Response to Referee #2

This manuscript uses satellite observations from CALIPSO to evaluate Arctic cloud cover in ECHAM6. The authors found that low liquid cloud cover in the Arctic is biased high over surfaces covered by snow and ice in the default version of the model. They investigate two potential reasons for the high bias the strength of surface heat fluxes and the impact of the Wegener-Bergeron-Findeisen (WBF) process. The authors conclude that surface heat fluxes are too strong in the default version of the model and that they can instead decrease their high bias in Arctic low liquid cloud cover by allowing for slight supersaturation with respect to ice in their cloud cover scheme, which in turn impacts the WBF process in ECHAM6. I have numerous concerns about the manuscript that are primarily related to the methodology and conclusions drawn by the authors. My comments are below.

We thank the reviewer for the constructive comments.

Major comments

The description of the observational dataset does not contain a discussion of observational uncertainties associated with CALIPSO/GOCCP. Namely, lidar beam attenuation is particularly problematic in the Arctic, where many clouds are optically thick, liquid, low-lying and precipitate snow. When compared to ground-based observations in the Arctic, CALIOP cannot see clouds in the lowest few kilometers (see e.g. Liu et al. (2017)) and the difference with GOCCP can be quite substantial especially over the Greenland ice sheet (Lacour et al. (2017)). This was also noted to be problematic in Cesana et al. (2012), and mostly affects precipitating ice underneath optically thick liquid clouds. I worry that the authors claim of a high bias in low, liquid clouds in the Arctic and their comparison for ice clouds may be inaccurate for the aforementioned reasons. The disadvantage of ground-based remote sensing observations, of course, is their lack of spatial coverage. I would still, however, recommend that the authors incorporate Arctic ground-based remote sensing observations from a few sites collocated with GOCCP to get an idea of potential biases that might impact their conclusion.

In the revised version of the manuscript, a more detailed review of the uncertainties related to the GOCCP dataset is included (i.e. lidar attenuation by liquid clouds, cloud detection thresholds that might not be representative for Arctic region and also possible affects of spatio-temporal sampling of satellite data).
Nevertheless, we think that our conclusion of an overestimated low-level cloud fraction in ECHAM6 is still valid. The GOCCP dataset is based on satellite retrievals and is not directly comparable to ground observations or to model output. In order to make our model results comparable to the GOCCP dataset we use the COSP satellite simulator. In the revised manuscript, we compare modeled (ECHAM6+COSP minus ECHAM6) to observed (GOCCP minus ground based observations) cloud cover profile differences and see a similar underestimation for modeled clouds when using a satellite simulator (ECHAM6+COSP) compared to the cloud fraction form ECHAM6's cloud cover scheme. While comparing modeled and observed differences in cloud cover profiles is not an "apples-to-apples" comparison (because of different definitions of what is a cloud), this demonstrates that COSP derived cloud properties can mimic real world issues of the spaceborne lidar. Therefore, the reported overestimation of low-level clouds in the model is a "real" signal and not just due the observational issues in the GOCCP dataset.

Furthermore, the description of the observational dataset also does not mention the vertical resolution and criteria used for phase discrimination in the GOCCP product. Were daytime and nighttime data used? What timeframe was used? Were data before prior to the change in nadir-viewing angle used? How were oriented crystals handled?

In light of this remark by the reviewer, we revised the description of the CALIPSO-COCCP dataset. Section 2 now contains a more detailed description of the observational dataset (i.e. cloud detection thresholds, information on vertical resolution, phase discrimination). In Section 3, we now also state that we use monthly averaged data for the same timeframe as the model simulations using both, day- and nighttime overpasses. Concerning the change of the nadir pointing angle at the end of 2007, the period we used for evaluation of ECHAM6 (2007-2010) could be affected by that. This would mainly affect the retrieval of the cloud phase due to an effect on the depolarization ratios by horizontally oriented crystals. As COSP does not use any information on the shape of ice crystals from the model (as most models do not have information on the shape of the ice crystals), the effect of horizontally oriented crystals can be ignored at least from the model side.

