

Responses to Reviewer 1:

The authors present results for a new 2-moment microphysics scheme for activation of aerosols and nucleation of cloud droplets in the UK Met Office Unified Model. They demonstrate marginal improvements via validation of simulated marine stratocumulus clouds over the previous 1-moment scheme that provided ok results for the wrong reasons due to compensating precipitation processes. The presentation and analysis is sufficient, but I find it difficult to accept some of the justifications, simplifications, and assumptions made in the schemes being discussed. More specific details about this statement are below. Perhaps the authors can address my comments and shed additional light on the motivation for the efforts described herein.

We thank the reviewer for their helpful comments. We have identified improvements we can make to our article in response to these comments and those of the other reviewer. We will highlight the main changes in red in a new version we are ready to upload.

Page 3, Line 27 down to Page 4, Section 2: It's unclear why you would use a 2-moment scheme with aerosols and then use a saturation adjustment for supersaturation and condensation.

We appreciate that some people have assumed in the past that double-moment microphysics schemes include prognostic supersaturation (we found an example in the review by Guichard and Couvreux (2017)). We therefore agree with the reviewer that we should justify our choice more explicitly, and so we propose to include the following paragraph in our new draft of the manuscript:

“In our CASIM microphysics scheme, saturation adjustment is applied. It has sometimes been assumed that prognostic supersaturation is generally part of double-moment microphysics schemes (Guichard *et al.*, 2017). However, Shipway and Hill (2012) compared several single and double-moment bulk microphysics schemes including CASIM with a bin microphysics scheme in a single-column framework. The bin scheme treated supersaturation prognostically while most of the bulk schemes did not. The double-moment bulk schemes with saturation adjustment they tested were in closer agreement with the bin microphysics scheme than the single-moment schemes. Their conclusions were substantiated further by Hill *et al* (2015). While useful, prognostic supersaturation is not essential for a double-moment microphysics scheme to improve on a single-moment scheme.”

We understand that saturation adjustment can still lead to biases, which we discuss below. A version of the UM with prognostic supersaturation would be a valuable research tool that would allow us to reach shorter timesteps and higher spatial resolutions, and we would like to implement it in future. However, we are not trying to simulate convective invigoration by aerosols with high accuracy (which is where we feel prognostic supersaturation is most important in representing aerosol-cloud interactions in models), as we pointed out in the text.

There is clear value in a one-size-fits-most weather- and climate model microphysics parameterization that can work at degree-scale resolution down through almost all the numerical weather prediction scales to about 500m resolution. The configuration of the model we introduce in this paper will allow improved representations of (for example) precipitation, fog, or the Twomey effect in weather and climate simulations, almost all of which are run at lower spatial resolution than we use here. We are testing the model at the finest resolution we think the

microphysics with saturation adjustment should be able to handle for these clouds. We emphasise, though, that we present a new configuration of an existing numerical weather prediction (NWP) model, and the GLOMAP aerosol and CASIM cloud microphysics schemes have been characterized before.

If we used prognostic supersaturation, we also would need a timestep of order 2s which would be unfeasible in our climate and NWP models. In the model of Morrison and Grabowski (2008) with prognostic supersaturation, timesteps longer than 3 seconds were shown to lead to substantial biases at realistic updraft speeds. Lebo and Morrison (2014) similarly use 3 seconds. Other schemes, for example that of Phillips *et al* (2007), use microphysical substeps, and while the scheme in the RAMS model was demonstrated to work with 10s timesteps in a simple case (Walko *et al.*, 2000) even 10s is much too short for an NWP model.

In the abstract, we said previously that our improvements ‘reinforce our confidence in the ability of the model to simulate aerosol-cloud microphysical interactions’. To avoid giving the impression that the model is capable of representing *all* aerosol-cloud microphysical interactions, we plan to rephrase this to say ‘simulate the aerosol-cloud microphysical interactions it was designed to represent’.

In an aerosol-limited regime, like typical marine stratocu, what happens if you consume all the aerosols but still have vertical motion and generation of supersaturation? If you are not permitting supersaturation to be carried around, where does the excess vapor go? To the growth of existing droplets?

Yes, excess vapor would still lead to droplet growth. The only issue with saturation adjustment in itself is that this condensation would occur too quickly, and therefore at the wrong altitude, if the timesteps are short. However, we show that for our model the timesteps are (almost always) not too short: the relaxation time (given in our paper, at page 20, line 33 of the posted discussion article) for supersaturation is almost always shorter than the timestep (Figure S13, top-left).

However, we agree that the potential biases would still be worth investigating in future work.

Will those just grow in-definitely via saturation adjustment?

No, because the water vapor will be depleted. To a reasonable approximation, for the 20s timesteps used in this model, the same amount of water condenses whether one uses saturation adjustment or prognostic supersaturation.

It seems like this could truly cause a problem with reliable prediction of latent heat release and droplet growth. I think more explanation or justification is needed beyond a reference to Morrison and Gettleman (2008).

We agree, and we propose to point out that the relaxation time constraint is satisfied in our introduction, and explain the significance of this. We also propose to add a citation to the review of cloud-resolving model microphysics schemes by Khain *et al* (2015) to back up our assertion that saturation adjustment is commonly applied, even at cloud-resolving resolution.

If supersaturations in our case study were higher, as in deep convective clouds, the relaxation time would be longer, and we would be more likely to incur biases if we did not lengthen the timestep and reduce the spatial resolution, possibly to the point that the convection would not be sufficiently well resolved. Lebo *et al* (2012) suggest one minute is a sufficiently long timestep for saturation adjustment to work in deep convective clouds. We have not shown our model is

suitable for simulations of deep convection at our 500m spatial resolution, and we pointed this out in our conclusions, at line 30 on page 21 of the posted discussion article.

It is important to note that in this work, as in many operational systems, a sub-grid cloud scheme is employed to represent sub-grid inhomogeneities in RH and cloud (even at 500m resolution). This means that cloud can exist under sub-saturated conditions. At the end of each timestep, we therefore intend ‘saturation adjustment’ to mean that the grid-cell mean relative humidity is not constrained to be 100%, but must instead be a grid-mean relative humidity below 100%. We appreciate that this wasn’t as clear as it should have been in the discussion paper, so we plan to clarify it during revision. We will also avoid using the term ‘cloud-resolving’, as this could misleadingly suggest there is no sub-grid cloud scheme. The minimum relative humidities are fixed parameters in the model which may vary with altitude.

In this way, the sub-grid cloud scheme allows us to account, to some extent, for unresolved vertical velocities when we handle condensation of water vapor. Most models with prognostic supersaturation we are aware of don’t represent sub-grid fractional cloud cover or the variability of supersaturation within a grid cell— there is, as far as we know, only one prognostic variable for supersaturation, and (for example) no second prognostic for the standard deviation of its spatial distribution. Therefore, at kilometer-scale resolution, the grid-mean supersaturation these models calculate may not be representative of the real supersaturation within different parts of the grid cell, as the updrafts from which the prognostic supersaturation is calculated are not resolved. This potential source of bias may be large – as model grid cell size is increased, at some point it would become larger than biases from saturation adjustment – and it is avoided by using saturation adjustment with a sub-grid cloud scheme.

Page 6, Line 33: Why do the fine grid regional domains go all the way up to 40km? Seems like you could reduce the number of vertical levels substantially and improve runtime if you topped things out in the lower stratosphere, especially since you’re not simulating deep clouds.

We agree the number of vertical levels is excessive for this case. The high model top is the default in regional UK Met Office operational NWP configurations such as UKV or RAL (Bush *et al.*, 2020) and we did not need to change it for this study.

Page 11, Lines 8-9: Could these high peaks be resulting from saturation adjustment? If you suddenly force new droplets to form and/or vapor growth of existing droplets in order to use up all saturation in one time-step, perhaps this is shocking the system and creating a sudden spike in drops and perhaps also a spike in latent heat release, buoyancy, and W. This is one particular reason to move away from saturation adjustment schemes.

We think this is possible, as occasionally the grid-mean relaxation time exceeds the model timestep. However, this is a rare occurrence. We propose to add a sentence “It is possible that these peaks correspond to the few grid cells where our saturation adjustment assumption has led to excessive release of latent heat, temporarily strengthening the updrafts.”

Page 12, Lines 9-10: Perhaps the under-predicted RH is causing less simulated cloud fraction shown in figure 3. Maybe you’re just not getting enough of an area that can generate clouds due to lack of moisture. Any thoughts?

This is possible, though we cannot be sure the isolated measurement in a single column of the atmosphere is representative. See our next comment.

Page 15, Lines 6-7: I would suspect you're not getting enough simulated convection. Your simulated cloud area is a lot less than the satellite viewed cloud cover. Perhaps this again goes back to the simulation being too dry?

