General assessment

This study uses a pedon-scale soil physical model - called DynSoM-2D – to describe how small-scale heterogeneity affects soil thermodynamics and surface heat fluxes in a specific type of permafrost topography (non-sorted circles), with a particular focus on the effect of heterogeneity-induced lateral heat fluxes. This is an important topic, since most land surface models fail to represent subgrid-scale heterogeneity in permafrost-affected areas, thereby leading to an inaccurate representation of permafrost dynamics and biogeochemistry. The approach, tools and simulation design employed in this study are highly relevant for addressing these questions. The discussion is particularly interesting and provides an insightful interpretation of the results. The code is clear and well commented, making it easy to read.

Unfortunately, there are some major problems. First, the quantitative analysis is based on a single year, which is insufficient for deriving robust conclusions about the effect of soil heterogeneity on soil temperature. While the authors briefly explain their decision to conduct a single-year analysis (l.193), the results are highly dependent on the specific atmospheric conditions of that year. More robust results could be obtained by including several years in the analysis (e.g. a 10-year climatology or a longer timeseries), especially given that this modelling study is not limited by a lack of data.

Second, (inter- and intra-) model ranges are highly sensitive to extreme values and are not absolute metrics. They lack a reference point, such as a mean or median, against which they can be compared (for example, an increase in the intra-model range does not have the same impact on permafrost physics and biogeochemistry at high temperatures or close to 0°C). Although extreme values can be informative, other statistical metrics (e.g. mean/standard deviation or median/interquartile range) are a more robust choice for describing the results.

Finally, as this is a new model that has not been described elsewhere, a much more detailed description is required, particularly with regard to snow, hydrology and the calculation of soil thermal properties. In particular, the description lacks the key equations that underpin the results presented in this study. Some specific points are detailed in the "Specific Comments".

Overall, this is an interesting study that contributes to a growing body of literature on an important and challenging topic. It would definitely be a valuable contribution for the land surface modelling community. However, major revisions are required in a number of areas.

Specific Comments

- The title is quite technical, making it difficult to guess what the article is about at first glance (for instance, it does not mention "soil heterogeneity"). I do not have any other suggestions, but the article would benefit from a clearer title.
- l.101 : The principles to which the authors are referring to are unclear. Maybe just keep the following sentence (l.102) or combine both ?
- l.108-109 : Could you give the soil layer thickness for all layers ?
- l.111-112 : Bulk density is missing from the list (needed in pedotransfer functions of Wösten at al. (1999)).
- l.115 : What about the other soil parameters besides soil porosity ?

- l.118 : The variables driving the model should be specified. Also specify how rainfall and snowfall are separated from total precipitation.
- l.140 : Please provide the equations used to calculate soil thermal conductivity and capacity as these parameters largely control soil temperature dynamics.
- Given the crucial role of soil water content and frozen/unfrozen fractions on soil temperature, the soil hydrological scheme should be described (including boundary conditions).
- l.156 “The heat is recalculated to a skin layer/surface temperature” is not very precise and should be clarified.
- l.160-169 : The equations governing snow cover dynamics should be provided, along with the values of snow conductivity, capacity and any other parameters controlling snow dynamics.
- l.174 : For clarity, please recall that the soil column extends below 1m. If I have understood the code correctly, the soil column is 50m deep, but this should be clearly explained in the manuscript.
- l.175 : Which soil properties are used at depths below 1m ?
- Fig.1 : The vertical axis needs a unit/label.
- l.178 : Vegetation growth is excluded but it is unclear whether vegetation is present in the simulations. If so, please describe the type of vegetation present.
- Table 2 : The percentages on the left do not add up to 100% for A-B.
- l.188 : The authors should provide evidence that the soil temperatures and soil moisture are close to equilibrium after the 50-year spinup (with no long-term drift).
- l.191-193 : Analysing a single year does not produce robust quantitative results, as these depend on the atmospheric forcing of that particular year (e.g. the amount of precipitation influencing both snow cover and soil thermal properties via soil moisture). The aggregation error could be different if a different year had been chosen. Why not use a longer time period ? I do not understand the need to “preserve the current atmospheric signal” (l.192). The results would be more robust if a longer period was included in the analysis, for example daily or monthly averages over 10 years, or the analysis of a longer timeseries if the authors prefer not to use time averages.
- l.201 : Please mention how θ^{het} is calculated. Horizon-wise average ?
- l.203 “In general, the horizon-wise averaged heat conductivity is lower in the heterogeneous soil than in the homogeneous soil” → This statement is not supported by the data in Table 3 where $\kappa^{het} > \kappa^{hom}$ for both the subsoil and the total soil.
- l.208-210 : I do not think it can be concluded from Fig.A1c&e that “the calculated albedo is hardly affected”. The authors should provide data to support this statement.
- l.234 : What does “per day” mean ? I suppose it refers to daily values but this should be clarified.
- l.255 : “at the simulated distance of 1m” → unclear what this refers to.
- l.257 : “After freezing” → unclear, please precise.
- Using DOY in Fig.2 and months in Fig.3 makes it difficult to compare the two figures. Adding months to Fig.2 would improve understanding.
- l.276 : “reflecting differences in soil moisture content that affect surface albedo” → albedo data are missing to support this statement.
- l.300 : What does “from below” refer to ? The freezing front ?
- Fig.4 : A statistical test, such as a t-test, is required in order to assess the significance of soil temperature differences.
- Fig.4 : The difference in soil temperature is likely to be dominated by snow and ice in winter, resulting in values close to zero from February to June. Maybe focusing on the period when soil temperature differences are significant (July to January) would help to assess the effect of

