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 #1

General comments: The authors present analysis of new very high resolution simulations over the EU domain for one day near peak emissions in 1985 and one day in the present decade with relatively low emissions. The authors carefully analyze the high-resolution simulations using satellite and ground-based data. They find that AOD differences, and Nd differences between 1985 and 2013 are reproduced. Changes in cloud macrophysics are too small relative to natural variability to observe. The authors derive an ERFaci for the global mean using scaling with traditional GCMs. The paper is a very nice analysis of cutting-edge new simulations and provides an interesting new evaluation of ERFaci.
We thank the reviewer for their thoughtful summary of our study.

I have the following **major issues** with the paper:

A lot of the paper is given over to ground-based remote sensing. This is fine, but it is not a field that I am very familiar with and I recommend that a reviewer who is an expert be nominated to comment on this. However, I am concerned by the characterization of standard deviation as uncertainty in comparing observations and models (as discussed in specific comments) and I think this needs to be explained more clearly. I am not sure that the authors have made a meaningful comment about the adjustment strength, besides the fact that adjustments are small compared to meteorological variability and are hard to see in one day of data- which doesn't preclude them being important to ERFaci. Critically, I think the scaling to the global ERFaci could be done better (see specific comments below) by expanding the number of GCMs and by showing that the relationship is linear.
We thank the reviewer very much for his/her interesting comments. We will proceed to address all of them in the following specific comments.

Specific comments:

Pg1 Ln11: I kind of follow what the authors are trying to say here, but it is a little easy to lose track. I would suggest not using reference and perturbed to refer to 2013 and 1985 in the abstract. It will be easier to follow which conditions are consistent and inconsistent. Effectively it sounds like the model needs the appropriate year of aerosol data to get the right output, which could be said more succinctly.
We thank the reviewer for his/her appreciation. The sentence has been simplified in the abstract.

Pg2 Ln6: What is large LWP? Is it in or area-mean LWP? I am not sure what I should really be taking away from this result. Is it really a key result that needs to be shown in the abstract?
We really think this has to be said in the abstract since it is an important result. Following your suggestion, we have added "(LWP >200 g m^{-2})" in the sentence, and clarified it is in-cloud LWP (consistent with the observations).

Pg2 Ln16: The results in Rosenfeld 2019 are no longer accurate. There is an errata that revokes most of the findings of the original paper.
We thank the reviewer for the observation. The reference has been removed from the paper.

Pg3 Ln9 "to what extent".
The typo has been corrected.

Pg3 Ln17: Split this into two sentences. CCN is changed and INP is not. Direct effects are not considered. This is really confusing. How is AOD being evaluated if the direct effect isn't considered?
We thank the reviewer for his/her comment, the sentence has been divided into two and the information extended to be more clear since the direct effect (and the semi-direct) is, in fact, considered but no changes are made to it in the different simulations carried out. That means that the changes in the CCN are not affecting the direct (and semi-direct) effect radiative balance in our simulations. We are able to evaluate the AOD because of the additional offline calculations based on COSMO-MUSCAT, where the CCN number concentration of the multi-modal size distribution at a fixed supersaturation is calculated according to Abdul-Razzak and Ghan (2000), as explained in section 2.2.

Pg7 Ln17: Please discuss Song et al 2018 (https://www.geosci-modeldev.) in the context of using COSP for ICON at this resolution. What is used to drive COSP in this case? Is there subgrid variability assumed?
We thank the reviewer for the Song et al. (2018) reference; this study highlights very well the importance of properly including subgrid variability for GCM evaluation via COSP. However, the conclusions of Song et al. (2018) mainly concern the preferred usage of a model-specific sub-grid information (allowed in COSPv2) instead of the COSP sub-column generator, in order to accurately account for the GCM sub-grid cloud and hydrometeor variability (1.9x2.5deg simulations were used to reach these conclusions). This is not relevant for ICON-LEM, where no sub-grid variabilities of cloud and aerosol properties are considered. Consequently, COSP was used without subcolumn varibility and was driven directly by the grid-level ICON outputs. The goal is here to apply the satellite retrieval algorithm on a pixel level, mainly to reproduce the instrumental sensitivity limitations (i.e. which clouds are too thin to be detected, where does the signal saturate?). Note that the 156-m icosahedral outputs were aggregated into a 1-km^2 lat-lon grid (fitting the MODIS resolution) prior to being used in COSP, but we still decided to not to include sub-grid (< 1km) cloud variability to stay consistent with the MODIS retrieval algorithm, which performs its retrievals by ignoring sub-pixel variabilities.

A new sentence has been added into the "Observation / Satellite-based" subsection: "No sub-column variability is used in COSP, consistently with the lack of sub-pixel variability in MODIS retrievals".

Pg9 Ln9: Typo- sentence needs to be reordered. Maybe "AOD is only available over the North Sea region for xx% of retrievals." To reduce ambiguity.
The sentence has been reformulated.

Pg9 Ln17: The authors show a systematic difference in the mean CCN profiles from observations and the CCN used to drive the model. I think the authors are somewhat misusing the uncertainty range. Don't you want uncertainty in the mean, not just the variability, which is what this shows? Shouldn't these be standard error in the mean? Ultimately it seems like there is a 10-30% overestimate in CCN relative to the observations (I assume the standard error in the mean is small). Can the authors convert that to an overestimate in Nd using the nucleation scheme, which is the more relevant quantity in this study?
We appreciate the reviewer's suggestion. In fact, we chose to show the median and the 25-75th percentiles in purpose because we think they better show the variability of the AOD in the region, rather than the mean and the standard deviation. In this sense, we have removed the word "uncertainty" in Table 2 caption to avoid misunderstandings. Regarding the Nd inquiry, it is discussed in Fig. 3.

Pg 12 Ln8: I am not sure that 10% change in LWP is small (am I reading table 3 right?). It's certainly true that variability in LWP due to meteorological variability is large, but this doesn't really tell us anything about the radiative forcing induced by adjustments.
The reviewer is right, 10 % change is not a small change, however it is too small for detection and attribution of LWP changes by satellite, considering current retrieval uncertainties, therefore the change in LWP is not detectable by MODIS on the studied case, and this is what we mean in the sentence. We clarify that the LWP change translates into a substantial systematic effect on the radiation balance and, thus, the aerosol effective radiative forcing.

Pg12 Ln11 Is the cutoff for large LWP? Does this just mean not thin clouds?
The 200 g m^{-2} value refers to the analysis of Fig. 5. At the large-LWP tail of the PDFs, an increase of high LWP values (higher than about 200 g m-2) clearly appears in the perturbed simulation by comparison to the reference and satellite observations. As explained in this paragraph, we attribute this adjustment effect to invigoration of convective clouds as a consequence of higher Nd. This observation is of particular interest because such adjustments could in principle be detectable based on MODIS-like satellite retrievals. We clarify this in the revised manuscript.

Pg12 Ln18: How many more days of simulation would you need to beat down the noise and be able to see the LWP perturbation clearly?
We could only speculate, since such a method hasn't been applied yet to large-domain large-eddy simulations. In GCM analyses, even for nudged simulations, yearlong integrations are necessary. In an LES we believe a shorter analysis is sufficient due to the very much larger amount of independent columns.

Pg 20 Ln29 I think the authors are just calculating the ERF over Europe versus the global mean and coming up with a scaling factor. I think a better approach would be to plot ERF_EU_1985-ERF_EU_2013 versus ERF_global_mean_PD for each CMIP5 model. The way that the authors are doing this assumes linearity in this relationship, which is not necessarily true since the EU in 1985 is so polluted. Based on Carslaw et al. (2013), I am not sure that this calculation should really reduce uncertainty much, but Carslaw et al. (2013) paper implies strong non-linearity in the relationship between local ERF and global-mean ERF. If the authors could increase the number of GCMs beyond 4 and show that the relationship is linear this would be a more robust calculation.
How do the authors deal with the direct effect not being calculated in the simulations for this comparison since it will be in the GCMs (Pg3 Ln17)?

We thank the reviewer for highlighting that this point requires more attention; reviewer #2 had a very similar concern. The reviewer indeed was right that our previous analysis was overly superficial. Fortunately in the meanwhile, the new 6th Coupled Model Intercomparison Project (CMIP6) provided output from the new multi-model ensemble. This is very valuable to the problem here in question since the part of CMIP6 that addresses the radiative forcing (the RFMIP) has one simulation that allows to diagnose the transient ERF due to aerosols. From this new output, we were now able to assess the scaling in a more thorough way. We explain now in the revised manuscript in much more detail the revised procedure to scaling the forcing, and – more importantly perhaps still in response to this reviewer remark – we much better highlight and quantify the uncertainties. The new approach also allows to better isolate the aerosol-cloud interactions.

Carslaw, K. S., Lee, L. A., Reddington, C. L., Pringle, K. J., Rap, A., Forster, P. M., . . . Pierce, J. R. (2013). Large contribution of natural aerosols to uncertainty in indirect forcing. Nature, 503(7474), 67-71. doi:10.1038/nature12674

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.

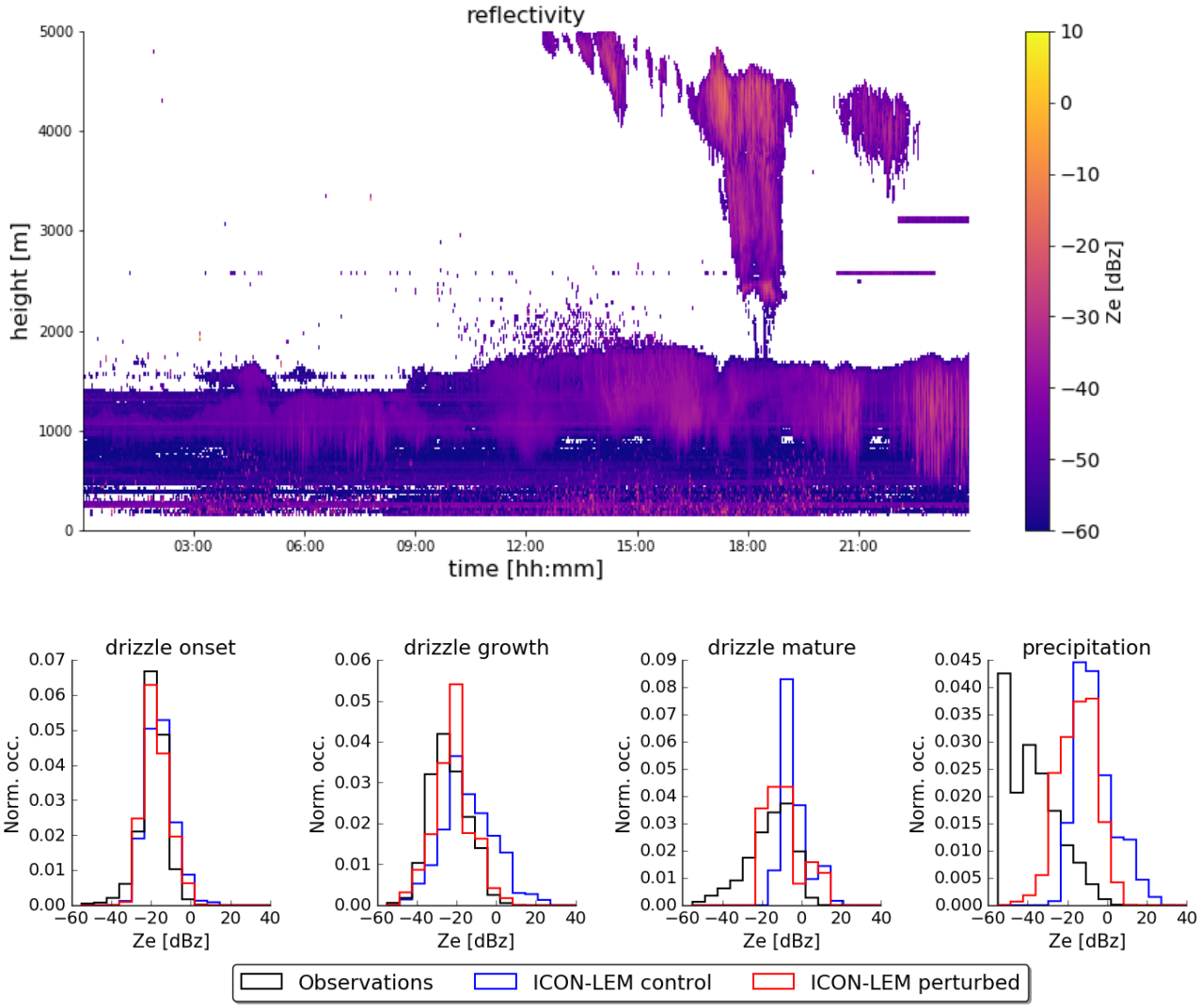
Figure 5 and Line 6-7 of Page 11: The authors suggest that the difference in PDF between the model and MODIS is originated from the sensitivity of the MODIS. However, the geographical distribution of cloud simulated by the models are largely different from that observed based on Fig. 9. I think that 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.

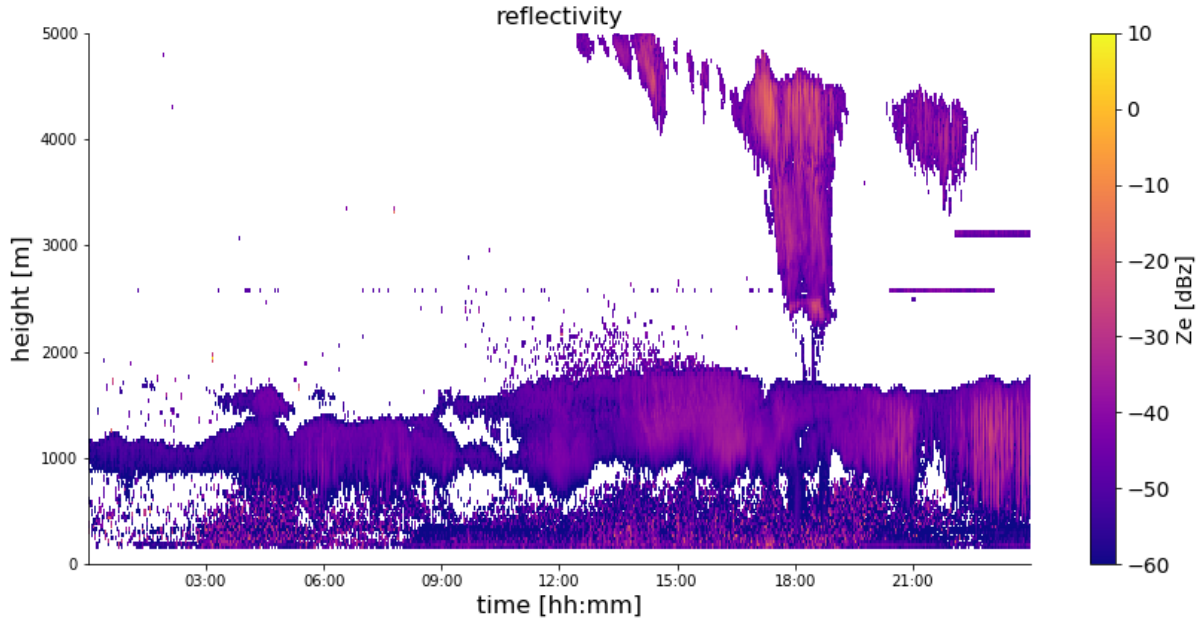
Line 12-13 of Page 16: The author said the small reflectivity values of for the precipitation observations are due to noise by insects. If the authors know the signal is not originated from the precipitation, the author should remove the noise data.
We appreciate the reviewer’s suggestion. Following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

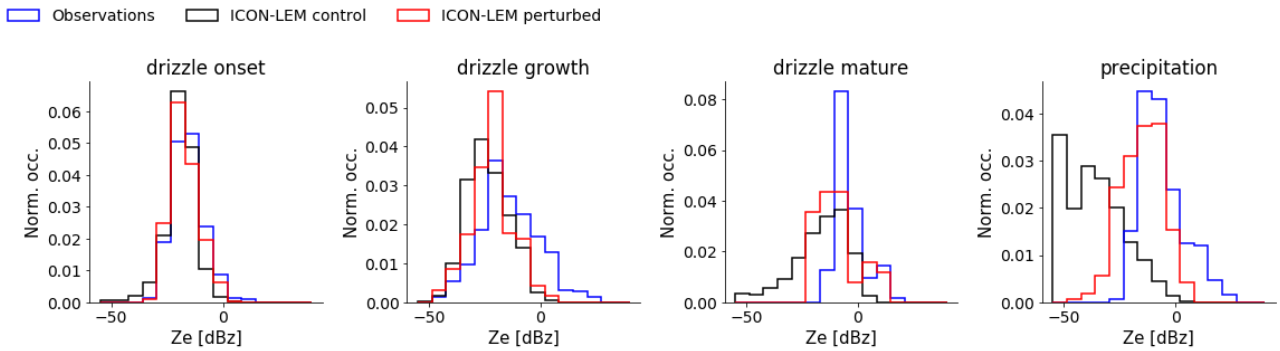
Here you can compare the two new plots of reflectivity and the old and new figures:

Without clutter removal:



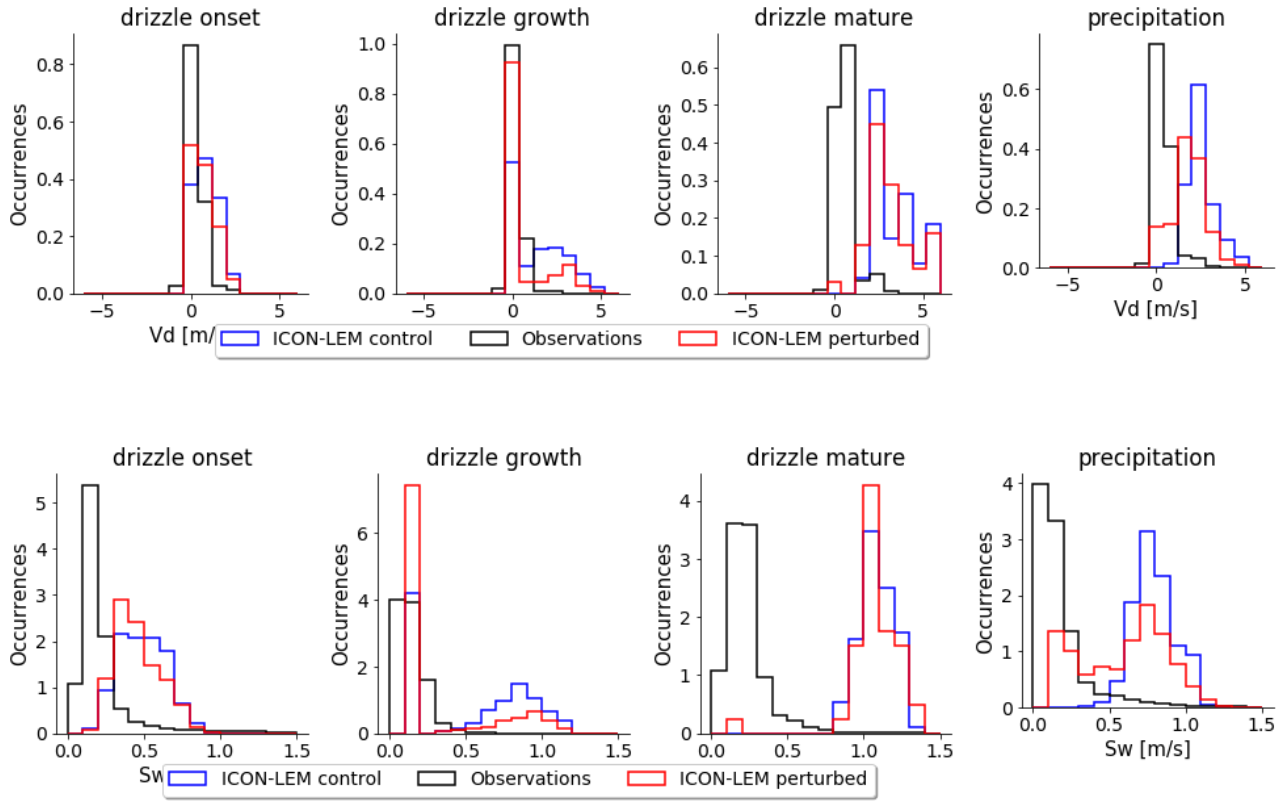
With clutter removal:





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.