This is a possibility. The small domain size means there is a substantial effect from boundary conditions too. We re-ran a simulation at 500m resolution with the values of the critical relative humidity in our sub-grid cloud scheme reduced by 10%, to a minimum of 80% instead of 90%. With this crude tuning we were able to produce more cloud, as shown in the center two plots of the figure below (the left two plots are the default as in Figure 3 of our paper, the center two are with the reduced critical relative humidity). We prefer to keep the standard model configuration for the main results in our paper, however, as this tuning is not likely to be relevant in other case studies and we feel it is more useful to demonstrate the performance of a well-documented model configuration. However, we will also add the figure to the supplementary materials and refer to it in the paper text.

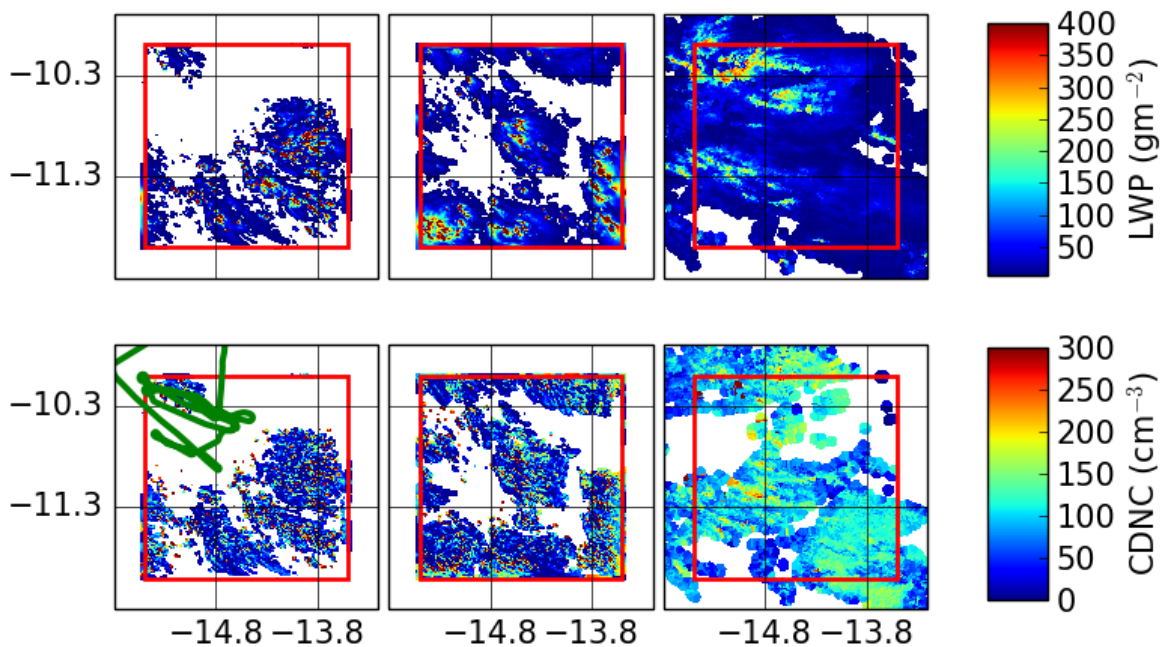


Figure 1: Comparison of the 500m-resolution model cloud liquid water path (LWP) and cloud droplet number concentration (CDNC) to MODIS AQUA observations (right) on 19 August 2017 at 1200. The leftmost plots show the default model and the center plots show the model when the critical relative humidity in the sub-grid cloud scheme is reduced by 0.1. Isolated MODIS pixels are enlarged during regridding as described in the manuscript text, and white space in the plot of cloud droplet number concentration may indicate both cloud-free areas and failed retrievals.

Your vertical velocity distributions in Figure 7 show that the simulations produce FAR fewer strong updrafts with very few instances of simulated $W > 2$ m/s. This would also help explain why your cloud cover is much less than observed. This could certainly be a model resolution

issue and perhaps these simulations should have been run with $DX < 200\text{m}$ in order to get better resolved updrafts.

We have a sub-grid cloud scheme, so in principle we shouldn't need to fully resolve vertical velocities in order to produce clouds.

Page 15, Lines 14-15: This again looks like it could be an issue with saturation adjustment in combination with activating aerosols with such a scheme.

We believe saturation adjustment would be unlikely to lead to such biases in our case. The existing text of the paper does acknowledge the inaccuracies associated with saturation adjustment in the introduction, and we discuss shortcomings in the activation scheme in Section 6 in the discussion article.

Page 18, Line 3: The underestimation of large drops could be due to the required fit to a gamma distribution. It could also have something to do with rain drop breakup being too aggressive when large oblate drops attempt to form via collisions

Yes, we agree – thank you. We will add in the point about the gamma distribution.

8. Page 21, Lines 23-24: In light of the comment regarding Phillips et al. (2007), I have to wonder why efforts are being made to improve a scheme with admittedly known limitations rather than adapting the model to something that avoids saturation adjustment and permits prognostic supersaturations. Many models have already moved in this direction.

We are using an operational NWP and climate model suitable for global to regional scales, and we test here it at a horizontal resolution that is close to the highest spatial resolution the model is run at operationally. In such a model, saturation adjustment is an appropriate approximation. Our aim is not to build a model specifically for kilometre-scale simulations of clouds.

9. Page 23, Lines 3-4: Here you state that the “ratcheting” effect should cause the 2-moment scheme to produce more droplets than the 1-moment scheme. Many times, 1-moment schemes hold the number concentration fixed, so your statement isn't really broadly true

This is a good point that we overlooked. We plan to adjust the text to read:

“However, the ratcheting means the double-moment scheme should produce more droplets on average than a single-moment scheme, provided the number concentration of droplets is diagnosed from the aerosol concentration in the single-moment scheme. For example, if the diagnostic UKCA activation scheme were fed by the same updraft speeds and the same aerosol as a double-moment scheme, fewer droplets on average would be produced.”

10. Page 25, Lines 23-25: If the PDFs are different, this is likely due to your choice of gamma distribution shape parameter. You would get a different solution if you change the breadth of the distribution.

We meant the PDFs of how CDNC varies in space, rather than the droplet size distribution function. Sorry for the confusion, we will amend the text to make this clear.

Responses to Reviewer 2

The writing of the manuscript is sloppy. Many statements are not in the format of scientific writing. Also, there are many confusing statements and confusing use of the terms. Simulations are not clearly described. There are inconsistent and undefined names used in figures legends,

contradicted statements, lack of clear description of data sampling/processing approach to compare with the model simulation. A lot of my detailed comments below are about these problems. I do not think I captured all of the problems. In my opinion, the manuscript was in some draft mode and not ready for submission.

We thank the reviewer for taking the time to make the detailed suggestions below, which we have addressed. In addition, we have tried to make other things in the manuscript text clearer following the examples highlighted by the reviewer. We will highlight the main changes in red in a new uploaded version of the manuscript.

The organization of manuscript needs improvement. For example, Sections 6 and 7 are purely the description of model developments and they are quite long. They should be moved to the Section 3 or 4 to have model developments described together.

We agree, and plan to move these sections, distributing some of the text elsewhere in the manuscript where appropriate. We also substantially restructured the description of the model in our latest draft.

It is an overstatement that the effect of subgrid vertical velocities on activation is accounted. Basically what the authors did was to lower the threshold of grid mean updraft speeds used for activation. The statement makes people think that they connect the activation with the subgrid vertical velocity spectrum calculated from turbulence to get this done. Using a grid mean value of updraft speed, does not justify to say “subgrid vertical velocity”. This has to be clarified throughout the paper. Otherwise it would be misleading.

As well as lowering the threshold of grid mean updraft speeds, we also examined the procedure of Malavelle *et al* (2014) for correcting updraft speeds to account for unresolved turbulence, and while we didn't actually implement the procedure, we did derive a reasonable ad-hoc tuning factor. We agree that the ad-hoc nature of our correction, which would only work for our specific case, may not have been as clear as it should have been, and we will adjust the text appropriately.

(4) The model configuration and model simulations are not clearly described. Suggest to use a table to clearly show the configurations of major simulations. Some confusions arise from the misuse of abbreviations, for example, in some places, UKCA and CASIM are used for microphysics schemes, while they also used for different aerosol activation schemes. In the first sentence of Section 9, one of them is referred to as an aerosol scheme?

We agree our previous labels were confusing. We propose to make numerous improvements and include a table at the start of section 5, including the simulations at their three resolutions, and revise the acronyms. We occasionally used UKCA and GLOMAP interchangeably, which could cause confusion, and we also used ‘UKCA activation scheme’ to refer to the diagnostic implementation of Abdul-Razzak & Ghan (2000) by West *et al* (2014) and CASIM activation scheme to refer to the prognostic implementation. We now use ARG (DIAG) in place of ‘UKCA activation’ and ARG (PROG) in place of ‘CASIM activation’ throughout the paper to avoid this confusion.

