The authors would like to thank Dr. Matsui for his helpful and constructive suggestions. We have addressed each of his comments below. The reviewer comments are in black font, while the authors' responses are in blue for contrast. The * next to line numbers refers to the line number from the revised manuscript with track changes.

Summary:

This study conducted hundreds of sensitivity experiments of idealized cloud-resolving simulations in order to understand the effect of environmental parameters upon aerosol-sea-breeze convection interactions in tropics. Overall, set up of comprehensive sensitivity experiments, and statistical analysis (statistical emulation and variance-based analysis) are appealing aspects of this manuscript. However, the problem of this manuscript is that the figures are not summarizing and highlighting the physics very well. Although physics explanations are all reasonable, it's hard to extract essence from the figures. More specific major comments are described below. This paper has quite potential if revision goes well. So, I request "major revision" at this point.

Major Comments:

Parameter ranges (Table 2): I understand that some of them are derived from previous studies (Igel et al. 2018, Park et al. 2020). Are these perturbations in a realistic range? What do these perturbed ranges statistically mean? For example, soil saturation fraction between 0.1 and 0.9 are ranges from Savannah to tropical rainforest. Is this range typically happening in the real world of tropical coast regions? This question is also related to analysis of Fig 9a and 9b. When you compare different environmental factors, you should understand the natural ranges of these parameters, and should normalize/standardize them. Otherwise, you cannot state soil moisture has the largest impact on aerosol-cloud interactions.

Following one is my old paper that compared aerosol and thermodynamic impacts on low clouds by measuring 95%-frequency ranges of aerosol index and lower-tropospheric stability (Fig 3) in order to discuss relative importance.

Matsui, T., H. Masunaga, S.M.Kreidenweis, R.A.Pielke, Sr, W.-K. Tao, M. Chin, Y. Kaufman (2006), Satellite-based assessment of global warm cloud properties associated with aerosols, atmospheric stability, and diurnal cycle, *Journal of Geophysical Research– Aerosol and Clouds*. 111, D17204, doi:10.1029/2005JD006097.

You don't need to re-set new ranges of parameters for another hundreds of simulations, because you can just use a statistical emulator to estimate the relative impact of different parameters in standardized range. But, you have to understand statistical distributions of these parameters in the real world to understand "typical (one/two standard deviation)" ranges. With the standardized ranges of environmental parameters, you can state which parameters are important or not.

This is a great point. However, when we first started these studies, we could find very little in the literature regarding the maximum and minimum values of the parameters impacting sea breeze circulations, let alone their distributions. This is one of the reasons we performed these idealized tests across a wide range of values, as the emulator can then be used to select appropriate responses to appropriate parameter values. This said, we certainly strived to select a range of reasonable values, as described in Igel et al. (2018). We added the following statement.

> *Lines 117–121*: The range in the selected variables were sourced from the sea breeze literature, the reasons for which were described in detail in Igel et al. (2018). While statistical distributions of the parameters of interest are very difficult to find in the literature, plausible parameter ranges for tropical regions were assigned to each of the variables tested. Should new observations provide greater constraints on the range of these parameters, the emulator approach (described below) allows for an assessment of the responses of the sea breeze convective system to the range of parameter values of interest.*

We added also the following sentence in section 5.2.2 regarding the soil moisture values and why we tested 0.1 to 0.9 soil saturation fraction.

> *Lines 385–390*: It should be noted that sandy clay loam's observed soil saturation fraction varies from 0.25 to 0.75 along coastal equatorial Africa in June, July, and August (Rodell et al., 2004). Here we have extended the range tested to 0.1 through 0.9 to encompass slightly drier and wetter soil conditions in addition to those reported by Rodell et al. (2004) to take into account potentially more extreme conditions anticipated with changing climates. The corresponding variance-based sensitivity analyses for the daytime cumulus convection stratified by these three different soil regimes are shown in the second to fourth bar graphs from the left in Figures 9a and 9b.*

Section 5.1 and 5.2 (Figures 5-7): I don't quite understand why you plot clean-polluted differences in zig-zag form, because simulation ID in X-axis does not represent physics at all. There should be a more effective way to represent this statistical representation. For example, histograms (clean, polluted, and clean-polluted) would be better to represent statistical differences, distributions, and significance of these sensitivity overall. Same issue also applies to Fig 10, too.

Thank you for this great suggestion. We initially plotted these figures as a function of simulation ID as in that way we could track the variability in aerosol cloud interactions as a function of each specific environment (identified by simulation ID) and hence address our question of environmental modulation of aerosol effects. However, we ultimately decided to remove the detailed environmental information and hence we agree that it no longer makes sense to plot these as a function of simulation ID. We have now revised Figures 5–7 and Figure 10 as you have suggested. These revised figures are shown below, along with the new text relevant to these figures.

