

## Reply to comments from reviewer #1

### General comments:

1. *The manuscript contains many typos and errors that the team of authors should really have addressed together through proof-reading before submission for review. I have listed these in my specific comments.*

**Reply :** Thank you very much for pointing these out! We have corrected all the highlighted issues and proof-read the new manuscript more carefully.

2. *I believe there are some inconsistencies in the references to ensemble members. I have noted these in my comments on Section 4.2.*

**Reply :** We are very sorry that there were inconsistencies in ensemble member references in the original manuscript. We have very carefully checked they are ok in the new version.

3. **Figure ordering:** *I found the order of figures somewhat counterintuitive and hard to follow. I had to lay my printed copy of the figures and the supplementary figures out next to each other in order to follow the arguments made in the text. Whilst I appreciate that there are already a lot of figures in the paper, I would suggest moving figure S8 to the main paper if possible. Further, the figure ordering in the main paper is not logical. I appreciate that it is difficult to optimally order figures when investigating concurrent sensitivities, but I would recommend placing the order of current Figures 7,6,5 as such. On P10 L4, the reference to Fig 7a, I had to jump ahead several Figures in order to see this. On P10 L34 I had to jump back again to Fig. 5, which is referenced for the first time after Fig.s 7 and 6 are discussed in detail. Why put Fig 5 in its current location? You would make it much easier for the reader if it appeared after 6 and 7. Page 11, 1<sup>st</sup> paragraph: again, you refer to Fig 9 and then immediately after to Fig 11, and Fig 10 is not even mentioned until page 13.*

**Reply :** Thank you for your suggestions. We have re-ordered the figures so that they appear in the sequence they are mentioned in the text. Also, we changed the partitioning of figures between the main paper and the SI to have only the most crucial figures in the main paper following your suggestions and those of reviewer 2. For example, former SI Fig 8a is moved to the main text (new Fig. 8 b).

### 4. General comments on figures:

*Many of the figures have lines joining the points representing each ensemble member. This is misleading, as the abscissa on these figures show ensemble members (a discrete dataset) and not continuous data. I recommend removing these lines.*

**Reply :** We have removed these lines from all figures as suggested.

*Figure SI 9 – I tried very hard to understand this Figure, but it many things in it don't make sense to me. See notes under my comments referring to individual figures.*

**Reply :** We are were sorry for the poor description of this figure. The caption and labelling has been improved, so it should be comprehensible now.

### 5. Section 2:

*Stochastic physics – are stochastic physics used in the regional model as well as the global model? Are stochastic physics used in the full set of ensemble runs? (Are you using stochastic physics as well as perturbed initial conditions?) What kind of stochastic physics are used? Which schemes and which parameters? Etc. This needs a little more explanation if you are discussing a study which aims to capture meteorological variability.*

**Reply :** Stochastic physics are only used for the re-run of the global operational ensemble and not in the regional ensemble simulations. This has been clarified in the model set-up description (section 2: p. 4, l. 23-25 & p. 5, l. 9-11). Since stochastic physics are only used to derive the perturbed initial and boundary conditions, they do not influence the actually used ensemble data

and we therefore refrain from a more detailed description of the stochastic physics for brevity (reference is provided for interested reader!).

### **6. Section 3.1:**

*P6 L13 – Was the model microphysics output passed offline through the same radar algorithm as the Radarnet data? If not, could part of the difference be because the online UM dBZ calculation is different from the dBZ calculation in the Radarnet algorithm?*

**Reply :** No. The radar reflectivity is computed from the modelled hydrometeor properties online, i.e. within the model. For the modelled surface precipitation, we are using the direct model output, i.e. surface precipitation is not diagnosed from the modelled reflectivity fields. The model assumes only Rayleigh scattering and does not assume particles have a single phase, e.g. partially melted hydrometeors do not exist in the modelling world. Conversely, the radar algorithm does not account for sub-cloud evaporation of rain, which has been shown in previous work to affect retrieved surface precipitation rates. Therefore the differences, could be due not only to a deficiency in the model microphysics, but also to issues with the radar-derived surface precipitation or radar reflectivity calculation in the model.

We added the following to clarify this point (p. 6, l. 17-20):

“While the model derived surface precipitation is the sedimentation flux at the surface, the radar derived surface precipitation is computed from the low-level radar reflectivity according to Harrison et al. (2009). Accordingly, the modelled and radar-derived surface precipitation products involve different assumptions. For example, sub-cloud evaporation is not taken into account in the retrieval of surface precipitation rates from observational data.”

### **7. Section 4.1:**

*P9 L5-6: “These members have a higher cloud fraction” – do you know why this is the case for these members?*

**Reply :** As discussed in section 4.2 these members have a higher boundary layer moisture convergence, which most likely explains the higher cloud fraction. The text has been modified as follows (p. 8, l. 27-30):

“Only in ensemble members 1 and 2 the temperature difference remains smaller than 1.5 K (SI Fig. 7 c). These members have a higher cloud fraction in the morning (not shown), which is likely related to a relatively large large-scale moisture convergence. The higher cloud fraction reduces radiative heating of the land surface explaining the smaller peak land-sea temperature difference.”

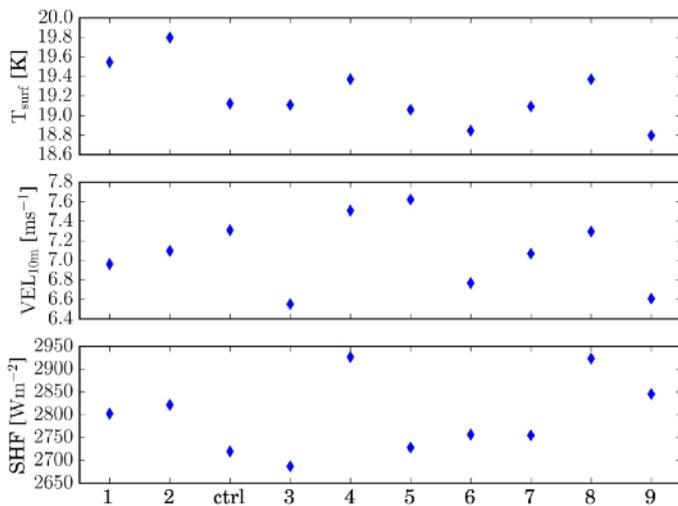
*P9 L17-18: Is this also related to the cloudiness (higher cloud fraction in these members)?*

**Reply :** The smaller wind speed and more southerly wind direction should affect the propagation of the sea-breeze front inland and therefore changes the low-level convergence. However, this relation is not very strict, as e.g. ensemble member 5 to 8 also have similarly high low-level convergence. The latter do not have a particularly large cloud fraction. Therefore, we think the cloud fraction variability is dominated by the differences in large-scale convergence. Although of course differences in sea-breeze convergence strength impact cloud fraction, but these differences are much smaller than those in large-scale convergence.

### **8. Section 4.2:**

*P10 L11-12: Members 4 and 7 have a particularly large surface sensible heat flux – can you explain why? It doesn't seem like they stand out in terms of cloudiness (Fig. 6).*

**Reply :** The high sensible heat flux for these members is a combination of high surface temperature and high surface wind speed (Fig. 1). The surface temperature in these ensemble members is already in the upper range at 9 UTC, i.e. before the onset of significant radiative heating. Also the relatively large surface wind-speed is consistent with the upstream profiles. Therefore the large sensible heat fluxes is likely related to changes in the initial and boundary conditions and not so much related to differences in cloudiness. Since the sensible heat flux is not discussed anymore in the revised version of the manuscript, there are no alterations regarding this issue in the text.



**Figure 1.** Mean surface temperature (top), 10 m wind velocity (middle) and sensible heat flux (bottom) for the different ensemble members.

equilibrium level pressure corresponds to the overall tendency in the mean cloud top height: For example, ensemble 5 with the smallest mean cloud top height has the largest equilibrium level pressure and ensemble 2/9 with the largest mean cloud top height have the smallest equilibrium level pressure. Modified text (p. 9, l. 23-26):

“Mean cloud top height varies by about 750 m between ensemble members (Fig.8) with largest (smallest) values for ensemble member 2 and 9 (5). Variations in mean cloud top height are in general consistent with those of the equilibrium level pressure (SI Fig. 10 c): For example, the equilibrium level pressure in ensemble member 5 is largest, while members 2 and 9 have the smallest equilibrium level pressure.”

*P10 L25-26: ensemble members 1,2,5,8 have a relatively large fraction of deep clouds – I don’t see this. What about e.g. member 6 (Fig SI 8a)?*

**Reply :** The reviewer is correct ensemble member 6 has the largest deep cloud fraction. This sentence is not part of the revised manuscript anymore, because we tried to shorten the manuscript as requested by reviewer 2.

*P10 L29: changes in condensate generation, i.e. air mass lifting – have you looked at the dynamical convergence to see if this is the case?*

**Reply :** We are not sure what the reviewer mean with “dynamical convergence“. The large-scale boundary layer convergence and the low-level convergence have been discussed in section 4.1 and 4.2. Their variability corresponds in general very well with the variability in G.

*P10 L32:*

- “member 8 has a relatively large PE” – I disagree with this. Many others have a greater PE, e.g. 1, ctrl, 6 (Fig. 7)
- “and the largest fraction of clouds with tops above 4.3 km” – I also disagree with this. The largest fraction of clouds with tops above 4.3 km is seen in member 6 (Fig SI 8a). I think in this sentence perhaps the authors mean to refer to member #6, not member #8? Then I agree with the statements made in the sentence.

**Reply :** This statement refers indeed to ensemble member 6 and has been corrected accordingly (p. 9, l. 34).

*P11 L1-4: this final section is not particularly well-explained and no relevance is given. Can you say anything about the processes and impact or importance?*

**Reply :** We have expanded this section as follows (p. 10, l. 8-15):

*P10 L22: largest (smallest) values for ensemble members 8(4) – I find it hard to see by eye on this Figure, but doesn’t this actually apply to members 9(5) not 8(4)?*

**Reply :** We are sorry for the incorrect referencing of ensemble members in the text. The reviewer is of course right and we have corrected the text (p. 9, l. 24; checking the numerical values indicates ensemble number 2 has the highest mean cloud top height very closely followed by ensemble member 9).

*P10 L22-23: Really? I find this hard to see (Fig 6d vs Fig SI 7c)*

**Reply :** The overall tendency in the mean

“Mean reflected shortwave radiation ranges from 130 W m<sup>-2</sup> to 155 W m<sup>-2</sup> (Fig. 11 a). The reflected shortwave is influenced by the cloud cover and the cloud droplet number concentrations. The largest (smallest) outgoing shortwave flux is predicted for the ensemble members with the largest (smallest) cloud fraction, i.e. ensemble 1 (8). Since the CDNC variability is small (Fig. 5), the variations in cloud fraction between ensemble members is dominating the variability of outgoing shortwave radiation. Changes in outgoing longwave radiation are on the order of 3 W m<sup>-2</sup> (Fig. 11 b). The outgoing longwave radiation is influenced by the surface temperature, the cloud top height and the cloud fraction. While differences in the cloud top height distribution contribute to the variability in outgoing longwave radiation, variations in the clear sky outgoing longwave radiation dominate the overall variability due to the relatively small cloud fraction (SI Fig. 13).“

*P11 L3-4: Largest (smallest) values occur for ensemble members 2(7) – I only just agree with this. Do you mean member 1 not member 2 for the largest SW radiation values and largest CF?*

**Reply :** Yes, this is member 1. We changed the text accordingly (p. 10, l. 10).

**Specific comments and typos:**

*P1 L16: “consider” -> “considered”*

*P2 L3: “climate system”*

*P2 L4: “The main issues...”*

*P2 L5: “... on model grid scales several orders of magnitude larger, and the...” P2 L6: “In the last few decades” / “In recent decades”*

*P2 L6-7: “the modification of cloud properties has been studied in particular”*

*P2 L7-8: “... and the relation between particle number concentrations and radiation” – this whole sentence feels quite clumsy.*

*P2 L13: “necessitated by” -> “necessary because of”*

*P2 L15: “changes to simulated for individual clouds” – simulated what?*

**Reply :** Thank you for pointing these out. All issues have been fixed as suggested.

*P2 L17: You could also include a reference here to the 2012 paper by Seoung Soo Lee where placing an aerosol perturbation in the mesoscale domain of a simulation led to intensification of convection within an MCS but suppressed precipitation in the larger-scale domain. (Reference provided at the end of this set of comments)*

**Reply :** We have added this reference (p. 2, l. 14-19):

“These interactions can at least partly compensate the large changes simulated for individual clouds (e.g. Lee, 2012; Seifert et al. 2012). In a case-study of tropical deep convection, Lee (2012) found that locally invigorated convection in polluted conditions induces stronger large-scale subsidence resulting in an overall suppression of precipitation on a cloud-system scale. Seifert et al. (2012) demonstrated with simulations extending over three summer seasons that aerosol perturbations can produce large local changes in precipitation, while not significantly changing the mean precipitation.“

*P2 L29: Southern Great Plains*

*P3 L6: What do the authors mean by “cloud-induced changes to large-scale forcing”? Does this refer to large-scale circulation and / or synoptic forcing, or something else?*

*P3 L8: “has also” -> “also has”*

*P3 L10: relay -> rely*

*P3 L11: rises -> raises*

*P3 L12: datasets*

*P3 L12: has recently been demonstrated*

*P3 L14: Southern Great Plains*

*P3 L23: in future forecasting systems*

*P3 L28: 30<sup>th</sup>*

*P3 L34: baseline*

*P4 L2: a precipitation -> precipitation*

*P4 L4: the observed aerosol*

**Reply :** Thank you for pointing these out. All issues have been fixed as suggested.

*P4 L7: convective invigoration hypothesis needs a description and / or citation*

**Reply :** We added a reference to the Rosenfeld et al. (2008) paper (p. 4, l. 7)

*P4 L10: investigate whether the*

*P4 L17: Section number missing. (should this be Section 5?)*

*P4 L23-24: The way this sentence is written doesn't quite make sense.*

*P4 L25: do you mean "9 members are selected from", not "selected for"?*

*P4 L32: mesoscale*

*P5 L3-4: repetition of "current study"; you could just say "to our main conclusions".*

*P5 L6: In addition to (delete comma)*

*P5 L11: h a grid -> horizontal grid ?*

*P5 L32: datasets*

*P6 L3: peninsula (remove capital P)*

*P6 L16: have also reported*

*P6 L15: underdispersive over longer*

*P6 L31: smaller if (delete comma)*

*P7 L21, 23, 26: dewpoint*

*P7 L31: ensemble members*

*P9 L20: similar, with a well-mixed*

*P9 L25: temperate -> temperature*

*P10 L3,8: mesoscale*

*P10 L4: "G is very well correlated" – have you actually correlated this (or can you)?*

*P10 L4-5: Figures 4a and 7a are difficult to compare as they are on different pages*

*P10 L14: convergence*

*P10 L18: areal*

*P10 L26: an about 20% -> about a 20%*

*P10 L 10-31: refer to Fig SI 8a*

*P11 L3: largest (smallest) cloud fraction – please refer to Fig 6c.*

*P11 L4: distribution of cloud top heights (Fig. 11b) – you also need to refer to the Fig. showing CTH.*

*P 11 L8: "low", "high": open quotations are the wrong way wrong (LaTeX `` not "?)*

*P 11 L9: "which have a factor of 10 lower and higher aerosol number concentrations, respectively, than the standard profile"*

*P11 L10: altitudes*

*P11 L10: The mean and effective radius – mean what? Mean radius and effective radius? P11 L12: the first section of this study?*

*P11 L13: ensemble members*

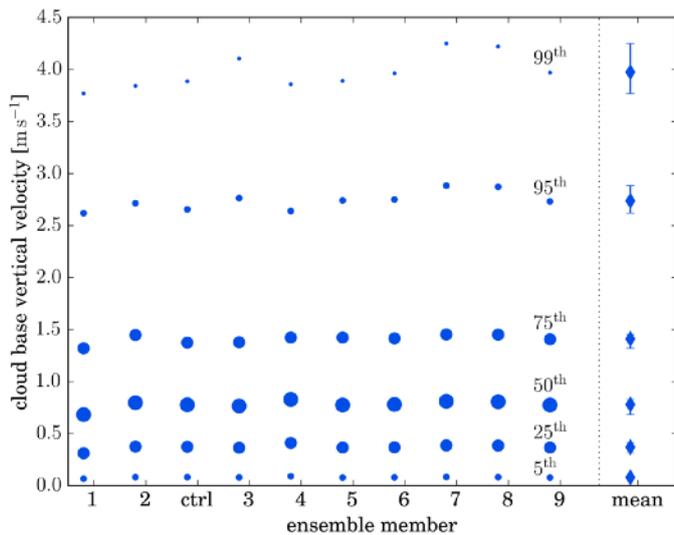
**Reply :** Thank you for pointing these out. All issues have been fixed as suggested and sentences have been reformulated where necessary.

*P11 L17; Figure 5 should be moved, as discussed in the major comments*

**Reply :** The figures have been reordered (s. reply to general comment 3).

*P11 L21: "suggest only minor changes in the cloud-base vertical velocity distribution" – can you plot this distribution? Doesn't this contradict the previous statement made about convergence?*

**Reply :** The cloud-base vertical velocity distribution is shown in Fig. 2 (new SI Fig. 9b). This confirms our hypothesis of small changes in the cloud-base vertical velocity distribution. While these changes are small, the variability in average cloud-base updraft between ensemble members is still larger than the difference in average boundary-layer top moisture content, which is



**Figure 2.** Cloud base vertical velocity distribution for the simulations with the standard aerosol profile considering all grid points at cloud base and a positive vertical velocity for the time period 09 - 19 UTC.

merging with other updraft cores resulting in overall fewer, but larger clouds. Also, energetic constraints potentially limit an increase in overall lifting and cloud fraction. “

*P11 L30: cloud top height increases*

*P12 L2: “ensemble members 1,2,7,8, and 9” (missing space between “and9”)*

**Reply :** Thank you for pointing these out. All issues have been fixed as suggested.

*P12 L2: “ensemble members 1,2,7,8, and 9” – this is also true for member 4*

**Reply :** Sorry again for the confusion with the ensemble member numbers. This list should read 1, 2, 4, 7 and 9 and has been corrected accordingly (p. 11, l. 18-19).

*P12 L3: “does not increase further (members 1 and 2)” – doesn’t member 2 increase?*

**Reply :** The median mean cloud top height does, while the time-average mean cloud top height does not increase from the standard to the high aerosol run. We have clarified this in the text (p. 11, l. 18-19).

*P12 L5: higher than 4.3 km shows only*

*P12 L7: aerosol scenario is likely (remove comma)*

*P12 L8: “maximum” (open quotation incorrect way round)*

*P12 L12: “only a small change”*

*P12 L14: “-4 – 2.5%” - this notation is confusing. Do you mean -4% to -2.5%, or -4% to +2.5%?*

**Reply :** Thank you for pointing these out. All issues have been fixed as suggested.

*P12 L18: I do not understand Figure SI9.*

**Reply :** We apologise for the poor presentation of this figure. We have improved the legend and axis labels in the figure as well as the caption (new SI Fig. 14 d).

*P12 L21: Can you plot delta G and delta L instead of G, L?*

**Reply :** We could plot  $\Delta G$  and  $\Delta L$  instead of  $G$  and  $L$ , which would make it easier to discern aerosol-induced changes. However, this would make the plot less useful to understand the meteorological variability.  $\Delta G$  and  $\Delta L$  were/are shown in previous Fig. 10 (new Fig. 13). We add a plot showing  $\Delta G$  and  $\Delta L$  in the format of previous Fig. 7b in the SI (SI Fig. 14 b and 15 b).

all that we claimed earlier. Note that the discussion of the latter has been removed from the manuscript to meet demands for shortening the text from reviewer 2.

*P11 L24: “the number of cells decreases with increasing background aerosol concentration, but the cell area increases” – this is interesting! Can you explain why this happens?*

**Reply :** We can only speculate about the physical reason for this behaviour, which has been done in the first part of the study. We added a brief summary of the hypothesis the revised version (p. 11, l. 8-11):

“It has been hypothesised in the first part of this study, that the slower conversion of condensate to precipitation in high aerosol conditions allows clouds to grow larger and

merging with other updraft cores resulting in overall fewer, but larger clouds. Also, energetic constraints potentially limit an increase in overall lifting and cloud fraction. “

P12 L29: “seized” -> “sized”

P 12 L32: simulations in the standard

**Reply** : Thank you for pointing these out. All issues have been fixed as suggested.

*P12 L34: in Figure 7c in my printed copy, member 7 also looks like it has no change*

**Reply** : The change in precipitation for ensemble member 7 is very small. The numeric values indicate that accumulated precipitation slightly decreases when aerosol concentrations are increased from the low to the standard scenario. This is consistent with the position of ensemble member 7 off the one-to-one line in Fig. 13. In contrast, ensemble members 6 and 8 fall almost exactly on the one-to-one line in Fig. 13.

*P12 L34: “The latter have a relatively small decrease of PE and comparatively large delta G” – but member 1 also has a decrease in PE and delta G, but a decrease in precip in standard vs low scenarios, and is outside the shading in Figure 10.*

**Reply** : The four ensemble members with the smallest change in PE (from low to standard) are the control, member 6, member 1 and member 8, in this sequence.  $\Delta G$  for ensemble member 1 is the second smallest (joint position with ensemble member 3). In contrast, ensemble member 8, which has a similar change in PE, has the third largest  $\Delta G$ . Also note that changes in PE operate on G, which is much larger in ensemble member 1 than 8, and not on  $\Delta G$ . Hence, a relatively small change in PE is significant for ensemble member 1, while  $\Delta G$  still dominates for ensemble member 8 despite a similar absolute change in PE. The text has been modified to make this clearer (p. 12, l. 18-21):

“Exceptions are the control simulation with a small increase in precipitation and ensemble members 6 and 8 with no change in accumulated surface precipitation. These ensemble members have a relatively small decrease of PE as well as a relatively large  $\Delta G$  and G compared to ensemble members with a similar change in PE (e.g. compare ensemble member 1 and 8).”

*P13 L1: “comparatively large delta G” – this is hard to see from Fig 7a. Can you plot delta G and delta L instead of G and L?*

**Reply** :  $\Delta G$  and  $\Delta L$  are shown in previous Fig. 10 (new Fig. 13). We also added a figure showing  $\Delta G$  and  $\Delta L$  in the format of previous Fig. 7 to the SI. See also reply to comment on P12 L21.

