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 our responses qualify us to submit a revised version of the manuscript.

Response to Referee Comment 1 (RC1)

The manuscript of Groos et al. presents a dataset on ground temperature data: At 5 + 14 locations in the Ethiopian highlands, ground temperature was monitored at three depths (5) or near the surface (14 stations) for nearly three years. The manuscript contains a first analysis of the data with focus on the occurrence of frost, annual and seasonal temperature variations. I am not an expert in the area of ground temperature dynamics. Thus, my comments are mostly restricted to the overall content of the paper, its structure and technical details. That being said, the authors convincingly lay out the potential use of the data set. Moreover, the logistic and technical challenges of the respective measurements in this remote area underline its uniqueness.
We appreciate that the reviewer acknowledges the usefulness and uniqueness of the presented ground temperature dataset from the Ethiopian Highlands.
However, with effectively only three station providing true temperature profiles (albeit 3 depths only), and timespan of just 3 years and many data gaps (although reasonably-well filled), the data set is not excessively rich. Therefore, I strongly suggest to include the concomitant measurements at the meteorological stations into the dataset. The paper often relates to them (e.g. ll), and even includes their analysis (p. 12)., as they greatly enhance its value or seems even mandatory (as stated in the Discussion) for its analysis. I was even somewhat surprised not to find them. As for meteo-data of secondary provider that cannot be included, at least the respective reference would of great help to the potential user.
We would like to stress that 5 sites (and not only 3 sites as stated by the reviewer) provide “true” temperature profiles, although some longer gaps in the time series had indeed to be filled. In comparison with datasets from the lowlands or other continents, the presented data set may “not [be] excessively rich”, but in view of the logistical and technical challenges related to measurements in remote and high-mountain environments, we think the presented ground temperature dataset (comprising multi-annual time series from different depths and elevations between 3500 and 4400 m) from this data-scarce region is comprehensive enough to be of interest for the wider research community. Apart from Kilimanjaro, a similar dataset does not exist for any other high mountain on the entire continent.

We agree that the meteo data from the automatic weather stations are of benefit for the interpretation and analysis of the ground temperature data, but we cannot publish the meteo data along for multiple reasons: 1. The data from the weather stations are not automatically transferred from the study site to the database of the research unit and some of the data still need to be saved and postprocessed. 2. The meteo data available so far are currently analysed by PhD students within the research unit. Once the meteo data are processed and analysed (we don't know when this will be exactly the case), they will be made publicly available as well. However, before the final publication of the meteo dataset, personal access could be granted on request to researchers who are interested in the data. We added a respective statement regarding the data availability in the text (see Section 6 Data availability, last paragraph).

The core data of the publication are available from a Zenodo repository. However, they seem incomplete in some aspects, while redundant in others. I recommend some restructuring:

- Naming the folders "raw data" (i.e. log files) and "processed data" would seem more intuitive to me.

- Some files (e.g. Hourly_Ground_Temperatures_Corrected.csv) have a deviating formatting of the date column. Highly impractical for automatic use / scripts

- Information_Sheet_Data_Correction.ods: For each logger, add a column "flag" with indicator(combinations) for "interpolated to full hour", "missing/filtered", "interpolated/gap filled based on logger nn"

This potentially makes the other files (except, perhaps, Information_Sheet_Data_Interpolation) obsolete.

- The duplication of *.txt and *.ods seems unnecessary. I suggest using txt for the actual data, ods for the meta-data

- Please add a GIS file (or at least table with the station coordinates), the information of Table 1 plus further site attributes (vegetation, etc.)

- Please add some reference to the ESSD-manuscript (e.g. with an overall readme.txt and/or in the Zenodo desciption field)

Thanks for the helpful advice. We have made the following changes according to your suggestions to restructure and modify the repository:
- we renamed the mentioned folders to "raw_data" and "processed_data"
- we removed all redundant or unnecessary files (e.g. "Hourly_Ground_Temperatures_Corrected.csv" and "Information_Sheet_Data_Correction.ods")
- we combined all post-processed data into one file (i.e. table) and added a "flag" column to provide additional information on the treatment of each hourly measurement of each time series (see comment further below)
- we included a readme-file
- we included a GeoPackage (GIS file) with the requested attributes
- we added a reference (DOI) to the ESSD-manuscript

