Author's response to comments of Referee #1

AC: We thank the anonymous referee for the detailed review of our manuscript.

RC1: General comments

The discussion paper describes a modelling scheme to inversely estimate the near-surface atmospheric thermal states from the thaw depth information under permafrost conditions, primarily employing the idealized relationship derived from the Stefan formula, and discussed the applicability of the model to infer the past thermal conditions from the relict periglacial features, namely, active layer thickness. The authors demonstrated the efficiency of the scheme to inversely estimate the temperature characteristics, such as mean annual air temperature, and mean air temperatures of coldest and warmest months, from the active layer thickness observed at the Antarctic and Arctic sites.

The relationship expressed as the Stefan solution, Equation (1) in the text, is widely known for its useful simplicity but also tendency for biases when applied to real observations, as partly stated in the discussion paper. Still, the author proposed a new and intriguing idea to apply the relationship as a modelling framework to infer the paleo-thermal conditions at the formation time of the currently relict periglacial feature, which can be very relevant to geoscientific modelling within the scope of GMD, as well as of paleoclimatology and cryosphere-related science. In the current form of the paper, however, the explanation and evaluation of the modelling scheme are confusing, or poorly written in term of model description, and the title and the target of the paper show substantial mismatches with the current structure of the paper. Thus, the reviewer believes that the manner conducting the model evaluation, as well as overall organization of the presented text need substantial revisions (appreciating the effort that the manuscript went overall rearrangements for a model description paper), as well as the way the model application was conducted and presented in the Result section.

AC: We agree with the referee and admit that there is a mismatch between the title and content of the original version of the manuscript. Originally, it was meant to be a proof-of-concept study showing that the model performs well on present-day data, providing its best possible validation, which was to demonstrate that it could also reasonably derive past temperature conditions, but now we recognize that its real application on palaeo-periglacial features is necessary. Consequently, we also intend to include in the revised version of the manuscript a palaeo-air temperature reconstruction using a palaeo-active-layer thickness and to compare its outputs with reconstructions based on other proxy records and/or model products. This will also bring changes of the manuscript structure, which will be done in accordance with the referee's suggestions given here and below.

RC1: Two major suggestions are:

1. Move the subsection 5.4 to the Introduction section to describe the previous studies, or motivation of the study, and reconstruct the whole text to fit to a description paper of a model to be used for paleo-temperature reconstructions.

AC: Agreed. This subsection will be removed from the discussion and its parts will be incorporated into the introduction of the revised version of the manuscript, which will also be made less general and more related to the aims of the manuscript as suggested.

RC1: 2. Add a model validation case. Use only those terms and variables that would be available and used in paleo application cases (for example, seting P to 365 days, and Aa as described in 5.2.3, etc.), and see how the computed atmospheric thermal states compare to the observed. Also, sensitivity tests for parameters (eg, physical properties; thawing n-factor) to evaluate the range of variations in the computed temperatures would be very informative addition to discussions in 5.2.

AC: We intend to include in the revised version of the manuscript a palaeo-air temperature reconstruction using a palaeo-active-layer thickness and to compare its outputs with reconstructions based on other proxy records and/or model products. A section containing present-day data will be retained in the revised version of the manuscript in order to provide model validation using analogous data available for the palaeo-cases and to perform sensitivity tests.

RC1: Specific comments
Abstract:
The title and abstract claim that this paper introduces a presented model is to be used for paleo-temperature reconstructions. The reviewer feels that the authors' intention and the current structure and the way the model was run and evaluated have a large gap, and suspect that the current evaluation against modern temperature records could not serve properly to judge the model's ability when applied to the past periglacial features. The performed evaluation seemed to use information that could not be available for the paleo cases.