*P13 L10: “Exceptions are ensemble members 3,4 and 5...” – you should point out that the behavior in each of these members is different from each other. For the (a) low to standard and (b) standard to high aerosol scenarios, member 3 has an (a) decrease and (b) increase, member 4 has an (a) decrease and (b) decrease, and member 5 has an (a) increase and (b) decrease.*

**Reply** : The text has been modified to: “Only in ensemble member 3 does the mean precipitation rate not decrease further in the high aerosol scenario, while in ensemble members 4 and 5 the decrease between the standard and the high aerosol scenario is comparable to the decrease between the low and standard scenario.” (p. 12, l. 30-32)

*P14 L11: two sections*

*P14 L20: datasets*

*P14 L29: realisations*

*P15 L4: distribution in different cloud top height classes*

*P 15 L11: Precipitation formation is known...*

*P1 L21: accordingly displays*

*P15 L29: very similarly to*

**Reply** : Thank you for pointing these out. All issues have been fixed as suggested.

*P15 L34-35: “the liquid water path (...) shows little sensitivity to the aerosol scenario” – actually, there is a decrease in LWP (Fig 9b) which is not that much weaker than the increase in CWP (Fig 9a) – this indicates even more strongly than you currently state that the FWP must increase!*

**Reply :** The text has been modified to read: “However, the liquid waterpath (condensate in the cloud and rain category) shows relatively little sensitivity in its median value, while the mean liquid water path generally decreases with increasing aerosol concentrations.” (p. 15, l. 29-31)

*P16 L9: mesoscale*

*P17 L2: “perfect” (open quotations incorrect way round)*

*P17 L3: only slightly different*

*P18 L4: exact number is*

*P18 L7: several 100 ensemble members -> several hundreds of ensemble members*

**Reply :** Thank you for pointing these out. All issues have been fixed as suggested.

*P18 L11: Why is low-high so different from low-standard and standard-high?*

**Reply :** This is mainly because the aerosol perturbation is a factor 10 larger, if the low and high aerosol scenario are considered, than in either low to standard or standard to high scenario. The larger amplitude in the aerosol perturbations results in larger aerosol-induced changes. However, the meteorological variability is in all combinations the same. Accordingly, the “signal-to-noise” ratio is larger in the low-high combination than any other. This has been included in the discussion: “The number of samples required depends on the amplitude of the aerosol perturbation (low and high aerosol scenario versus low/high and standard scenario) as well as the location in the aerosol space (different for increase or decrease relative to the standard aerosol scenario).” (p. 18, l.8-10)

*P18 L13: Accumulated precip stands out here – are you able to explain why? (It’s the only one that needs fewer observations for an increase of number concentration above the standard scenario).*

**Reply :** For all considered variables except P the aerosol-induced changes are smaller (or identical for CDNC) for an increase of aerosol concentrations above the standard aerosol scenario than for decreasing aerosol concentrations. In part 1, it is hypothesised that thermodynamical constraints lead to the saturation of the aerosol effect for high aerosol conditions. In contrast, for accumulated precipitation the aerosol-induced change increases with increasing aerosol concentrations. This is primarily due to PE changes. In part 1, we hypothesise that this change in PE is due to a larger export of condensate into the stratiform region with less active microphysics prompted by the thermodynamic limitations on cloud top height.

The text has been modified to (p. 18, l. 10-15): “In general, more observations are required for an increase of aerosol number concentrations above the standard scenario, which is related to the thermodynamic constraints on aerosol-induced changes in the considered case discussed in Miltenberger et al. (2018). The only exception is accumulated surface precipitation, for which fewer observations are required for an increase above the standard scenario. This reflects the larger aerosol-induced signal in accumulated precipitation for increased compared to decreased aerosol concentrations.”

*P18 L14: “the thermodynamic constraints on aerosol-induces changes...” – constraints for this particular case, or general constraints?*

**Reply :** While there are very likely thermodynamic constraints on aerosol-induced changes in many situations, the conclusions are of course only valid for the investigated case. We have modified the sentence to reflect this (p. 18, l. 12).

*P18 L15: allows us to put the aerosol-induced changes / allows the aerosol-induced changes to be put*

**Reply :** Thank you for pointing this out. Fixed as suggested.

### **Comments on individual figures:**

*Fig 4: Don’t join these points with lines*

Reply : Done.

Fig. 5:

- *Can you comment on why these are so invariant?*

Reply : As discussed on p. 9, l. 15-17 of the manuscript the CDNC is not very variable across the ensemble, because the aerosol concentrations are identical in all members and the cloud base vertical velocity distribution in the different ensemble members is not strongly differing (see also Fig. 2 in this reply).

- *What do the colours represent?*

Reply : Added legend.

- *I think Figure 5 should appear AFTER Figure 6 (given the ordering of discussion in the manuscript)*

Reply : see reply to general comment 3.

Fig. 6: *“cloud fraction is the fraction of the domain for which” (add “the”, remove comma)*

Reply : Done.

Fig 7:

- *Fig. 7a would be clearer if you plotted delta G and delta L instead of absolute values*

Reply : We could plot  $\Delta G$  and  $\Delta L$  instead of  $G$  and  $L$ , which would make it easier to discern aerosol-induced changes. However, this would make the plot less useful to understand the meteorological variability.  $\Delta G$  and  $\Delta L$  were/are shown in previous Fig. 10 (new Fig. 13). We add a plot showing  $\Delta G$  and  $\Delta L$  in the format of previous Fig. 7b in the SI (SI Fig. 14b and 15b).

- *Don't join these points with lines*

- *“the last column in each panel”*

- *Caption: “means” – do you mean ensemble mean, or ensemble means?*

Reply : Done. The caption has been modified to clarify the raised point.

Fig 10:

- *What do the open versus the filled circles represent?*

- *Legend: should “high processing” be “high aerosol”?*

- *Caption: “black symbols” – I can't see any black symbols on Fig 10*

- *Caption: downward / upward triangles: I can't see any of these on Fig. 10*

Reply : Done. The caption has been modified to clarify the raised points.

Fig. SI 2: *“The distributions consider cloudy...”*

Fig. SI 3: *caption L3: “observational data”*

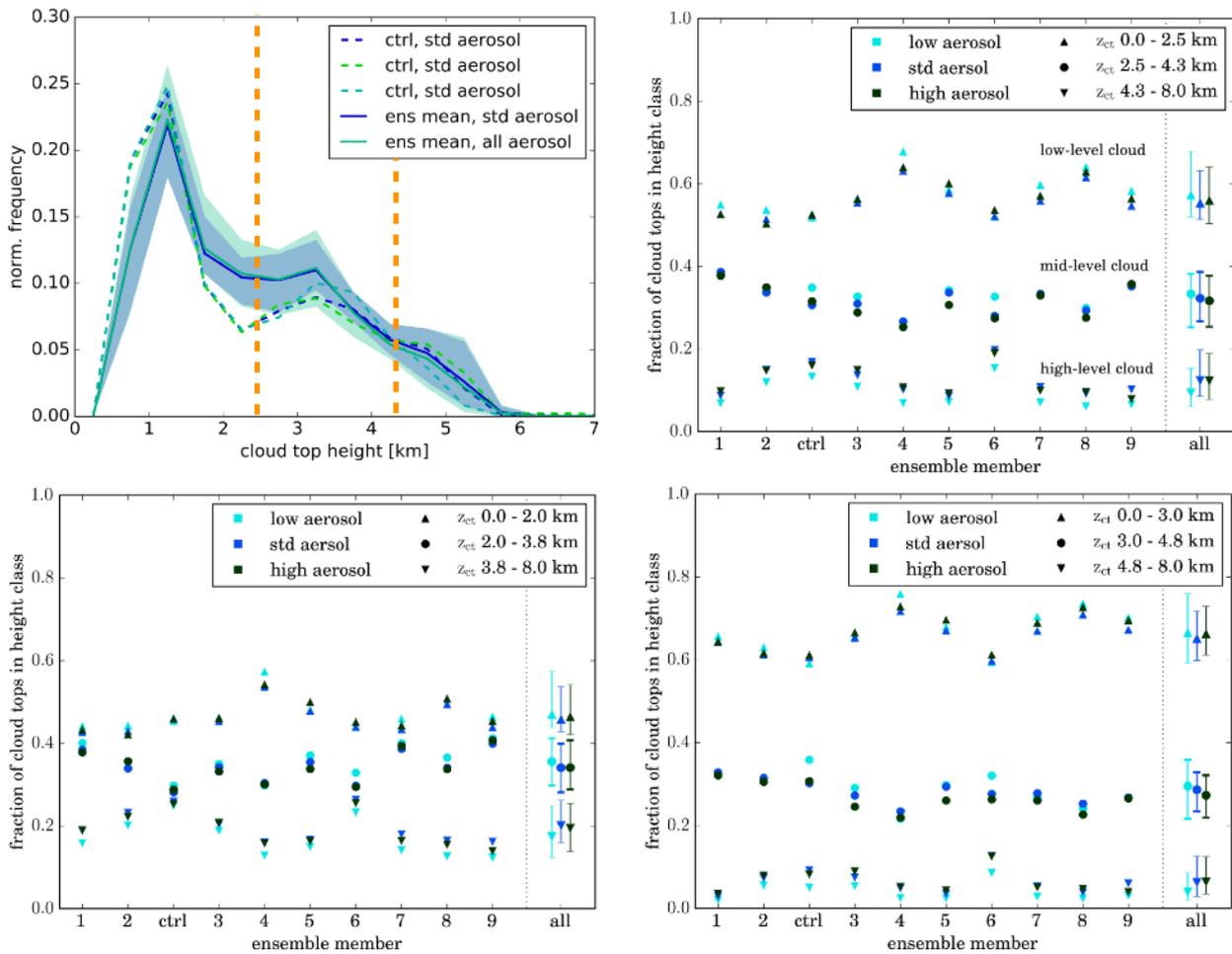
Reply : Done.

Fig. SI 4: *I find the dark blue lines hard to distinguish in my printed copy*

Reply : The colour scheme has been adapted.

Fig. SI 7: *7b: where are the points for ctrl data on the CIN and CAPE Figures?*

Reply : Thanks for spotting this. The data was actually missing from the plot. It has been included in the revised version.



**Figure 3.** Cloud top height distribution (a) and fractions of cloud top in specific altitude ranges (b-d). The thresholds for these ranges are modified by  $\pm 500$  m in (c) and (d) compared to those used in the manuscript and (b).

Fig SI 8:

- How sensitive is this figure to how you choose to define low / medium / deep cloud tops?

**Reply :** The altitude bands were chosen to reflect cloud top ranges with a different response to aerosol perturbations (Fig. 3 a). The overall behaviour of the cloud top height fractions does not differ significantly for variations by  $\pm 500$  m in the thresholds (Fig. 3 b-d), although the absolute fraction values of course do.

- It would be worth placing labels on the Figure with “low”, “med”, “deep” near the relevant set of points, just to make it clearer for the reader.  
 - Again, I don’t think these points should be joined with lines.

**Reply :** Done.

Fig SI 9:

- I found it almost impossible to understand this Figure. Are condensation and deposition shown separately, or combined? What are the symbols? What do IG and IL refer to? Also, as mentioned in the major comments, I don’t think that the points representing the ensemble members should be joined with lines. This is not a continuous dataset. (My printed copy also has different linestyles in the Figure, which are not explained, but I suggest to remove the lines entirely).

**Reply :** We have improved the legend, axis labels and caption.

- Caption: ... and deposition rates

Reply : Done.

**Comments on tables:**

Table 1:

- *Caption: variable (columns)... aerosol scenarios (rows) – aren't these the other way round? (Don't the rows show the variables and the columns the aerosol scenario?)*
- *What do the bold numbers in the table mean?*

Reply : The caption has been modified accordingly.

Table 2:

*Why does the low-high scenario need so few samples compared to low-standard or standard-high?*

Reply : see reply to specific comment on P18 L11

## Reply to comments from reviewer #2

*From my point of view, the main article strength is that it is able to disentangle (to a certain extent) whether the cloud and precipitation effects are due to the meteorological variability or due to aerosol background concentration initial conditions, at least for the case study of mixed-phase convective clouds. Moreover, I think the choice of an increase and decrease in passive aerosol concentration by a factor of 10 is appropriate for the simulations with perturbed aerosol profiles. On the contrary, the main weakness is that there is too much description of the case study until the interesting conclusions are reached. That, at my understanding, makes the reading too much detailed and tedious to follow. Therefore, I would recommend to reduce the number of figures (or move them to the SI) and get to the point on the important findings and conclusions (basically sections 5, 6 and 7) sooner.*

**Reply :** The figures have been reconsidered and there have been significant changes to the distribution of figures between the main text and the SI as well as to the ordering of the figures. All figures are now ordered according to their mentioning in the text. Sections 3 and 4 have been shortened to streamline the text and focus on the key results. However, these sections are still important for the understanding of the results and for providing context on the ensemble performance as well as the magnitude of the meteorological changes, so we believe these sections belong into the manuscript

*Also, I would suggest to the authors to use the significance results presented in Table 1 (unpaired) all over the whole discussion text, since it has important implications whether a result is significant or not. For instance, in section 5.4. Radiation I would add that only OSR results are significant (and only those regarding the comparison between low-standard and low-high aerosol concentrations) and not OLR results, thus the reader do not have to wait until the end of the paper (section 7) to know that some of those differences described before are in fact not significant.*

**Reply :** We have included the statistical analysis in the text in section 5 and 6. These sections contain now references to Table 1.

*Besides, there are many other technical issues and questions which I am listing in the following Specific comments and Typos.*

**Reply :** Thank you very much for pointing these out! We have corrected all the highlighted issues and proof-read the new manuscript more carefully.

### **Specific comments:**

*Section 1: The introduction is appropriate since it is explaining the nowadays main issues, providing the necessary state-of-art, and introducing the contribution of the present study.*

- *Page 2, lines 6-9: however it is true that in the last decades was a large increase in anthropogenic aerosol emissions, I would add that “the emissions have decreased in the last decades (in comparison the 80s-90s maximums) thanks to the introduction of pollution policies in the developed countries in the Northern Hemisphere”.*

**Reply :** Thank you for pointing this out. For sake of brevity, we modified the text to say that anthropogenic aerosol emissions have changed significantly over historic period without specifying any trends (p. 2, l. 6-8): “In recent decades, the modification of cloud properties by aerosols has received particular attention, as anthropogenic aerosol emissions have changed strongly over the historic period.”

*Section 2:*

- *Page 4, lines 20-27: the fact that 9 ensemble members were selected is repeated 3 times in only 8 lines, please consider rewriting the paragraph. Moreover, could you explain how they were chosen among the 33 global ensemble members? (see comment in Fig. 1)*

**Reply :** The selection procedure was already described on p. 4, l. 20-27. We have reformulated this paragraph to make the description clearer and to make the text more concise (p. 4, l. 27-31):

"The selection of the ensemble members for dynamical downscaling is based on the time-series of moisture convergence and moist static energy convergence computed over the regional model domain from the global model fields (Fig. 1). These timeseries are then used to construct a similarity matrix by summing the Euclidean distances of moisture convergence and moist static energy convergence. Using the the algorithm by J. H. Ward (1963) 9 clusters are defined and from each cluster the member closest to the mean cluster time series is chosen for downscaling."

- Page 5, line 11: please clarify the sentence "[...] nested simulations with a h a grid [...]".

Reply : Sentence was reformulated.

- Page 5, lines 13-14: consider the necessity of repeating the same set of references for CASIM since they have been cited already in page 3, lines 33-34.

Reply : We have removed these references here.

- Page 5, lines 17-18: How and why moisture conservation is enforced?

Reply : The methodology for moisture conservation is described in detail in the two cited papers by Aranami et al.. It is beyond the scope of this paper to provide a description of this methodology.

#### Section 4:

- Page 10, lines 9-10: could you clarify why "the surface moisture flux adds some modifications to the boundary layer moisture budget, e.g. ensemble members 3 and 4, and members 7 and 8, respectively", because I am not able to see it in the figures.

Reply : If the surface moisture flux is included the order of ensemble members (if ordering from largest to smallest moisture flux) changes. For example, the lateral PBL moisture flux (red points in Fig. 4 of the manuscript) suggests ensemble member 3 has a larger moisture flux than member 4. However, the surface moisture flux in member 4 is larger than that of member 3 (green symbols in the same figure). Hence, if the total boundary layer moisture flux is considered member 4 has a larger flux than member 3. The text has been modified to (p. 8, l. 15-16):

"The surface moisture flux adds some modifications to the boundary layer moisture budget, e.g. compare total and lateral moisture convergence for ensemble members 3 and 4 and member 7 and 8, respectively."

- Page 10, line 11: from my interpretation of figure 4b, I think ensemble members cntl, 4 and 8 have a particularly large surface sensible heat flux.

Reply : We are really sorry that in several instance throughout the paper the wrong ensemble member numbers were referred to in the text. This is one of the instances. The reviewer is of course right that here it should read control, ensemble member 4 and 8 (instead of members 4 and 7). The section on surface heat fluxes has been removed from the manuscript to shorten the paper as suggested by the reviewer, so these changes are not actually applied in the new manuscript.

- Page 10, lines 18-19: could you add references/citations to the cloud top height threshold based on the condensed water content, and to the cloud fraction based on the condensed water path?

Reply : The chosen threshold for cloud top height (condensed water content larger than  $10^{-6}$  kg kg<sup>-1</sup>) is typically used in modelling studies to reflect detectability in observational studies and avoid issues with very small numeric values in models (e.g. Fridlind et al. 2010). The condensed waterpath threshold ( $10^{-3}$  kg m<sup>-2</sup>) is derived from this value: For a column to be classified as cloudy the minimum condensed water content needs over about 1000 m altitude range, i.e. for about 10 model levels (all considerations based on low-level values of gridspacing and atmospheric density). The condensed water path threshold is below the estimated lower detection limit of microwave satellite instruments (0.02 kg m<sup>-2</sup>, Grosvenor et al. 2017). These references have been added to the revised version of the manuscript (p. 9, l. 20-21).

- Page 10, line 22: Are not ensemble members 2 (instead of 8) and 5 (instead of 4) those with the largest and smallest mean cloud top height, respectively? (for standard aerosol case).

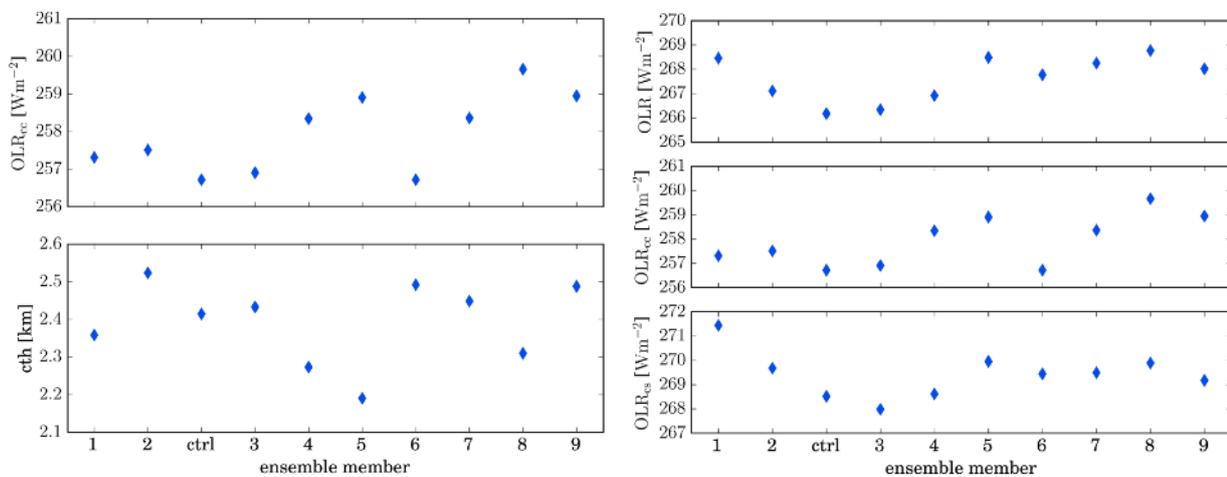
Reply : Yes, of course. The text has been corrected accordingly (p. 9, l. 24).

- Page 10, line 25: based on figure 4a, I would say that ensemble members 2, 4 and 8 are those with high surface moisture fluxes (and not 4 and 7). On the other hand only ensemble members 4 and 8 have larger low-cloud fraction.

Reply : The reviewer is right, this should refer to ensemble members 4 and 8. This section is not part of the revised manuscript anymore.

- Page 10, line 28: based on Fig. 7d, I think that the variation in PE is higher than 5 %.

Reply : We mean the difference between any to PE values does not exceed 0.05. We recognise that the percent notation introduces confusion. To avoid this confusion and address comments from reviewer 1 regarding this sentence, the new text reads: "In contrast, PE does not vary systematically with the large-scale convergence (Fig. 9 b)." (p. 9, l. 30-31)



**Figure 1.** Mean outgoing longwave radiation for cloudy grid points (top left) as well as mean cloud top height (bottom left) and mean cloud fraction (bottom right). The panels on the right show the total mean outgoing longwave (top) and outgoing longwave from cloudy (middle) and clear sky (bottom) gridpoints. Only simulations with the standard aerosol conditions are shown.

- Page 10, line 32: based on Fig. 7d, I think ensemble member 6 has a relatively large PE (instead of 8).

Reply : Yes, this has been corrected (p. 9, l. 34).

- Page 11, line 3: based on Fig. 11a, "the largest (smallest) values occur for ensemble member 1 (8)" (instead of 2 (7)).

Reply : Yes, this has been corrected (p. 10, l. 10).

#### Section 5:

- Page 11, lines 3-4: please, recheck the sentence regarding the relation between outgoing longwave radiation and the cloud top heights since for instance in Fig. 6 it is seen that ensemble members 4 and 5 have similar mean CHT but on the other hand large differences on OLR. How do you explain that?

Reply : Thank you for pointing this out. A more careful analysis shows that there is a relatively good correspondence of the mean OLR from cloud grid points and the mean cloud top height (left panels of Fig. 1). The match is of course not perfect, as OLR is very sensitive to cloud top temperature ( $\sim T^4$ ) and therefore changes in the distribution matter as well, which are not reflected in considering the mean cloud top height. There is also a relatively strong variation in the clear sky

outgoing long wave radiation, which is caused by different surface temperatures, cloud positions, e.g. more or less cloud over the ocean, and water vapour path in the different ensemble members (Fig. 1, right panels). The domain average outgoing longwave radiation is a combination of these two contributions, but is strongly weighted towards the clear sky outgoing longwave due to the overall small cloud fraction ( $\approx 0.15$ , Fig. 2).