The authors note that ECHAM6 mixes too strongly in the Arctic and instead decide to turn to the models parameterization of the WBF process instead to attempt to remedy the bias in Arctic cloud cover. To this end, the authors increased the efficiency of the WBF process by decreasing the threshold of in-cloud ice water mixing ratio required to activate the depositional growth of ice. However, it appears that the authors are unaware that ECHAM6 (Lohmann and Neubauer(2018)), like many other climate models (Komurcu et al. (2014), Cesana et al.(2015), McCoy et al. (2016)), underestimates the proportion of liquid to ice in mixed-phase clouds. Decreasing the efficiency of the WBF process would only exacerbate this underestimate (Tan and Storelvmo (2016), Lohmann and Neubauer (2018)), which could also affect the climate sensitivity of the model (Tan et al. (2016), Lohmann and Neubauer (2018)).

Citing Lohmann and Neubauer (2018), the reviewer states that ECHAM6, like many other climate models, underestimates the proportion of liquid to ice in mixed-phase clouds. We would like to point out that Lohmann and Neubauer (2018) did not use the ECHAM6 Stevens et al. (2013), but used ECHAM6-HAM2 Zhang et al. (2012). Even though both models share a lot of their physical parameterizations, they significantly differ in the microphysical parametrizations. While ECHAM6 employs a single-moment scheme, ECHAM6-HAM2 uses a more sophisticated double-moment scheme. Even though both microphysical schemes stem from a common predecessor, they considerably vary in a lot of microphysical processes. One has therefore be careful when comparing ECHAM6-HAM2 to ECHAM6. Figure 3 in Lohmann and Neubauer (2018) shows the fraction of supercooled liquid clouds for ECHAM6-HAM2 (as well as for a number of sensitvity studies) on a global average. While on global average ECHAM6-HAM2 might underestimate this fraction, this figure does not show the fraction of supercooled liquid clouds in the Arctic. Komurcu et al. (2014) provides zonal-mean averages of supercooled liquid cloud fraction for different cloud top temperatures for ECHAM6-HAM2 (see their Figure 4) and for temperatures at or below -30° C, ECHAM6-HAM2 overestimates the amount of supercooled liquid clouds for high latitudes, even though by not much.
Figure 5 in Cesana et al. (2015) provides a similar zonal-mean, temperature binned supercooled liquid cloud fraction for MPI-ESM Giorgetta et al. (2013), which is the coupled version of ECHAM6, and a similar overestimation of supercooled liquid shows for MPI-ESM in the Arctic (compared to GOCCP at temperatures below -30° C). This overestimation of liquid cloud fraction in the lower part of the mixed-phase temperature regime is consistent with the fact that the overestimation of liquid cloud is only simulated in winter (DJF) and spring (MAM) where such cold temperatures can occur in high latitudes. Additionally, while being positively biased in high latitudes, MPI-ESM slightly underestimates the amount of supercooled liquid in the clouds in the mid-latitudes and in the tropics (see their Figure 6) even though not by much.

Thus, although the bias in cloud cover might be remedied, the partitioning of cloud phase would be further exacerbated. I would recommend the authors to look into how cloud thermodynamic phase is affected in the model before retuning the WBF process, which previous studies have already shown to be too efficient in climate models, including ECHAM6.

