We thank the reviewer for her/his comments and suggestions which have contributed to improving the manuscript significantly. In the following, we reply to each point raised, with our answers and changes to the manuscript given in bold.

On behalf of all authors,

Sebastian Westermann

Anonymous Referee #1

This study makes use of the satellite remote sensing and the weather reanalysis data to produce land surface temperatures (LST), which is an important component for modeling the permafrost thermal state, and correspondingly the permafrost extent. This work addresses how LST could be used by the semi-empirical simple permafrost model to calculate present permafrost extent and, more importantly, ground temperatures. I have several suggestions and comments that could be used to improve the current manuscript. I also have several questions. Overall, I enjoyed reading the manuscript and will suggest for the publication after revision.

In the Introduction section, I suggest to mention similar studies by Panda et al, (2014) and Luo et al, (2014). In previous studies, researchers used the GIPL1 model. It will be useful to mention how the GIPL1 is different from GryoGrid1.

We have inserted an additional paragraph in the Introduction, which refers to GIPL1 and the two mentioned studies.

Section 2.4. It is not clear why authors used ERA-Interim and not MERRA-Land or CRUNCEP or any other reanalysis dataset.

Our analysis was performed with ERA-interim, and we have not tested any other reanalysis data set, which we consider beyond the scope of the presented study. We have added an additional sentence on ERA-interim to Sect. 2.3, and paragraphs discussing the uncertainty of ERA-interim to Sect. 4.1.2.

Section 2.5. Authors mentioned the importance of snow and used snow fall (SF) parameter in Table 1 as a turning coefficient for the input parameters. It is not clear, what are the units for SF? Is it normalized or non-dimensional?

It is in m of water equivalent, we have added the information to the caption of the table.

It looks to me that SF is the major coefficient affecting the range in SD shown on Figure 2. If so, then some areas in the mountains could accumulate a lot of snow. Does that mean that we should expect higher uncertainty in certain regions? Adding more background on where we should expect high/low uncertainties in SD will be useful.

In areas with high average snow depth (i.e. high snow fall), the spatial variability of winter ground surface temperatures can indeed be expected to be higher if one assumes

that wind redistribution leads to a pattern of bare-blown spots and snow drifts. At the same time it also plays a role how cold the winters are, i.e. how many FDD are accumulated. We have inserted a new paragraph in Sect. 2.7, which explains the motivation for the different ranges of nf and rk.

Authors mentioned that input parameters in Table 1 are drawn from previous studies. I suggest expanding this by adding more background information on how and why the ranges for nt and rk have been selected. How are the ranges from Table 1 applicable to other geographic regions?

A new paragraphhas been inserted in Sect. 2.7, which explains the motivation for the different ranges of nf and rk. From the study area, there are only few published values for nf and rk, but more background information is provided in Sect. 2.7 of the revised version. It must remain unclear in how far the ranges are applicable to other regions and this issue should be investigated further. While the provided ranges give a satisfactory match with borehole temperatures in the study region, it is not unlikely that other ranges must be assumed e.g. for more continental climate conditions. In Sect. 4.3, we state that "it may become necessary to refine the parameterizations for the ranges of nf and rk, possibly by introducing functional dependences on other environmental variables."

P758. L24. More background is needed on how the cloudy scenes were identified (manually or by using some sort of algorithm)?

Cloudy regions are automatically masked out by the MODIS cloud detection, so that such data are not contained in the employed MODIS LST products. An explanatory sentence has been inserted: "Cloudy regions are automatically detected and removed by the MODIS cloud mask (Frey et al., 2008)".

P 762. L 5. Why is assumed to be equal to 1?

MODIS LST provides the radiative skin temperature, which can be considered equal to the "ground surface temperature" if a dense canopy is lacking. In densely forested areas, one should investigate if there is a "surface offset" between radiative skin and ground surface temperature in long-term FDD and TDD. However, such conditions are not typical for the permafrost domain in the study area. We have inserted: "In case of a dense canopy, where MODIS LST may rather represent top-of-canopy instead of surface temperatures, a different nt may be required, but such conditions rarely occur in permafrost areas in the study region."

Figure 2. Typically TSP temperatures are measured at depth of 20 m. The TTOP calculates temperatures at the top of the permafrost table. It is not clear how did authors compared those two temperatures? If temperatures calculated by GryoGrid1 get extrapolated to the deeper depth then that has to be mentioned in the manuscript.

This is a very valid point, and we are grateful to the reviewer for pointing this out. The borehole temperatures compiled in International Permafrost Association (2010) are

extremely inhomogeneous and provided different depths, as well as for different time periods or points in time. We have used the single one value for the ground temperature provided in International Permafrost Association (2010) for comparison, without any interpolation to a specific depth. However, we emphasize that CryoGrid 1 delivers "Mean Annual Ground Temperatures", which is a hypothetical model temperature that is attributed to some depth at the top of the permafrost (Appendix A).

The mismatch/"error" of the in-situ measurements is estimated as follows: assuming steady-state conditions as in the derivation of the TTOP equation, the vertical temperature gradient $\Delta T/dz$ is equal to $-F/K$, where F is the geothermal heat flux and K the thermal conductivity of the ground. For typical values of $F=-50mW/m^2$ and $K=2W/mK$, the gradient is 0.25K/10m, so a temperature provided from a depth of e.g. 20m may be around 0.5K warmer than at the top of the permafrost at e.g. 0.5m depth. However, ground temperatures in many boreholes have also been warming in the past decade (Romanovsky et al., 2010), and a borehole temperature provided for e.g. 20 m depth may be strongly influenced by climate conditions before 2002, i.e. the time before satellite measurements of LST are available.

With the presented equilibrium approach, this problem cannot be solved. However, we argue that the direction of the error is not necessarily systematic, and generally smaller than the accuracy of 2.5K. We have A) clarified in Sect. 3.1 what was done in the comparison, and B) inserted a new subsection 4.1.1 discussing the above issue. A part of 4.2 has been moved to this new subsection.

In the conclusion section, I suggest to add a paragraph describing how this regional study can be carried over globally? What researchers need to know when use GryoGrid1 in Mongolia or Russia?

An additional statement has been inserted in the Conclusion section: "…, but uncertainties related to satellite-based land cover maps and associated ranges of model parameters must be investigated and resolved prior to global application."

Adding one more figure with calculated uncertainties (SD corresponding to Figure 4) will be extremely useful as well as describing the uncertainty map.

A new Figure 5 visualizing the standard deviation of all model realizations has been inserted, and it is explained in detail in a new paragraph under 3.1.

References

Panda SK, Marchenko SS, Romanovsky VE. 2014. High-resolution permafrost modeling in Denali National Park and Preserve. :1-46.Natural Resource Technical Report NPS/CAKN/NRTR-2014/858. National Park Service, Fort Collins, Colorado. [https://irma.nps.gov/App/Reference/Profile/2208990]

Luo D L, Jin H J, Marchenko S, Romanovsky V. 2014. Distribution and changes of active layer thickness (ALT) and soil temperature (TTOP) in the source area of the Yellow River *using the GIPL model Science China Earth Science, 09/2014; 57(8):1834–1845. DOI: 10.1007/s11430-014-4852-1*

Additional corrections:

There has been an error in the derivation of the TTOP equation in the Appendix. The way it was defined, it is incorrect to set MAGT=T_{TOP}, since the latter refers to the winter temperature at the bottom of the active layer. MAGT is the weighted average of winter and summer temperatures at the bottom of the active layer. This has been corrected by modifying Eq. A1, and adding Eq. A2.