We have included Fig. 2 in the SI and altered the text as follows (p. 10, l. 13-15): “While differences in the cloud top height distribution contribute to the variability in outgoing longwave radiation, variations in the clear sky outgoing longwave radiation dominate the overall variability due to the relatively small cloud fraction (SI Fig. 13).“

- *Page 11, line 5: why the section title says “identical meteorological initial and boundary conditions” when in fact here are discussed the differences between ensemble members with different meteorological initial and boundary conditions (as stated in line 13 in the same page)?*

**Reply :** We acknowledge that the section title is misleading. This section discusses aerosol-induced changes in the high and low aerosol scenario relative to the ensemble member with the same meteorological and the standard aerosol scenario. We have renamed the section “Aerosol-induced cloud property changes in different meteorological ensemble members (paired meteorology)“

- *Page 11, lines 24-25: does it mean that there are less clouds but larger?*

**Reply :** Yes exactly. This is included in the revised text (including also changes made in response to reviewer 1): “It has been hypothesised in the first part of this study, that the slower conversion of condensate to precipitation in high aerosol conditions allows clouds to grow larger and merging with other updraft cores resulting in **overall fewer, but larger clouds.**“ (p. 11, l. 7-9)

- *Page 12, line 2: according to SI Fig. 8a, I think ensemble member number 4 is missing from the list of ensemble members where the change in low cloud top fraction is dominant.*

**Reply :** Yes, this has been corrected. (p. 11, l. 15-17): “In the control run and ensemble members 4 and 6, the decrease of the mean cloud top height is due to a reduction in the medium altitude fraction, while in all other members changes in the low cloud top fraction dominate.“

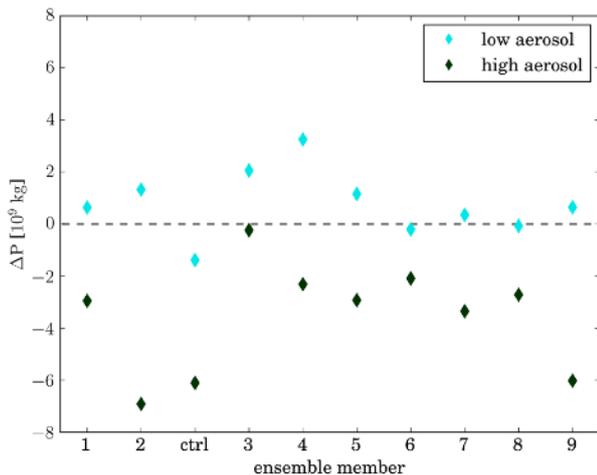
- *Page 12, lines 32-34: Is not ensemble member number 7 also fitting in the exception list? At least this is what I can see from Fig. 7c. For this reason, I recommend changing the graphic color palette or enlarging the figure. Anyway, what do you think is the reason why the control simulation has higher surface precipitation with the low aerosol scenario?*

**Reply :** In ensemble member 7 surface precipitation is decreasing slightly. We agree this is hardly visible from the previous Fig. 7c, but it is clear from its position of the one-to-one line in previous Fig. 10. Previous Fig. 7c has been moved to the appendix and includes now a plot of  $\Delta P$ , which should make this even clearer (new SI Fig. 11).

As for the increase in the surface precipitation in the control simulation: We think this is due to the relative large increase in G (largest of all ensemble members) and the increase in PE. In fact, the control is the only ensemble member, for which PE increases from the low to the standard aerosol scenario. Without a detailed analysis of the thermodynamic, latent heating, and hydrometeor profiles similar to Miltenberger et al. (2018), it is difficult to speculate on the physical processes driving these changes. However, such an analysis is beyond the scope of the paper.

- *Page 12 line 34 and page 13 line 1. Could you check the affirmation again? I do not see the comparatively large condensate gain in Figure 7a.*

**Reply :** This point has not been well made in the previous manuscript, the text has been altered to: “These ensemble members have a relatively small decrease of PE as well as a relatively large  $\Delta G$  and G compared to ensemble members with a similar change in PE (e.g. compare ensemble member 1 and 8).“ (p. 12, l. 19-21)



**Figure 2.** Difference in accumulated precipitation between the low and high aerosol scenario, respectively, and standard scenario.

-Page 13, line 1: please check the following inconsistency: you stated “Accordingly, the precipitation increase for these members...” when you just said in page 12 line 34 that “[...] members 6 and 8 with no change in accumulated surface precipitation.”

**Reply :** If the numeric values are considered all have a slight increase in precipitation (Fig. 2). This figure has been included in the SI. However, changes in member 6 and 8 are very small, so we have changed the text to: “Accordingly, the precipitation response in these cases is either dominated by  $\Delta G$  (control) or  $\Delta G$  and  $\Delta PE$  are of equal importance (member 6 and 8), as also indicated by their position in the shaded area in Fig. 12.” (p. 12, l. 21-23)

- Page 13, line 2: I do not see clear the sentence “For the other members, the change in PE dominates over changes in condensate production” because the change in PE is also large for ensemble members 6-8, and for some of the other members it is actually not that large.

**Reply :** The amplitude of the  $\Delta PE$  and  $\Delta G$  does not in alone indicate, which one is more important, since  $\Delta PE$  operates on  $G$  and not on  $\Delta G$  only. The part of the  $\Delta G - \Delta L$  parameter space, in which changes in  $\Delta G$  dominate is highlighted by the shaded areas in new Fig. 13. The derivation for this is presented in Appendix A of Miltenberger et al. (2018). All points except those representing the control and ensemble members 6 and 8 fall clearly outside this area.

We have clarified the basis for our conclusion in the revised manuscript: “This response is in all ensemble members dominated by PE changes (points outside the shaded area).” (p. 12, l. 25/26)

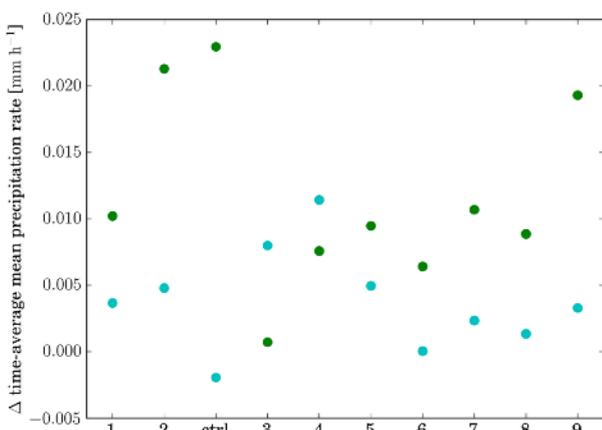
- Page 13, line 3: consider adding “(decrease)” after “precipitation response”.

**Reply :** The sentence has been altered to “If the aerosol concentration is enhanced beyond the standard scenario, the precipitation decreases in all ensemble members (points above the one-to-one line).” (p. 12, l. 24/25)

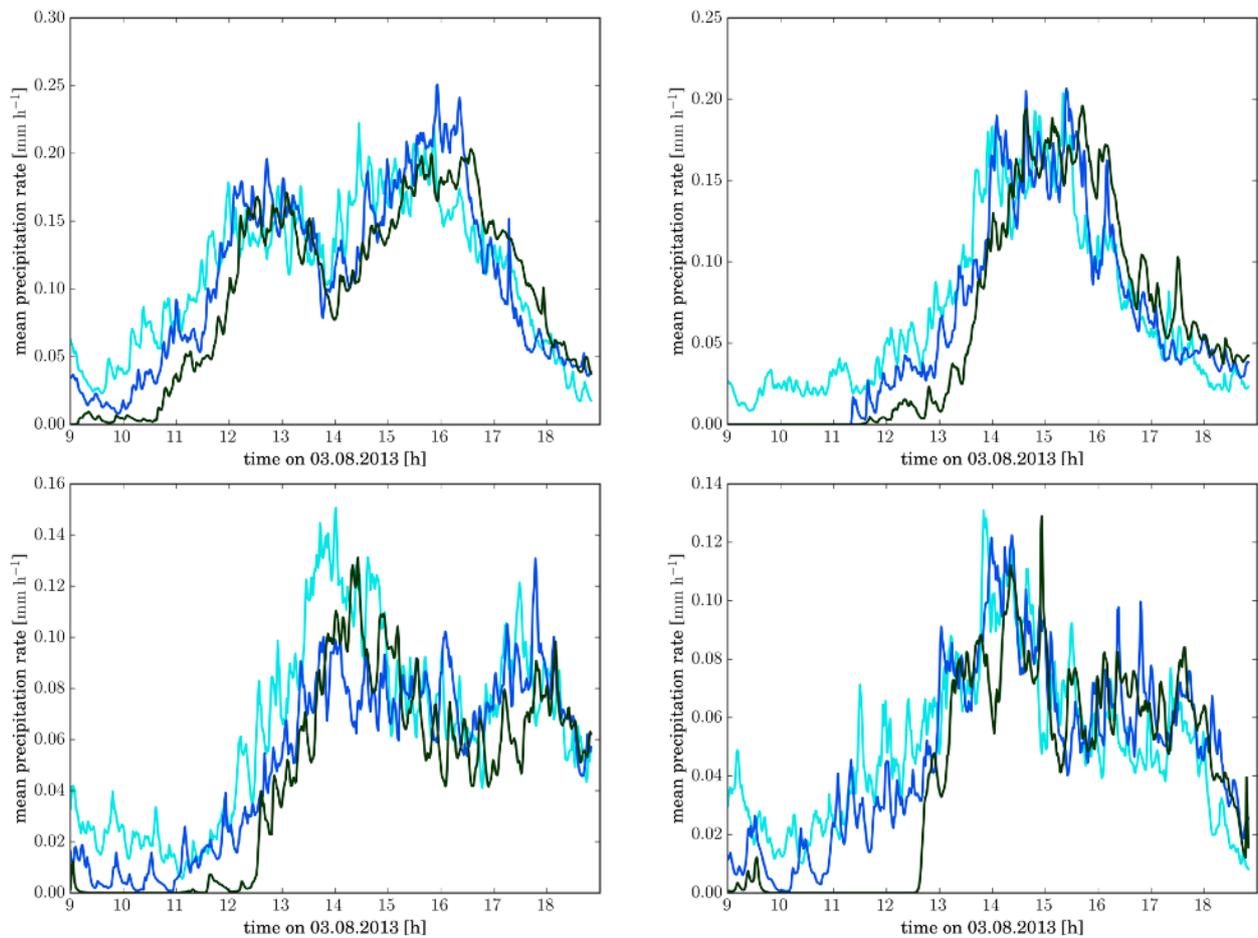
- Page 13, line 10: I am not sure if ensemble member number 5 falls in this exception list, could you please check it again? Additionally, from figure 8a it is remarkable that for some ensemble members (3, 4 and 8) the mean precipitation turns to zero with the highest aerosol concentration scenario, perhaps you would like to highlight it into the discussion.

**Reply :** We think ensemble member 5 belongs in the exception list, as the precipitation change is almost symmetric for an increase or decrease of aerosol concentrations (Fig. 4). The text has been modified to make this clearer: “Only in ensemble member 3 does the mean precipitation rate not decrease further in the high aerosol scenario, while in ensemble members 4 and 5 the decrease between the standard and the high aerosol scenario is comparable to the decrease between the low and standard scenario.” (p. 12, l. 30-31)

Thanks for pointing out that 75<sup>th</sup> percentile of mean accumulated precipitation is turning zero on some of the ensemble members for the high aerosol scenario. Since the distribution shown in these plots represents the temporal variability, this indicates an increased frequency of 10 min intervals without significant precipitation. The



**Figure 3.** Change in the average mean precipitation rate (cyan: low - std aerosol scenario, green: std-high aerosol scenario).



**Figure 4.** Time series of mean precipitation for ensemble member 1 (top left), 3 (top right), 4 (bottom left) and 8 (bottom right) for all three aerosol scenarios (cyan: low, blue: standard, green: high).

complete suppression of precipitation in members 3, 4 and 8 is due to much later onset of surface precipitation (Fig. 4). Note that higher percentiles in most runs are zero in most of the ensemble members. As discussed for the control run in the first part of the presented study, clouds are mostly have lower cloud tops during the morning period and aerosol-induced changes have been found to be larger in this period. As this is extensively discussed in Miltenberger et al. (2018), we do not include this here. In particular, as splitting the response in different time periods would make the article more lengthy, while reviewers have already asked to shorten the manuscript.

- Page 13, lines 12-14: *I do not think “all percentiles up to and including the 75<sup>th</sup> percentile show an increase with the aerosol concentration” for all ensemble members, since in ensemble member 4 the standard aerosol is lower than the other two and in member 7 the high aerosol is lower than the low and standard aerosol concentration.*

**Reply :** The reviewer is right. The text has been modified accordingly (p. 12. l. 33-34): “All percentiles up to and including the 75<sup>th</sup> percentile show an increase from the low to the high aerosol concentration, while the 99<sup>th</sup> percentiles are generally smallest (largest) for the high (standard) aerosol scenario.”

- Page 13, lines 16-17: *consider adding to the discussion the fact that, as other studies have shown, enhanced aerosol scenarios suggest more freezing processes inside clouds and invigoration, provably due to longer cloud lifetime.*

**Reply :** In the first part of the study, we have shown that at least for simulations very similar to the control run enhanced freezing does not contribute to convective invigoration in the investigated clouds. We therefore refrain from citing the convective invigoration hypothesis based on enhanced freezing here.

- Page 13, lines 33 and 34: I agree with the sentence “This change is consistent with the increased CDNC and small impact of the aerosol scenario on the cloud fraction”, however, the change (decrease) in the CF shown in Fig. 6 would cause the opposite effect. How do you explain that? (The same reasoning applies in the sentence in page 15, lines 7-8).

**Reply :** The changes in cloud fraction (between 2 and 9 % depending on ensemble member) are very small compared to the changes in CDNC (factor  $\sim 7$ ). Therefore the latter dominate the change in reflected shortwave radiation. The revised manuscript explains this in more detail (p. 13, l. 20-23):

“This change is consistent with the aerosol-induced change in CDNC and the cloud albedo effect (Twomey, 1977). The co-occurring decrease of cloud fraction under high aerosol conditions (between 2 and 9 % for a factor 10 aerosol change) counteracts the CDNC effect, but the cloud albedo effect dominates due to the large amplitude of the CDNC change (about a factor 7 for a factor 10 aerosol change).“

- Page 13, lines 30-34, and page 14, line 1-2: how do you know that radiative signal presented here is mainly due to CDNC changes and not due to an increase in aerosol scattering (the so-called ‘direct effect’)? Moreover, consider referring here to the ‘indirect effect’ or ‘cloud albedo effect’ and adding “Twomey, 1974” citation reference. Consider also adding “(increase)” after “due to CDNC changes”.

**Reply :** The aerosol direct effect is not included in the model simulations. We have added the suggested reference and altered the text according to the other suggests (see reply to previous comment, p. 13, l. 20-23).

#### Section 6:

- Page 14, line 16: I would rather prefer “changes follow a similar pattern for each meteorological ensemble member” than “changes are similar for each meteorological ensemble member”.

**Reply :** Thank you for this suggestion. The text has been altered accordingly (p. 14, l. 7).

- Page 14, lines 11-19: I miss the authors saying something regarding changes in WP and LWP in this paragraph.

**Reply :** Change in WP and LWP are discussed on p. 15, l. 33ff (old manuscript) and p. 15, l. 28-33 (revised manuscript).

- Page 14, lines 28-29: why the authors only consider the time frame from 9 to 19 UTC while the model was run from 0 to 24h on the 3/8/13? Is it due to meteorological reasons or because of the model spin-up time period? Moreover, why in Fig. 12 is used the time average 10 - 19UTC? Is it a typo?

**Reply :** We only use the model output between 9 and 19 UTC, since this is the time period of main convective activity. Also, the first few hours of the simulation may be affected by model spin-up and later on some high-level cirrus clouds are advected into the domain, which influence the domain integrated cloud variables. We do not want to incorporate these, as they are not related to the convective clouds along the sea-breeze convergence and are mainly dominated by the boundary conditions for the innermost nest.

Fig. 12 actually shows the cloud properties for the 9 to 19 UTC time frame as all other plots. The caption has been corrected accordingly.

- Page 14, line 29: please change “Figs. 5-11” for “Figs. 5-9 and 11”. Moreover, is Fig. 10 done with the data from 0 to 24h or with data from 9 to 19h?

**Reply :** The list of figures has been adapted. All plots contain only data from 9 - 19 UTC. We state this more clearly now in section 2 model and data (p. X5 l. 15/16): “All simulations are run for 24 h. If not stated otherwise, the analysis presented in this paper focusses on the time period between 9 and 19 UTC, i.e. the time period of main convective activity.“

- Page 14, line 30: the variables are not plotted in a box-plot but in an error bar type plot.

**Reply :** Yes, of course. Thank you for pointing this out.

The sentence has been corrected (p. 14, l. 20-24): “If instantaneous realisations of the different (domain-averaged) variables would be considered (box-plots on left side of the figures), the variability would be much larger than suggested by the domain mean plots (right side of the plots).“

- Page 14 line 35 and page 15, line 1: could it be related to a longer cloud lifetime?

**Reply :** We are not sure what the reviewer is referring to. Is it possible that this comment refers to p. 15, l. 35/p. 16, l. 1. If so, a longer lifetime could of course influence the frozen fraction as well, but it is not possible to investigate this effect in our simulations, we do prefer not to comment on any cloud lifetime changes.

**Section 7:**

- Page 16, lines 17-19: I suggest changing the sentence since some of the cloud properties stated here are poorly modified by the aerosol perturbations (e.g. cloud fraction), not modified considering all perturbations (e.g. condensation gain  $G$  is not significant for standard-high comparison), and not modified if unpaired cases are considered (e.g. precipitation rate or cloud fraction are only significant for paired cases)

**Reply :** The sentence has been modified to: “Changes in aerosol concentrations can potentially modify cloud field properties, e.g. cell number and size, cloud depth, cloud fraction, and the domain-wide condensate budget (condensate gain and loss, precipitation rate).“ (p. 16, l. 14)

- Page 16, lines 22-34: Could you explain in more detail how the significance analysis of paired and unpaired was done? How do you pair the ensemble members? I do not really see the advantage of the paired significances with so few ensemble members and it looks to me confusing, if not misleading, since as you already say in page 16, lines 26-27, the statistical analysis is based on a very small sample, which affects the validity of several assumptions. Therefore, I personally prefer the results with unpaired cases because the sample is already too small to be paired and because the results with unpaired cases better reflect the results and error bars (spread) shown in all figures and in particular in Fig 12, even at the expense of having less significant results.

**Reply :** The ensemble members are paired according to their meteorological initial conditions, i.e. are computed by assuming at the cloud properties for the three simulations with different aerosol but identical meteorological initial conditions are not independent. The statistical analysis of paired ensemble members tests the significance of aerosol-induced changes, if aerosol perturbations are considered for identical meteorological conditions but for a number of cases. Hence this reflects the approach taken by most previous modelling studies. We think the difference in the significance between the unpaired and paired ensemble is interesting and may offer an explanation as to why ACI is often found more pronounced in model, case-study based analysis as compared to observational studies.

We agree that the sample size is an issue in the presented work, which is aggregated by the pairing of ensemble members for statistical analysis. However, we decide to still use show this data, as it is consistent with the physical analysis of changes and therefore we think broad picture painted by the statistical analysis is not severely affected by sample size issues.

We modified the introductory text in section 5 (p. 10, l. 22-27):

“In this section, we compare the aerosol signal in the different meteorological ensemble members, i.e. the difference in realisations with different aerosol scenarios but identical aerosol initial and boundary conditions. Thereby we test the robustness of aerosol-induced changes to small perturbations in the meteorological conditions. To quantify the significance of aerosol-induced changes we use a two-sided t-test for ensemble members paired according to meteorological conditions (Table. 1). Using paired ensemble members reflects the interdependence of cloud properties in realisations with different aerosol but identical meteorological initial and boundary conditions.“

And in section 7 (p. 16, l. 20-22):

„First, the idealised situation where the meteorological initial conditions are identical for different aerosol perturbations is assessed by pairing ensemble members according to the meteorological initial conditions. This is equivalent to testing the statistical significance of the differences between realisations with different aerosol scenarios and identical meteorological initial and boundary conditions.

*Page 18, lines 31-32: could you give an example of those “variables closely related to aerosol concentrations” and for those “variables that are linked to aerosol concentrations by a series of complex processes” which apply to the investigated case?*

**Reply :** Yes certainly. New text: “Consistent with previous studies, we find that the aerosol signals in variables closely related to aerosol concentrations, such as for example cloud droplet number concentrations, are easier to retrieve than for variables that are linked to aerosol concentrations by a series of complex processes, such as for example accumulated surface precipitation.” (p. 18, l. 32-34)

*I would suggest changing the section 4, 5 and 6 titles since they are not helpful for understanding the article structure. In my opinion, it would be easier for the reader if they are rewritten somehow in that way:*

- *Section 4: Comparison of the cloud-properties results among 10 different meteorological ensemble members (unperturbed aerosol profiles)*
- *Section 5: Analysis of the results regarding aerosol-induced changes (3 aerosol concentration scenarios) among 10 ensemble members*
- *And section 6: Cloud-adjustments attribution (due to initial meteorological and boundary conditions or aerosol concentration loads)*

**Reply :** We agree that the previous section titles were not ideal. They have been altered to:  
Section 4: Cloud property variability in the meteorological ensemble (standard-aerosol scenario only)  
Section 5: Cloud property changes between ensemble members with different aerosol and identical meteorological initial and boundary conditions  
Section 6: Contribution of aerosol and meteorology perturbations to overall cloud property variability

*References:*

- *Fan et al. 2016: has the DOI link repeated.*
- *Sheffield et al. 2015: has the DOI reference repeated?*
- *Tao et al. 2012: please check if the reference between the DOI and the year should be there.*

**Reply :** Thank you for finding these. The references have been corrected.

*Figures:*

*General comments on the figures:*

*Generally speaking, the way the figures are presented is a bit chaotic. First of all, they are not correctly ordered, and secondly there are too many. I suggest the following improvements:*

- *Re-organize all the figures in order of appearance in the text. For example: figures 6 and 7 are referenced in the text before 4 and 5 have been, and figure 11 before figure 10.*

**Reply :** Figures have been re-order to match their mentioning in the text. Note that some figures have been moved to the SI and the order they are mentioned in the text has also be somewhat changed.

- *Remove linking lines between ensemble members from the following figures: 4b, 7a and b, SI 8a, and SI 9.*

**Reply :** Done.

- Consider changing the color palette for the different aerosol load runs in the following figures, since it is difficult to differentiate them: 4b, 7, SI 8a, and SI 9.

**Reply :** We refrain from changing the colour palette, as it is identical to the one used in the already published first part of the manuscript. However, to make it more easy to distinguish the differences between aerosol scenario, we have supplied difference plots (low / high scenario - std scenario).

*Comments on figures from the main discussion paper:*

- Fig. 1:

- In the caption it is stated that 9 ensemble members were chosen from 33. Could you give more information on that? How they were chosen? Which criteria were used?

**Reply :** The selection of ensemble members for downscaling is described in section 2 (p. 4, l. 27-32). We added a reference to this description in the caption.

- Could you change the x-axis ticks in a way that both graphs (a and b) have the same.

**Reply :** Done.

- Fig. 3: please add the data information in the last column "mean" in c and d, otherwise remove it from the graphs.

**Reply :** Columns have been removed.

- Fig. 4: consider removing it or moving it to the SI since it is only cited once in the paper (and actually only Fig. 4a), the latent heat flux (Fig. 4b) is not used in for the discussion, and the sensible heat flux figure is only used once. Moreover, I think the caption is wrong because it does not match with any of the three graphs and legends.

**Reply :** Figure 4 b and c have been removed from the manuscript, as they are not discussed in the revised text. The caption has been corrected.

- Fig. 5: since it does not show big differences between ensemble members I would suggest moving it to the SI.

**Reply :** We have moved Fig. 5 b to the SI. Fig. 5a remains in the manuscript, as it is the only variable with a clear separation from between aerosol scenarios. This is itself an important point of the manuscript. Also, we think this plots helps the reader to understand the following plots of this type, which are more messy due to the larger meteorology-induced variability.