The reviewer is correct that even though the bias in liquid cloud fraction might be remedied by a stronger WBF processes, the effects of this measure on the actual (mass) phase partitioning (IWC/(LWC+IWC)) might be different. To this end, we follow the reviewer's advice and look into how cloud thermodynamical phase is affected before retuning the model. There is no observational product that can provide both, liquid and ice water content, on a large enough scale to compare it to a GCM. This is also the reason why all the studies cited by the reviewer are trying to mimic frequency ratio fraction of the cloud phase that can be provided by CALIOP. A possible approach to evaluate cloud phase would be to look at liquid/ice water path which can be derived from MODIS. As stated in the introduction, using passive spaceborne sensors might be problematic due to the environmental conditions and also due to fact the Arctic clouds are often mixed-phase clouds, which further complicates the retrieval of cloud microphysical properties (Khanal and Wang, 2018). To obtain at least a rough estimate of how the ice (mass) fraction is affected by a stronger by a stronger WBF process in ECHAM6, we added a plot of temperature-binned average ice fraction over the North Atlantic and over Siberia (Figure 6 in the revised manuscript). For the ice fraction in Siberia, we find quite low ice fraction ($\sim 70\%$) in the temperature range between -25° C and -10° C. Comparing this to in-situ observation of ice fraction as provided by Korolev et al. (2017) such a "plateau" is not visible. Figure 5-14 in Korolev et al. (2017) shows a more gradual increase in ice fraction (decrease in liquid fraction) with decreasing temperature (which can be seen in the bins for high/low ice fraction) and we think that the more or less constant ice fraction in the model over Siberia is another indication of an overestimated amount of liquid clouds over snow/ice covered surface as has been stated in the manuscript. As the ice fractions from in-situ observations and the ice fractions from the model are on a completely different spatial scale, one nevertheless has to be careful when doing such a comparison. As we have shown in our conclusion, the TOA shortwave CRE seems to be biased low in MPI-ESM which might be another hint that there is more liquid water in the clouds, which would make them less reflective, so we think that a slightly stronger efficiency of the WBF and therefore an higher ice (mass) fraction can be justified.

Why do the authors choose to focus on the WBF process? Why not ice nucleation for example, which also plays an important role in Arctic radiation (Prenni et al. (2007), Xie et al. (2013))?

The reason why we focused on the WBF is twofold. Firstly, it has to be a process that is able to efficiently reduce the amount of cloud liquid water. We conducted a number of sensitivity studies and modified the strength of all processes that can affect the liquid water content and we found the WBF to be by far the most efficient one. It also can be seen from table 4 and 5 in Klaus et al. 2012 that only the WBF process (γ_{thr}) and the collection of cloud droplets by snow (γ_4) are able to do so. Not included in this table is heterogeneous freezing of cloud droplets, but we found that increasing its efficiency did not lead to strong enough reduction in liquid cloud cover over snow and ice covered surfaces. Secondly, what makes it appealing to tune this process is the fact that it is strongly simplified in ECHAM6. Due to efficiency in tuning the amout of ice in clouds, modifying the strength of this process is also often used to tune the model to bring it into radiative balance. This can be seen from the fact that this parameter can vary up to an order of magnitude for different horizontal resolutions in ECHAM6. These considerations are now explained in more detail in the revised manuscript.

The authors note that although there were improvements to Arctic low liquid cloud cover by increasing the efficiency of the WBF process, total cloud fraction remained overestimated. To this end, the authors then modified the cloud cover scheme to allow for slight supersaturation with respect to ice in the model (their NEW experiments). The authors seem to point out in the main text that cloud although some of the high bias in low-cloud fraction is reduced in their NEW simulations, new low-biases in low-cloud cover are introduced. Although improvements to the high bias in low-cloud fraction were highlighted in the abstract and conclusions, they authors fail to mention that there appears to be a simultaneous introduction of a new low bias in low-cloud cover. In fact, this low bias in Arctic low-cloud fraction was already shown for the CAM5 model (Kay et al. (2016)), which allows for supersaturation with respect to ice (Gettelman et al. (2010)). Therefore, the author's parameterization does not seem to entirely solve the problem of the high bias in low-clouds in the Arctic, and the problem now reduces to an issue known to already exist in another model..

