Reply to RC 1

*P6 L21-22: "**Nevertheless,* *previous evaluation studies of convection-permitting ensemble simulations have* *also* *reported precipitation forecasts to be underdispersive over longer evaluation periods (e.g. Romine et al., 2014; Schwartz et al., 2014) "*
P7 L24: "but the radiosonde passed through clouds"
P10 L16: "the variations... is dominating" -> "the variations... dominate"
*P 13 L29: you have a missing figure reference (**Fig. reffig:cdnc* *)*
SI P5 Fig 4: "red (orange) curves"

Reply: Corrected as suggested. Thank you!

SI P8 Fig 7a: On screen I find it very hard to distinguish between the colors used for 12 UTC and for 14 UTC. (Fig 7b looks much better).

Reply: The new version of the plot uses the same color scale as Fig. 7b.

SI P9 Fig 9b: what do the sizes of the circles represent?
SI P11 Fig 11c: what do the sizes of the circles represent?

Reply: The sizes of the circles in both Figures only vary to make it easier to distinguish points representing the same percentiles. We have clarified this in the caption.

SI P10 Fig 10: "variability"

Reply: Done.

Reply to RC 2

I am concerned because it seems that the authors have not corrected some of the errors found by the two reviewers in the first review iteration process. For instance, there are still some typos in the figure numbers, and there are still figures with lines connecting the dots of different ensemble members (which we agreed has no sense and that must be removed, e.g. Fig. 9a).

Reply: We have removed the lines in Fig. 9a now. The wrong figure references on p. 10 l. 11 and p. 13 l. 14 have been corrected.

Besides, this review iteration has been wearisome due to the fact that most of the pages and lines referenced in the answer to reviewer document are not correct (from the 12th answer until the end, the reference to page and line number are wrong), and because some of the changes in the main paper are not marked in green (e.g.
p. 2, l. 7: "received particular attention"
p. 2, l. 8: "strongly over the historic period"
p. 5, l. 6: "0000 UTC"
p. 5, l. 8: "horizontal"
p. 5, l. 19: "in this study is configured"
p. 9, l. 15: "4.2. Cloud-property variability" (title)
p. 9, l. 24: "(e.g., Fridlind et al., 2010)
p. 9, l. 25: "areal" and "(e.g., Grosvenor et al., 2017)"
p. 9, l. 27: "However"
p. 10, l. 25: "respectively"
p. 36, l. 1: "0900")
and none of the changes in the supplement, which makes me think perhaps many other changes were not marked neither.

Reply: Thank you very much for checking the supplied material so carefully! We are really sorry for the inconsistencies, but unfortunately cannot correct these in hindsight.

On top, the number of figures has not been reduced, as suggested. Instead some of the figures have been combined as sub-figures into single figures, sometimes without any compelling reason.

Reply: The total number of Figures as given by the figure numbers has not been reduced. However, if sub-panels are included the number of plots has been reduced from 29 (original manuscript) to 19 (revised version). This excludes Fig. 2, which has been added on request of RC 1. The number of figures is similar in both versions mainly because the previous Fig. 5 has been split into three different figures to add new Fig. 8b (on request of RC1) and still have a meaningful set of sub-panels in each figure. Previous Fig. 10 and Fig. 9 b, d have been combined into a single figure, as they are the only figures pertaining to the condensate budget analysis, and to minimise the number of figures in the manuscript.

Specific issues:

Although the authors used the significance results presented in Table 1 all over the article, as I suggested, I still have some considerations to add:

- It would be fair to remark that some statements are only valid for the paired ensemble members tests and add "and only for paired ensembles" to the sentences in:
p. 11, l. 29-30 ("The change of cloud top height is only significant for an increase in aerosol concentrations from the low to the standard scenario"),
p. 12 l. 8-9 ("Aerosol-induced modifications of CR and G are only significant for a decrease of aerosol concentrations relative to the standard scenario").
p. 14, l. 5 ("Aerosol-induced modifications to the outgoing radiative fluxes are significant at the 5 % level").

Reply: We do not think this makes sense, because the unpaired test is only introduced in section 6 and all the references given are from section 5. Also, it is very clearly stated in the introduction to section 5 that all statistical results in this section are from paired tests.

- I think they mean 95 % level of confidence instead of 5 %, since 5% level of confidence would be very low.

Reply: In the text we only use the significance level (and not the confidence level). The significance level is given correctly as 5 %.

Figure 1 provided in the "Reply to comments form reviewer #2" has the OLR_cc plot twice (bottom left and middle right), and the capture is not correct, since it refers to the mean cloud fraction (bottom right), which is not shown in the figure. Additionally, I recommend the authors not to label the figures this document in the same way as in the paper but with roman numbers or letters (e.g. Fig I, II… or Fig. a, b...) since it becomes very confusing to have figure numbers for the old version paper, for the new version and for the reply figures.

Reply: We are sorry for the inconsistency between the caption and the actual figure as well as the confusion caused by similar figure references in the reply document and the manuscript. However, we are not able to correct this in hindsight.

Furthermore, there are figures wrong referenced in the text, in particular "SI Fig. 12" in page 13 line 14 should read "SI Fig. 13". In fact, SI Fig. 13 is not cited at all in the text (I think it should be cited at least in p. 16 l. 3), as well as SI Fig. 14b and SI Fig. 15b which are not cited neither.

Reply: We corrected the reference (p. 13 l. 15) and added references to SI Fig. 13 on p. 16 l. 3 and 4, as suggested. Also, SI Fig. 14 b is now referenced on p. 12 l. 4 and p. 15 l. 17. SI Fig. 15 is referenced on p. 15 l. 24.

