General response

We would like to thank the associate editor for obtaining three valuable reviews and the anonymous referees for their thorough and constructive comments on our manuscript. We are pleased that all three referees are convinced of the usefulness of the data set and generally support the publication of the manuscript if their concerns are properly addressed. On the following pages, we respond to the reviewers’ comments point by point. The reviewers’ comments are highlighted in grey and the responses in white. We hope that the responses qualify us to submit a revised version of the manuscript.

Response to Referee 3 (RC3)

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 temperarures. 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.
We thank the reviewer for his constructive feedback and his contribution to improve the manuscript.
Despite the overall clarity if 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.
We address all concerns point by point below.
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.
Thanks for the comment. The loggers were originally numbered in the order they were installed. Since we did not know in advance which locations would be suitable and could be revisited on a regular basis, the numbering seems unsystematic. We did not want to change the numbering later on

as this could lead to confusion in the long term. However, we agree with all reviewers that a systematic numbering would help the potential end users. We therefore decided to renumber the loggers on a geographical basis – more or less from northwest to southeast. Furthermore, we now use one number per site and differentiate between different depths using the letters h(igh), m(iddle), and l(ow) as suggested by reviewer 2. The type of logger could already be identified before by the first two characters: GT = Geotest dataloggers; TM = Tempmate dataloggers.

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.

Thanks for the hint. We included some general information on the ground/soil type in Section 2 "Study area": *"Most of the basaltic and trachytic rocks, especially on the plateau, are covered by a fine regolith layer that is only sparsely vegetated (see Fig. 2). In areas where soils have developed, Andosols are the most widespread soil type (Lemma et al., 2019)."*
Moreover, we added the respective information in the Section 3.2 "Ground temperature monitoring" and Section 3.3. "Data post-processing".

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.

Some of the requested information (e.g. R^2 and RMSE) were already provided in the supplementary file "Information_Sheet_Data_Gap-Filling.ods" in the Zenodo repository (see Section 6 "Data availability"). We expanded this table and included also the distance to the predictor logger as well as the individual linear regression models (i.e. equations). Moreover, we created scatterplots (measured vs. modelled ground temperature) for all time series that were gap-filled (see new Fig. B1). Another figure (R^2 and RMSE vs. distance to predictor logger) was included in the Appendix (see new Fig. B2). It shows that the RMSE tends to slightly increase with increasing distance between the predictor and target logger while R^2 is also very high when a logger from a further location is chosen for the gap-filling procedure. Figure B3 was added to provide examples of gap-filled hourly ground temperature time series at different depths. As we used statistical (simple linear regression models) for the gap-filling procedure, local differences in the ground type and moisture are considered implicitly to a certain degree. However, we didn't measure ground moisture and, thus, are not able to characterise local differences. Nevertheless, the high R^2 of more than 0.8 at most sites (see Fig. B1) suggests that the local differences are of minor relevance (apart from maybe loggers TM07m and TM11t).

I also have some more specific comments:
- In the title, the data period (start-end) should be indicated.

We refrain from indicating a static observation period in the title as some measurements start in

2017 and some in 2018. Moreover, the low-cost loggers were collected while the scientific dataloggers continue measuring.
- 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.
We highlighted some of the results in the abstract more specifically.
- 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.
We are not convinced that a reference to the climate classification of Köppen-Geiger makes sense here as it does not include an appropriate category for the tropical mountains (in Eastern Africa). However, we revised the section regarding the rainy seasons accordingly.
- In page 14, line 8, it doesn't make sense the mention of the endemic giant root rat. The reference should be contextualized.
We revised the sentence to provide a bit more context: *"In the Bale Mountains, distributed meteorological, ecological, and ground temperature data are of particular interest to better understand the relationship between spatial ground property variations and the scattered distribution of Erica trees across the Sanetti Plateau (Miehe and Miehe, 1994; Lemma et al., 2019; Mekonnen et al., 2019) as well as the scattered occurrence of endemic giant root rats (Tachyoryctes macrocephalus) that spend most of their time in large underground nests in the afro-alpine belt (Vlasatá et al., 2017)."*
- Figure 1 could be improved if the type of font/type/size of letters discriminate the information that they relate to.
We are not quite sure if that's what you meant, but we coloured the labels of the weather stations and data loggers to match the colour of the respective symbols in the map.
- in fig 3, as mentioned before the acronyms of the loggers should be modified in order to simplify the reading.
All logger acronyms are revised according to the reviewers' suggestions (see comment above and responses to reviewer 1 and 2)