- Fig. 7: as said before, I recommend changing the color palette of the graphic or enlarging the figure.

**Reply :** We refrain from changing the colour palette, as it is identical to the one used in the already published first part of the manuscript. However, to make it more easy to distinguish the differences between aerosol scenario, we have supplied difference plots (low / high scenario - std scenario).

- Fig. 10: please change the legend with the symbols that appear in the graph (i.e. there are no squares in the graph). Also the caption is wrong since there are no "black symbols", "downward pointing triangles" or "upward pointing triangles".

**Reply :** Legend has been modified and the text in the caption has been corrected.

- Fig.12: this is a really interesting and helping figure. just want to say that adding a legend or a caption explanation regarding the colors, as well as regarding the acronyms used in the x-axis, would improve it. Please, also include if CDNC is at the cloud base or at the top.

**Reply :** Thank you for the positive comments on this figure. The additional information has been included in the caption.

Comments on the SI figures:

- *SI Fig. 1: I would include this figure in the main paper (not in the supplement) since it helps the reader quickly identifying the region on the model simulations have been done and how they look like.*

**Reply :** This figure is now in the main paper (new Fig. 2).

- *SI Fig. 4: it needs some improvements since it is not intuitive. I would suggest plotting in different colors the temperature and the dew temperature profiles, from both model and observational data.*

**Reply :** The suggested improvements have been made.

- *SI Fig 6 and 7 captions: "The box plots represent the temporal variability of each variable" should read "The box plots represent the temporal variability of the variable" or "The box plots represent the temporal variability of the variable for each ensemble member".*

**Reply :** Done.

- *SI Fig. 6: why ensemble 4 with low aerosol look so different from the others? Is there any apparent reason?*

**Reply :** Thanks for spotting this. There was an error in the plotting routine affecting in particular this specific ensemble member. The plot has been corrected.

- *SI Fig. 7: CAPE for cntl simulation is missing. And please, either add values for the last column or remove "mean" from the graphic.*

**Reply :** Thanks for spotting this. The plot has been corrected.

- *SI Fig. 9: I suggest adding into the caption the description of IG and IL (from the legend) as well as "[...] condensation (C) and deposition rate (D) [...]".*

**Reply :** The caption as well as the plot itself have been improved and the inconsistent/unclear notation have been cleaned up.

Tables:

- *Table 1: as mentioned before, I would remove the unpaired results.*

**Reply :** see reply to comment on Page 16, lines 22-34

Typos:

*Page 4, line 17: "section ??" should read "section 5".*

*Page 7, line 24 and line 28: "15 UTC" should read "15.20 UTC".*

*Page 10, line 17: "coherent areas" or "consecutive areas"?*

*Page 10, line 18: "arial fraction" should read "areal fraction".*

*Page 11, line 9: the word "and" is missing between "lower" and "respectively".*

*Page 11, line 13: the word "part" is missing between "first" and "of this study".*

*Page 12, line 5: remove "is" from the sentence.*

*Page 13, line 10: remove "also" from the sentence.*

*Page 13, line 32: add the word "by" after "cloud top and".*

*Page 13, line 34: change "increasing" by "increased".*

*Page 14, line 11: "section" should read "sections".*

*Page 18, line 21: "aerosol-induce" should read "aerosol-induced".*

**Reply :** Thank you very much for pointing these out! We corrected all issues as suggested.

# Aerosol-cloud interactions in mixed-phase convective clouds. Part 2: Meteorological ensemble.

Annette K. Miltenberger<sup>1</sup>, Paul R. Field<sup>1,2</sup>, Adrian A. Hill<sup>2</sup>, Ben J. Shipway<sup>2</sup>, and Jonathan M. Wilkinson<sup>2</sup>

<sup>1</sup>Institute of Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, United Kingdom

<sup>2</sup>Met Office, Exeter, United Kingdom

*Correspondence to:* Annette K. Miltenberger (a.miltenberger@leeds.ac.uk)

**Abstract.** The relative contribution of variations in meteorological and aerosol initial and boundary conditions to the variability in modelled cloud properties are investigated with a high-resolution ensemble (30 members). In the investigated case, moderately deep convection develops along sea-breeze convergence zones over the southwestern peninsula of the UK. A detailed analysis of the mechanism of aerosol-cloud interactions in this case has been presented in the first part of this study (Miltenberger et al., 2018).

The meteorological ensemble (10 members) varies by about a factor of 2 in boundary layer moisture convergence, surface precipitation, and cloud fraction, while aerosol number concentrations are varied by a factor of 100 between the three considered aerosol scenarios. If ensemble members are paired according to the meteorological initial and boundary conditions, aerosol-induced changes are consistent across the ensemble. Aerosol-induced changes in CDNC, cloud fraction, cell number and size, outgoing shortwave radiation, instantaneous and mean precipitation rates, and precipitation efficiency are statistically significant at the 5 % level, but changes in cloud top height or condensate gain are not. In contrast, if ensemble members are not paired according to meteorological conditions, aerosol-induced changes are statistically significant only for CDNC, cell number and size, outgoing shortwave radiation, and precipitation efficiency. The significance of aerosol-induced changes depends on the aerosol scenarios compared, i.e. an increase or decrease relative to the standard scenario.

A simple statistical analysis of the results suggests that a large number of realisations (typically > 100) of meteorological conditions within the uncertainty of a single day are required for retrieving robust aerosol signals in most cloud properties. Only for CDNC and shortwave radiation small samples are sufficient.

While the results are strictly only valid for the investigated case, the presented evidence combined with previous studies highlights the necessity for careful consideration of intrinsic predictability, meteorological conditions, and co-variability between aerosol and meteorological conditions in observational or modelling studies on aerosol indirect effects.

*Copyright statement.* The works published in this journal are distributed under the Creative Commons Attribution 3.0 License. This licence does not affect the Crown copyright work, which is reusable under the Open Government Licence (OGL). The Creative Commons Attribution 3.0 License and the OGL are interoperable and do not conflict with, reduce, or limit each other.

## 1 Introduction

Clouds and precipitation are an integral part of the atmospheric system relevant for weather and climate. Considerable uncertainty remains in our understanding and modelling of clouds and their interaction with other parts of the climate system. The main issues are an incomplete physical understanding of cloud microphysical processes, a lack of quantitative formulations representing microphysical processes on model grid scales which are typically several orders of magnitude larger than the process scales, and the many non-linear interactions between different components of the system. In recent decades, the modification of cloud properties by aerosols has received particular attention, as anthropogenic aerosol emissions have changed strongly over the historic period.

Many modelling studies have investigated the impacts of an aerosol change on either isolated clouds or larger cloud fields, but found different responses of the studied clouds depending on environmental conditions, model formulations, duration of simulations, and domain size (recent reviews by Tao et al., 2012; Altaratz et al., 2014; Rosenfeld et al., 2014; Fan et al., 2016). Recent studies have highlighted that it is necessary to simulate entire cloud fields over long periods in order to quantify a climate relevant aerosol signal (e.g. Grabowski, 2006; van den Heever et al., 2011; Seifert et al., 2012) due to interactions between clouds and their thermodynamic environment. These interactions can at least partly compensate the large changes simulated for individual clouds (e.g. Lee, 2012; Seifert et al., 2012). In a case-study of tropical deep convection, Lee (2012) found that locally invigorated convection in polluted conditions induces stronger large-scale subsidence resulting in an overall suppression of precipitation on a cloud-system scale. Seifert et al. (2012) demonstrated with simulations extending over three summer seasons that aerosol perturbations can produce large local changes in precipitation, while not significantly changing the mean precipitation.

The highly non-linear nature of convective cloud dynamics and microphysics calls for the use of large ensembles due to a potentially rapid growth of small perturbations to the system (e.g. Wang et al., 2012). While the importance of predictability limits has been acknowledged in weather forecasting, its implications for the evaluation of cloud microphysics parameterisations or the quantification of aerosol-cloud interactions has only been acknowledged in a few studies (Grabowski et al., 1999; Khairoutdinov and Randall, 2003; Zeng et al., 2008; Morrison and Grabowski, 2011; Morrison, 2012). To our knowledge, the first study to highlight the importance of intrinsic predictability for cloud microphysics evaluation and aerosol-cloud interactions is Grabowski et al. (1999). Along with changes to various parameters in the cloud microphysics, cloud-radiation interaction, and CCN number concentrations, they applied random perturbations to the large-scale forcing, the surface fluxes, and nudging timescale in their 2D simulations of deep tropical convection. Khairoutdinov and Randall (2003) investigated the sensitivity of convective clouds over the ARM Southern Great Plains site to the choice of cloud microphysical parameterisations and perturbed initial conditions. While they found the mean hydrometeor profile and cloud fraction to be strongly dependent on the chosen cloud microphysical scheme, the variability of cloud fraction, precipitable water, and surface precipitation induced by different microphysical schemes was similar to those resulting from perturbed initial conditions. In a similar modelling framework to Grabowski et al. (1999), Morrison and Grabowski (2011) also applied random perturbations to simulations of deep tropical convection based on the Tropical Warm Pool International Cloud Experiment. They found a large variability in

top of atmosphere radiative fluxes between ensemble members generated by modest perturbations to the boundary layer temperature structure. In this case, therefore a large ensemble with 240 members was required to retrieve a robust aerosol-induced signal in the top of atmosphere radiative fluxes. In their ensemble, surface precipitation was insensitive to aerosol changes. The simulations in these studies use large-scale forcing time series, which provide realistic time variations in forcing, but do not allow for a two-way interaction of the clouds with the large-scale forcing. While this avoids the even larger complexity of cloud-induced changes to large-scale circulation, it is ultimately necessary to include this interaction in order to quantify the impact of uncertainties in cloud microphysical processes or changes in aerosol concentration on the atmospheric system.

The relative importance of meteorological and aerosol conditions for cloud properties also has implications for obtaining observational evidence of aerosol-cloud interactions. Many observational studies of aerosol-induced changes in cloud properties need to rely on correlations between bulk parameters (e.g. Devasthale et al., 2005; Koren et al., 2010; Gryspeerdt et al., 2014), which raises the question of co-variability and coincidence (e.g. Stevens and Feingold, 2009; Gryspeerdt et al., 2014). The importance of cloud dynamics in observational datasets has recently been demonstrated by Sena et al. (2016). The study analysed the correlation of aerosol, cloud dynamics, and a range of cloud properties for shallow warm-phase clouds over the ARM Southern Great Plains site. They showed that the variability of cloud radiative properties was dominated by cloud dynamics rather than cloud microphysical properties.

One approach to investigate the role of intrinsic predictability and the relative importance of aerosol and meteorological variability is the use of convection-permitting ensemble systems. Ensemble forecasting is now an important component of operational forecasting and is increasingly used at convection-permitting or even higher spatial resolutions (e.g. Bowler et al., 2008; Marsigli et al., 2014; Beck et al., 2016). The use of convection-permitting ensemble forecasts provides a means for assessing the magnitude of aerosol-induced changes in the context of variations in the cloud and precipitation evolution due to perturbations in the meteorological conditions, which are consistent with the uncertainty in available meteorological observation. Besides offering insight into the questions of robustness and observability of aerosol-induced changes, the ensemble approach explores whether perturbations of the aerosol environment should be included in future forecasting systems for quantitative precipitation forecasts.

In the present study, we investigate the robustness and relative importance of aerosol-induced changes in mixed-phase, sea-breeze related convective cloud in high-resolution ensemble simulations with perturbed meteorological and aerosol initial and lateral boundary conditions. The case was selected from the CONvective Precipitation Experiment (COPE) that was conducted over the southwestern peninsula of the UK in 2013 (Blyth et al., 2015; Leon et al., 2016). On the selected day (3<sup>rd</sup> August 2013) deep convective clouds with maximum cloud top heights of about 5 km developed in the late morning along converging sea-breeze fronts. The line of convective clouds remained roughly stationary along the main axis of the peninsula until the late afternoon. Generally, new cells formed at the southwestern tip of the peninsula and merged into larger cloud clusters while propagating northeastwards along the line. Simulations of this case were conducted with the Unified Model (UM) at a spatial resolution of 250 m using the newly developed Cloud-AeroSol Interacting Microphysics Module CASIM (Shipway and Hill, 2012; Hill et al., 2015; Grosvenor et al., 2017; Miltenberger et al., 2018). The comparison of the baseline simulation with observational data and the sensitivity of cloud properties to aerosol perturbations was presented in the first part of this

study (Miltenberger et al., 2018) and is briefly summarised here: Increasing aerosol concentrations suppress precipitation in the morning. With progressing organisation of the clouds along the sea-breeze fronts, the response transitions into precipitation enhancement. In the early phase, precipitation decreases continuously with aerosol concentration (0.1 to 30 times observed value), while in the afternoon the largest accumulated precipitation occurs with the observed aerosol profile. Limitations on cloud deepening from a mid-tropospheric stable layer were hypothesised to inhibit a further increase of precipitation for aerosol number concentrations larger than the observed values. Vertical velocities increase in the convective core regions with aerosol concentrations. However, contrary to the convective invigoration hypothesis (e.g. Rosenfeld et al., 2008) changes in latent heat release are dominated by changes in the warm-phase part of the cloud with very small changes above the 0 °C line. It was hypothesised that accompanying changes in the cloud field structure (fewer, larger cells with increasing aerosol) were important for the changes in latent heat release from condensation.

In this paper, we extend the analysis of Miltenberger et al. (2018) by including simulations with perturbed meteorological conditions in the analysis. With the combined perturbed meteorology and aerosol initial condition ensemble we investigate whether the aerosol-induced changes are (i) robust to and (ii) significant relative to small changes in the meteorological initial conditions. The paper is structured as follows: Section 2 provides details on the model set-up and observational data used in this study. The ensemble simulations are compared to observational data in section 3. In section 4, we discuss the variability of cloud properties in the perturbed meteorology only ensemble, while the impact of aerosol perturbations on clouds and precipitation for individual ensembles members is assessed in section 5. Finally, the results from the full ensemble, i.e. including perturbations to meteorology and aerosols, are presented in section 6. The findings are summarised in section 7.

## 2 Model and data

The initial condition ensemble discussed in this paper is constructed by downscaling **selected members** from the operational global ensemble system of the Met Office (MOGREPS, Bowler et al. (2008)) over the southwestern peninsula of the UK. The global model ensemble is recomputed from the Met Office operational analysis and initial condition perturbation for 1800 UTC on 02. 08. 2013. The global model version and set-up used for the operational forecast in 2013 are employed for the re-run (UM, vn8.2, PS31 configuration, N400 resolution, i.e.  $\approx 33$  km in mid-latitudes). **This includes stochastic physics as described in Bowler et al. (2009). The control run (no initial condition perturbations applied) and 9 global ensemble members provide the initial and boundary conditions for the higher resolution regional simulations. The control run is included in the term “ensemble members” if not stated differently. The selection of the ensemble members for dynamical downscaling is based on the time-series of moisture convergence and moist static energy convergence computed over the regional model domain from the global model fields (Fig. 1). These time-series are then used to construct a similarity matrix by summing the Euclidean distances of moisture convergence and moist static energy convergence. Using the algorithm by Ward (1963) 9 clusters are defined and from each cluster the member closest to the mean cluster time series is chosen for downscaling.** Note, that while this procedure provides a sampling of different time series, it does not necessarily retain the statistical properties of the global ensemble. It is known that convection-permitting ensembles constructed by downscaling global ensemble members

do not represent the mesoscale error characteristic correctly (e.g. Saito et al., 2011; Berner et al., 2011). As a result convection-permitting ensemble forecasts are often underdispersive (e.g. Romine et al., 2014; Schwartz et al., 2014). Our high-resolution ensemble will hence represent some unknown fraction of the true meteorological uncertainty for the studied day. Most likely the meteorological uncertainty is underestimated in the current study. Although the ensemble selection and initialisation of the ensemble should be improved in future studies, we do not think that this is a strong caveat to our main conclusions.

Regional simulations with a grid spacing of 1 km (500 by 500 grid points) are started at 0000 UTC on 03. 08. 2013 from the 10 selected global model runs. These simulations provide the initial and boundary conditions for simulations in a second set of nested simulations with a horizontal grid spacing of 250 m (900 by 600 grid points). For the regional simulations, we use the UM version 10.3 (GA6 configuration, Walters et al. (2017)) with the CASIM module. In contrast, to the global ensemble, we do not use the stochastic physics module for the regional ensemble, as we aim to investigate the role of initial condition uncertainty. The model set-up for the regional simulations is identical to the set-up described in Miltenberger et al. (2018). The control simulations are identical to the simulations used in Miltenberger et al. (2018) with the only difference that the simulations discussed here use the cloud droplet number predicted by CASIM instead of a prescribed value for the computation of the radiative fluxes. In all regional simulations, moisture conservation is enforced according to Aranami et al. (2014, 15 2015). All simulations are run for 24 h. If not stated otherwise, the analysis presented in this paper focusses on the time period between 0900 and 1900 UTC, i.e. the time period of main convective activity. Also note that ensemble members have been sorted according to the large-scale moisture convergence computed from the fluxes at the domain boundary: ensemble member 1 has the largest large-scale moisture convergence and ensemble member 9 the smallest.

Cloud microphysical processes are parameterised within the CASIM module which in this study is configured as a double-moment microphysics scheme with five different hydrometeor categories. The CASIM module can represent the interactions between aerosol fields and cloud microphysical properties. For the ensemble simulations, we use the so-called “passive aerosols” mode: aerosol fields are used for droplet activation and ice nucleation, but are not altered by cloud microphysical processes. The impact of this choice on the representation of aerosol-cloud interactions is discussed in Miltenberger et al. (2018). Aerosol initial and lateral boundary conditions are derived from aircraft data as described in Miltenberger et al. (2018). In the following, simulations with the aerosol profile derived from observations are referred to as “standard aerosol” simulations. Additional simulations of each ensemble member are performed with perturbed aerosol profiles, for which aerosol number densities and mass mixing ratio are multiplied at all altitudes by a factor of 10 (“high aerosol”) and 0.1 (“low aerosol”), respectively. Hence, the mean aerosol radius is retained in the perturbed profiles. Accordingly, the entire ensemble with perturbed meteorological and aerosol initial conditions has 30 members in total.

For the evaluation of the ensemble with the standard aerosol profile, we use the same set of observations as in the first part of this study. These include radiosonde and aircraft data from the COPE field campaign and data from the operational radar network. Details about these datasets can be found in Miltenberger et al. (2018).

### 3 Evaluation of ensemble simulations

#### 3.1 Radar reflectivity and surface precipitation

In all ensemble simulations, a convergence line develops roughly over the centre of the peninsula in the early afternoon (Fig. 2). The convective clouds are associated with convergence zones along sea-breeze fronts. However, the members vary in the amount of clouds and there are some differences in the location and the orientation of the main cloud line. These differences are not specific to the time instance shown in Fig. 2, but persist throughout the simulations. Differences between meteorological ensemble members are further discussed in section 4, while we focus here on the comparison of the ensemble to the observational data.

Consistent with the similar meteorological evolution of the ensemble members, the domain-average precipitation has a similar temporal evolution with increasing values during the morning hours and maximum values between 1300 UTC and 1600 UTC (Fig. 3). Domain-average precipitation rates from the control forecast (dashed blue line) are mostly within the spread of the ensemble members (blue shading), although the ensemble-mean domain-average precipitation rate is about a factor of 2 smaller than the control during the period of main convective activity (1200 -1700 UTC). The spread of the ensemble including aerosol perturbations (cyan shading) is not much larger than the ensemble spread based on perturbed meteorological conditions alone, particularly after about 1430 UTC. The ensemble mean is almost identical for both ensembles. The domain-average precipitation rates derived from radar (Radarnet IV, Harrison et al. (2009); MetOffice (2003)) fall mostly outside the spread of the ensemble. This indicates that either the ensemble is underdispersive or that there are issues with the radar derived surface precipitation. While the model derived surface precipitation is the sedimentation flux at the surface, the radar derived surface precipitation is computed from the low-level radar reflectivity according to Harrison et al. (2009). Accordingly the modelled and radar-derived surface precipitation products involve different assumptions, e.g. on sub-cloud evaporation, which has been shown to affect radar-detrieved surface precipitation rates (e.g. Li and Srivastava, 2001). Nevertheless, also previous evaluation studies of convection-permitting ensemble simulations have reported precipitation forecasts to be underdispersive over longer evaluation periods (e.g. Romine et al., 2014; Schwartz et al., 2014), as not all sources of uncertainty are taken into account. For example, structural or parametric uncertainty in the model physics is not considered and perturbations to the initial and boundary conditions may not be fully representative of the true uncertainty. The incorporation of perturbations to the aerosol initial and boundary conditions does not improve the comparison. However, the underdispersivity of the ensemble does not strongly impact the major conclusions of our study, as we interpret the meteorological uncertainty as a lower limit of meteorological variability in the discussion (section 7).

The underestimation of domain-average precipitation in the ensemble is, similarly to the results in Miltenberger et al. (2018), caused by a combination of a too small precipitating area fraction and too low occurrence frequency of medium precipitation rates in precipitating areas (not shown). While the observed in-cloud precipitation rate distribution is outside of the ensemble spread (SI Fig. 1), the simulated distributions of column-maximum radar reflectivity and of low-level (750 m agl) radar reflectivity agree well with the observed distribution (SI Fig. 2). In contrast to the domain-average time-series, the ensemble spread in the radar reflectivity distributions increases significantly if aerosol perturbations are considered in addition to meteorological

initial condition perturbations (SI Fig. 2). However, the ensemble mean distributions are almost identical for members with and without aerosol perturbations.

5 The 3D radar composite available for this case provides information about the vertical structure of the clouds. Here we compare the simulated and observed altitude of the highest occurrence of a radar reflectivity larger than 18 dBZ, which is frequently used in radar products to measure cloud depth (e.g. Lakshmanan et al., 2013; Scovell and al Sakka, 2016). The ensemble mean is closer to the observed evolution than the control run (within 200 m, SI Fig. 3). The inclusion of perturbed aerosol initial and boundary conditions has only a small impact on the ensemble mean height of the 18 dBZ contour (maximum difference:  $\pm 100$  m). Also, for other reflectivity thresholds (5 – 25 dBZ), the observed mean height is within the ensemble spread and the difference to the ensemble mean is generally smaller than 500 m (not shown).

### 10 3.2 Radiosonde data

Thermodynamic profiles are available at two hourly intervals from radiosondes released at Davidstow (50.64 °N, 4.61 °W). These profiles are compared to the thermodynamic structure of the closest model grid column from the simulation with the standard aerosol profile (SI Fig. 4). The overall (out of cloud) structure of the temperature and dewpoint temperature profiles are similar to the observed structure for all times and ensemble members. The observed temperature profile generally falls within the ensemble spread except between 550 – 400 hPa. Also, the observed dewpoint temperature profile is generally contained within the ensemble spread, with the exception of a lower observed humidity below 900 hPa at 1520 UTC. All ensemble members have a stable layer between 5 – 6 km altitude, which is an important factor for the cloud top height distribution (Miltenberger et al., 2018). Ensemble members differ mainly in the humidity above 600 hPa, with the altitude of the driest point in this layer varying by about 100 hPa.