(5) I do not think 500m resolution is fine enough to simulate stratocumulus clouds explicitly. The turbulence is very difficult to simulate at this resolution because it is partially resolved and a portion of it needs to be parameterized but model is difficult to know how much. This might be a reason for the poor simulation of PBL height. I would suggest run a test with resolution smaller than 250 m with a smaller domain to see how the simulated cloud and updraft can be improved, particularly the inversion height.

As we use a sub-grid cloud scheme, we are not technically trying to simulate the clouds explicitly. However, we did use the term ‘cloud-resolving’ in the text, which we appreciate may mislead the reader to suggest that the clouds *are* simulated explicitly. We will replace it.

We were able to make modest improvements to the simulation of the clouds by tuning the critical relative humidity in the sub-grid cloud scheme, please see our response to reviewer 1.

While the global and 500m simulations have higher PBL heights than the 7km resolution model, the variability between the three simulations is relatively small. It takes about 4 hours for air masses to advect from the edge of the 500m-resolution model domain to the center, assuming an 8m/s horizontal wind speed. As this is not very long, we think the biases in the PBL height in the 500m model are largely due to the biases in the lower-resolution driving simulations. Therefore running a higher-resolution inner simulation is unlikely to produce a substantial improvement.

The biased inversion height is partially an issue with the reanalysis that our global model is nudged to, as well as the PBL scheme. Our nudging is not too constraining: we nudge to horizontal wind, not temperature, above approximately 2000m altitude, but we note the PBL height is higher than 2000m here. We added ERA-interim and ERA5 reanalysis temperature profiles to our existing temperature profile figure (shown as Figure 2 in these replies, below), to demonstrate the bias in the reanalysis. We then plan to add the global model to Figure 4 in the main text of the paper, to show that it, too is biased, in line with the reanalysis. We will make appropriate comments in the manuscript text. We still don’t know why the 7km-resolution model produces a slightly lower PBL height than the global model.

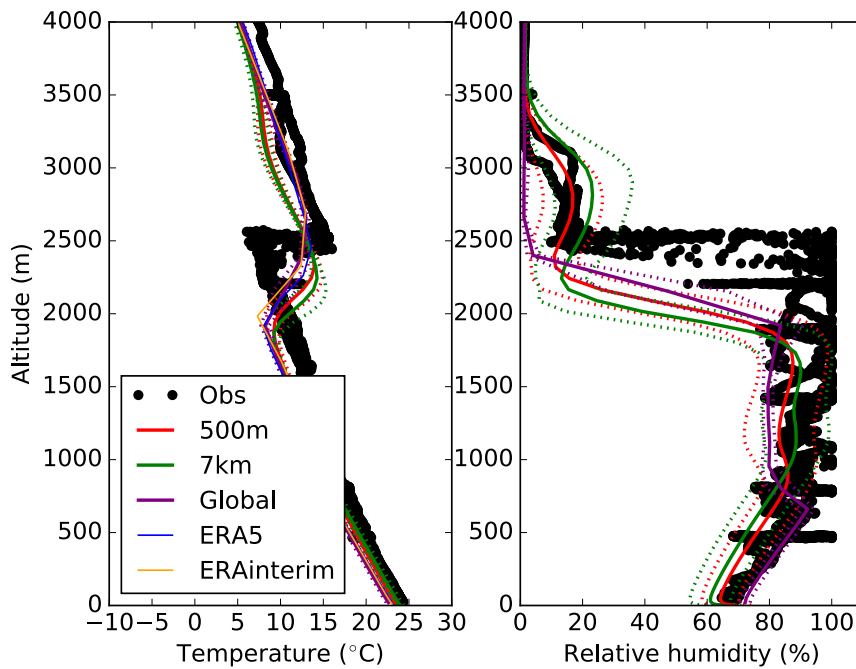


Figure 2: A copy of the temperature profiles (left) shown in the manuscript with ERA5 and ERA-interim reanalysis temperature profiles superposed, together with the simulated relative humidity

in the three models (right). Dotted lines indicate one standard deviation across the model grid cells sampled.

After several unsuccessful attempts to improve the 7km and global model performance, we tried running the global model nudged only above 4000m altitude (starting from 11 August), with a tuning of entrainment parameters in the boundary layer scheme, and we do see a modest 100m improvement in the boundary layer height compared to the reanalysis (Figure 3 in these replies, below). We conclude that the global model *can* be tuned to perform marginally better than either ERA-interim or ERA5 in this case, although it still underestimates the boundary layer height. The improvement in performance in this region may well be accompanied by a degradation elsewhere – we didn't check. As in the case of our tuning of the critical relative humidity (response to reviewer 1), we prefer to retain our default simulations in the paper, to avoid hiding the shortcomings of the well-defined CMIP6 or RA1 climate and weather prediction models by tuning.

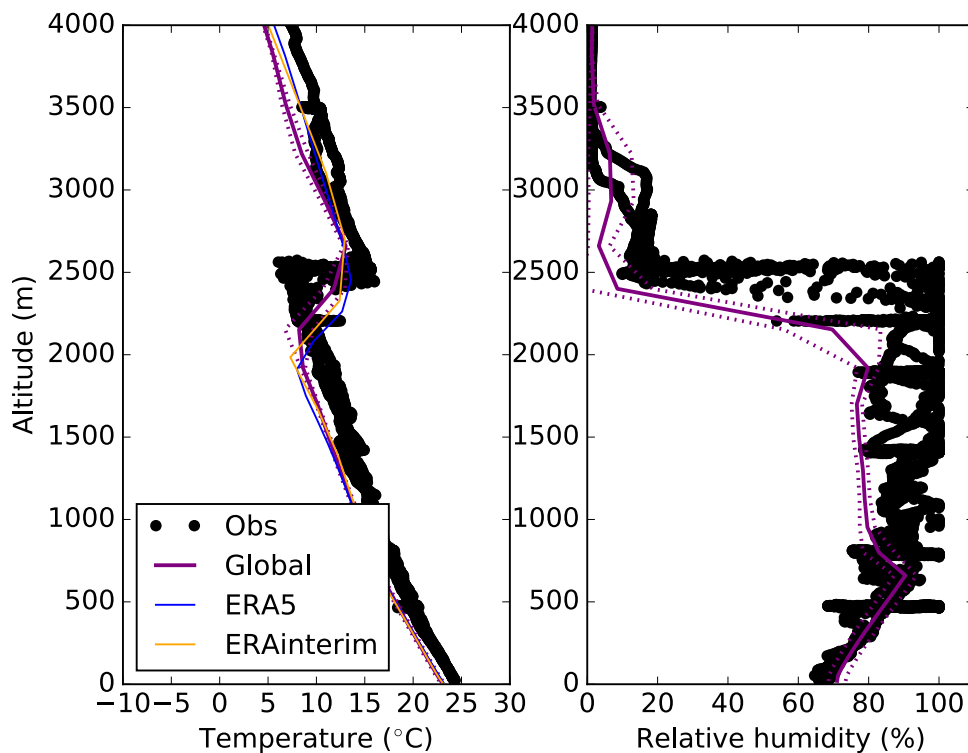


Figure 3: Temperature and relative humidity profiles in the tuned global model (purple), nudged only above 4000m altitude, compared to FAAM aircraft observations and, for temperature, ERA-5 and ERA-interim reanalysis in the domain of our 500m-resolution model. Dotted lines indicate one standard deviation across the model grid cells sampled.

Currently, the simulations that authors presented did not do well in simulating the clouds and had a huge problem with the Aitken mode of aerosols as well. This might not justify an acceptance of the paper due to these problems. If you do not want to further look at the aerosol problem, at least try to provide a good simulation of the clouds or find the major reasons leading to the large model-observation discrepancy

The overestimated Aitken mode number is a problem with the global climate model, as shown by Mulcahy et al (2020), page 33 and 34 (the GC3.1 climate model has identical aerosol microphysics to our model configuration). We are preparing a global modeling manuscript

(Ranjithkumar, Gordon, Carslaw *et al*, in prep.) to study the issue more closely with the UKESM configuration, which has coupled online oxidation chemistry but is otherwise similar. We believe the overestimated concentration is the result of excessive new particle formation in the upper troposphere. We show the concentration of Aitken-mode particles in the global model at high altitudes in Figure 4 below, which we propose to include in supplementary text.

In total, including the nucleation mode (all particles above 3nm diameter), we simulate about 20000 particles per cubic centimeter at STP in the upper troposphere at 10km altitude, which is about a factor of four higher than measurements from the ATom campaign published by Williamson *et al* (2019) suggest. We note our Figure 4 shows ambient concentrations in the Aitken mode only. At 10km altitude, our SO₂ concentrations are about 40ppt, which is about a factor of two too high, presumably due to underestimated scavenging of SO₂ in deep convective clouds, erroneous emissions or deposition, or in-cloud aqueous reactions. Since new particle formation is non-linear in gas concentrations, we think the SO₂ may well explain the discrepancy in Aitken-mode aerosol number concentration.

