1. Author's responses to Reviewer comments

We gratefully thank the three Reviewers for their suggestions and comments. Below we present our response to each Reviewer. General comments are considered first, which are followed by point-to-point answers to specific comments provided by each of the Reviewers. (The Reviewer comment is shown in *italics* highlighted with yellow, which is followed by our response).

In the process of revising our analysis, a small bug was found from the code for autoconversion which on rare occasions caused the CNDC to be increased. Luckily, it had an almost negligible impact on the results (e.g. impacts on global-mean radiation fluxes were on the order of 0.01 W m^{-2}). The bug has been corrected for all the results presented in the manuscript, and this results in some very small differences in some of the reported numerical values as compared to the initial submission.

As a general note about the revised manuscript, much of Section 4 has been revised, with corresponding changes in the Conclusions. Section 3 also includes some more specific corrections, which are outlined in our replies below.

1.1 Reviewer #1

General comments

1.
The manuscript needs to better illustrate the differences between simulations in a quantitative way (probably with zonal mean figures) as noted below.

2.
In addition, given the prominence of autoconversion as a process for the results, it is probably necessary to show some process rates for autoconversion.

3.
Vertical velocity distributions and activation rates would also be useful to explain some of the more confusing aspects of the relationship between vertical velocity variance and cloud drop number noted in the text.

The Reviewer highlighted the need for more quantitative illustration of the results: our results are now presented in zonal means. The process rates for autoconversion and cloud activation are now also shown (Figures 2 and 3 in the revised manuscript), and used to provide more detailed understanding of several processes leading to our results. The revised manuscript also includes more discussion about the influence of the subgrid vertical velocity distribution

Specific comments

1.
P15524,L6: should the ACI be negative?

Yes, this has been corrected.

2.
P15228,l25: what type of cloud is the high cloud cover due to? You are using the ISCCP simulator, so you can discriminate thin and thick clouds, and high and low clouds? Also in Tontilla et al 2013? Or is that a feature of the nudging?

The high total cloud cover is not a feature of nudging; similar features were present in Tonttila et al. 2013 as well, where nudging was not used. In general, the cloud-top pressure is too low in our model runs as compared to ISCCP-data, suggesting an overestimation in high-level cloud cover. However, low level cloud cover at mid-latitudes is also larger than in ECHAM5 simulations without the HAM2 aerosol module. Moreover, comparison of the global mean cloud radiative effects with CERES-EBAF data (presented in Table 4 in the revised manuscript) shows too strong cloud radiative effect for both SW and LW radiation in our model, which would be consistent with an overestimation of both high- and low-level clouds.

3.
P15529,L29: note that reducing autoconversion might also impact accretion. What is the relative importance of these processes in ECHAM-HAM?

Autoconversion plays the key role for the differences between the model versions considered in the manuscript, yet it is agreed that autoconversion and accretion rates do interact. Their relative importance is also strongly controlled by model tuning which, however, has not been considered for accretion rate in this manuscript.

4.
P15530,L6: but above you said sensitivity was low Inge S.H. and higher in the NH. Please clarify.

The following explanation will be added to the revised manuscript:

"In the southern hemisphere the autoconversion rate is sensitive to changes in CDNC due to the generally low CCN concentration over the oceans. Thus, the slightly increased CDNC shown by ACT is accompanied by increased LWP as compared to REF, since reduced droplet size reduces the amount of water that is converted to drizzle and rain. In comparison, in the northern hemisphere subtropics and mid-latitudes, CDNC is lower in ACT than in REF, especially over land, but LWP is similar to REF. This likely relates to the low sensitivity of autoconversion to small changes in CDNC in regions with high CCN concentration"

5.
P15330,l25: again, do the earlier simulations show this with ECHAM ham (high cloud amounts and high tau). Or is this a feature of nudging? From Tontilla et al 2013 it looks like cloud amounts stay high.

Also: perhaps you should show microphysical process rates for autoconversion here.