20 The height of the 0 ° level and the lifting condensation level corroborate the good agreement between observed and modelled profiles for the duration of the simulation and all aerosol scenarios: maximum deviations are about 300 m for the 0 ° level height and 400 m in the lifting condensation level (SI Fig. 5). While the observed lifting condensation level falls within the ensemble spread (except at 1520 UTC), the observed 0 ° level is generally outside the ensemble spread (except at 1350 UTC, but radiosonde passed through clouds).

25

Overall the ensemble reflects the cloud and precipitation evolution, as well as thermodynamic structure indicated by observational data. However, the ensemble does not improve on the performance of the control run. Overall the ensemble performance provides confidence that the most important physical mechanisms are well enough represented to conduct aerosol perturbation experiments.

## 4 Cloud property variability in the meteorological ensemble (standard-aerosol scenario only)

Given the overall similar meteorological situation in the ensemble members, i.e. a line of convective clouds forming along sea-breeze convergence zones, the main impact of the perturbed meteorological initial conditions should be (i) perturbations to vertical lifting and hence condensation, and (ii) the vertical cloud structure by modifications to the vertical wind shear and the thermodynamic profiles. The discussion in this section focusses on the meteorological ensemble with the standard aerosol profile. Differences in the large-scale moisture convergence, upstream thermodynamic profiles and sea-breeze strength are discussed in section 4.1. The resulting variation in cloud-properties is described in section 4.2.

### 4.1 Large-scale convergence and condensate formation

The meteorological ensemble members have been selected on the basis of the moisture and moist static energy convergence (section 2), as the large-scale moisture convergence should influence the amount of lifting and hence condensate formation. The large-scale convergence is diagnosed from the moisture fluxes at the domain boundaries. Here, we focus on the boundary-layer moisture convergence, which is most relevant for the cloud-base mass flux. As expected ensemble members have very different boundary layer moisture convergence (Fig. 4 a, red symbols). Consistent with this variability, the condensate gain  $G$ , i.e. the domain-integrated condensation and deposition rate, varies across ensemble members with decreasing values for members with smaller large-scale boundary-layer moisture convergence (Fig. 4 b). The correspondence between moisture convergence and  $G$  is further improved if the net moisture flux at the top of the boundary layer is considered (Fig. 4a, blue symbols), which is diagnosed from the sum of the moisture flux at the domain boundaries (red symbols) and the surface moisture flux (green symbols). The surface moisture flux adds some modifications to the boundary layer moisture budget, e.g. compare total and lateral moisture convergence for ensemble members 3 and 4 and member 7 and 8, respectively.

In addition to the large-scale moisture convergence, differences in meteorological initial and boundary conditions could also result in different local convergence patterns, i.e. sea-breeze strength. Differences in sea-breeze strength between ensemble members can contribute to the variability of  $G$  across the meteorological ensemble. The main controlling factors for sea-breeze strength are the temperature difference between sea and land, the large-scale wind direction relative to the coastline, and the background wind speed (e.g. Estoque, 1961; Miller et al., 2003). Golding et al. (2005) and Warren et al. (2014) have demonstrated the importance of differential heating of the land surface and the interaction with the background wind field for stationary convergence lines and associated convective activity over the southwest peninsula of the UK. The profiles of the wind components, temperature and specific humidity are shown in SI Fig. 6, the variation in the land-sea temperature gradient in SI Fig. 7 a, and the “low-level” convergence, i.e. the integrated convergence of the 10 m windspeed over the peninsula, as indicator of the sea-breeze strength in SI Fig. 7 b.

The temperature difference between land and sea increases from 0.9 – 1.4 K in the morning to 1.8 – 2.0 K by noon. Only in ensemble members 1 and 2 the temperature difference remains smaller than 1.5 K (SI Fig. 7 a). These members have a higher cloud fraction in the morning (not shown), which is likely related to a relatively large large-scale moisture convergence. The higher cloud fraction reduces radiative heating of the land surface explaining the smaller peak land-sea temperature difference.

The wind speed in the boundary layer varies between about  $8 \text{ ms}^{-1}$  and  $11 \text{ ms}^{-1}$ , and increases to values of  $13 - 18 \text{ ms}^{-1}$  at 4 km altitude (SI Fig. 6 a). The wind direction is generally from the south-west with a variability of about  $10^\circ$  and a shift towards a more easterly direction at higher altitudes (SI Fig. 6 b).

The low-level convergence consistently increases towards noon as is expected for sea-breeze systems (SI Fig. 7 b). Overall there are only small differences in the time-integrated low-level convergence between ensemble members. This suggests that neither the variability in the land-sea temperature difference ( $\approx 0.5 \text{ K}$ ) nor the variability in the low-level wind speed ( $\approx 3 \text{ ms}^{-1}$ ) and direction ( $\approx 10^\circ$ ) has a significant impact on the sea-breeze strength.

Other variables in the initial conditions important for cloud and precipitation formation are the temperature and moisture profiles (SI Fig. 6 c and d). The temperature structure in all ensemble members is very similar, with a well-mixed boundary layer below  $800 \pm 200 \text{ m}$ , an almost moist-adiabatic temperature gradient up to 500 hPa, and a layer of almost constant temperature between 500 hPa and 450 hPa. As a result of the small variation in the temperature profile, the average and maximum CAPE values are similar for all ensemble members ( $100 - 160 \text{ J kg}^{-1}$ , SI Fig. 8). Also, variations in the moisture content are small with difference between ensemble members smaller than  $0.5 \text{ g kg}^{-1}$  for all altitudes. The altitude of the driest point in the profile varies by about 100 hPa between ensemble members (SI Fig. 4).

## 4.2 Cloud-property variability

The different meteorological initial and boundary conditions result in different boundary-layer moisture convergence, thermodynamic and moisture profiles, and wind shear as discussed in the previous sections. These changes can impact cloud properties, cloud field structure, and precipitation formation.

The cloud droplet number concentration is almost invariant across ensemble members (Fig. 5) suggesting relatively small differences in the cloud-base vertical velocity distribution (SI Fig. 9 b). Cloud-top cloud droplet number concentrations also display little variability between meteorological ensemble members (SI Fig. 9 a).

The cloud field structure is described in terms of the cloud fraction, cell number and mean size, and cloud top height. Cells are defined as coherent areas with a column maximum radar reflectivity larger than 25 dBZ. Cloud top height is defined by the highest model level with a condensed water content larger than  $1 \text{ mg kg}^{-1}$  (e.g., Fridlind et al., 2010). Cloud fraction is calculated as the areal fraction of the domain with condensed water path larger than  $0.001 \text{ kg m}^{-2}$  (e.g., Grosvenor et al., 2017).

Cell number (Fig. 6 a) and cloud fraction (Fig. 7) in general decrease with decreasing boundary-layer moisture convergence and condensate gain. However, variations in mean cell size are quite small (Fig. 6 b). Mean cloud top height varies by about 750 m between ensemble members (Fig. 8) with largest (smallest) values for ensemble member 2 and 9 (5). Variations in mean cloud top height are in general consistent with those of the equilibrium level pressure (SI Fig. 10 c): For example, the equilibrium level pressure in ensemble member 5 is largest, while members 2 and 9 have the smallest equilibrium level pressure. The distribution between low, medium and high cloud tops varies by about a 20 % between the ensemble members (Fig. 8 b).

Precipitation formation is described by the condensation ratio CR and the precipitation efficiency PE. These describe the fraction of the incoming moisture flux that is converted to condensate (CR) and the fraction of the condensate gain that is converted to surface precipitation (PE), respectively. As expected, CR varies strongly across ensemble members and in gen-

eral decreases with decreasing large-scale moisture convergence (Fig. 9 a). In contrast, PE does not vary systematically with the large-scale convergence (Fig. 9 a). Ensemble member 4 has a significantly lower PE than the other ensemble members, which is likely related to the high fraction of shallow clouds with cloud tops below 2.5 km and a therefore small contribution of mixed-phase processes to domain-wide precipitation formation. Conversely, ensemble member 6 has a relatively large PE and the largest fraction of clouds with tops above 4.3 km. The relatively small differences in precipitation efficiency (between 0.17 and 0.27) are consistent with the almost invariant cloud droplet number concentrations for all ensemble members (Fig. 5). The combined effect of CR and PE results in a variation of about a factor 1.5 in the mean precipitation rate (Fig. 10 a) and the accumulated precipitation (SI Fig. 11). The precipitation variability corresponds in general to the variations in large-scale moisture convergence with some modulations by the different precipitation efficiencies (e.g. compare ensemble member 2 and control or ensemble members 5 and 6). Variations in the mean condensed water path are consistent with variations in the condensate generation between ensemble members (SI Fig. 10 b), i.e. the condensed water path decreases in members with smaller moisture convergence and  $G$ .

Mean reflected shortwave radiation ranges from  $130 \text{ W m}^{-2}$  to  $155 \text{ W m}^{-2}$  (Fig. 11a). The reflected shortwave is influenced by the cloud cover and the cloud droplet number concentrations. The largest (smallest) outgoing shortwave flux is predicted for the ensemble members with the largest (smallest) cloud fraction, i.e. ensemble 1 (8). Since the CDNC variability is small (Fig. 5), the variations in cloud fraction between ensemble members is dominating the variability of outgoing shortwave radiation. Changes in outgoing longwave radiation are on the order of  $3 \text{ W m}^{-2}$  (Fig. 11 b). The outgoing longwave radiation is influenced by the surface temperature, the cloud top height and the cloud fraction. While differences in the cloud top height distribution contribute to the variability in outgoing longwave radiation, variations in the clear sky outgoing longwave radiation dominate the overall variability due to the relatively small cloud fraction (SI Fig. 12).

## 5 Cloud property changes between ensemble members with different aerosol and identical meteorological initial and boundary conditions

The simulation of each meteorological ensemble member was conducted with three different aerosol profiles: a so-called “standard” aerosol scenario, which was derived from aircraft observations, and “low” and “high” aerosol scenarios, which have a factor 10 lower and higher aerosol number concentration, respectively. The impact of the perturbed aerosol profiles on cloud and cloud field properties as well as precipitation formation in the control simulation has been discussed in the first part of this study. In this section, we compare the aerosol signal in the different meteorological ensemble members, i.e. the difference in realisations with different aerosol scenarios but identical meteorological initial and boundary conditions. Thereby we test the robustness of aerosol-induced changes to small perturbations in the meteorological conditions. To quantify the significance of aerosol-induced changes we use a two-sided t-test for ensemble members paired according to meteorological conditions (Table. 1). Using paired ensemble members reflects the interdependence of cloud properties in realisations with different aerosol but identical meteorological initial and boundary conditions. Significance is tested at the 5 %-level.

## 5.1 Cloud droplet number concentration

The cloud-base CDNC is shown in Fig. 5 for all ensemble members and aerosol profiles. All ensemble members show a consistent increase of the cloud-base CDNC by about a factor 7 between the low (standard) and the standard (high) aerosol scenarios. Also, the aerosol-induced change in cloud-top CDNC is similar in all meteorological ensemble members with a change by about a factor 5.5 for each factor 10 increase in the background aerosol concentrations (SI Fig. 9 a). The small differences between ensemble members suggest only minor changes in the cloud-base vertical velocity distribution. **The aerosol-induced changes in CDNC are highly significant (Table 1).**

## 5.2 Cloud field structure

The cloud field structure is described in terms of cloud fraction, cell number and size, and cloud top height. **The number of cells decreases with increasing background aerosol concentrations in all ensemble members (Fig. 6 a). Conversely, the cell area increases (Fig. 6 b). The changes in cell number and area largely compensate, so that the cloud fraction displays little sensitivity to the aerosol scenarios with changes being smaller than 0.01 (Fig. 7). Although small, the aerosol-induced change in cloud fraction is consistent across all ensemble members. It has been hypothesised in the first part of this study, that the slower conversion of condensate to precipitation in high aerosol conditions allows clouds to grow larger and to merge with other updraft cores resulting in overall fewer, but larger clouds. Also, energetic constraints potentially limit an increase in overall lifting and cloud fraction. The changes in cell number, mean cell area and cloud fraction are all significant for paired meteorology reflecting the consistency in the sign of the aerosol-induced changes across the ensemble members (Table 1).**

The mean cloud top height is shown in Fig. 8 a and **the fraction of cloud tops in different altitude ranges in Fig. 8 b.** In all ensemble members, the mean cloud top height increases from the low to the standard aerosol scenario. The increase in mean cloud top is due to an increase in the fraction of cloud tops higher than 4.3 km. **In the control run and ensemble members 4 and 6, this is accompanied by a reduction in the medium altitude fraction, while in all other members changes in the low cloud top fraction dominate.** For an increase in aerosol number concentration above the standard aerosol scenario, the **time-average mean cloud top height (diamonds in Fig. 8 a)** does not increase further (members 1 and 2) or even decreases (members 4, 5, 6, 7, 8, 9). The decrease in mean cloud top height for the latter is mainly due to an increase in the fraction of clouds with low cloud tops. The fraction of clouds with high cloud tops shows only very small changes between the simulations with standard and high aerosol profiles. The small change in cloud top height is likely related to the presence of a mid-tropospheric stable layer, which is present in all ensemble members and limits cloud depths (section 4). Most larger convective cells have reached this “maximum” cloud top height for the standard aerosol scenario **and hence no further deepening occurs in the high aerosol scenario. The change of cloud top height is only significant for an increase in aerosol concentrations from the low to the standard scenario, while it is not significant for a further increase in aerosol concentrations (Table 1).**

### 5.3 Condensed water budget and precipitation formation

The condensation ratio displays only very small changes between different aerosol scenarios (Fig. 9 a). Accordingly, the condensate gain  $G$  changes only by 0 – 4 % between the low and the standard aerosol scenario and by –4 % to 2.5 % between the standard and high aerosol scenario (SI Fig. 14 a). As discussed in Miltenberger et al. (2018), the asymmetry in the response to increased and decreased aerosol concentrations is likely related to the thermodynamic limitations on cloud deepening. Changes in domain-wide condensation and deposition contribute to change in condensate gain  $\Delta G$  (SI Fig. 14 c and d). Condensation contributes most to the increases between the low and the standard aerosol scenario, while changes in condensation and deposition contribute about equally to  $\Delta G$  between the standard and high aerosol scenario. **Aerosol-induced modifications of CR and  $G$  are only significant for a decrease of aerosol concentrations relative to the standard scenario (Table 1).**

**The precipitation efficiency PE is more sensitive to aerosol changes than CR and decreases continuously with aerosol concentrations (Fig. 9 a). The change in PE is larger for increasing than decreasing aerosol concentration relative to the standard aerosol scenario. The stronger decrease in PE from the standard to the high aerosol scenario compared with the low and standard aerosol scenario is consistent with a higher lateral condensate transport to the stratiform region when cloud deepening becomes limited by thermodynamic constraints, as hypothesised by Miltenberger et al. (2018). With further cloud deepening limited by the upper-level stable layer, more condensate is transported into the stratiform area reducing the residence time of the condensate in the active convective core region. In contrast, cloud deepening is larger and changes in lateral condensate transport smaller when the low and standard aerosol scenario are compared. Therefore, the slower conversion of condensate to precipitation-sized hydrometeors in the standard aerosol scenario can be partly balanced by a longer residence time in the convective core region. This hypothesis is discussed in more detail in Miltenberger et al. (2018). Consistent with the larger amplitude and consistent sign, the changes in PE are significant for both a decrease and an increase of aerosol concentrations relative to the standard scenario (Table 1).**

**The changes in the condensate budget result in a modification of the accumulated surface precipitation as illustrated in Fig. 9 b. The diagram displays changes in condensate gain  $\Delta G$  and condensate loss  $\Delta L$  and is discussed in detail in Miltenberger et al. (2018). Increasing the aerosol concentrations from the low to the standard scenario results in a precipitation decrease in most ensemble members (points below the one-to-one line). Exceptions are the control simulation with a small increase in precipitation and ensemble members 6 and 8 with no change in accumulated surface precipitation (points on the one-to-one line). These ensemble members have a relatively small decrease of PE as well as a relatively large  $\Delta G$  and  $G$  compared to ensemble members with a similar change in PE (e.g. compare ensemble member 1 and 8). Accordingly, the precipitation response in these cases is either dominated by  $\Delta G$  (control) or  $\Delta G$  and  $\Delta PE$  are of equal importance (member 6 and 8), as also indicated by their position in the shaded area in Fig. 9 b. For the other members, the change in PE dominates over changes in condensate production, as indicated by their position outside the shaded area in Fig. 9 b. If the aerosol concentration is enhanced beyond the standard scenario, the precipitation decreases in all ensemble members (points above the one-to-one line). This response is dominated by PE changes in all ensemble members (points outside the shaded area). Differences in accumulated precipitation are significant, if ensemble members are paired according to meteorology (Table 1).**

The decrease in accumulated precipitation is accompanied by a reduced mean precipitation rate with increasing aerosol concentrations (Fig. 10a). For most ensemble members the change is larger between the standard and high aerosol scenario than between the standard and low aerosol scenario. **Only in ensemble member 3 does the mean precipitation rate not decrease further in the high aerosol scenario and in ensemble members 4 and 5 the decrease between the standard and the high aerosol scenario is comparable to the decrease between the low and standard scenario. The percentiles of the precipitation distribution increase for all percentiles up to and including the 75<sup>th</sup> percentile from the low to the high aerosol concentration (SI Fig. 11 c). The 99<sup>th</sup> percentiles are generally smallest (largest) for the high (standard) aerosol scenario. The only exception is ensemble member 4, for which the standard aerosol scenario has the smallest 99<sup>th</sup> percentile.**

The condensed water path in the domain is a result of the condensate generation and the timescale of condensate conversion to precipitation. Parcel model considerations suggest a longer timescale for precipitation formation under enhanced aerosol concentrations. Therefore, an increase in the condensed water path is expected with increasing aerosol concentrations. Indeed, the mean condensed water path in most ensemble members increases with aerosol concentrations (Fig. 10 b). This is the result of small decreases in the liquid water path (cloud and rain species) and a larger gain in the mass of the frozen hydrometeors (ice, snow, and graupel species) (SI Fig. 12), consistent with a slower conversion of cloud droplets to rain drops and accordingly a larger mass transport across the 0 ° C level. The only ensemble member displaying a different pattern is ensemble member 9, for which the total condensed, the solid, and the liquid water path decrease from the standard to the high aerosol scenario. **This ensemble member has also the largest reduction in precipitation efficiency.** In addition, for ensemble member 9 the mean cloud top height and the fraction of clouds with cloud tops larger than 4.3 km decreases compared to the standard aerosol scenario. These changes are consistent with a lower condensed water path, as reduced cloud top heights indicate a smaller vertical displacement of the air parcels and accordingly less condensate generation. The decrease in precipitation efficiency is likely linked to these changes as the longer timescale for conversion of cloud droplets to precipitation seized hydrometeors is not compensated by a longer residence time in the cloud due the reducing vertical extend of the clouds. **Aerosol-induced changes in condensed, liquid and frozen water path are significant in the paired meteorology tests (Table 1).**

#### 5.4 Radiation

The response of cloud radiative properties to changes in aerosol concentrations is climatologically important, but not well constrained mainly due to the impact of aerosols on both cloud fraction and cloud lifetime. The reflected shortwave radiation is affected by the size and number of the hydrometeors close to cloud top and by the cloud fraction. **The outgoing shortwave flux increases for higher aerosol concentrations in all ensemble members (Fig. 11a). This change is consistent with the aerosol-induced change in CDNC (Fig. reffig:cdnc) and the cloud albedo effect (Twomey, 1977). The co-occurring decrease of cloud fraction under high aerosol conditions (between 2 and 9 % for a factor 10 aerosol change) counteracts the CDNC effect, but the cloud albedo effect dominates due to the large amplitude of the CDNC change (about a factor 7 for a factor 10 aerosol change).** Note, that the radiative signal presented here does not fully take into account potential changes in radiative properties of the ice-phase species, as the effective diameter of the latter is diagnosed from the ice water content.

The outgoing longwave radiation is mainly influenced by the surface temperature, the cloud fraction, and the cloud top tem-

perature. The mean outgoing longwave radiation shows only a small sensitivity to the aerosol scenario for all meteorological ensemble members (Fig. 11b). The small discernible trend of decreasing mean outgoing longwave radiation with increasing aerosol ( $< 0.5 \text{ W m}^{-2}$ , standard to high aerosol scenario) is consistent with the small increase in mean cloud top height (Fig. 8 a).

5 **Aerosol-induced modifications to the outgoing radiative fluxes are significant at the 5% level.**

## 6 **Contribution of aerosol and meteorology perturbations to overall cloud property variability**

In the previous two sections, the response of cloud properties to perturbations of the aerosol or meteorological initial and boundary conditions has been discussed separately. The 10 meteorological ensemble members vary in the large-scale moisture convergence, thermodynamic profile, and wind shear. The variation in the large-scale moisture convergence is most important for the cloud field properties, e.g. cloud fraction, cell number, condensate generation, and accumulated precipitation (section 4). The mean cloud top height varies between ensemble members according to the different thermodynamic profiles. Aerosol-induced changes follow a similar pattern for each meteorological ensemble member (section 5). An increase in aerosol number concentration translates to a larger CDNC, mean cell area, and outgoing shortwave radiation, while the cell number and precipitation efficiency decrease with increasing aerosol concentrations. The mean cloud top height, condensate generation, and outgoing longwave radiation display only very small changes in response to altered aerosol concentrations.

To detect aerosol-induced changes in cloud properties or precipitation formation with observational datasets, it is important to separate changes resulting from different meteorological conditions from changes resulting from different aerosol concentrations. This is necessitated by the co-variability of aerosol and meteorological conditions in the real atmosphere. The question of the relative importance of meteorological and aerosol initial and boundary conditions for the cloud field structure and precipitation formation is also important for operational numerical weather prediction and the future design of ensemble prediction systems. Here, we use the combined meteorological and aerosol initial condition ensemble, i.e. combining the discussion of the two previous section, to address the question of the relative importance of aerosol and meteorological variability for the COPE case. The discussion will focus on changes in the 10 hour mean properties of the cloud field between 0900 UTC and 1900 UTC. The mean value of the considered variable is displayed along with its spread from the meteorological ensemble members on the right side of Figs. 5-11 for each aerosol scenario (different colours). If instantaneous realisations of the different (domain-averaged) variables would be considered (box-plots on left side of the figures), the variability would be much larger than suggested by the domain-mean time-averaged plots (right side of the plots). For a quantitative assessment we use the p-values from two-sided t-tests for the full ensemble, i.e. not pairing ensemble members according to the meteorological initial conditions as in section 4.

30 The cloud droplet number concentration at either cloud base or cloud top is strongly influenced by the assumed aerosol scenario, but varies little between the different meteorological members (Fig. 5, SI Fig. 9 a). As a result, a clear aerosol signal remains present even when the meteorological variability is taken into account. The aerosol-induced CDNC change remains highly significant at the 5% level in the unpaired tests (Table 1).