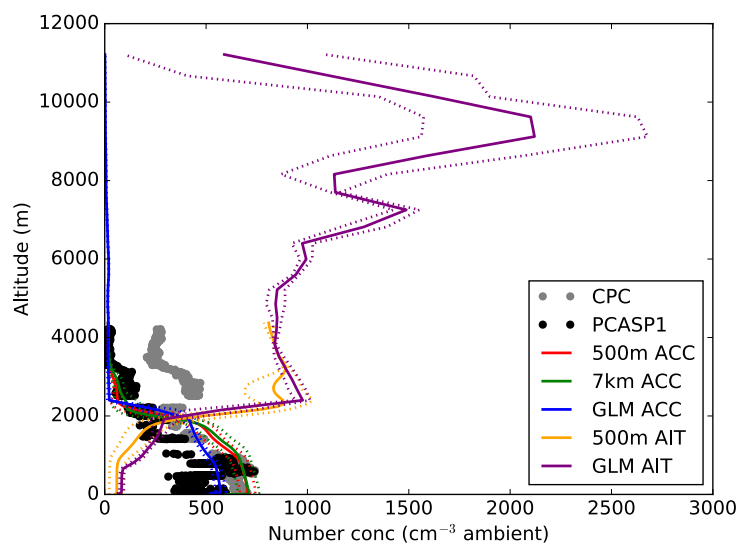


Figure 4: vertical profiles of aerosol number concentration in the Aitken (AIT) and accumulation (ACC) modes of the regional models with 500m and 7km resolution, and the global model (GLM), compared to the measurements from the aircraft CPC and PCASP instruments. Compared to the manuscript, we also plot simulated number concentrations from the global model at higher altitudes, showing the excessive new particle formation in the upper troposphere.

At our simulated cloud top, about 2100m altitude, our Figure 4 shows the total Aitken mode concentration exceeds the accumulation mode concentration. However, Figure 3 in our manuscript shows that the simulated Aitken mode diameter is about 30nm at 1900m altitude, and 40nm at 2550m, so only a fraction of the Aitken mode will be large enough to activate. Moreover, most cloud droplets are not formed at the cloud top height, rather nearer cloud base, where the concentration in the Aitken mode is negligible. We think the biased Aitken mode may be partly responsible for the overestimated droplet concentration, and the overestimated width in the spatial distribution of droplets, at 2215m altitude, shown in Figure 8, and we will add a comment to that effect in the text. We will also make some relevant comments in the text that was in Section 5.3 of the discussion article to accompany the new supplementary figure.

1. Need to describe explicitly the model configuration/resolution instead of using something like RA1, GA7.1

We agree, and will restructure the description to make it clearer, add additional references to our emissions datasets, the requested summary table, and some additional detail. We prefer to keep the references to standard configurations to avoid lengthening the manuscript excessively.

2. P1 Line 18-21, text is contradicted with the text on p9 Line 12. If you are testing the aerosol and chemistry component of the model at a higher resolution than has been attempted before, why do you need to specific a kappa value for activation? Then at P9Line 25-27, it is said volume-weighted hygroscopicity is passed to CASIM for activation. Very confusing.

We use the term ‘volume weighted hygroscopicity’ when we should have said the ‘volume-weighted kappa value’ – since each component of aerosols (e.g sulphate, organic carbon) has its own assigned kappa value, and the overall kappa value of the aerosol is the volume-weighted average of the kappa values of the components. Sorry, this will be corrected. However, we still need to specify a kappa value, because we still use the activation parameterization of Abdul-Razzak and Ghan (2000).

P8-9, it is not clear why ARG is implemented differently between the 7 km and 500m resolutions.

The reason is that the 7km and the global simulations still use single-moment cloud microphysics, while the 500m-simulation uses double-moment cloud microphysics (the CASIM scheme). This will be clarified in the new text.

4. P.9 Line 5-10, First, is ARG applicable to 500-m resolution since it was developed based on cloud parcel model with timesteps for global climate models?

We did not identify any physical reason why ARG should not be run in models with higher horizontal resolution (and the vertical resolution in our model is also closer to that of a climate model than to a large eddy simulation.) We don’t believe the reviewer meant that the cloud parcel model itself could have used timesteps from global models (~20 minutes). In fact, we would expect ARG to perform better at higher spatial resolution, because a higher fraction of the sub-grid variability in aerosol concentrations which, unlike the variability in updraft speeds, are not usually accounted for, will be resolved.

The ARG parameterization or similar parameterizations have been applied successfully in cloud-resolving models, or models with similar spatial resolution, before, as summarised in Table 3 of Ghan *et al* (2011). We will note this in the text.

Figure 3 showed that cloud droplet numbers from 500 m resolution are worse compared with 7-km resolution, indicating the scheme might not work well for very high-resolution.

It’s true that the cloud droplet numbers are not in good agreement with observations, but we do think the underlying activation parameterization should work fine at this resolution, as discussed above. We think the biases in the cloud droplet concentration are due at least in part to biases in the accumulation-mode aerosol concentrations. In Figure 3, the 500m simulations use CASIM double-moment microphysics, and the cloud droplet numbers come from the “CASIM activation scheme” (which we propose to relabel ARG(PROG)) while those in the center from the 7km simulations use Wilson and Ballard single-moment microphysics and the “UKCA activation scheme” that we propose to relabel ARG(DIAG), so unfortunately the activation schemes cannot be directly compared in this figure.

Second, based on your description here about accounting for subgrid velocity effect on activation, you are not using any subgrid velocity. If you only use a grid mean value of updraft speed, which means you are only account for the impact of resolved updraft, not subgrid vertical velocity. This has to be clarified throughout the manuscript because the writing gave me impression that a particular method is employed to account for the subgrid velocity effect on activation in this study until I read here.

In our default simulation, we do not use a subgrid velocity. We will clarify this in the manuscript. We do, however, show results from a simulation where we scale the grid-box mean vertical velocity to illustrate what a simulation that accounted for sub-grid effects might look like. Malavelle *et al* (2014) derive a scaling factor for the variability in vertical velocity with the appropriate dependence on model grid resolution. We assert in this paper that, if we average over many model timesteps and spatial grid cells, scaling the grid-mean updraft velocity should have the same effect as scaling the variability in vertical velocity, provided, in line with Malavelle *et al*, that enough of the variability in updraft speed is resolved by the grid. For no bias to result, the vertical velocity would need to vary independently amongst the model grid cells, and from one time-step to the next, and it would need to have a Gaussian distribution when many grid cells and timesteps were combined. This is not a very good approximation, but the simulation is still a reasonable demonstration of the effect expected from a sub-grid velocity distribution.

5. P9 Line 34-35, it is confusing to say “CASIM microphysics code has the capability to simulate aerosol microphysical process” just because they simulate in-cloud removal and aerosol resuspension. Those processes are called aerosol-cloud interaction processes, not aerosol microphysical processes (instead they are cloud microphysical processes)

We agree that our use of “processing” here is confusing, as not all readers may be familiar with the jargon. We will change the text to read ‘cloud microphysical processing of aerosol’– thank you.

Figure 3 showed that the model did not capture the observed cloud well. At least try to figure out potential reasons for the large discrepancy between the model and observation

Please see our response to reviewer #1 and major comment #5.

P10 Line 4-7, I am confused. Here you said In the UKCA code, aerosol resuspension is not accounted, but I think UKCA is said as just an implementation method for ARG activation scheme. If it is a different microphysics scheme, why you need a different microphysics scheme from CASIM? Also, didn’t you use coupled GLOMAP-CASIM for this study? Why “a coupled GLOMAP-CASIM double-moment model that includes aerosol microphysical and chemical processing is deferred to future work

We will change the acronyms, as outlined in our response to major comment #4.

8. Suggest use a table to clearly describe the configurations of aerosol and microphysical processes of the simulations for each domain

This is a good suggestion, we will add the table.

9. Figure 4, What is ACC and AIT? They are not defined. From Figure 4, the model did not capture the inversion well. Is it the problem of resolution?

We will define the abbreviations in the caption. For the inversion, please see above.

10. P12 Line 5-7, confusing. What boundary layer height are you talking about?

Ascension Island is not in our simulation domain. During the case study period, the boundary layer height (BLH) at Ascension Island is 500m lower than it is in the domain of the 500m-resolution model. In the 7km resolution model, the simulated BLH at Ascension Island is 300m lower than it is in the domain of the 500m-resolution model. We agree it was written confusingly -we will clarify the text.

11. Sections 5.3 and 5.4, How did you sample the data from the simulations to compare with the aircraft observation? Need some description either in the figure captions or in the text.

For Figure 4, we used a horizontal mean of the 500m-model domain at midday UTC on 19 August 2017, after removing all grid cells within 10km of the domain boundaries.