AC: As stated above, originally, it was meant to be a proof-of-concept study, but we intend to include in the revised version of the manuscript a palaeo-air temperature reconstruction using a palaeo-active-layer thickness and to compare its outputs with reconstructions based on other proxy records and/or model products. Still, a section containing present-day data will be retained in the revised version of the manuscript in order to provide model validation using analogous data available for the palaeo-cases and to perform sensitivity tests.

RC1: P. 2, ll. 46-48: This comment is related to the above one. If the paper is intended to introduce a model, it should show clear structure of the model, preferably with a simple schematic diagram, to show, for example, what are the input variables to the model; what are the parameters to be set or assumed; and what are the output variables that the model produce. The current paper appears merely to demonstrate a Stefan-based calculation scheme using the observed temperature and relevant data. Thus, if it is intended to evaluate the model to infer air temperature characteristics from past periglacial evidence, it should demonstrate the way the model would perform when applied to paleo cases (See the related comments in the "Result" section). An example of a modelling scheme for paleo application would be like:
[Input] active layer thickness
[Parameters to be determined, assumed, or deduced] thermal conductivity (thawed), wetness, thawing n-factor, length of the period (fixed at 365 days), annual air temperature amplitude
[Outputs] thermal conditions and related information (MAAT, MATW/CM, MATT/FS, Lt, Lf…)

AC: Please note that Table 1 in the original version of the manuscript shows what variables are inputs (upper section) and what variables are outputs (lower section) and it is incorporated in the section 2.1 describing driving parameters of the model.
As stated above, originally, it was meant to be a proof-of-concept study, but we intend to include in the revised version of the manuscript a palaeo-air temperature reconstruction using a palaeo-active-layer thickness and to compare its outputs with reconstructions based on other proxy records and/or model products, which should also clearly show how the input parameters should be chosen. Still, a section containing present-day data will be retained in the revised version of the manuscript in order to provide model validation using analogous data available for the palaeo-cases and to perform sensitivity tests.

RC1: P. 4, l. 84 (and others), "amplitude":
This is merely a suggestion. Use of a word "range" to denote a margin between the lowest and highest values, maybe useful to distinguish *Aa* and *Aa*/2 in the text.

AC: Agreed. A collocation "annual air temperature range" will be used in the revised version of the manuscript instead of "annual air temperature amplitude".

RC1: P. 5, ll. 98-106: Estimation of MAAT and *Aa* appears the key, or the central part of the model when applied to the paleo settings when no a-priori thermal knowledge is available. Under the current modelling framework, when one assumes a sinusoidal annual temperature change, and has the *Ita* value as the area under the curve for the positive values, MAAT and *Aa* can be determined independently. In the current form, it is not clear if *Aa* is a parameter or an output variable in the model. So, it definitely needs more elaboration to describe how to calculate (or estimate) MAAT and/or *Aa* (with this in mind that this modelling scheme is to be used for the paleo applications).

AC: Please note that numerous other palaeo-air temperature reconstructions, for instance those based on glacier mass-balance modelling, also frequently utilized present-day climatology (that is, including present-day annual air temperature range) combined with MAAT and/or precipitation perturbations (though these were frequently not listed as model driving parameters) or adjusted annual air temperature range to derive most plausible palaeo-climate scenarios. Consequently, we believe that the presented scheme is meaningful.
Unfortunately, MAAT and A_a cannot be determined independently because I_{ta} depends on both of them, that is, I_{ta} can achieve identical values for various combinations of MAAT and A_a. For instance, I_{ta} equals 526 °C d if MAAT and A_a is -4 °C and 20 °C, respectively, but also if it is -8 °C and *ca*. 29.8 °C. Consequently, A_a must be a model parameter in order to estimate MAAT and other air temperature characteristics. However, please note that A_a does not define their precise values. Moreover, it is believed to undergo substantially lower temporal variations than, for instance, MAAT, and thus A_a for palaeo-applications may be approximated by its present-day value at a site with relict periglacial features or it may be adjusted based on other regional proxies. We will stress it more in the revised version of the manuscript.