OM distribution on these differences, and would avoid pulling the median down artificially due to the near-zero ΔT_{soil} during half of the year ?
- Section 3.4 lacks quantification. Statistical metrics and tests should be used.
- l.315-317 : “By aggregating...lower OM content” → not sure if this is significant. A statistical test is needed.
- l.317-318 : “this relationship...from M2 to M0” → based on which metrics ?
- l.356-357 : This is counter-intuitive as OM generally insulates the ground against warmer air temperatures in permafrost regions (Zhu 2019, Loranty 2018). However this can be explained by the higher thermal conductivity in columns with a higher OM content, due to higher soil moisture, as the authors explain. A figure showing the 2D evolution of soil moisture (vertical and horizontal, similar to Fig.3) would support this statement.
- l.375-376 : “The only time when snow actually has an effect on simulated surface heat fluxes is at the end of the snow melt period” → rather on the difference of surface heat fluxes between homogeneous and heterogeneous configurations than on surface heat fluxes themselves (which are always impacted by snow) ?
- l.409 : “which will be consistent for any aggregation error.” → I do not think this conclusion can be drawn from a single-site study. Such a conclusion would require simulations for other sites, as the amplification of the aggregation error by lateral fluxes probably depends on the spatial distribution of soil texture and OM content.
- l.443-446 : I do not understand why the amplification of the aggregation error by lateral fluxes provides more confidence that these results would remain for other sites.
- l.448 : In the context of land surface modelling, there are no global maps of soil texture and OM at a 10 cm resolution. Could you briefly describe how these results could help improving land surface models ?
- l.453 : “we assume that it will remain at other sites and under changed environmental conditions” → I do not think this assumption can be made based on a single-site study.
- Fig. A3 and Fig.A4 are not referenced in the main text. They should either be included or removed.

Technical corrections

- Bonan 2019 is cited for various aspects of the model and it would be helpful to provide more specific references for each citation (e.g. refer to the relevant chapter in the book).
- l.105 : Please give the section number.
- l.123 : “phase change” → water phase change
- l.126 : Please refer to the specific chapter/equations from Bonan (2019).
- l.131-134 : Please precise the units.
- l.136 : Please give the section number.
- l.138 : Please precise the units.
- l.154 : Please refer to the specific chapter/equations from Bonan (2019).
- l.171 : Please specify which figure from Gentsch et al. (2015) you are referring to.
- Table 2 : Please define O, A, B and C in the table legend.
- Figure 1 : Please precise in the legend what the red dash-dot lines refer to (separation topsoil, subsoil and deep soil ?).
- Table 3 : Please add units.
- Fig.2 (b) : The peak of maximum difference looks cropped.
- l.234 : “In general” should be removed.
- deVrese et al. Paper is now published and is no longer a preprint.

References

Loranty, M. M., Abbott, B. W., Blok, D., Douglas, T. A., Epstein, H. E., Forbes, B. C., Jones, B. M., Kholodov, A. L., Kropp, H., Malhotra, A., Mamet, S. D., Myers-Smith, I. H., Natali, S. M., O'Donnell, J. A., Phoenix, G. K., Rocha, A. V., Sonnentag, O., Tape, K. D., and Walker, D. A.: Reviews and syntheses: Changing ecosystem influences on soil thermal regimes in northern high-latitude permafrost regions, Biogeosciences, 15, 5287–5313, https://doi.org/10.5194/bg-15-5287-2018, 2018.

Wösten, J., Lilly, A., Nemes, A., and Le Bas, C.: Development and use of a database of hydraulic properties of European soils, Geoderma, 90, 169–185, 1999.

Zhu, D., Ciais, P., Krinner, G. *et al.* Controls of soil organic matter on soil thermal dynamics in the northern high latitudes. *Nat Commun* **10**, 3172 (2019). https://doi.org/10.1038/s41467-019-11103-1