Comments to the paper:
Sensitivities and uncertainties of modeled ground temperatures in mountain environments

By S. Gubler, S. Endrizzi, S. Gruber and R.S. Purves

General comments:

This paper contains a sound treatment of sensitivities and uncertainties evaluated with the model GEOtop representing a well-known and sophisticated model, which can be applied well to several questions related to alpine permafrost. The paper presents an interesting contribution to an important research topic within the geoscientific research community. Appropriate and process-based estimates of sensitivities, errors and uncertainties of heat and mass flux modeling results are more and more important in cryospheric research.
Although, I have some rooted objections against some selected statements in the paper, it represents an excellent and careful high quality work, which will be probably have a major impact and will be some sort of a state-of-the-art (role model) for other model evaluations in alpine permafrost research. Especially, the model evaluation procedure containing recommendations about the discretization parameters in the model are very important and done in a sound manner. Therefore, I would recommend this paper for publication with some changes considering some of the points below:

- Page 802: The authors choose several ground types for their modeling exercise including clay, sand, silt, peat, gravel and rock writing that liquid water influences the thermal conductivity. I fully agree with this statement. However, dealing with alpine permafrost equally important is the air filled part in the ground materials with high porosity such as coarse blocks, which are covering large parts of alpine permafrost areas.
- Page 811/812: I strongly disagree with some statements in chapter 5.1. The authors view is, in my opinion, strongly biased by a selective modelers perspective. If the authors are tuning their model to the measurements, the output variability will be reduced automatically and implicitly they assume that the model represent the real physical processes, which is in their case definitively not true. The authors state that 'These findings can in turn inform future measurement campaigns by quantifying the benefit of an individual measurement'. This statement is of major impact and I find it very dangerous because based on this statement the authors claim that they were able to model all important processes. This is certainly not the case here as the authors probably know by themselves neglecting e.g. important processes in ground materials with high porosity containing large amounts of air with several different processes such as air circulation, long-wave radiative heat transfer between blocks etc. not included in their model. As long as a model study is not able to take all processes into account, I find it somehow foolhardy to make such a statement in a paper. I recommend that the authors reformulate this part of the paper taking into account the suggestion above.

- Since English is not my first language, neither grammar nor spelling of the manuscript was reviewed in detail. However, my impression is that the paper is very well written and no changes have to be made.

Specific comments:

1. Page 792, line 9: Why only hydraulic properties? In coarse debris, it is a remarkable percentage of air within the pore spaces. It has to be taken into account!
2. Page 792, line 13-15: These results are already, since decades known, so this is not a really exciting result of your study. The result of your study is the careful and really well done evaluation of errors and uncertainties evaluated by your model. Therefore, you should concentrate in the abstract on these results!
3. Page 793, line 21: Before the authors start writing about model evaluation, it has to be considered that a model is a) a strong abstraction of the real system, b) it is always a simplification and c) it is influenced strongly by the modeler's perception of the system, which is often strongly biased. These effects have to be taken into account when writing about model evaluation, because uncertainties are often influenced by the biased perception of the modeler.
4. Page 800, line 3: Please give some references for your applied setting of the thresholds for rain and snow.
5. Page 801, line 18: What is the basis of this precipitation correction factor? Please describe this more detailed.
6. Page 801, line 19 to 24: Probably the determination of the fetch distance would be necessary. However, as the mountain topography as well as local surface condition is influencing the local turbulent heat fluxes considerable also the heights of 0.5 to 16 m seems to me somewhat arbitrary chosen. The authors can also select a range between 0.01 and 800 m, which would probably better reflect the range of the real system based on knowledge from balloon soundings in alpine valleys.
7. Page 803, line 3: When the authors define ground as a volume below earth surface, then they have to include the part, which is filled with air too. Otherwise their simulations can not really be compared with alpine permafrost conditions (see general comment above)
8. Page 803, line 17: What do the authors mean with estimated?
9. Page 804, line 8: What the authors mean with *plausible* parameter values. Please be more specific.
10. Page 804, line 17: Again, what is a *plausible* range? Please be more specific.
11. Page 804, line 18: Here one of the fundamental problems in the study is directly introduced -> parameters are determined based on literature review and/or expert opinion -> this approach will automatically lead to a strongly biased model setting. How can the authors prevent this? I do not have a solution but I would expect that the authors discuss this aspect to show that they are aware of this serious problem.
12. Page 804, line 19: What is a *local* sensitivity or better what is the contrary of a local sensitivity -> a *general* sensitivity? Please define this better.
13. Page 805, line 3: The authors write that 'all parameters are assumed independent'. I do not really understand this assumption, because the authors are of course aware that this is certainly never true! May they can support this assumption by a more fundamental explanation, because writing that the study setting is 'synthetic, spatial autocorrelation of the parameters are therefore not taken into account' is

not enough to justify this approach. Especially not if the authors at the end of the paper want to justify that they model study can be used for an improved measurement concept in the field.

14. Page 811, line 24: I strongly disagree with this statement (see general comment above)
15. Page 813, line 19: I strongly agree with the statement that it is important to evaluate individual processes separately if used in impact models. I would like to remember that the authors would have had the opportunity to do such independent evaluations, because at their research site 16 years of energy balance measurements exist to perform such an independent evaluation! Such evaluations were already performed in the past by Stocker-Mittaz et al. (2002), which is maybe worth to mention here.
16. Page 814, line 10 to 18: Yes, again I strongly agree but see comments above.
17. Page 814, line 23&24: The authors state ‘ Parametric uncertainty of MAGT at different depth is almost constant’. Is this result not simply caused by the assumptions made by the authors for the model definition?
18. Page 815, line 9 to 11: In general, I agree with this statement. However, if you miss some processes in the model, the model is producing nonsense although with the help of your measurements and the corresponding tuning of the model produces good results!
19. Page 815, line 25: see comment 15.
20. Page 816, line 10: I would add to this sentence the following ‘after an in-depth evaluation of all processes in the field’.

Technical comments:

1. Page 792, line 7: uncertainties instead of uncertainty
2. Page 796, line 22: in general the use of expression is global radiation or total shortwave radiation instead of global shortwave radiation.
3. Page 833, Fig. 5, Page 835, Fig. 7, Page 836, Fig. 8: The labeling of the axis in these Figures are too small.