We assume a) that the duration of the flight is sufficiently short that the observations may be compared to this temporal snapshot of the model. The aircraft entered the simulation domain at 1103 UTC and left it at 1310 UTC. We will point this out in the text, in a new introduction to the Section 5 of the discussion article. We also assume b) that variability in aerosol concentrations, temperature and relative humidity in the simulation output across the domain is quite small. The accumulation mode number varies from 600/cc to 800/cc in the BL (Figure 2) so this does introduce an error. However, because the simulated and observed clouds are not in the same place in the domain, we think it is better to use a spatial average than to sample the simulation along the track of the aircraft. In the figure caption, we plan to say: “Variability across the model domain is assumed here to be small, but, for accumulation-mode aerosol, it may be inferred from Figure 2.”

Figure 5: we will point out the simulated data are “a mean over the domain at 1200 UTC. Only aircraft data from straight-and-level legs at the specified altitude are shown.”

Figure 6: in the caption we will additionally specify: “We include all model grid cells at the specified altitude within the simulation domain with a liquid water content greater than 0.01gkg^{-1} at the instant of 1200 UTC on 19 August in the histogram, except for the grid cells within 10km of the domain boundary.”

Figure 7: We will add “Model data are from instantaneous simulation output of the 500m and the 7km resolution simulations, in both cases across the domain of the 500m simulation, at 1200 UTC on 19 August 2017.”

12. Section 5.5, should we expect 500m resolution is fine enough to simulate stratocumulus clouds? See my major comments #4

Please see response to major comment #5.

13. P15 Line 9-11, do you still have cloud fraction for 500-m resolution? If so, how is the cloud fraction determined and is the way to determine cloud fraction is the same between the 7 and 0.5 km resolution?

We do have cloud fraction, determined diagnostically using the scheme of Smith (1990), and it is the same at both model resolutions. This is now included in the model description. We now avoid using the term ‘cloud-resolving’ as this may have implied we did not have a cloud fraction.

14. Figure 8, which simulation does “Model” denote? What are UKCA and UKCA 7here? I could not find definition of these terms anywhere. In the Table 1, there are simulation names of CASIM, UKCA 500m, UKCA 7km, UKCA global. None of them are consistent with the names used in Figure 8.

We changed the labels throughout to ARG(PROG), ARG (DIAG) etc, including on the figure caption, to reflect the prognostic double-moment cloud droplet number in the activation scheme used with CASIM microphysics, or the diagnostic activation scheme used with the single-moment microphysics of Wilson and Ballard (1999). The orange and red curves in Figure 8 are

both from the 500m simulation – orange is the diagnostic activation scheme we previously called ‘UKCA activation’ and now call ARG(DIAG), while red is the activation scheme for cloud droplets in the double-moment CASIM microphysics scheme, now called ARG(PROG). The green curve is the UKCA (now ARG(DIAG)) activation scheme at 7km resolution.

15. The poor simulations of cloud droplet distribution might indicate that ARG scheme might not be applicable to those resolutions, as I commented above.

On the applicability, please see our response to minor comment #4. Our paper focuses on the ARG(PROG) scheme, but the ARG(DIAG) scheme also has shortcomings, some of which are documented in the supplement. So far, however, we have only identified shortcomings in the implementation of the scheme, and not in the underlying ARG algorithm. These shortcomings are combined with biases in the simulated aerosol concentration.

16. Figure 9, which simulation did you compare here? Such information should be clearly described in the figure caption.

We agree, and propose to add ‘The simulations are from the 500m-resolution model (the only simulation with double-moment cloud microphysics).’

17. P20, 11-14, which activation scheme are you talking about here? There are two activation schemes - UKCA and CASIM as discussed above

CASIM (now labelled ARG (PROG)); this will be clarified.

18. The organization of Section 6-8 is strange since Sections 6 and 7 are purely the description of model developments. They should be moved to the Section 3 or 4 to have model developments described together.

This will be done.

19. Section 9, in-cloud activation is thought as secondary nucleation, which should increase droplet number. Here it decreases CDNC. I think here it means different from what people usually think. It is just a way to treat activation with accounting for existing droplets? Please clarify. Is it added to the primary nucleation (cloud-base nucleation)? If so, how can it reduce CDNC?

In our original model, as in simulations published by Grosvenor et al (2017), Miltenberger et al (2018), and various others, “primary nucleation” was applied irrespective of whether or not a cloud is present already in the grid box. We understand this is also done in other models, such as WRF with Morrison & Gettelman (2008) microphysics. When we apply “secondary nucleation”, which is indeed a way to account for existing droplets, we apply it instead of primary nucleation, and therefore the overall droplet number goes down because the supersaturation used in the activation scheme is lowered by the existing droplets.

We describe how we combine ‘primary’ and ‘secondary’ nucleation in partially cloudy grid boxes as defined by our cloud fraction scheme. It is the way of doing this combination that we feel makes our change non-trivial. We will modify some of the introductory text to make our changes clearer.

20. Section 9, first sentence: it is a confusing sentence. Which one is two-moment aerosol scheme and which is the two-moment cloud microphysics? Based on what I read, UKCA and CASIM were also used for naming different activation schemes.

We agree, and will relabel the activation schemes.

21. P26, Line 15-16, I do not think it is scientific writing by saying “model emits smoke..
We will adjust the text to ‘In our simulations, smoke and marine aerosols are emitted...’

22. P28, Line 1, First I do not agree that the model performs well based in the results shown. Second, even if it indeed performs well, please state perform well in what quantity and what aspect. It is not a scientific writing to only state “the model performs well”.

We agree, and propose to remove this comment and the subsequent text, and integrate the content later with a more quantitative statement, as the model performance is discussed more specifically in the paragraph that follows, starting ‘Our evaluation of meteorology and of aerosol and cloud microphysics highlighted some areas where the model performs well but also some shortcomings.’

23. P28, 7, I have no idea what you mean about “cloud deck is just under 200km across”
The approximate spatial extent of continuous cloud cover in the horizontal direction is just under 200km. We did not understand the source of confusion here, but we are happy to amend the text if needed.

References for review replies

- Abdul-Razzak, H. & Ghan, S. J. A parameterization of aerosol activation: 2. Multiple aerosol types *J. Geophys. Res. Atmos.* **105**, 6837–6844. (2000).
- Bush, M., Allen, T., Bain, C., Boutle, I., Edwards, J., Finnenkoetter, A., Franklin, C., Hanley, K., Lean, H., Lock, A., et al. The first Met Office Unified Model–JULES Regional Atmosphere and Land configuration, RAL1 *Geosci. Model Dev.* **13**, 1999–2029. (2020).
- Ghan, S. J., Abdul-Razzak, H., Nenes, A., Ming, Y., Liu, X., Ovchinnikov, M., Shipway, B., Meskhidze, N., Xu, J. & Shi, X. Droplet nucleation: Physically-based parameterizations and comparative evaluation *J. Adv. Model. Earth Syst.* **3**, (2011).
- Grosvenor, D. P., Field, P. R., Hill, A. A. & Shipway, B. J. The relative importance of macrophysical and cloud albedo changes for aerosol-induced radiative effects in closed-cell stratocumulus: insight from the modelling of a case study *Atmos. Chem. Phys.* **17**, 5155–5183. (2017).
- Guichard, F. & Couvreux, F. A short review of numerical cloud-resolving models *Tellus, Ser. A Dyn. Meteorol. Oceanogr.* **69**, (2017).
- Hill, A. A., Shipway, B. J. & Boutle, I. A. How sensitive are aerosol-precipitation interactions to the warm rain representation? *J. Adv. Model. Earth Syst.* **7**, 987–1004. (2015).
- Khain, A. P., Beheng, K. D., Heymsfield, A., Korolev, A., Krichak, S. O., Levin, Z., Pinsky, M., Phillips, V., Prabhakaran, T., Teller, A., et al. Representation of microphysical processes in cloud-resolving models: Spectral (bin) microphysics versus bulk parameterization *Rev. Geophys.* **53**, 247–322. (2015).
- Lebo, Z. J. & Morrison, H. Dynamical Effects of Aerosol Perturbations on Simulated Idealized Squall Lines *Mon. Weather Rev.* **142**, 991–1009. (2014).
- Lebo, Z. J., Morrison, H. & Seinfeld, J. H. Are simulated aerosol-induced effects on deep convective clouds strongly dependent on saturation adjustment? *Atmos. Chem. Phys.* **12**, 9941–9964. (2012).
- Malavelle, F. F., Haywood, J. M., Field, P. R., Hill, A. A., Abel, S. J., Lock, A. P., Shipway, B.