Although there is a stronger meteorology-induced variability in the cell number and mean cell size (Fig. 6) and the predicted range of values overlap for different aerosol scenarios, the aerosol signal is clearly detectable in these variables and aerosol effects remain significant also in the unpaired test (Table 1). However, if the cloud fraction, the mean cloud top height or the distribution in different cloud top height classes is considered the meteorological variability dominates (Fig. 7, 8). Hence, changes in cloud fraction, mean cloud top height and deep cloud fraction are not significant at the 5% level, if ensemble members are not paired according to meteorological conditions (Table 1). Considering previous arguments on convective invigoration, it is interesting to note that the cloud top height varies only very little with aerosol scenario, but is sensitive to relatively small changes in meteorological conditions. Consistent with the small changes in cloud top height, no significant differences in outgoing longwave radiation exist between the aerosol scenarios (Fig. 11 b, Table 1). For the outgoing shortwave radiation a stronger aerosol signal is retained above the meteorological variability due to the large impact of aerosol concentrations on CDNC (Fig. 11a). However, this signal is not statistically significant for an aerosol increase beyond the standard aerosol scenario (Table 1).

Precipitation formation is known to be strongly influenced by dynamical and microphysical processes. Miltenberger et al. (2018) used an analysis of the water budget to separate the contributions from cloud dynamics and microphysics to aerosol-induced changes. As expected, condensation ratio CR and condensate gain G vary strongly between different meteorological ensemble members, but show little sensitivity to the aerosol scenario (Fig. 9 a, SI Fig. 14 a). The small dependency of the condensate gain on the aerosol number concentration may be a result of using a saturation adjustment scheme for the condensation in our model. Previous studies using a prognostic supersaturation found the condensation rates to be dependent on the CDNC number concentration (e.g. Lebo et al., 2012; Lebo, 2014; Sheffield et al., 2015). However, due to the thermodynamic constraints on integrated condensation, we do not expect this will have a strong impact on the overall behaviour of the condensate gain. In contrast to CR, the precipitation efficiency PE displays a relatively small systematic dependency on the large-scale moisture convergence and a large dependency on the aerosol scenario (Fig. 9 a). However, the condensate loss L varies strongly across meteorological ensemble members due to its close relation with the condensation gain (SI Fig. 15 a). This co-variability is discounted for in PE. However, still only the aerosol-induced PE change between the standard and high aerosol scenario is significant at the 5% level for unpaired ensemble members.

The aerosol-induced change in accumulated precipitation is the combined result of the changes in condensation ratio and precipitation efficiency. While the accumulated surface precipitation in most meteorological ensemble members decreases with increasing aerosol scenario, these differences are much smaller than the variability of accumulated surface precipitation between meteorological ensemble members (SI Fig. 11). The meteorological variability is due to large differences in the condensate gain, which is directly related to the variability in large-scale moisture convergence. The aerosol signal is much larger for an increase of the aerosol concentrations beyond the standard aerosol scenario due to a significantly larger change in the precipitation efficiency. The mean precipitation rate behaves qualitatively very similarly to the accumulated precipitation (Fig. 10 a). Consistently, neither changes in mean precipitation rate nor accumulated precipitation are statistically significant (Table 1).

The ensemble-mean condensed water path increases with increasing aerosol concentrations, if all hydrometeor types are con-

sidered (Fig. 10 b). However, the liquid water path (condensate in the cloud and rain category) shows relatively little sensitivity in its median value, while the mean liquid water path generally decreases with increasing aerosol concentrations. This indicates an increasing frozen water path with increasing aerosol concentrations consistent with a longer timescale for precipitation formation. The aerosol-induced change in both variables are much smaller than the variability induced by different meteorological initial and boundary conditions and are hence not significant at the 5% level (Table 1).

## 7 Discussion and Conclusions

High-resolution ensemble simulations ( $\Delta x = 250$  m) with perturbed aerosol and meteorological initial and boundary conditions were performed for convection forming along sea-breeze convergence zones over the southwestern peninsula of the UK. The relative importance of perturbations in meteorological (10 members) and aerosol initial conditions (3 for each member) for various cloud properties and precipitation formation is analysed over a forecast lead-time of 10 – 20 h. The 10 different meteorological ensemble members develop similar mesoscale flow patterns with a sea-breeze convergence zone establishing over the centre of the peninsula. As a result of the different lateral boundary conditions, the large-scale boundary layer moisture convergence and the accumulated condensate gain vary by a factor 2 and the accumulated surface precipitation by a factor 2.5 between ensemble members. The average cloud fraction differs by up to 0.1 between the meteorological ensemble members. This meteorological variability is compared to changes in cloud properties induced by a factor 10 increase and decrease of aerosol number concentrations relative to the standard scenario. While the perturbations to the meteorological initial conditions reflect at best the uncertainty for the investigated case, changes in aerosol number concentration by a factor 100 are probably even larger than what could be expected for the climatological variability. Changes in aerosol concentrations can potentially modify cloud field properties, e.g. cell number and size, cloud depth, cloud fraction, and the domain-wide condensate budget (condensate gain and loss, precipitation rate). Aerosol-induced changes are consistent across the ensemble suggesting that the physical mechanism discussed by Miltenberger et al. (2018) is robust against small changes in meteorological initial conditions. The variability of cloud field properties across the ensemble is summarised in Fig. 12. The possibility of discerning aerosol-induced differences in various cloud metrics relative to realistic meteorological variability is assessed in the following. First, the idealised situation where the meteorological initial conditions are identical for different aerosol perturbations is assessed by pairing ensemble members according to the meteorological initial conditions. This is equivalent to testing the statistical significance of the differences between realisations with different aerosol scenarios and identical meteorological initial and boundary conditions. For the paired ensemble members, a factor 10 increase or decrease in aerosol concentrations introduces statistically significant changes (at the 5% level) in CDNC, cloud fraction, cell number and size, outgoing shortwave radiation, instantaneous and mean precipitation rates, and precipitation efficiency (Table 1). Note that the statistical analysis is based on a very small sample, which affects the validity of several assumptions. However, since the statistical results agree qualitatively with the physical analysis, we use the significance values as a helpful diagnostic for summarising the results. Aerosol-induced changes in accumulated precipitation are only significant for an increase of aerosol

concentrations beyond the standard scenario. An analysis of the condensed water budget suggests that for a decrease in aerosol concentrations, a smaller condensation ratio is balanced by an increasing precipitation efficiency. In contrast, for higher aerosol concentrations than in the standard scenario, the precipitation response is dominated by a strong decrease in precipitation efficiency with little change in the condensation ratio due to the thermodynamic constraints on cloud top height.

5 Secondly, we can use the simulations to assess our ability to discern aerosol-cloud effects for the situation where meteorological initial and boundary conditions are similar but subject to observational uncertainty. This would represent a “perfect” observational campaign where the meteorological conditions each day are only slightly different (convergence within a factor 2) and large perturbations to aerosol concentrations occur (factor 10 – 100). This scenario is replicated by analysing aerosol-induced changes in the full ensemble without pairing ensemble members according to meteorological initial conditions. For  
10 the un-paired ensemble, only aerosol-induced changes in CDNC, cell number and size, outgoing shortwave radiation, and precipitation efficiency are statistically significant (Table 1). For some of these variables, the changes are significant only for a decrease or an increase of aerosol number concentration relative to the standard scenario. For all other investigated variables (cloud fraction, cloud top height, condensation ratio, domain-average precipitation rate, condensed water path, and liquid water path) the variability resulting from different meteorological initial and boundary conditions is equal to or larger than the  
15 aerosol-induced changes.

The ensemble data can be used for a rough estimate of the number of observations that are required for retrieving a robust aerosol signal from observational data for sea-breeze convection. For this analysis we assume (i) the aerosol scenario and meteorology are independent, (ii) the ensemble is representative of the meteorological variability, (iii) the meteorological variability can be described by a Gaussian distribution, and (iv) observational data is perfect. While it is difficult to a priori estimate  
20 the impact of these assumptions on the analysis, we expect the analysis to provide a lower limit of the required number of observations due to the following reasons: In contrast to assumption (i), aerosol and meteorological conditions are likely to be correlated (e.g. Brenguier et al., 2003; Naud et al., 2016; Wood et al., 2017) reducing the observed section of the phase-space. Secondly, the meteorological variability in the ensemble simulations is not representative of the climatological variability of meteorological conditions for sea-breeze convection over the the southwestern peninsula of the UK, which can be assumed  
25 to be much larger (Golding, 2005). Lastly, observational data will not be perfect due to measurement errors and spatial and temporal sampling issues (e.g. Schutgens et al., 2017). All these issues will likely increase the number of required samples compared to the values suggested by our analysis.

With the assumptions listed above, a Gaussian distributions representing the meteorological variability is defined for the low and the high aerosol scenario. For each variable, the mean value across all ensemble members with the same aerosol scenario is  
30 used as the mean of the Gaussian distribution. The standard deviation  $\sigma$  of the Gaussian distribution is defined by assuming the value range (minimum and maximum) across the ensemble members equals  $4\sigma$  ( $3\sigma$  and  $5\sigma$  are tested as well). Then  $10^4$  realisations with  $n$  samples are drawn from the Gaussian distributions for the low and high aerosol scenario separately. The number of samples  $n$  can be interpreted as the number of times the same day is observed, as the statistical analysis presented here uses either daily average or accumulated variables. However, it may be possible to interpret the necessary number of observations  
35 also as number of individual observations, e.g. from satellite overpasses, if subsequent observations are not auto-correlated, i.e.

are from different cloud lifecycles. If snapshots are used, it may be necessary to take into the possibly different life cycle stages of the observed cloud field (e.g. Luo et al., 2009; Witte et al., 2014). However, given the limited number of ensemble members in our analysis, an assessment of this effect is beyond the scope of our study. For each of the  $10^4$  realisations, we test the hypothesis that the low and high aerosol scenario are not equal with a two-sided t-test. The resulting distribution of p-values gives the probability that a significant aerosol-signal can be retrieved from a sample of  $n$  observations with low and high aerosol conditions each (SI Fig. 16). The number of days required to have a 95 % chance to observe a significant aerosol-induced change in various cloud properties is listed in Table 2. This required number of observations only gives an approximate indication, as the exact number is sensitive to the assumptions made regarding the presentation of the meteorological variability in the ensemble, e.g. whether the ensemble spread corresponds to  $3\sigma$ ,  $4\sigma$ , or  $5\sigma$ . It is important to note that the statistical analysis has the strong caveat of being based on a rather small ensemble. To obtain robust statistics a much larger ensemble with several hundreds ensemble members would be required, which is currently beyond the computational resources available. However, we think the analysis provided here gives some general indication of the scale of observations required as the statistics confirm the impressions gained from the physical analysis of the ensemble members. Our analysis indicates that a small sample  $n \leq 10$  is sufficient for variables such as the CDNC and outgoing shortwave radiation, while a large sample often exceeding 100 is required for variables such as cloud fraction, cloud top height, or accumulated precipitation. The number of samples required depends on the amplitude of the aerosol perturbation (low and high aerosol scenario versus low/high and standard scenario) as well as the location in the aerosol space (different for increase or decrease relative to the standard aerosol scenario). In general, more observations are required for an increase of aerosol number concentrations above the standard scenario, which is related to the thermodynamic constraints on aerosol-induced changes in the considered case discussed in Miltenberger et al. (2018). The only exception is accumulated surface precipitation, for which fewer observations are required for an increase above the standard scenario. This reflects the larger aerosol-induced signal in accumulated precipitation for increased compared to decreased aerosol concentrations.

While the meteorological ensemble allows us to put the aerosol-induced changes in cloud properties into the context of changes related to meteorological variability, the considered changes in meteorology are fairly small (section 4). Even if the represented meteorological variability is assumed to be representative of all possible meteorological conditions on the investigated day, they do not cover the full range of meteorological conditions that could occur for convection along sea-breeze convergence zones. However, even this very conservative estimate on the meteorological variability is for many variables on the same order of magnitude or larger than the aerosol-induced changes. We expect that the number of samples required to retrieve a statistically robust aerosol-induced change would increase, if the climatological variability of the meteorological conditions is considered. The results presented in this paper certainly only pertain to the specific cloud type investigated and the relative magnitude of aerosol and meteorological related changes in cloud properties may be different for other cloud types. This will be investigated in future studies. In addition to the results presented here, some previous studies have highlighted the importance of considering the intrinsic predictability of investigated cases before drawing conclusions about the significance of aerosol-induced changes in cloud properties (Grabowski et al., 1999; Khairoutdinov and Randall, 2003; Zeng et al., 2008; Morrison and Grabowski, 2011; Morrison, 2012). These studies used prescribed large-scale conditions and applied random perturbations to thermody-

5 namic fields throughout the simulations. The present study complements their analysis by considering the impact of changes in the large-scale conditions, which are small compared to observational uncertainties and much smaller than the expected variability in meteorological categories used to retrieve aerosol signals from general circulation models (e.g. Bony et al., 2004; Zhang et al., 2016). Consistent with previous studies, we find that the aerosol signals in variables closely related to aerosol concentrations, such as for example cloud droplet number concentrations, are easier to retrieve than for variables that are linked to aerosol concentrations by a series of complex processes, such as for example accumulated surface precipitation. From the limited number of studies available, the set of variables in either category appears to vary for different cloud types and geographic location. However, our and previous studies all suggest that aerosol-induced change in surface precipitation are very difficult to retrieve reliably (e.g. Khairoutdinov and Randall, 2003; Morrison and Grabowski, 2011).

10 The evidence presented in the to-date very limited number of studies considering the relative impact of meteorological and aerosol conditions on cloud properties suggest that it is crucial to carefully consider intrinsic predictability, meteorological conditions, and co-variability between aerosol and meteorological conditions in modelling and observational studies of aerosol indirect effects. While these aspects have been highlighted by Stevens and Feingold (2009) and Feingold et al. (2016), only a few modelling studies have investigated these aspects and there is a clear need for future studies extending the analysis to other cloud types and meteorological scenarios. An improved knowledge and quantification of these aspects is mandatory for progress in our understanding of aerosol-induced changes in cloud properties and for retrieving observational evidence thereof.

15

*Data availability.* Model data is stored on the tape archive provided by JASMIN (<http://www.jasmin.ac.uk/>) service. Data access to Met Office data via JASMIN is described at <http://www.ceda.ac.uk/blog/access-to-the-met-office-mass-archive-on-jasmin-goes-live/>.

*Author contributions.* All authors contributed to the development of the concepts and ideas presented in this paper. B. J. Shipway developed the CASIM microphysics code. A. A. Hill, J. M. Wilkinson, P. R. Field and A. K. Miltenberger contributed to the further development of the CASIM code. A. K. Miltenberger and P. R. Field set up the model runs. A. K. Miltenberger performed the model simulations and analysis, and wrote the majority of the manuscript, along with input and comments from all co-authors.

20

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

*Acknowledgements.* We thank the COPE research team for collecting observational data, Alan Blyth for very useful discussion input and Jill Johnson for the insightful discussions on the statistical analysis. We acknowledge use of the Monsoon/NEXCS system, a collaborative facility supplied under the Joint Weather and Climate Research Programme, a strategic partnership between the Met Office and the Natural Environment Research Council. Further we acknowledge JASMIN storage facilities (doi : 10.1109/BigData.2013.6691556) as well as

25

FAAM, CEDA, BADC, and the Met Office for providing data. Funding for this study has been provided by the University of Leeds and the U.K. Natural Environment Research Council under grant NE/J023507/1. We thank two anonymous reviewers for their valuable feedback.

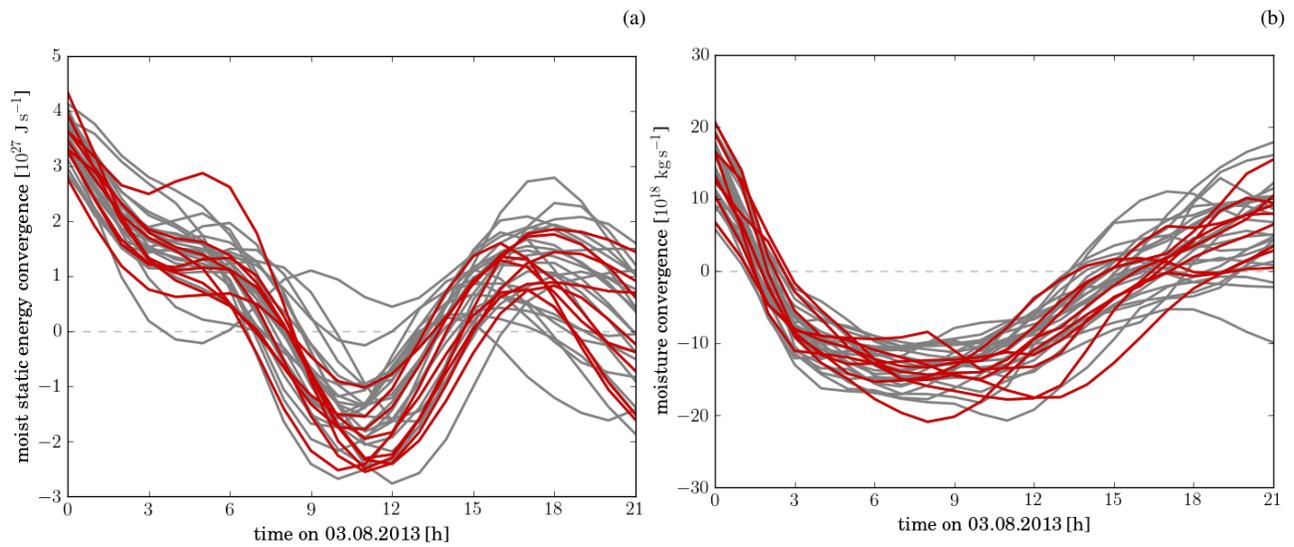
## References

- Altaratz, O., Koren, I., Remer, L. A., and Hirsch, E.: Review: Cloud invigoration by aerosols - coupling between microphysics and dynamics, *Atmos. Res.*, 140–141, 38–60, <https://doi.org/10.1016/j.atmosres.2014.01.009>, 2014.
- Aranami, K., Zerroukat, M., and Wood, N.: Mixing properties of SLICE and other mass-conservative semi-Lagrangian schemes, *Q. J. R. Meteorol. Soc.*, 140, 2084–2089, <https://doi.org/10.1002/qj.2268>, 2014.
- 5 Aranami, K., Davies, T., and Wood, N.: A mass restoration scheme for limited-area models with semi-Lagrangian advection, *Q. J. R. Meteorol. Soc.*, 141, 1795–1803, <https://doi.org/10.1002/qj.2482>, 2015.
- Beck, J., Bouttier, F., Wiegand, L., Gebhardt, C., Eagle, C., and Roberts, N.: Development and verification of two convection-allowing multi-model ensembles over Western Europe, *Q. J. R. Meteorol. Soc.*, 142, 2808–2826, <https://doi.org/10.1002/qj.2870>, 2016.
- 10 Berner, J., Ha, S.-Y., Hacker, J. P., Fournier, A., and Snyder, C.: Model uncertainty in a mesoscale ensemble prediction system: stochastic versus multiphysics representations, *Mon. Wea. Rev.*, 139, 1972–1995, <https://doi.org/10.1175/2010MWR3595.1>, 2011.
- Blyth, A. M., Bennett, L. J., and Collier, C. G.: High-resolution observations of precipitation from cumulonimbus clouds, *Meteorol. Appl.*, 22, 75–89, <https://doi.org/10.1002/met.1492>, 2015.
- Bony, S., Dufresne, J. L., Le Treut, H., Morcrette, J. J., and Senior, C.: On dynamic and thermodynamic components of cloud changes, *Clim. Dyn.*, 22, 71–86, <https://doi.org/10.1007/s00382-003-0369-6>, 2004.
- 15 Bowler, N. E., Arribas, A., Mylne, K. R., Robertson, K. B., and Beare, S. E.: The MOGREPS short-range ensemble prediction system, *Q. J. R. Meteorol. Soc.*, 134, 703–722, <https://doi.org/10.1002/qj.234>, 2008.
- Bowler, N. E., Arribas, A., Beare, S. E., Mylne, K. R., and Shutts, G. J.: The local ETKF and SKEB: upgrades to the MOGREPS short-range ensemble prediction system, *Q. J. R. Meteorol. Soc.*, 135, 767–776, <https://doi.org/10.1002/qj.394>, 2009.
- 20 Brenguier, J.-L., Pawlowska, H., and Schüller, L.: Cloud microphysical and radiative properties for parameterization and satellite monitoring of the indirect effect of aerosol on climate, *J. Geophys. Res.*, 108, 8632, <https://doi.org/10.1029/2002JD002682>, 2003.
- Devasthale, A., Kruger, O., and Grassl, H.: Change in cloud-top temperatures over Europe, *IEEE Geoscience and Remote Sensing Letters*, 2, 333–336, <https://doi.org/10.1109/LGRS.2005.851736>, 2005.
- Estoque, M. A.: A theoretical investigation of the sea breeze, *Q. J. R. Meteorol. Soc.*, 87, 136–146, <https://doi.org/10.1002/qj.49708737203>, 25 1961.
- Fan, J., Wang, Y., Rosenfeld, D., and Liu, X.: Review of Aerosol–Cloud Interactions: Mechanisms, Significance, and Challenges, *J. Atmos. Sci.*, 73, 4221–4252, <https://doi.org/10.1175/JAS-D-16-0037.1>, 2016.
- 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. USA*, 113, 5812–5819, <https://doi.org/10.1073/pnas.1514035112>, 2016.
- 30 Fridlind, A. M., Ackerman, A. S., Petch, J., Field, P. R., Hill, A., G.McFarquhar, G., Xie, S., and Zhang, M.: ARM/GCSS/SPARC TWP-ICE CRM intercomparison study, NASA Tech. Memo., NASA TM-2010-215858, 24 pp., 2010.
- Golding, B.: Boscastle and North Cornwall post flood event study - meteorological analysis of the conditions leading to flooding on 16 August 2004, Forecasting Research Technical Report 459, Met Office, 2005.
- Golding, B., Clark, P., and May, B.: The Boscastle flood: Meteorological analysis of the conditions leading to flooding on 16 August 2004, 35 *Weather*, 60, 230–235, <https://doi.org/10.1256/wea.71.05>, 2005.
- Grabowski, W. W.: Indirect impact of atmospheric aerosols in idealised simulations of convective-radiative equilibrium, *J. Climate*, 19, 4664–4682, <https://doi.org/10.1175/JCLI3857.1>, 2006.

