #3

Referee comment on "An hourly ground temperature dataset for 16 high-elevation sites (3493–4377 m a.s.l.) in the Bale Mountains, Ethiopia (2017–2020)" by Alexander R. Groos et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-268-RC3>, 2021

The manuscript "**A multiannual ground temperature dataset covering sixteen high elevation sites (3493–4377 m a.s.l.) in the Bale Mountains, Ethiopia**" by Groos et al presents a dataset from 29 ground temperature loggers that were installed at 16 sites and operate since 2017. The data was measured from 2 to 50cm depth, using single-channel miniloggers, UTL-3 Scientific Dataloggers providing high resolution and accuracy temperatures, and tempmate.®-B2 providing low accuracy temperatures. The measurements were obtained at hourly intervals. The data that is made available in the repository Zeondo is relevant for research and applications, mainly due to the scarcity of ground temperature observations from mountains and highlands in Africa. It has the potential to be used for soil climate characterization, geomorphic dynamics analysis, as well as to locally validate climate modeling or thermal remote sensing imagery. However, the dataset shows several relevant gaps that occurred due to logger malfunction. These gaps were filled by simple statistical techniques based on correlation. Both the original and the full infilled series are presented. The series were characterized using statistical measures, including some widely used indexes in environmental studies, but the results are not analysed in depth. The manuscript is, in general, well organized, clear and the language is precise. I must, however, note that English is not my native Language.

Despite the overall clarity of the manuscript, I think that there are some significant issues that the authors must address to improve the manuscript and the applicability of the dataset before it may be accepted for publication.

The coding system used to identify the data is cumbersome. It is based solely on the logger IDs (given as numbers, which are in some cases, non-consecutive spatially). I strongly suggest using an ID system allowing to identify the location, depth of measure and the type of logger. This would facilitate data analysis.

The characterization of the soil at the different sites where the instruments were installed is highly needed and the manuscript needs this information before its acceptance. This

issue is of most importance, both for the analysis of the datasets, but also because data from different sites are used to estimate missing temperature values.

The procedures used for data-filling must be described in more detail. In particular, all correlations, p-values and also scatterplots of soil temperatures between the sites (reference vs corrected) should be presented. The limitations of the procedures should be also described, especially in what concerns to the distance between sites and the potential impacts of the types of vegetation cover, soils and topography in the data. It seems that some depths were infilled with regression using temperature profiles from other sites. How were soil type and moisture differences considered for the procedure? These are factors that may strongly influence temperature change with depth and dynamics and they are especially relevant in environments with ground frost.

I also have some more specific comments:

- In the title, the data period (start-end) should be indicated.

- The abstract should be improved. For instance, in line 9, the authors mention that the series provide new insights, but what they say is too vague. They should concretize some of the results of the temperature analysis that they have done.

- In the last paragraph of page 3, It should be mentioned the type of climate according to Köppen classification, and it should be mentioned why the wet season is bimodal. It is confusing to mention that there are two rainy seasons and only one dry season. It could be mentioned that this is a transitional regime and that the two peaks in precipitation follow the movement of the sun and it can also be indicated the importance of the Intertropical Convergence.

- In page 14, line 8, it doesn't make sense the mention of the endemic giant root rat. The reference should be contextualized.

- Figure 1 could be improved if the type of font/type/size of letters discriminate the information that they relate to.

- in fig 3, as mentioned before the acronyms of the loggers should be modified in order to simplify the reading.

