# Replies to referee comments on “Investigating the thermal state of permafrost with Bayesian inverse modeling of heat transfer”

*The Cryosphere*, `10.5194/egusphere-2022-630`

**RC:** ***Referee’s Comment***, AR: Authors’ Response, ☐ Manuscript Text

## 1. Response to referee 1

We would first like to thank the referee for their time and effort in providing valuable feedback to improve our work. While we acknowledge the referee’s general criticisms regarding the study design, we believe that our method, as well as the data that we have available, are strong enough to still draw meaningful conclusions about how the thermal state of permafrost is responding to long-term changes in climate conditions. We provide detailed responses to the referee’s individual comments below.

### 1.1. Issue 1: Research design and uncertainties

**RC:** ***The research design has some issues making it unclear if the main conclusions are attributed to the physical processes or the modeling uncertainties. First, the available depths of borehole data are not the same. At the two colder sites (Samoylov and Barrow, Fig 4b and 4c), both sites have deep borehole data although Barrow does not have shallow borehole data. In contrast, neither of the two warmer sites (Fig 4d and 4e) has deep borehole data. This could be the main reason causing the much larger temperature variability (Fig 4d and 4e), more scattered relationships in Fig 5, and more observed uncertainties in Fig 6. Therefore, the majority of conclusions made by comparing colder and warmer sites are not convincing. One or more warm sites with deep borehole data are needed to validate this study’s conclusions. It is also worth performing the inverse modeling again on the Samoylov site excluding its deep borehole data to see if its thermal behavior stays the same or changes toward the warm sites.***

AR: We acknowledge that the disparity between available borehole temperature measurement depths is a concern. It is important to note, however, that this is not an intentional aspect of the study design but rather a limitation of the available data. It is common for researchers and practitioners to install automated temperature sensor instrumentation in the upper one meter of the ground since this is generally achievable without heavy drilling equipment. High quality, automated instrumentation of deep boreholes is unfortunately relatively rare and instead research teams typically collect manual measurements once per year (often in the summer when the borehole can be easily located). The data from the Barrow North Meadow Lake site featured in this work are an example of such measurements. As mentioned in the text, these annual measurements cannot be compared to mean annual temperatures recorded in instrumented boreholes (such as the other three sites) above the depth of zero annual amplitude (ZAA) due to the effects of seasonal variation. This is why we only use the manual measurements from Barrow at 20 m and below, as this is presumed to be deep enough that seasonal variation should be negligible.

Despite this limitation, we argue that our conclusions are justified for three primary reasons:

Firstly, the availability of deeper measurements should largely only affect uncertainty in the initial temperature profile at the beginning of the simulation period (i.e. at the year 2000); this is because deeper measurements help better constrain the range of plausible temperature profiles after the spin-up period (1979-1999). The impact on later years where observations are available in the upper $10\,\mathrm{m}$ will necessarily be less significant since, after the first $5$ to $10$ years, the climate signal will dominate the initial condition.

Secondly, one of the main reasons why we limit the analysis of energy contents to the upper $10\,\mathrm{m}$ is because this is the range in which all sites (excepting Barrow) have measurements available. While it is true that the temperature profile at the beginning of the simulation period would have some impact on the resulting distribution of observed trends, we would expect this effect to be mostly limited to temperature (i.e. a wider range of initial temperatures would spread out the distribution along the x-axis in Figure 5 from the manuscript). It should not affect the underlying relationship between temperature and latent heat, which is the central interest of this study.

Lastly, while it is true that the availability of deeper borehole measurements will affect the resulting spread of temperature predictions across the ensemble, we do not agree that this weakens the conclusions drawn from comparing the cold and warm sites. On the contrary, it is actually a strength of our method (and Bayesian methods more generally) that the posterior distribution meaningfully reflects uncertainty due to differences in data availability between sites and therefore allows us to make inferences despite these limitations of the available data.

