Norrköping, 2019-06-12

Response to Reviewer #2

We thank the reviewer for providing constructive remarks. We have tried to incorporate your suggestions in the revised manuscript. Please find below a point by point response to them.

Major comments:

1. The observational uncertainty is very poorly quantified. CloudSat suffers from main issues with undersampling light precipitation events (as the authors indicate), but no uncertainty is assigned to that. I wonder why the authors did not consider working with other CloudSat data (including internal quality flags of the CloudSat algorithm, and/or radar-derived reflectivity).

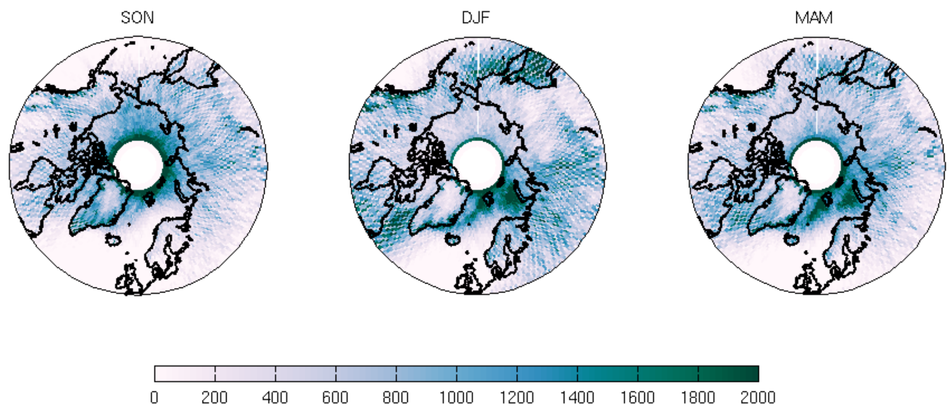
We thank the reviewer for this comment, which is also raised by the Reviewer #1. Instead of repeating the same response here, we kindly refer the Reviewer #2 to our response given to the Reviewer #1 on the same topic.

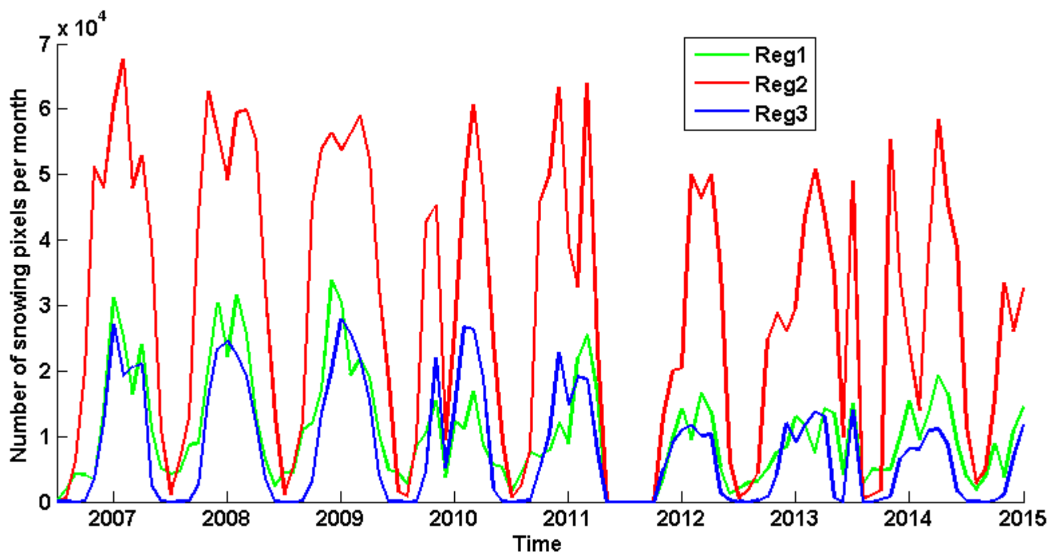
2. Statistical robustness and process analysis. If working with percentiles, while only considering <10 years of data (so only a sample of <30 months for each season), I wonder how significant these percentiles are. In addition, the authors should be extremely careful in translating a low observed monthly snowfall rate to an inadequate treatment of light snowfall by the models; it might very well be that this monthly snowfall rate comes from just one large snowfall event! Although I realize that CloudSat frequency is only monthly for this gridded product, this connection from observation to model physical processes should be much more expanded upon; for example, the authors can look into the distribution of snowfall in the models in those low-snowfall months, to confirm whether or not these are associated with just low-snowfall events or not. Without such a detailed analysis, the reader remains 'in the dark' about the exact defiency in the models. Note that recent literature suggests that models produce 'too little, too frequent' snowfall (e.g. https://journals.ametsoc.org/doi/pdf/10.1175/JCLI-D-16-0666.1); does that apply do these models as well?

We do actually share the reviewers concern. However, please keep in mind that CloudSat remains the only source of snowfall estimates covering nearly the entire Arctic. Although we do not have multidecadal data, we believe about 10 years of data are sufficient enough to capture the first order features, seasonality and spatial variability in snowfall over the Arctic. This has therefore been the focus of our evaluations. The climate models in question are being used for the next IPCC assessments and it therefore becomes necessary to work with the data we currently have to understand the model performance in simulating snowfall.

With regard to the concern if the individual snowfall events can impact monthly snowfall estimates, we show below 1) the spatial distribution of the total number of snowfall pixels available at the original 1x1 deg grid, accumulated for each season studied here from 2006 to 2014, and 2) monthly time series of the number of snowing pixels accumulated over the three

selected regions shown in Fig. 5.
It can be safely concluded that the monthly averages during the SON, DJF och MAM months are not represented by just a few strong events.





