Responses to the interactive comments on "Detection and attribution of aerosol-cloud interactions in large-domain large-eddy simulations with ICON" by Montserrat Costa-Surós et al. to

Anonymous Referee #2

Review of the manuscript numbered ACP-2019-850 Title: "Detection and attribution of aerosol-cloud interactions in large-domain large eddy simulations with ICON" written by Montserrat Costa-Surós et al. Manuscript number: "acp-2019-850". Decision: "Major revision"

In this study, the authors conducted numerical simulations using large-eddy simulation mode of ICON (ICON-LEM) covering a large calculation domain (whole area of Germany) with fine grid resolution (156 m). They evaluated the ICON-LEM through the comparison between the results of satellite and ground-based observations and those of ICON-LEM. They also tried to detect and attribute the signal of the aerosol effects on the cloud properties through the sensitivity experiment with changing aerosol. From their analyses, the authors indicated that the signal of the cloud aerosol interaction is only seen in the cloud number concentration and liquid water path larger than 200 g m-2. I think that the nesting simulation using the LES model covering such large domain has never conducted, and this is one of the unique points of this study. This study can be a basis of the numerical weather prediction with such fine grid resolution, and a basis of the numerical studies targeting on aerosol-cloud interaction by "real-case (nesting) simulation" with fine grid resolution. So, I evaluate the authors' efforts to conduct this study. However, most of the analyses conducted in this study can be done by the results of the simulation with "coarse" grid resolution. So, the manuscript has room to be modified as described below. Based on the descriptions shown above, my decision is "major revision", and I encourage the authors to modify the manuscript.
We thank the reviewer for his/her suggestions. However, we consider that the study could not have been done with the coarser resolution since according to Stevens et al. (2020) there is a clear benefit of using high-resolution simulations (horizontal resolution of 156 m) in comparison to coarser ones (315 m and 625 m) for cloud-related studies.

Stevens et al., 2020. The Added Value of Large-eddy and Storm-resolving Models for Simulating Clouds and Precipitation, J. Meteor. Soc. Japan., in press (doi: 10.2151/jmsj. 2020-021).

General Comment:

1: As I mentioned before, I evaluate the author's efforts to conduct simulation with fine grid resolution covering such large calculation domain. However, most of the analyses conducted by this study can be done by results of the simulation with coarse grid resolution. The analyses, which can only be done by the results with fine grid resolution, are required. Such analyses extend the value of this study. Entrainment around cloud edge, supersaturation and therefore CCN around the cloud base, and turbulence structure are examples of such analyses (Please do not misunderstand, entrainment, supersaturation, and turbulence are examples).
The reviewer of course is right that similar studies can be performed with coarse-resolution models. Most aerosol-cloud interaction studies so far, in fact, use general circulation models at 1 million times (150 x 150 km² rather than 150 x 150 m²) coarser horizontal resolution. The point here is that the LEM is much better at resolving the relevant cloud processes (cg. Stevens et al., 2020). The high resolution here was also needed for a unique detection-attribution assessment, as it went down nearly to the instrumental scale.

2: The author concluded that the signal of the aerosol-cloud interaction is difficult to be detected in terms of the cloud cover, cloud top height, cloud bottom height, liquid water path smaller than 200 g m-2. However, is this conclusion applicable for other cases? Based on the previous numerical simulation like Khain et al. (2008), the impacts of the aerosol perturbation on the clouds is

dependent upon the meteorological field. I understand that the simulations for other cases using ICON-LEM require huge amount of computational resources, and it is not necessary to conduct the simulations. However, the author should add comments about whether the conclusion of this study is applicable for other cases or not with referring previous studies.
The reviewer is right that it is a clear limitation of our study that only one day over one domain was simulated. However, as stated in Section 2.1 (page 4) the selected date (2 May 2013) covered a wide range of cloud- and precipitation regimes (see Fig. 1, which illustrates the cloud conditions, based on satellite data). The conditions of that day allowed us to study at the same time low, mid, high, and convective clouds, as well as different types of precipitation (see section 3.5). A statement has been added at the end of the Conclusions section that further studies are needed for longer periods and other regions to corroborate, falsify or extend the conclusions.