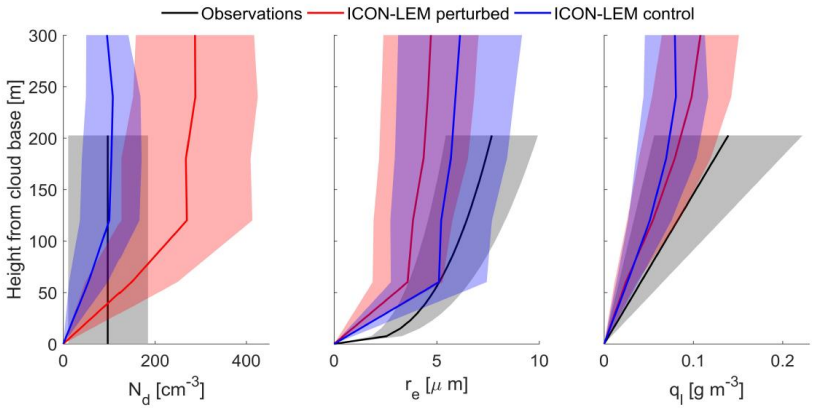
Following the reviewer's suggestion, figure 6 has been updated including all stations available. The following table show the 15 stations that have been added in order to update the statistics for plotting figure 6:

Station name	**Latitude (°N)**	**Longitude (°E)**
JOYCE	50.909	6.414
KIT_HOPE	50.897	6.464
De_Bilt	52.100	5.180
Lindenberg	52.210	14.110
HOPE-1	50.933	6.388
HOPE-2	50.908	6.413
HOPE-3	50.899	6.459
HOPE-4	50.880	6.414
HOPE-5	50.897	6.399
HOPE-6	50.864	6.425
HOPE-7	50.894	6.402
HOPE-8	50.873	6.461
HOPE-9	50.925	6.423
HOPE	50.884	6.451
Muenchen_EDDM	48.354	11.775

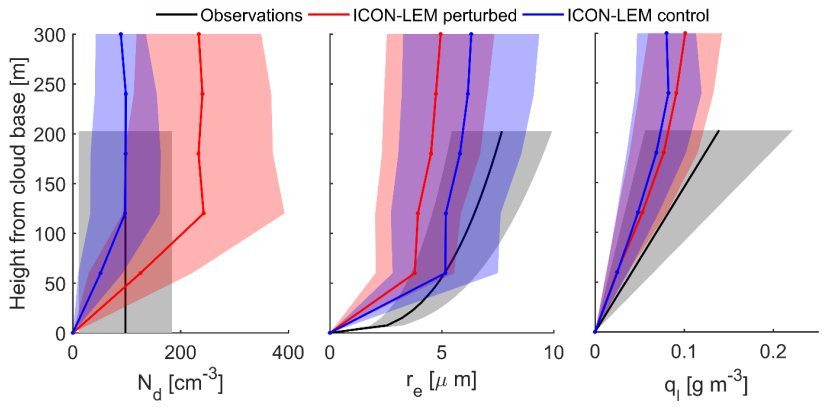
Incorporating these additional stations means, two-fold analysis of "De_Bilt", "Lindenberg" and "Muenchen_EDDM" stations (because we were already using a 'twin'-dataset of each station: "Cabauw", "RAO", "Muenchen-Oberschleissheim", correspondingly), and a 12-fold analysis of the HOPE-Jülich area (so far, we were using the gridpoint "LACROS-HOPE").

The following plots show the comparison of the original figure 6 (a) and the updated figure 6 including all stations (b) for the analysis of the liquid clouds:

(a) Selected meteogram stations (as in submitted manuscript)



(b) All meteogram stations (as requested by reviewers)



As the reviewer can see, there's almost no difference between the two figures. Therefore, the manuscript has not been changed in the view of these new results.

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/.

List of changes made in the original manuscript Costa-Surós et al. submitted on ACP

Page 1, line 11:
Original: It first is demonstrated, using satellite aerosol optical depth retrievals available for both 1985 and 2013, that the aerosol fields for the reference conditions and also for the perturbed ones, as well as the difference between the two, were consistent in the model and the satellite retrievals.
Revised: It is first demonstrated ~~using satellite aerosol optical depth retrievals available for both 1985 and 2013,~~ that the aerosol fields ~~for the reference conditions and also for the perturbed ones, as well as the difference between the two, were consistent in the model and the satellite retrievals.~~ used in the model are consistent with corresponding satellite aerosol optical depth retrievals for both 1985 (perturbed) and 2013 (reference) conditions.

Page 2, line 7:
Original: An exception to this is the fact that at large liquid water path, the control simulation matches the observations, while the perturbed one shows too large LWP.
Revised: An exception to this is the fact that at large liquid water path (LWP >200 g m−2), the control simulation matches the observations, while the perturbed one shows too large LWP.

Pag 2, line 14:
Original: Clouds and aerosols contribute the largest uncertainty to estimates and interpretations of the Earth's changing energy budget (Boucher et al., 2013).
Revised: According to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), clouds and aerosols are largest contributors to uncertainty estimations ~~Cloud and aerosols contribute the largest uncertainty to estimates~~ and interpretations of the Earth's changing energy budget (Boucher et al., 2013).

Page 2, line 16:
Original: In particular, aerosol-cloud interactions continue to be a challenge for climate models and consequently for climate change predictions (Stevens and Feingold, 2009; Feingold et al., 2016; Seinfeld et al., 2016; Rosenfeld et al., 2019).
Revised: In particular, aerosol-cloud interactions continue to be a challenge for climate models and consequently for climate change predictions (Stevens and Feingold, 2009; Feingold et al., 2016; Seinfeld et al., 2016; ~~Rosenfeld et al., 2019~~).

Page 3, line 12:
Original: The key idea is to assess to which extent the model-simulated aerosol-cloud interaction effects might be detected and attributed in comparison to various observational dataset.
Revised: The key idea is to assess to what ~~which~~ extent the model-simulated aerosol-cloud interaction effects might be detected and attributed in comparison to various observational dataset.

Page 3, line 19:
Original: Note that in the present study only modifications to the CCN concentration have been taken into account, and not to ice nucleating particles (INP), neither to scattering, nor to the absorbing aerosol properties (aerosol-radiation interactions, in previous literature referred to as direct and semi-direct aerosol effects).
Revised: Note that in the present study only modifications to the CCN concentration have been taken into account, and not to ice nucleating particles (INP) ~~neither to~~. As well, the changes made to the CCN are not affecting the scattering or ~~and nor to~~ the absorbing aerosol properties (aerosol-radiation interactions, in previous literature referred to as direct and semi-direct aerosol effects), but only the aerosol-cloud interactions (in previous literature also called first aerosol indirect effect,

cloud albedo effect, or Twomey effect, and the cloud adjustments, such as the cloud lifetime effect or other rapid responses).

Page 4, line 7:
Sentences added: CCN are prescribed in the study (see next Section for more details) as temporally and spatially varying fields read in from offline calculations. The model version used here, however, has been updated to allow for the consumption scavenging of CCN, at droplet activation, a CCN is scavenged. For replenishment, CCN concentrations outside clouds are then relaxed back to the prescribed distribution with a relaxation time scale of 10 minutes.

Page 4, line 10:
Sentence added: In the vertical, 150 levels are used, with grid stretching towards the model top at 21 km. The minimal layer thickness is 20 m near the surface and the lowest 1000 m encompass 20 layers.

Page 4, line 19:
Sentence added: More details about the weather conditions are given in Heinze et al. (2017).

Page 4, line 20:
Original: [...] simulations were initialized at 0 h UTC on 2 May 2013 from boundary conditions from the European Centre for Medium-Range Weather Forecast (ECMWF) analysis."
Revised: […] simulations were initialized at 0 h UTC on 2 May 2013 from ~~from boundary conditions from the European Centre for Medium-Range Weather Forecast (ECMWF) analysis.~~ simulations with the COSMO-DE model at 2.8 km resolution, and were run in three nests from coarse (624 m horizontally) to intermediate (312 m) to fine (156 m) resolution as described in Heinze et al. (2017).

Page 4, line 22:
Sentence added: In particular, the model output used in the study was on one hand, the so-called meteograms which were temporarily high-resolved (9 s) and available at 36 station locations; and on the other, 2D (at 1 min resolution) and 3D (available every 30 min and at 150 vertical levels) whole domain data fields.

Page 7, line 5:
Original: Details on the used hygroscopicity parameters as well as the derivation of number size distributions from the simulated speciated aerosol mass can be found in Genz et al. (2019).
Revised: The calculation of the CCN number concentration (number of activated aerosol particles) follows Hande et al. (2016) and Genz et al. (2019), using the parameterization by Abdul-Razzak and Ghan (2000) for multi-modal size distributions. Details on the used hygroscopicity parameters as well as the derivation of number size distributions from the simulated speciated aerosol mass can be found in Genz et al. (2019). For the comparison to available observations, the CCN number concentration field at a fixed supersaturation is calculated. However, in order to provide CCN fields for ICON-LEM, time varying 3D fields of the CCN number concentration at different constant updraft velocities was required. Abdul-Razzak and Ghan (2000) relate the aerosol composition and updraft velocity to the maximum supersaturation during an air parcels ascend, which in the end determines the number of activated aerosol particles. The calculated CCN fields were then used in the ICON-LEM simulations replacing the fixed assumed CCN number concentration/distribution.

Page 7, line 24:
A new sentence has been added: No sub-column variability is used in COSP, consistently with the lack of sub-pixel variability in MODIS retrievals.

Page 8, line 3:
A new sentence has been added: for measurements during the HOPE campaign (Section 3.1)

Page 8, line 6:
A new sentence has been added: and allow a climatological assessment (Section 3.4)

Page 8, line 15:
Original: The cloud radar (8.6 mm wavelength) also based in the (MOL-RAO)
Revised: The cloud radar (8.6 mm wavelength) based in the Meteorological Observatory Lindenberg – Richard Aßmann-Observatorium (MOL-RAO)

Page 8, line 29:
Sentence added: The CBH 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^{-1}.

Page 9, line 12:
Original: Note that since AVHRR retrieves AOD only in cloudlesscases over sea, only a rather small fraction of the time a valid retrieval is available (Table 2).
Revised: Note that since AVHRR retrieves AOD only in cloudless cases over sea, AOD is only available over the North sea and Baltic sea region for a small fraction of the time (between 10-12 % in 1985, and 35-38 % in 2013, ~~a rather small fraction of the time a valid retrieval is available~~ Table 2).

Page 9, line 28:
Original: Above the boundary layer, the overestimation increases to more than a factor of 2.
Revised: Above the boundary layer, the overestimation increases to more than a factor of 2 since the model tends to overestimate the vertical mixing between boundary layer and free troposphere (Heinold et al., 2011a).

Page 9, Table 2 caption:
Original: Uncertainty ranges are provided as 25th and 75th percentiles of regional variability of the temporal mean AOD.
Revised: ~~Uncertainty ranges are provided as~~ 25th and 75th percentile ranges of regional variability of the temporal mean AOD are provided in brackets.

Page 10, line 6
Original: 3.2 Mean vertical profiles of number and mass concentrations [...]
Revised: 3.2 Mean vertical hydrometeor profiles of number and mass concentrations [...]

Page 10, line 7:
Original: Domain-averaged profiles of [...]
Revised: Domain-averaged hydrometeor profiles of [...]

Page 10, line 9:
Original: relative increase of 146.9 %
Revised: relative increase of 147 %

Page 11, line 2:
Original: Distributions of liquid water path […]
Revised: Normalized frequency of occurrence distributions of liquid water path […]

Page 11, line 4:
Original: In turn, graupel number and mass concentrations are higher in the cleaner environment in low to mid altitudes (3 – 4 km).
Revised: In turn, graupel number and mass concentrations are slightly higher in the cleaner environment in low to mid altitudes(3 – 4 km).

Page 11, line 7:
Original: […] compared to MODIS (Table 4). A reason could be the MODIS instrument sensitivity, since optically […]
Revised: […] compared to MODIS (Table 4). This can be partly explained by a too large range simulated by the model, which in turn can be partly related to a difference in observed and simulated spatial distribution of clouds at the MODIS observation times. However, a part of the difference in the range of the distributions can be very likely attributed to the MODIS instrument characteristics, since optically […].

Page 12, line 7:
Original: […] smooth PDF.
Revised: […] smooth distribution.

Page 12, line 12:
Original: An exception is at large LWP (larger than 200 g m^{-2}), where […].
Revised: An exception is at large LWP (larger than 200 g m^{-2}, see Fig. 5), where […].

Page 12, line 18:
Sentence added: The systematic change in LWP, even if small compared to weather variability, implies a substantial contribution to the aerosol effective radiative forcing.

Page 13, Figure 4 caption:
Note added: Also note that in the left panel the black solid line is over the blue solid one, and the blue and black dashed lines are also one over the other.

Page 15, line 14:
Original: the Meteorological Observatory Lindenberg – Richard Aßmann-Observatorium (MOL-RAO)
Revised: the MOL-RAO

Page 15, Figure 6 caption:
Note added: Note that measurements can only provide profiles up to 200 m into the cloud due to the strong extinction of the Lidar signal

Page 16, line 8:
Original: The simulated values of reflectivities, [...], fall into the range of the observations of MOL-RAO radar, [...]"
Revised: The mean simulated value of the reflectivities, [...], fall into the range of the observation of MOL-RAO radar, [...].

Page 16, line 12:
Original: A closer look into the radar signal suggests that the small reflectivity values for precipitation observations are due to insects detected by the radar.
Revised: A closer look into the radar signal suggests that the small reflectivity values for precipitation observations are due to insects detected by the radar, despite a clutter removal filter was applied to the radar spectra from 0 to 1600 m AGL as pre-processing.

Page 16, Table 5:
Original:

Table 5. Median as well as 25th and 75th percentiles of the cloud radar reflectivity (Ze) distributions on 2 May 2013 for the four classes described in Fig. 7.

	Drizzle onset Mean [25th to 75th]	Drizzle growth Mean [25th to 75th]	Drizzle mature Mean [25th to 75th]	Precipitation Mean [25th to 75th]
MOL-RAO cloud radar	–18.4 [–22.0 to –15.5]	–24.4 [–30.5 to –17.2]	–13.9 [–22.7 to –7.3]	–40.1 [–51.2 to –30.4]
ICON-LEM control (C2R)	–16.2 [–20.5 to –11.7]	–15.6 [–22.3 to –5.2]	–4.9 [–7.0 to –2.3]	–8.1 [–13.5 to –0.9]
ICON-LEM perturbed (P2R)	–17.8 [–21.5 to –13.4]	–21.5 [–27.4 to –16.9]	–9.9 [–15.1 to –5.5]	–14.4 [–21.6 to –8.2]

Revised :

Table 5. Median as well as 25th and 75th percentiles of the cloud radar reflectivity (Ze) distributions on 2 May 2013 for the four classes described in Fig. 7.

	Drizzle onset Mean [25th to 75th]	Drizzle growth Mean [25th to 75th]	Drizzle mature Mean [25th to 75th]	Precipitation Mean [25th to 75th]
MOL-RAO cloud radar	–18.4 [–22.1 to –15.4]	–24.2 [–30.3 to –17.0]	–14.2 [–23.1 to –7.5]	–38.2 [–49.7 to –28.2]
ICON-LEM control (C2R)	–16.2 [–20.5 to –11.7]	–15.6 [–22.3 to –5.2]	–4.9 [–7.0 to –2.3]	–8.1 [–13.5 to –0.9]
ICON-LEM perturbed (P2R)	–17.8 [–21.5 to –13.4]	–21.5 [–27.4 to –16.9]	–9.9 [–15.1 to –5.5]	–14.4 [–21.6 to –8.2]

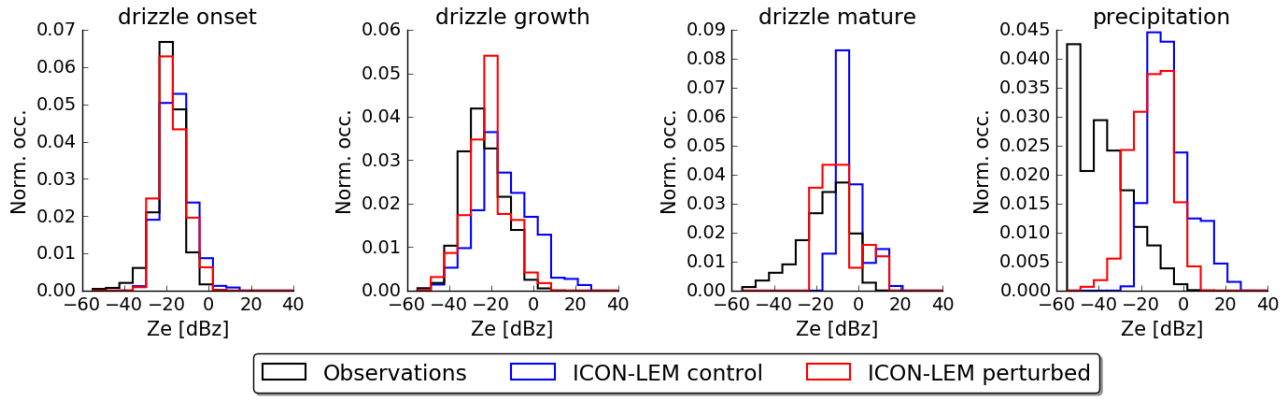
Page 16, line 14-18:
Original: The comparison between cloud reflectivities (and Doppler velocities and spectrum width) to the corresponding forward modeled variables from ICON-LEM simulations shows a fair agreement between observations and model. However, despite the effort to make model and data comparable, and despite the rather strong signal in the model, this is not yet a useful tool for detection and attribution of differences between control (C2R) and perturbed simulations (P2R) in this case. This tool is at a preliminary stage and can be used to evaluate microphysical schemes to observations (Acquistapace, 2017).
Revised: In conclusion, despite the effort to make model and data comparable using the forward operator, and despite the rather strong signal in the model, no detection and attribution of an aerosol signal could be achieved. In the future more comparisons are needed as difference at one grid point only could arise from a different sampling of cloud life cycle.

Page 16, line 22:
Original: […] are compared with high-resolution ceilometer measurements (15 s temporal resolution) from the DWD ceilometer network.
Revised: […] are compared with high-resolution ceilometer measurements (15 s temporal resolution) from the DWD ceilometer network (please see Section 2.3.2 for details).

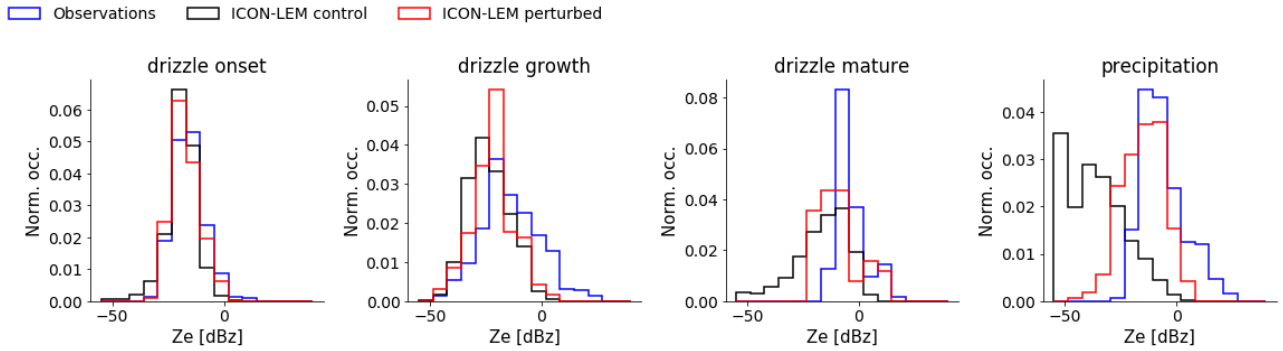
Page 16, line 24:
Sentence added: The problem is not due to issues with the initial- or boundary conditions, as the discrepancy to the reference observations is larger in the outer nests (312 and 624 m resolution, respectively; result not shown here).