RC1: P. 5, ll. 105-106, "However, potential drawbacks can be easily handled if exclusively permafrost-related features are examined.": It is not clear what is intended to say.

AC: Since the solution is designed to be used in permafrost environments, problems may occasionally arise in situations where seasonally frozen ground coexists under negative mean annual air temperature (MAAT) because permafrost–seasonal frost boundary rarely coincides exactly with MAAT of 0 °C. So the model should be applied on those features that indisputably formed in the presence of permafrost. Nonetheless, the statement will be changed in the revised version of the manuscript in order to be more understandable.

RC1: P. 6, Figure 2. It would be very user-friendly to describe how to draw relevant information from the figure, for most of the GDM reader won't be familiar with this nomogram.

AC: We will extend the caption and briefly describe how to draw information from the figure in the revised version of the manuscript.

RC1: P. 7, Table 2, and "Model validation": Please comment on the applicability to "Berry Hill slopes" in the text (other James Ross sites appear on the flat locations).
Applicability of the Stefan solution may be limited to those sites with high vertical heterogeneity, and it is mentioned in the text for the Alaskan site. Applicability to sites with large lateral flows of heat or water would also be limited.

AC: First of all, it should be noted that for consistency we stick to long-established site names on James Ross Island. Admittedly, the name "Berry Hill slopes" may be a little misleading because the site actually occurs on a

slightly-inclined surface in the foothills of the "Berry Hill", which has a gradient of 5–10° (Hrbáček, F., Nývlt, D., Láska, K.: Active layer thermal dynamics at two lithologically different sites on James Ross Island, Eastern Antarctic Peninsula, Catena, 149, 592–602, http://dx.doi.org/10.1016/j.catena.2016.06.020, 2017), and thus lateral flows of heat or water are supposed to be limited or small. We will mention it in the revised version of the manuscript.
We agree that the applicability of the Stefan solution may be limited at sites having high vertical heterogeneity and those experiencing a certain degree of lateral flows of water or heat. However, we believe it is necessary to test the model in various settings, including those that does not perfectly meet its assumptions, and evaluate its actual performance there. Actually, the minority of locations where the Stefan solution has been applied to estimate the thickness of the active layer can be considered as "ideal".

RC1: P. 7, ll. 128-129, "The Stefan equation": Does it mean Eq. (1)? If so, please add the notation (similar to p. 10, l. 212). Also, it is not clear, what are the difference between the results of what this sentence means, and what the later demonstration of the model in "Result" section.

AC: Yes, it means Eq. (1). The corresponding notations will be added in the revised version of the manuscript.
The sentence was to state that earlier thaw-depth estimates using the Stefan solution (not estimates of air temperatures based on active-layer thickness) were among the most accurate ever on James Ross Island, while those from the Alaskan sites were among the worst, and just for this reason we believe that these locations are well suited to evaluate the model performance. It will be changed in the revised version of the manuscript in order to be more understandable.

RC1: P. 8, ll. 162-163: It is curious if the results from the "successive wet and dry weighing" and the TDR probes are consistent to each other, or were independently done and not comapred (eg, Alaskan sites were solely done by the former, and the James Ross by latter). Is it possible to mention the representative of the results to be applied to the entire thawed layer?

AC: Only one of these methods was employed at each study site. Alaskan sites were solely done by successive wet and dry weighing, which was also applied at Abernethy Flats and Johann Gregor Mendel sites on James Ross Island, while the rest of the locations, that is, Berry Hill slopes and Johnson Mesa, was done by a calibrated TDR probe. We will consider using solely the outputs of successive wet and dry weighing in the revised version of the manuscript because it is now available from all the sites on James Ross Island.
Please note that the representativeness of the measurements was already documented by earlier publications cited in the original version of the manuscript (Zhang, 1993; Romanovsky and Osterkamp, 1997; Hrbáček et al., 2017a; Hrbáček and Uxa, 2020), which estimated active-layer thickness there.