To validate our arguments here, we followed the suggestion of the referee and ran an additional set of simulations for the Samoylov site with the measurements below $10\,\mathrm{m}$ omitted from the inference procedure. The results are presented below. It is clear that the primary impact of omitting the deeper borehole observations for Samoylov on long-term change in the energy partitions (Fig. R1) is on the sensible heat content. This is due to the additional uncertainty in the initial temperature profile at the beginning of the simulation period (Fig. R2). The effect on latent heat, as well as the relationship between changes in temperature and latent heat and/or permafrost thaw, is negligible (Fig. R3), with the linear relationship between latent heat and temperature staying the same within the bootstrap margin of error: $(2.1 \pm 0.9)\,\mathrm{MJ/K}$ with all sensors depths to $(2.2 \pm 0.3)\,\mathrm{MJ/K}$ excluding depths below $10\,\mathrm{m}$. The change in the linear relationship between active layer thickness and temperature was also negligible: $(0.011 \pm 0.007)\,\mathrm{m/K}$ to $(0.014 \pm 0.002)\,\mathrm{m/K}$.

We propose that this lack of sensitivity to the initial temperature offset is exactly because of (i) the historically cold temperature of deep permafrost at Samoylov Island and (ii) the freezing characteristics of sandy soils which allow for minimal unfrozen water at temperatures well below the freezing point. Thus, even the warmest plausible initial temperatures, given the larger

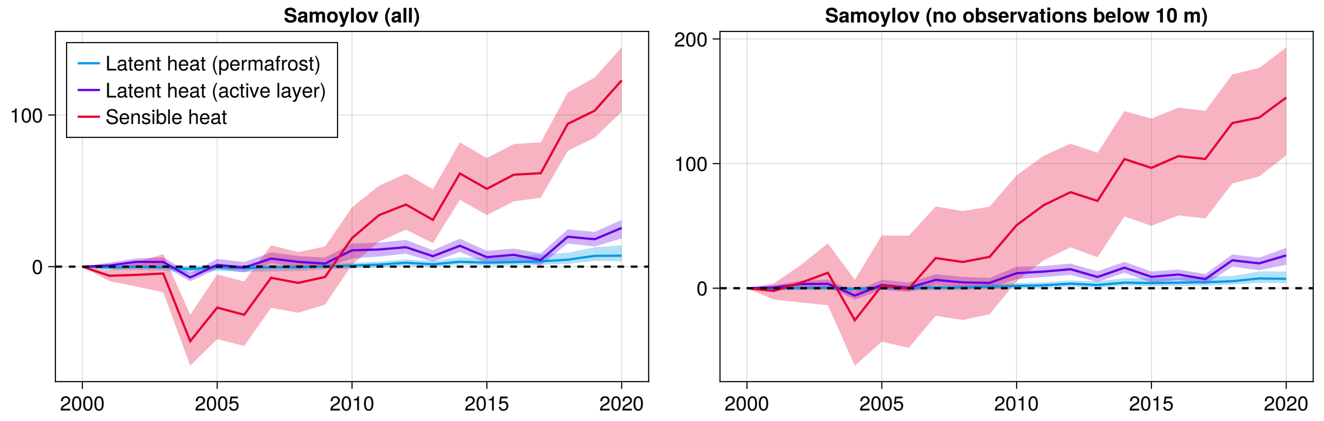



**Figure R1.** Total change in energy partitions (upper 10 m of soil profile) since the beginning of simulation period. Energy is partitioned into three categories: Latent heat in frozen grid cells (i.e. cells with maximum annual temperature $T_{\max} < 0\,°\mathrm{C}$), latent heat in the active layer ($T_{\max} \geq 0\,°\mathrm{C}$), and sensible heat. Solid lines show the median energy change while the shaded regions show the 95% CCI over the ensemble.

uncertainty range, are cold enough that there is minimal change in latent heat. This is consistent with the central argument of our paper: soil characteristics and historical climatology are crucial to understanding what observed changes in ground temperature can tell us about the thermal state of permafrost.

