Comment on esurf-2022-39
Anonymous Referee #1

Referee comment on "Simulating the effect of subsurface drainage on the thermal regime and ground ice in blocky terrain, Norway" by Cas Renette et al., Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2022-39-RC1, 2022

The manuscript presents a modeling study of permafrost temperatures and ground ice with and without (lateral) drainage of water from the model domain. The hypothesis is that well-drained blocky terrain may result in colder ground conditions and the subsurface drainage of water should therefore not be ignored in permafrost models. The model experiment is nicely designed, and the results clearly show the impact of drainage in different ground substrates. The results should be of interest for anyone interested in permafrost processes, especially in mountainous and blocky terrain. While I find the study interesting and the modeling results valuable, I have some concerns about the validation procedures and the general structure of the paper and to some extent the setup of the simulations.

Major comments:

- The idea of the model experiment – testing the impact of drainage on the ground thermal regime – is very interesting and the approach seems sound and straightforward. However, the model validation and the links to the two sites in Norway are weak. For one of the sites there are observed temperatures available to compare with the model results but these data are not used to validate the model results in a robust way. Instead, validation is only carried out visually and no objective statistics are used to evaluate the model results. For the other site, there are no data on ground temperatures but only observed ground surface temperatures. The validation and comparison to observations from the Norwegian sites thereby adds next to nothing to the modeling study. My suggestion would be to either remove these sites from the manuscript, to add some statistical analysis based on the available ground temperature data, and/or to find a rock glacier site with available ground temperature data which could be used for a real validation of model results.
- The methods section would benefit from restructuring and extensive editing, because it is rather confusing as currently written.
Model description. It would be helpful to start the model description with a presentation of the mesh used: is this a vertical column (1D) mesh? What is the thickness and cell size of the domain? Some of this information is available in other parts of the methods section, but starting off by clarifying such basic facts would help the reader to envision the model setup.

Snow model. A list of processes included in the model is provided, but there is no explanation of how these are represented in the model. If these are explained in previous publications, please clearly refer to those for the specific processes listed. For example, “the physical effect of wind drift on the snowpack” is included, according to the text, but later it is stated that the redistribution of snow by wind is phenomenologically represented using the snowfall factor. So, what effects of wind on the snowpack are actually included? Does the snowfall factor allow for variable distribution of snow on the domain, or is there also a redistribution of snow over time?

The description of the different model runs (validation, equilibrium, and transient) is rather confusing and split up in different sub-sections of the methods description. I would suggest to clarify the purpose of the model runs and explain them in a separate sub-section. The validation runs are runs validated based on visual fit, which is not a very robust metric for model validation. It seems like the purpose of validation runs was to determine (calibrate?) the proper sediment stratigraphy for each site. But how were other parameter values determined? Table 1 presents some physical properties, but not e.g. Kh. Snowfall factor is for some reason tested in the equilibrium runs (why?). In general, the purpose of testing the snowfall factor sensitivity is rather unclear to me. The outcome of this test is not mentioned in the abstract, but the current text does not suggest that this is pure model calibration. L230 suggests that snowfall factors were also tested in validation runs, but this is not mentioned previously.

The forcing data used for spinup and simulation runs are also not presented in a clear manner. The spinup and initialization procedure and data vary between the sites and the runs and the motivation for the choice of spinup periods is not always clear. Perhaps a table presenting the basic details of each run would be more helpful. Also not clear in the current text: Is forcing data downloaded and downscaled for the two sites (e.g., two separate series of data)? What is the time resolution (daily? monthly?) of the data used for simulations and for validation of simulation results?

Minor comments

L155-164: Seepage face is at atmospheric pressure, but does that mean that the seepage face is located at Zwt? Could that sentence just be removed and eq. 1 explains all? Where are Kh values from? Eq. 1 mirrors Darcy’s law but values seem made up, and dz is based on the grid z. Explain the rationale behind this model and why it was chosen. (alternatively, if Kh and dlat are unknown, these two parameters could have been replaced by one single parameter.)

L282: But table 1 does not include any tested case with a porosity of 0.2!! Either the methods section does not completely describe what was done in the validation runs or there is a typo here?
L296-312: The model results suggest that the differences in model setups have very little impact on MAGST and that overall there seems to be a slight cold bias in the model. What then does this validation add to our understanding, if the purpose of the study is to investigate the impact of drainage on ground temperatures and ground ice?

Section 4.1: This section would greatly benefit from some objective measure of model fit instead of just some conclusions from visual inspection of plotted temperature curves.

Section 4.2: It is not surprising that snow insulates the ground from wither cooling, generally leading to higher ground temperatures. More relevant than a comparison of snowfall factors, would be to check if observed snow depths and ground temperatures are recreated with the forcing data used here. When reading this section, I get curious about how much snow is generally observed at these two sites and does the forcing data capture the range of observed snow, or is a snowfall factor scaling needed to "correct" the forcing data.