The overview data analysis in the paper provide an adequate first insight into the data. As this is not

the core of a data publication, I recommend not extending them. Please consider shortening by excluding the analysis of the effect of slope and orientation. However, I made some suggestions how to make improve some points, if these are considered essential. Some conclusions drawn should probably be formulated more carefully, as they do not seem to be fully backed up by the analysis.
We decided to keep the analysis of the effect of slope and orientation in the discussion as this section is important to understand the spatial variations of ground temperature in the Bale Mountains. Moreover, this section is relatively short anyway. We considered the other recommendations to rephrase some of the conclusions more carefully.
The formal quality of the paper is high, with logical structure, adequate style and (mostly) helpful figures. As a technical suggestion I strongly urge to decrease the file size of the resulting PDF - 20 MB for a text document is inacceptable. I assume this comes from photos with excessively high resolution.
Thanks for the hint. We noticed the large file size too late. Figures and photos have been compressed (where possible bitmaps were converted to vector graphics) in the revised version to reduce the file size of the PDF to a few MB.
The dataset and paper merits publication, especially when supplemented with the meteo-data. However, I recommend moderate revision before. Further details in the annotated PDF.
We outlined before why we are not unable to publish the meteo-data along. Please find below our responses to the annotations in the PDF.
Page 1, Line 6: Why "climate" in this context? The current dataset concerns soil temperatures. Would "temperature observations" be more appropriate? Does "on the continent" mean "in Africa"? If so, I suggest to state this explicitly.
We replaced “climate observations” by “temperature observations” and “on the continent” by “on the African continent”
Page 1, Line 7: Please add monitored depths for temperature.
Monitored depths were added.
Figure 1: The upward-pointed triangle is commonly used to demark peaks. Please consider another symbol for the weather stations
Symbol for weather stations was changed (from triangle to square)
Page 4, Line 6: sorry to be nitpicking: Would't "cold phase" be more appropriate than "cooling", as those formations testify the former, not (necessarily) the latter?
Yes, “cold phase” would be more appropriate. Revised accordingly.
Page 5, Line 16: What is "pan" in this context?
“logger-pan-to-USB cable” was replaced by “logger-to-USB adapter»
Page 5, Line 30: Please add some kind of overview or table to illustrate the selection and choice of locations. Could also be added e.g. as some colour code to the labels of Fig. 3
Sorry, but we do not understand this comment as an overview of the distribution of the GT and TM data loggers is already provided in Fig. 1 and Table 1.
Page 5, Line 32: specify the temperature range where this RMSE was obtained

The comparative measurement was performed at 12~°C and at 4~°C in a fridge over seven hours with logger GT00 as reference. The temperature range for the accuracy of the TM data loggers is -10 °C to 65 °C as stated in Section 3.1 “Data loggers”. We added this information here as well.
Figure 2: Panels a) and b) do not seem to convey any necessary information apart from the general impression of the landscape. Consider removing or reducing to one. c) and d) seem redundant, too. It would be more informative how the readily installed site looked like, i.e. what (if any) disturbance was necessary to access the loggers for downloading the data. I suggest removing e) in favor for an enlarged version of f), showing only the logger and the measuring tape.
Panel (a) shows the characteristic relief and vegetation of the Sanetti Plateau and panel (b) shows the northern slope of Mount Wasama with its sparse vegetation cover. We think that such overview photos are essential for potential end users to get an impression of the afro-alpine environment in which the ground temperature loggers were installed. However, we restructured panels (c-f) following your suggestions. Panels (c-e) show a sequence of the logger installation process. Panel (f) was enlarged.
Figure 3: Please indicate which loggers continue measuring, e.g. by arrows at the end of the bars. In line graphs, try including vector graphics in favor of bitmaps.
We added arrows to indicate which loggers continue measuring. The bitmaps (Figures 3-5) have been replaced by vector graphics. Figures 6-7 were not replaced by vector graphics, but the resolution was reduced as well to reduce the size of the PDF.
Caption Table 1: I suggest to write "29" instead (as with all number above 12).
Revised accordingly.
Page 9, Line 10: I cannot see this period of interpolation in figure 3. I would expect the bar of GT07 having a grey part at the start. Please check.
We reserved the grey colour for “real” data gaps in the time series. The first part of the time series of GT05t (former GT07) was indeed adjusted to account for the wrong installation depth. Since this adjustment only slightly affected the temperature amplitude and not the long-term ground temperature variations, we do not think it’s necessary to emphasize this moderate modification in Fig. 3.
Page 9, Line 17: I suggest not to use the term "interpolation" here, as this is different from actual interpolation of "adjacent" values. Please consider using the introduced term "gap filling" consistently.
We replaced “interpolation” by “gap filling” throughout the manuscript
Page 9, Line 34: meteorolgoical => meteorological
Corrected
Page 10, Line 2: I don't understand the tenses used in this sentence. The data had not been available so far, but are available now, I expect. Please clarify.
We rephrased this sentence for clarification. The meteorological data from the Bale Mountains are shared within the research unit we are part of for internal usage, but they have yet neither been published nor made publicly available as some PhD students are still working on the data: “The meteorological data are stored in an on-demand processing database system - currently with