Page 17, Figure 7:

Original:



Revised:



Page 17, Figure 7 caption:

Original: […] as in Acquistapace et al. (2017) [...]

Revised: […] as in Acquistapace et al. (20197) [...]

Page 19, line 16:

Sentence added: A part of the difference in simulated and observed cloud types can be attributed to a difference in the spatial distribution of the cloud types at the MODIS overpass time.

Page 19, line 18:

Original: Figure 10 displays PDFs of Nd, [...]

Revised: Figure 10 displays normalized frequency of occurrence distributions of Nd, [...]

Page 20, line 29:

Original: In order to put this number into context, we assessed the aerosol ERF from four different models from the 5th Coupled 30 Model Intercomparison Project (CMIP5; Taylor et al., 2012) for which the relevant output diagnostics were available to infer the time series of the aerosol ERF. This time series of the ERF over the industrial period was computed by Kretzschmar et al. (2017). The average aerosol ERF (both, aerosol-cloud and aerosol-radiation interactions; the latter are not considered in our current modelling study) for the annual average 1850 to 2000 (present-day minus pre-industrial) as global annual mean was -1.2 W m−2. The aerosol ERF over central Europe (the domain investigated here) in May, averaged over 1983 to 1987, was -4.0 W m−2. This latter is larger than the global-annual average present-day vs. pre-industrial, since it is over a region with a large local aerosol perturbation which is even much larger than present-day minus pre-industrial, and the solar zenith angle in May is larger than in the annual mean. In total, a scaling factor of 3.4 is obtained. The ICON-LEM thus implies a global annual mean ERF due to aerosol-cloud interactions in 2000 of about -0.8 W m−2.

Revised: In order to put this number into context, we assess the ERF computed by general circulation models. Within the 6th Coupled Model Intercomparison Project (CMIP6; Eyring et al., 2016), the Radiative Forcing Model Intercomparison Project (RFMIP; Pincus et al., 2016) defined a simulation dedicated to the assessment of the transient historical effective radiative forcing, the simulation "RFMIP-ERF-HistAerO3". There are four models in the CMIP6 archive that submitted output for these simulations, namely the CanESM5 (Swart et al., 2019), the GFDL-CM4 (Held et al., 2019), the MIROC6 (Tatebe et al., 2019), and the NorESM2-LM (Bentsen et al., 2013; Kirkevåg et al., 2018), which supplied 3, 1, 3, and 2 ensemble members, respectively. The ERFaci is approximated by using the change in cloud radiative effect (CRE, the difference between all-sky and clear-sky top-of-atmosphere net radiation flux density here taken in the solar spectrum only) between two time periods (Quaas et al., 2009). When evaluating the difference in solar CRE of the individual years 2013 and 1850, one obtains as the difference in global annual mean, a multi-model mean of -0.81 Wm^{-2}, with an inter-model standard deviation of 0.34 Wm^{-2} (using all ensemble members). For a five-year average difference(2010 to 2014 and 1850 to 1854; the periods are not centered around the specific years because the simulations run from 1850 to 2014; Pincus et al., 2016), the values for the global annual mean are -0.71 ± 0.35 W m^{-2}. For the domain of the ICON-LEM simulation, and only using May (only monthly output is available), the signal, defined as the difference 1985 minus 2013 is much more noisy since it is averaged much less. The mean and standard deviation are -4.69±13.05 Wm^{-2}. To assess the uncertainty, we computed also the change in solar CRE for the months April and May (since the actual day is early May), averaged over the five year-periods 1983 to 1987 minus 2010 to 2014, and a larger domain (10°W to 30°E, 40° to 60°N). This yields a smaller value and much smaller standard deviation of -2.55±2.99 Wm^{-2}. The scaling factor for the five-year and bigger European domain is 3.6; the one for the single years and ICON-LEM domain is 5.8. The uncertainty in these scaling factors obtained from the GCMs is very large. Nevertheless it may be instructive to know that the forcing for May, considering the large difference in aerosol levels over Europe between 1985 and 2013 is a factor of 4 to 6 larger than considering the global ERFaci between 2013 and 1850. The -2.6 W m^{-2} obtained in this simulation (Table 7) thus would imply a global, annual mean ERFaci for 2013 vs. 1850 of between -0.4 and -0.7 Wm-2.

Page 21, Fig. 10 caption:
Original: Figure 10. Probability distributions of […]. The numbers of the right side show the number of pixels used for the PDF calculation.
Revised: Figure 10. Normalized frequency of occurrence distributions of […]. The numbers of the right side show the number of pixels used for the normalized frequency of occurrence distribution calculation.

Page 22, line 26:
Sentence added: Although the simulations in this study are limited to one day over the domain of Germany, this work shows the great potential of combining these new high resolution simulations with a large set of observations for the detection and attribution of aerosol-cloud interactions. In the future this work should be complemented by extended analyses for longer time periods and more regions to further improve our understanding of cloud-aerosol interactions.

Page 22, line 31:
Sentence added: The authors thank two anonymous reviewers for their constructive remarks.

Detection and attribution of aerosol-cloud interactions in large-domain large-eddy simulations with ICON

Montserrat Costa-Surós[1], Odran Sourdeval[2,3], Claudia Acquistapace[1], Holger Baars[4], Cintia Carbajal Henken[5], Christa Genz[2,4], Jonas Hesemann[6], Cristofer Jimenez[4], Marcel König[4], Jan Kretzschmar[2], Nils Madenach[5], Catrin I. Meyer[7], Roland Schrödner[4], Patric Seifert[4], Fabian Senf[4], Matthias Brueck[8], Guido Cioni[8], Jan Frederik Engels[9], Kerstin Fieg[9], Ksenia Gorges[9], Rieke Heinze[8], Pavan Kumar Siligam[9], Ulrike Burkhardt[10], Susanne Crewell[1], Corinna Hoose[6], Axel Seifert[11], Ina Tegen[4], and Johannes Quaas[2]

[1]Universität zu Köln
[2]Universität Leipzig
[3]Université de Lille, CNRS, UMR 8518 – Laboratoire D'Optique Atmospherique
[4]Leibniz Institute for Tropospheric Research
[5]Institute for Space Sciences, Freie Universität Berlin
[6]Karlsruhe Institute of Technology
[7]Jülich Supercomputing Centre, Forschungszentrum Jülich
[8]Max Planck Institute for Meteorology
[9]Deutsches Klimarechenzentrum
[10]Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre
[11]Deutscher Wetterdienst

Correspondence: Montserrat Costa Surós (mcostasu@uni-koeln.de)

Abstract.

Clouds and aerosols contribute the largest uncertainty to current estimates and interpretations of the Earth's changing energy budget. Here we use a new-generation large-domain large-eddy model, ICON-LEM, to simulate the response of clouds to realistic anthropogenic perturbations in aerosols serving as cloud condensation nuclei (CCN). The novelty compared to previous studies is that (i) the LEM is run in weather prediction mode and with fully interactive land surface over a large domain, and (ii) a large range of data from various sources are used for the detection and attribution. The aerosol perturbation was chosen as peak-aerosol conditions over Europe in 1985, with more than five-fold more sulfate than in 2013. Observational data from various satellite and ground-based remote sensing instruments are used aiming at a detection and attribution of this response. The simulation was run for a selected day (2 May 2013) in which over the selected domain of central Europe a large variety of cloud regimes was present.

It is first demonstrated that the aerosol fields used in the model are consistent with corresponding satellite aerosol optical depth retrievals for both 1985 (perturbed) and 2013 (reference) conditions. In comparison to retrievals from ground-based lidar for 2013, CCN profiles for the reference conditions were consistent with the observations, while the ones for the 1985 conditions were not.

Similarly, detection-and-attribution was successful for droplet number concentrations: the ones simulated for the 2013 conditions were consistent with satellite as well as new ground-based lidar retrievals, while the ones for the 1985 conditions were outside the observational range.

For other cloud quantities, including cloud fraction, liquid water path, cloud-base altitude, and cloud lifetime, the aerosol response was small compared to their natural variability. Also, large uncertainties in satellite and ground-based observations make the detection-attribution difficult for these quantities. An exception to this is the fact that at large liquid water path (LWP $> 200\,\mathrm{g\,m^{-2}}$), the control simulation matches the observations, while the perturbed one shows too large LWP.

The model simulations allowed to quantify the radiative forcing due to aerosol-cloud interactions, as well as the adjustments to this forcing. The latter were small compared to the variability and showed overall a small positive radiative effect. The overall effective radiative forcing (ERF) due to aerosol-cloud interactions (ERFaci) in the simulation was dominated thus by the Twomey effect and yielded for this day, region, and aerosol perturbation -2.6 $\mathrm{W\,m^{-2}}$. Using general circulation models to scale this to a global-mean present-day vs. pre-industrial ERFaci yields a global ERFaci of -0.8 $\mathrm{W\,m^{-2}}$.

1 Introduction

According to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), clouds and aerosols are largest contributors to uncertainty estimations and interpretations of the Earth's changing energy budget (Boucher et al., 2013). In particular, aerosol-cloud interactions continue to be a challenge for climate models and consequently for climate change predictions (Stevens and Feingold, 2009; Feingold et al., 2016; Seinfeld et al., 2016). Changes in aerosol burden have an effect on cloud microphysical properties, with more aerosol leading to more numerous cloud droplets. The resulting change in cloud albedo (Twomey, 1974) implies a radiative forcing due to aerosol-cloud interactions.

Cloud adjustments to aerosol-cloud interactions are manifested as changes in horizontal (cloud fraction) and vertical extent (manifest as liquid water path, LWP) of cloudiness, with consequent impact on the Earth's radiation budget and, thus, climate. With increased aerosol loading and thus increased drop concentrations, at constant LWP, droplets are smaller. One hypothesis for a subsequent adjustment is that the precipitation rates are reduced, implying that the cloud lifetime increases (Albrecht, 1989), and consequently, LWP and cloud fraction increase. But, at the same time, other adjustment processes, such as responses of the cloud mixing and evaporation (Ackerman et al., 2004) occur that partly act in the opposite direction (Stevens and Feingold, 2009; Mülmenstädt and Feingold, 2018; Gryspeerdt et al., 2019). Recently, Toll et al. (2019) have found strong observational evidence that aerosols cause a weak average decrease in the amount of water in liquid-phase clouds compared with unpolluted clouds, since the aerosol-induced cloud-water increases and decreases partially cancel each other out. The different adjustments take place at the same time and become more complex with ice and mixed-phase processes. Because different effects can compensate each other (Stevens and Feingold, 2009; Mülmenstädt and Feingold, 2018), it is difficult to observe isolated cloud effects.

It is nevertheless key to observe how clouds behave as aerosols are perturbed (e.g., Quaas, 2015). Given the large variability of clouds, however, and the plethora of processes involved, it seems best to combine detailed observations with high-resolved modelling (e.g., Mülmenstädt and Feingold, 2018).

Cloud-resolving simulations are a very useful tool to investigate aerosol-cloud interactions, and much of the progress in process-level understanding is from conducting sensitivity studies with such models, and analysing model output in detail (e.g., Ackerman et al., 2004; Khain et al., 2005; Xue and Feingold, 2006; Sandu et al., 2008; Small et al., 2009; Feingold et al., 2010; Fan et al., 2013; Seifert et al., 2015; Gordon et al., 2018). New capabilities now emerge with the possibility to perform cloud-resolving simulations over large domains, with realistic boundary conditions at the land surface and driven by the large scale flow in numerical-weather prediction mode (e.g., Heinze et al., 2017; Miltenberger et al., 2018).

Following this idea, in the present study, mechanisms of aerosol-cloud interactions are analysed and evaluated, using observations, and making use of a set ICOsahedral Non-hydrostatic Large Eddy Model (ICON-LEM) simulations over Germany. The key idea is to assess to what extent the model-simulated aerosol-cloud interaction effects might be detected and attributed in comparison to various observational datasets. This is done in terms of quantification of the impact on cloud macro- and microphysics, precipitation, and radiation, as a response of the modification of CCN concentrations, the small aerosol particles necessary for water vapor to condensate and form cloud droplets. Two CCN input configurations were used in the high-resolved ICON-LEM simulations (156 m horizontal resolution), which were performed over Germany on a selected date (2 May 2013). For the control simulation, CCN concentrations as estimated for 2 May 2013 from a detailed aerosol model were used. For the perturbation, CCN concentrations valid for 1985 were selected. At this time, pollution in Europe was at its peak, about four times higher than presently (Smith et al., 2011, see later for more details). Note that in the present study only modifications to the CCN concentration have been taken into account, and not to ice nucleating particles (INP). As well, the changes made to the CCN are not affecting the scattering or the absorbing aerosol properties (aerosol-radiation interactions, in previous literature referred to as direct and semi-direct aerosol effects), but only the aerosol-cloud interactions (in previous literature also called first aerosol indirect effect, cloud albedo effect, or Twomey effect, and the cloud adjustments, such as the cloud lifetime effect or other rapid responses). The methodology section provides a description of the model (section 2.1) and the CCN perturbation (section 2.2) as well as the observations (section 2.3) used in this study.

The results have been divided into several sections. In order to evaluate the CCN concentrations, in section 3.1 aerosol optical depth (AOD) and CCN are compared with remote sensing observations from satellite and from the ground. After that, the mean vertical profiles of mass and number concentration from the model simulations are revised in section 3.2 and the differences found between perturbed and control simulations are introduced. In the following sections, these differences are explored in comparison to several kinds of observations: liquid-cloud microphysics are compared to satellite and ground-based remote sensing observations in sections 3.3 and 3.4; and the aerosol effects on precipitation, as well as cloud boundaries and cloud cover, and their perturbation, are analyzed and discussed in sections 3.5 and 3.6, respectively. Details on the sensitivity to the CCN perturbation in different cloud regimes are discussed in section 3.7. The simulated effects on the radiation budget are quantified in section 3.8. The results are summarized and the conclusions are discussed in section 4.

2 Methodology

2.1 Model and simulation setup

The simulations are run with the ICON model, the atmospheric model jointly developed by the German Meteorological Service (Deutscher Wetterdienst, DWD) and the Max Planck Institute for Meteorology for numerical weather prediction (Zängl et al., 2015) and climate simulations (Giorgetta et al., 2018), respectively. Within the High Definition Clouds and Precipitation for Climate Prediction (HD(CP)^2) project, another configuration of the model was developed to perform large-eddy simulations (Dipankar et al., 2015), whose physics package is used in the present investigation. A detailed model description is provided by Heinze et al. (2017); their study also provides a thorough model evaluation with various observations including those in the the HD(CP)^2 Observational Prototype Experiment (HOPE) (Macke et al., 2017). The ICON-LEM includes parameterizations for land surface processes, sub-grid turbulence (3D Smagorinsky turbulence), cloud microphysical processes, and radiative transfer. A key feature is an advanced two-moment liquid- and ice-phase bulk microphysics scheme (Seifert and Beheng, 2006): CCN are prescribed in the study (see next Section for more details) as temporally and spatially varying fields read in from offline calculations. The model version used here, however, has been updated to allow for the consumption scavenging of CCN, at droplet activation, a CCN is scavenged. For replenishment, CCN concentrations outside clouds are then relaxed back to the prescribed distribution with a relaxation time scale of 10 minutes. The simulation is performed in a limited-area set-up of Germany. It is run in three one-way nested resolutions and the triangular discretisations at each of the three domains in which ICON-LEM is run correspond to 625 m, 312 m, and 156 m horizontal resolution. This analysis focuses on the results at the highest resolution (156 m). In the vertical, 150 levels are used, with grid stretching towards the model top at 21 km. The minimal layer thickness is 20 m near the surface and the lowest 1000 m encompass 20 layers.

Due to the large computational cost (1 simulated hour consumes 13 real time hours on 300 nodes, equivalent to about 100.000 EUR per simulated day full economic cost of computing time), only one single day was chosen for the model study. Based on the evaluation results from Heinze et al. (2017), the date of 2 May 2013 has been selected. The key reason is that a wide range of cloud- and precipitation regimes was present at this day (Fig. 1 illustrates the cloud conditions, based on satellite data). High pressure prevailed over Germany on that date, with low to mid-level convective clouds, that were locally produced. Specifically, at 10 h UTC shallow convection started and finally lead to a peak of deeper precipitation-forming convection at 17 h UTC. In the southern and eastern part of the domain, stronger convection occurred accompanied with thick cloud layers in the afternoon along with a frontal passage. However, the convective clouds were shifted to higher latitudes in the model simulations. More details about the weather conditions are given in Heinze et al. (2017).

Four simulations were conducted (Table 1). The control (C2R) and perturbed (P2R) simulations were initialized at 0 h UTC on 2 May 2013 from simulations with the COSMO-DE model at 2.8 km resolution, and were run in three nests from coarse (624 m horizontally) to intermediate (312 m) to fine (156 m) resolution as described in Heinze et al. (2017). For analysis, the daytime period between 8 h and 20 h UTC is chosen. In particular, the model output used in the study was on one hand, the so-called meteograms which were temporarily high-resolved (9 s) and available at 36 station locations; and on the other, 2D (at 1 min resolution) and 3D (available every 30 min and at 150 vertical levels) whole domain data fields. The control simulation

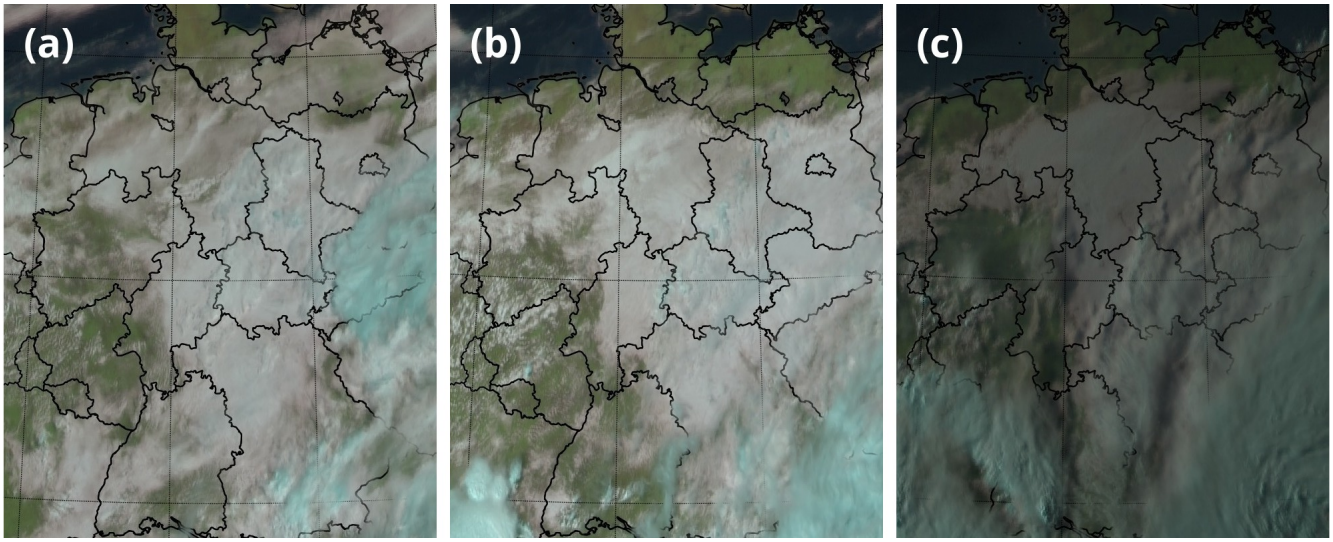


Figure 1. Temporal sequence of satellite images at (a) 9 UTC, (b) 13 UTC and (c) 17 UTC on 2 May 2013 as natural color composite of the 0.6-, 0.8-, and 1.6-μm channels and the high-resolution visible channel from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument on board the geostationary Meteosat satellite. Liquid clouds appear in shades of white and an increase in ice content shifts the color of the clouds towards cyan.

(C2R) used prescribed CCN distributions from 2013. The sensitivity simulation (P2R) that is analysed here in comparison to the control simulation (C2R) used prescribed CCN distributions from 1985 (see next section).