In all of these figures, we constructed histograms for rOn-500 and rOn-2000, as well as the rOn-2000 minus rOn-500 differences, in order to emphasize aerosol-induced changes. For the histograms of rOn-2000 minus rOn-500, we added a zero vertical line to make it clear as to whether the aerosol impacts produced a positive or negative response. We also added blue (rOn-500) and red (rOn-2000) vertical lines on the histograms to show the ensemble-median value of corresponding characteristics. The bin width of each histogram is marked on each panel. We have edited the associated text in the manuscript as follows:

> *Lines 242–254*: These reductions in sensible and latent heating will negatively impact the convective boundary layer by limiting both the heating and moistening of this layer. Specifically, less moisture will be available for convection via evapotranspiration, evaporation, and condensation, which is reflected in the reduction of the ensemble-median surface latent heat flux in the polluted case (Figure 5a), as well as the differences between the polluted and clean ensembles, by far the majority of which are negative (Figure 5b). The surface-based mixed layer depth, defined here as the level above the surface at which the vertical gradient of the potential temperature first exceeds 2 K km^{-1}, decreases in rOn-2000 compared with rOn-500 due to this reduction in surface sensible heat flux and associated turbulent mixing. Histograms of the mean surface based-mixed layer depth ahead of the sea breeze front in rOn-500 and rOn-2000 and their differences are shown in Figures 5c and 5d, respectively. In Figure 5c, the mean mixed layer depth distribution shifts toward lower values with the change from rOn-500 to rOn-2000. This aerosol-induced decrease in mixed layer depth is also evident in the reduced ensemble-median values (vertical lines in Figure 5c) in rOn-2000 compared with rOn-500. Figure 5d further indicates that the mixed layer in each member of the rOn-2000 ensemble is shallower than the corresponding mixed layer in the rOn-500 ensemble, with all of the rOn-2000 minus and rOn-500 values being negative.*

> *Lines 255–266*: While mixed layer depth is a valuable indicator of instability in the boundary layer and hence the depths of shallow cumulus, CAPE is a more pertinent assessment of instability for the deep convective clouds driven by the sea breeze convergence. As shown in Figure 5e, most of the simulations in both ensembles have averaged mixed-layer CAPE values close to zero which is in keeping with the fact that only a handful of the ensemble members produce deep convection. The ensemble-median values are slightly reduced with enhanced aerosol loading, from 7.6 to 7.1 J kg^{-1}. Figure 5f also demonstrates that the differences in CAPE between rOn-500 and rOn-2000 may be positive or negative but are mostly quite small in magnitude. The exceptions to this are the magnitudes for those cases that produce deep convection, where the CAPE values may be as high as 2102 J kg^{-1} and exceed marginal CAPE. The aerosol-induced differences are all negative and range in magnitude from 8 to 115 J kg^{-1}, except for one ensemble pair with a positive difference of 50 J kg^{-1}. Therefore, while the variations in CAPE with aerosol loading appear to be small in magnitude for most members of the ensembles, they may play a discriminating role in aerosol impacts on deep convective updraft velocities for those cases that do support deep convection. This is discussed further in section 5.2.1.*

Lines 266–285: We now turn our attention to the vertical lift provided by the convergence along the sea breeze front in all of the simulations. Classical sea breeze theory dictates that, to first order, the faster the sea breeze moves, the further inland the sea breeze travels during the day, the stronger the convergence along the sea breeze front, and hence the greater the vertical lift along the front. Here we examine the maximum inland extent of the sea breeze front, as well as the maximum updraft velocity found within $\pm$ 1 km of the algorithm identified surface location of sea breeze front during the afternoon (1200 LT–1800 LT). The maximum inland extent of the sea breeze front is identified as the last inland location of the sea breeze front detected by the sea breeze front algorithm (Igel et al., 2018). We assess the maximum vertical velocity within $\pm$ 1 km of the surface location of the front in order to account for any forward bulging or backward tilting of the frontal boundary in relation to the identified location of the front at the surface. The distribution of the maximum sea breeze inland extent shows a shift towards lower values with enhanced aerosol loading (Figure 5g), evident in the decrease in the ensemble median with enhanced aerosol loading (Figure 5g). It is also obvious from Figure 5h that the sea breeze extent is less in rOn-2000 than rOn-500 for each and every one of the ensemble pairs, thus demonstrating the significant role of aerosol loading and the direct effect on this baroclinic circulation, and the subsequent forcing of deep convection. The distribution of the maximum updraft velocities (and hence lift) along the sea breeze front shows a shift towards reduced updraft velocities in more polluted environments, as is demonstrated by the small reduction in the ensemble median of rOn-2000 compared with rOn-500. However, in spite of the robust response of the maximum inland extent of the sea breeze to aerosol loading (Figure 5h), the impacts of enhanced aerosol loading on the maximum frontal velocities do not always produce a negative vertical velocity response (Figure 5j). This suggests that while the environment does not appear to modulate the direct impacts of aerosols on the sea breeze dynamics and inland extent, it may locally modulate aerosol impacts on the updraft velocities, possibly through aerosol indirect processes and/or changes to CAPE.*