3: The description about how to couple the aerosol and clouds in the ICON-LEM is not enough. The coupling of the aerosol and cloud is sensitive to the aerosol cloud interaction simulated by the model. In my understanding based on the manuscript, the number concentration of CCN calculated through the results of the COSMO-MUSCAT and the parameterization of Abdul-Razzak and Ghan (2000: AD2000) was given to the microphysical model of Seifeld and Beheng (2006: SB06) in ICON-LEM, and feedback of the cloud to the aerosol field was not calculated like off-line coupling in this study. Is this right? Or is the feedback explicitly calculated? The feedback of the cloud to aerosol (e.g., wet deposition) can reduce the aerosol and CCN number concentration. So, there is a possibility that one of the main conclusions of this study: "signal of the aerosol cloud interaction is limited to the number concentration of clouds (Nd) and LWP larger than 200 g m-2" could be change when the aerosol coupled on-line. Of course, I understand that off-line coupling is good as a first step, but I suggest the authors to add more detailed description about how to couple the aerosol and cloud in ICON-LEM (e.g., how to use CCN number concentration by AD2000 in SB06 with equation).
The reviewer is right in this. We agree that a more detailed description in the text of this work is beneficial for the reader, therefore we added a more extensive statement on this in Section 2.1 to explain that the aerosol is prescribed, but we improved the model by allowing for the CCN sink on activation.
Based on the aerosol species mass modeled by COSMO-MUSCAT, the parameterization described by Abdul-Razzak and Ghan, 2000 is used to calculate time varying 3D fields of the CCN number concentration for a set of updraft velocities. The translation from aerosol mass into aerosol number is done according to Hande et al., 2016, assuming average number size distribution for the different aerosol species. The CCN fields are then used in ICON-LEM.

4: The discussion about the radiative forcing is poor. The authors discussed the radiative forcing for global scale through the scaling of the radiative forcing over the Germany. However, this discussion is unreasonable for the estimation of the global radiative forcing. I think that the discussion about the global radiative forcing is not necessary for this manuscript.
The reviewer raises an important point here that also was raised by reviewer #1. We now substantially expanded the explanations how we obtain the scaling factor. We felt this discussion is important after discussions at a conference where we showed preliminary results: parts of the audience misunderstood the top-of-atmosphere radiation effect over Europe 1985 to 2013 as an aerosol ERF which they compared to the usually-quoted global numbers. This of course is not correct, and so we wanted to help the reader with this short additional computation to understand what the global implications are.

Major Comment:

Line 14 of page 2: Start writing of abstract and introduction are exactly same. . . I suggest the author to change the start writing of the introduction.

We thank the reviewer for the observation. The paragraph in the Introduction section has been changed.

Line 9-10 of Page 4: There are no information about the vertical grid spacing. As well as the horizontal grid spacing, the vertical grid spacing is highly sensitive to the activation of the cloud around the cloud bottom. The author should add the information about the vertical grid spacing.
The reviewer is right that this is important information. We added the information about the vertical resolution to the model description.

Line 10-11 of Page 4: The detail information about the computational resources is not necessary.
We agree with the reviewer that is not quite necessary. However, we feel some explanation is required why we only simulate a single day, and to some readers this information maybe useful.

Line 12-13: The authors describe the weather condition of target day at this part. The weather map of the target day is helpful for readers to clarify the location of high pressure and frontal system.
The reviewer is right. We now refer to a former publication (Heinze et al. 2017) for more explanation.

Line 15-16 of Page 4: In my understanding, the resolution of ECMWF analysis data is much coarser than ICON-LEM, and it is not suitable for the initial and boundary condition for the simulation with fine grid resolution. The author should be added the detail information of the initial and boundary condition (e.g., resolution, temporal interval, the physical variables used for the initial and boundary condition). In addition, if the initial and boundary condition is much coarser than ICON-LEM model, how do the authors drive the sub-grid scale turbulence? Was the small-scale turbulence, which can be resolved by ICON-LEM but cannot be resolved by ECMWF data, reasonably reproduced after the spin-up time (after 8 hours)?
The reviewer raises an important point which needed clarification in the text. The text was overly unclear and short on this aspect. We now clarify that indeed it was driven by the COSMO-DE run at 2.8 km and run at three different nests; and refer to the Heinze et al. (2017) paper for more detail.