Some sentence citations of the altered new version paper text are not correctly written in the answers document ("Reply to comments form reviewer #2"):
- You said the altered text in p.10 l. 13-15 l. 18-20 is "[…] (SI Fig. 13)" and it says "[…] (SI Fig. 12)" in the paper text.
- You said you renamed section 5 as "Aerosol-induced cloud property changes in different meteorological ensemble members (paired meteorology)", but in fact the new title which appears in the paper text is "Cloud property changes between ensemble members in different aerosol and identical meteorological initial and boundary conditions".
- You said in p. 12 l. 21-23 l. 27-29 "[…] Fig. 12", but in fact in the text says "[...] Fig. 9B".
- Answering to my question regarding the PE change you refer me to Fig. 13 which it does not exist!
- Figure 3 of the "Reply to comments from reviewer #2" is not mentioned anywhere, I guess you identified it as "Fig. 4" in your reply.,

Reply: We are sorry for these inconsistencies!

Regarding the answer to my question "Page 13, lines 12-14: I do not think "all percentiles up to and including the 75 th 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", first of all, you answered that the text has been modified but the sentence you wrote in the answer document is not the same written in green in the paper. And secondly, I think the sentence is still not correct, since it is not true for ensemble number 7, consider changing it for "almost all ensemble members".

Reply: Thank you for pointing this out. We have modified the sentence as suggested (p. 13 l. 7/8).

After, you answered that "the aerosol direct effect is not included in the model simulations", then, I think it should be stated somewhere in the methodology since it has major effects on the radiative balance latter on.

Reply: We have added this to the "Data and Methods" section.

Some comments on the figures:

Figure 9:
I would change the sentence "The dots represents [...]" for "The dots and diamonds represent [...]".
I would change the words "open"/"closed" for "filled"/"unfilled".
Please add a filled symbol in the legend referring to "without advective term" simulations to make it complete.

Reply: Done.

Figure SI 4: reference to the orange curves is missing. Please, consider adding "(orange)" after "red".

Reply: Done.

Figure SI 12: it could be improved by plotting all three variables in a single plot instead of three (with different colors or symbols). Then, the three variables will be in the same x-y axis and the magnitude of each of them will be much more obvious than is now, and therefore better showing how the impact of aerosols in clear sky is prevailing over cloudy skies.

Reply: Done. Thank you for the suggestion.

Typos:

- p. 4, l. 31: "[…] from each cluster the member closest to the mean cluster […]" should read "[…] from each cluster the closest member to the mean cluster […]".
- p. 9, l. 29: reference to non-existing figure 10c should read "Fig. 10".
- p. 13, l. 29: "Fig. reffig:cdnc"

Reply: Done.

Regarding all the other technical issues and questions I am satisfied with the answers provided by the authors.

I recommend the article publication in ACP after minor revision.

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

Annette K. Miltenberger[1], Paul R. Field[1,2], Adrian A. Hill[2], Ben J. Shipway[2], and Jonathan
M. Wilkinson[2]

[1]Institute of Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, United Kingdom
[2]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.

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 acknowledge 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

5   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

10   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

15   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

20   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.

25   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 (3rd August 2013) deep convective clouds with maximum cloud top heights of about $5\,\text{km}$ developed in the late morning along converg-

30   ing 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 $250\,\text{m}$ 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

35   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

5 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 $0\,^{\circ}$C line. It was hypothesised that accompanying changes in the cloud field structure (fewer, larger cells with increasing aerosol) were important

10 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

15 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

20 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. $\approx 33\,\mathrm{km}$ in mid-latitudes). This includes stochastic physics

25 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

30 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 closest member 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

4

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. Note, that the aerosol direct effect is not inclued in the simulations. In all regional simulations, moisture conservation is enforced according to Aranami et al. (2014, 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, previous evaluation studies of convection-permitting ensemble simulations have also 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 ($750\,\mathrm{m\,agl}$) radar reflectivity agree well with the observed distribution (SI Fig. 2). In contrast to the domain-average precipitation time-series, the ensemble spread in the radar reflectivity distributions increases significantly if aerosol perturbations are considered in addition

6

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

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 $18\,\mathrm{dBZ}$, 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 $200\,\mathrm{m}$, SI Fig. 3). The inclusion of perturbed aerosol initial and boundary conditions has only a small impact on the ensemble mean height of the $18\,\mathrm{dBZ}$ contour (maximum difference: $\pm\,100\,\mathrm{m}$). Also, for other reflectivity thresholds ($5-25\,\mathrm{dBZ}$), the observed mean height is within the ensemble spread and the difference to the ensemble mean is generally smaller than $500\,\mathrm{m}$ (not shown).

3.2   Radiosonde data

Thermodynamic profiles are available at two hourly intervals from radiosondes released at Davidstow ($50.64\,^\circ\,\mathrm{N}$, $4.61\,^\circ\,\mathrm{W}$). 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 $550-400\,\mathrm{hPa}$. Also, the observed dewpoint temperature profile is generally contained within the ensemble spread, with the exception of a lower observed humidity below $900\,\mathrm{hPa}$ at 1520 UTC. All ensemble members have a stable layer between $5-6\,\mathrm{km}$ altitude, which is an important factor for the cloud top height distribution (Miltenberger et al., 2018). Ensemble members differ mainly in the humidity above $600\,\mathrm{hPa}$, with the altitude of the driest point in this layer varying by about $100\,\mathrm{hPa}$.

The height of the $0\,^\circ$ 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 $300\,\mathrm{m}$ for the $0\,^\circ$ level height and $400\,\mathrm{m}$ 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 $0\,^\circ$ level is generally outside the ensemble spread (except at 1350 UTC, but the radiosonde passed through clouds).

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. 4 a, 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 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 $10 \, \mathrm{m}$ 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 $0.9 - 1.4 \, \mathrm{K}$ in the morning to $1.8 - 2.0 \, \mathrm{K}$ by noon. Only in ensemble members 1 and 2 the temperature difference remains smaller than $1.5 \, \mathrm{K}$ (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.