We will add these additional results to the appendix of the revised manuscript, and we will also include additional discussion to highlight the robustness of our results to the discrepancies in available measurements between sites.

**RC:** ***Line 374 seems to demonstrate depth alone cannot explain the variability. However, the statement is not strong because 82 cm is too small on a 10 m scale. Also, the observations of Bayelva also have less variability than those of Parson's Lake, which likely explains the less variability in the modeled temperature at Bayelva.***

AR: We agree with the referee's assessment here and have removed this assertion from the text.

**RC:** ***The authors do have a full section 5.6 to discuss the limitations. While these limitations are definitely important, the current research design is not strong enough to support the conclusions even neglecting other uncertainties.***

AR: We believe that the additional results presented in figure R1-R3 validate the study design and support the central arguments of our paper. The limitations detailed in section 5.6 are, as the referee states, important. However, as also argued in the main text, we believe that the model and study design are still strong enough to support the main conclusions.

**RC:** ***Secondly, section 5.3 discusses the role of surface conditions on ground warming based on the n-factor change before and after 2005. Again, uncertainties can be the main driver because no borehole observations are available to constrain the model before 2005. This is another key point made based on the comparison of two data not having the same conditions.***

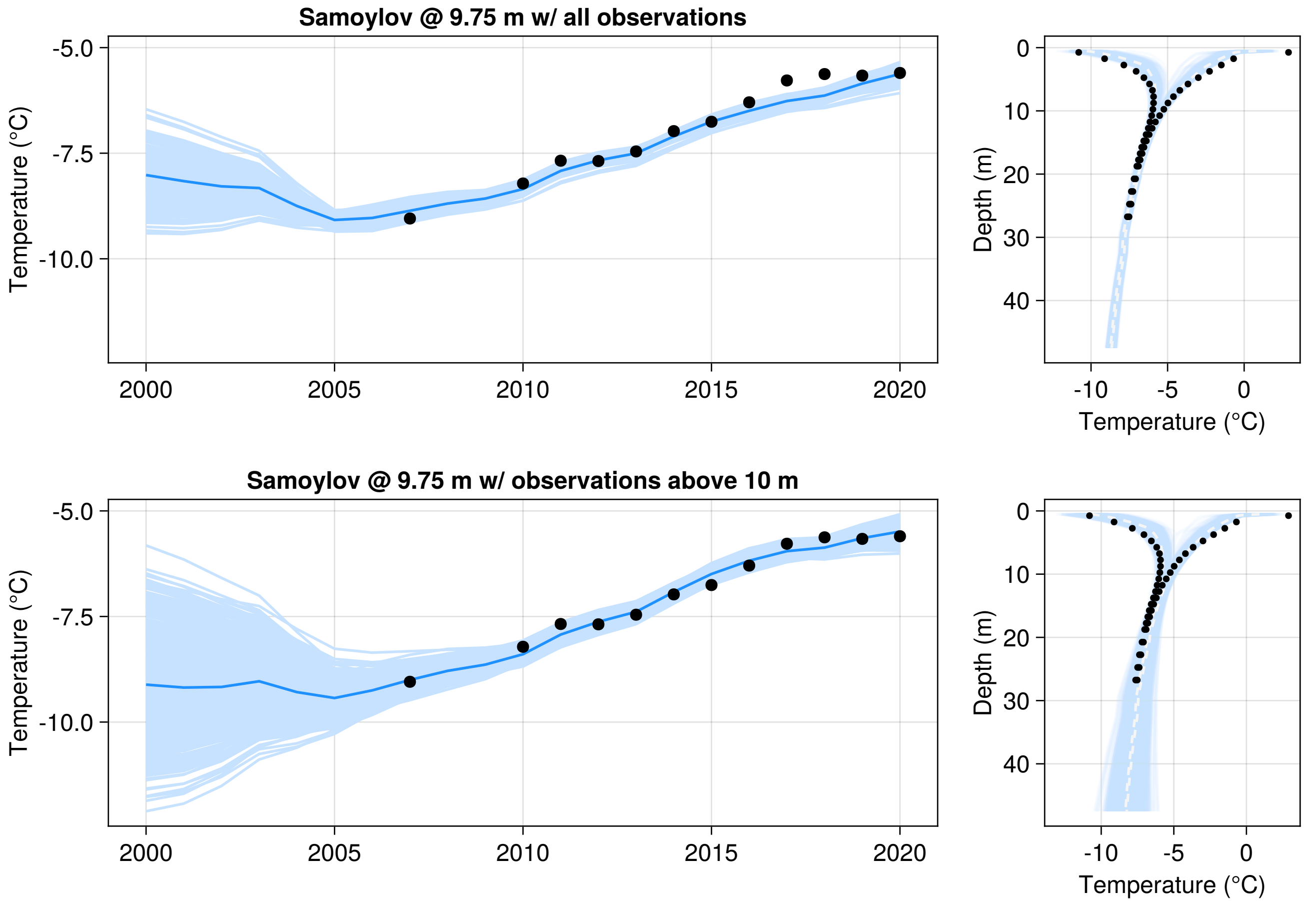
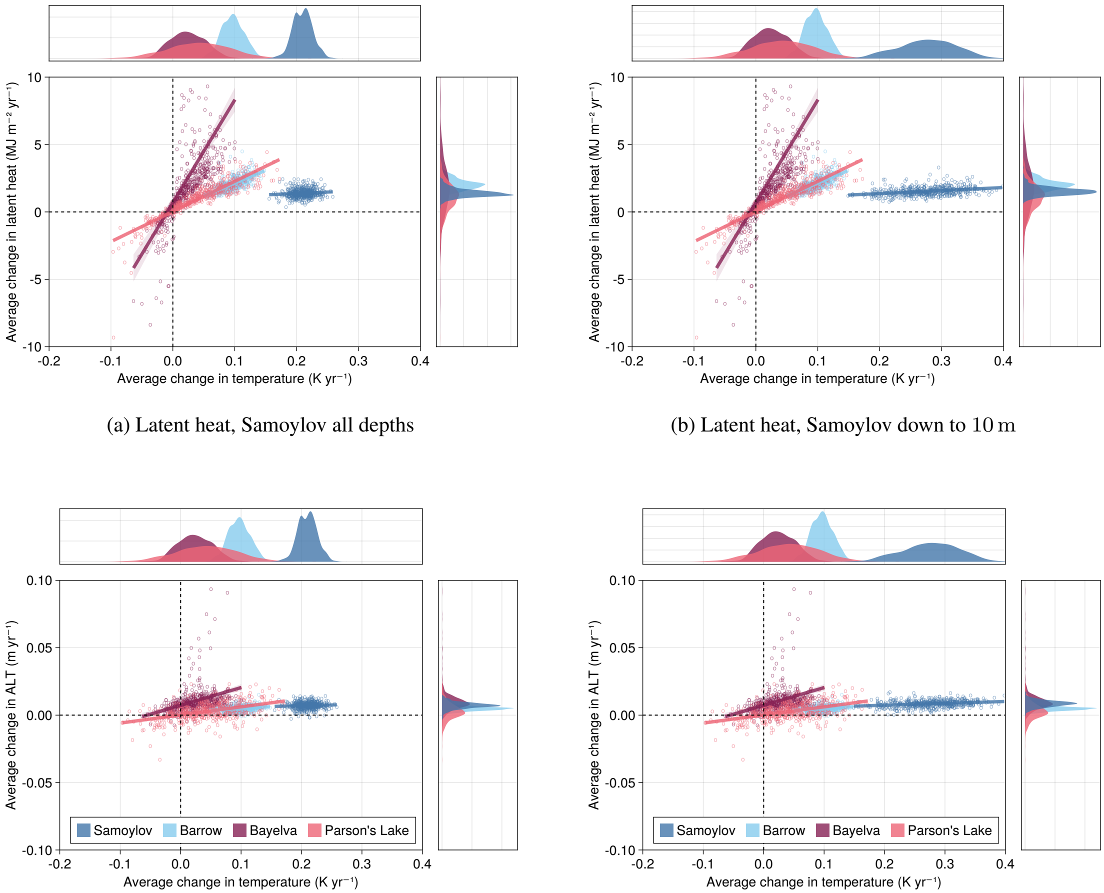