The extra pair of simulations (C1R and P1R, respectively) was conducted in a setup in which the ICON standard description of cloud optical properties in the radiation scheme was applied. In this approach, cloud optical thickness as input to the radiation is computed on the basis of solely the cloud liquid water mixing ratio, without taking into account variable cloud droplet concentrations. This second pair of simulations with CCN for 2013 and 1985 thus does not account for the radiative forcing due to aerosol-cloud interactions (in previous literature also called first aerosol indirect effect, cloud albedo effect, or Twomey effect), but only for cloud adjustments (such as the cloud lifetime effect or other rapid responses). In turn, the first pair of simulations (C2R and P2R, respectively) allows to calculate the full effective radiative forcing due to aerosol-cloud interactions, including the Twomey effect. The difference between the two pairs of simulations thus allows to isolate the latter.

2.2 CCN for 2013 and 1985

The anthropogenic aerosol emissions in the Northern Hemisphere exhibited a strong increase since the industrialization starting in the early 18th century and accelerating from the 19th century onwards. Over Europe and North America, they reached a maximum in the 1980s, and declined since then (Smith et al., 2011; Cherian et al., 2014). Over Europe, the peak was at about 40 Tg sulfur per year. After that, they decreased to early 1900s levels because of the introduction of air quality policies, and due to economic restructuring in Eastern Europe. Over the course of about 30 years from the mid-1980s (represented here

Table 1. Summary of simulations for 2 May 2013. Control ("C") runs are with (low) 2013 aerosol concentrations, Perturbed ("P") runs, with (high) 1985 aerosol. The first set takes into account two moments of the cloud particles size distributions in radiation ("2R"), the specific mass and number, whereas the second one, only one moment ("1R"), which is the specific mass.

Simulation acronym	Microphysics scheme	Clouds in radiation	Aerosol conditions
C2R	2-moment	Mass and number	2013
P2R	2-moment	Mass and number	1985
C1R	2-moment	Only mass	2013
P1R	2-moment	Only mass	1985

by the year 1985) and the mid-2010s (represented here by the year 2013), a very substantial change in aerosol concentrations occurred (Smith et al., 2011).

A prerequisite to realistically simulate the aerosol-cloud interactions is a realistic representation of the aerosol concentrations and their capacity to serve as CCN. Compared to the earlier model version (Heinze et al., 2017; Hande et al., 2016), new time-varying 4D distributions of CCN concentrations were generated from the emissions valid for 2013 and for the peak-aerosol conditions around 1985. The simulation imposing these latter CCN concentrations is hereafter called "perturbed", in comparison to the "control" run which employs CCN concentrations for 2013. The difference between peak-aerosol in 1985 and 2013, rather than, e.g. a comparison of 2013 conditions to pre-industrial aerosol, has been chosen for two reasons: (i) the perturbation is much larger, since the emissions in 2013 were much closer to pre-industrial than to 1985 levels, and (ii) some observations for 1985 are available to assess the aerosol fields.

The CCN distributions were created with the regional coupled model system COSMO-MUSCAT, which consists of two online-coupled codes: the chemistry transport model Multi-Scale Chemistry Aerosol Transport (MUSCAT) (Wolke et al., 2004, 2012) and the Consortium for Small-scale Modelling (COSMO) model, which is the the operational weather forecast model of the German Meteorological Service (Deutscher Wetterdienst, DWD) and other national meteorological services in Europe (Baldauf et al., 2011; Schättler et al., 2014). The COSMO model, which is a non-hydrostatic meteorological model and solves the governing equations on the basis of a terrain-following grid (Baldauf et al., 2011), provides MUSCAT with all required meteorological fields. Based on these fields, MUSCAT calculates the transport and transformation processes which include advection, turbulent diffusion, and physico-chemical conversion of particles and trace gases in the air, as well as sink processes (sedimentation, dry and wet deposition) (Knoth and Wolke, 1998; Wolke et al., 2012). The emissions of anthropogenic primary particles and precursors of secondary aerosols is prescribed using emission fields from the European Monitoring and Evaluation Programme (EMEP, 2009; http://www.emep.int/). The emission fluxes of natural primary aerosols (e.g. desert dust, primary marine particles), are computed online within the model depending on meteorological fields (surface wind speed, precipitation) (Heinold et al., 2011b).

For the present study, two periods of the year 2013 were simulated with COSMO-MUSCAT, coinciding with the measurement campaigns of the HOPE experiment (Macke et al., 2017). It includes measurements around JOYCE (Löhnert et al., 2015) for the period 26 March – 19 June, and at Melpitz from 1 – 30 September. In order to estimate the aerosol concentrations in 1985, the concentrations of 2013 were scaled using scaling factors for black carbon (BC), sulfate (SU), ammonium sulfate (AS), and ammonium nitrate (AN) (Genz et al., 2019). The scaling factors were derived based on the emission ratios between 2013 and 1985 for the different species. Genz et al. (2019) estimate that the concentration for BC, SU, and AS was larger in 1985 by factors of 2.0, 5.3, and 3.9, respectively, compared to the 2013 concentrations.

Due to the high SO_2 emissions in the 1980s, Genz et al. (2019) assumed that there was more than enough sulfate available in order to consume all ammonia to form ammonium sulfate. Therefore, ammonium nitrate is assumed to play a negligible role in the 1980s and consequently its concentration for 1985 was set to zero. Natural aerosol species (sea salt, mineral dust, and organic carbon) are assumed unchanged between the two simulations.

In section 3.1 the AOD derived from the model simulations is compared to satellite AOD retrievals. Since the modeled AOD is calculated at 0.5 μm, it was scaled to the 0.63 μm wavelength (taking into account Ångström coefficients for the different species) at which AVHRR retrieves AOD. The calculation of AOD from the modeled results was done following Meier et al. (2012). Moreover, for both HOPE campaigns, the modeled CCN number concentrations are compared against measurements with a PollyXT lidar systems (Engelmann et al., 2016). In the offline calculation based on COSMO-MUSCAT, the CCN number concentration of the multi-modal size distribution at a fixed supersaturation is calculated according to Abdul-Razzak and Ghan (2000). The calculation of the CCN number concentration (number of activated aerosol particles) follows Hande et al. (2016) and Genz et al. (2019), using the parameterization by Abdul-Razzak and Ghan (2000) for multi-modal size distributions. Details on the used hygroscopicity parameters as well as the derivation of number size distributions from the simulated speciated aerosol mass can be found in Genz et al. (2019). For the comparison to available observations, the CCN number concentration field at a fixed supersaturation is calculated. However, in order to provide CCN fields for ICON-LEM, time varying 3D fields of the CCN number concentration at different constant updraft velocities was required. Abdul-Razzak and Ghan (2000) relate the aerosol composition and updraft velocity to the maximum supersaturation during an air parcels ascend, which in the end determines the number of activated aerosol particles. The calculated CCN fields were then used in the ICON-LEM simulations replacing the fixed assumed CCN number concentration/distribution.

2.3 Observations

2.3.1 Satellite-based

AOD as retrieved in the PATMOS-x Advanced Very High Resolution Radiometer (AVHRR) retrievals uses the 0.63 μm channel. It is only retrieved over sea and in clear-sky cases. The product is a daily average of different NOAA satellites with AVHRR aboard. For 1985 the daily average contains NOAA 7, 8 and 9; and for 2013 data from NOAA 15, 18 and 19 and METOP-B satellite are used.

To compare distributions of liquid cloud properties from the ICON-LEM simulations to satellite observations, the off-line diagnostic tool Cloud Feedback Model Intercomparison Project Observation Simulator Package (COSP; Bodas-Salcedo et al., 2011; Swales et al., 2018) was applied to the ICON simulation output. COSP allows for consistency between cloud properties simulated by ICON and retrieved by satellite observations such as the Moderate Resolution Imaging Spectroradiometer (MODIS). For 2 May 2013, collection-6 cloud products MOD06/MYD06 (Platnick et al., 2015, 2017) from four MODIS satellite overpasses within the domain (MODIS-Aqua at 11.45 h UTC and 13.20 h UTC, MODIS-Terra at 9.55 h UTC and 11.35 h UTC) were analyzed. The ICON-COSP simulations were temporally and spatially matched to the satellite observations as well as regridded to the MODIS data resolution of 1 km. No sub-column variability is used in COSP, consistently with the lack of sub-pixel variability in MODIS retrievals. Only cloudy satellite pixels with assigned liquid cloud phase as well as good quality and solar zenith angles below 50° were considered, to exclude uncertain/problematic cloudy retrievals.

For liquid clouds, cloud droplet number concentration (N_d) is derived from cloud effective radius (r_e) and cloud optical thickness (τ_c) as in Quaas et al. (2006); where α = 1.37·10^{-5}m$^{-0.5}$:

$$N_\mathrm{d} = \alpha\, \tau_\mathrm{c}^{0.5}\, r_\mathrm{e}^{-2.5}, \tag{1}$$

an approach that assumes an adiabatic growth of clouds (Grosvenor et al., 2018).

2.3.2 Ground-based

A comprehensive set of active and passive remote sensing instruments is part of the Leipzig Aerosol and Cloud Remote Observations System (LACROS) (Bühl et al., 2013). Specifically, its multiwavelength-Raman-polarization lidar provides backscatter and extinction profiles almost continuously (Baars et al., 2016). In this study it serves to retrieve aerosol profiling variables, and CCN number concentration, r_e, liquid water content (q_l), and N_d, which are compared with the model output. The CCN concentration profiles of the lidar measurements were calculated using the method described in Mamouri and Ansmann for measurements during the HOPE campaign (Section 3.1). Profiles of liquid cloud microphysical properties were derived from ground-based remote sensing using the recently established dual-field-of-view (DFOV) lidar techniques (Grosvenor et al., 2018). Such observations are available at Leipzig, Germany (51.3°N, 12.4°E), since 2013 and provide information about aerosol-cloud-interaction processes (Schmidt et al., 2014) and allow a climatological assessment (Section 3.4). Originally, the observations were based on the DFOV Raman lidar technique (Schmidt et al., 2013), which can only be applied to nighttime lidar observations in order to reduce effects of the solar background on the measurements of Raman-scattering lidar returns from nitrogen molecules. Progress that was made recently in the accuracy of polarization measurements with lidar (Jimenez et al., 2019a) allows to apply an alternative technique for profiling of liquid-cloud microphysical properties even during daytime (Jimenez et al., 2017, 2019b). In this novel DFOV-polarization approach, the liquid water content, q_l, profile is assumed to increase adiabatically with height from the cloud bottom, while the N_d remains constant. An inversion scheme exploits a non-ambiguous relation between r_e and the extinction coefficient, both dependent on N_d and q_l, with the single-FOV depolarization ratio and the relative depolarization, quantities that a DFOV polarization lidar can measure (Jimenez et al., 2017, 2019b).

The cloud radar (8.6 mm wavelength) based in the Meteorological Observatory Lindenberg – Richard Aßmann-Observatorium (MOL-RAO), is well suited for the study of thin, low-reflectivity clouds such as non-drizzling and drizzling stratocumulus clouds, due to its high sensitivity. The cloud radar transmits linear polarized radiation at 35.5 GHz and simultaneously receives the co- and cross-polarized backscattered signal. Observations in zenith mode are used, with an integration time of 1 s and a 256 point Fourier transform for generating the Doppler spectrum. Forward simulations of radar Doppler spectra and their corresponding moments have been performed from the ICON-LEM simulation output using the radar forward simulator included in the Passive and Active Microwave radiative TRAnsfer (PAMTRA) framework (Maahn, 2015). The moments of the synthetic Doppler spectra are derived in the same way as for the observations (Acquistapace et al., 2017), and then the drizzle detection criterion described in Acquistapace et al. (2019) is applied to them. Only liquid clouds are selected following Cloudnet criteria (Illingworth et al., 2007). In section 3.5 the ICON-LEM forward simulated reflectivities have been compared to the radar observations for four different categories of drizzle-cloud droplets-rain drops.

The ceilometer network of DWD provides backscatter measurements and cloud base height (CBH) retrievals from Jenoptik ceilometers (model CHM15k) at very high temporal resolution (15 s) up to 15 km above ground at 51 stations across Germany (Wiegner and Geiß, 2012; Wiegner et al., 2014; Martucci et al., 2010, http://www.dwd.de/ceilomap) within the ICON-LEM domain. The CBH 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^{-1}. CBH, cloud occurrence as an estimator of cloud cover (CC) (see methodology in Costa-Surós et al., 2013), and cloud persistences (CP) derived from the ceilometer network are compared with those simulated at the same locations with the ICON-LEM. The methodology to derive the CP is the following: the time resolution from 51 CBH timeseries from ceilometer observations (15 s resolution) was changed to 1 min, which matches the output frequency from ICON-LEM diagnostics. All three 12-hour (8-20 h) timeseries (ceilometer observations, control and perturbed simulations) were splitted into 30 min intervals. For each of these 30 min intervals the CBH between 0 – 3000 m was checked, and if there was a CBH it was flagged as cloudy and the neighbour cloudy pixels were assigned to one cloud, so a cloud persistence for each single cloud of the interval was assigned. Then each 30-min interval was classified according to its cloud cover as "clear-sky" (0 – 5 %), "few clouds" (5 – 25 %), "scattered clouds" (25 – 50 %), "broken clouds" (50 – 87 %), or "overcast sky" (87 – 100 %). After that, a normalized histogram was plotted using the cloud persistences from the time intervals classified as "few", "scattered" and "broken" clouds, i.e. cloud cover between 5 – 87 % (see section 3.6).

3 Results

3.1 Evaluation of the perturbed and unperturbed CCN and aerosol distributions

A comparison of the AOD over the North and Baltic seas as retrieved from AVHRR for 1985 and 2013 and AOD simulated by COSMO-MUSCAT is shown in Fig. 2 and summarized in Table 2. Note that since AVHRR retrieves AOD only in cloudless cases over sea, AOD is only available over the North sea and Baltic sea region for a small fraction of the time (between 10-12 % in 1985, and 35-38 % in 2013, Table 2). Approximately double AOD levels on average are observed in March–June 1985

in comparison to March–June 2013. The average AOD and also its change between 1985 and 2013 are well captured by the model simulation considering the observation and modelling uncertainty, and natural variability. The geographical distribution over sea (Fig. 2) also shows consistency between the simulation and the satellite retrievals with strong differences between the polluted spring 1985 and the much cleaner spring 2013. Since not very many data-points go into the average from the satellite, the distribution is more noisy compared to the simulation output that is available anytime. Despite this, it is evident that both, observations and model consistently show the expected spatial gradient with larger AOD near the coast and less aerosol over open sea.

Table 2. Median aerosol optical depth from AVHRR and COSMO-MUSCAT from 26 March to 19 June in 1985 and 2013. 25^{th} and 75^{th} percentile ranges of regional variability of the temporal mean AOD are provided in brackets. In parentheses mean number of days (out of the possible 86 days) per grid box with valid AVHRR retrieval.

	1985		2013	
	North Sea (9) Median [25^{th} – 75^{th}]	Baltic Sea (11) Median [25^{th} – 75^{th}]	North Sea (33) Median [25^{th} – 75^{th}]	Baltic Sea (30) Median [25^{th} – 75^{th}]
AVHRR	0.25 [0.22 – 0.30]	0.30 [0.27 – 0.33]	0.14 [0.13 – 0.17]	0.14 [0.13 – 0.15]
COSMO-MUSCAT	0.22 [0.17 – 0.25]	0.30 [0.27 – 0.32]	0.09 [0.08 – 0.10]	0.11 [0.11 – 0.12]

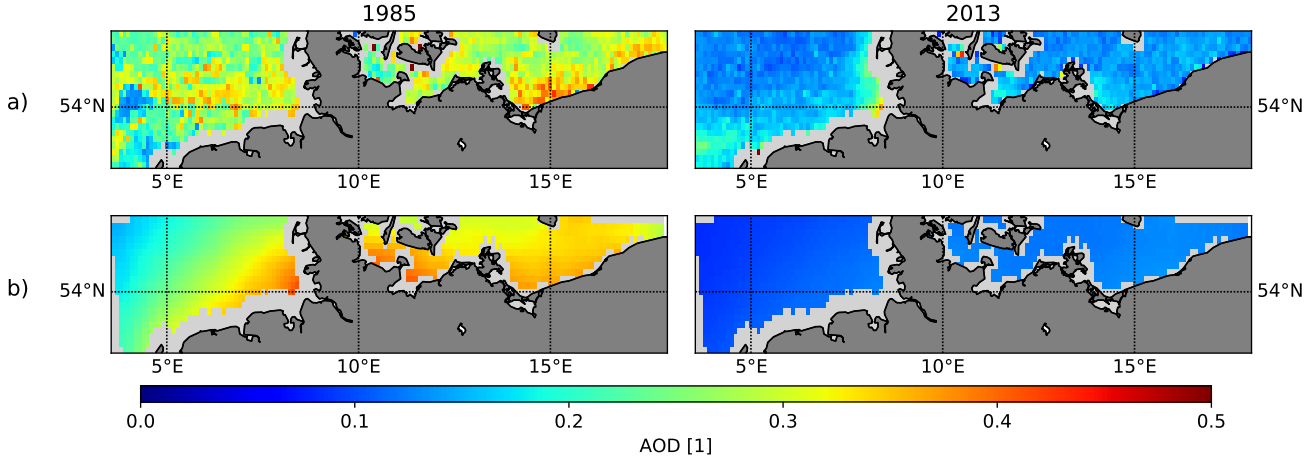


Figure 2. Mean AOD retrieved from a) AVHRR (top) and simulated by b) COSMO-MUSCAT (bottom) for 1985 (left column) and 2013 (right). The comparison period covers 26 March to 19 June for each of the years.

Profiles of CCN at 0.2 % supersaturation from both COSMO-MUSCAT and from PollyXT lidar retrievals are shown in Fig. 3. The location and time period assessed is the HOPE campaign (Krauthausen, Germany) from 26 March to 19 June 2013. Model and retrieval agree to within the uncertainty in the boundary layer. The mean profiles for 2013 are overestimated by the model by 10–30 % in the boundary layer (from the surface to 1800 m in spring 2013, and 1200 m in fall 2013). Above the boundary layer, the overestimation increases to more than a factor of 2 since the model tends to overestimate the vertical mixing between boundary layer and free troposphere (Heinold et al., 2011a). The estimated CCN concentrations for 1985 are inconsistent with the observations, they exceed the concentration in 2013 by a factor of 2–5 in the boundary layer and up to an order of magnitude in the free troposphere. Similar results are found for the HOPE-Melpitz campaign during September 2013 (not shown).

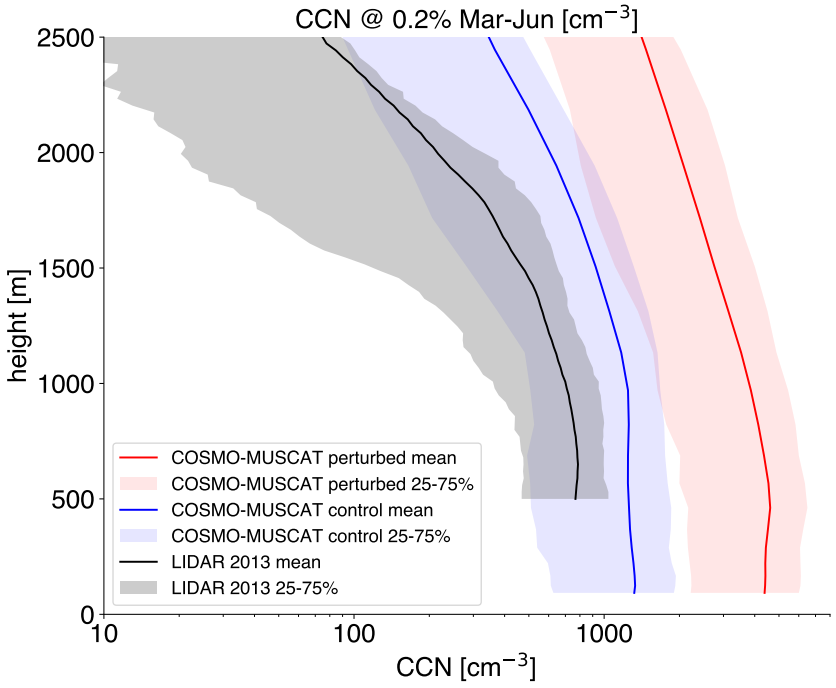


Figure 3. Comparison of mean vertical profile of CCN number concentration at 0.2% supersaturation simulated by COSMO-MUSCAT (blue: 2013, red: 1985 aerosol conditions) and retrieved from the PollyXT Lidar measurements using the algorithm of Mamouri and Ansmann (black), from 26 March to 19 June 2013 during HOPE campaign in Krauthausen, Germany. The shaded area depicts the 25–75 percentiles of the set of single profiles included in the average. A factor 2 of uncertainty is considered for the lidar retrievals, as outlined in Mamouri and Ansmann and Ansmann et al. (2019).

