Reply to reviewer 2

We thanks the reviewer for the positive and encouraging review and the comments that helped us to improve the manuscript.

Major comments:

Reviewer: *1. I suggest the authors have a look at the following papers...*

Reply: We have been aware of the listed publications, except for the Morrison and Grabowski (2011, ACPD) paper which has been submitted to ACPD only a few weeks before we submitted our paper, and has still been in open discussion at that time. The literature on aerosol-cloud effects is vast and it is hardly possible to give a full review of all relevant aspects, and at the same time keep the paper concise and short. In an earlier draft of our paper we had a paragraph with a discussion of the other Grabowski and Morrison papers, but they focus mostly on aerosol effects on the radiation budget. The effects on precipitation are difficult to judge in their study, because of their model setup which is basically radiative-convective equilibrium. As our investigation is mostly concerned with effects on precipitation, we decided not to include a detailed discussion of radiative effects. We now include some references and a short discussion in the revised version of our paper.
It is not the aim of our paper to get into a discussion of observation-based estimation of aerosol-cloud effects. Therefore we will not go into further detail concerning the Koren et al. (2010) work.

Reviewer: *Comment on implicit numerical schemes*

Reply: Standard deep convection tests have been done extensively, over and over again, with the COSMO model. If anything, the model produces too strong updrafts (as does the WRF model). The Klemp and Wilhelmson papers are actually not a good benchmark, because their model was overly diffusive.

Reviewer: *The discussion of the homogeneous freezing (bottom/top of 20209/20210) is unclear to me. What is meant by the sentence starting with 'The competition of...' and 'fictitious downdraft' in particular?*

Reply: We would like to refer the reviewer to the Kärcher and Lohmann (2002) as well Kärcher et al. (2006, KHL06 hereafter) papers on the parameterization of homogeneous freezing of supercooled liquid aerosol particles. The term 'fictitious downdraft velocity' has been used by KHL06 to describe pre-existing ice. It is, in the prognostic equation for particle density, equivalent to a reduction of the updraft velocity (i.e., a reduction in the cooling) in the nucleation parameterization. By 'competition' it is referred to the fact that the pre-existing ice from heterogeneous nucleation reduces the supersaturation over ice, especially the supersaturation exceeding the threshold for homogeneous freezing. This reduces the number of homogeneously nucleated ice particles. This chain of processes is sometimes called 'negative Twomey-Effekt' (Kärcher, B. and U. Lohmann, 'A Parameterization of cirrus cloud formation: Heterogeneous freezing'. J. Geophys. Res. 108 (D14), 4402, doi:10.1029/2002JD003220, 2003; see also the textbook by Cotton, Bryan and van den Heever, 'Storm and Cloud Dynamics', 2010, p. 639).

Reviewer: *4. and 5. The discussion at the bottom of p. 20214. As shown in Morrison and Grabowski ACPD paper, changes in the cloud microphysics can explain observed changes of the cloud field without invoking changes in the updraft strength (in fact, convection became slightly weaker, not invigorated in the polluted case in that paper). To support the authors statement, one should look at the maximum velocity within the cores, not the core depth. ...*

Reply: Cloud core depth is, in our opinion, the best diagnostic to show such an effect, although we understand that it might also be affected by microphysical effects. Looking at statistics of updraft strength it is very difficult to distinguish the different stages of deep convection (developing, mature, dissipating). For the aerosol-cloud dynamical effects the developing stage which is the shortest of the three, is the relevant one. This makes it difficult to isolate the effect properly when looking at updraft statistics, in contrast the cloud core depth provides the necessary integration over time over the developing stage.

Figure 1 show histograms of vertical velocity and although they show a small increase in the frequency of stronger updrafts for 'high CCN' they are, in our opinion, quite inconclusive, because the effect is very small. This does, however, not falsify or question our statements about aerosol effects on the dynamics of deep convection. The details of this interaction have been revealed and discussed by many idealized studies of single deep convective cells (e.g. Teller and Levin 2006, Seifert and Beheng 2006b, and many others).

Nevertheless, we agree with the comment that our analysis of the dynamical feedback is not fully conclusive, and although we might have the data, we currently don't have the quantitative tools to settle this discussion. In our opinion a tracking and lifecycle analysis of the deep convection would be necessary to answer those question. We do have 15 min output data of the 3D vertical velocity fields, and this would be sufficient to do this. But unfortunately, we don't have the methods and resources available to do this investigation now.

Reviewer: *6. The impacts on surface temperature... and cold pools... and squall lines*

Reply: Even during summertime, large meso-α-scale squall lines which are driven by their own cold pools, as they occur frequently over the continental United States, are very rare for Germany. Note that the term 'squall line' has originally been introduced by Bjerknes as a line of convective cells associated with a cold front. This definition fits much better for Europe and such lines of cells do occur in our simulations, but their dynamics is very much slaved to the large-scale synoptic situation in which they are embedded. Such situations are, of course, much more complicated than the idealized simulations of Morrison et al. (2009). An exceptional example of a strong meso-α-scale squall line from 2007 is documented and discussed in Baldauf et al. (2011). This squall line is maybe dominated by cold pool dynamics and therefore it shows a strong sensitivity to microphysics. We have been surprised (and disappointed, to be honest) that the three years 2008, 2009 and 2010 have not shown one single squall line case of similar intensity and similar sensitivity to CCN/IN assumptions. Therefore we come to the conclusion that this type of squall lines is just too rare for Germany to play a major role for the precipitation statistics. Of course, it would be better to have 10-20 years of data instead of only three. Nevertheless, we conclude that the results concerning strong cold pool driven squall lines do not necessarily apply to the convection which prevails over Germany.