8

The wind speed in the boundary layer varies between about $8\,\mathrm{m\,s^{-1}}$ and $11\,\mathrm{m\,s^{-1}}$, and increases to values of $13 - 18\,\mathrm{m\,s^{-1}}$ at $4\,\mathrm{km}$ 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 ($\approx 0.5\,\mathrm{K}$) nor the variability in the low-level wind speed ($\approx 3\,\mathrm{m\,s^{-1}}$) 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 $800 \pm 200\,\mathrm{m}$, an almost moist-adiabatic temperature gradient up to $500\,\mathrm{hPa}$, and a layer of almost constant temperature between $500\,\mathrm{hPa}$ and $450\,\mathrm{hPa}$. As a result of the small variation in the temperature profile, the average and maximum CAPE values are similar for all ensemble members ($100 - 160\,\mathrm{J\,kg^{-1}}$, SI Fig. 8). Also, variations in the moisture content are small with difference between ensemble members smaller than $0.5\,\mathrm{g\,kg^{-1}}$ for all altitudes. The altitude of the driest point in the profile varies by about $100\,\mathrm{hPa}$ 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 at cloud base 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 $25\,\mathrm{dBZ}$. Cloud top height is defined by the highest model level with a condensed water content larger than $1\,\mathrm{mg\,kg^{-1}}$ (e.g., Fridlind et al., 2010). Cloud fraction is calculated as the areal fraction of the domain with condensed water path larger than $0.001\,\mathrm{kg\,m^{-2}}$ (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 $750\,\mathrm{m}$ between ensemble members (Fig. 8 a) 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): 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 general decreases

9

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 $2.5\,\mathrm{km}$ 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 $4.3\,\mathrm{km}$. 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 a and b). 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 (Fig. 10 b), i.e. the condensed water path decreases in members with smaller moisture convergence and G.

Mean reflected shortwave radiation ranges from $130\,\mathrm{W\,m^{-2}}$ to $155\,\mathrm{W\,m^{-2}}$ (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 dominate the variability of outgoing shortwave radiation. Changes in outgoing longwave radiation are on the order of $3\,\mathrm{W\,m^{-2}}$ (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 $4.3\,\mathrm{km}$. 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-troposheric 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).

11

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 and b). 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 ΔG (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 ΔG 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 ΔG and condensate loss ΔL 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 ΔG 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 ΔG (control) or ΔG and ΔPE 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 75th percentile from the low to the high aerosol concentration for almost all ensemble members (SI Fig. 11 c). The 99th 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 99th 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. 13), consistent with a slower conversion of cloud droplets to rain drops and accordingly a larger mass transport across the $0\,^{\circ}\mathrm{C}$ 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 $4.3\,\mathrm{km}$ 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. 5) 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.

13

The outgoing longwave radiation is mainly influenced by the surface temperature, the cloud fraction, and the cloud top temperature. 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 ($< 0.5 \, \mathrm{W \, m^{-2}}$, standard to high aerosol scenario) is consistent with the small increase in mean cloud top height (Fig. 8 a).

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.

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 and b). 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). 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 b). 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).

15

The ensemble-mean condensed water path increases with increasing aerosol concentrations, if all hydrometeor types are considered (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 (SI Fig. 13 a). The frozen water path increases with increasing aerosol concentrations for most ensemble members (SI Fig. 13 b) consistent with a longer timescale for precipitation formation. Theaerosol-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

16

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

5   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.

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

10   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 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

15   (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 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

20   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 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.

25   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 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.

30   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 used as the mean of the Gaussian distribution. The standard deviation σ 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

35   of samples n can be interpreted as the number of times the same day is observed, as the statistical analysis presented here uses

17

either daily average or accumulated variables. However, it may be possible to interpret the necessary number of observations 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

5 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,

10 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

15 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,

20 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.

25 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

30 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

35 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 thermodynamic 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

5   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 geo-

10   graphic 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).

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

15   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.

19

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.

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.

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.

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.

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, 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.

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, 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, 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, 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., Choularton, 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, 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.

22

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.

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, 10 503–10 523, https://doi.org/10.5194/acp-11-10503-2011, 2011.

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, 10 504–10 511, 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.

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.

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.

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.

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, 11 301–11 318, 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.

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, 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., Augligné, 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-14-20   6729-2014, 2014.