As already mentioned in the text, the high cloud cover is present in all our runs using the McICA and the stochastic cloud generator in conjunction with ECHAM5-HAM2, also in Tonttila et al. 2013. Earlier papers using the McICA and stochastic cloud generator in standard ECHAM5 (without HAM) did not show this feature. It is not caused by nudging. This will be elaborated in the revised manuscript. The process rates for autoconversion and cloud activation are given in Figures 2 and 3, respectively, in the revised manuscript.

6.
P15531, L3: these need to be more quantitative. Perhaps zonal means would be better?

7.
P15531,L11: the difference is not easily visible. See comment above.

8.
P15332,L1: can you demonstrate with zonal mean difference plots? I also think an analysis of the microphysical process rates would be wise.

9.
P15532,L5: why not show the microphysical process rates for autoconversion and other processes?

11.
P15541: figure 2 and 3 would be better as zonal means. In figure 2 I cannot see any differences. These could be made more quantitative for figure 3 with zonal means. Could maybe separate land and ocean as well.

Since the same general issues are repeated in comments 6, 7, 8, 9 and 11 here, we provide a collective response to them. All the figures are now given in zonal means and new figures 2 and 3 have been included for autoconversion and nucleation rates.

10.
P15533,L5: useful to state that ACI seem to follow deltaLWP, not deltaCDNC.

12.
P15543: figure 4. Again, this is hard to see any quantitative differences. Perhaps showing a zonal mean on the same plot would be better. Where are differences? Assume this is mentioned in the text.

The plot has been converted to zonal means. Clearly, the strongest differences in the aerosol indirect effect are seen for the northern mid-latitudes, collocated with the significant differences in the anthropogenic perturbation of cloud properties.

The following discussion will be added in the revised manuscript:
"The zonal mean net AIE is shown in Figure 6. The bulk of the difference between REF and ACACT occurs in the midlatitudes of the northern hemisphere. It is also noted that the global distribution of AIE follows rather tightly the anthropogenic perturbation in LWP, which along with the difference in the global mean AIE highlights the importance of how autoconversion is calculated in the model."

1.2 Reviewer #2

1.
The authors should clarify how the model configurations REF, ACT, ACACT differ from the configurations REF, SUBW, SUBWRT, W_ADJ1, W_ADJ2 in T2013.

REF and ACACT are similar to REF and SUBW respectively, except that the former ones are now run for both pre-industrial and present-day conditions, and they are performed as nudged runs, unlike in Tonttila et al 2013. This will be described more clearly in the manuscript

2.
There is no mention of retuning for radiation balance. If REF is in radiation balance, then ACACT must not be. I would suggest to add a retuned version of ACACT to the comparison. In T2013, the retuning involved adjusting the autoconversion scaling factor. There is ample evidence in the literature that altering autoconversion can have a large impact the magnitude of the indirect effect, so this should be discussed and investigated.

A retuned model configuration ACACTRT has been added in the revised manuscript. The autoconversion rate scaling factor was reduced from 3.0 to 1.5. This results in quite similar radiation budget in REF and ACACTRT in pre-industrial conditions (net radiation balance within 0.3 W m^{-2}, and also similar LWP). The net aerosol indirect effect (AIE) in ACACTRT is -1.37 W m^{-2}, which is approximately 0.09 W m^{-2} stronger than in the un-tuned ACACT configuration. Thus, the difference in AIE between ACACTRT and REF is -14 %, which is smaller than that between ACACT and REF (-19% after the bug correction mentioned above), but it is still a significant effect.

The question of the initial basic state of the different model configurations is an interesting one, yet rather difficult to answer definitively. On one hand, when we are tuning the model closure parameters, we are directly tampering with the model physics, thus adding additional effects on the model results besides the ones we are actually interested in. On the other hand, the argument about differing model basic state is of course valid. Thus, we interpret the result such that the 19 % reduction of AIE in the un-tuned model setup represents the direct impact of the subgrid parameterizations, whereas the 14 % reduction after retuning more closely resembles an operational implementation. Discussion on this will be added to the concluding section (Section 5) of the revised manuscript.

It is also worth noting that to a large extent the change in AIE due to subgrid treatment of microphysics eventually stems from the change in autoconversion rate, and is thus related to precipitation formation and LWP (more in-depth discussion of this will be added to Section 4 of the revised manuscript. Please consult also our responses to comments 8 and 11 of Reviewer #3). This is consistent with earlier studies about the importance of precipitation formation in representing the anthropogenic aerosol effects in large-scale models.