- J. & McBeath, K. A method to represent subgrid-scale updraft velocity in kilometer-scale models: Implication for aerosol activation *J. Geophys. Res. Atmos.* **119**, 4149–4173. (2014).
- Miltenberger, A. K., Field, P. R., Hill, A. A., Rosenberg, P., Shipway, B. J., Wilkinson, J. M., Scovell, R. & Blyth, A. M. Aerosol–cloud interactions in mixed-phase convective clouds – Part 1: Aerosol perturbations *Atmos. Chem. Phys.* **18**, 3119–3145. (2018).
- Morrison, H. & Gettelman, A. A New Two-Moment Bulk Stratiform Cloud Microphysics Scheme in the Community Atmosphere Model, Version 3 (CAM3). Part I: Description and Numerical Tests *J. Clim.* **21**, 3642–3659. (2008).
- Morrison, H. & Grabowski, W. W. Modeling Supersaturation and Subgrid-Scale Mixing with Two-Moment Bulk Warm Microphysics *J. Atmos. Sci.* **65**, 792–812. (2008).
- Mulcahy, J., Johnson, C., Jones, C., Povey, A., Scott, C., Sellar, A., Turnock, S., Woodhouse, M., Andrews, M., Bellouin, N., et al. Description and evaluation of aerosol in UKESM1 and HadGEM3-GC3.1 CMIP6 historical simulations *Geosci. Model Dev. Discuss.* 1–59. (2020).
- Phillips, V. T. J., Donner, L. J. & Garner, S. T. Nucleation Processes in Deep Convection Simulated by a Cloud-System-Resolving Model with Double-Moment Bulk Microphysics *J. Atmos. Sci.* **64**, 738–761. (2007).
- Shipway, B. J. & 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. (2012).
- Smith, R. N. B. A scheme for predicting layer clouds and their water content in a general circulation model *Q. J. R. Meteorol. Soc.* **116**, 435–460. (1990).
- Walko, R. L., Cotton, W. R., Feingold, G. & Stevens, B. Efficient computation of vapor and heat diffusion between hydrometeors in a numerical model *Atmos. Res.* **53**, 171–183. (2000).
- West, R. E. L., Stier, P., Jones, A., Johnson, C. E., Mann, G. W., Bellouin, N., Partridge, D. G. & Kipling, Z. The importance of vertical velocity variability for estimates of the indirect aerosol effects *Atmos. Chem. Phys.* **14**, 6369–6393. (2014).
- Williamson, C. J., Kupc, A., Axisa, D., Bilsback, K. R., Bui, T. P., Campuzano-Jost, P., Dollner, M., Froyd, K. D., Hodshire, A. L., Jimenez, J. L., et al. A large source of cloud condensation nuclei from new particle formation in the tropics *Nature*. **574**, 399–403. (2019).
- Wilson, D. R. & Ballard, S. P. A microphysically based precipitation scheme for the UK meteorological office unified model *Q. J. R. Meteorol. Soc.* **125**, 1607–1636. (1999).

Summary of main changes.

Additional justification and clarifications concerning saturation adjustment
 Perform and document, mainly in supplement, some tuned simulations to improve model-observation agreement as described in review replies.
 Relabel activation schemes throughout for clarity
 Merge section 6 and section 7 and exchange with section 5.
 Substantial updates to model description and evaluation sections.
 Additional table summarizing model configuration
 Numerous clarifications in figure captions, including in the supplementary information.

Text changes are highlighted in red in the marked-up version attached.

Improving aerosol activation in the double-moment Unified Model with CLARIFY measurements

Hamish Gordon^{1,2}, Paul R. Field^{1,3}, Steven J. Abel³, Paul Barrett³, Keith Bower⁴, Ian Crawford⁴, Zhiqiang Cui¹, Daniel P. Grosvenor¹, Adrian A. Hill³, Jonathan Taylor⁴, Jonathan Wilkinson³, Huihui Wu⁴, and Ken S. Carslaw¹

¹School of Earth and Environment, University of Leeds, LS2 9JT, United Kingdom

²Engineering Research Accelerator, Carnegie Mellon University, Forbes Avenue, Pittsburgh 15213, United States

³Met Office, Fitzroy Road, Exeter, EX1 3PB, United Kingdom

⁴Department of Earth and Environmental Sciences, University of Manchester, M13 9PL, United Kingdom

Correspondence to: Hamish Gordon hamish.gordon@cern.ch

Abstract. Representing the number and mass of cloud and aerosol particles independently in a climate, weather prediction or air quality model is important in order to simulate aerosol direct and indirect effects on radiation balance. Here we introduce the first configuration of the UK Met Office Unified Model in which both cloud and aerosol particles have ‘double-moment’ representations with prognostic number and mass. The GLOMAP aerosol microphysics scheme, already used in the HadGEM3 climate configuration, is coupled to the CASIM cloud microphysics scheme. We demonstrate the performance of the new configuration in ~~cloud-resolving~~ **high-resolution** simulations of a case study defined from the CLARIFY aircraft campaign in 2017 near Ascension Island in the tropical south Atlantic. We improve the physical basis of the activation scheme by representing the effect of existing cloud droplets on the activation of new aerosol, and we also ~~attempt to account for~~ **discuss** the effect of unresolved vertical velocities. **We show that neglect of these two competing effects in previous studies led to compensating errors but realistic droplet concentrations.** The first of these improvements should be applicable to the representation of aerosol activation in other microphysics schemes. While these changes lead only to a modest improvement in model performance, they reinforce our confidence in the ability of the model **microphysics code** to simulate **the** aerosol-cloud microphysical interactions **it was designed to represent**. Capturing these interactions accurately is critical to simulating aerosol effects on climate.

1 Introduction

Shallow marine clouds are an important source of uncertainties in climate forcing and sensitivity. Representing aerosol effects on these clouds is a priority for climate modelling efforts worldwide. In this paper, we describe ~~cloud-resolving~~ model simulations in a 2° by 2° region of the tropical south Atlantic Ocean near Ascension Island. The simulations are performed with the UK Met Office Unified Model (UM), with double-moment aerosol microphysics driving double-moment bulk cloud microphysics at 500 m horizontal resolution. **The cloud and aerosol microphysics parameterizations form, or are intended to form in future, part of the atmosphere model code used for climate simulations in Coupled Model Intercomparison Project (CMIP) experiments and for operational weather forecasts across the Unified Model partnership. The rest of the model code is also used**

in the HadGEM3-GC3.1 configuration for CMIP6 and in current operational numerical weather prediction configurations. As well as testing the aerosol and chemistry component of the model at a higher resolution than has been attempted before, we study and suggest improvements to the performance of the aerosol activation scheme. We evaluate the second day of a two-day long simulation against CLARIFY (CLouds and Aerosol Radiative Impacts and Forcing) aircraft measurements on 19 August 2017.

A series of recent field campaigns (Zuidema et al., 2016) have focused on the tropical south-east Atlantic ocean, which hosts one of the planet’s largest stratocumulus decks, and is the destination for much of the biomass burning aerosol that originates from central and southern Africa. The prevailing winds, which are south-easterly in the boundary layer and easterly in the free troposphere, advect smoke over the ocean where slow subsidence causes the smoke to mix with the clouds. The smoke can have large direct, semi-direct and indirect radiative effects on the regional climate (Costantino and Bréon, 2013; Lu et al., 2018; Gordon et al., 2018). Here we focus on the indirect effect of the aerosols.

In a previous study, Gordon et al. (2018) evaluated aerosol transport and microphysics in global and convection-permitting simulations of the south-east Atlantic. They built on earlier work with the same aerosol microphysics scheme employed at high spatial resolution by Planche et al. (2017). In this paper, we start from an updated version of the same atmospheric model, and increase the resolution further, to approach the cloud-resolving scale. We also increase the sophistication of the cloud microphysics scheme from single-moment to double-moment, in order to study aerosol-cloud interactions in more detail. The resulting model is the first configuration of the Unified Model with fully double-moment aerosol and cloud microphysics. Double-moment microphysics, including prognostic cloud droplet number concentration, is important to enable good representations of processes such as aerosol activation and droplet settling at high spatial and temporal resolution. The coupling to the double-moment interactive aerosol microphysics scheme enables aerosol-induced variability in the droplet number concentration. We evaluate the aerosol and cloud microphysics in the new model configuration in this paper, paying particular attention to shortcomings of the simulations that are specific to aerosol-cloud interactions or to simulating aerosols at high spatial resolution. We also highlight some underlying issues with aerosols in the coarse-resolution global climate model that drives our regional simulations, which we will address in future work.

Compared to single-moment cloud microphysics schemes, double-moment schemes have been shown previously to improve the representation of stratiform rain in NWP simulations (Morrison et al., 2009) and to reduce a range of biases in high-resolution climate models (Seiki et al., 2015). The CASIM (Cloud AeroSol Interacting Microphysics) double-moment microphysics scheme we use here is that published previously (Shipway and Hill, 2012; Grosvenor et al., 2017; Miltenberger et al., 2018; Stevens et al., 2018) and Furtado et al. (2018). In the last of these manuscripts, Furtado et al. (2018) evaluate CASIM for deep convective clouds and compare it to a reduced, single-moment version of the same scheme, and to the different single-moment cloud microphysics scheme (Wilson and Ballard, 1999) currently used in the operational version of the model. For the case they study, the CASIM double-moment microphysics initially performed better than the single-moment schemes, although these gave comparable performance to the double-moment scheme when tuned.