restricted access to the members of the joint Ethio-European Research Unit 2358 "The Mountain Exile Hypothesis" (Wöllauer et al., 2020). The data will be made publicly available at some point in the future (see Section 6)”
Page 10, Line 26: I suggest to prepare a more informative plot on the performance of the *gap filling procedure*: installation depth vs. R2 (or RMSE). Colour the resulting dots by the length of the time series reconstructed this way.
We included two additional plots in the appendix (B) of the revised manuscript to better illustrate the performance of the gap-filling procedure. One shows scatter plots of measured vs. modelled ground temperatures for each gap-filled time series (see Fig. B1) and the other shows R^2 and RMSE vs. distance between the predictor logger and gap-filled logger (see Fig. B2).
Page 10, Line 29: I suggest to locate the cross-comparison exercise to the beginning of this section, since consistency of TMs and GTs is a prerequisite for all subsequent steps. I expect this was performed on non-gapfilled data? Please confirm.
We shifted the concerned paragraph to the beginning of this section. Yes, the cross-comparison was performed on non-gapfilled data. We added this information here.
Page 10, Line 31: Please report RMSE and length of time series used in this comparison.
The RMSE was 1.7 °C and the time series spanned four month. Both information are now included.
Page 11, Line 4: For all depths?
This statement is based on measurements from a depth of 2 cm. We added this note.
Page 11, Line 7: please specify, which stations are these (Logger IDs)
We specified the stations (GT03, GT04, GT05) in the revised version.
Page 11, Line 10: specify the logger-IDs
We specified the logger location (GT02) in the revised version.
Page 11, Line 22: I agree that data analysis is not the focus of this paper. Still, providing a plot of T_ground vs. T_air and colouring the dots by season could help to illustrate the described phenomenon. In the current way, this all very plausible, but cannot really be seen by just looking at the data.
As ground temperatures are in principle controlled by meteorological parameters as well as ground properties (e.g. porosity) and conditions (e.g. moisture), a bivariate statistic (e.g. T_ground vs. T_air) does not provide much insight in the processes governing the ground temperature variations. We think a multivariate statistical analysis would be interesting/necessary, but that’s beyond the scope of this manuscript.
Page 11, Line 26: Clarify at which depths.
2 cm. Clarified in the revised version.
Page 11, Line 28: Replace "superficial" with "near-surface"
Replaced.
Page 11, Line 33: Would it be possible to give an estimate on the maximum frost penetration depth based on the gradient between 2 and 10 cm? If so, I suggest to state this value here.
We used the two complete time series at 2 and 10 cm at location GT05t (southern Sanetti Plateau)

to estimate the maximum frost penetration depth assuming a linear ground temperature gradient (see new Fig. C2). Based on this simple analysis, we can conclude that frost usually does not penetrate deeper than 5 cm into the ground. The mean frost penetration depth is about 3 cm.
Page 11, Line 35: Please use protected whitespace between number and unit to prevent unfortunate linebreaks.
Thanks, we tried to use protected whitespaces, but we forgot it here. It’s now corrected.
Page 12, Line 3: tend => tends
Corrected.
Page 12, Line 4: replace "constant" with "(near) zero"
Done.
Page 12, Line 6: I suggest to treat the altitude gradients om meant T first, as they concern the primary data. After this, move to profile gradients and seasonal varaibility.
We restructured the paragraph accordingly.
Page 12, Line 12-14: hard to understand, consider rephrasing.
Rephrased
Page 12, Line 20: replace "long-term" with "annual". I cannot make any inference on diurnal amplitude from the given plots. If you want to make this point, consider plotting the (smoothed) daily temperature range of the stations over time.
We replaced “long-term” with “annual” and added a plot (Fig. 7d) to illustrate differences in the diurnal amplitude.
Page 12, Line 30: Please state the intended update interval of the repository.
We note in the revised version that we intend to update the repository on an annual basis.
Page 13, Line 5: Please clarify if you mean "denser compared to solitary measurements" (which I suppose) or "denser than the network presented here".
We clarified in the revised version that we mean a “network consisting of numerous loggers”
Page 13, Line 9: replace "final" with "post-processed
Modified accordingly.
Page 13, Line 12-13: "sporadic permafrost" seems to be a contradiction in terms. Rephrase.
“Sporadic permafrost” is an established term in the scientific community for landscapes where permafrost represents 0/10 to 50 % of the total area. So it has a spatial and not a temporal meaning. https://ipa.arcticportal.org/publications/occasional-publications/what-is-permafrost
Page 13, Line 14: replace "order" with "range"
Modified accordingly.
Page 13, Line 20: Neither *surface* measurements nor an analyses of diurnal temperature amplitude have been shown.
“at the ground surface” was replaced by “ground temperature amplitude near the surface” We added two exemplary plots of the diurnal temperature amplitude in the revised manuscript and appendix (Fig. 7c and Fig. B3).