Wood, R., Stemmler, J. D., Rémillard, 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.
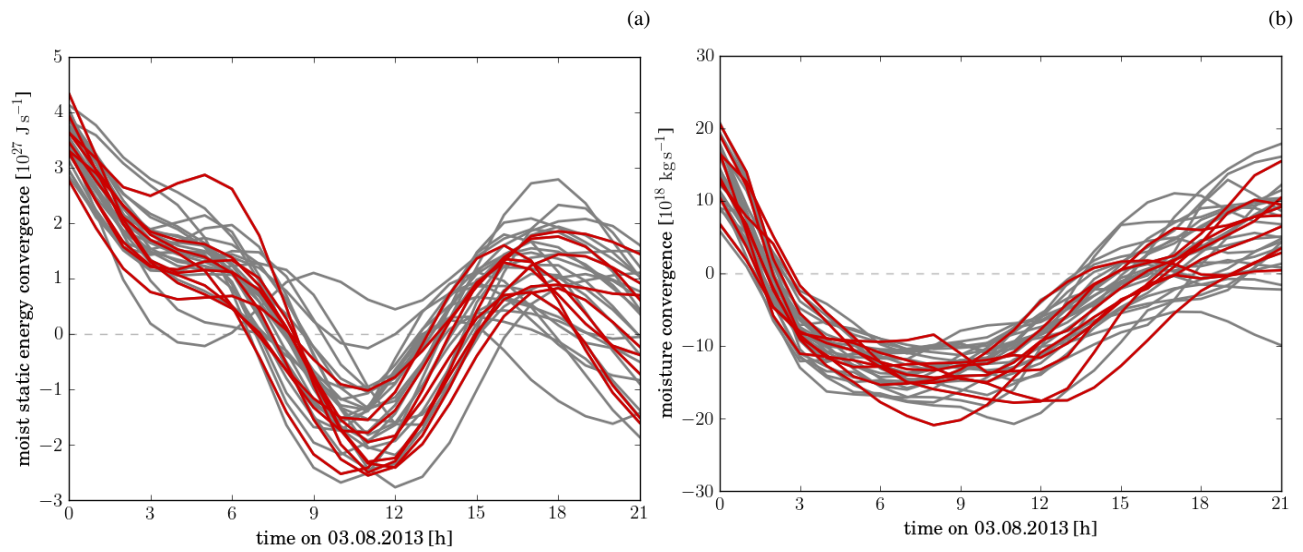
Figure 1. Convergence of moist static energy (a) and moisture (b) across the 1 km domain computed from the global ensemble. The grey lines show all 33 ensemble members in the global ensemble and the red lines the 9 members selected for the regional ensemble simulations. The selection procedure is described in section 2.
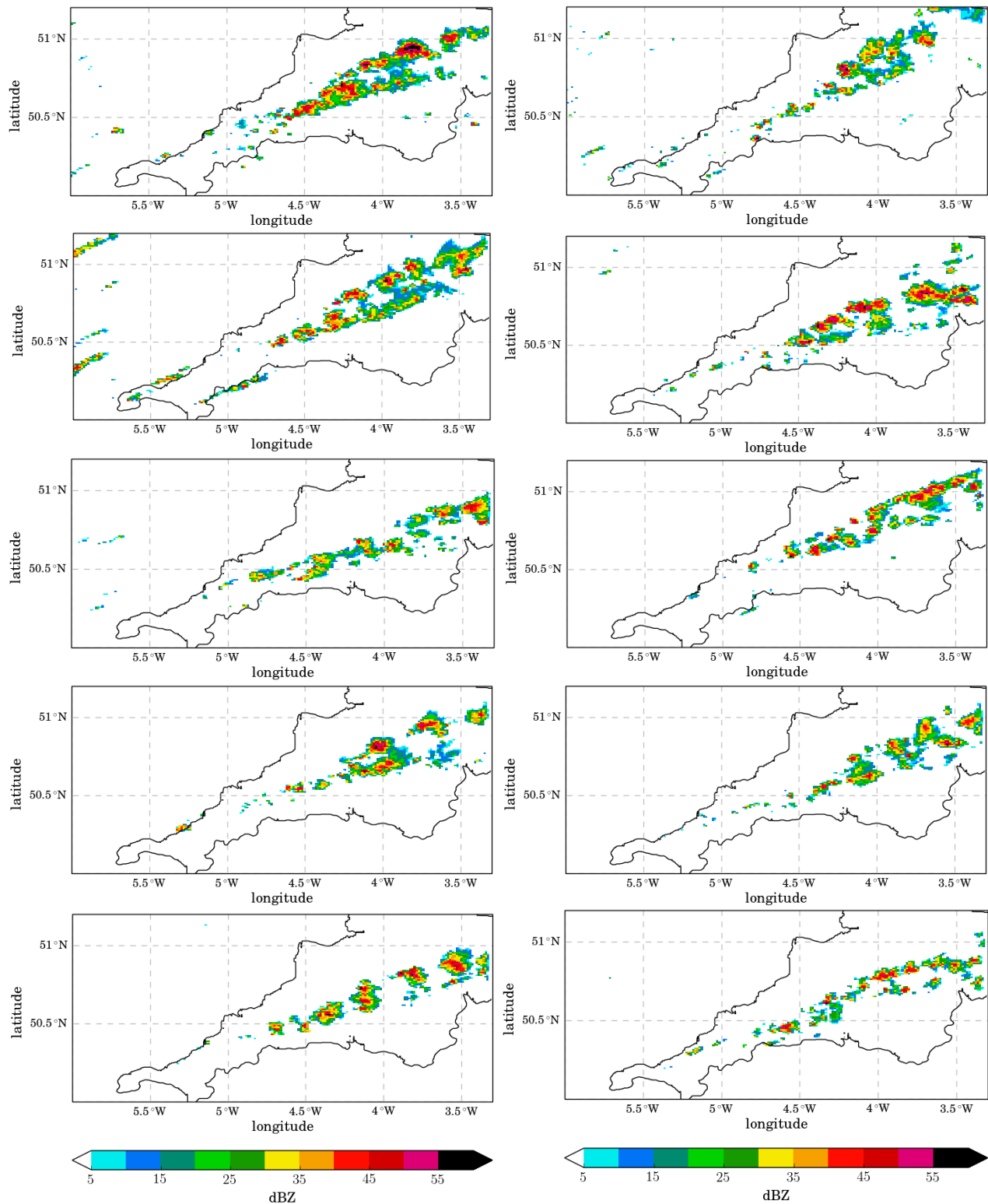