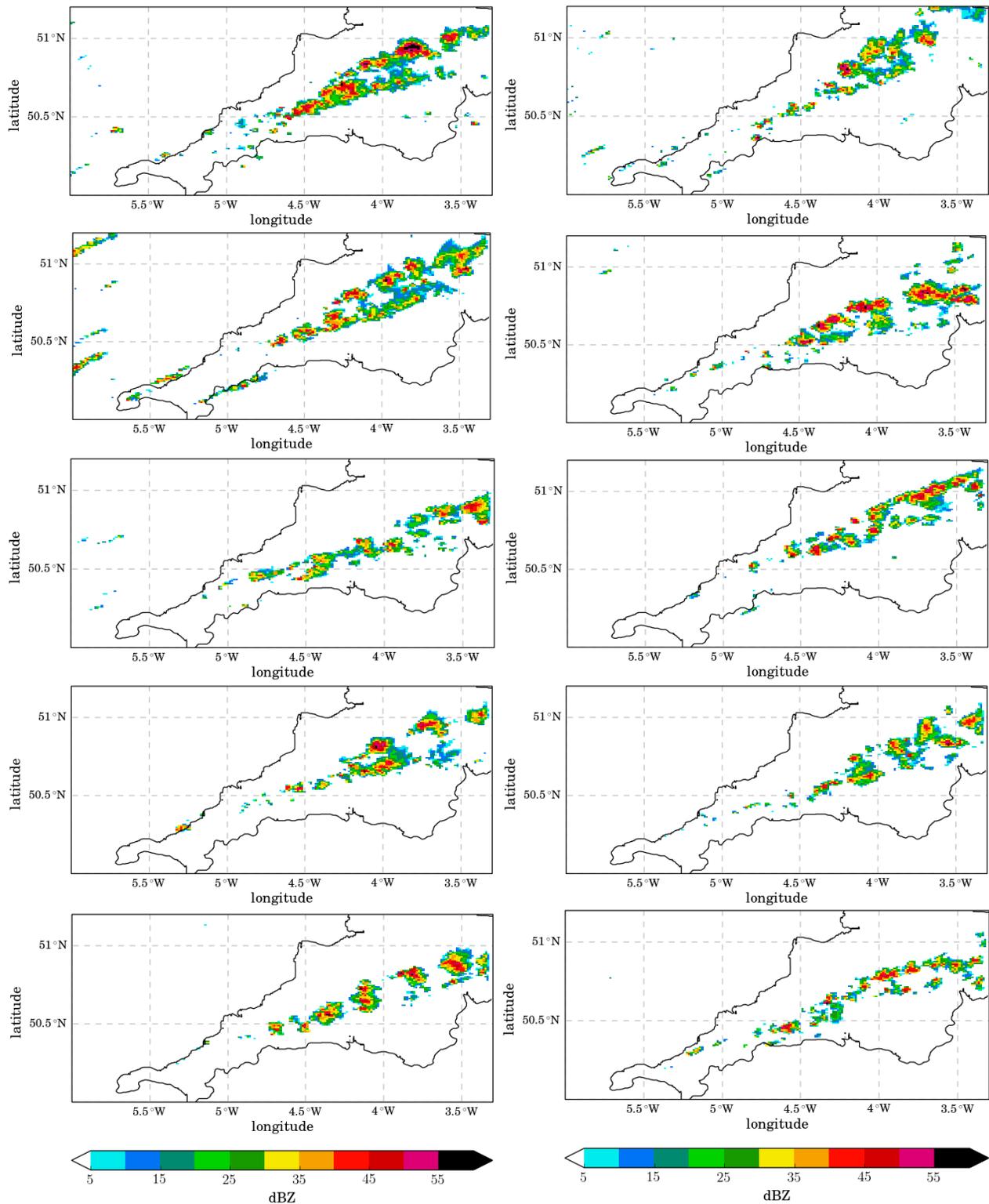
- Grabowski, W. W., Wu, X., and Moncrieff, M. W.: Cloud Resolving Modeling of Tropical Cloud Systems during Phase III of GATE. Part III: Effects of Cloud Microphysics, *J. Atmos. Sci.*, 56, 2384–2402, [https://doi.org/10.1175/1520-0469\(1999\)056<2384:CRMOTC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<2384:CRMOTC>2.0.CO;2), 1999.
- Grosvenor, D. P., Field, P. R., Hill, A. A., and Shipway, B. J.: The relative importance of macrophysical and cloud albedo changes for aerosol induced radiative effects in stratocumulus: insight from the modelling of a case study, *Atmos. Chem. Phys.*, 17, 5155–5183, <https://doi.org/10.5194/acp-17-5155-1017>, 2017.
- Gryspeerdt, E., Stier, P., and Partridge, D. G.: Links between satellite retrieved aerosol and precipitation, *Atmos. Chem. Phys.*, 14, 9677–9694, <https://doi.org/10.5194/acp-14-9677-2014>, 2014.
- Harrison, D. L., Scovell, R. W., and Kitchen, M.: High-resolution precipitation estimates for hydrological uses, *Proceedings of the Institution of Civil Engineers - Water Management*, 162, 125–135, <https://doi.org/10.1680/wama.2009.162.2.125>, 2009.
- Hill, A. A., Shipway, B. J., and Boutle, I. A.: How sensitive are aerosol-precipitation interactions to the warm rain representation?, *J. Adv. Model. Earth Syst.*, 7, 987–1004, <https://doi.org/10.1002/2014ms000422>, 2015.
- Khairoutdinov, M. F. and Randall, D. A.: Cloud Resolving Modeling of the ARM Summer 1997 IOP: Model Formulation, Results, Uncertainties, and Sensitivities, *J. Atmos. Sci.*, 60, 607–625, [https://doi.org/10.1175/1520-0469\(2003\)060<0607:CRMOTA>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<0607:CRMOTA>2.0.CO;2), 2003.
- Koren, I., Feingold, G., and Remer, L. A.: The invigoration of deep convective clouds over the Atlantic: aerosol effect, meteorology or retrieval artifact?, *Atmos. Chem. Phys.*, 10, 8855–8872, <https://doi.org/10.5194/acp-10-8855-2010>, 2010.
- Lakshmanan, V., Hondl, K., Potvin, C. K., and Preignitz, D.: An improved method for estimating radar echo-top height, *Wea. Forecasting*, 28, 481–488, <https://doi.org/10.1175/WAF-D-12-00084.1>, 2013.
- Lebo, Z. J.: The Sensitivity of a Numerically Simulated Idealized Squall Line to the Vertical Distribution of Aerosols, *J. Atmos. Sci.*, 71, 4581–4596, <https://doi.org/10.1175/JAS-D-14-0068.1>, <https://doi.org/10.1175/JAS-D-14-0068.1>, 2014.
- Lebo, Z. J., Morrison, H., and Seinfeld, J. H.: Are simulated aerosol-induced effects on deep convective clouds strongly dependent on saturation adjustment?, *Atmos. Chem. Phys.*, 12, 9941–9964, <https://doi.org/10.5194/acp-12-9941-2012>, 2012.
- Lee, S. S.: Effect of Aerosol on Circulations and Precipitation in Deep Convective Clouds, *J. Atmos. Sci.*, 69, 1957–1974, <https://doi.org/10.1175/JAS-D-11-0111.1>, 2012.
- Leon, D. C., French, J. R., Lasher-Trapp, S., Blyth, A. M., Abel, S. J., Ballard, S., Barrett, A., Bennett, L. J., Bower, K., Brooks, B., Brown, P., Charlton-Perez, C., Choullarton, T., Clark, P., Collier, C., Crosier, J., Cui, Z., Dey, S., Dufton, D., Eagle, C., Flynn, M. J., Gallagher, M., Halliwell, C., Hanley, K., Hawkness-Smith, L., Huang, Y., Kelly, G., Kitchen, M., Korolev, A., Lean, H., Liu, Z., Marsham, J., Moser, D., Nicol, J., Norton, E. G., Plummer, D., Price, J., Ricketts, H., Roberts, N., Rosenberg, P. D., Simonin, D., Taylor, J. W., Warren, R., Williams, P. I., and Young, G.: The Convective Precipitation Experiment (COPE): Investigating the origins of heavy precipitation in the southwestern United Kingdom, *B. Am. Meteorol. Soc.*, 97, 1003–1020, <https://doi.org/10.1175/BAMS-D-14-00157.1>, 2016.
- Li, X. and Srivastava, R. C.: An Analytical Solution for Raindrop Evaporation and Its Application to Radar Rainfall Measurements, *J. Appl. Meteor.*, 40, 1607–1616, [https://doi.org/10.1175/1520-0450\(2001\)040<1607:AASFRE>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<1607:AASFRE>2.0.CO;2), 2001.
- Luo, Z., Liu, G. Y., Stephens, G. L., and Johnson, R. H.: Terminal versus transient cumulus congestus: A CloudSat perspective, *Geophys. Res. Lett.*, 36, <https://doi.org/10.1029/2008GL036927>, 2009.
- Marsigli, C., Montani, A., and Paccagnella, T.: Perturbation of initial and boundary conditions for a limited-area ensemble: multi-model versus single-model approach, *Q. J. R. Meteorol. Soc.*, 140, 197–208, <https://doi.org/10.1002/qj.2128>, 2014.
- MetOffice: 1 km Resolution UK Composite Rainfall Data from the Met Office Nimrod System, 2003.

- Miller, S. T. K., Keim, B. D., Talbot, R. W., and Mao, H.: Sea breeze: Structure, forecasting, and impacts, *Rev. Geophys.*, 41, <https://doi.org/10.1029/2003RG000124>, 1011, 2003.
- 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.
- 5 Morrison, H.: On the robustness of aerosol effects on an idealized supercell storm simulated with a cloud system-resolving model, *Atmos. Chem. Phys.*, 12, 7689–7705, <https://doi.org/10.5194/acp-12-7689-2012>, 2012.
- Morrison, H. and Grabowski, W. W.: Cloud-system resolving model simulations of aerosol indirect effects on tropical deep convection and its thermodynamic environment, *Atmos. Chem. Phys.*, 11, 503–523, <https://doi.org/10.5194/acp-11-10503-2011>, 2011.
- 10 Naud, C. M., Posselt, D. J., and van den Heever, S. C.: Aerosol optical depth distribution in extratropical cyclones over the northern hemisphere oceans, *Geophys. Res. Lett.*, 43, 504–511, <https://doi.org/10.1002/2016GL070953>, 2016.
- Romine, G. S., Schwartz, C. S., Berner, J., Fossell, K. R., Snyder, C., Anderson, J. L., and Weisman, M. L.: Representing Forecast Error in a Convection-Permitting Ensemble System, *Mon. Wea. Rev.*, 142, 4519–4541, <https://doi.org/10.1175/MWR-D-14-00100.1>, 2014.
- Rosenfeld, D., Lohmann, U., Raga, G. B., O’Dowd, C. D., Kulmala, M., Fuzzi, S., Reissell, A., and Andreae, M. O.: Flood or drought: How do aerosols affect precipitation?, *Science*, 321, 1309–1313, <https://doi.org/10.1126/science.1160606>, 2008.
- 15 Rosenfeld, D., Andreae, M. O., Asmi, A., Chin, M., de Leeuw, G., Donovan, D. P., Kahn, R., Kinne, S., Kivekäs, N., Kulmala, M., Lau, W., Schmidt, K. S., Suni, T., Wagner, T., Wild, M., and Quaas, J.: Global observations of aerosol-cloud-precipitation climate interactions, *Rev. Geophys.*, 52, 750–808, <https://doi.org/10.1002/2013RG000441>, 2014.
- Saito, K., Hara, M., Kunii, M., Seko, H., and Yamaguchi, M.: Comparison of initial perturbation methods for the mesoscale ensemble prediction system of the Meteorological Research Institute for the WWRP Beijing 2008 Olympics Research and Development Project (B08RDP), *Tellus A*, 63, 445–467, <https://doi.org/10.1111/j.1600-0870.2010.00509.x>, 2011.
- 20 Schutgens, N., Tsyro, S., Gryspeerdt, E., Goto, D., Weigum, N., Schulz, M., and Stier, P.: On the spatio-temporal representativeness of observations, *Atmos. Chem. Phys.*, 17, 9761–9780, <https://doi.org/10.5194/acp-17-9761-2017>, 2017.
- Schwartz, C. S., Romine, G. S., Smith, K. R., and Weisman, M. L.: Characterizing and Optimizing Precipitation Forecasts from a Convection-Permitting Ensemble Initialized by a Mesoscale Ensemble Kalman Filter, *Wea. Forecasting*, 29, 1295–1318, <https://doi.org/10.1175/WAF-D-13-00145.1>, 2014.
- 25 Scovell, R. and al Sakka, H.: A point cloud method for retrieval of high-resolution 3D gridded reflectivity from weather radar networks for air traffic management, *J. Atmos. Ocean Tech.*, 33, 461–479, <https://doi.org/10.1175/JTECH-D-15-0051.1>, 2016.
- Seifert, A., Köhler, C., and Beheng, K. D.: Aerosol-cloud-precipitation effects over Germany as simulated by a convective-scale numerical weather prediction model, *Atmos. Chem. Phys.*, 12, 709–725, <https://doi.org/10.5194/acp-12-709-2012>, 2012.
- 30 Sena, E. T., McComiskey, A., and Feingold, G.: A long-term study of aerosol–cloud interactions and their radiative effect at the Southern Great Plains using ground-based measurements, *Atmos. Chem. Phys.*, 16, 301–318, <https://doi.org/10.5194/acp-16-11301-2016>, 2016.
- Sheffield, A. M., Saleeby, S. M., and van den Heever, S. C.: Aerosol-induced mechanisms for cumulus congestus growth, *J. Geophys. Res. Atmos.*, 120, 8941–8952, <https://doi.org/10.1002/2015JD023743>, 2015.
- 35 Shipway, B. J. and Hill, A. A.: Diagnosis of systematic differences between multiple parametrizations of warm rain microphysics using a kinematic framework, *Q. J. R. Meteorol. Soc.*, 138, 2196–2211, <https://doi.org/10.1002/qj.1913>, 2012.

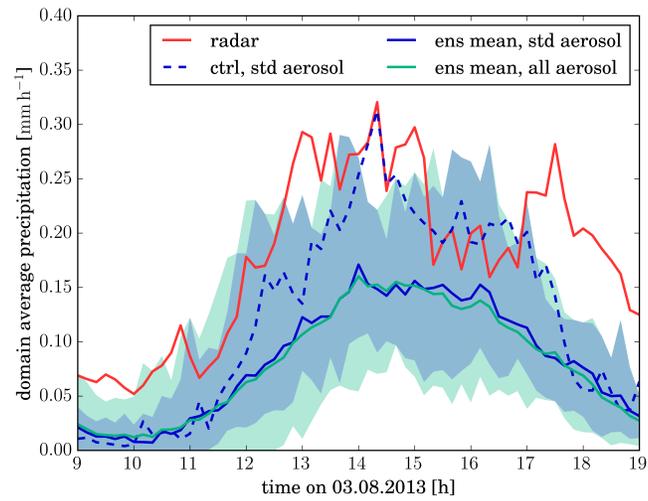
- Stevens, B. and Feingold, G.: Untangling aerosol effects on clouds and precipitation in a buffered system, *Nature*, 461, 607–613, <https://doi.org/10.1038/nature08281>, 2009.
- Tao, W.-K., Chen, J.-P., Li, Z., Wang, C., and Zhang, C.: Impact of aerosols on convective clouds and precipitation, *Rev. Geophys.*, 50, RG2001, <https://doi.org/10.1029/2011RG000369>, 2012.
- 5 Twomey, S.: The influence of pollution on the shortwave albedo of clouds, *J. Atmos. Sci.*, 34, 1149–1152, [https://doi.org/10.1175/1520-0469\(1977\)034<1149:TIOPOT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1977)034<1149:TIOPOT>2.0.CO;2), 1977.
- van den Heever, S. C., Stephens, G. L., and Wood, N. B.: Aerosol indirect effect on tropical convective characteristics under conditions of radiative-convective equilibrium, *J. Atmos. Sci.*, 68, 699–718, 2011.
- Walters, D., Boutle, I., Brooks, M., Melvin, T., Stratton, R., Vosper, S., Wells, H., Williams, K., Wood, N., Allen, T., Bushell, A., Copsey, D.,  
 10 Earnshaw, P., Edwards, J., Gross, M., Hardiman, S., Harris, C., Heming, J., Klingaman, N., Levine, R., Manners, J., Martin, G., Milton, S., Mittermaier, M., Morcrette, C., Riddick, T., Roberts, M., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Vidale, P. L., Wilkinson, J., Willett, M., Woolnough, S., and Xavier, P.: The Met Office Unified Model global atmosphere 6.0/6.1 and JULES global land 6.0/6.1 configurations, *Geosci. Model Dev.*, 10, 1487–1520, <https://doi.org/10.5194/gmd-10-1487-2017>, 2017.
- Wang, H., Auglign e, T., and Morrison, H.: Impact of microphysics scheme complexity on the propagation of initial perturbations, *Mon. Wea.*  
 15 *Rev.*, 140, 2287–2296, <https://doi.org/10.1175/MWR-D-12-00005.1>, 2012.
- Ward, J. H.: Hierarchical Grouping to Optimize an Objective Function, *J. Amer. Statist. Assoc.*, 58, 236–244, 1963.
- Warren, R. A., Kirshbaum, D. J., Plant, R. S., and Lean, H. W.: A ‘Boscastle-type’ quasi-stationary convective system over the UK Southwest Peninsula, *Q. J. R. Meteorol. Soc.*, 140, 240–257, <https://doi.org/10.1002/qj.2124>, 2014.
- Witte, M. K., Chuang, P. Y., and Feingold, G.: On clocks and clouds, *Atmos. Chem. Phys.*, 14, 6729–6738, [https://doi.org/10.5194/acp-14-](https://doi.org/10.5194/acp-14-6729-2014)  
 20 6729-2014, 2014.
- Wood, R., Stemmler, J. D., R emillard, J., and Jefferson, A.: Low-CCN concentration air masses over the eastern North Atlantic: Seasonality, meteorology, and drivers, *J. Geophys. Res.*, 122, 1203–1223, <https://doi.org/10.1002/2016JD025557>, 2017.
- Zeng, X., Tao, W.-K., Lang, S., Hou, A. Y., Zhang, M., and Simpson, J.: On the sensitivity of atmospheric ensembles to cloud microphysics in long-term cloud-resolving model simulations, *J. Meteor. Soc. Jap.*, 86A, 45–65, <https://doi.org/10.2151/jmsj.86A.45>, 2008.
- 25 Zhang, S., Wang, M., Ghan, S. J., Ding, A., Wang, H., Zhang, K., Neubauer, D., Lohmann, U., Ferrachat, S., Takemura, T., Gettelman, A., Morrison, H., Lee, Y. H., Shindell, D. T., Partridge, D. G., Stier, P., Kipling, Z., and Fu, C.: On the characteristics of aerosol indirect effect based on dynamic regimes in global climate models, *Atmos. Chem. Phys.*, 16, 2765–2783, <https://doi.org/10.5194/acp-16-2765-2016>, 2016.



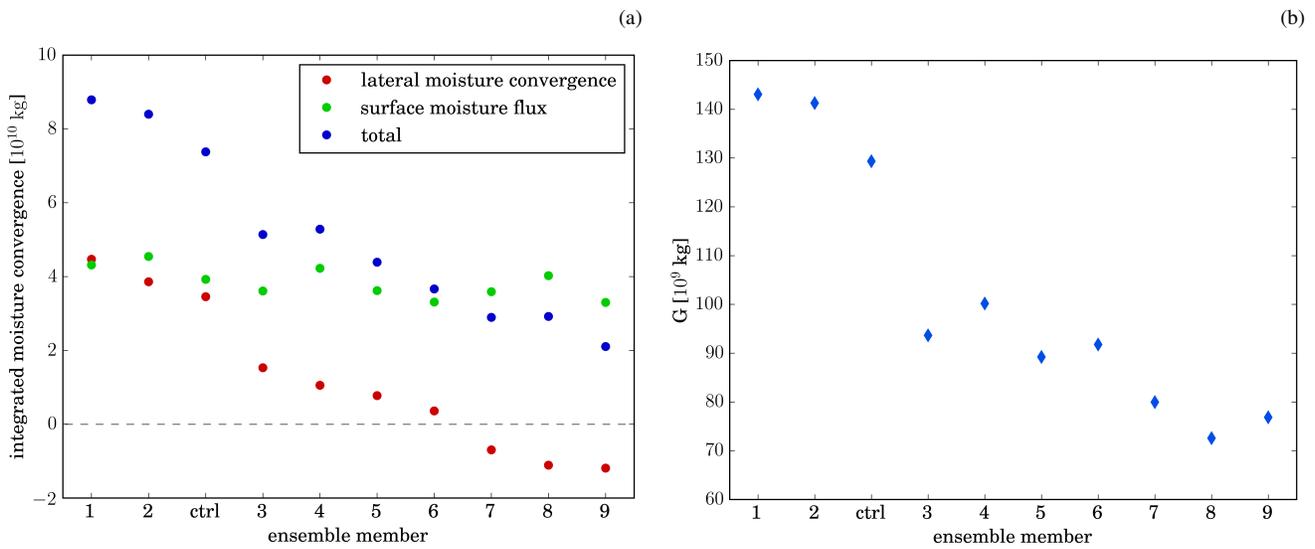
**Figure 1.** Convergence of moist static energy (a) and moisture (b) across the 1 km domain computed from the global ensemble. The grey lines show all 33 ensemble members in the global ensemble and the red lines the 9 members selected for the regional ensemble simulations. The selection procedure is described in section 2.



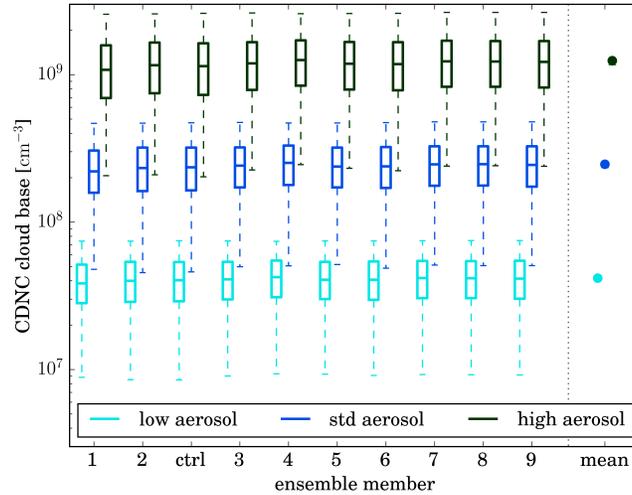
**Figure 2.** Column maximum radar reflectivity over 250 m domain at 1400 UTC from the control simulation (top left) and the 9 ensemble members using the standard aerosol profile.