In climate simulations with resolutions coarser than around 0.5 degrees, updraft velocities in shallow clouds are almost entirely unresolved and convection is parameterized. The activation of aerosols to form cloud droplets requires the supersaturation

of water vapor relative to aerosols and hydrometeors to be diagnosed or parameterized. Typically, supersaturation is calculated by imposing an updraft speed (or a distribution of updraft speeds) derived from diagnostics of the sub-grid turbulence rather than the grid-box mean updraft speed, which is close to zero. A single cloud droplet number concentration per grid box is thus produced and used in the prediction of rain rates (via an autoconversion parameterization) and cloud albedo.

5 As the resolution of the simulation is increased into the 'terra incognita' or 'gray zone' (Wyngaard, 2004), a higher fraction of the turbulence in the boundary layer is resolved, until, at the large eddy simulation (LES) scale of order 10 m, we assume for the purposes of this paper that the spatial variability of prognosed updrafts would be a good representation of reality. The turbulence starts to be resolved when the effective grid resolution is below about four times the height of the boundary layer (Honnert et al., 2011), typically 4 – 8 km. There is therefore a point at which it is no longer necessary to use an updraft
10 speed diagnosed from sub-grid turbulence in the activation scheme, and the grid-box mean can be used instead. Prior CASIM simulations (Grosvenor et al., 2017; Miltenberger et al., 2018; Furtado et al., 2018) have also used grid-box mean updraft speeds to activate aerosols at horizontal resolutions ranging from 250 m to 1 km, and a similar approach has been taken in Regional Atmospheric Modeling System (RAMS) and some Weather Research Forecasting (WRF) simulations (Saleeby and Cotton, 2004; Thompson, 2016). The scale invariance of activation schemes has been tested before down to horizontal resolutions of
15 1 km (Possner et al., 2016). However, these resolutions are still much coarser than typical LES resolution, and therefore the full variability in updraft speeds will not be resolved. By comparing **near-cloud-resolving** and LES simulations, Malavelle et al. (2014) developed a bootstrapping parameterization which enables the estimation of the fraction of the variance in updraft that is resolved.

In existing clouds, activation of new droplets will often be negligible, but not always, for example if the updraft strengthens
20 towards the top of the cloud. In a detailed model, Pinsky and Khain (2002) suggested that in-cloud activation leads to a bimodal cloud droplet size spectrum, and is an important factor in accelerating rain formation by broadening the droplet spectrum (Segal et al., 2003; Heymsfield et al., 2009). In a deep cloud case study, the large eddy simulations of Fridlind et al. (2004) suggested that aerosol concentrations in the boundary layer were too low to explain observed droplet number concentrations, and therefore aerosols must be entrained from the mid- and upper troposphere and activated inside the clouds. More recent detailed modeling
25 studies have continued to investigate secondary activation (Khain et al., 2012; Fan et al., 2018). To capture fully effects of in-cloud activation on the droplet size distribution, a detailed microphysics scheme is needed, such as a size-bin-resolving scheme or a super-droplet model, and to accurately simulate the full dynamical response of deep clouds to aerosols, very high spatial resolution may be needed, which in turn requires relatively short timesteps. If the model timestep is shorter than the 'relaxation timescale', **the timescale for** supersaturation production **is to be** balanced by condensation onto existing droplets, a bias will
30 occur unless supersaturation is represented as a prognostic variable (Khain and Lynn, 2009; Fan et al., 2012; Lebo et al., 2012; Grabowski and Jarecka, 2015; Grabowski and Morrison, 2017). This timescale is typically a few seconds, depending on the vertical velocity and droplet spectrum. Prognostic supersaturation is desirable in simulations requiring accurate calculations of the latent heat released by condensation (Grabowski, 2007), but the short timesteps needed mean it is very expensive to treat supersaturation prognostically (Arnason and Brown, 1971; Morrison and Grabowski, 2008), although this can be mitigated to
35 some extent with semi-analytic approaches (Clark, 1973; Hall, 1980). The relaxation timescale is short (less than 10 seconds)

in the case of thick, polluted clouds with many or large droplets and/or low updraft speeds, and long in the case of very clean clouds or strong updrafts (Kogan and Martin, 1994). If, however, the model timestep is longer than the relaxation timescale (and we verify that this is the case in our study), one can then assume that supersaturation produced during a timestep leads to condensation immediately, so that the relative humidity is 100% at the end of the timestep. This assumption, termed ‘saturation adjustment’, may lead to biases in the latent heat released by condensation and in the evaporation of clouds. However, it is much simpler and computationally cheaper than treating supersaturation prognostically, and it has been widely applied for decades at the cloud-resolving scale (Khain et al., 2015) as well as in lower-resolution weather prediction and climate models.

In our CASIM microphysics scheme, saturation adjustment is applied. It has sometimes been assumed that prognostic supersaturation is generally part of double-moment microphysics schemes (e.g. Guichard and Couvreux, 2017). However, Shipway and Hill (2012) compared several single and double-moment bulk microphysics schemes including CASIM with a bin microphysics scheme in a single-column framework. The bin scheme treated supersaturation prognostically while most of the bulk schemes did not. The double-moment bulk schemes with saturation adjustment they tested were in closer agreement with the bin microphysics scheme than the single-moment schemes. Their conclusions were substantiated further by Hill et al. (2015). While useful, prognostic supersaturation is not essential for a double-moment microphysics scheme to improve on a single-moment scheme.

In our model, on each timestep the activation scheme is re-run assuming there are no existing droplets. If the new droplet concentration is greater than the existing droplet concentration, the old droplet concentration is overwritten by the new. Changes to the cloud fraction in the gridbox may also change the droplet concentration, as discussed later. A similar procedure is followed in the widely used Morrison and Gettelman (2008) scheme. In this study, we follow the suggestion by Korolev (1995); Ghan et al. (2011), and others to improve on this procedure by accounting for existing cloud droplets when new droplets are activated, assuming a supersaturation that results from a balance between production (updraft) and loss (condensation on existing droplets) that there exists an equilibrium between supersaturation production by updraft and condensational growth of existing droplets. This assumption can be seen as a natural extension of the saturation adjustment assumption, and should apply in the same conditions. When it was tested in WRF-chem with Morrison et al. (2009) cloud microphysics, Yang et al. (2015) found that it improved simulated wet scavenging. Our implementation is aware of and consistent with our simulated sub-grid cloud fraction. We emphasise that we seek here to improve the accuracy of our existing model, but our improvements will not enable it to compete with detailed studies of aerosol-cloud interactions that employ spectral bin microphysics or prognostic supersaturation.

We attempt to improve the activation scheme further by accounting for considering unresolved sub-grid-scale updraft velocities, as a first step towards implementing the parameterization of Malavelle et al. (2014). We examine the suitability of the parameterization of Malavelle et al. (2014) for our model, but we are not yet able to implement it explicitly, and instead we derive an ad-hoc correction factor suitable for our case study. We show, unsurprisingly, that accounting for existing cloud droplets in the activation scheme reduces the cloud droplet number while our correction factor for sub-grid-scale updrafts increases it. This may explain why previous studies with CASIM (Grosvenor et al., 2017) have successfully produced realistic droplet concentrations with realistic CCN, despite underestimating the updraft and ignoring existing droplets. We examine

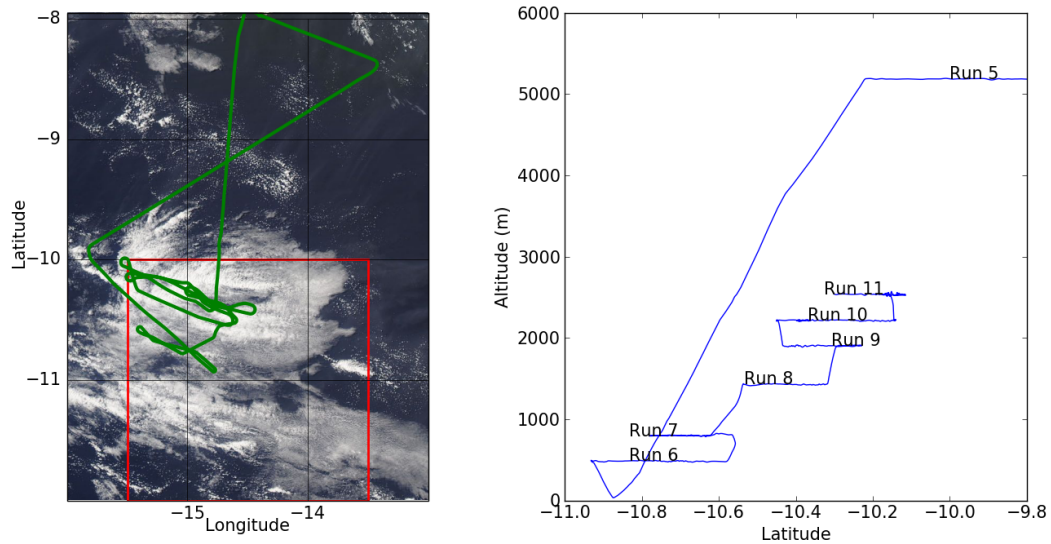


Figure 1. Flight pattern of the CLARIFY flight on 19 August 2017 used in the evaluations presented here, represented by the green line. The left plot is superposed on imagery from MODIS on the TERRA satellite that corresponds to the morning of the flight, showing the cloud deck that was sampled. The red box corresponds to the domain of the 500 m-resolution simulation. The right panel shows the path of the aircraft as a function of height and latitude.

the implications of our improvements for the cloud droplet spatial distribution and size distribution and for rain formation. We evaluate our simulations against CLARIFY aircraft measurements in 2017 to confirm they are realistic and to identify directions for future developments.