In conclusion, the imposed aerosol concentrations for 1985 (perturbed simulations) and 2013 (control simulations) match well the distribution and mean values from the satellite retrievals over clear-sky ocean, and it is evident that only the 2013 aerosol yields CCN profiles that are consistent with lidar retrievals from 2013.

3.2 Mean vertical hydrometeor profiles of number and mass concentrations: control vs. perturbed simulations

Domain-averaged hydrometeor profiles of number and mass concentration from model output on 2 May 2013 from 8 h to 20 h UTC are shown in Fig. 4. The vertical profile of cloud droplet concentrations closely corresponds to the introduced CCN disturbance (vertical integrated relative increase of 147 %, Table 3). In the following sections (3.3 and 3.4) it is explored to which extent this perturbation is also seen by satellites and from ground-based remote sensing in terms of N_{d} retrievals. The total-water mass difference at about 2 km altitude (vertically integrated relative increase by 8.8 %) is mainly due to decreased rain in the perturbed simulation (-12.3 %). In the following sections, the changes in the LWP and liquid water content (q_{l}) are investigated.

A slightly higher homogeneous cloud droplet freezing for the perturbed simulation is triggered by upward transport of cloud droplets. In turn, graupel number and mass concentrations are slightly higher in the cleaner environment in low to mid altitudes (3 – 4 km). On the contrary, at higher altitudes (6 – 10 km) the graupel mass is higher in the perturbed simulation.

Table 3. All-sky domain mean vertically integrated changes between the perturbed and control simulations (P2R – C2R) for number and mass concentrations for water species and for the total water as temporal average from 8 h UTC to 20 h UTC. The absolute changes are given along with the temporal standard deviation of the domain-mean changes. The numbers in brackets are the relative changes.

Variable	Absolute (relative) all-sky vertically integrated domain-mean day-mean differences (P2R – C2R)	
	Number concentration [m^{-2}]	Mass concentration [$\mathrm{kg\ m}^{-2}$]
Total water	$6.9 \cdot 10^{10} \pm 1.3 \cdot 10^{10}$ (147 %)	$3.8 \cdot 10^{-3} \pm 7.7 \cdot 10^{-3}$ (0.9 %)
Cloud water	$6.9 \cdot 10^{10} \pm 1.3 \cdot 10^{10}$ (147 %)	$7.6 \cdot 10^{-3} \pm 4.4 \cdot 10^{-3}$ (8.8 %)
Cloud ice	$9.7 \cdot 10^{5} \pm 1.0 \cdot 10^{7}$ (0.4 %)	$2.8 \cdot 10^{-4} \pm 7.1 \cdot 10^{-4}$ (0.9 %)
Rain	$-3.7 \cdot 10^{6} \pm 1.5 \cdot 10^{6}$ (-25.5 %)	$-5.2 \cdot 10^{-3} \pm 2.5 \cdot 10^{-3}$ (-12.3 %)
Snow	$2.7 \cdot 10^{4} \pm 5.6 \cdot 10^{4}$ (2.0 %)	$1.7 \cdot 10^{-4} \pm 5.5 \cdot 10^{-4}$ (1.2 %)
Graupel	$-379.3 \pm 1.62 \cdot 10^{4}$ (-1.3 %)	$9.1 \cdot 10^{-4} \pm 6.5 \cdot 10^{-3}$ (0.1 %)
Hail	-1.9 ± 32.4 (0.1 %)	$-3.1 \cdot 10^{-3} \pm 3.1 \cdot 10^{-4}$ (0.6 %)

3.3 Liquid-cloud microphysics in comparison to satellite data

Normalized frequency of occurrence distributions of liquid water path (LWP) and cloud droplet number concentration (N_{d}) from the reference (C2R) and perturbed (P2R) ICON-LEM simulations as well as the corresponding satellite retrievals (MODIS)

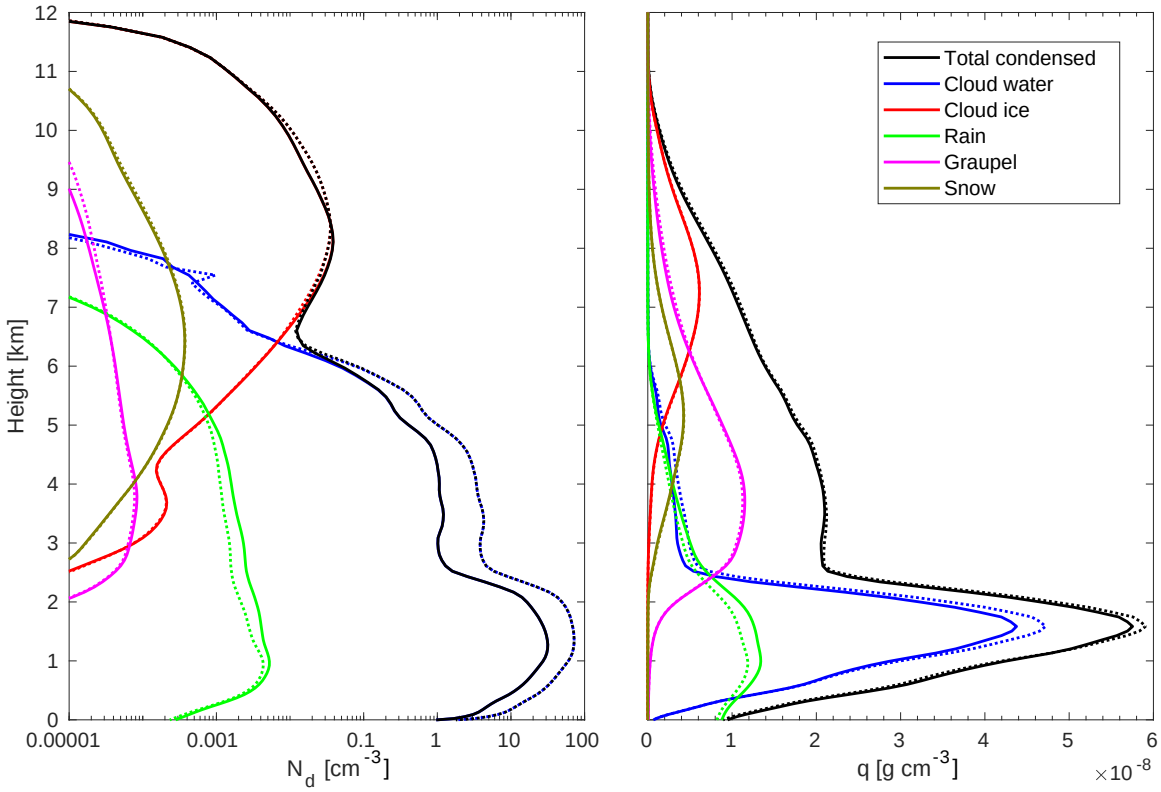


Figure 4. Vertical profile of (all-sky) domain mean number concentration (left) and mass concentration (right) of total water and individual particle species on 2 May 2013 from 8 h to 20 h UTC. The control (C2R) simulation is plotted as solid lines and the perturbed (P2R) simulation as dotted lines. Note the logarithmic x-axis for the number concentration panel. Hail not shown due to very low values. Also note that in the left panel the black solid line is over the blue solid one, and the blue and black dashed lines are also one over the other.

are shown in Fig. 5. The maximum peak of occurrence and 50 % percentile values for N_d compare well between the ICON control simulation (C2R) and MODIS, although the distribution of N_d simulated by ICON-LEM is much broader, resulting in lower and higher 25 % and 75 % percentile values, respectively, compared to MODIS (Table 4). This can be partly explained by a too large range simulated by the model, which in turn can be partly related to a difference in observed and simulated spatial distribution of clouds at the MODIS observation times. However, a part of the difference in the range of the distributions can be very likely attributed to the MODIS instrument characteristics, since optically very thin clouds are not observed or give problematic retrievals (Grosvenor et al., 2018), and for optically very thick clouds the measurements can go into saturation (note that N_d is computed from the cloud optical thickness and cloud effective radius, see Eq. 1). The simulated N_d distribution for the perturbed simulation (P2R) is shifted to significantly higher values. A factor of about 2 between ICON-control and ICON-perturbed is reflected in the percentile values.

For LWP values larger than about 10 g m^{-2}, MODIS and ICON compare well, both showing occurrence peak values between 100 and 200 g m^{-2}. However, ICON (control and perturbed) has higher frequency of low LWP values, resulting in lower 25%, 50%, and 75% percentile values. The model also does not show the bi-modal distribution as MODIS. This bi-modal distribution is due to having two different cloudy scenes in the different overpasses, which happen at different times of the day, i.e. a cloudy scene with optically thinner clouds and thus lower LWP, and then scenes with optically thicker clouds and higher LWP. However, even if the ICON-LEM output is sampled along the MODIS swath, it does not show this distinct behaviour, but a smooth distribution. The difference between LWP from the ICON reference simulation (C2R) and ICON perturbed simulation (P2R) is small compared to the LWP variability, and small in comparison to the model bias with respect to the MODIS retrievals. It is nevertheless a systematic increase in LWP that is simulated, even if it is small as also expected from recent investigations of satellite data (Malavelle et al., 2017; Toll et al., 2017; Gryspeerdt et al., 2019; Toll et al., 2019). An exception is at large LWP (larger than 200 g m^{-2}, see Fig. 5), where the control simulation is much smaller than the perturbed one, and closer to the satellite retrievals. This is firstly consistent with the expectation that an increase in N_d leads to an invigoration of convective clouds and, hence, deeper clouds with higher LWP in the tail of the LWP-distribution, where the convective cloud cores can be found. It, secondly, reflects the fact that the thick, precipitating clouds most strongly respond to precipitation delay in response to the CCN perturbation.

Table 4. Median, 25$^{\mathrm{th}}$ and 75$^{\mathrm{th}}$ percentiles of liquid water path (LWP) and cloud droplet number concentration (N_d) from MODIS and ICON-LEM distributions.

	N_d [cm^{-3}] Median [25$^{\mathrm{th}}$ – 75$^{\mathrm{th}}$]	LWP [g m^{-2}] Median [25$^{\mathrm{th}}$ – 75$^{\mathrm{th}}$]
MODIS	170 [97 – 283]	112 [35 – 203]
ICON-LEM-COSP control	188 [64 – 416]	68 [23 – 158]
ICON-LEM-COSP perturbed	339 [113 – 808]	70 [24 – 165]

In conclusion, the influence of the perturbed aerosol on N_d clearly can be detected and attributed in comparison to satellite retrievals, but the systematic increase in LWP is too small in comparison to natural variability and model bias to be detected and attributed, except for large LWP ($>$ 200 g m^{-2}). The systematic change in LWP, even if small compared to weather variability, implies a substantial contribution to the aerosol effective radiative forcing.

3.4 Liquid-cloud microphysics in comparison to ground-based remote sensing

Profiles of several cloud microphysical variables (N_d, r_e, and q_l) were retrieved from ground-based remote sensing as explained in section 2.3.2. In order to derive comparable profiles from ICON-LEM that are best suited for evaluation against the lidar observations, the temporarily high-resolved (9 s) meteogram output was used. All profiles from 21 of the 36 stations for which meteogram output of ICON-LEM was available (Table A1), were in a first step screened for the occurrence of the presence

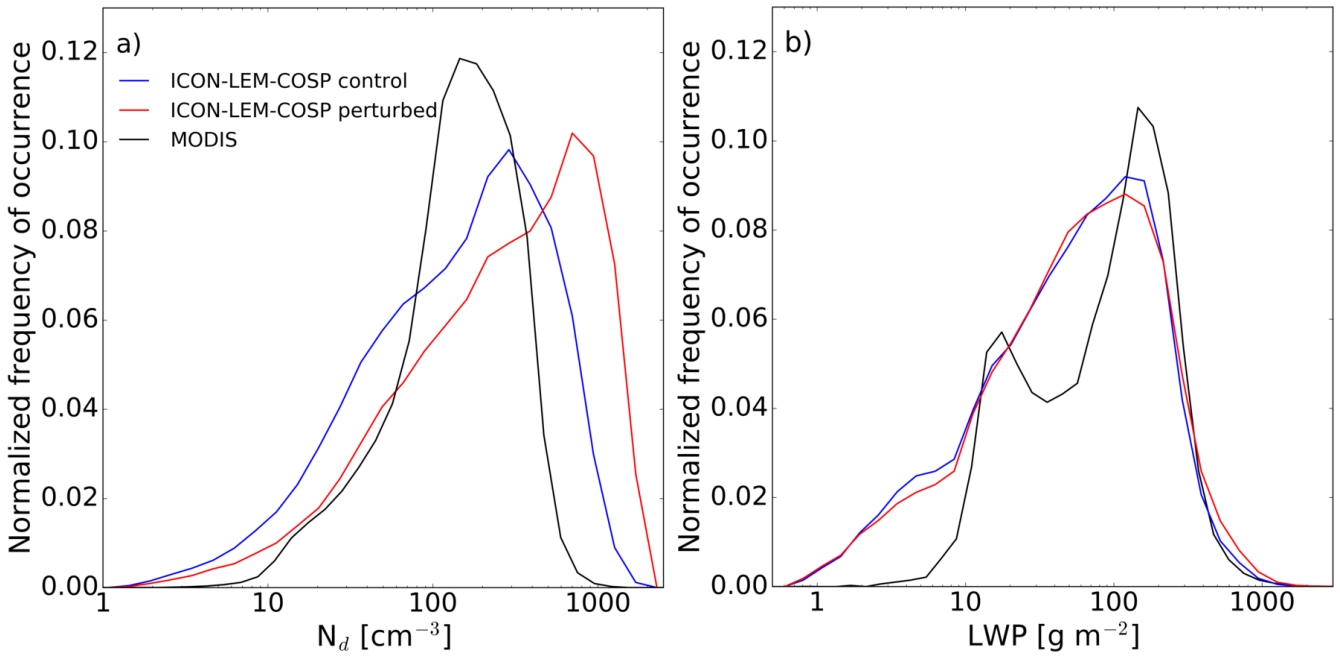


Figure 5. Normalized frequency of occurrence distributions computed from ICON-LEM-COSP data from the 2 May 2013 control (C2R, blue) and perturbed (P2R, red) simulations, matched with satellite observations from MODIS (black) obtained at the four satellite overpass times on the same day for a) N_d and b) LWP.

of hydrometeors. The remaining 15 stations were not considered in the analysis because they were too close (approx. within 20 km) to an already considered station, which would lead to an unwanted weighting of the statistics towards a certain region of the ICON-LEM domain. This was specifically the case for the region of the HOPE campaign (Macke et al., 2017), for which output of 13 stations is available. Vertically continuous sequences of hydrometeors were classified as cloud layers to which cloud base height and cloud top height were assigned. Each identified cloud layer was in a subsequent processing step filtered in such a way, that (i) precipitation was absent, (ii) ice was absent, (iii) LWP $> 150\,\mathrm{g\,m^{-2}}$, (iv) cloud depth $< 500\,\mathrm{m}$, and (v) $1000\,\mathrm{m} <$ cloud base height $< 4000\,\mathrm{m}$. Such constraints are similar to the properties of cloud layers which were observed with the DFOV-polarization lidar at Leipzig. The Cloudnet processing suite (Illingworth et al., 2007) operated based on cloud radar, lidar and microwave radiometer observations at Leipzig, was used to identify these conditions for the periods when valid DFOV-polarization observations were made.

For the evaluation of the ICON-LEM simulations, approx. 40 h of DFOV polarization lidar observations with 3 min temporal resolution (i.e., 800 profiles) were averaged to yield mean vertical profiles of N_d, q_l, and r_e. The cloud periods considered were distributed over 27 days in the spring-summer-autumn seasons of 2017, and over heights between 1 and 4 km above ground level. The resulting profile from the observations and the outputs from the ICON-LEM model for 1985 and 2013 aerosol conditions of 2 May 2013 are presented in Fig. 6. The results corroborate the satellite-based conclusions: for N_d, the simulation

with the 1985 CCN is inconsistent with the lidar observations, and the one with the 2013 CCN matches the retrievals to within the uncertainty. A similar conclusion is drawn for the effective radius. For q_l, in turn – as found before during the comparison to satellite retrieved LWP – the model shows only little aerosol impact, and both realizations compare almost equally to the observations.

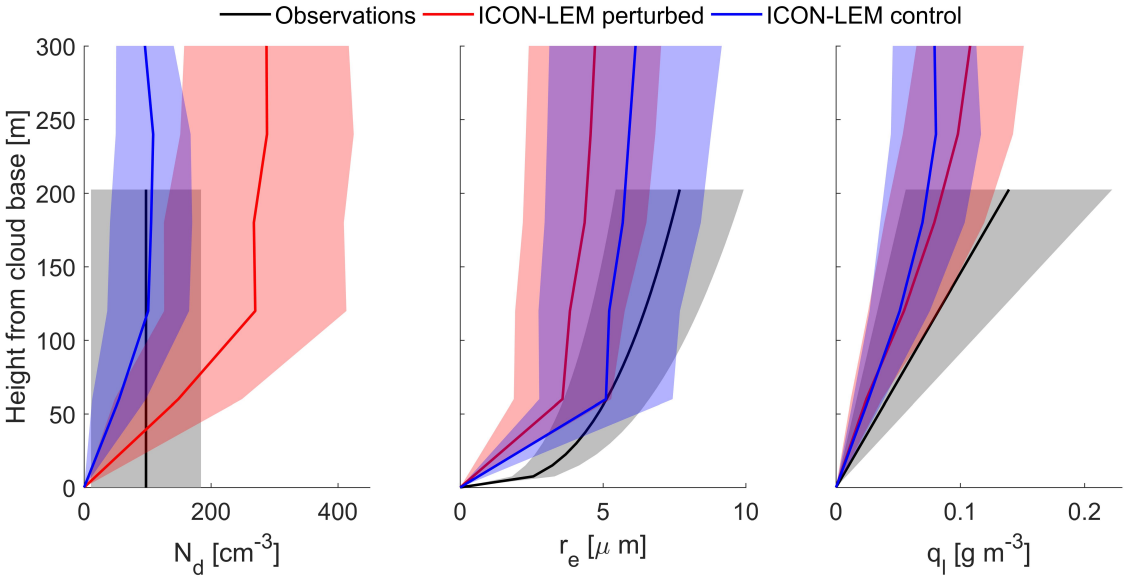


Figure 6. Averaged cloud droplet number concentration (N_d), cloud droplet effective radius (r_e), and liquid water content (q_l) profiles retrieved from Lidar (in black, and the temporal variablility as shading in grey) and in ICON-LEM control (in blue) and perturbed (in red) simulations, as a function of height above cloud base. Note that measurements can only provide profiles up to 200 m into the cloud due to the strong extinction of the Lidar signal.

3.5 Effects on precipitation

As discussed earlier, the model simulates large reductions in rain mass and number concentrations, and increases in cloud liquid mass and number concentrations in the perturbed vs. control simulation (Fig. 4). To enquire into the role of the autoconversion process and find out to which extent these changes might be observable, forward simulations of control and perturbed meteogram profiles at the site of the MOL-RAO have been performed using the PAMTRA tool (Maahn et al., 2015). After that, the drizzle detection criterion described in Acquistapace et al. (2019) has been applied to the forward simulated Doppler radar moments that were generated with PAMTRA. This approach is similar to the one presented in Rémillard et al. (2017). An ice cloud mask has been applied to the model output in order to filter out the ice pixels and therefore applying the drizzle detection criterion only to the pixels corresponding to liquid in the cloud mask, in both observations and model output.