Figure 2. Column maximum radar reflectivity over 250 m domain at 1400 UTC from the control simulation (top left) and the 9 ensemble members using the standard aerosol profile.
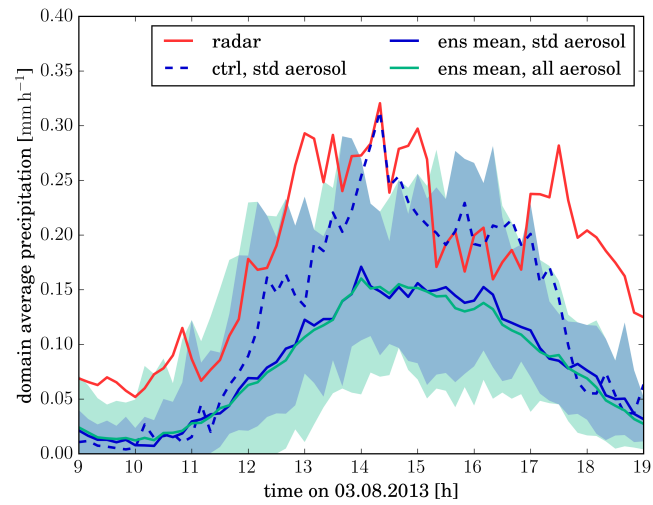
Figure 3. Comparison of domain-mean surface precipitation from model simulations and radar observations (red line). Values from the control simulation with the standard aerosol profile are shown by the dark blue dashed line. The mean (envelope) of all ensemble members using the standard aerosol profiles 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).
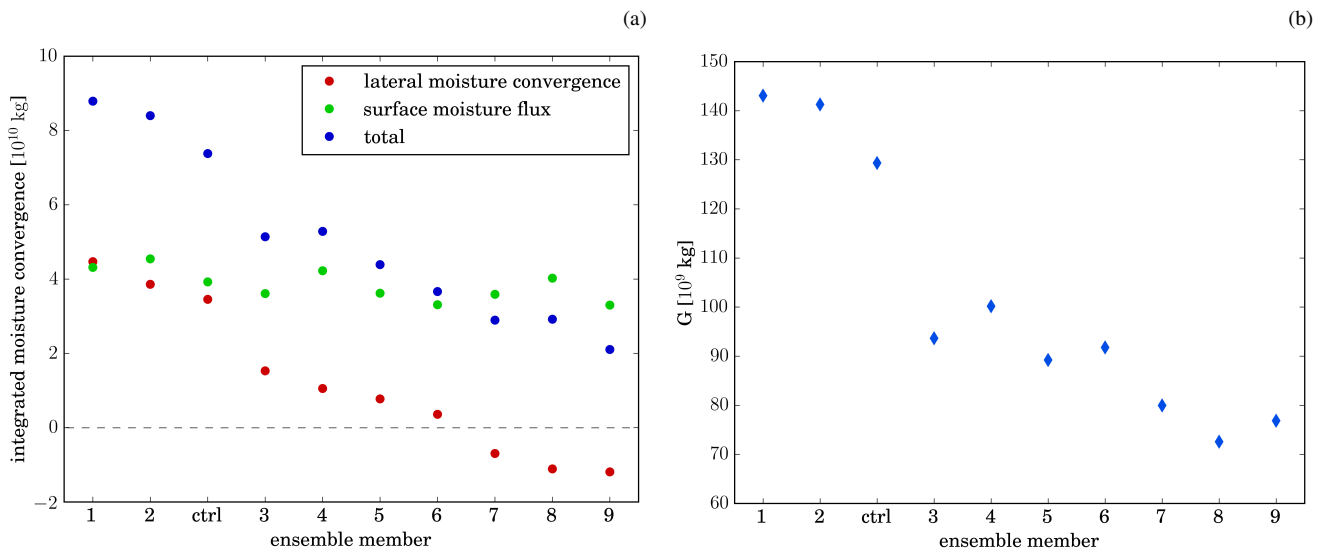
27