Line 5-6 of Page 7: As I mentioned in the general comment, the detail descriptions of about how to couple the COSMO-MUSCAT's aerosol and ICOM-LEM are necessary. The detail information about the treatment of the CCN using equations is helpful for readers.
We agree. A more detailed answer is given above as response to general comment #3. We included further information on the calculation of CCN within COSMO-MUSCAT and its usage within ICON-LEM in the revised manuscript.

Fig. 2 and Table 2: The AOD simulated with CCN of 2013 is smaller than that observed by satellite. What is the reason of the underestimation of AOD?
The reason for the deviation between model and observation is not known. In the particular case it is not necessarily an underestimation of the model, but could also be an overestimation by the AOD retrieval. The uncertainty of a single retrieved AOD value is 0.2 (see Zhao et al., 2017), which gives a large relative uncertainty for today's rather clean conditions. If there is a bias after averaging over up to 33 days per pixel is not known. The model naturally also has uncertainties. It seems, that most of the difference occurs near the coast pointing to uncertainties in the exact emissions, e.g. the amount of ships coming (or at least their emissions) out of Hamburg harbour are perhaps underestimated. Most anthropogenic emissions, such as ship tracks, need to be treated on an averaged basis (e.g., monthly or annual average broken down to the integration time step increments).

Zhao, Xuepeng; and NOAA CDR Program (2017): NOAA Climate Data Record (CDR) of AVHRR Daily and Monthly Aerosol Optical Thickness (AOT) over Global Oceans, Version 3.0. doi:10.7289/V5BZ642P.

Line 21-22 of Page 9: What is the reason of the overestimation of aerosols above the boundary layer? Is the overestimation affects the conclusion of the manuscript? I require the authors to add some comments.
Predicting vertical profiles of CCNs with COSMO-MUSCAT is particularly challenging above the planetary boundary layer, and therefore less accurate since the model tends to overestimate the vertical mixing between boundary layer and free troposphere (this information has been added into the manuscript). However, we do not think this affects the conclusion of the manuscript since even if there is an overestimation the values are inside the observations range of uncertainty.

Line 14-15 of Page 10: The authors indicate that graupel number and mass simulated by clear case are higher at height of 3 – 4 km, but the difference between solid and dotted pink line in Figure 4 is too small to be identified.
The reviewer is right that this is indeed not a very clear signal. The sentence has been changed accordingly in the revised manuscript.

Line 2 of Page 11: I think that "Distributions of liquid water path" should be "Probability density frequency (PDF) of liquid water path". Is this right?
We thank the reviewer for his/her suggestion. In this case the area below the curve is not unitary, so it cannot be called "PDF". However, these are normalized frequency of occurrence distributions and can still be interpreted as a kind of probability if one assumes that the distribution is representative for many cases.