**Figure R2.** Mean annual ground temperatures at $9.75\,\mathrm{m}$ depth (the deepest sensor after truncation) at Samoylov along with temperature profiles in the final simulation year (2020) for all ensemble members. The blue line shows the ensemble mean. The dotted white line on the temperature profile plots shows the ensemble median. Note the recent change in the long-term warming trend at this depth within the last five years cannot be resolved by our current model due to the bidecadal n-factor parameterization. This change is actually not due to air temperature but rather changes in snow cover, in particular substantial early-season snowfall in the winter of 2016/2017 (Boike et al., 2019).

AR: We have removed this paragraph from the text since it is, in hindsight, overly speculative given the limitations of the current study design and available data. We do not agree, however, that this is a particularly important point of the manuscript. The discussion here was largely tangential and was intended only to comment on the potential hazards of extrapolating recently observed trends in deep ground temperatures backwards in time.




(a) Latent heat, Samoylov all depths

(b) Latent heat, Samoylov down to 10 m

(c) Active layer thickness, Samoylov all depths

(d) Active layer thickness, Samoylov down to 10 m

**Figure R3.** Joint densities of modeled mean annual change in latent heat (a-b) and active layer thickness (c-d) vs. mean annual change in ground temperature across sites for both the normal Samoylov run with all borehole depths (a,c) and the additional Samoylov run with only borehole observations up to 10 m (b,d). Note that omitting the deeper sensors largely only affects the spread of observed temperature trends and not the relationship with changes in latent heat. In all four plots, a small number of points from Bayelva exceed the upper y-axis limit and thus are not shown.

### 1.2. Issue 2: Method description

**RC:** ***The manuscript has a large space describing the modeling method but most contents are too technical and not accessible to people who are in the cryosphere community but do not have expertise in numerical modeling, inversion, and Bayesian method, etc. The authors focus too much on the advanced topics of the method but completely missed the information on the basic idea of the applied method. Also, in many cases, the authors only cite some references without explicitly describing the methods, which makes the readers difficult to follow or understand.***

AR: We thank the referee for this valuable feedback. Although we already attempted in the original manuscript to keep technical language and details to a minimum, we recognize that this communication gap is one of the primary challenges of interdisciplinary research and that there are certainly further improvements that can be made to the existing text. We respond to each of the referee's specific concerns below.

**RC:** ***Section 3.1. The introduction of Bayesian inference involves too many technical terminologies. Please consider adding supporting sentences to make it easier for people not familiar with the Bayesian method to understand it.***

This section was intended to provide a basic introduction to the ideas of Bayesian inference for those not familiar with such methods, so it is of course important that it is accessible for this audience. We have revised the paragraph in question as follows:

> The Bayesian approach to statistics provides a natural framework for inferring unobserved quantities of interest while simultaneously accounting for their associated uncertainties [...]. This is accomplished ~~by applying Bayes ruleto some observed and unobserved variables, $Y$ and $X$, respectively~~via Bayes rule:
>
> $$p(X|Y) = \frac{p(Y|X)p(X)}{p(Y)} \quad \text{with} \quad p(Y) = \int_{x \in X} p(Y|X=x)p(X=x)dx, \tag{1}$$
>
> which can be seen a generic formula for obtaining the so-called *posterior distribution* of an unobserved quantity $X$ ~~*a posteriori*~~ given observations $Y$ ~~from some sampling distribution or *likelihood* $p(Y|X)$. The prior distribution~~. The *prior distribution* $p(X)$ ~~encodes information about $Y$ known~~ reflects our pre-existing uncertainty about $X$ *a priori* ~~and plays a crucial role in the Bayesian inference workflow.~~ (i.e. before observing $Y$) while the *likelihood* $p(Y|X)$ measures how well the model's predictions agree with the observations, $Y$. In this work, $Y$ are temperature measurements, typically sampled over time and/or space, whereas $X$ are unknown model parameters or unobserved physical quantities such as soil properties, thaw depth, or the ratio of sensible to latent heat. The overall objective is then to obtain the posterior distribution, $p(X|Y)$, of these unknown parameters given the temperature measurements which quantifies not only the best-fitting parameter settings but also the associated modeling uncertainties.

**RC:** ***Lines 140-144. Need to briefly explain the bias correction procedure.***

AR: We have added the following clarifying clause to briefly elaborate on the bias correction procedure of Piani et al. (2010):

> The bias correction procedure for air temperature follows closely the empirical quantile mapping method of [...] in which the empirical quantiles of both the model (reanalysis) and observational data are computed over some reference period (in this study, we use the full time period for which observations are available); the model data are then mapped to the corresponding quantiles of the observations.

**RC:** ***Line 150. Need to briefly explain the numerical procedures and parameterizations of CryoGrid.***

AR: We have added two additional missing pieces of information to the appendix section, namely the parameterization of the thermal conductivity, $k_T(z,t)$, and heat capacity, $C_T(z,t)$, functions. We will also add parameter tables to the supplement enumerating the relevant constituent conductivities and heat capacity parameters which are treated as constants in this study in order to avoid colinearity with the soil composition parameters.

**RC:** ***Section 3.6 This section introduces a key methodology EKS. It presents the advantages of EKS over MCMC and EKI without explaining the basic theory/idea of EKS in the first place. Again, this makes researchers not familiar with EKS very difficult to follow and understand it.***

AR: We will add a few sentences describing the theory behind EKS as well as one or two of the key equations from Garbuno-Inigo et al. 2020.

**RC:** ***Line 250. Need briefly explain what a mean vector from Garbuno-Inigo et al. 2020 is.***

AR: As discussed in the text, the observed temperatures are assumed to be generated according to equation (9), i.e:

$$T_{\text{obs}} = (h_T \circ f)(\boldsymbol{\theta}) + \eta$$

where $f$ is the forward model evaluated at $\boldsymbol{\theta}$, $h_T$ is the mapping function which extracts and aggregates the modeled temperatures, and $\eta \sim \mathcal{N}(0, \Sigma_T)$ is the observational noise. For the purposes of constructing the EKS algorithm, the model predictions can thus be equivalently seen as being sampled from a Gaussian distribution centered at $T_{\text{obs}}$ (this follows from moving $\eta$ to the left hand side of the above equation). Since we assume the observation noise to be independent across space and time, we can flatten the two-dimensional temperature field into a vector which thus constitutes the mean of this Gaussian distribution, hence the term "mean vector". However, we acknowledge that this term is non-standard and possibly confusing; in light of this, we have now rephrased this sentence in the revised manuscript:

> ~~The~~ We use the observed mean annual ground ~~temperatures,~~ temperatures from each borehole site as the observations, i.e. $T_{\text{obs}}$ ~~, are used as the observation mean vector for the Ensemble Kalman Sampler described in [...]~~in Eq. 7, for the Ensemble Kalman Sampling algorithm.

## 1.3. Issue 3: Quantiative analysis in discussion

**RC:** ***The discussion needs more quantitative and specific analysis. When interpreting the results, the authors only briefly propose possible factors without explaining how would these factors impact the results.***

AR: We agree with the referee's assessment and will revise the discussion section accordingly. More specifically, we will add concrete numbers to reinforce the statements regarding model biases in Sec. 5.2 as well as for the statements discussing correlations and effect sizes in Sec. 5.4.