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.
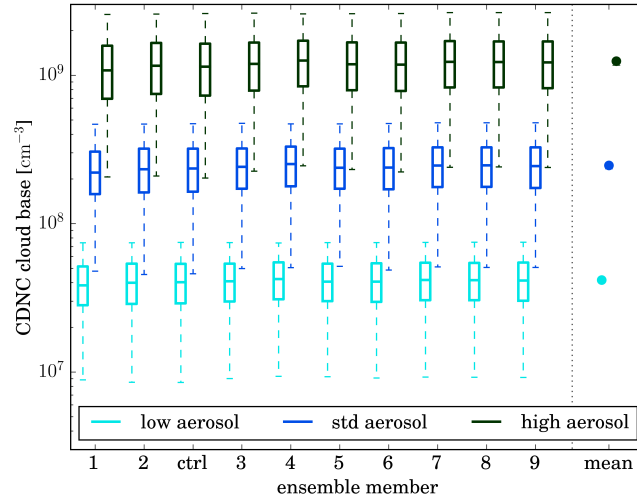
Figure 5. Cloud droplet number concentration (CDNC) at cloud base for different ensemble members (abscissa) using different aerosol profiles (colours). CDNC at cloud base is computed as the average CDNC within 500 above the lowest point in each grid column that has a cloud or ice mass mixing ratio larger than $1\,\mathrm{mg\,kg^{-1}}$. The horizontal line inside the boxes indicates the mean CDNC, the upper and lower edges the 25th and 75th percentile, respectively, and the whiskers the 1st and 99th 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.
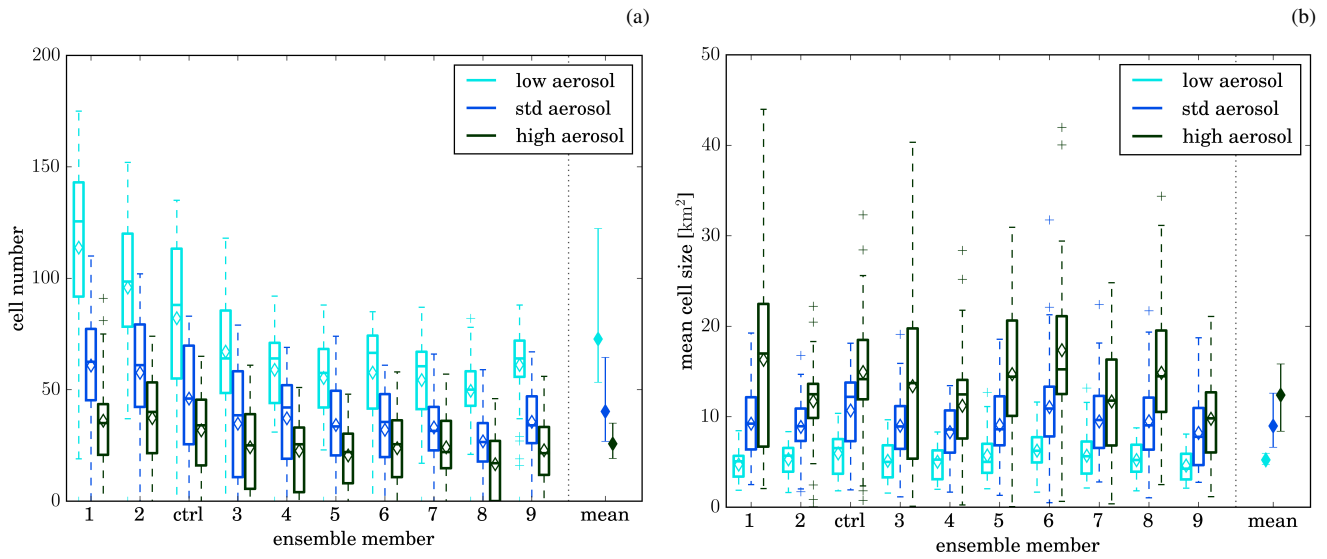
Figure 6. Cell number (a) and mean cell size (b). Cells are defined as continuous areas of column maximum radar reflectivity exceeding $25\,\mathrm{dBZ}$. The horizontal line inside the boxes indicates the time mean value, the upper and lower edges the 25th and 75th percentile, respectively, and the whiskers the 1st and 99th 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 $1\,\mathrm{g\,m^{-2}}$. The horizontal line inside the boxes indicates the time mean value, the upper and lower edges the 25th and 75th percentile, respectively, and the whiskers the 1st and 99th 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.
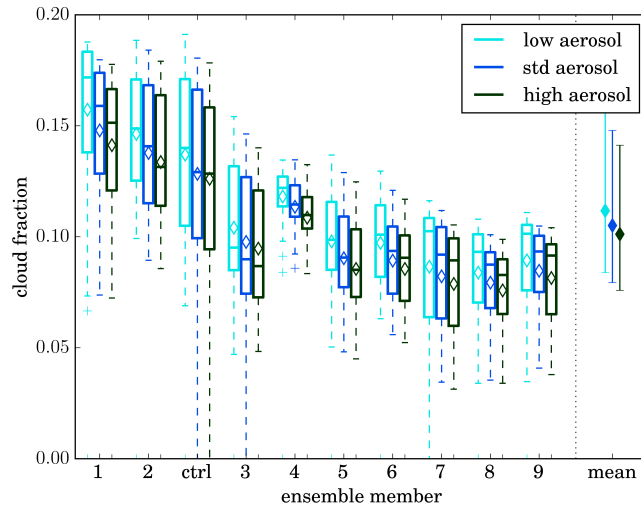