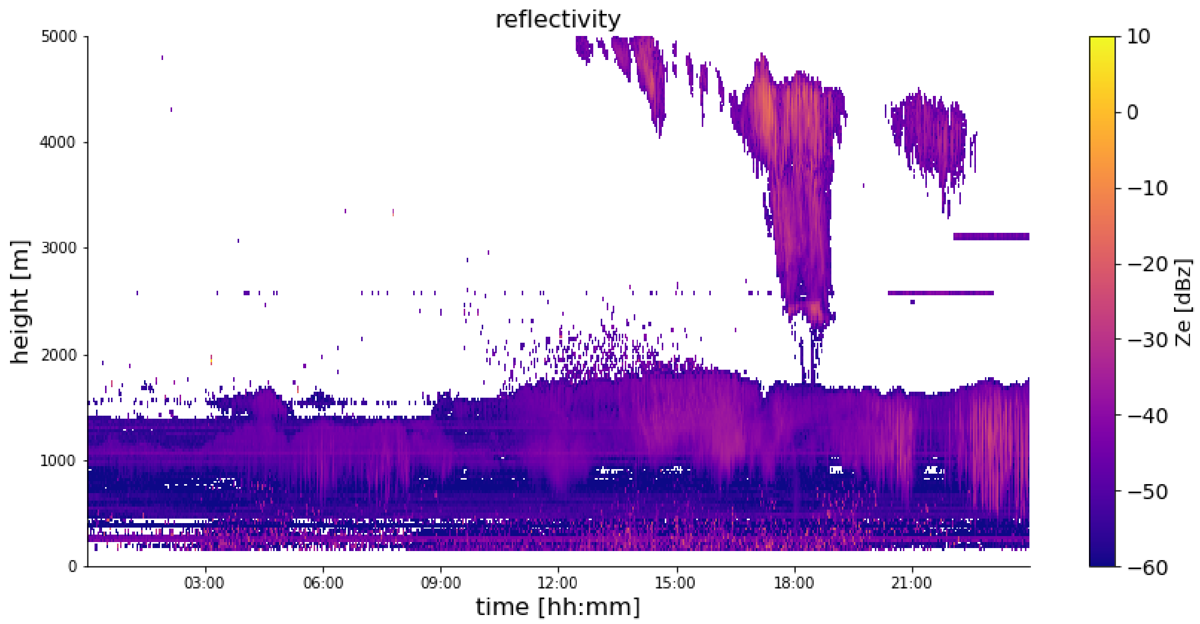
Figure 5 and Line 6-7 of Page 11: The authors suggest that the difference in PDF between the model and MODIS is originated from the sensitivity of the MODIS. However, the geographical distribution of cloud simulated by the models are largely different from that observed based on Fig. 9. I think that such difference in the geographical distribution has impacts on the PDF shown in Figure 5.
The reviewer is right, the statement too readily reads as if we blame the discrepancy entirely on the data. Instead, we now first write that the first obvious reason is that the model simulation is far from perfect, but then still remind the reader that also the retrievals are not 100% reliable.

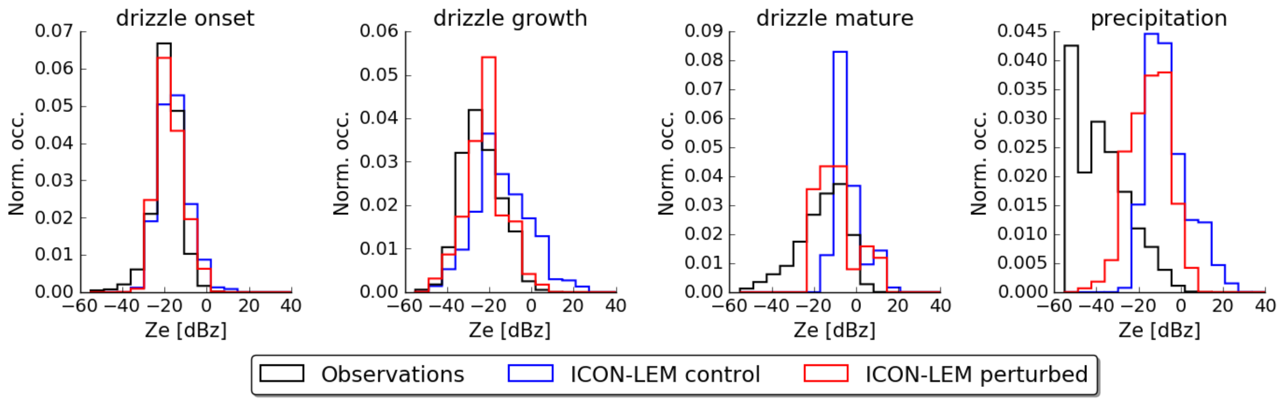
Line 8-7 of Page 16: "simulated value of reflectivities fall into the range of the observations of MOL-RAO radar" should be "mean simulated value of the reflectivities fall into the range of the observation. . .".
We thank the reviewer for his/her correction, the sentence has been changed accordingly.

Line 12-13 of Page 16: The author said the small reflectivity values of for the precipitation observations are due to noise by insects. If the authors know the signal is not originated from the precipitation, the author should remove the noise data.
We appreciate the reviewer's suggestion. Following your comment, we have re-processed the data with the ground clutter removal filtering turned on from 0 to 1600 m AGL.