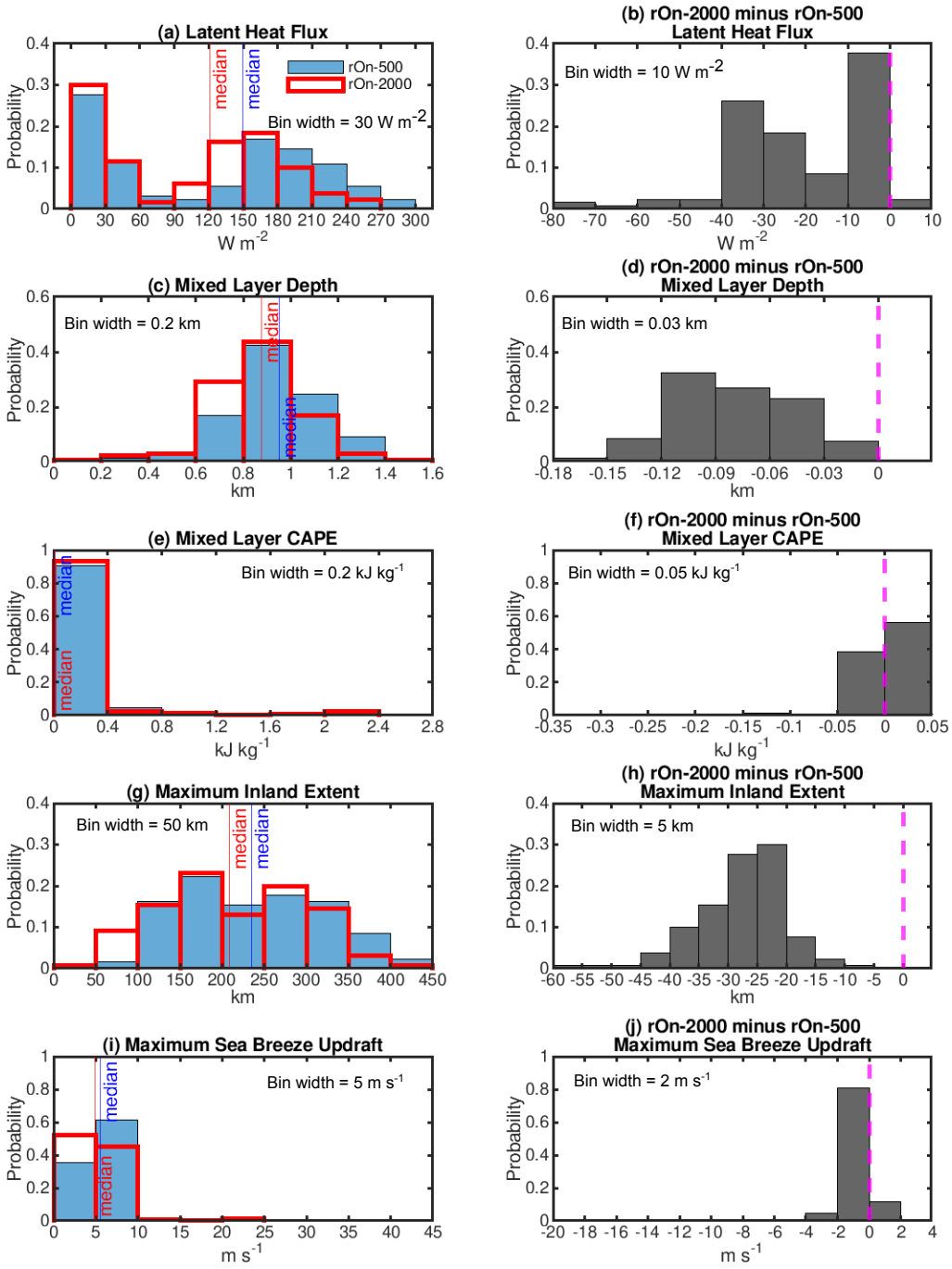

Figure 5. Histograms of (a) the land-surface latent heat flux (W m^{-2}), (c) the surface-based mixed layer depth (km), and (e) the mixed-layer CAPE (kJ kg^{-1}), all of which are averaged for each ensemble from the western domain edge to 50 km ahead of the algorithm-identified sea breeze front between 1200–1800 LT; (g) the maximum inland extent of the sea breeze front (km); and (i) maximum updraft velocities at the algorithm-identified sea breeze front$\pm$1 km (m s^{-1}). The median values of each characteristic are marked by the red and blue thin vertical lines. The light blue shading and the red lines represent rOn-500 and rOn-2000 ensembles, respectively. Figures (b, d, f, h, and j) are histograms of the differences in the corresponding fields shown in (a, c, e, g, and i) arising due to aerosol loading (rOn-2000 minus rOn-500). The dashed magenta lines indicate where the difference between rOn-2000 and rOn-500 is zero. The bin width of each histogram is marked at the upper corner of each panel.