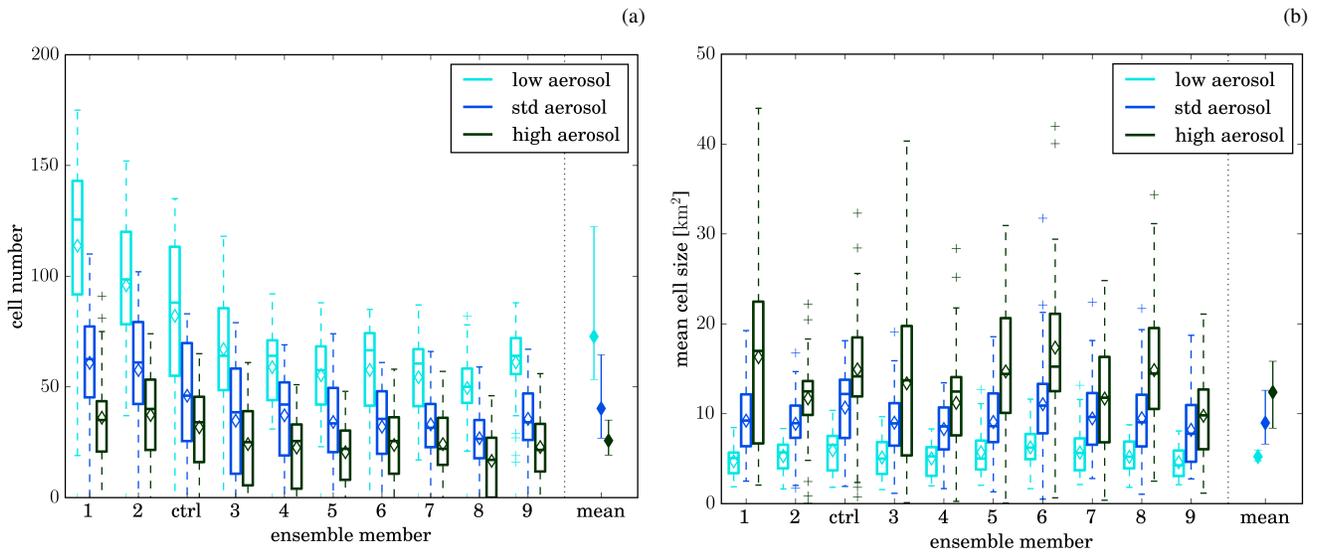
**Figure 3.** Comparison of domain-mean surface precipitation from model simulations and radar observations (red line). Values from the control simulation with the standard aerosol profile are shown by the dark blue dashed line. The mean (envelope) of all ensemble members using the standard aerosol profiles is shown by the dark blue solid line (shading) and those of all ensemble members irrespective of the used aerosol profile by the solid cyan line (shading).



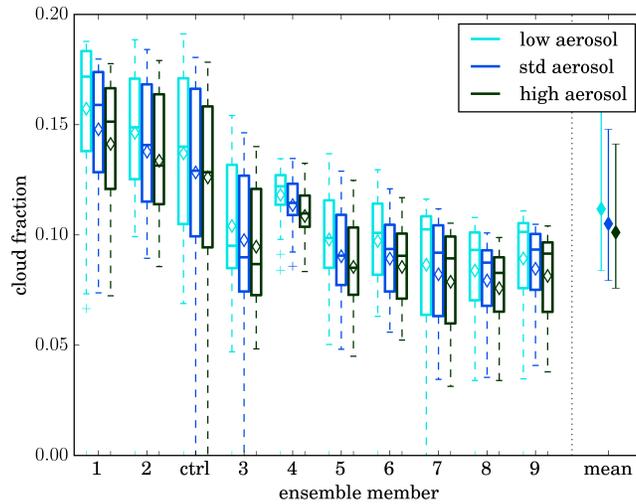
**Figure 4.** Panel (a) shows the time-integrated net (blue), lateral (red), and surface moisture flux (green) over the model domain in the boundary layer for each ensemble member. Panel (b) shows the time-integrated condensate gain  $G$  for the different ensemble members.



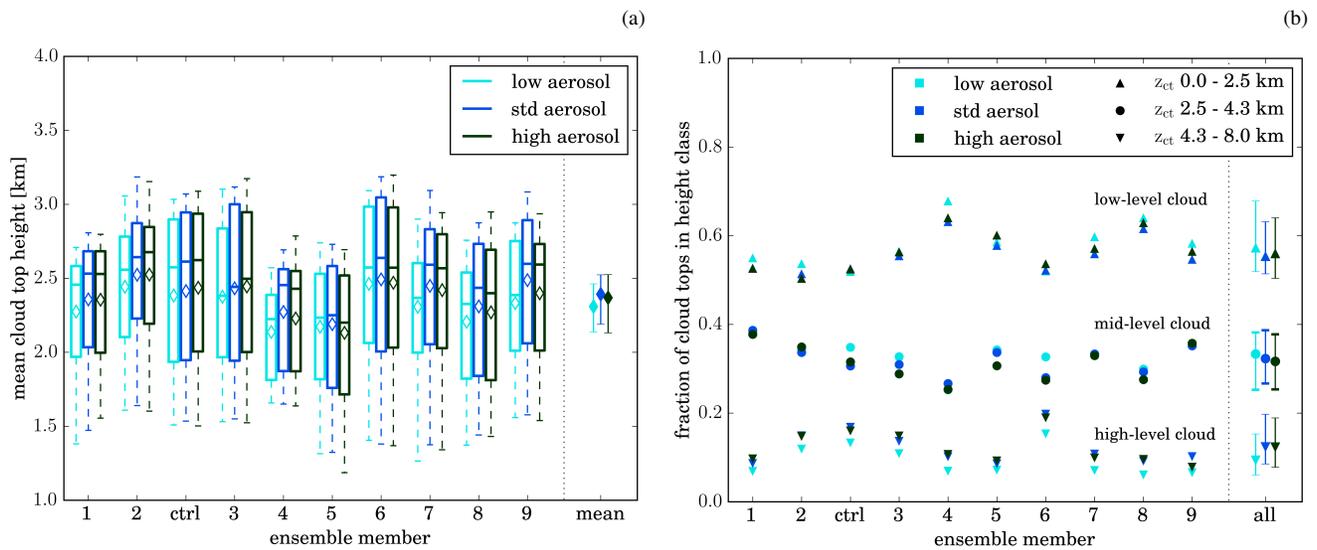
**Figure 5.** Cloud droplet number concentration (CDNC) at cloud base for different ensemble members (abscissa) using different aerosol profiles (colours). CDNC at cloud base is computed as the average CDNC within 500 above the lowest point in each grid column that has a cloud or ice mass mixing ratio larger than  $1 \text{ mg kg}^{-1}$ . The horizontal line inside the boxes indicates the mean CDNC, the upper and lower edges the 25<sup>th</sup> and 75<sup>th</sup> percentile, respectively, and the whiskers the 1<sup>st</sup> and 99<sup>th</sup> percentile. The statistics are computed over all qualifying grid points in the domain between 0900 UTC and 1900 UTC and therefore reflect the spatial and temporal variability of CDNC. The last column provides the distribution of the ensemble means, with the dot representing the average of the ensemble means and the bars the spread of the ensemble means.



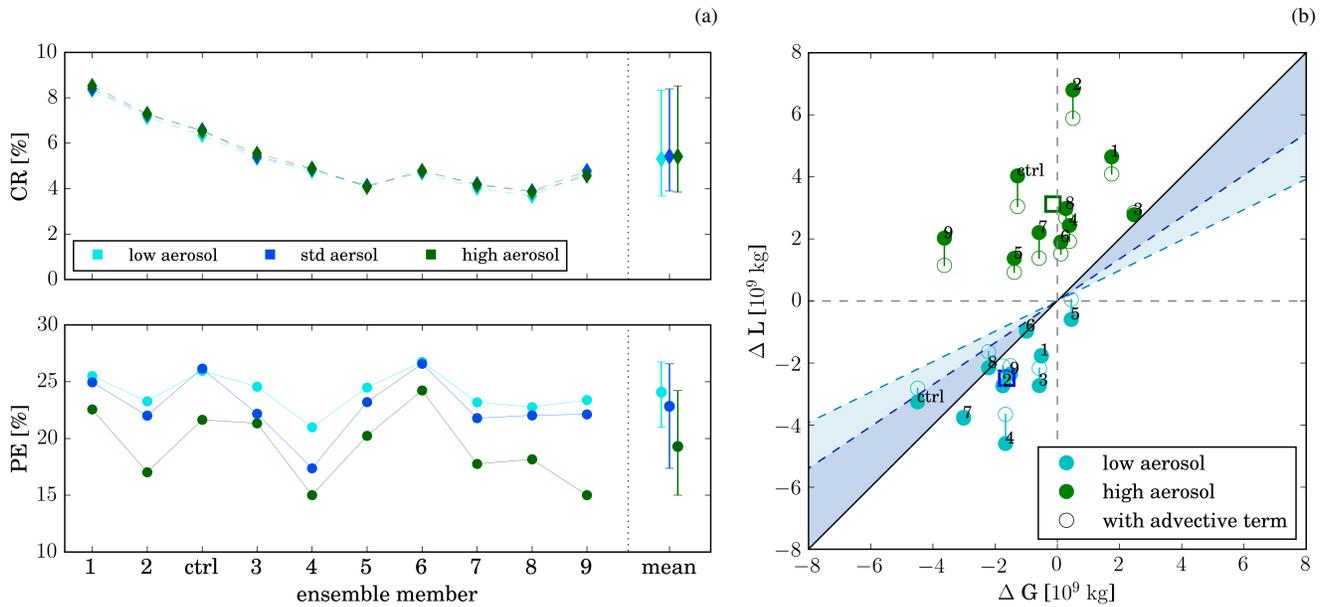
**Figure 6.** Cell number (a) and mean cell size (b). Cells are defined as continuous areas of column maximum radar reflectivity exceeding 25 dBZ. The horizontal line inside the boxes indicates the time mean value, the upper and lower edges the 25<sup>th</sup> and 75<sup>th</sup> percentile, respectively, and the whiskers the 1<sup>st</sup> and 99<sup>th</sup> percentile. These statistics reflect the temporal variability of the considered variables. The last column in each panels provides the distribution of the ensemble means, with the dot representing the average of the ensemble means and the bars the spread of the ensemble means.



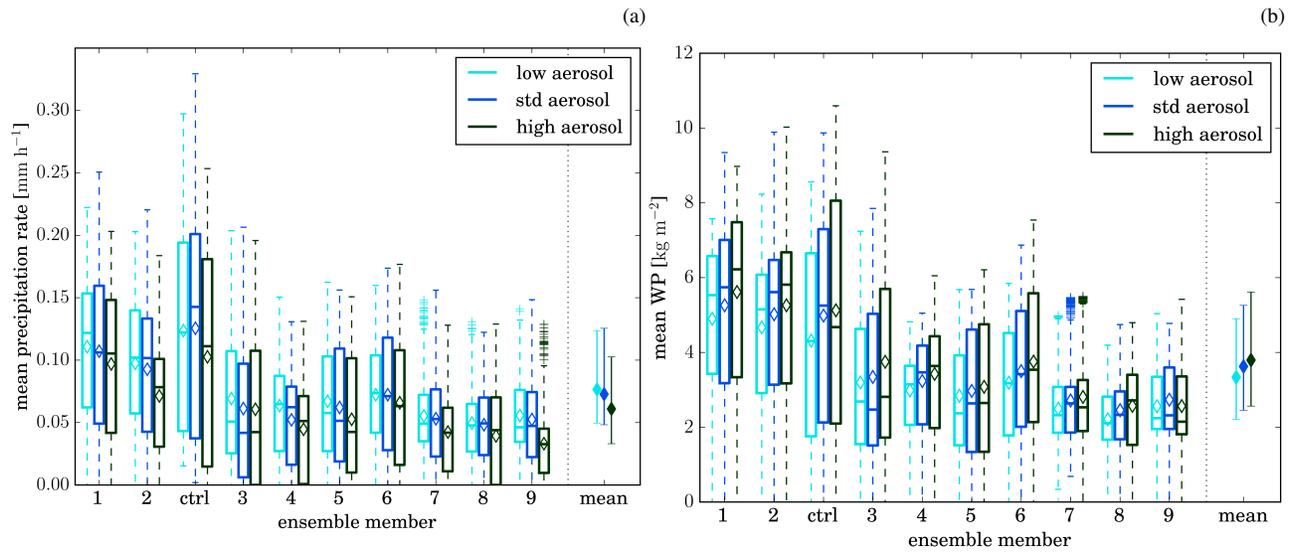
**Figure 7.** Cloud fraction in the different ensemble members. Cloud fraction is the fraction of the domain for which the condensed water path is larger than  $1 \text{ g m}^{-2}$ . The horizontal line inside the boxes indicates the time mean value, the upper and lower edges the 25<sup>th</sup> and 75<sup>th</sup> percentile, respectively, and the whiskers the 1<sup>st</sup> and 99<sup>th</sup> percentile. These statistics reflect the temporal variability of the considered variables. The last column in each panels provides the distribution of the ensemble means, with the dot representing the average of the ensemble means and the bars the spread of the ensemble means.



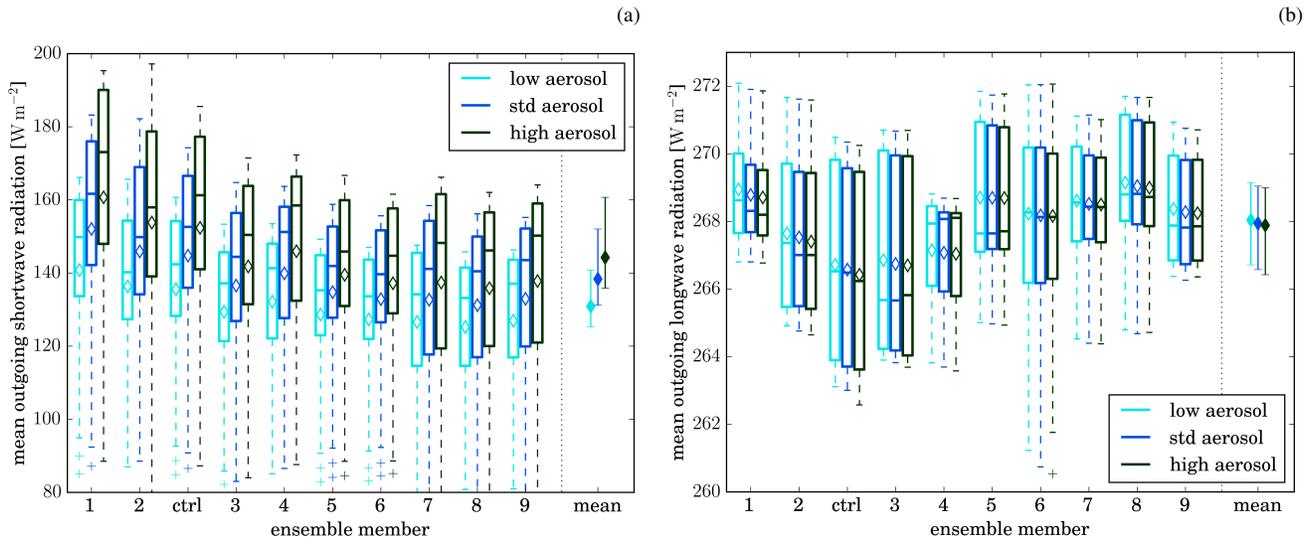
**Figure 8.** Mean cloud top height (a) and fraction of clouds with cloud tops specific altitude bands (b). Cloud top height is the height of the highest vertical level in each grid column with a condensate mass mixing ratio larger than  $1 \text{ mg kg}^{-1}$ . The horizontal line inside the boxes indicates the time mean value, the upper and lower edges the 25<sup>th</sup> and 75<sup>th</sup> percentile, respectively, and the whiskers the 1<sup>st</sup> and 99<sup>th</sup> percentile. These statistics reflect the temporal variability of the considered variables. The last column in each panels provides the distribution of the ensemble means, with the dot representing the average of the ensemble means and the bars the spread of the ensemble means.



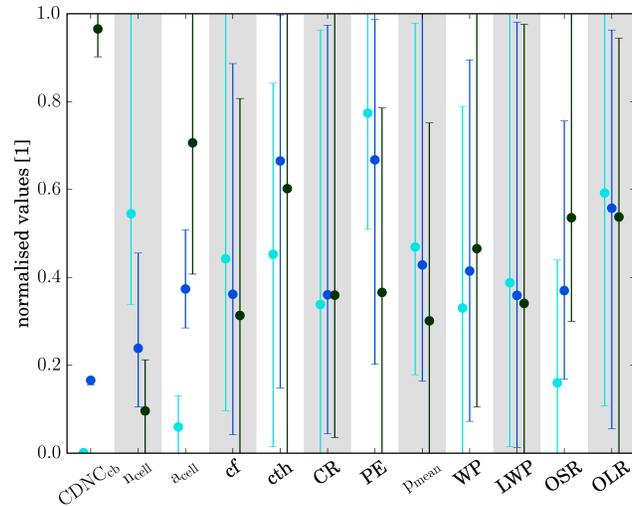
**Figure 9.** (a) Condensation ratio and precipitation efficiency for the different ensemble members and aerosol scenarios. The last column in each panel provides the distribution of mean values from each ensemble member: The dot represents the mean over all ensemble members and the bars represent the range between the largest and smallest mean value. (b)  $\Delta G$  in relation to  $\Delta L$  for ensemble members paired according to the meteorological initial conditions.  $\Delta G$  and  $\Delta L$  are computed for simulations with the high (green symbols) and low (cyan symbols) aerosol profile relative to the simulations with the standard aerosol profile. The closed symbols represent  $\Delta L$  and  $\Delta G$  values computed over the regional model domain, while the open symbols include advective fluxes of condensate at the domain boundary in the loss term  $\Delta L$ . The blue (cyan) shaded area indicates the region in the phase-space for which changes in  $\Delta G$  dominate the precipitation response using the minimum (maximum) precipitation efficiency from the ensemble with standard aerosol conditions. The open square shows the average response across the ensemble members.



**Figure 10.** Domain-average precipitation rate (a) and mean condensed water path (b). The horizontal line inside the boxes indicates the time mean value, the upper and lower edges the 25<sup>th</sup> and 75<sup>th</sup> percentile, respectively, and the whiskers the 1<sup>st</sup> and 99<sup>th</sup> percentile. These statistics reflect the temporal variability of the considered variables. The last column in each panels provides the distribution of the ensemble means, with the dot representing the average of the ensemble means and the bars the spread of the ensemble means.



**Figure 11.** Outgoing shortwave (a) and longwave (b) radiative flux at the top of the atmosphere, i.e.  $\approx 40$  km. The horizontal line inside the boxes indicates the time mean value, the upper and lower edges the 25<sup>th</sup> and 75<sup>th</sup> percentile, respectively, and the whiskers the 1<sup>st</sup> and 99<sup>th</sup> percentile. These statistics reflect the temporal variability of the considered variables. The last column in each panels provides the distribution of the ensemble means, with the dot representing the average of the ensemble means and the bars the spread of the ensemble means.



**Figure 12.** Summary of variability in time average (0900 – 1900 UTC) cloud properties induced by variations in meteorological initial conditions (bars) and aerosol initial conditions (colours; cyan: low aerosol scenario, blue: standard aerosol scenario, green: high aerosol scenario). Each variable has been normalised such that the minimum and maximum values in the entire ensemble (aerosol and meteorology) map to the value range  $[0, 1]$ . The variables displayed are cloud-base cloud droplet number  $CDNCC_b$ , number of cells  $n_{cell}$ , mean cell area  $a_{cell}$ , cloud fraction  $cf$ , mean cloud top height  $cth$ , condensation ratio  $CR$ , precipitation efficiency  $PE$ , average precipitation rate  $p_{mean}$ , mean condensed water path  $WP$ , liquid water path  $LWP$ , mean outgoing shortwave radiation  $OSR$ , and mean outgoing longwave radiation  $OLR$ .

**Table 1.** p-values from two-sided t-test with the null hypothesis of no change in the variable (rows) between two aerosol scenarios (columns) for all ensemble members. The results for ensemble members paired according to meteorological conditions and un-paired members are provided. **Bold numbers indicate statistical significance at the 5 %-level.**

	low - standard		standard - high		low - high	
	paired	unpaired	paired	unpaired	paired	unpaired
CDNC cloud base	<b>8.31e-16</b>	<b>1.23e-15</b>	<b>7.52e-15</b>	<b>1.09e-14</b>	<b>5.26e-15</b>	<b>6.15e-15</b>
cloud fraction	<b>0.00154</b>	0.643	<b>0.00310</b>	0.737	<b>0.00190</b>	0.431
cell number	<b>8.99e-6</b>	<b>0.00178</b>	<b>1.49e-5</b>	<b>0.00591</b>	<b>7.68e-6</b>	<b>8.33e-5</b>
cell area.	<b>3.94e-9</b>	<b>9.91e-9</b>	<b>6.03e-5</b>	<b>0.000326</b>	<b>1.05e-6</b>	<b>1.25e-6</b>
cloud top height	<b>0.000243</b>	0.325	0.549	0.678	0.104	0.204
deep cloud fraction	<b>2.61e-6</b>	0.222	0.465	0.914	<b>2.63e-5</b>	0.263
mean precipitation rate	<b>0.0123</b>	0.748	<b>0.000555</b>	0.313	<b>3.16e-5</b>	0.174
PE	<b>5.78e-3</b>	0.273	<b>1.31e-4</b>	<b>0.0145</b>	<b>1.32e-5</b>	<b>9.18e-4</b>
CR	<b>0.00140</b>	0.874	0.878	0.994	<b>0.0164</b>	0.8823
G	<b>0.00501</b>	0.896	0.803	0.991	<b>0.0363</b>	0.906
L	<b>1.15e-4</b>	0.794	<b>0.000190</b>	0.753	<b>2.08e-5</b>	0.571
P	<b>0.0144</b>	0.701	<b>0.000372</b>	0.248	<b>2.84e-5</b>	0.120
condensed WP	<b>0.000258</b>	0.5323	<b>0.00748</b>	0.730	<b>0.000176</b>	0.342
frozen WP	<b>1.13e-5</b>	<b>0.0159</b>	<b>0.000341</b>	0.222	<b>9.17e-6</b>	<b>0.00192</b>
liquid WP	<b>0.00450</b>	0.848	<b>0.0152</b>	0.905	<b>0.000477</b>	0.756
cloud WP	<b>2.34e-6</b>	0.144	<b>6.99e-6</b>	0.396	<b>2.81e-6</b>	<b>0.031</b>
OSR	<b>6.80e-7</b>	<b>0.0154</b>	<b>9.27e-7</b>	0.113	<b>7.63e-7</b>	<b>0.000799</b>
OLR	<b>8.33e-5</b>	0.817	<b>0.00576</b>	0.894	<b>0.000373</b>	0.717

**Table 2.** Number of observation days to obtain a statistically significant (at the 5 % level) aerosol signal in 95 % of all cases. The main value assumes the spread of the meteorological ensemble members equals  $4\sigma$ , while the values in brackets use  $5\sigma$  and  $3\sigma$ .

Variable	aerosol within factor 100 (low to high scenario)	aerosol factor 10 lower (low to standard scenario)	aerosol factor 10 higher (standard to high scenario)
CDNC	< 10 (< 10, < 10)	< 10 (< 10, < 10)	< 10 (< 10, < 10)
cloud fraction	90 (50, 140)	250 (130, 340)	480 (320, 860)
cloud top height	60 (40, 100)	110 (70, 190)	540 (350, 950)
outgoing SW	< 10 (< 10, < 10)	20 (< 10, 20)	30 (20, 50)
outgoing LW	460 (290, 810)	1110 (710, 1960)	3350 (2180, 6000)
ac. precipitation	30 (20, 60)	420 (290, 790)	60 (40, 90)