Figure 8. Mean cloud top height (a) and fraction of clouds with cloud tops specific altitude bands (b). Cloud top height is the height of the highest vertical level in each grid column with a condensate mass mixing ratio larger than $1\,\mathrm{mg\,kg^{-1}}$. The horizontal line inside the boxes indicates the time mean value, the upper and lower edges the 25th and 75th percentile, respectively, and the whiskers the 1st and 99th 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.
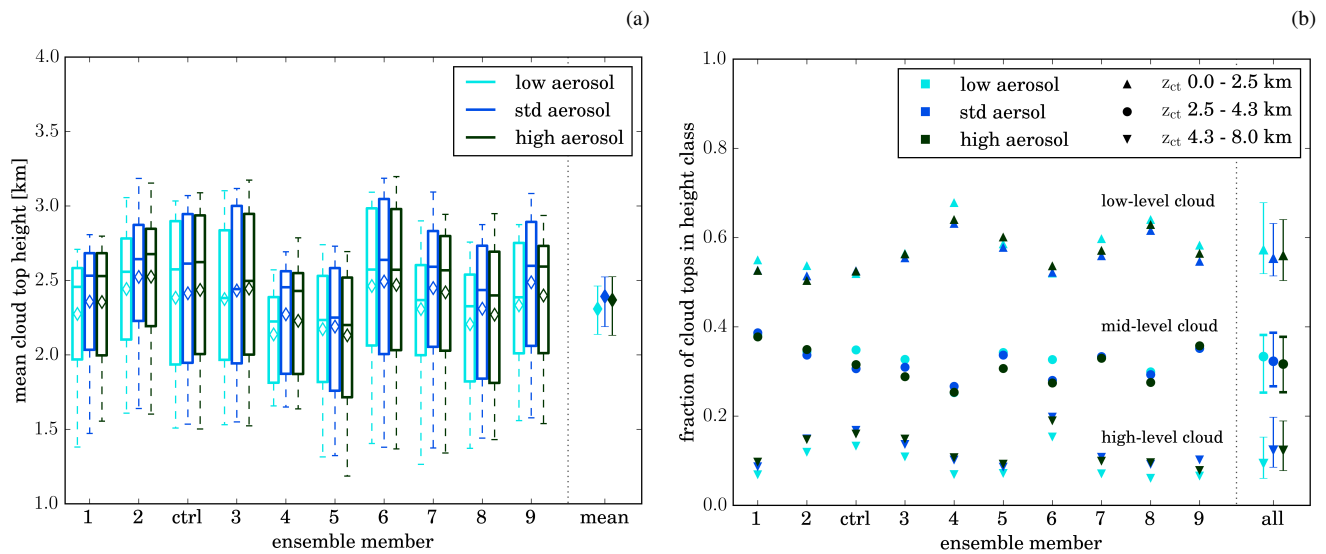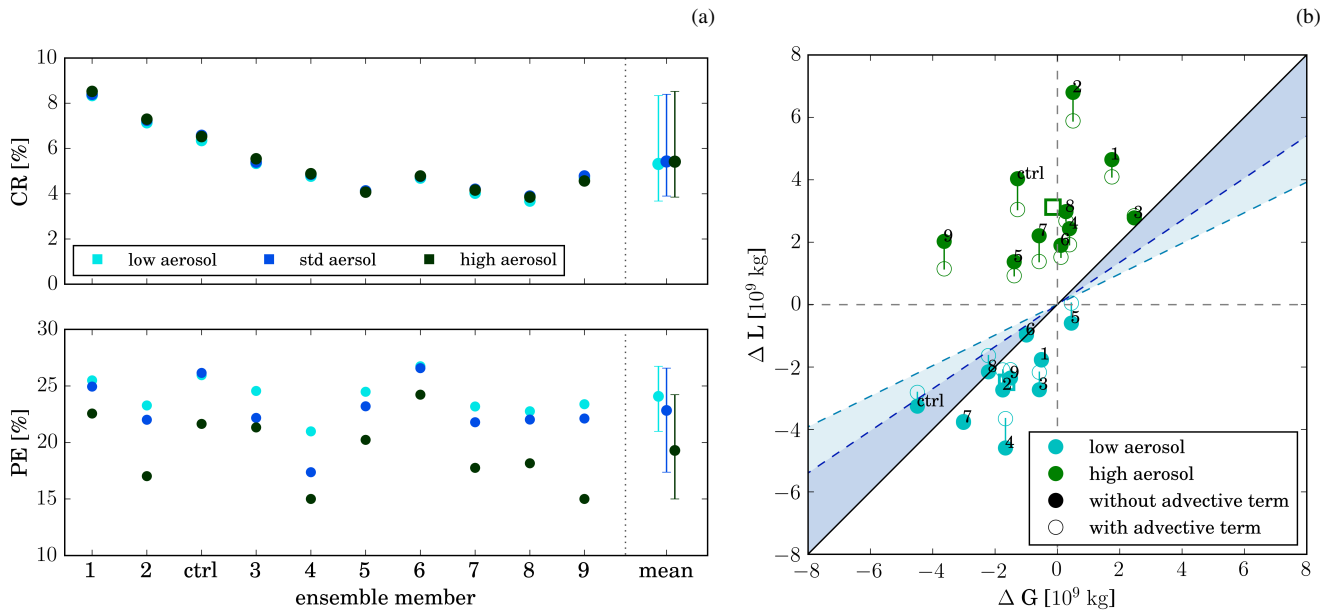
Figure 9. (a) Condensation ratio and precipitation efficiency for the different ensemble members and aerosol scenarios. The last column in each panel provides the distribution of mean values from each ensemble member: The dot represents the mean over all ensemble members and the bars represent the range between the largest and smallest mean value. (b) ΔG in relation to ΔL for ensemble members paired according to the meteorological initial conditions. ΔG and ΔL 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 filled symbols represent ΔL and ΔG values computed over the regional model domain, while the unfilled symbols include advective fluxes of condensate at the domain boundary in the loss term ΔL. The blue (cyan) shaded area indicates the region in the phase-space for which changes in ΔG dominate the precipitation response using the minimum (maximum) precipitation efficiency from the ensemble with standard aerosol conditions. The unfilled square shows the average response across the ensemble members.
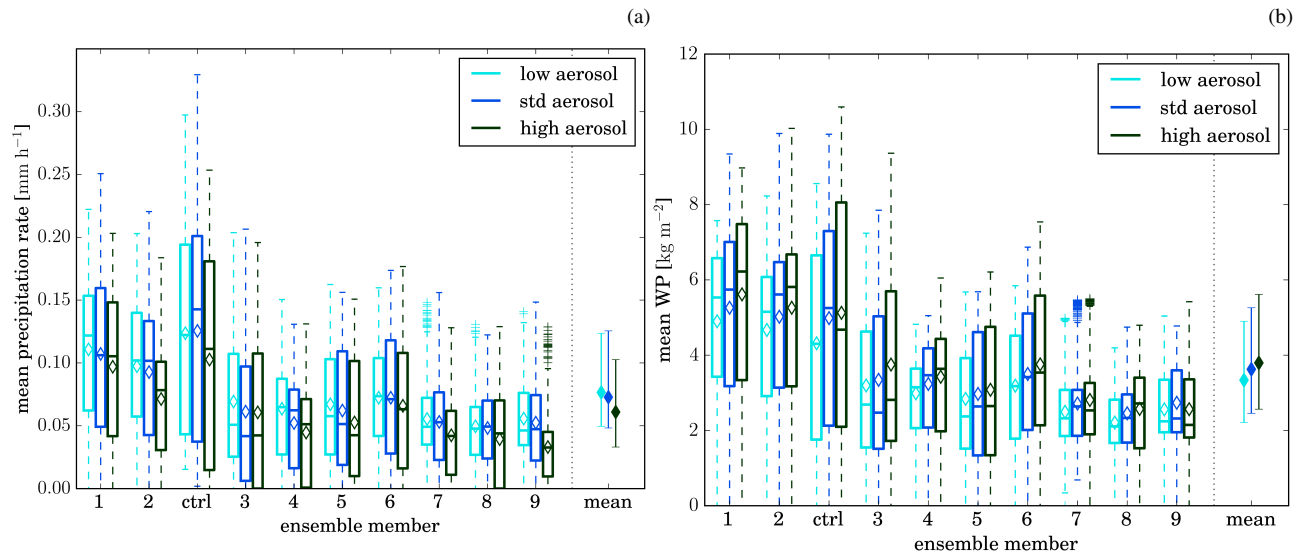
33