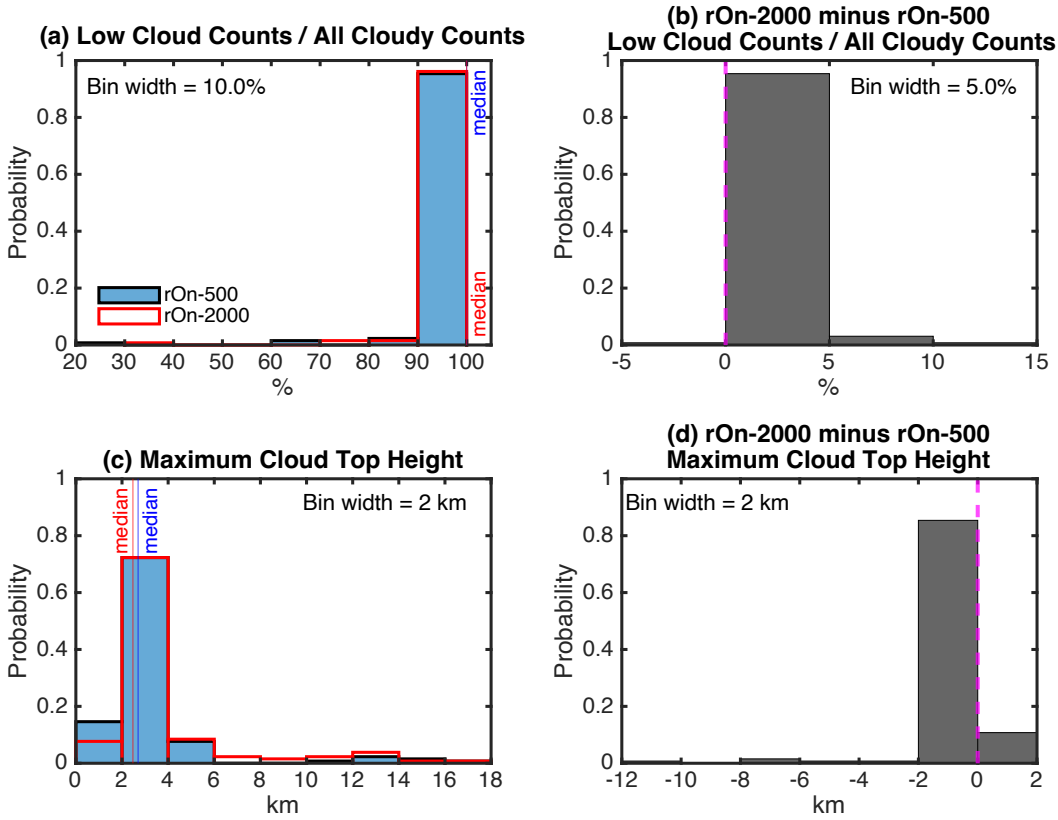

Figure 6. Similar to Figure 5 but for histograms of the (a) low cloud (cloud top height < 4 km) fractional contribution to the total number of cloudy columns (low cloud columns / all cloudy columns) and (c) the maximum cloud top height. The right column (b and d) represents rOn-2000 minus rOn-500 difference histograms for the corresponding fields (a and c) in the left column. The dashed magenta lines indicate where the difference between rOn-2000 and rOn-500 is zero. The median values of each characteristic are marked with vertical lines in the left column. The bin width of each histogram is marked at the upper corner of each panel.

Lines 300–309: We examine the impacts of aerosol loading on cloud top height in two ways. First, we examine the frequency distribution of the fraction of low (cloud top height < 4 km) cloudy columns to the total number of cloudy columns for all simulations (low cloud columns / all cloudy columns) (Figure 6a). Among the 130 simulations in each ensemble, there are 104 and 113 simulations with low clouds only in rOn-500 and rOn-2000, respectively. The vast majority of ensemble members in both ensembles are therefore dominated by low clouds, as demonstrated by the ensemble-median values of 100% (Figure 6a). In other words, only shallow convective clouds develop both ahead of and along the sea breeze front in most of the ensemble members, and while all of the environmental conditions tested here support the development of the sea breeze-initiated shallow convective mode, most do not support the development of the sea breeze-initiated deep convective mode. The difference between the two ensembles (Figure 6b) shows that in the majority of the simulations, the low cloud fraction stays the same or is weakly increased with enhanced aerosol loading.*

Lines 310–312: Second, we analyse the distribution of maximum cloud top heights (Figure 6c), and the differences as a result of aerosol loading (Figures 6d). The maximum cloud top height is determined during the afternoon hours (1200-1800 LT) anywhere over land.*

