Interactive comment on "Snowfall distribution and its response to the Arctic Oscillation: An evaluation of HighResMIP models in the Arctic using CPR/CloudSat observations" *by* Manu Anna Thomas et al.

Anonymous Referee #1

Received and published: 15 April 2019

This article presents a comparison of simulated snow accumulation from high resolution model simulations contributing to HighResMIP, and satellite estimates of snowfall based on retrievals from CloudSat-CPR. The topic is interesting, and the comparison itself is novel, with respect to the evaluation of snow simulation in this new set of high-resolution modelling experiments. However, I have some concerns about the description of the methods, and some of the interpretation, which are described below.

1. Inadequate methodological description. I found that the methods used were not adequately described. For instance, how was the regridding of CloudSat measurements

to 1 x 3 degrees performed? Were all instantaneous profiles within a grid box simply averaged together, or something else? On P3 L25 it states that model data were regridded to 1x1 degree, so how were comparisons made between model and CloudSat if the grids are different? Further, P4 L18 states that 1x3-degrees provides "sufficient" sample size, but how was "sufficient" determined, and what constitutes "robust"? Was the comparison of mm snow water equivalent (SWE), or snow depth? If the former, then the conversion method and treatment of snow density should be included.

Another concern is the treatment of missing data: on P12 L5 the 2011 battery failure is discussed; however, we are aware of two other battery failures [September 2009 - December 2009] and [January 2011], but these periods are seemingly unaffected in Fig.6. Was any gap-filling/interpolation required during these times?

2. Set of models used for comparison. When the authors mentioned an evaluation of low, versus high, resolution models, I expected to see at least a couple of "typical" CMIP5-class GCMs/ESMs included, to provide a reference for how/if these high resolution simulations represent an improvement for the simulation of snow. The low/high res sample evaluated here is somewhat artificial, since even the "low res" models are among the highest resolution simulations one would find in CMIP5. I would strongly recommend that the authors include one, or more, CMIP5-era simulations, to provide some more context for the HighResMIP results.

3. Figure quality. I found the maps very difficult to read, because of the small size and choice of colour scale. In general, it is very challenging to ask readers to evaluate by eye how well a set of model simulations compares to a reference map. I strongly recommend condensing these maps into standard model v obs diagnostics, for instance using Taylor Diagrams. This would provide a much more rigorous evaluation of each model than can be provided by eye, and would make it easier for the authors to add more models to the comparison (see point 2 above). In addition, the CloudSat map panels have a data-wrapping issue at the date line.

4. Uncertainties. The devil is in the details with this type of comparison, and I would have appreciated a much more thorough discussion of the various sources of uncertainty that are present in these results. A much-needed addition to Fig.6 could be a credible interval for each estimate, considering sampling, instrumental and model uncertainty. Due to sampling, I would expect to see much larger uncertainties on the CloudSat snow estimates coming from more southerly latitudes (e.g. Reg#3).