3.
Panels in Figures 2, 3, and 4 are very small and difficult to read. In the difference panels, most regions are probably not statistically significantly different from one configuration to another. Maybe it would be better to plot zonal averages and then highlight which regions of the zonal averages are statistically significant.

These figures are replaced by zonal mean plots. Please refer to our responses to Reviewer #1. The zonal distribution of statistical significance was analysed and is commented in Section 4 of the revised manuscript accordingly.

4.
West et al. (2014, doi:10.5194/acp-14-6369-2014) found a strong relationship between the variance of the subgrid vertical velocity distribution and the magnitude of the indirect effect.

The citation will be included in the revised manuscript.

5.
P15525, lines 1-6: this is an incomplete description of the state-of-the-art. A very large number of climate models do not use a single effective vertical velocity for activation, but rather explicitly integrate over a vertical velocity distribution. This was first proposed in 1997 and has been adopted in many contemporary climate models (see for example dois: 10.1029/97JD00703, 10.1029/96JD03087, 10.1029/2005JD006300, 10.1175/2010JCLI3945.1). ACT follows the same basic idea.

We will account for this comment, and add the suggested references, in the Introduction of the revised manuscript.

6.
P15526, lines 16-17. Even if one were to assume that all the TKE was confined to vertical motions (which is physically impossible), the upper bound on the proportionality coefficient would be 1.41 (sqrt(2)). Is the 1.68 value simply treated as a tuning parameter?

Agreed. The value 1.68 is indeed considered as a type of tuning parameter, as also discussed in Tonttila et al. (2013). For clarity, a short discussion is now included also in the current manuscript. It states that the value 1.68, although unphysical as the Reviewer states, is selected in order to match the magnitude of the GCM grid-cell mean vertical velocity with the effective velocity in REF, in order to isolate the effects of subgrid variability alone. This point will be elaborated on in Section 2 of the revised manuscript.

7.
P15526: choosing sigma to be the same as the single effective velocity in REF almost automatically guarantees that CDNC will be smaller with subgrid variability than without, since the majority of sample points will have velocities smaller than the effective velocity.

Please note that we choose sigma so that the <u>mean vertical velocity over the positive side of the PDF</u> matches the single effective vertical velocity, i.e. the mean magnitude of the subgrid vertical velocity samples used for cloud activation approximately matches the effective vertical velocity. For a Gaussian distribution, with the distribution peak at zero, the mean over the positive side is ~0.79*sigma. Thus sigma needs to be larger than the effective vertical velocity in REF in order to match the average magnitude of the vertical velocity for cloud activation, and to isolate the impact of subgrid variability. As it is suggested in the text, in many cases it still results in smaller CDNC with subgrid variability than without, since small velocities still gain more weight in terms of the mean CDNC, if CCN concentration is high enough so that cloud activation is sensitive to variations in vertical velocity.

8.
Section 2: mention the number of sub-columns and the additional associated cost compared to REF.

We use 50 subcolumns which increases the computational cost by 20-25 % on a Cray XC30 computer. This will be mentioned at the end of section 2 in the revised manuscript.

9.
Table 2: add CERES-EBAF observation for SWCRE and LWCRE. Also add net TOA radiation values.

We have separated the cloud properties to Table 2 and radiation quantities to new Tables 3 and 4 in the revised manuscript. Table 3 shows the radiation quantities for the pre-industrial runs of each model configuration. Table 4 shows present-day radiation quantities, also including the CERES-EBAF observations.

10.
In T2013, Sect 6, there is a brief discussion about an imposed minimum cloud drop number of 40 cm-3 in ECHAM5.5. If this minimum value is still being imposed, it would be relevant to discuss it in the present manuscript.

The issue of minimum CDNC will be noted in Section 5 of the revised manuscript.