Lines 318–328: The suppression of sea breeze convective intensity in rOn-2000, when compared with rOn-500, is evident in the reduction in the ensemble-median of the maximum cloud top height (Figure 6c). Negative values in Figure 6d imply that the maximum cloud top height decreases with enhanced aerosol loading in the vast majority of simulations, most of which apply to low clouds (< 4 km AGL). However, there are some cases in which the cloud top heights increase in the presence of enhanced aerosol loading (Figure 6d). It is evident from Figure 6c that most of these enhancements in cloud top height with aerosol loading occur in association with the deep convective mode (> 7 km AGL). This is in spite of the fact that while there are 12 cases with a deep convective mode in rOn-500, only 7 of these 12 cases have a deep convective mode when aerosol loading is enhanced in rOn-2000. Altogether, enhanced aerosol loading results in reduction in cloud top height of the low clouds (< 4 km AGL), but in a mixed response in the deep convective mode. As such, it appears that the impacts of increased aerosol loading on shallow cloud top heights are relatively robust and occur independently of the initial environment, whereas aerosol impacts on the deep convective cloud top heights vary as a function of the environment, and hence are environmentally modulated.*

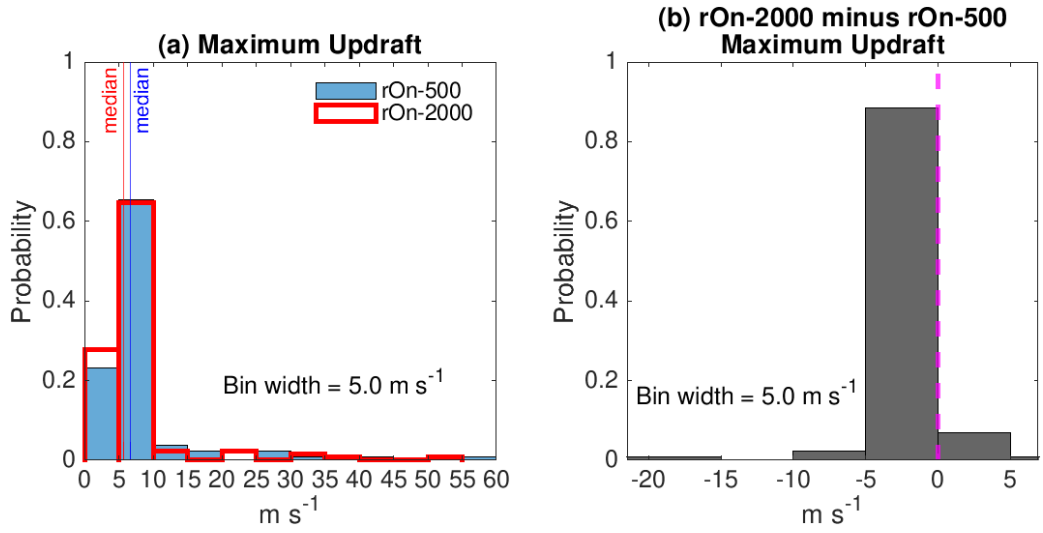

Figure 7. Similar to Figure 5 but for (a) histograms of the maximum updraft velocity in rOn-500 and rOn-2000 and (b) a histogram of the maximum updraft velocity differences arising from aerosol loading (rOn-2000 minus rOn-500).

Lines 335–338: Overall aerosol-induced suppression of the maximum updraft velocities is evident by the decrease in the ensemble-median values of the maximum updraft velocities (Figure 7a). However, Figure 7b implies that enhanced aerosol loading may produce either weaker or stronger updraft velocities depending on the initial environmental conditions and suggests that such aerosol-induced responses are environmentally modulated.*

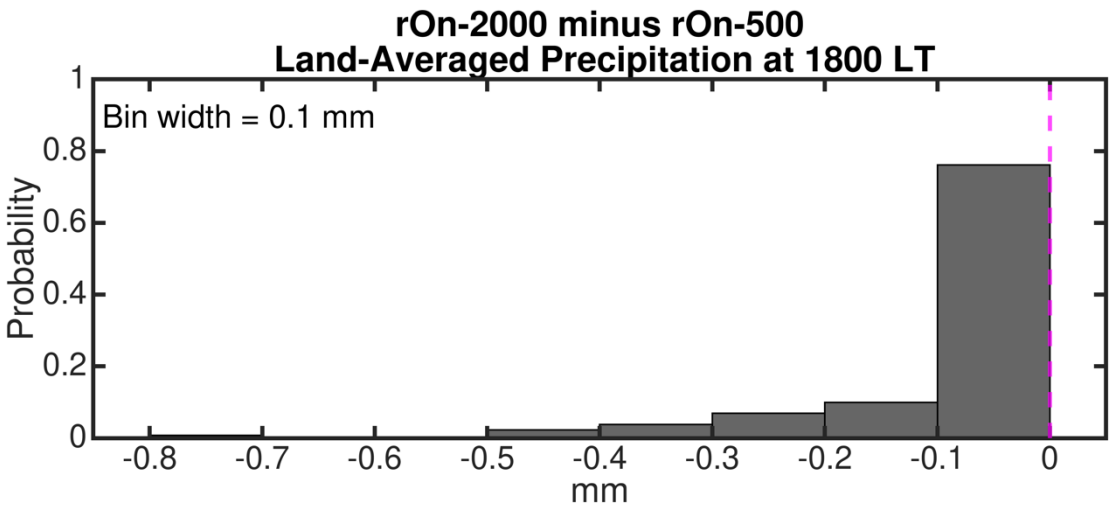

Figure 10. A histogram of the difference between rOn-500 and rOn-2000 land-averaged accumulated surface precipitation at sunset (1800 LT) for simulations that produce at least 0.1 mm of the land-averaged accumulated precipitation in rOn-500. The dashed magenta line indicates where the difference between rOn-2000 and rOn-500 is zero. The bin width of the histogram is marked at the upper left corner.