The distributions for the different classes of drizzle development are shown in Fig. 7 and summarized in Table 5. The post-processed model results shown in Fig. 7 are consistent with the raw results summarized in Fig. 4: the perturbed simulation

shows reflectivities that are shifted towards smaller values for all drizzle/precipitation classes, compared to the control simulation. Vice versa, the control simulation produces larger reflectivities, hence larger drops at every stage of drizzle formation, and produces larger raindrops. The mean simulated values of the reflectivities, from both control and perturbed, fall into the range of the observations of MOL-RAO radar, except for the precipitation where both runs overestimate the reflectivities. This overestimation is also visible from the mean Doppler velocities (proxy of fall speed of hydrometeors) and spectrum width distributions (not shown). Mean Doppler velocities are too high (due to too large drops compared to reality) and this produces also broader spectra, giving larger spectrum widths compared to what observed. A closer look into the radar signal suggests that the small reflectivity values for precipitation observations are due to insects detected by the radar, despite a clutter removal filter was applied to the radar spectra from 0 to 1600 m AGL as pre-processing.

Table 5. Median as well as 25th and 75th percentiles of the cloud radar reflectivity (Ze) distributions on 2 May 2013 for the four classes described in Fig. 7.

	Drizzle onset Mean [25th to 75th]	Drizzle growth Mean [25th to 75th]	Drizzle mature Mean [25th to 75th]	Precipitation Mean [25th to 75th]
MOL-RAO cloud radar	–18.4 [–22.1 to –15.4]	–24.2 [–30.3 to –17.0]	–14.2 [–23.1 to –7.5]	–38.2 [–49.7 to –28.2]
ICON-LEM control (C2R)	–16.2 [–20.5 to –11.7]	–15.6 [–22.3 to –5.2]	–4.9 [–7.0 to –2.3]	–8.1 [–13.5 to –0.9]
ICON-LEM perturbed (P2R)	–17.8 [–21.5 to –13.4]	–21.5 [–27.4 to –16.9]	–9.9 [–15.1 to –5.5]	–14.4 [–21.6 to –8.2]

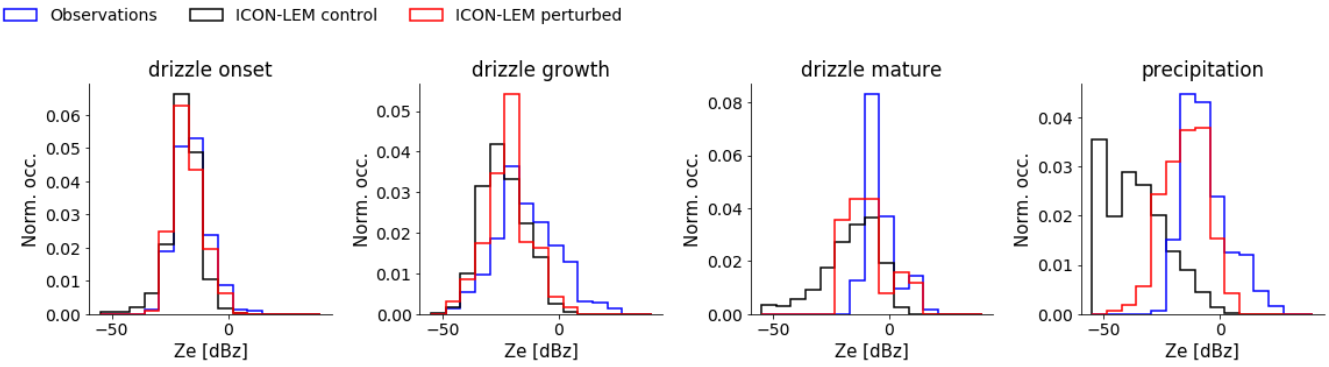


Figure 7. Normalized distributions for the different classes of drizzle development in Meteorological Observatory Lindenberg – Richard Aßmann-Observatorium (MOL-RAO) on 2 May 2013. a) drizzle onset, corresponding to the small non-precipitating drizzle drops, larger than cloud droplets but not big enough to fall yet – this is only a signature in the skewness of the Doppler spectrum, not in reflectivity; b) drizzle growth, which contains the drops big enough to modify the spectra shape; c) drizzle mature, which is the drizzle precipitating inside the cloud; and d) precipitation, which is the class of precipitation below cloud base. The radar observations are processed as in Acquistapace et al. (2019), and ICON-LEM meteogram output at the MOL-RAO site is processed using the PAMTRA forward operator and processed analogously. Black: observations; blue: control simulation (C2R); red: perturbed simulation (P2R).

In conclusion, despite the effort to make model and data comparable using the forward operator, and despite the rather strong signal in the model, no detection and attribution of an aerosol signal could be achieved. In the future more comparisons are needed as difference at one grid point only could arise from a different sampling of cloud life cycle.

3.6 Cloud macrophysics: cloud boundaries, cloud cover and cloud persistence

ICON-LEM cloud base height (CBH), as well as the calculated cloud cover (CC) and cloud persistence (CP) based on the CBH measurements, are compared with high-resolution ceilometer measurements (15 s temporal resolution) from the DWD ceilometer network (please see Section 2.3.2 for details). In Table 6, mean CBH and CC have been calculated over the 51 stations for 2 May 2013. On average, the model produces less clouds than observed and the cloud base heights are too low in comparison to the ones measured by the ceilometer network. The problem is not due to issues with the initial- or boundary conditions, as the discrepancy to the reference observations is larger in the outer nests (312 and 624 m resolution, respectively; result not shown here). However, the cloud variability is very large, and the model output still is consistent with the data to within the uncertainty range. For both variables, the differences between control (C2R) and perturbed (P2R) simulations are so small, in comparison to the observations, that no significant deviations can be detected given the simulation and observation uncertainties. The absolute (relative) differences between perturbed and control simulations are for the mean CBH –4 m (–0.37 %) and for the CC 0.4 % (0.70 %). The same tendency is found in the all-domain simulation means (Table 7) where the perturbed ICON-LEM shows on average higher cloud tops and bases in comparison to the control simulations. The difference between mean cloud top pressures is –263 Pa (0.35 %) and between mean cloud base pressures, –140 Pa (0.17 %). The perturbed simulation also shows higher total cloud cover by 0.16% (relative difference: 0.20 %) in the domain average (Table 7). Cloud fraction can also be assessed from MODIS satellite data and compared to the COSP-processed ICON-LEM output (as in Section 3.3 for LWP and N_{d}). The domain-average cloud fraction for the four MODIS overpasses is 0.84, compared to 0.49 and 0.50 for the control and perturbed ICON-LEM simulations, respectively. Despite the very different observational approaches and spatiotemporal sampling, thus, the general conclusion is confirmed: ICON simulates less clouds than observed (a result that also has been noted by Heinze et al., 2017), and shows a positive response of cloud fraction to more aerosol.

Figure 8 shows normalized CBH distributions for the ceilometer network and for both ICON-LEM control (C2R) and perturbed (P2R) simulations, which overestimate very low CBH (< 500 m), and underestimate higher CBH (751 – 1751 m). Possibly the ceilometer network is not able to detect the latter reliably. Further analysis of 30 min periods exhibit that both simulations overestimate "clear sky" cases (0 - 5 %) and underestimate "overcast" skies (87 - 100 %). "Few" (5 - 25 %) and "scattered" (25 - 50 %) skies are also overestimated, and "broken clouds" (50 - 87 %) are slightly underestimated (not shown). Cloud persistence analysis for the group of few, scattered, and broken cloud conditions (i.e. cloud cover between 5 % to 87 %) and with bases below 3000 m shows also small differences between control and perturbed distributions (Fig. 8). None of the simulations are able to capture the observed distributions for the short-lived clouds (less than 5 min.), while the longer-lived clouds are overestimated.

The comparison of CBH, CC and CP between model and ceilometer measurements shows systematic differences between either model simulation and the observations. Detection and attribution of differences between control (C2R) and perturbed

Table 6. Cloud base height (CBH), cloud persistence (CP), and cloud cover (CC) cover as measured at 51 ceilometer stations in Germany and from ICON-LEM output (2 May 2013 from 8 h to 20 h UTC), as well as CC retrieved from MODIS satellite and ICON-LEM output postprocessed with the COSP simulator. Uncertainty ranges are provided as 25^{th} and 75^{th} percentile ranges.

	DWD ceilometer network measurements			MODIS
	CBH [m]	CP [min]	$CC_{ceilo.}$ [%]	CC_{MODIS} [%]
	Mean [25^{th} – 75^{th}]	Mean [25^{th} – 75^{th}]	Mean [25^{th} – 75^{th}]	Mean [25^{th} – 75^{th}]
Observations	1454 [774 – 1724]	4.9 [1.0 – 6.0]	80.1 [63.9 – 99.5]	90.4 [88.1 – 92.9]
ICON-LEM control (C2R)	1089 [412 – 1423]	6.8 [2.5 – 9.0]	57.1 [32.2 – 84.4]	52.8 [46.8 – 62.7]
ICON-LEM perturbed (P2R)	1085 [407 – 1428]	6.7 [3.0 – 9.0]	57. 5 [32.4 – 83.8]	53.8 [48.1 – 63.2]

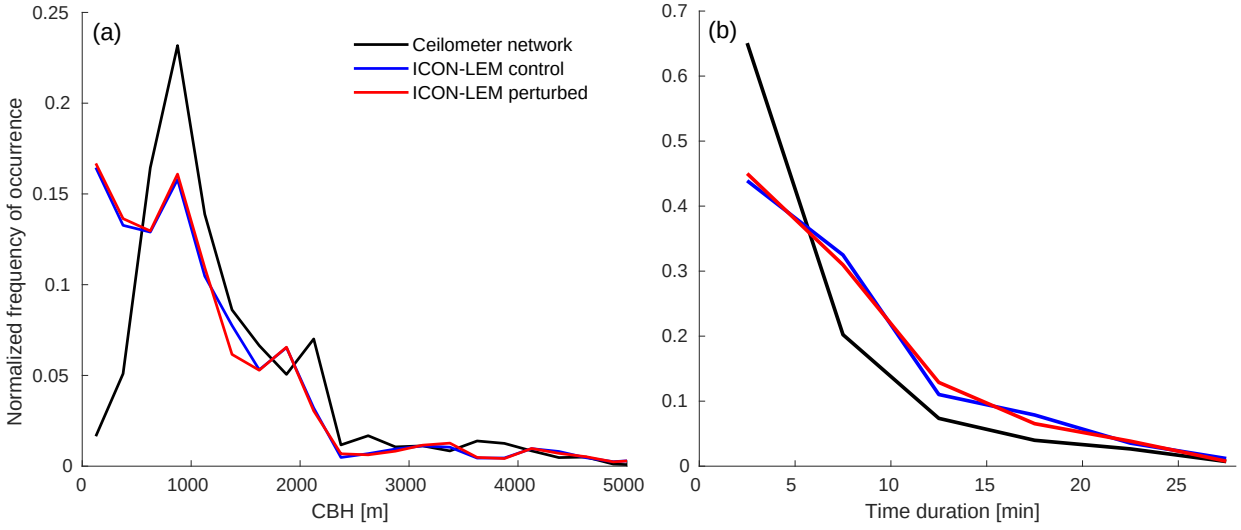


Figure 8. Left: Normalized cloud base height distribution (bin size: 250 m) of 51 ceilometer stations over Germany on 2 May 2013 (8 – 20 h UTC). Right: Normalized cloud lifetime (bin size: 5 min) for low-level clouds (<3 km) of few, scattered and broken cloud conditions (i.e. cloud cover between 5 and 87 %). Data from the DWD ceilometer network in Germany (black), and ICON-LEM C2R (blue) and P2R (red) simulations.

simulations (P2R) is not feasible in this case. This is mostly because the effect of the aerosol perturbation on these quantities is small compared to the model bias.

3.7 Sensitivity to cloud regimes

The International Satellite Cloud Climatology Project (ISCCP) regime classification (Rossow and Schiffer, 1991) utilizes cloud top pressure and τ_c to define different cloud regimes. A consistent diagnostics is applied to the ICON output. From the top of

the atmosphere down in each column, the first model grid point which has a condensed mass of cloud water plus cloud ice above a threshold of 0.01 g kg^{-1} is utilized to obtain the cloud top pressure (threshold from van den Heever et al., 2010).

An offline computation is used for the total column optical thickness. Cloud top pressure and total column optical depth categorize the clouds into the nine categories, namely: Cumulus, Stratocumulus, Stratus (low level clouds), Altocumulus, Altostratus, Nimbostratus (mid level clouds), Cirrus, Cirrostratus (high clouds), and deep convective clouds. We applied the classification to the perturbed (P2R) and reference (C2R) model simulation. The classified cloud regime distributions are very similar for the perturbed and control simulations (Fig. 9 shows, as an example, the distribution at 11:45 h UTC). Applying the ISCCP classification to MODIS satellite data yields larger low and mid-level cloud coverage, compared to the model, and less convective cloud coverage. The low and mid-level clouds observed by MODIS are classified as ISCCP-Cirrus and Cirrostratus. A convective region (clouds with high column optical thickness and high cloud tops) in the south-east is miss-classified by the model, compared to MODIS, while the convective region in the north-east is over-represented in the model. A part of the difference in simulated and observed cloud types can be attributed to a difference in the spatial distribution of the cloud types at the MODIS overpass time. Both Cirrus and convective clouds are most abundant in late afternoon and evening when no MODIS satellite observations were available.

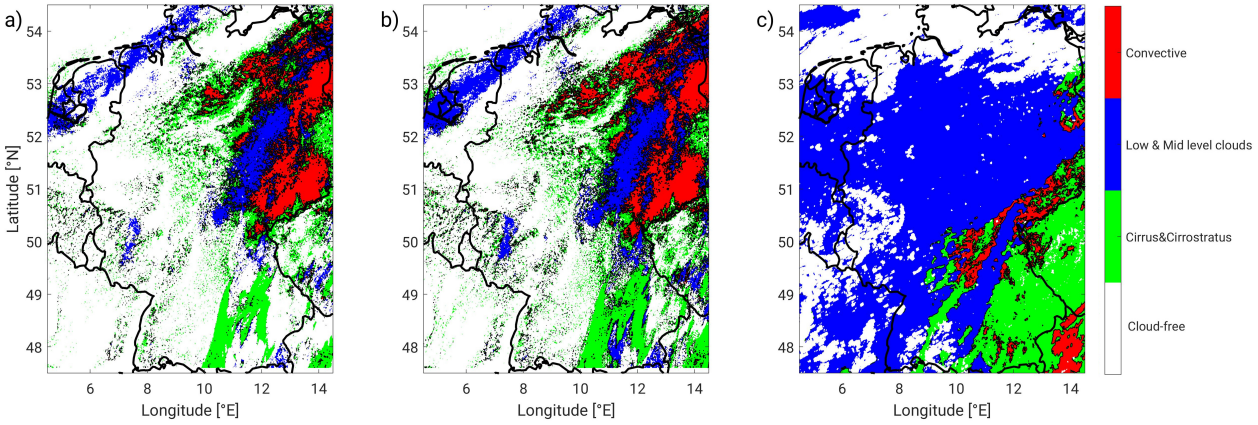


Figure 9. Cloudy regimes after application of the ISCCP classifications to ICON-LEM (a) control, (b) perturbed model output, and (c) MODIS at 11:45 h UTC (MODIS overpass time).

Figure 10 displays normalized frequency of occurrence distributions of N_{d}, cloud water path, and cloud optical thickness for the entire domain for low and mid-level, convective and high (Cirrus and Cirrostratus) clouds from ICON-COSP simulator output and MODIS retrievals at the four times of MODIS overpasses (section 2.3). On the one hand, liquid cloud droplet number concentration does show sensitivity for low and mid-level cloud, since control simulations fits well to MODIS liquid number concentration and perturbed simulation has higher values (already seen in section 3.3). On the other hand, cloud water path is barely sensitive to the CCN perturbation for all cloud regimes considered in the analysis (consistent with the result in section 3.3). However, the control mean values are slightly closer to the observations in all cases. Besides, cloud optical

thickness show a notable increase in the perturbed simulation for low and mid-level clouds, compared to the control simulation. On average, the control simulation is closer to MODIS retrievals than the perturbed simulation. For high and convective clouds, there is not much difference between control and perturbed simulations, and they partly fit to MODIS distribution.

Looking at the different cloud regimes, we can conclude that (1) N_{d} is a suitable variable for detection and attribution of changes of liquid clouds (low and mid-level clouds), (2) there is a potential use of CWP and COT for detection and attribution specifically for low and mid-level clouds, (3) the ice concentrations are too similar in the control and perturbed ICON-LEM simulations and so do not allow for an attribution of an aerosol signal of convective and high clouds (Cirrus and Cirrostratus), at least regarding CWP and COT variables.

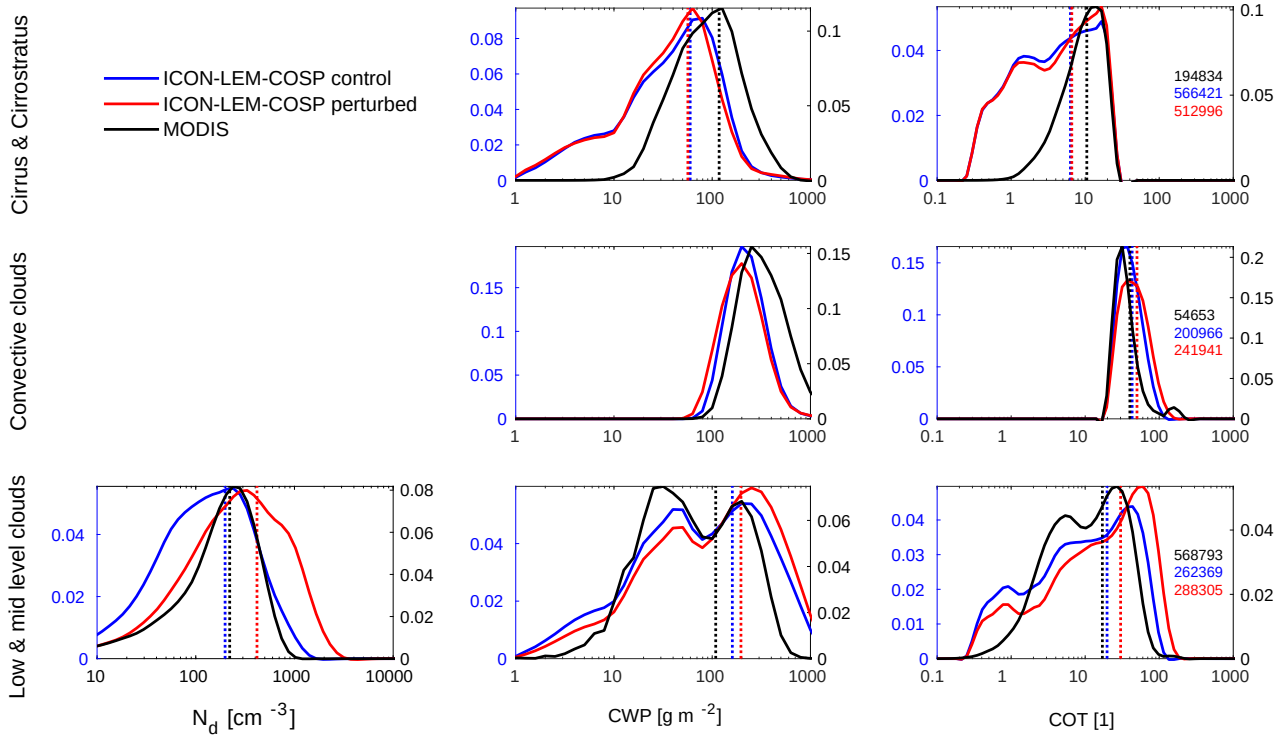


Figure 10. Normalized frequency of occurrence distributions of droplet number concentration (left column, only for low, liquid clouds), cloud water path (liquid plus ice water, center column) and cloud optical thickness (right column) of Cirrus and Cirrostratus (top row), convective clouds (second row), and low and mid-level clouds (bottom row) from ICON-COSP simulator and MODIS (in total four overpasses on 2 May 2013). Left y-axis are for ICON-LEM-COSP and right y-axis for MODIS observations. Straight-dotted lines display means. The numbers on the right side show the number of pixels used for the normalized frequency of occurrence distribution calculation.