1.3 Reviewer #3

Major comments
1.
The differences in the basic states in three configurations. The differences in the basic states of three configurations are large. For example, LWP is 67.4 g/m2 in ACT and reduces to 50.2 g/m2 in ACACT, about 25% decrease. The same large differences are also true for column-integrated droplet number concentrations. Therefore when the authors discuss the differences in aerosol indirect effects, the large differences in the basic states need to be accounted for. The relative difference therefore may be more meaningful. For example, although increase in in-cloud CDNC at the 890 hPa is the largest in REF (36.42/cm3), with smaller increase in ACT (31.83 /cm3) and ACACT (30.7/cm3), the relative increase (compared to the PI CDNC) is the largest in ACACT (44.5%), with smaller relative increase in REF (41.7%), and Act (40.5%). The same can be applied to LWP and aerosol indirect forcing as well. When the relative differences are used, the picture can be quite different. Accordingly, many discussions in Section 4 and 5 will need to be revised.

2.
The physical mechanisms behind simulated changes in anthropogenic aerosol effects on LWP needs to be better understood, especially between ACACT and ACT. The large difference in LWP change from PI to PD between ACACT and ACT is probably one of the most important results of this manuscript, but the reason behind this is not clear at all. On the other hand, if the relative change is used, the difference is much more moderate, and not sure whether the difference is still statistically significant.

The differences in the initial model states have now been accounted for by 1) analysing both the absolute and relative changes between PI and PD conditions and 2) by including a new retuned model version, where the radiation balance and e.g. LWP are quite close to those in REF in PI conditions. The new results are considered in Section 4 and the Conclusions. We have also added a more detailed explanation about the underlying reasons for especially the differences in the LWP change from PI to PD. A more detailed list of changes can be found in our responses to the Specific comments.

Specific comments

1.
line 17, page 15524: Results from climate model simulations that account for cloud-scale motion (Wang et al., 2011, doi:10.5194/acp-11-5431-2011; Wang et al., 2012, doi: 10.1029/2012GL052204) also contribute to the weaker aerosol indirect forcing estimate, as discussed in IPCC AR5. Wang et al. (2012) is particularly relevant to this paper, as that paper is also about understanding aerosol indirect effect differences in different models, and how differences in cloud microphysics might help to explain model differences.

The references will be included in the revised manuscript.

2.
Lines 22-24, page 15525: Not sure I would agree with this statement ("a significant part of model-based overestimation of aerosol indirect effect can be explained by omitting subgrid variability in cloud microphysical processes") and a similar statement in the abstract. I think the paper needs to provide more evidence to support this statement. One thing is that tthe differences in the basic states are needed to be accounted for. Another thing is that the physical mechanisms that lead to the large reduction of aerosol indirect forcing from ACT to ACACT needs to be further explored.

A new model run with retuned autoconversion scaling parameter has been added to the manuscript. It shows that even after retuning the net indirect radiative effect due to subgrid variability is approximately 14 % weaker than in REF. We interpret this so that for an operational setup the impact of subgrid parameterizations is most likely at least 14 %. Please refer also to our answer to comment no. 2 of Reviewer #2.

The difference in the indirect radiative forcing between ACT and ACACT follows quite closely the difference in LWP between those model configurations, which in turn is seen to be tightly coupled with the treatment of autoconversion, and for which we can provide an explanation. Regarding this, please refer to our answer to comments 8 and 11 below.

Furthermore, we have moderated the wording of the above-mentioned sentence to reflect the fact that retuning slightly reduces the difference in indirect effects between ACT and REF. It now reads: "These simulations demonstrate directly that omitting subgrid variability in cloud microphysics contributes to the overestimation of model-based aerosol indirect effect". A corresponding change will be included in the abstract.

3.
Lines 21-23, page 15526: It is not clear to me how the subgrid distribution of CDNC is purely determined from the subgrid distribution of vertical velocity. To my knowledge, grid-mean CDNC is a prognostic variable in ECHAM5-HAM2, which accounts for both source and sink terms such as droplet activation, advection and precipitation. So how can the PDF of CDNC is directly determined by the PDF of sugbrid vertical velocity?