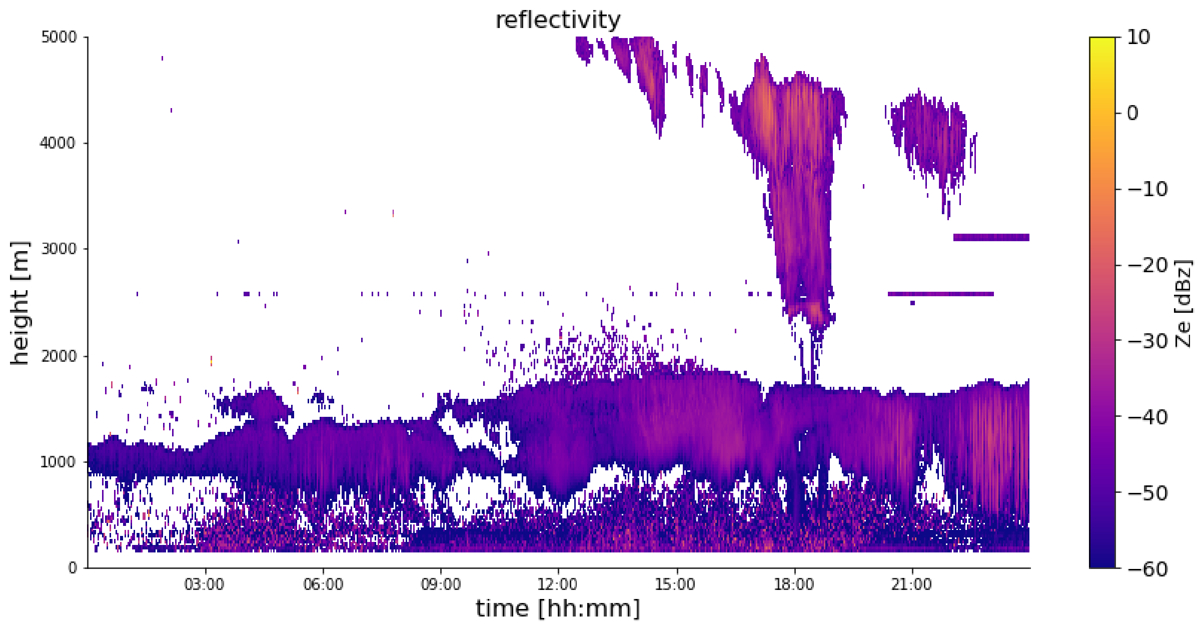
Here you can compare the two new plots of reflectivity and the old and new figures:

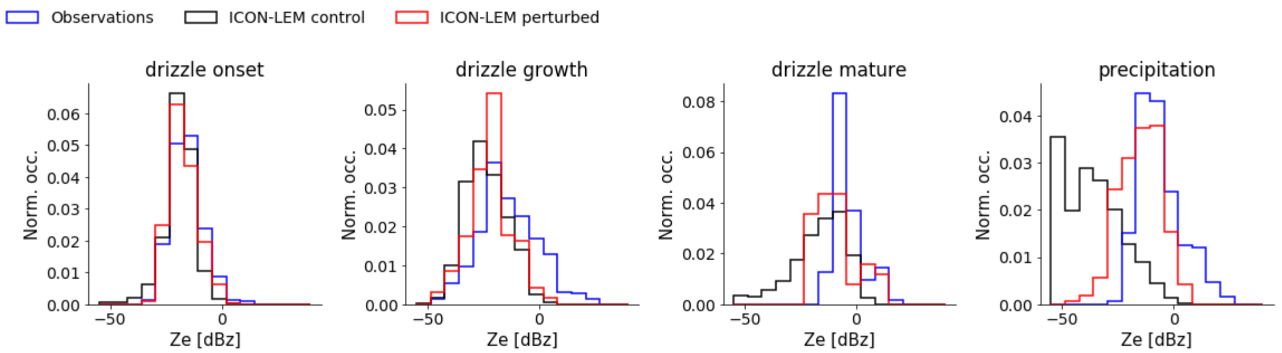
Without clutter removal:





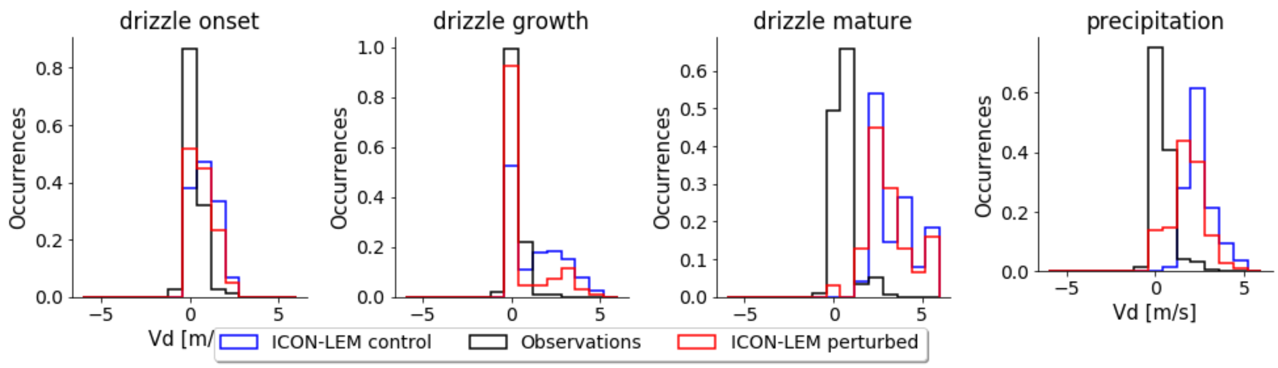
With clutter removal:

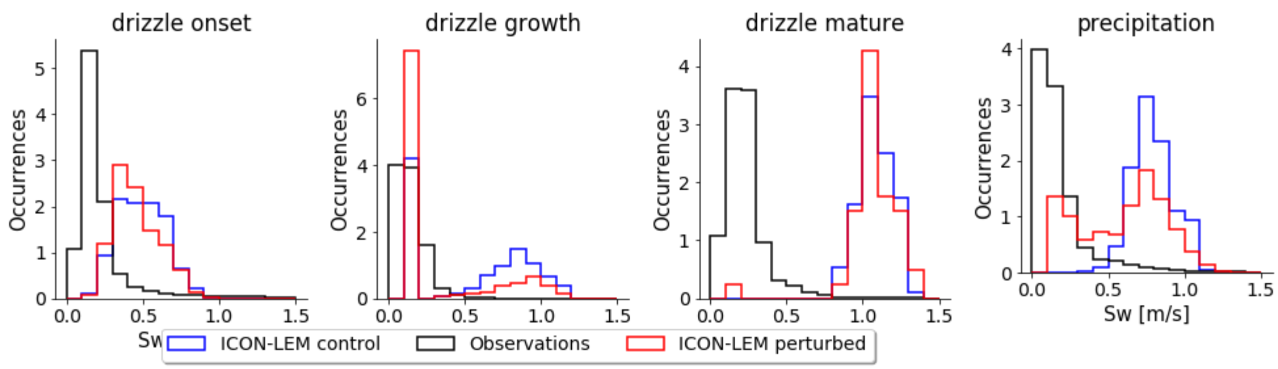




The ground clutter removal diminished the height of the precipitation peak of reflectivity (4th graph in black) at -60 dbz, however it did not shift the distribution.

By looking at higher spectral moments (see following graphs of mean Ddoppler velocity, Vd, and spectrum width, Sw), one notices that the signal has to be due to clutter, which has not entirely been removed since on that particular day there was a lot of clutter. Observed mean Doppler velocity values are very small, close to zero, indicating non-precipitating targets, typical of ground clutter, and spectrum width values are also very small, around zero. They indicate very narrow spectra shapes, which again are typical of clutter.





The text, the table and the figure in this section has been changed accordingly.

Line 14-18 page 16: I think this paragraph is not necessary.
The reviewer is right, the conclusion is not very specific. We substantially shortened the paragraph to one sentence in the revised version.