Page 13, Line 24-27: I don't see how your analyses "reveals" this: While this assumption is very plausible, without measurements of net radiation and soil moisture this remains a hypothesis. Please correct.
We weakened our statement and replaced “reveals” by “suggests” in the revised version.
Page 13, Line 27-28: not shown
We added a respective figure (Fig. 7c). See previous comment before.
Page 13, Line 31-33: References to respective studies or examples would strengthen this claim.
References to respective studies were added
Page 13, Line 35-36: I am very skeptical how even a very dense network of temperature stations alone should be able to provide high resolution maps of precipitation and (air?) humidity. Please weaken this statement.
This was a misunderstanding. We didn’t intend to claim that maps of precipitation or air humidity can be inferred from temperature data (alone). We just wanted to provide a selection of potential methods that can be used to create maps from numerous point data. We modified the sentence as follows: “Geo-statistical and machine learning techniques have been applied in other studies to create high-resolution maps of temperature, precipitation, and humidity for Mount Kilimanjaro on the basis of meteorological data from different sites”
Page 14, Line 3-5: As these will be highly site-specific, I find this claim quite optimistic. I'd rather envision regionalizing soil temperatures (or profiles) with the help of remote sensing. Please comment.
We did not claim that the statistical relationship between air and ground temperature at one site (e.g. the Bale Mountains) is transferable to other sites (in the Ethiopian Highlands). We just wanted to outline the following: “For mountain regions where both air and ground temperature are measured simultaneously, their statistical relationship can principally be used to generate air temperature maps from remotely-sensed land surface temperatures”
Page 14, Line 8: remove quotes.
Quotes were removed.
Page 14, Line 29: this seems to be "gap filled" rather than "interpolated"? File is possibly obsolete (see extra remarks)
“interpolated” was replaced by “gap-filled” (throughout the manuscript) We merged the files “Hourly_Ground_Temperatures_Corrected" and “Information_Sheet_Data_Correction” to one file (“Hourly_Ground_Temperatures”) using a flag to provide additional information regarding each measurement (i.e. if it was interpolated to the full hour, gap-filled, corrected, removed, etc.). See Section 6 “Data availability)
Page 15, Line 2-3: "...with the data download date encoded in the filename". Please specify delimiter character.
We added the suggested information.
replace "superficial" with "near surface"
The sentence was rephrased (see comment below).
Having measured at a few dozen sites, I find it a bit strong to deny the existence of permafrost

"anywhere" in the EH. Please consider weakening your statement or strengthening this claim by better presenting that you monitored all/the most frost-prone sites existing.
We weakened the statement as follows: “Moreover, the data confirm the frequent occurrence of nocturnal ground frost in the afro-alpine belt to a depth of about 5 cm (Fig. C2). However, the mean annual ground temperatures of more than 7 °C on the highest peaks suggest that permafrost in the southern Ethiopian Highlands is either absent or restricted to isolated patches.”
strong claim. The performance of regional climate models will hardly be evaluated on ground temperature data.
The sentenced was rephrased as follows: “The dataset may serve a wide range of scientific applications, ranging from the validation of remote sensing products and the modelling of spatial ground temperature variations to the investigation of certain natural processes such as the formation of periglacial landforms or the geographic distribution of species that live underground such as the giant root rat.”
Figure 5e: Replace line plot by bar plot, as daily prcp is not a continuous variable
The line plot was replaced by a bar plot.
Figure 6, headline first column: replace "annual" with "whole year"
“annual” was replaced by “full year”
Figure 6, Caption, Line 2: add "dry" and "rainy", respectively.
“dry” and “rainy” was added in the caption
Figure 7, Caption: replace "temperature variations" with "smoothed temperature". Add dashed line to legend.
The caption was modified accordingly and the meaning of the colour-coded dashed lines was included in the figure caption as well.