3.8 Radiative implications

The effective radiative forcing due to aerosol-cloud interactions (ERFaci) has been estimated by subtracting the control simulation (C2R) domain averages from the perturbed (P2R) ones (Table 7). For the simulated case, the ERFaci is -2.62±1.80 W m^{-2} in the TOA net solar radiation ($R^{\mathrm{s}}_{\mathrm{toa}}$), and 0.21±0.40 W m^{-2} for the TOA net thermal radiation ($R^{\mathrm{t}}_{\mathrm{toa}}$). Consistent with the expectation, the negative forcing in the solar spectrum is slightly reduced by a positive forcing in the terrestrial spectrum (Heyn et al., 2017).

Thanks to the extra simulations that didn't use the number concentration in the radiation transfer computations (C1R and P1R), which also have two different 4D CCN concentration distributions (for 1985 and 2013), the adjustments to the RFaci, as far as they operate via cloud- and precipitation microphysical and dynamical changes, were quantified. In these simulations only the adjustments associated to aerosol forcing are responsible for the radiation changes. The results are noisy signals: the average changes in cloud fraction and LWP are not different from zero to within the temporal variability (not shown). On average, cloud fraction is simulated to decrease slightly (-0.17 %±0.40 %). This result is surprising and different from what is seen in satellite statistics (Gryspeerdt et al., 2019). Further analysis is ongoing. The consequence of the decreasing cloud cover, which is more important radiatively than the increase in LWP, is a positive radiative effect of +0.23 ± 1.24 W m^{-2}. As difference between the ERFaci and the adjustments, RFaci, or the cloud albedo effect (Twomey, 1974), is obtained as -2.85 W m^{-2}.

In order to put this number into context, we assess the ERF computed by general circulation models. Within the 6th Coupled Model Intercomparison Project (CMIP6; Eyring et al., 2016), the Radiative Forcing Model Intercomparison Project (RFMIP; Pincus et al., 2016) defined a simulation dedicated to the assessment of the transient historical effective radiative forcing, the simulation "RFMIP-ERF-HistAerO3". There are four models in the CMIP6 archive that submitted output for these simulations, namely the CanESM5 (Swart et al., 2019), the GFDL-CM4 (Held et al., 2019), the MIROC6 (Tatebe et al., 2019), and the NorESM2-LM (Bentsen et al., 2013; Kirkevåg et al., 2018), which supplied 3, 1, 3, and 2 ensemble members, respectively. The ERFaci is approximated by using the change in cloud radiative effect (CRE, the difference between all-sky and clear-sky top-of-atmosphere net radiation flux density here taken in the solar spectrum only) between two time periods (Quaas et al., 2009). When evaluating the difference in solar CRE of the individual years 2013 and 1850, one obtains as the difference in global annual mean, a multi-model mean of -0.81 W m^{-2}, with an inter-model standard deviation of 0.34 W m^{-2} (using all ensemble members). For a five-year average difference (2010 to 2014 and 1850 to 1854; the periods are not centered around the specific years because the simulations run from 1850 to 2014; Pincus et al., 2016), the values for the global annual mean are -0.71 ±0.35 W m^{-2}. For the domain of the ICON-LEM simulation, and only using May (only monthly output is available), the signal, defined as the difference 1985 minus 2013 is much more noisy since it is averaged much less. The mean and standard deviation are -4.69±13.05 W m^{-2}. To assess the uncertainty, we computed also the change in solar CRE for the months April and May (since the actual day is early May), averaged over the five year-periods 1983 to 1987 minus 2010 to 2014, and a larger domain (10°W to 30°E, 40° to 60°N). This yields a smaller value and much smaller standard deviation of -2.55±2.99 W m^{-2}. The scaling factor for the five-year and bigger European domain is 3.6; the one for the single years and ICON-LEM domain is

5.8. The uncertainty in these scaling factors obtained from the GCMs is very large. Nevertheless it may be instructive to know that the forcing for May, considering the large difference in aerosol levels over Europe between 1985 and 2013 is a factor of 4 to 6 larger than considering the global ERFaci between 2013 and 1850. The -2.6 W m^{-2} obtained in this simulation (Table 7) thus would imply a global, annual mean ERFaci for 2013 vs. 1850 of between -0.4 and -0.7 W m^{-2}.

4 Discussion and Summary

This study used a new type of large-eddy simulation which was carried out over a very large domain and driven as a numerical weather prediction with realistic initial and boundary conditions, including an interactive land surface. A large set of observational data from various sources is used aiming for detection and attribution. Four simulations with ICON-LEM model over Germany were carried out with different time-varying prescribed CCN concentration distributions. The 4D cloud condensation nuclei concentration inputs generated with COSMO-MUSCAT for 2 May 2013 and 1985 were demonstrated to be consistent with the satellite retrievals of AOD by several AVHRR instruments on different NOAA satellites. Furthermore, the control simulation (C2R) results are consistent with the CCN profile as retrieved from ground-based lidar at two sites in 2013, while the ones for the perturbed simulation (P2R), with 1985 conditions, are not.

In terms of cloud quantities, it was demonstrated that detection and attribution of the aerosol-induced changes of the droplet number concentration is also possible. The simulated cloud droplet number concentration for the 2013-aerosol simulation (C2R) is consistent with MODIS satellite retrievals, while the perturbed simulation (P2R) results are on average twice higher. An assessment for N_{d} also was possible from ground-based active remote sensing thanks to a new lidar retrieval technique. The result confirmed the conclusions on the basis of the satellite data, namely that N_{d} using the 2013 aerosol is consistent with the 2013 observations, while the perturbed-run output is not.

The other cloud quantities examined included cloud liquid water path, cloud fraction, cloud base height, and cloud lifetime. Satellite data and network ground-based remote sensing were used as observational reference data. For all of these quantities, the ICON model simulated systematic changes between the perturbed (P2R) and control (C2R) aerosol runs. However, in each case, the difference between either model simulation and the observations was larger than the difference between the simulations in response to the different aerosol conditions. In addition, the natural cloud variability was large compared to the signal. A possible exception is that at large LWP (> 200 g m^{-2}), the control simulation was consistent with the satellite retrievals, while the perturbed simulation showed too large LWP. Small changes between perturbed and control simulations are found in the surface rain rate domain average mean (–2.6 %). However, a detection and attribution even with detailed radar observations was impossible. The sensitivity analysis of different variables in different cloud regimes may well depend on the synoptic situation. It cannot be excluded that in other synoptic situations the sensitivity in mixed phase clouds and maybe even high clouds may be significant.

The cloud changes lead to an increase in the cloud albedo, with changes in the solar radiation (ERFaci) at the TOA of -2.62 W m^{-2}. Thanks to a model sensitivity study, the RFaci could be quantified as -2.85 Wm^{-2}. Using information from global

models, this can be scaled up to the global scale, and the present-day vs. pre-industrial time frame, implying a global ERFaci of -0.8 $\mathrm{W\,m^{-2}}$.

Although the simulations in this study are limited to one day over the domain of Germany, this work shows the great potential of combining these new high resolution simulations with a large set of observations for the detection and attribution of aerosol-cloud interactions. In the future this work should be complemented by extended analyses for longer time periods and more regions to further improve our understanding of cloud-aerosol interactions.

Table 7. Domain mean differences between perturbed and control simulation pair (P2R – C2R). The absolute changes are given along with the temporal standard deviation of the domain mean changes. The numbers in brackets are the relative changes.

Variable	Absolute (relative) domain mean difference (P2R – C2R)
Total cloud cover	0.16 ± 0.37 % (0.20 %)
Liquid water path	$7.42 \pm 4.09\,\mathrm{g\,m^{-2}}$ (11.1 %)
Cloud droplet no. conc. (N_d)	$218 \pm 31\,\mathrm{cm^{-3}}$ (143 %)
Rain rate	$-3.28 \pm 8.46\,\mathrm{g\,m^{-2}\,h^{-1}}$ (–2.6 %)
TOA net solar radiation	$-2.62 \pm 1.80\ \mathrm{W\,m^{-2}}$ (–0.58 %)
TOA net terrestrial rad.	$0.21 \pm 0.40\ \mathrm{W\,m^{-2}}$ (–0.09 %)
Cloud top pressure	-263 ± 180 Pa (–0.35 %)
Cloud base pressure	-140 ± 155 Pa (–0.17 %)

Appendix A

Author contributions. CCH, PS, UB, SC, CH, IT, AS and JQ conceived the study with input and revisions from all authors. MCS, OS, CA, HB, CCH, CG, JH, CJ, MK, NM, RS, PS, FS performed the analysis with contributions from the other authors. CG, JK, CIM, MB, GC, JFE, KF, KG, RH and PKS updated the ICON model with the necessary revisions for this study, created necessary input, performed and postprocessed the simulations. MCS, OS and JQ with input from all authors wrote the manuscript.

Competing interests. The authors declare that they have no conflict of interests.

Acknowledgements. We thank Christine Knist and Ulrich Görsdorf from the Lindenberg Meteorological Observatory – Richard Assmann Observatory for providing us with data from the microwave radiometers and radar, respectively; Werner Thomas from the Deutscher Wetterdi-

Table A1. List of stations with ICON-LEM meteogram output used in the scope of this study.

Station name	Latitude (°N)	Longitude (°E)
LACROS_HOPE	50.880	6.415
RAO	52.210	14.128
Cabauw	51.854	4.927
ETGB_Bergen	52.810	9.930
EDZE_Essen	51.400	6.960
Greifswald	54.100	13.400
ETGI_Idar-Oberstein	49.700	7.330
ETGK_Kuemmersbruck	49.430	11.900
Meiningen	50.560	10.380
Muenchen-Oberschleissheim	48.250	11.550
Norderney	53.710	7.150
Schleswig	54.530	9.550
Stuttgart	48.830	9.200
Bayreuth	49.979	11.681
Nordholz	53.778	8.668
Ziegendorf	53.311	11.837
Frankfurt_EDDF	50.035	8.555
Duesseldorf_EDDL	51.288	6.769
Hamburg_EDDH	53.633	9.994
Berlin_Tegel_EDDT	52.560	13.288
LACROS_Leipzig	51.353	12.435

enst – Hohenpeißenberg Meteorological Observatory for providing us with high-resolution ceilometer data; and Björn Stevens from the Max Planck Institute for Meteorology for his advice in the "Radiative implications" section. This work is funded by the German Federal Ministry of Education and Research (BMBF) within the framework programme "Research for Sustainable Development (FONA)", www.fona.de, through the research programme "HD(CP)² - High Definition Clouds and Precipitation for Climate Prediction", under the numbers FKZ 01LK1209C, 01LK1212C, 01LK1501E, 01LK1502I, 01LK1503A, 01LK1503E, 01LK1503F, 01LK1503G, 01LK1503H, 01LK1504A and 01LK1507A. The authors gratefully acknowledge the computing time granted through JARA-HPC on the supercomputers JUQUEEN and JURECA at Forschungszentrum Jülich.

The authors thank two anonymous reviewers for their constructive remarks.

References

Abdul-Razzak, H. and Ghan, J.: A parametrization of aerosol activation. 2. Multiple aerosol types, J. Geophys. Res., 105, 6837–6844, https://doi.org/10.1029/1999JD901161, 2000.

Ackerman, A. S., Kirkpatrick, M. P., Stevens, D. E., and Toon, O. B.: The impact of humidity above stratiform clouds on indirect aerosol climate forcing, Nature, 432, 1014–1017, https://doi.org/10.1038/nature03137, 2004.

Acquistapace, C., Kneifel, S., Löhnert, U., Kollias, P., Maahn, M., and Bauer-Pfundstein, M.: Optimizing observations of drizzle onset with millimeter-wavelength radars, J. Tech., 10, 1783–1802, https://doi.org/10.5194/amt-10-1783-2017, 2017.

Acquistapace, C., Löhnert, U., Maahn, M., and Kollias, P.: A New Criterion to Improve Operational Drizzle Detection with Ground-Based Remote Sensing, J. Atmos. Oceanic Technol., 36, 781–801, https://doi.org/10.1175/JTECH-D-18-0158.1, 2019.

Albrecht, B.: Aerosols, Cloud Microphysics, and Fractional Cloudiness, Science, 245, 1227–1230, https://doi.org/10.1126/science.245.4923.1227, 1989.

Ansmann, A., Mamouri, R.-E., Hofer, J., Baars, H., Althausen, D., and Abdullaev, S. F.: Dust mass, CCN, and INP profiling with polarization lidar: Updated POLIPHON conversion factors from global AERONET analysis, Atmos. Meas. Tech., in press, https://doi.org/10.5194/amt-2019-98, 2019.

Baars, H., Kanitz, T., Engelmann, R., Althausen, D., Hesse, B., Komppula, M., Preissler, J., Ansmann, A. and, W. U., Lim, J.-H., Young Ahn, J., Stachlewska, I., Amiridis, V., Marinou, E., Seifert, P., Hofer, J., Skupin, A., Schneider, F., Bohlmann, S., Foth, A., Bley, S., Pfüller, A., Giannakaki, E., Lihavainen, H., Viisanen, Y., Kumar Hooda, R., Nepomuceno Pereira, S., Bortoli, D., Wargenr, F., Mattis, I., Janicka, L., Markowicz, K., Achtert, P., Artaxo, P., Pauliquevis, T., Souza, R., Prakesh Sharma, V., van Zyl, P., Beukes, J., Sun, J., Rohwer, E., Deng, R., Momouri, R.-E., and Zamorano, F.: An overview of the first decade of PollyNET: an emerging network of automated Raman-polarization lidars for continuous aerosol profiling, Atmos. Chem. Phys., 16, 5111–5137, https://doi.org/10.5194/acp-16-5111-2016, 2016.

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., and Reinhardt, T.: Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities, Mon. Weather Rev., 139, 3887–3905, https://doi.org/10.1175/MWR-D-10-05013.1, 2011.

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I. A., Hoose, C., and Kristjánsson, J. E.: The Norwegian Earth System Model, NorESM1-M â€“ Part 1: Description and basic evaluation of the physical climate, Geoscientific Model Development, 6, 687–720, https://doi.org/10.5194/gmd-6-687-2013, 2013.

Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J. L., Klein, S. A., Zhang, Y., Marchand, R., Haynes, J. M., Pincus, R., and John, V. O.: COSP: Satellite simulation software for model assessment, Bull. Am. Meteorol. Soc., 92, 1023–1043, https://doi.org/10.1175/2011BAMS2856.1, 2011.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S. K., Sherwood, S., Stevens, B., and Zhang, X. Y.: Clouds and Aerosols. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2013.

Bühl, J., Seifert, P., Wandinger, U., Baars, H., Kanitz, T., Schmidt, J., Myagkov, A., Engelmann, R., Skupin, A., Heese, B., Klepel, A., Althausen, D., and Ansmann, A.: LACROS: the Leipzig Aerosol and Cloud Remote Observations System, Proc. SPIE 8890, Remote Sensing of Clouds and the Atmosphere XVIII; and Optics in Atmospheric Propagation and Adaptive Systems XVI, 889002, 6, https://doi.org/10.1117/12.2030911, 2013.

Cherian, R., Quaas, J., Salzmann, M., and Wild, M.: Pollution trends over Europe constrain global aerosol forcing as simulated by climate models, Geophys. Res. Lett., 41, 2176–2181, https://doi.org/10.1002/2013GL058715, 2014.

Costa-Surós, M., Calbó, J., González, J., and Martin-Vide, J.: Behavior of cloud base height from ceilometer measurements, Atmos. Res., 127, 64–76, https://doi.org/10.1016/j.atmosres.2013.02.005, 2013.

Dipankar, A., Stevens, B., Heinze, R., Moseley, C., Zängl, G., Giorgetta, M., and Brdar, S.: Large eddy simulations using the general circulation model ICON, J. Adv. Model. Earth Syst., 7, https://doi.org/10.1002/2015MS000431, 2015.

Engelmann, R., Kanitz, T., Baars, H., Heese, B., Althausen, D., Skupin, A., Wandinger, U., Komppula, M., Stachlewska, I. S., Amiridis, V., Marinou, E., Mattis, I., Linné, H., and Ansmann, A.: The automated multiwavelength Raman polarization and water-vapor lidar PollyXT: the neXT generation, Atmos. Meas. Tech., 9, 1767–1784, https://doi.org/10.5194/amt-9-1767-2016, 2016.

European Monitoring and Evaluation Programme: EMEP, http://www.emep.int/, 2019.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fan, J., Leung, L. R., Rosenfeld, D., Chen, Q., Li, Z., Zhang, J., and Yan, H.: Microphysical effects determine macrophysical response for aerosol impacts on deep convective clouds, Proc. Natl. Acad. Sci., 110, E4581–E4590, https://doi.org/10.1073/pnas.1316830110, 2013.

Feingold, G., Koren, I., Wang, H., Xue, H., and Brewer, W. A.: Precipitation-generated oscillations in open cellular, Nature, 466, 849–852, https://doi.org/10.1038/nature09314, 2010.

Feingold, G., McComiskey, A., Yamaguchi, T., Johnson, J. S., Carslaw, K. S., and Schmidt, K. S.: New approaches to quantifying aerosol influence on the cloud radiative effect, Proc. Natl. Acad. Sci., 113, https://doi.org/10.1073/pnas.1514035112, 2016.

Genz, C., Schrödner, R., Heinold, B., Henning, S., Baars, H., Spindler, G., and Tegen, I.: Estimation of Cloud Condensation Nuclei number concentrations and comparison to in-situ and lidar observations during the HOPE experiments, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-742, 2019.

Giorgetta, M. A., Brokopf, R., Crueger, T., Esch, M., Fiedler, S., Helmert, J., Hohenegger, C., Kornblueh, L., Köhler, M., Manzini, E., Mauritsen, T., Nam, C., Raddatz, T., Rast, S., Reinert, D., Sakradzija, M., Schmidt, H., Schneck, R., Schnur, R., Silvers, L., Wan, H., Zängl, G., and Stevens, B.: ICON-A, the Atmosphere Component of the ICON Earth System Model: I. Model Description, J. Adv. Model. Earth Syst., pp. 1638–1662, https://doi.org/10.1029/2017MS001242, 2018.

Gordon, H., Field, P., Abel, S., Dalvi, M., Grosvenor, D., Hill, A., Johnson, B., Miltenberger, A., Yoshioka, M., and Carslaw, K.: Large simulated radiative effects of smoke in the south-east Atlantic, Atmos. Chem. Phys., 18, 15 261–15 289, https://doi.org/10.5194/acp-18-15261-2018, 2018.

Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., Alexandrov, M. D., Bennartz, R., Boers, R., Cairns, B., Chiu, C., Christensen, M., Deneke, H., Diamond, M., Feingold, G., Fridlind, A., Hünerbein, A., Knist, C., Kollias, P., Marshak, A., McCoy, D., Merk, D., Painemal, D., Rausch, J., Rosenfeld, D., Russchenberg, H., Seifert, P., Sinclair, K., Stier, P., B. van D., Wendisch, M., Werner, F., Wood, R., Zhang, Z., and Quaas, J.: Remote sensing of cloud droplet number concentration in warm clouds: A review of the current state of knowledge and perspectives, Rev. Geophys., 56, 409–453, https://doi.org/10.1029/2017RG000593, 2018.

Gryspeerdt, E., Goren, T., Sourdeval, O., Quaas, J., Mülmenstädt, J., Dipu, S., Unglaub, C., Gettelman, A., and Christensen, M.: Constraining the aerosol influence on cloud liquid water path, Atmos. Chem. Phys., 19, 5331–5347, https://doi.org/10.5194/acp-19-5331-2019, 2019.

Hande, L. B., Engler, C., Hoose, C., and Tegen, I.: Parameterizing Cloud Condensation Nuclei concentrations during HOPE, Atmos. Chem. Phys., pp. 12 059–12 079, https://doi.org/10.5194/acp-16-12059-2016, 2016.

