Response to anonymous Referee #2

Thank you very much for your remarks. We have tried to incorporate all your suggestions. Please find below a point by point response to your comments.

Overview:

The goal of this model evaluation study is to compare the cloud radiative effects between standard-resolution and high-resolution climate models. The authors find that cloud radiative effects differ strongly between models, but not nearly as much between runs with different resolutions of the same model. The authors conclude that the apparent insensitivity to increased atmospheric horizontal resolution indicates that physical parameterizations play a dominant role in determining the behavior of cloud-radiation feedbacks. The manuscript touches an important topic and is worthy of publication. I did not find any major flaws that would require substantive revisions. The article could benefit from a final discussion on what the findings mean for model development (e.g., Should we focus on better parameterizations? Is it useful to run high-resolution climate models if biases aren't really improved compared to lower resolution models?)

A following paragraph is added to the end of the 'Conclusions' section:
"From this study it is clear that the well-known issue of the large biases in SW CREs over the polar regions during the melt season does not improve by increasing the resolution of the models chosen here. This would require improvements not only in the parameterization schemes involving the microphysical properties of clouds, but also in the surface description. Analysis of the spatial pattern of the TOA SW CREs during winter reveal that different cloud regimes are affected drastically with a change in resolution in MPI-ESM and HadGEM3 models. For example, the Hi-res HadGEM3 model show an overestimation over the convective ITCZ regions compared to its Std-res counterpart and this may have an impact on SST resampling and thus convection. On the other hand, the Hi-res MPI-ESM overestimates the CREs over the Southern Oceanic stratocumulus region and this may have an impact on the cloud fraction. The observed differences can be attributed to the lack of tuning in higher resolution versions. Though the models tend to simulate the spatial variability in cloud radiative response to ENSO and NAO variability, they vary widely in the magnitude of the response. The CRE biases associated with the NAO phase are smaller compared to those with the ENSO phase. Although some improvements can be seen regionally, it is difficult to identify patters that hold across all models. Hence, it can be concluded that improving the physical parameterization schemes rather than increasing the resolution is perhaps important in better simulating the CREs. However, it has to be noted that these are atmospheric only simulations and the impact may be different in the presence of coupled climate models."

General Comments:

1. Especially the abstract and the beginning of the manuscript are well written. However, the manuscript becomes tedious to work through after about page 6. I think this may be just a reflection of the topic, since model validation studies tend to be tedious.

Indeed, the nature of the manuscript, being an evaluation study, makes it a bit tedious to read at a first glance. Please note that we have chosen to discuss the results both at the TOA and the surface, while discussing the individual SW and LW components of CREs as well. Furthermore, we investigate both ENSO and NAO variabilities during their two phases. While we believe that laying out such a detailed analysis might help interested readers to make further evaluations/comparisons (targeting individual piece of information), it has unavoidably

resulted into the manuscript being a bit tedious. We have tried to balance the simplicity and information content as smoothly as we possibly could.

2. One of the points that makes the article tedious to read is that the authors switch between “low resolution” and “std-res” models. It would be easier if just one qualifier is chosen, in this case probably std-res (standard resolution).

This has been corrected in the revised manuscript.

3. The quality of the figures could be better. For example, the multi-panel map plots have lots of white space between the two columns, and lots of space is occupied by large colorbars. The individual panels could be enlarged at the expense of the wasted space.

Following the reviewer suggestion, the figures are re-plotted to be clearer.

Specific comments:
1. page 1, line 10: no comma after whereas
2. page 1, line 18: EU-funded
3. page 4, line 23: add “the” before Niño3.4 index
4. page 26, lines 11-12: add “the” before coastal Antarctic and Arctic

All the above suggestions are incorporated in the manuscript.