**RC:** ***Paragraph 255. I may miss something but I did not get the purpose of this paragraph. It states that the prior distribution over model parameters is important but does not explain what was done to improve performance.***

AR: The purpose of this paragraph was just to provide some basic motivation. We have rearranged this and the following paragraph to make this more clear:

> The prior distribution over model parameters, $p(\phi)$, ~~is of crucial importance to our method~~also plays a key role in the inversion procedure. Some parameters in the heat transfer model, such as soil composition, will cause the resulting optimization problem on $\phi$ to be under-constrained, since there may be more than one possible combination of soil components which have similar thermal properties. Additionally, incorporating prior knowledge about plausible parameter values allows us to reduce the amount of computational effort wasted on physically implausible ~~or incoherent~~ model configurations that may arise from random sampling.
>
> ~~EKS assumes the $m$ unconstrained parameters $\gamma(\phi) = \psi \in \Psi \subseteq \mathbb{R}^m$ to follow a multivariate Gaussian distribution, $\psi \sim \mathcal{N}(\mu_\psi, \Sigma_\psi)$, where $\gamma : \Phi \to \Psi$ is a bijective function which maps the $m$-dimensional possibly constrained parameters $\phi \in \Phi$ to their unconstrained values on the real line.~~ We define our priors in the constrained parameter space $\Phi$ in order to more easily incorporate physically meaningful information about each site. We define suitable parameter priors for each site based on published field measurements and soil core analyses; full details on choices of priors for each site are in Appendix [...]
>
> EKS assumes the $m$ unconstrained parameters $\gamma(\phi) = \psi \in \Psi \subseteq \mathbb{R}^m$ to follow a multivariate Gaussian distribution, $\psi \sim \mathcal{N}(\mu_\psi, \Sigma_\psi)$, where $\gamma : \Phi \to \Psi$ is a bijective function which maps the $m$-dimensional (and possibly constrained) parameters $\phi \in \Phi$ to their unconstrained values on the real line.

**RC:** ***Line 356. This sentence does not explain why Samoylov has deep soil temperature warming faster than the air temperature. Factors other than air temperature should be included here.***

AR: We agree that the other factors are also important to highlight. This was actually explained further in the following paragraph, but the connection was not necessarily obvious as written. We have revised these two paragraphs to make this point more clear:

> This is consistent with the results of our analysis which show mean air temperature trends ranging from $0.09\,\mathrm{K\,yr^{-1}}$ at Parson's Lake to $0.11\,\mathrm{K\,yr^{-1}}$ on Samoylov Island. ~~The large difference~~ This discrepancy in observed permafrost temperature trends between these two sites, despite similar changes in air temperature, indicates considerable uncertainty in how permafrost is responding to the changing ~~climate. Furthermore, the observation that deep permafrost on Samoylov Island is most likely warming faster than air temperature suggests that changes in air temperature alone cannot always fully explain permafrost warming.~~
>
> ~~These results motivate our inverse modeling study by demonstrating clear, localized differences and substantial uncertainty in how the permafrost thermal regime responds to long-term changes in air temperature. The discrepancies in the apparent relationship between long-term changes in air and permafrost temperatures suggest that other factors are at play, such as surface conditions (e. g. snow cover) and variability in~~ climate that is likely attributable to other factors. For example,

thicker and/or lower density snow cover can accelerate permafrost warming by insulating the ground against rapid drops in air temperature characteristic to autumn and early winter thereby delaying the refreezing of the active layer [...]. Additionally, soil thermal properties ~~. These factors~~such as the bulk conductivity and freezing characteristics due to soil texture can also play a significant role in modulating the effects of surface temperature changes [...]. Both of these factors, among others, can significantly affect energy uptake in the subsurface, and ultimately, the current and future thermal state of permafrost in Arctic regions [...]. ~~We believe~~

The results of this trend analysis motivate our inverse modeling study by demonstrating clear, localized differences and substantial uncertainty in how the permafrost thermal regime at these four sites is responding to long-term changes in air temperature. We argue that this can be at least partially attributed to ~~the latent heat effect, in addition to soil thermal properties, both of which are a major source~~ factors affecting the uptake of latent heat in the subsurface such as soil freezing characteristics as well as historical climatology. We discuss in the following sections how our inverse modeling results suggest that both of these factors are major sources of uncertainty in making inferences about the ~~subsurface thermal regime [...]~~changing thermal state of permafrost.