It would be very interesting to repeat our simulations with a similar setup over the continental U.S., but at DWD we cannot easily motivate to spent computing time on that as it would not directly benefit our NWP system, especially as the computational domain would have to be much larger.

Although we very much appreciate the comment by the reviewer, one should be very careful (a) when transfering subjective experience from one continent to another, because the meteorological situation might be very different (b) to transfer results for very idealized simulation to more complex real case situations.

Reviewer: *7. Importance of large-scale forcing...*

In Europe, in contrast to the continental U.S., deep convection during summertime is much stronger slaved to the large-scale forcing, as we have already elaborated on in the answer to the previous question. The first-author, who has been responsible for monitoring and testing the day-to-day performance of COSMO-DE for many years, has investigated many situations and cases, and the effects of details of the surface forcing and/or properties are most often only relevant on the meso-γ-scale. This is also confirmed by the pre-operational COSMO-DE EPS (ensemble prediction systems) which can objectively measure the different effects of errors/uncertainties of boundary conditions, initial conditions and model physics. In most cases, especially in convective situations, the large-scale forcing, i.e., boundary conditions, are the dominant source of error on the meso-α- and meso-β-scale. Only in weakly-forced situations, which are roughly 1-2 weeks per summer season, the surface forcing (and model errors related to boundary layer physics) become dominant. Therefore this statement that large-scale forcing plays a dominant role is not just speculation, but builts upon many years of experience and objective analysis.

Minor comments:

Line 16, p. 20210: on should be one.

In my version this reads 'a set of hindcasts'. It looks like technical editing of ACPD changed this to 'on set'.

Figure 7 is supposed to show temperature, but the vertical axis label says precipitation. Are Figs. 7 and 10 simply switched?

Yes, the plots of Figs. 7 and 10 have been mixed-up during technical editing.

Are color labels correct in Fig. 8? If so, the core depth seems to be larger for some but not all high-CCN runs compared to similar low-CCN runs.

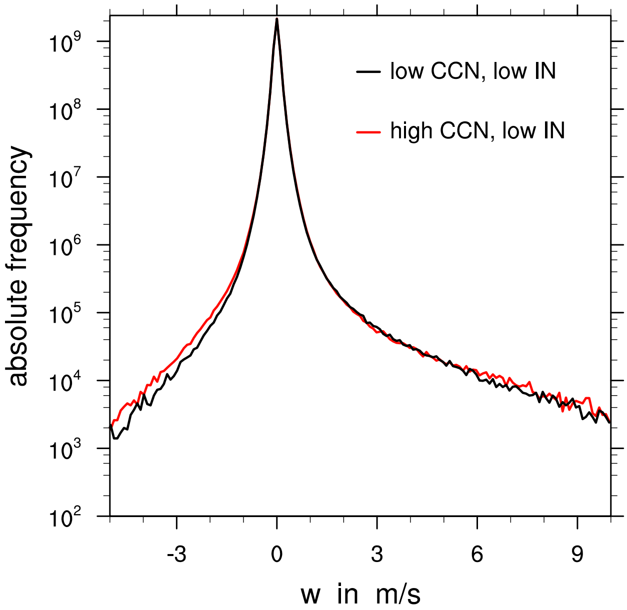
The core depths is larger for all high-CCN runs (long dashed) when compared to the low CCN counterparts (short dashed). Only for the high IN cases (red lines) there is one data point (at 24 h) where the two lines are touching each other.

Overall, the figures were plotted too small to legitimately evaluate them.

We suggest that we will discuss with ACP technical editing whether the plots of Figs. 7, 8, and maybe 10 can be extended to cover the full paper width.

List of Figures

a) unconditional sampling



b) conditional sampling with CAPE > 0 and LWC > 0

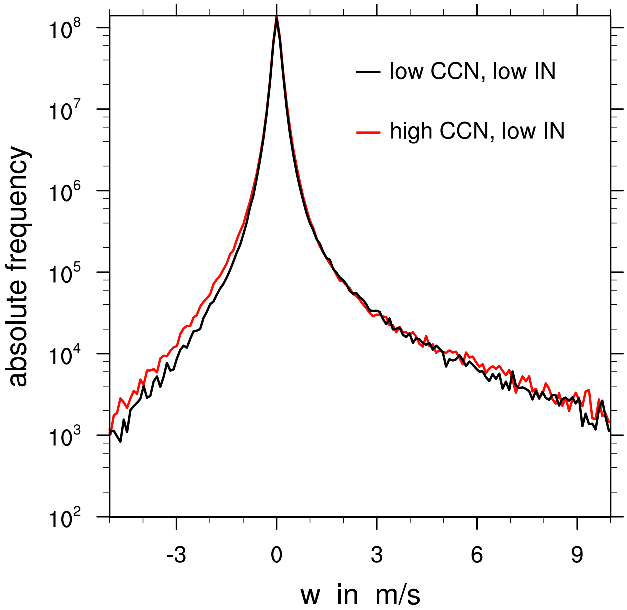


FIG. 1. Histograms of vertical velocity between a height from 5000 m to 7500 m above ground, all data from 2008, 2009 and 2010. Data is from the full evaluation domain using 00 and 12 UTC simulations sampled every 3 hours from 6-15 h simulation time (as Fig. 4 of the manuscript). (a) unconditional sampling (b) only grid-points with CAPE > 0 and liquid water content (LWC) > 0.