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 25th and 75th percentile, respectively, and the whiskers the 1st and 99th 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. ≈ 40 km. The horizontal line inside the boxes indicates the time mean value, the upper and lower edges the 25th and 75th percentile, respectively, and the whiskers the 1st and 99th 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.
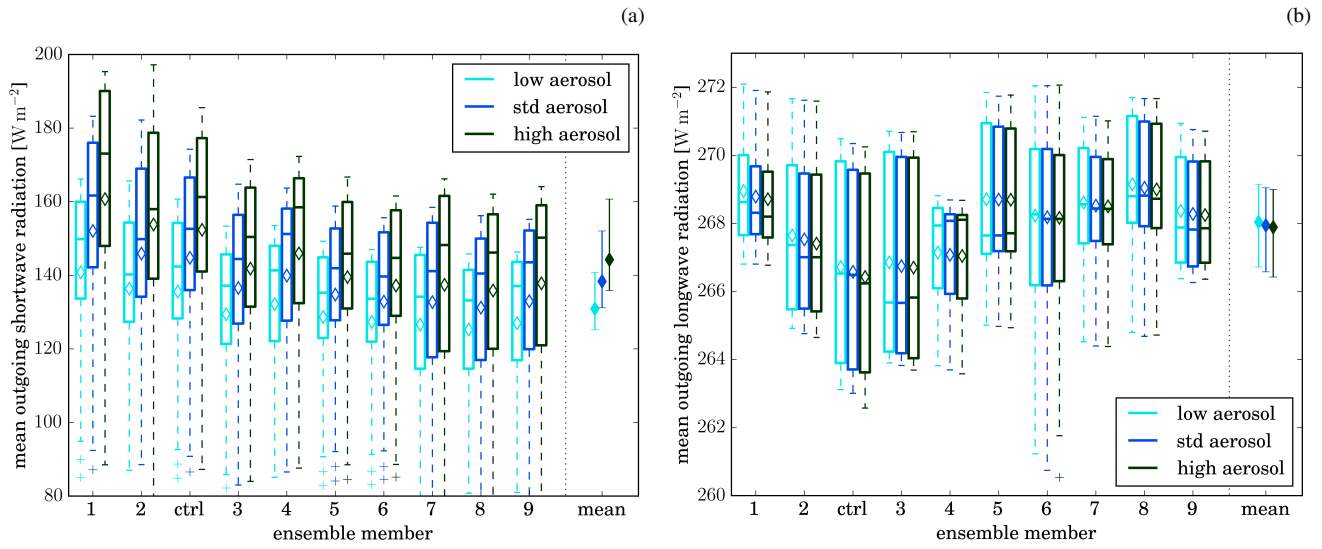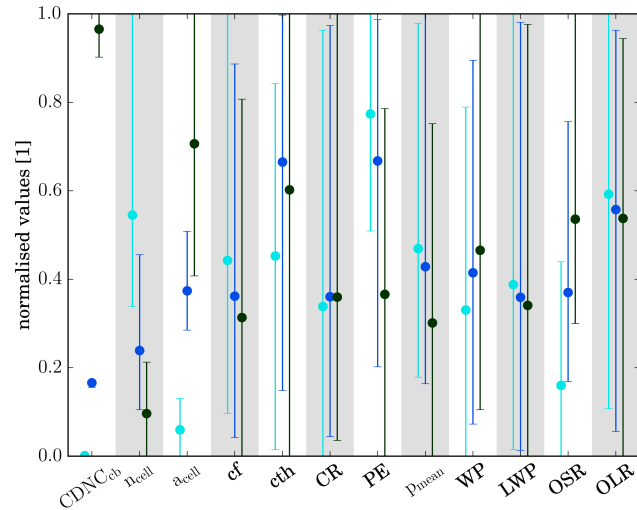
Figure 12. Summary of variability in time average $(0900 - 1900\ \text{UTC})$ cloud properties induced by variations in meteorological initial conditions (bars) and aerosol initial conditions (colours; cyan: low aerosol scenario, blue: standard aerosol scenario, green: high aerosol scenario). Each variable 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 CDNC_{cb}, 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	**8.31e-16**	**1.23e-15**	**7.52e-15**	**1.09e-14**	**5.26e-15**	**6.15e-15**
cloud fraction	**0.00154**	0.643	**0.00310**	0.737	**0.00190**	0.431
cell number	**8.99e-6**	**0.00178**	**1.49e-5**	**0.00591**	**7.68e-6**	**8.33e-5**
cell area.	**3.94e-9**	**9.91e-9**	**6.03e-5**	**0.000326**	**1.05e-6**	**1.25e-6**
cloud top height	**0.000243**	0.325	0.549	0.678	0.104	0.204
deep cloud fraction	**2.61e-6**	0.222	0.465	0.914	**2.63e-5**	0.263
mean precipitation rate	**0.0123**	0.748	**0.000555**	0.313	**3.16e-5**	0.174
PE	**5.78e-3**	0.273	**1.31e-4**	**0.0145**	**1.32e-5**	**9.18e-4**
CR	**0.00140**	0.874	0.878	0.994	**0.0164**	0.8823
G	**0.00501**	0.896	0.803	0.991	**0.0363**	0.906
L	**1.15e-4**	0.794	**0.000190**	0.753	**2.08e-5**	0.571
P	**0.0144**	0.701	**0.000372**	0.248	**2.84e-5**	0.120
condensed WP	**0.000258**	0.5323	**0.00748**	0.730	**0.000176**	0.342
frozen WP	**1.13e-5**	**0.0159**	**0.000341**	0.222	**9.17e-6**	**0.00192**
liquid WP	**0.00450**	0.848	**0.0152**	0.905	**0.000477**	0.756
cloud WP	**2.34e-6**	0.144	**6.99e-6**	0.396	**2.81e-6**	**0.031**
OSR	**6.80e-7**	**0.0154**	**9.27e-7**	0.113	**7.63e-7**	**0.000799**
OLR	**8.33e-5**	0.817	**0.00576**	0.894	**0.000373**	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)