In the revised version of the manuscript, we try to more clearly point out why a temperature-weighted scheme for saturation vapor pressure in combination with an increased efficiency of the WBF process introduces an negative bias in low clouds. As the amount of low-level ice clouds remains more or less constant for different values of γ_{thr}, the amount of liquid clouds strongly decreases and therefore also the amount of clouds in general. The decrease in liquid clouds is mainly caused by the more efficient WBF processes which more efficiently turns liquid into ice clouds over continents compared to oceanic regions, it also affects clouds there. In the standard setup of ECHAM, liquid clouds are already biased low in those regions which is even further enhanced by a more effective WBF process. As liquid clouds seem to react rather sensitively to a more effective WBF process, only minor changes of γ_{thr} can have strong effects on the amount of liquid clouds and we think that setting γ_{thr} to $2.5 \cdot 10^{-6}$ $\mathrm{kg\,m^{-3}}$ is already the best choice to improve WBF process. This value is the best compromise between improving cloud cover over snow and ice covered surfaces by simultaneously not further worsen clouds in other regions.

Also, although their temperature-weighted scheme for saturation vapor pressure may be new to the ECHAM6 model, it is not a new concept to climate models. Please cite previous work that have used similar weighting schemes in the calculation of saturation vapor pressure.

In the revised version of the manuscript, we now cite previous work that have used similar weighting schemes in the calculation of saturation vapour pressure.

Section 3: It seems to me that there is a chicken and egg game when using observations of the vertical profiles of temperature and humidity to establish a cause for high bias in low liquid clouds in the model. Low-clouds can in turn affect temperature and relative humidity, so how can one establish the cause for the low-cloud bias?

The reviewer is correct that no causal relationship can be established between a positive bias in low-level temperature and humidity and a positive cloud cover bias. Nevertheless, we believe that such biases in temperature and humidity can be an indicator of an overestimated cloud cover due to this two-way relationship that has been stated by the reviewer. We mainly used this comparison of vertical profiles to show that the reported cloud cover bias is not just due to possible uncertainties in GOCCP but is a real model problem. On request by the other reviewer, we additionally show data from ERA-Interim to also have information on temperature and humidity profiles on a wider spatial scale to show that there is a difference between snow/ice covered regions and water/open land. Looking at relative humidity, ECHAM6 seems to generally overestimate it over the continents, but this overestimation is most strongly pronounced in those regions we observed the strongest positive biases in low-level clouds, which make us confident that this overestimation actually exists.

Minor comments

Abstract, line 9: Phase partitioning" typically refers to mass ratio or frequency ratio defined as liquid/(liquid + ice) in mixed-phase clouds within a grid cell or specified domain. Here, the authors refer to the ratio of total low liquid cloud cover to total cloud cover. I recommend changing the terminology to avoid confusion.

We replaced "Improvements in the phase partitioning of Arctic low-level clouds" with "Improvements on the overestimated Arctic low-level liquid cloud cover"

I suggest changing the title of Section 2.1 to GOCCP" to reflect the fact that this CALIPSO-derived product was used in the analysis.

In the revised manuscript, we replaced all instances of CALIPSO with GOCCP and completely revised section describing GOCCP.

Page 2, lines 20-23: I would also mention the advantage that active satellites are also able to provide vertical profiles of clouds.

We mentioned that actives satellites can provide vertical profiles of clouds which cannot be provided by passive satellites.

Page 5, lines 10-13: If the mid-level cloud bias is similar to the low-cloud bias because of how low- and mid-level clouds are defined, then shouldnt that mean that the bias in mid-level clouds for JJA should resemble the bias for high clouds? It does not appear to.

We misinterpreted the similarity of the mid-level cloud bias to the low-cloud bias and our explanation does not hold. We therefore looked into the vertical profile of clouds and at the altitude of the threshold for low-, mid- and high-clouds (see attached figure). The thresholds themselves vary only a little between summer and winter. The actual cause for the seasonal variation of the mid-cloud bias can be attributed to the vertical position of the generally overestimated high-clouds in ECHAM6. The vertical extent of the troposphere is influenced by the atmospheric temperature which cause the cirrus clouds to be present at lower altitudes in winter. The similarity to low-cloud stems form the fact the temperatures are colder over snow and ice covered surfaces, which cause the cirrus clouds to be simulated at even lower altitudes and therefore contributed more the mid-level clouds compared to oceanic regions. We correct our false claim in the revised manuscript.