Indeed, the grid-scale CDNC is the result of source and sink terms, but the subgrid distribution of CDNC operates in the subcolumn space, which is stochastic. Thus, the subgrid variability is a diagnostic property generated for each timestep. Of course, during that timestep the distribution is also affected by the subgrid autoconversion in the case of ACACT and ACACTRT. This is done also in part for computational reasons, since having each subcolumn prognostic would increase the number of tracers to an unpractical level.

The following explanation will be added to Section 2 of the revised manuscript:
"The subgrid vertical velocity samples from the PDF are used to calculate cloud droplet activation, which yields the distribution of CDNC in the stochastic subcolumn space. Note that the subcolumn CDNC distribution is treated as a a diagnostic property, while a prognostic formulation (Lohmann et al., 1999) is retained for the grid-scale mean CDNC."

4.
Lines 25-26, page 1526: Please also elaborate how the model account for the correlations between LWC and CDNC in their subcolumn generator (or, more precisely, the correlation between LWC and the subgrid vertical velocity). This can be elaborated either here or in the description of the case of ACACT on page 15527.

Since these parameterizations operate on turbulent stratiform clouds, we do not assume any correlation between the cloud properties and vertical velocity, following the discussion in Tonttila et al. 2013 and e.g. Morales and Nenes 2010. For cumulus clouds the situation would be different since in that case the in-cloud thermodynamics are important in driving the updrafts.

We will add the following sentence in Section 2:
"Since our focus is on stratiform clouds, the vertical motions to be parameterized are highly turbulent and thus presumably weakly correlated with the thermodynamical properties of the cloud (in contrast to convective cumulus clouds), as also noted in e.g. Morales and Nenes (2010). Therefore, we do not assume any correlation between vertical velocity (and thus CDNC) and LWC."

5.
Page 15527, the case of ACT: Please clarify whether the subgrid CDNC in ACT is used in radiation calculation.

Yes, it is. This is stated explicitly in the revised manuscript.

6.
Figure 1c and 1d, page 15529: I understand the decrease in CDNC from REF to ACT, as activated droplet number concentration increases non-linearly with increasing vertical velocity and this non-linearity is mainly caused by the competition of water vapor from more activated droplets. However, there is also a significant decrease in CDNC around 60S from REF to ACT. This is not clear to me.

7.
Line 20, page 15529: "even slightly increased CDNC". The increase at around 60S is quite significant. What causes this increase?

The most likely explanation is that as shown by Figure 3 in the revised manuscript, the nucleation rate is slightly increased in ACT and ACACT over the southern hemisphere oceans, where CCN concentration is low, as compared to REF. This is in contrast to regions with higher CCN concentration, where subgrid treatment decreases the cloud droplet nucleation rate. While virtually all suitable accumulation and coarse mode particles are activated already at very weak updrafts when the aerosol concentration is low, ACT and ACACT showed an additional boost in the number of activated aerosols in small Aitken mode particles due to some subcolumns having a very strong vertical velocity compared to the effective velocity in REF.

The slight increase in CDNC in ACT then stands out, because the autoconversion rate is very similar between ACT and REF, while ACACT displays stronger autoconversion due to the subgrid calculations. The stronger autoconversion compensates for the effects of cloud activation in ACACT, resulting in lower CDNC at 60 S, than what is seen for ACT. This point will be discussed in Section 3 of the revised manuscript.

9.
Lines 10-17, page 15531: see the major comment #1. As the basic states are quite different in three configurations, the absolute difference in CDNC can be misleading sometimes, and a relative difference can be more meaningful. See the approach used in Wang et al. (2012). If the relative difference is used here, the picture can be quite different (see the major comment #1). Accordingly, many statements and discussions in this section will need to be revised. For example, I do not think you can attribute 80% of this difference to subgrid cloud droplet activation alone. I also do not think you can conclude that "the type of autoconverion is not important for the anthropogenic perturbation in CDNC in our model".

10.
Lines 5-15, page 15532: see my last comment and the major comment #1. The same argument can be applied to LWP as well. The relative change in LWP is 11.0% in ACTACT, and 12.8% in ACT. The difference between ACACT and ACT is therefor much moderate than 35% cited in the paper based on the absolute change.

