# Interactive comment on "Toward improved parameterization of a macro-scale hydrologic model in a discontinuous permafrost boreal forest ecosystem" by Abraham Endalamaw et al.

**Anonymous Referee #1**

The paper introduces a fine-scale parameterization scheme based on a landscape model using 30-m DEM data. The paper futher assesses the hydrological impacts of the fine-scale parameterization scheme compared to a coarse one based on outdated datasets (e.g., FAO soil data and SNAP vegetation cover) in a perfmafrost-dominated 100-km2 catchment. The results show that the streamflow estimates from the small-scale parameterization match the observations slightly better (but still rather poorly), which is kind of obvious. Of course using 9-km FAO data to parameterize a 100-km2 catchment is not going to work and I don't see why anyone would think otherwise. It is also a shame that the approach is tested in only one very small catchment, how much better does the approach work in considerably larger catchments?

In the abstract, it is stated that the two parameterizations "capture most the peak and low flows with similar accuracy in both sub-basins" and then after it states that "on average, the small-scale parameterization improves the total runoff simulation approximately by up to 50% in the LowP sub-basin and 10% in the HighP sub-basin". Which one is true?

On page 13 it is stated that there is a "lack of high spatial resolution soil data in the region". The SoilGrids250m dataset might perhaps be useful (http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0169748).

Page 13: "All soil parameters are regridded to 1/64th degree". How? Nearest neighbour? Bilinear?

So the catchments cover like a single grid cell of the 9-km resolution FAO map. Then how come there is finer-scale spatial variability visible in Figure 4a?

Page 19: Why is only the large-scale parameterization used for the calibration? Wouldn't it be more fair to re-calibrate for both the small- and large-scale parameterizations?

Might be worth mentioning that lumped catchment values are derived using the calibration.

"Values between 1.0 and 0.0 are widely considered to be acceptable levels of model performance". This is not true. Although it depends on the situation, I suppose values >0.5 can generally be considered acceptable.

Figure 6a: The observed streamflow time series look kind of strange, in 2006 in particular the streamflow looks truncated. Could this be due to ice blockage or?