Lines 434–436: Figure 10 displays a histogram of the aerosol-induced differences in the land-averaged accumulated surface precipitation at 1800 LT between rOn-2000 and rOn-500.*

Lines 439–443: Figure 10 distinctly shows that the accumulated precipitation is reduced in all of the rOn-2000 precipitating ensemble members when compared with their corresponding counterparts in rOn-500, thereby demonstrating that the enhanced aerosol loading leads to an overall reduction in surface precipitation produced by the sea breeze system, irrespective of the environment.*

Section 5.2.1 (Figure 8): You mentioned that "It is clear from this figure that….", but these scatter diagrams are not clear to me for comparison reasons. You may create a probability density grid scatter diagram (instead of dots), and you may plot clean-polluted. Or, at least, you may overlay scatter plots of clean and polluted like Fig 9e-f, and conduct some statistical process to mention "significant" or "clear" differences between clean and polluted cases.

Thank you for the suggestion. We tried a range of different graphical representations, including the probability density grid scatter diagram and the overlaid scatter plot, but found that the original scatter plots are the still the clearest representation of these trends. However, in order to better highlight the changes occurring between 297 and 299 K of the initial boundary layer potential temperature on the maximum updraft velocity, we added the 10 m s^{-1} horizontal line to show that while 5 out of 12 deep convective simulations with the maximum updraft velocity greater than 10 m s^{-1} in rOn-500 have initial boundary layer potential temperature between 297 and 299 K, the only simulations in rOn-2000 with the maximum updraft velocity significantly greater than 10 m s^{-1} occur for initial boundary layer potential temperature greater than 299 K. We have also added the following text:

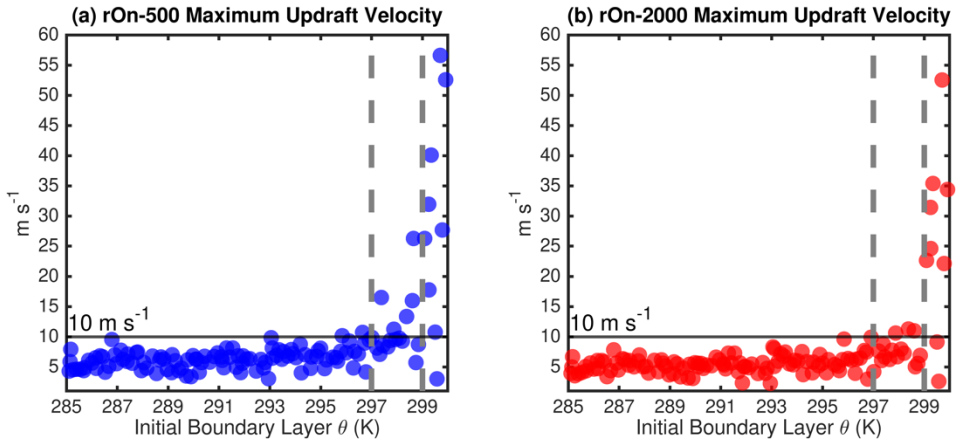

Figure 8. Pairwise scatterplots for the maximum updraft velocity (m s^{-1}) versus the initial boundary layer potential temperature (K) for the (a) rOn-500 and (b) rOn-2000 ensembles. The vertical dashed grey lines refer to the potential temperature thresholds described in the text.

Lines 354–362: Figure 8 shows scatter plots relating the maximum updraft velocity to the initial boundary layer potential temperature for all 130 ensemble members in rOn-500 (Figure 8a) and rOn-2000 (Figure 8b). It is clear from this figure that the sea breeze-initiated deep convective mode (updrafts are greater than 10 m s^{-1}) in rOn-500 occurs in all of the ensemble members in which the initial boundary layer potential temperature is 297 K or greater, and in which the mixed layer CAPE is greatest (not shown). However, in rOn-2000 the threshold above which the deep convective mode occurs is 299 K, which is 2 K greater than that in rOn-500. For instance, in rOn-500 while 5 out of 12 deep convective simulations with the maximum updraft velocity greater than 10 m s^{-1} have initial boundary layer potential temperatures of between 297 and 299 K (Figure 8a), the*

only simulations in rOn-2000 with the maximum updraft velocity greater than 10 m s^{-1} occur for initial boundary layer potential temperatures greater than 299 K (Figure 8b).