The relative differences between PI and PD for CDNC and LWP are now shown in figures 4 and 5, respectively and considered in the text. Even though the subgrid parameterizations do clearly reduce the CDNC for both PI and PD conditions even after retuning, it is true that the relative change between PI and PD is not very significant. This is discussed in Section 4 of the revised manuscript.

For LWP, the relative PD-PI changes are more considerable and statistically significant, although indeed not as strong as the 35% difference obtained using the absolute change. In the retuned model (ACACTRT) the absolute LWP change is about 19% lower than in REF, with very similar initial LWP between the two in PI conditions. Taking the difference in the relative LWP change between ACACT and REF gives 18 %. These results will be discussed in Section 4 of the revised manuscript.

8.
Lines 1-7, page 15530: I am a little bit surprised by the difference in LWP between REF and ACT. The almost identical LWP in the NH is particularly puzzling. Does this mean that LWP in the NH is dominated by those over oceans in your model? Is this result consistent with Tonttila et al. (2013)?

11.
Lines 5-15, page 15532: the differences in LWP between ACTAC and ACT clearly needs more explanation. It is not immediately clear to me why accounting for the subgrid variation of LWC and CDNC leads to smaller LWP change in ACACT. Given that the relative difference is now quite moderate, I am not sure whether the relative difference is still statistically significant.

First, regarding comment 8, over the midlatitudes of both hemispheres the LWP is much higher over oceans than over continents (revised Figure 1), which is expected. However the difference between ACT and REF is quite small in both. One explanation for the NH, that is most heavily influenced by anthropogenic aerosol in present-day conditions, is that CDNC is generally high, so that the difference in CDNC between ACT and REF would yield rather small differences for the cloud water removed by autoconversion (both use grid-mean values of CDNC and cloud water to calculate the autoconversion rate!). This is backed up by Figure 2 in the revised manuscript, which shows the autoconversion rates.

Second, regarding the PI-PD change in LWP in ACACT and ACT (comment 11), the smaller change in LWP due to the subgrid treatment of autoconversion is also visible in the relative change. The smaller change in ACACT is likely connected to the spread of the subgrid values of CDNC, which can easily be larger than the difference in the grid-mean CDNC between ACACT and ACT. The mechanism can be explained as follows: First, in PD conditions, cloud activation is limited by the available CCN less frequently than in PI conditions. Therefore, the subgrid variability of vertical velocity plays a larger role in PD conditions, which results in a larger spread of the subgrid CDNC in PD conditions both for ACT and ACACT. Second, in ACACT the autoconversion rate is calculated using subgrid values of CDNC (and LWC), while ACT uses grid-mean values. It is expected that, due the non-linear dependence of autoconversion on CDNC, the consideration of subgrid variations in CDNC acts to increase the grid-mean autoconversion rate, and does so more effectively in PD conditions where the spread of CDNC is larger. This compensates for part of the decrease in autoconversion that is associated with the PD-PI change in the grid-mean CDNC. Consequently, the reduction in the autoconversion rate from PI to PD conditions is smaller for ACACT than for ACT. It is shown in Figure 2 of the revised manuscript that not only is the autoconversion rate consistently stronger (more negative) in ACACT than in ACT, but also the relative change between PI and PD is slightly weaker for ACACT, i.e. the decrease in autoconversion rate from PI to PD is indeed smaller for ACACT than for ACT in the northern hemisphere. This behaviour is seen throughout the lower troposphere and can thus explain the reduced anthropogenic change of LWP in ACACT. The difference in the PI-PD LWP change between ACT and ACACT is significant also in terms of the relative change at the 99 % confidence level.

The corresponding discussion will be included in Section 4 of the revised manuscript.

12.
Lines 16-23, page 15532: the smaller aerosol indirect effects in ACTAC can be partly explained by the smaller SWCRE in this case (-52.75 W m-2) than in REF (-55.92 W m-2).

The new retuned model configuration ACACTRT is used to investigate this issue as the SWCRE is quite similar to REF after retuning. Even for ACACTRT the global mean indirect effect is 14% smaller than in REF. This supports our basic conclusion: consideration of subgrid effects yields a relatively strong reduction in the model representation of the aerosol indirect effects. Please refer also to our response to comment no. 2 by Reviewer #2.