Two two figures are now added as Supplementary information in the revised version of the manuscript.

In Figure 7 of the manuscript, we have already investigated the distribution of snowfall in models versus CloudSat. We have expanded the discussion in the revised version, especially in conjunction with the recent study that the reviewer mentioned.

3. Choice of models: why were only these models chosen? Why only ocean-forced models and not fully-coupled models? Why is opted to not use a COSP-style radar simulator in the models? How does that impact the results? More details on the PRIMAVERA project are necessary. In any case, if only AMIP-style models are chosen, the authors should use the overlapping period between observations and models only.

This study was carried out as part of the PRIMAVERA project and the main aim was to evaluate those GCMs that would participate in the next IPCC assessments. The following line about the

PRIMAVERA project is added in the revised manuscript. “Hence, the main aim of this study is to evaluate the HighResMIP (High Resolution Model Intercomparison Project) simulations for CMIP6 (Haarsma et al., 2016) under the PRIMAVERA (PRocess-based climate sIMulation: AdVances in high resolution modelling and European climate Risk Assessment) project. This project is a European Union H2020 project wherein a total of 7 state of the art models are run at varying resolutions to understand the impact of resolution on different global climate processes.”.

As far as we know, unfortunately there isn’t any COSP-like simulator currently available that can be applied for comparing snowfall.

4. Methods, results, and discussion are mixed throughout the paper, which is somewhat confusing for the reader. A conclusion section with a long list reads like notes from a presentation, and should be converted a flowing text instead. More discussion should be added; why are these results important? How do these compare to other literature? What can we learn from it? What processes are lacking and/or inadequate in models?

The conclusion section is revised to provide more clarity on these issues.

Minor comments (P=page, L=line)

P1, L2: surface radiation budget

This is edited in the manuscript.

P1, L20: How would incoming radiation cool the surface?

‘... thereby cooling the surface’ has been removed.

P2, L34: irrelevant for this paper; Alpine snowfall is much more related to mid-latitude atmospheric dynamics, as well as orographic snowfall.

Indeed this study looks into the Alpine snowfall. It is however included as it carried out a model intercomparison between snowfall observations and a wide scale of models.

P3, L10: What is this project and what is its aim, and how does this paper fit in that?

The following sentences are added to the manuscript: Hence, the main aim of this study is to evaluate the HighResMIP CMIP6 simulations under the PRIMAVERA (PRocess-based climate sIMulation: AdVances in high resolution modelling and European climate Risk Assessment) project. This project is a European Union H2020 project wherein a total of 7 state of the art models are run at different resolutions to understand the impact of resolution on different global climate processes.

P3, L15: what is a snowfall flux? Mass or energy?

The snowfall flux is in kg/m2/s. The models output the snowfall estimates as snowfall flux, whereas, the CloudSat retrievals output snowfall accumulation. Hence, the model output are converted to snowfall accumulation for fair comparison. This is updated in the manuscript.

P3, L21: later, you discuss results from 2006-2015. Try to be clear
This is clearly addressed in the text. Though these simulations run from 1980-2015, for the computation of percentiles, a ten year period overlapping the observational period is considered. However, for the AO variability calculations, the whole period is considered. This is now clarified in the text.

P3, L27: possibly?
The word ‘possibly’ refers to the fact that some of the HighResMIP models used in this study consider graupel.

P4, L19: how do you end up with ten seasons? Clarify.
Sorry, that was a typo. I meant ‘the’ instead of ‘ten’.

P4, L20 to P5, L8: these are methods, not results. Consider changing your section titles to improve structure (Section 2: Data and Methods; Section 3: Results; Section 4: Discussion and Conclusions)
The revised manuscript is organized accordingly.

P5, L16: larger difference
P5, L29: extent
The above changes are done in the revised manuscript.

P5, L31: what does ‘negatively skewed’ mean?
When the models tend to overestimate the observed response, so, in the Gaussian curve, the model response would be towards the right of the observed line (as can be seen in Fig.7 for p10; the black line that corresponds to the CloudSat data lies on the x-axis), then the distribution is said to be negatively skewed.

P6, L9 and beyond: This belongs to the discussion section.
This is moved to the discussion section.

Figures: too small; add units and indicate seasons in Figure itself. Add sublabels (e.g. a, b, etc) to refer to in text.
The units are now given in the figure itself. The seasons are already mentioned in the figure. The sublabels, (a)-(d) are given for the models in the revised manuscript.

P10, L32: this is the main issue of this paper; it should be added to the discussion section and reflected/expanded upon – see Main Issue 2
The discussion section is revised in the manuscript.

P12, L5: this is part of Methods, and does not fit here
This is moved to the ‘Methodology’ section in the revised manuscript.

P15, L1-13: Methods.

Section 6: This has not been clearly introduced in the intro and feels obsolete. Why would this be relevant to this paper? Why not also look at impact of NAO, Arctic sea ice extent, GBI, and other indices?
The importance of AO and its relevance is mentioned in the Introduction section in the revised version, instead of mentioning it directly in Section 6. As was mentioned in Section 6, AO is the dominant mode of natural variability in the Arctic and has large impact on precipitation variability (mainly in the form of snow during the polar winters). Therefore, it is important that the models capture this response of snowfall to AO, at least to a first degree, to be able to reasonably represent Arctic climate variability.

We preferred to investigate AO over NAO, mainly because while NAO is regional (mainly affecting the Atlantic sector), the AO is considered to have an Arctic wide impact.

P17, L4: I would argue that this is definitely not ‘a wide range of models’
We look into 4 different models run at two resolutions each. Hence, the term ‘wide range’ of models.