**RC:** ***Paragraph 360. Besides only presenting the potential factors impacting the soil thermal states, I would include how they impact the thermal states. For example, how does the ground temperature change with air temperature giving increasing (or decreasing) snow thickness and soil thermal diffusivity?***

AR: We have added further discussion in the text to clarify the expected impacts on the thermal state (see the revised text above).

**RC:** ***Line 390. Please explain more about why latent heat is lost so that the temperature is warmer. Please also explain why drainage and evapotranspiration cause latent heat loss.***

AR: As discussed in section 5.5 in the original text, latent heat acts as both a heat sink during thawing and a heat source during freezing. The latter plays a particularly important role during the winter because it slows the propagation of cold surface temperatures thereby delaying the refreezing of the active layer (Romanovsky and Osterkamp, 2000). When unfrozen water is removed from the active layer due to drainage or evapotranspiration, the latent heat stored in this water is removed as well, and thus so is the heat source. The result is that cold temperatures can propagate faster in the winter-time since a higher fraction of this energy is diffused as sensible heat.

We will also include this additional explanation in the revised text.

**RC:** ***Section 5.6. It would be helpful if include some discussion about the expected changes after addressing each limitation.***

AR: We will add supporting sentences in this section to highlight the expected effect of addressing each limitation.

**RC:** ***Minor comments...***

AR: We thank the referee for these corrections. We have made the necessary revisions to the manuscript, including moving figure B1 to the main text.

## A. Additional comments from the authors

In the process of revising our simulation code to follow-up on the referee's concerns, we discovered some unrelated bugs in the configuration of the prior distributions for some of the parameters (specifically, the freeze curve parameters for the Parson's Lake site and the saturation level parameter for all sites). We fixed these errors and re-ran the simulations for all sites with the same random seed as used in the original simulations. We also increased the size of the ensemble from $256$ to $512$ to improve the robustness of our results as well as to facilitate the sensitivity analysis requested by the second referee.

## References

Boike, J., Nitzbon, J., Anders, K., Grigoriev, M., Bolshiyanov, D., Langer, M., Lange, S., Bornemann, N., Morgenstern, A., Schreiber, P., Wille, C., Chadburn, S., Gouttevin, I., Burke, E., and Kutzbach, L.: A 16-Year Record (2002–2017) of Permafrost, Active-Layer, and Meteorological Conditions at the Samoylov Island Arctic Permafrost Research Site, Lena River Delta, Northern Siberia: An Opportunity to Validate Remote-Sensing Data and Land Surface, Snow, and Permafrost Models, Earth System Science Data, 11, 261–299, https://doi.org/10.5194/essd-11-261-2019, 2019.

Piani, C., Weedon, G. P., Best, M., Gomes, S. M., Viterbo, P., Hagemann, S., and Haerter, J. O.: Statistical Bias Correction of Global Simulated Daily Precipitation and Temperature for the Application of Hydrological Models, Journal of Hydrology, 395, 199–215, https://doi.org/10.1016/j.jhydrol.2010.10.024, 2010.

Romanovsky, V. E. and Osterkamp, T. E.: Effects of Unfrozen Water on Heat and Mass Transport Processes in the Active Layer and Permafrost, Permafrost and Periglacial Processes, 11, 219–239, https://doi.org/10.1002/1099-1530(200007/09)11:3<219::AID-PPP352>3.0.CO;2-7, 2000.