Heinold, B., Tegen, I., Bauer, S., and Wendisch, M.: Regional modelling of Saharan dust and biomass-burning smoke. Part 2: Direct radiative forcing and atmospheric dynamic response, Tellus B: Chem. Phys. Meteorol., 63, 800–813, https://doi.org/10.1111/j.1600-0889.2011.00574.x, 2011a.

Heinold, B., Tegen, I., Schepanski, K., Tesche, M., Esselborn, M., Freudenthaler, V., Gross, S., Kandler, K., Knippertz, P., Müller, D., Schladitz, A., Toledano, C., Weinzierl, B., Ansmann, A., Althausen, D., Müller, T., Petzold, A., and Wiedensohler, A.: Regional modelling of Saharan dust and biomass-burning smoke. Part 1: Model description and evaluation, Tellus B Chem. Phys. Meteorol., 63, 781–799, https://doi.org/10.1111/j.1600-0889.2011.00570.x, 2011b.

Heinze, R., Dipankar, A., Henken, C. C., Moseley, C., Sourdeval, O., Xie, X., Adamidis, P., Ament, F., Barthlott, C., Behrendt, A., Blahak, U., Bley, S., Brdar, S., Brueck, M., Crewell, S., Deneke, H., Girolamo, P. D., Evaristo, R., Frank, C., Gorges, K., Hanke, M., Hansen, A., Hoose, C., Jahns, T., Kalthoff, N., Klocke, D., Kneifel, S., Knippertz, P., Kuhn, A., Laar, T. V., Macke, A., Maurer, V., Meyer, C., Muppa, S., Neggers, R., Orlandi, E., Pantillon, F., Pospichal, B., Scheck, L., Seifert, A., Senf, F., Siligam, P., Simmer, C., Steinke, S., Stevens, B., Wapler, K., Weniger, M., Wulfmeyer, V., Zhang, D., and Quaas, J.: Large-eddy simulations over Germany using ICON: A comprehensive evaluation, Q. J. R. Meteorol. Soc., 143, 69–100, https://doi.org/10.1002/qj.2947, 2017.

Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., Shevliakova, E., Winton, M., Zhao, M., Bushuk, M., Wittenberg, A. T., Wyman, B., Xiang, B., Zhang, R., Anderson, W., Balaji, V., Donner, L., Dunne, K., Durachta, J., Gauthier, P. P. G., Ginoux, P., Golaz, J.-C., Griffies, S. M., Hallberg, R., Harris, L., Harrison, M., Hurlin, W., John, J., Lin, P., Lin, S.-J., Malyshev, S., Menzel, R., Milly, P. C. D., Ming, Y., Naik, V., Paynter, D., Paulot, F., Rammaswamy, V., Reichl, B., Robinson, T., Rosati, A., Seman, C., Silvers, L. G., Underwood, S., and Zadeh, N.: Structure and Performance of GFDL's CM4.0 Climate Model, Journal of Advances in Modeling Earth Systems, 11, 3691–3727, https://doi.org/10.1029/2019MS001829, 2019.

Heyn, I., Block, K., Mülmenstädt, J., Gryspeerdt, E., Kühne, P., Salzmann, M., and Quaas, J.: Assessment of simulated aerosol effective radiative forcings in the terrestrial spectrum, Geophys. Res. Lett., 44, 1001–1007, https://doi.org/10.1002/2016GL071975, 2017.

Illingworth, A., Hogan, R., O'Connor, E., Bouniol, D., Brooks, M., Delanoé, J., Donovan, D., Eastment, J., Gaussiat, N., Goddard, J., Haeffelin, M., Baltink, H., Krasnov, O., Pelon, J., Piriou, J., Protat, A., Russchenberg, H., Seifert, A., Tompkins, A., G. van Z., Vinit, F., Willén, U., Wilson, D., and Wrench, C.: Cloudnet. Continuous Evaluation of Cloud Profiles in Seven Operational Models Using Ground-Based Observations, Bull. Amer. Meteor. Soc., 88, 883–898, https://doi.org/10.1175/BAMS-88-6-883, 2007.

Jimenez, C., Ansmann, A., Donovan, D., Engelmann, R., Malinka, A., Schmidt, J., and Wandinger, U.: Retrieval of microphysical properties of liquid water clouds from atmospheric lidar measurements: Comparison of the Raman dual field of view and the depolarization techniques, Proc. SPIE, 10429, 1042 907, https://doi.org/10.1117/12.2281806, 2017.

Jimenez, C., Ansmann, A., Engelmann, R., Haarig, M., Schmidt, J., and Wandinger, U.: Polarization lidar: an extended three-signal calibration approach, Atmos. Meas. Tech., 12, 1077–1093, https://doi.org/10.5194/amt-12-1077-2019, 2019a.

Jimenez, C., Ansmann, A., Engelmann, R., Malinka, A., Schmidt, J., and Wandinger, U.: Retrieval of microphysical properties of liquid water clouds from atmospheric lidar measurements: Dual field of view depolarization technique, Atmos. Meas. Tech., in preparation, 2019b.

Khain, A., Rosenfeld, D., and Pokrovsky, A.: Aerosol impact on the dynamics and microphysics of deep convective clouds, Q. J. R. Meteorol. Soc., 131, 2639–2663, https://doi.org/10.1256/qj.04.62, 2005.

Kirkevåg, A., Grini, A., Olivié, D., Seland, Ø., Alterskjær, K., Hummel, M., Karset, I. H. H., Lewinschal, A., Liu, X., Makkonen, R., Bethke, I., Griesfeller, J., Schulz, M., and Iversen, T.: A production-tagged aerosol module for Earth system models, OsloAero5.3 – extensions and updates for CAM5.3-Oslo, Geoscientific Model Development, 11, 3945–3982, https://doi.org/10.5194/gmd-11-3945-2018, 2018.

Knoth, O. and Wolke, R.: An explicit-implicit numerical approach for atmospheric chemistry-transport modeling, Atmos. Environ., 32, 1785–1797, https://doi.org/10.1016/S1352-2310(97)00476-7, 1998.

Löhnert, U., Schween, J. H., Acquistapace, C., Ebell, K., Maahn, M., Barrera-Verdejo, M., Hirsikko, A., Bohn, B., Knaps, A., O'Connor, E., Simmer, C., Wahner, A., and Crewell, S.: JOYCE: Jülich Observatory for Cloud Evolution, Bull. Am. Meteorol. Soc., 96, 1157–1174, https://doi.org/10.1175/BAMS-D-14-00105.1, 2015.

Maahn, M.: Exploiting vertically pointing Doppler radar for advancing snow and ice cloud observations, Ph.D. thesis, University of Cologne, http://kups.ub.uni-koeln.de/6002/, 2015.

Maahn, M., Löhnert, U., Kollias, P., Jackson, R. C., and McFarquhar, G. M.: Developing and Evaluating Ice Cloud Parameterizations for Forward Modeling of Radar Moments Using in situ Aircraft Observations, J. Atmos. Oceanic Technol., 32, 880–903, https://doi.org/10.1175/JTECH-D-14-00112.1, 2015.

Macke, A., Seifert, P., Baars, H., Beekmans, C., Behrendt, A., Bohn, B., Bühl, J., Crewell, S., Damian, T., Deneke, H., Düsing, S., Foth, A., Di Girolamo, P., Hammann, E., Heinze, R., Hirsikko, A., Kalisch, J., Kalthoff, N., Kinne, S., Kohler, M., Löhnert, U., Madhavan, B. L., Maurer, V., Muppa, S. K., Schween, J., Serikov, I., Siebert, H., Simmer, C., Späth, F., Steinke, S., Träumner, K., Wehner, B., Wieser, A., Wulfmeyer, V., and Xie, X.: The HD(CP)2 Observational Prototype Experiment (HOPE= - an overview, Atmos. Chem. Phys., 17, 4887–4914, https://doi.org/10.5194/acp-17-4887-2017, 2017.

Malavelle, F. F., Haywood, J. M., Jones, A., Gettelman, A., Clarisse, L., Bauduin, S., Allan, R. P., Karset, I. H. H., Kristjánsson, J. E., Oreopoulos, L., Cho, N., Lee, D., Bellouin, N., Boucher, O., Grosvenor, D. P., Carslaw, K. S., Dhomse, S., Mann, G. W., Schmidt, A., Coe, H., Hartley, M. E., Dalvi, M., Hill, A. A., Johnson, B. T., Johnson, C. E., Knight, J. R., O'Connor, F. M., Partridge, D. G., Stier, P., Myhre, G., Platnick, S., Stephens, G. L., Takahashi, H., and Thordarson, T.: Strong constraints on aerosol-cloud interactions from volcanic eruptions, Nature, 546, 485–491, https://doi.org/10.1038/nature22974, 2017.

Mamouri, R. E. and Ansmann, A.: Potential of polarization lidar to provide profiles of CCN- and INP-relevant aerosol parameters, Atmos. Chem. Phys., 16, 5905–5931, https://doi.org/10.5194/acp-16-5905-2016.

Martucci, G., Milroy, C., and O'Dowd, C. D.: Detection of Cloud-Base Height Using Jenoptik CHM15K and Vaisala CL31 Ceilometers, J. Atmos. Ocean. Technol., 27, 305–318, https://doi.org/10.1175/2009JTECHA1326.1, 2010.

Meier, J., Tegen, I., Mattis, I., Wolke, R., Alados Arboledas, L., Apituley, A., Balis, D., Barnaba, F., Chaikovsky, A., Sicard, M., Pappalardo, G., Pietruczuk, A., Stoyanov, D., Ravetta, F., and Rizi, V.: A regional model of European aerosol transport: Evaluation with sun photometer, lidar and air quality data, Atmos. Environ., 47, 519–532, https://doi.org/10.1016/j.atmosenv.2011.09.029, 2012.

Miltenberger, A. K., Field, P. R., Hill, A. A., Rosenberg, P., Shipway, B. J., Wilkinson, J. M., Scovell, R., and Blyth, A. M.: Aerosol–cloud interactions in mixed-phase convective clouds - Part 1: Aerosol perturbations, Atmos. Chem. Phys., 18, 3119–3145, https://doi.org/10.5194/acp-18-3119-2018, 2018.

Mülmenstädt, J. and Feingold, G.: The Radiative Forcing of Aerosol–Cloud Interactions in Liquid Clouds: Wrestling and Embracing Uncertainty, Curr. Clim. Chang. Reports, 4, 23–40, https://doi.org/10.1007/s40641-018-0089-y, 2018.

Pincus, R., Forster, P. M., and Stevens, B.: The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6, Geoscientific Model Development, 9, 3447–3460, https://doi.org/10.5194/gmd-9-3447-2016, 2016.

Platnick, S., King, M. D., Meyer, K. G., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Ridgway, B., and Riédi, J.: MODIS Cloud Optical Properties: User Guide for the Collection 6 Level-2 MOD06/MYD06 Product and Associated Level-3 Datasets, Version 1.0, available at: https://modis-images.gsfc.nasa.gov/_docs/ C6MOD06OPUserGuide.pdf (last access: 31 May 2016), 2015.

Platnick, S., Meyer, K. G., King, M. D., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Holz, R. E., Yang, P., Ridgway, W. L., and Riedi, J.: The MODIS Cloud Optical and Microphysical Products: Collection 6 Updates and Examples From Terra and Aqua, IEEE Trans Geosci Remote Sens., 55, 502–525, https://doi.org/10.1109/TGRS.2016.2610522, 2017.

Quaas, J.: Approaches to Observe Anthropogenic Aerosol-Cloud Interactions, Curr. Clim. Chang. Reports, 1, 297–304, https://doi.org/10.1007/s40641-015-0028-0, 2015.

Quaas, J., Boucher, O., and Lohmann, U.: Constraining the total aerosol indirect effect in the LMDZ and ECHAM4 GCMs using MODIS satellite data, Atmos. Chem. Phys., 6, 947–955, https://doi.org/10.5194/acp-6-947-2006, 2006.

Quaas, J., Ming, Y., Menon, S., Takemura, T., Wang, M., Penner, J. E., Gettelman, A., Lohmann, U., Bellouin, N., Boucher, O., Sayer, A. M., Thomas, G. E., McComiskey, A., Feingold, G., Hoose, C., Kristjánsson, J. E., Liu, X., Balkanski, Y., Donner, L. J., Ginoux, P. A., Stier, P., Grandey, B., Feichter, J., Sednev, I., Bauer, S. E., Koch, D., Grainger, R. G., Kirkevag, A., Iversen, T., Seland, Ø., Easter, R., Ghan, S. J., Rasch, P. J., Morrison, H., Lamarque, J.-F., Iacono, M. J., Kinne, S., and Schulz, M.: Aerosol indirect effects – general circulation model intercomparison and evaluation with satellite data, Atmospheric Chemistry and Physics, 9, 8697–8717, https://doi.org/10.5194/acp-9-8697-2009, 2009.

Rossow, W. and Schiffer, R.: ISCCP Cloud Data Products, Bull. Am. Meteorol. Soc., 72, 2–20, https://doi.org/10.1175/1520-0477(1991)072<0002:ICDP>2.0.CO;2, 1991.

Rémillard, J., Fridlind, A., Ackerman, A., Tselioudis, G., Kollias, P., Mechem, D., Chandler, H., Luke, E., Wood, R., Witte, M., and Ayers, J.: Use of cloud radar Doppler spectra to evaluate stratocumulus drizzle size distributions in large-eddy simulations with size-resolved microphysics, J. Appl. Meteorol. Climatol., 56, 3263–3283, https://doi.org/10.1175/JAMC-D-17-0100.1, 2017.

Sandu, I., Brenguier, J.-L., Geoffroy, O., Thouron, O., and Masson, V.: Aerosol Impacts on the Diurnal Cycle of Marine Stratocumulus, J. Atmos. Sci., 65, 2705–2718, https://doi.org/10.1175/2008jas2451.1, 2008.

Schättler, U., Doms, G., and Schraff, C.: A Description of the Nonhydrostatic Regional COSMO-Model. Part VII: User's Guide, Deutscher Wetterdienst, Offenbach, 2014.

Schmidt, J., Wandinger, U., and Malinka, A.: Dual-field-of-view Raman lidar measurements for the retrieval of cloud microphysical properties, Appl. Opt., 52, 2235–2247, https://doi.org/10.1364/AO.52.002235, 2013.

Schmidt, J., Ansmann, A., Bühl, J., Baars, H., Wandinger, U., Müller, D., and Malinka, A. V.: Dual-FOV Raman and Doppler lidar studies of aerosol-cloud interactions: Simultaneous profiling of aerosols, warm-cloud properties, and vertical wind, J. Geophys. Res. Atmos., 119, 5512– 5527, https://doi.org/10.1002/2013JD020424, 2014.

Seifert, A. and Beheng, K. D.: A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 1: Model description, Meteorol. Atmos. Phys., 92, 45–66, https://doi.org/10.1007/s00703-005-0112-4, 2006.

Seifert, A., Heus, T., Pincus, R., and Stevens, B.: Large-eddy simulation of the transient and near-equilibrium behavior of precipitating shallow convection, J. Adv. Model. Earth Syst., 7, 1918–1937, https://doi.org/10.1002/2015MS000489, 2015.

Seinfeld, J. H., Bretherton, C., Carslaw, K. S., Coe, H., DeMott, P. J., Dunlea, E. J., Feingold, G., Ghan, S., Guenther, A. B., Kahn, R., Kraucunas, I., Kreidenweis, S. M., Molina, M. J., Nenes, A., Penner, J. E., Prather, K. A., Ramanathan, V., Ramaswamy, V., Rasch, P. J., Ravishankara, A. R., Rosenfeld, D., Stephens, G., and Wood, R.: Improving our fundamental understanding of the role of aerosol-cloud interactions in the climate system, Proc. Natl. Acad. Sci., 113, 5781–5790, https://doi.org/10.1073/pnas.1514043113, 2016.

Small, J. D., Chuang, P. Y., Feingold, G., and Jiang, H.: Can aerosol decrease cloud lifetime?, Geophys. Res. Lett., 36, 1–5, https://doi.org/10.1029/2009GL038888, 2009.

Smith, S. J., Van Aardenne, J., Klimont, Z., Andres, R. J., Volke, A., and Delgado Arias, S.: Anthropogenic sulfur dioxide emissions: 1850-2005, Atmos. Chem. Phys., 11, 1101–1116, https://doi.org/10.5194/acp-11-1101-2011, 2011.

Stevens, B. and Feingold, G.: Untangling aerosol effects on clouds and precipitation in a buffered system., Nature, 461, 607–613, https://doi.org/10.1038/nature08281, 2009.

Swales, D., Pincus, R., and Bodas-Salcedo, A.: The Cloud Feedback Model Intercomparison Project Observational Simulator Package: Version 2, Geosci. Model Dev., 11, 77–81, https://doi.org/10.5194/gmd-11-77-2018, 2018.

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), Geoscientific Model Development, 12, 4823–4873, https://doi.org/10.5194/gmd-12-4823-2019, 2019.

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., Sudo, K., Sekiguchi, M., Abe, M., Saito, F., Chikira, M., Watanabe, S., Mori, M., Hirota, N., Kawatani, Y., Mochizuki, T., Yoshimura, K., Takata, K., O'ishi, R., Yamazaki, D., Suzuki, T., Kurogi, M., Kataoka, T., Watanabe, M., and Kimoto, M.: Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6, Geoscientific Model Development, 12, 2727–2765, https://doi.org/10.5194/gmd-12-2727-2019, 2019.

Toll, V., Christensen, M., Gassó, S., and Bellouin, N.: Volcano and Ship Tracks Indicate Excessive Aerosol-Induced Cloud Water Increases in a Climate Model, Geophys. Res. Lett., 44, 12,492–12,500, https://doi.org/10.1002/2017GL075280, 2017.

Toll, V., Christensen, M., Quaas, J., and Bellouin, N.: Weak average liquid-cloud-water response to anthropogenic aerosols, Nature, 572, 51–55, https://doi.org/10.1038/s41586-019-1423-9, 2019.

Twomey, S.: Pollution and the Planetary Albedo, Atmos. Environ., 8, 1251–1256, https://doi.org/10.1016/0004-6981(74)90004-3, 1974.

van den Heever, S. C., Stephens, G. L., and Wood, N. B.: Aerosol Indirect Effects on Tropical Convection Characteristics under Conditions of Radiative–Convective Equilibrium, J. Atmos. Sci., 68, 699–718, https://doi.org/10.1175/2010jas3603.1, 2010.

Wiegner, M. and Geiß, A.: Aerosol profiling with the Jenoptik ceilometer CHM15kx, Atmos. Meas. Tech., 5, 1953–1964, https://doi.org/10.5194/amt-5-1953-2012, 2012.

Wiegner, M., Madonna, F., Binietoglou, I., Forkel, R., Gasteiger, J., Geiß, A., Pappalardo, G., Schäfer, K., and Thomas, W.: What is the benefit of ceilometers for aerosol remote sensing? An answer from EARLINET, Atmos. Meas. Tech., 7, 1979–1997, https://doi.org/10.5194/amt-7-1979-2014, 2014.

Wolke, R., Knoth, O., Hellmuth, O., Schröder, W., and Renner, E.: The parallel model system LM-MUSCAT for chemistry-transport simulations: Coupling scheme, parallelization and applications, Adv. Parallel Comput., 13, 363–369, https://doi.org/10.1016/S0927-5452(04)80048-0, 2004.

Wolke, R., Schröder, W., Schrödner, R., and Renner, E.: Influence of grid resolution and meteorological forcing on simulated European air quality: A sensitivity study with the modeling system COSMO-MUSCAT, Atmos. Environ., 53, 110–130, https://doi.org/10.1016/j.atmosenv.2012.02.085, 2012.

Xue, H. and Feingold, G.: Large-Eddy Simulations of Trade Wind Cumuli: Investigation of Aerosol Indirect Effects, J. Atmos. Sci., 63, 1605–1622, https://doi.org/10.1175/JAS3706.1, 2006.

Zängl, G., Reinert, D., Rípodas, P., and Baldauf, M.: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core, Q. J. R. Meteorol. Soc., 141, 563–579, https://doi.org/10.1002/qj.2378, 2015.