RC1: PP. 8-10. Although not clearly written, it seems that the results shown in this section used some of the temperature information obtained from the observations (for example, P and Aa as shown in Table 2, which claims "model-driving" parameters). If the purpose of the paper is to demonstrate the ability to reconstruct the temperature conditions derived solely from the geomorphological evidences (that is, depth of the active layer), the evaluation should be done in the same manner as to be intended for the paleo cases. This means to run the model with the input (depth) and assumed parameters (thermal conductivity, wetness, thawing n-factor) only. Otherwise, it appears just a mere application of the calculation scheme using advantage of the present-day observations.

AC: As stated above, originally, it was meant to be a proof-of-concept study, but we intend to include in the revised version of the manuscript a palaeo-air temperature reconstruction using a palaeo-active-layer thickness and to compare its outputs with reconstructions based on other proxy records and/or model products, which should also clearly show how the input parameters should be chosen. The model validation based on present-day data will be

done in an analogous manner.
Please note that numerous other palaeo-air temperature reconstructions, for instance those based on glacier mass-balance modelling, also frequently utilized present-day climatology (that is, including present-day annual air temperature range) combined with MAAT and/or precipitation perturbations (though these were frequently not listed as model driving parameters) or adjusted annual air temperature range to derive most plausible palaeo-climate scenarios. Consequently, we believe that the presented scheme is meaningful.

RC1: Subsection "5.1 Model uncertainties,…":
This subsection needs to restructure the organization, and clearly reformulate sentences. There are many long sentences with unclean meaning (for example, ll. 222-224, 230-233, 235-237). Also, the discussion sometimes goes back and forth, right and left, with reservation and euphemism. One suggestion for a re-organization would be to first divide the discussion to "strength of the model" and "weakness of the model", and prioritize the issues under each of the categories.

AC: The discussion section in the revised version of the manuscript will be reorganized as suggested and sentences will be reformulated to make them clearer.

RC1: P. 10, l. 214, "Also, it assumes that the frozen layer is at 0 ^{O}C before thaw.": This looks pre-assumed in the derivation: that is, what is at stake is the temperature at the freezing interface between the thawed and frozen layers, which should be 0 °C.

AC: Please note that the Stefan equation does not consider heat conduction below the freeze-thaw plane and as such assumes that temperature is uniformly at 0 °C in the frozen zone (see e.g. Romanovsky, V.E., Osterkamp, T.E.: Thawing of the active layer on the coastal plain of the Alaskan Arctic, Permafrost and Periglacial Processes, 8, 1–22, https://doi.org/10.1002/(SICI)1099-1530(199701)8:1<1::AID-PPP243>3.0.CO;2-U, 1997; Kurylyk, B.L.: Discussion of 'A simple thaw-freeze algorithm for a multi-layered soil using the Stefan equation'by Xie and Gough (2013). Permafrost and Periglacial Processes, 26, 200–206, https://doi.org/10.1002/ppp.1834, 2015). Nonetheless, it will be changed in the revised version of the manuscript in order to be more understandable.

RC1: P. 11, 225-227: It seemingly depends on whether MAAT or MATWM (or MATTS) is at question (cf. Figure 3).

AC: Agreed. It will be changed in the revised version of the manuscript in order to be more understandable.

RC1: P. 11, l. 231: It is not clear what is meant by "which"

AC: It should relate to "the Stefan equation tends to deviate", but now we see that it can also ambiguously relate to "the peat-layer thickness in the active layer". It will be changed in the revised version of the manuscript in order to be more understandable.

RC1: P. 11, l. 236: Not a clear sentence. Are the sites "far from being saturated" those in James Ross Island, and sites of "two-layer" those in Alaska? What (or which) does "there" actually mean?

AC: Rather, it was supposed to be a general statement that the model performed well on present-day data even though some of the validation sites do not perfectly meet the assumptions for the application of the Stefan formula. It will be changed in the revised version of the manuscript in order to be more understandable.