Fig 11: Fig 11 does not summarize physics very well. It pretty much displays all cases. For example, if you compute clean-polluted differences in auto-conversion profiles, and you can create CFAD to summarize all cases in one plot for each microphysical process (melting of ice, ice-to-rain, rain-to-ice, cloud-to-rain, etc..), it would be nice, because Test ID does not show any information of environmental factors anyway. So far, it's too numerous and mechanical test ID. So, it's difficult to extract physics from this plot.

Thank you for the suggestion. As discussed above, our initial reasoning for doing this was in an attempt to keep track of the different environments producing the range in responses. However, given that we ultimately decided to eliminate tracking each of the specific environments, the original Figure 11 was less meaningful. We tried CFADs for each term, however, we found that these were not overly helpful in conveying our message. We ultimately wanted to clearly demonstrate that the 36 pairs with 36 different initial conditions all exhibit aerosol-induced decreases in cloud-to-rain and rain evaporation rates regardless of their different environments. As such, we decided to summarize rOn-2000 minus rOn-500 averaged process rates as shown below in Figure 11, in which all 36 pairs are plotted together. We also plotted the ensemble mean rOn-2000 minus rOn-500 values for cloud-to-rain and rain-to-vapor plots with thick gray lines, as all of the 36 pairs exhibit warm rain processes. While the cold rain processes (i.e., melting of ice, rain to ice, and ice to rain process) contributing to surface precipitation are included in Figure 11, a detailed discussion and representation of cold rain processes are not included in the revised manuscript, as not all 36 pairs exhibit cold rain processes.

Lines 442–455: The bulk microphysical processes contributing to the differences in the surface rainfall in the 36 precipitating ensemble pairs are shown in Figure 11. All of the following source and sink terms for rain are considered:*

1. *Cloud-to-rain: cloud water transferred to rain through collection (gain term; Figure 11a);*

2. *Rain-to-vapor: evaporation of liquid water from rain (loss term; Figure 11b);*

3. *Melting of ice: ice mass transferred to rain via thermodynamic melting as the ice species fall below the freezing level (gain term; Figure 11c);*
4. *Rain-to-ice: rainwater that is collected by ice species through riming (loss term; Figure 11d); and*
5. *Ice-to-rain: collisional ice melting due to collection of warmer rain (gain term; Figure 11e).*

These process rates are averaged across all grid points over the land domain between 1200 and 1800 LT. In most of the precipitating members, the cloud frequency over land is

heavily weighted by the daytime cumulus convection mode. As a result, the frequency and associated contributions made by the averaged mixed-phased process contributions are small, if even existent. However, in some members, as shown in Figures 11c–e, the averaged mixed-phase process contributions are greater than those of the warm-phase processes.

Lines 459–463: Similarly, average rain evaporation rates (i.e., rain-to-vapor) shown in Figure 11b are also greater in magnitude in rOn-500 than rOn-2000, implying that the population of less numerous but larger raindrops formed in rOn-2000 evaporate less readily. The production of populations of fewer but larger raindrops in polluted conditions has been observed previously (e.g., Altaratz et al. 2007; Storer and van den Heever 2013).*

Lines 466–469: Figures 11c–e show that enhanced aerosol loading primarily produces a reduction in all three of the cold rain processes contributing to the rain budget, with only minor increases in the profile with aerosol loading in some cases. However, given the small sample size, additional testing would be required before conclusive statements can be made regarding environmental modulation of cold phase processes.*