Page 5, line 20: This is an overstatement without formal proof. I would suggest replace is with appears to be.

Done.

Page 6, lines 19-22: This is an interesting hypothesis that may or may not be true. I would be more careful in emphasizing that the statement is speculative.

We try to more clearly formulate that this statement is speculative in the revised manuscript.

Page 8, line 13: Please add a reference for the WBF process and note the ways in which models simplify it (e.g. lack of dependence of vertical velocity). Please see Korolev (2007).

We added a reference for the WBF process at its first mentioning at the end of section 3. We also stated how ECHAM6 simplifies the WBF process due to its lack of dependence of vertical velocity.

Page 8, line 21: will" should go in front of depositional".

Done.

Page 10, Lines 11-12: Please specify that this the overestimate is with respect to GOCCP.

We now specify that the overestimation is with respect to GOCCP.

Page 11, lines 15-17: I disagree with this statement. The Karcher and Lohmann paper refers to cirrus clouds. In mixed-phase clouds, where liquid and ice clouds coexist and the WBF process occurs, the cloud may not necessarily glaciate immediately and will instead depend on how the liquid and ice are spatially distributed within the cloud (Tan and Storelvmo (2016)).

We removed the reference to the Karcher and Lohmann form our manuscript as it indeed refers more to cirrus clouds. Nevertheless, the way that mixed-phase clouds are parameterized in ECHAM6 will eventually cause any liquid water to be depleted quite quickly, as the condensation is the only process the can produce water in the mixed-phase temperature regime. As soon as there is enough cloud ice present and it exceeds γ_{thr}, condensation does not take place any more and any liquid water will quite quickly either freeze or evaporated. This can indeed be considered not physical as the presently used implementation of condensation/deposition does not allow for simultaneous growth of liquid and ice within a cloud. ECHAM6 also has no information on the subgrid distribution of liquid and ice within a cloud which might prevent this rather rapid depletion of liquid water.

Page 12, Line 17: reduce to" "reduce the"

Done.

Page 12, line 18: Please specify that supersaturation is with respect to ice.

We now specify that supersaturation is with respect to ice.

Figure 4: strength to strength

Done.

Figure 5: Please consider labelling the first value as the default value of the model in the legend of this figure for easy reference

Done.

Please remove all instances of the" in front of Arctic amplification".

Done.

References

Cesana, G., Waliser, D. E., Jiang, X., and Li, J.-L. F. (2015). Multimodel evaluation of cloud phase transition using satellite and reanalysis data. *Journal of Geophysical Research: Atmospheres*, 120(15):7871–7892, doi:10.1002/2014JD022932.

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Bttinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W., Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K. D., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K.-H., Claussen, M., Marotzke, J., and Stevens, B. (2013). Climate and carbon cycle changes from 1850 to 2100 in mpi-esm simulations for the coupled model intercomparison project phase 5. *Journal of Advances in Modeling Earth Systems*, 5(3):572–597, doi:10.1002/jame.20038.

Khanal, S. and Wang, Z. (2018). Uncertainties in modis-based cloud liquid water path retrievals at high latitudes due to mixed-phase clouds and cloud top height inhomogeneity. *Journal of Geophysical Research: Atmospheres*, 123(19):11,154–11,172, doi:10.1029/2018JD028558.

Komurcu, M., Storelvmo, T., Tan, I., Lohmann, U., Yun, Y., Penner, J. E., Wang, Y., Liu, X., and Takemura, T. (2014). Intercomparison of the cloud water phase among global climate models. *Journal of Geophysical Research: Atmospheres*, 119(6):3372–3400, doi:10.1002/2013JD021119.

Korolev, A., Fugal, J., Krämer, M., McFarquhar, G., Lawson, P., Wendisch, M., Borrmann, S., Abel, S. J., Schnaiter, M., Franklin, C., Crosier, J., Williams, E., Wang, Z., Axisa, D., Lohmann, U., Field, P. R., and Schlenczek, O. (2017). Mixed-Phase Clouds: Progress and Challenges. *Meteorological Monographs*, 58(Fahrenheit 1724):5.1–5.50, doi:10.1175/amsmonographs-d-17-0001.1.