RC1: "5.2 Driving data": Maybe a source of confusion in reading the discussion paper is that it is not clear what are the input (driving) data, what are the parameters or boundary conditions that set the calculations, and what are the output of the model to be applied for the paleo setting. What is discussed in this subsection is either parameters (ie, ground physical properties, thawing n-factor) or a part of the output (temperature amplitude), otherwise it does not make sense if the target temperature information is a driving data.

AC: Table 1 in the original version of the manuscript shows what variables are inputs (upper section) and what variables are outputs (lower section). Values of the input parameters should be obtained directly from relict periglacial structures, while those that cannot be obtained directly should be derived using empirical relations (transfer functions) or should rely on representative published data that allow a meaningful range of their values to be defined. Since we also intend to include in the revised version of the manuscript a palaeo-air temperature reconstruction, it should clearly show the above procedure.
Please note that numerous other palaeo-air temperature reconstructions, for instance those based on glacier mass-balance modelling, also frequently utilized present-day climatology (that is, including present-day annual air temperature range) combined with MAAT and/or precipitation perturbations (though these were frequently not listed as model driving parameters) or adjusted annual air temperature range to derive most plausible palaeo-climate scenarios. Consequently, we believe that the presented scheme is meaningful.

RC1: P. 13, l. 278, "it is possible to assess the extremes, between which the moisture likely occurred": Not clear what is meant.

AC: This should mean that it is possible to determine the range of values, which moisture could reach at the time when periglacial features formed. It will be changed in the revised version of the manuscript in order to be more understandable.

RC1: P. 14, l. 288, "these correlations": which correlations between what? Please describe clearly.

AC: This should mean correlations between thermal conductivity and other ground physical properties. It will be changed in the revised version of the manuscript in order to be more understandable.

RC1: "5.3 Implications for paleo-temperature reconstructions": It would be suggested to rename the subsection title, something like "Applicability to periglacial features".

AC: The subsection name will be changed or the whole subsection will be included into another one in the revised version of the manuscript, given that a palaeo-air temperature reconstruction will be added.

RC1: P. 15, ll. 329-331, "we hypothesize that their depth probably rather reflects the position of a transient layer where the contact between the active layer and the uppermost permafrost at the time of their formation oscillated": It is not clear what this sentence is meant.

AC: Transient layer is a transition zone that alternates in status between seasonally frozen ground and permafrost over sub-decadal to centennial time scales because of natural climate variability (Shur, Y., Hinkel, K.M., Nelson, F.E.: The Transient Layer: Implications for Geocryology and Climate-Change Science, Permafrost and Periglacial Processes, 16, 5–17, https://doi.org/10.1002/ppp.518, 2005), and this is what past periglacial features likely attest to. It will be changed in the revised version of the manuscript in order to be more understandable.

RC1: P. 15, l. 331 "the latter": It is not clear what is indicated.

AC: "the latter" was to be related to the transient layer mentioned in the previous sentence. It will be changed in the revised version of the manuscript in order to be more understandable.

RC1: P. 15, l. 347, “random-sampling methods”: how the methods work in the context? With no information, it is not possible to judge the adequacy of the methods.

AC: Agreed. Random sampling will be used to generate representative sets of input parameters for a palaeo-air temperature reconstruction, which will be added into the revised version of the manuscript.

RC1: “5.4 Progress over previous attempts”: the content of this subsection should be placed as “motivation” or “previous studies” in the Introduction, and the whole paper should be structured to “introduce and evaluate” a model to infer air temperature from the paleo-periglacial feature. It is strongly suggested that the overall organization and structure of the paper should be revised.

AC: This subsection will be removed from the discussion and its parts will be incorporated into the introduction of the revised version of the manuscript.

RC1: The authors gave proper credit to related work and clearly indicate their own new/original contribution. And the number and quality of references appear appropriate.