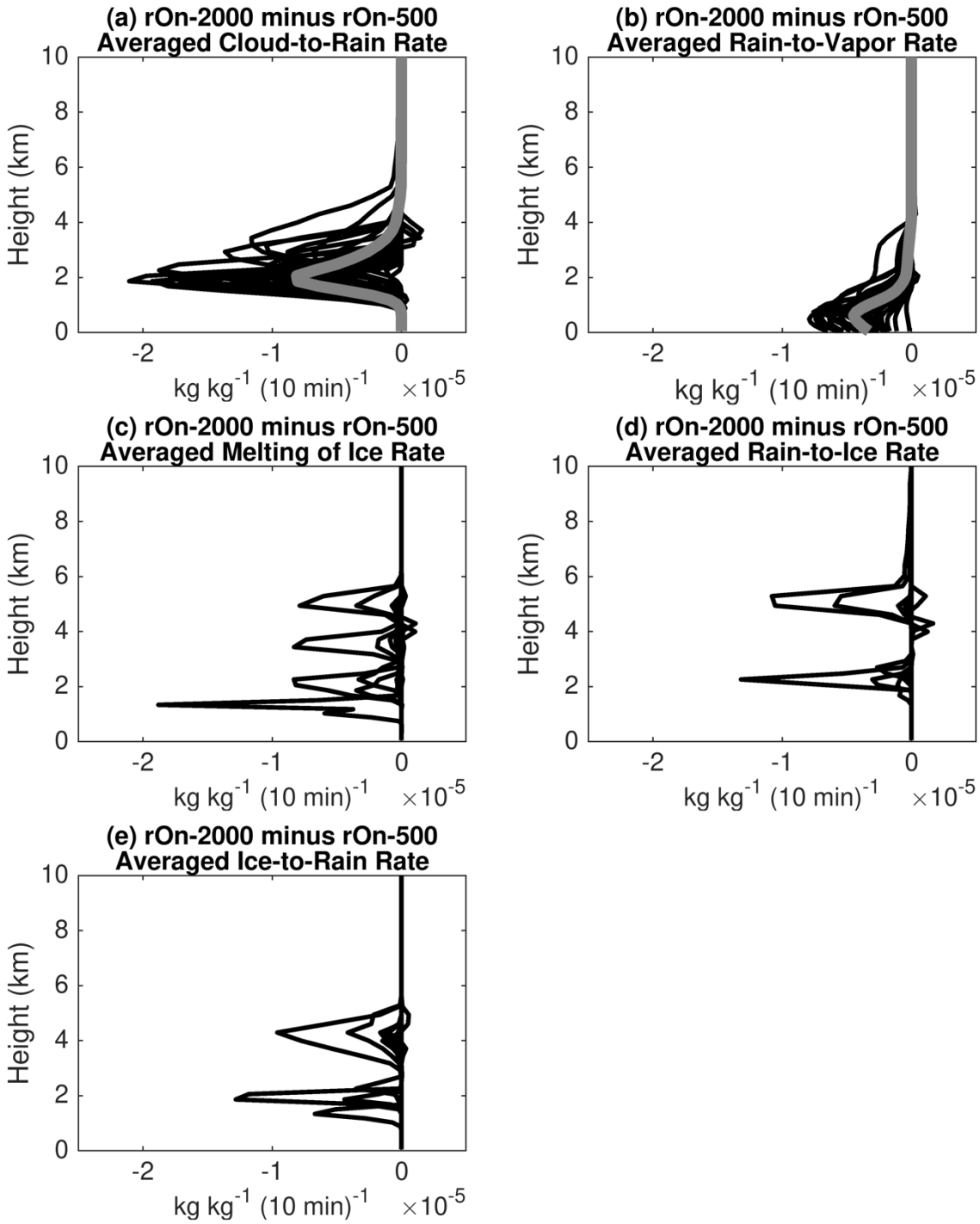

Figure 11. Aerosol-induced differences to the processes generating rain for the range of environmental conditions tested in these large ensemble experiments. Shown are the rOn-2000 minus rOn500 differences (a) cloud to rain; (b) rain to vapor; (c) melting of ice; (d) rain to ice; and (e) ice to rain rates (see the text for an explanation of these processes) using the 36 ensemble pairs that produce at least 0.1 mm of land-averaged surface accumulated precipitation in rOn-500. Process rates are averaged over the land domain between 1200 and 1800 LT. The thin black lines in all of the figures are from all 36 pairs, and the thick grey lines in a and b are the means of the rOn-2000 minus rOn-500 for the 36 pairs.

Minor Comments:

Resolution: Simulations are conducted with 1km grid spacing, and discussion of shallow-to-deep convection transition can be limited. I understand this is purely because of computational limitations with the many ensemble simulations. At least, you should mention this limitation somewhere in the manuscript.

We have added the following sentence in the revised manuscript.

> *Lines 537–540*: Thirdly, a grid spacing of 1 km was selected for these extensive ensembles given their high computational costs. Such a grid spacing will marginally resolve deep convective cloud systems but will under resolve the shallow convective mode. As computational capabilities are enhanced, a similar study should be conducted using grid spacings of O(100m).*

Line 37: I suggest ditch following sentence of this paper's topic "*Such organised tropical convection also plays an essential role in global climates via its impacts on planetary circulations such as the Walker circulation or the Madden-Julian Oscillation (Hendon and Woodberry, 1993; Zhang, 2005)*." This paper is not dealing with organized tropical convection.

It has been removed.

Line 63: "convectively" -> "convective"

It has been corrected.

Line 68: Suggest ditch "in the interest of focusing specifically on aerosol indirect effects". Sounds repetitive.

Done.

Line 70: "size and composition" -> "sizes and compositions"

Done.

Line 88: "theories" -> "hypothesis" Also apply the following sentences.

Thank you, we have made the change.

Line 139: Table 1 is not refered from sentences.

Table 1 was referred to in Section 2.1., Line 128.

Line 243: Add “and less surface turbulent heat flux” after “With less surface upwelling longwave radiation,”.

Done.

Line 245-246: Remove parenthesis.

It has been removed.

Line 247: “longwave radiation” -> “longwave radiation and surface turbulent heat flux”

Done.

Figures 3 and 4 (and related discussion) might be combined, since these are all surface impact and feedback.

Thank you for the suggestion. After considering this point, we have decided to keep Figures 3 and 4 as they are since Figure 3 is for showing land vs. ocean surface radiation and temperature responses, whereas Figure 4 is only for land surface heat fluxes.