Lohmann, U. and Neubauer, D. (2018). The importance of mixed-phase and ice clouds for climate sensitivity in the global aerosol–climate model echam6-ham2. *Atmospheric Chemistry and Physics*, 18(12):8807–8828, doi:10.5194/acp-18-8807-2018.

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E. (2013). Atmospheric component of the MPI-M earth system model: ECHAM6. *Journal of Advances in Modeling Earth Systems*, 5(2):146–172, doi:10.1002/jame.20015.

Zhang, K., O'Donnell, D., Kazil, J., Stier, P., Kinne, S., Lohmann, U., Ferrachat, S., Croft, B., Quaas, J., Wan, H., Rast, S., and Feichter, J. (2012). The global aerosol-climate model echam-ham, version 2: sensitivity to improvements in process representations. *Atmospheric Chemistry and Physics*, 12(19):8911–8949, doi:10.5194/acp-12-8911-2012.

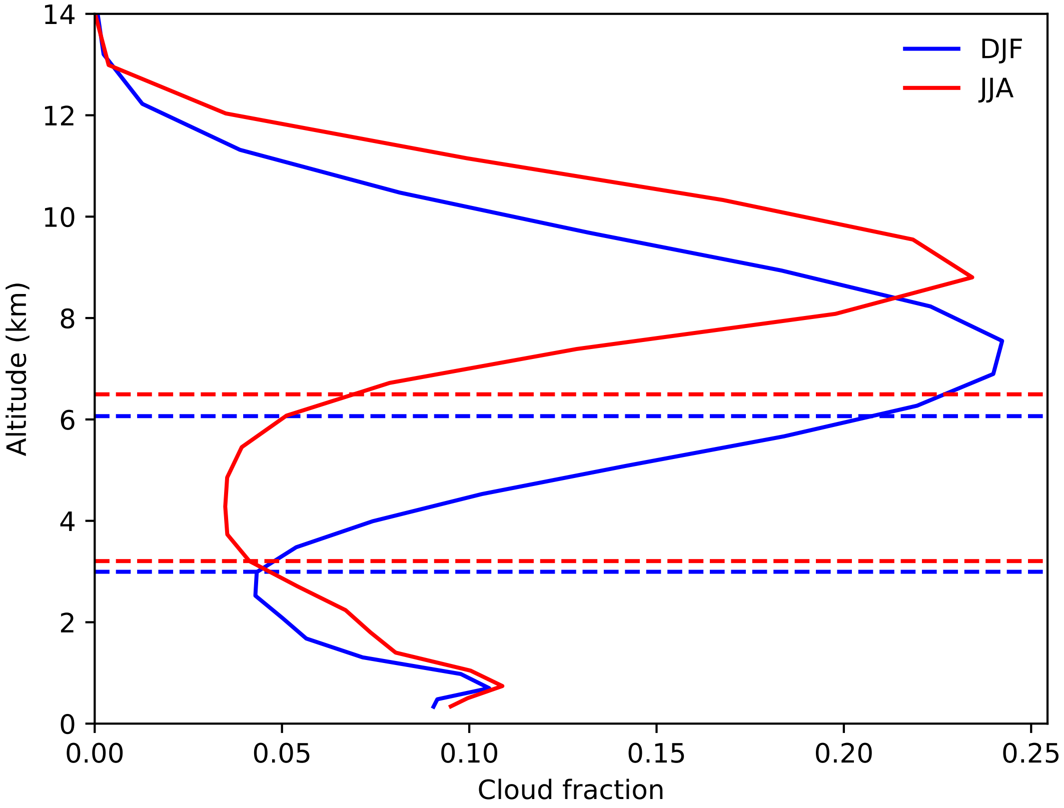

Figure 1: Vertical profiles of cloud cover for winter and summer in the Arctic as well as the thresholds for the low/mid/high classification.