Line 20 of Page 16. How did the authors determine the cloud base height and CC simulated by the model? Was this the output of COSP? Usually, the edge of the cloud in the model is determined by a threshold value of LWP or ql. The threshold value is sensitive to the cloud cover and cloud base height. The results in Figure 8 is also sensitive to the threshold value.
We thank the reviewer for pointing out that the text was not clear enough here. As explained in section 2.3.2, the cloud base height is an output from ICON-LEM diagnostics, which is determined as the lowest cloudy grid cell of each column. The threshold for determining a cloudy grid-cell in ICON-LEM is a sum of cloud water and cloud ice (q_c and q_i) larger than 10^{-8} kg/kg. We now point the reader to this in the revised version.

Table 6: The authors indicate that the ICON simulate less cloud than observation and CBH is lower than that observed (even though the simulated CBH is included the range of 25-75th of the observation). In my understanding, such difference in the simulated and observed one is usually not originated from the problems in the model used by inner nested domain (i.e. ICON-LEM), but from the data used for initial and boundary condition (i.e. ECMWF model). So, the author should check the data used for initial and boundary condition or results of outer domain (simulation with the grid spacing of 625 m and 312 m).
We thank the reviewer for sharing her/his expertise in model skill. In light of this remark, we also checked the data in the other domain resolutions (625 and 312 m) and they give less accurate results than the highest resolution (156 m). We now quote this additional result in the revised manuscript.

Figure 9: As I mentioned in the comment for Figure 5, the difference in the geophysical distribution of simulated cloud and observed one could have some contribution to the difference in PDF shown in Fig. 5.
The reviewer is right with this important remark, and we wrote this in the main text in the revised version.

Section 3.8: As I mentioned in the general comment, the discussion in this part is too rough. Of course, I understand the importance of the estimation of radiative forcing, but the estimation of global averaged ERFaci by the scaling of the results of regional model make readers misunderstanding.
Indeed the reviewer raises a good point that this was too rough. As explained above, without this extra bit, we had the experience that some (other) readers misunderstood the computed effects since they somehow compared the regional, 1985 vs. 2013 results to the global ERFaer they had in mind. So we now considerably added information to this section to make it unambiguous and easier to grasp.

Minor Comment:

Figure 1: The color scale (color bar) is helpful for the readers.
We thank the reviewer for his/her suggestion. We fully agree that, in general, a colorbar is very helpful for the reader to understand a shown figure. In our case, we apply a RGB-type mapping in which 3 fields are mapped to a red-green-blue space. This is a very common practice in satellite remote sensing with the advantage that a lot of information can be compressed into one image and it mirrors the way the human eye observes our colored environment. The disadvantage is that no single colorbar can be provided. In our case, we apply a variant of the natural colour RGB (please see https://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=PDF_RGB_QUICK_GUIDE_NCOL&RevisionSelectionMethod=LatestReleased&Rendition=Web) which combines MSG SEVIRI channels at 0.6, 0.8 and 1.6 micron.

Line 2 of page 5: Reference and detail information of ECMWF analysis data should be added in the list of the reference.
Also in response to the reviewer remark above, we now clarified where the boundary conditions come from and provide the reference to get the additional information in the revised manuscript.

Figure 4 left: For me, it is difficult to identify Black and blue line below the height of 6 km.
We agree with the reviewer. The reason is because the black solid line is over the blue solid one, and the blue and black dashed lines are also one over the other. That means that the most part of contribution to the total condensed water particles come from cloud droplets above 6 km (and from ice particles over 6 km) for both control and perturbed simulations. We have added this information in the figure caption in order to help the reader.

Line 6-9 of Page 14: The authors removed the data of the 15 stations because these stations are too close to other stations. I think that the averaged value of the close stations is better for the comparison with the model. The representativeness of the data of selected station is not always confirmed.
Unfortunately, the reprocessing takes longer then envisioned and we are still waiting for the result. In case any significant changes to our current results occur we will adapt the final manuscript.

Reference: Khain, A. P., N. BenMoshe, and A. Pokrovsky, 2008: Factors Determining the Impact of Aerosols on Surface Precipitation from Clouds: An Attempt at Classification. J. Atmos. Sci., 65, 1721–1748, https://doi.org/10.1175/.