2 Case study and aircraft measurements

- 5 The CLARIFY aircraft campaign took place from 16 August to 7 September 2017. The BAE-146 aircraft of the Facility for Airborne Atmospheric Measurements (FAAM) was based at Ascension Island during this period. Extensive sampling of biomass burning aerosol interacting with clouds was achieved during 24 flights usually of around 3.5-4 hours' duration. The aircraft flew in all directions around Ascension Island, according to where aerosol and cloud transitions could be identified or in order to pass under satellite tracks.
- 10 The research aircraft was fitted with a comprehensive suite of thermodynamic, radiometric, cloud physics and aerosol instrumentation for the CLARIFY field campaign. Ambient air temperature was measured using a non-deiced Rosemount/Goodrich type-102 total air temperature sensor. Atmospheric water vapour was measured using a WVSS-II near-infrared tunable diode laser absorption spectrometer fed from a standard flush mounted inlet (Vance et al., 2015). The temperature and humidity measurements are used to calculate relative humidity in cloud-free air. The temperature measurement in cloud was subject
- 15 to significant wetting effects that led to a cold bias, as illustrated by Heymsfield et al. (1979). The derived in-cloud relative

humidity is therefore set to 100% in this work. Zonal, meridional and vertical wind components were derived from the five-port turbulence probe located on the aircraft radome (Petersen and Renfrew, 2009; Barrett et al., 2019). On the transit at 5180 m altitude from Ascension Island to the cloud, the mean vertical velocity was -0.010 ms^{-1} . The small magnitude of this mean velocity compared to the in-cloud vertical velocities we discuss later suggests the probe is sufficiently well calibrated for our analysis so we do not subtract any baseline offset from the observed updrafts in the evaluation we present in Section 6.

The size distribution of aerosol particles was measured using a Passive Cavity Aerosol Spectrometer Probe (PCASP, Droplet Measurement Technologies) for nominal diameters between about 0.1 and $3 \mu\text{m}$. We applied a complex refractive index of $r_i = 1.54 - 0.027i$, appropriate for biomass burning aerosol during the CLARIFY time period (Peers et al., 2019) and recomputed the bin boundaries for the PCASP instrument. This resulted in changes to the locations of bin centres, compared to the nominal values from the manufacturer, of usually around 5% but sometimes up to 20% in the diameter range below $1 \mu\text{m}$. For smaller particles with diameters between 0.03 and 0.3 microns, we also used a Scanning Mobility Particle Sizer (SMPS). The total number concentration of aerosols with diameter larger than about 2.5 nm were measured with a TSI 3786 Condensation Particle Counter (CPC). Aerosol data were only used when in cloud- and precipitation-free air. We determined this using the standard deviation of raw power on the Nevzorov total water content probe (Korolev et al., 2013), where a power greater than 3.0 mW ($\sim 1.5 \times 10^{-4} \text{ gm}^{-3}$) indicates cloud conditions, following Barrett et al. (2019). An additional safety window of 5 s ($\sim 500 \text{ m}$) either side of positively identified cloud was applied to account for diffuse cloud edges and imperfect temporal and spatial synchronisation between probes and data recording systems. The cloud droplet size distribution (between 2 and $52 \mu\text{m}$ diameter) was measured with a Cloud Droplet Probe (CDP) that was calibrated with a ten point bead calibration (Rosenberg et al., 2012). Precipitation-sized particles were measured using a 2D-Stereo (2DS) probe (10 to $1280 \mu\text{m}$ diameter) and a Cloud Imaging Probe (CIP-100) probe ($100 \mu\text{m}$ to 6.4 mm diameter). Data from the CDP, 2DS and CIP-100 are combined to produce a composite PSD at a 1 Hz sampling frequency following the method of Abel and Boutle (2012). Elsewhere in this paper, the cloud drop number concentration (CDNC) and liquid water content (LWC) are calculated using the CDP data only. For cloud measurements, in-cloud conditions were determined using a liquid water content threshold of 0.01 g kg^{-1} calculated by integrating the cloud drop number size distribution from the CDP.

The most comprehensive sampling of a cloud feature took place on 19 August. The aircraft flew south of the island to sample a large precipitating cloud structure around 1.5° in size. Biomass burning aerosol was detected both within and just above the boundary layer. Cloud top height peaked at around 2.5 km altitude. The cloud deck was sampled as shown in Figure 1, in a series of five straight-and-level aircraft trajectories along a line of strong radar echoes seen on the aircraft weather radar.

Taken together, the aircraft observations described later in the paper suggest that the boundary layer is decoupled or cumulus-like, with a stratocumulus cap above large precipitating cumulus clouds beneath. The observed stratocumulus clouds could also be detrained remnants of cumulus. In the simulations, the cumulus sometimes seems to extend up to the top of the boundary layer. However, the size of the cloud feature is larger than average for stratocumulus-to-cumulus transition clouds, and the deepest cumulus appeared to be organized linearly. Based on this and on geostationary satellite imagery (not shown) we would describe it as the beginnings of a fish rather than sugar, flowers or gravel (Stevens et al., 2019).

In order to validate our simulations, we are also able to draw on surface measurements from the ARM site at Ascension Island, and satellite observations of cloud droplet concentration and liquid water path. To obtain these, the same procedure as used by Gordon et al. (2018) is followed.

3 Model setup

- 5 To establish a regional simulation at 500 m spatial resolution, we set up three configurations of the Unified Model (UM): a global model, a 7 km-resolution regional model, and the 500 m-resolution model. The code for all three models is almost identical, except for the cloud microphysics, as described later. The global model is used to produce lateral boundary conditions for the 7 km-resolution model, and this in turn provides boundary conditions to the 500 m model. The higher-resolution models do not feed back to the lower-resolution models (our setup is usually described as ‘one-way nesting’). To illustrate the setup, 10 simulated accumulation-mode aerosol number concentrations are shown in the 7 km model with the global model (top) and 500 m model (bottom) in Figure 2. The configuration of the three models is summarized in Table 1 and described in this Section.

- We use the RA1 (Bush et al, 2020) configuration of the UK Unified Model (UM) with some settings borrowed from the GA7.1 configuration, which is the global climate configuration submitted to CMIP6, labelled HadGEM3-GC3.1. Following 15 Gordon et al (2018), who built on idealized simulations by Planche et al (2017), we use the United Kingdom Chemistry and Aerosols (UKCA) module from the climate model, which has prognostic two-moment aerosol microphysics (GLOMAP-mode; Mann et al 2010). GLOMAP stands for Global Model of Aerosol Processes. As explained in the Introduction, we also use the CASIM two-moment bulk cloud microphysics module (Shipway and Hill, 2012).

- In order to simulate the region of interest at 500 m resolution, we start with Our global model setup is similar to that used 20 by Gordon et al. (2018) and identical to that used in the intercomparison study of Shinozuka et al. (2019). It follows the GA7.1 configuration (Walters et al., 2019), which is the global climate configuration submitted to CMIP6. The horizontal resolution is $0.8^\circ \times 0.55^\circ$ (N216) and there are 70 vertical levels from the surface to 85 km altitude. The horizontal winds are nudged to ERA-interim above the boundary layer. The global model is initialized from the model for August 2016 used by Gordon et al. (2018) and run through to 2017.

- 25 This global model drives a 7 km-resolution regional simulation centred at -16° latitude, 0° longitude, with 670 longitude grid-boxes and 450 latitude grid-boxes, and 140 vertical levels from the surface to 40 km altitude. This simulation is initialized on 17 August at 0000UTC from the global model, and run until the end of 19 August. The configuration is similar to the global model but uses UM version 10.8. We use the RA1 (Bush et al., 2019) configuration of the UM with some settings borrowed from the GA7.1 configuration, and UM version 10.8. During the three-day simulation, air masses in the boundary layer advect 30 from approximately the south-eastern corner to the north-western corner of the model domain. While the simulations at coarse resolution pass information to those at finer resolution, the reverse is not true. To illustrate the setup, the accumulation-mode aerosol number concentrations are shown in the 7 km simulation with the global model (top) and 500 m model (bottom) on 19 August at midday in Figure 2. Both the global and 7 km-resolution models use the single-moment cloud microphysics scheme

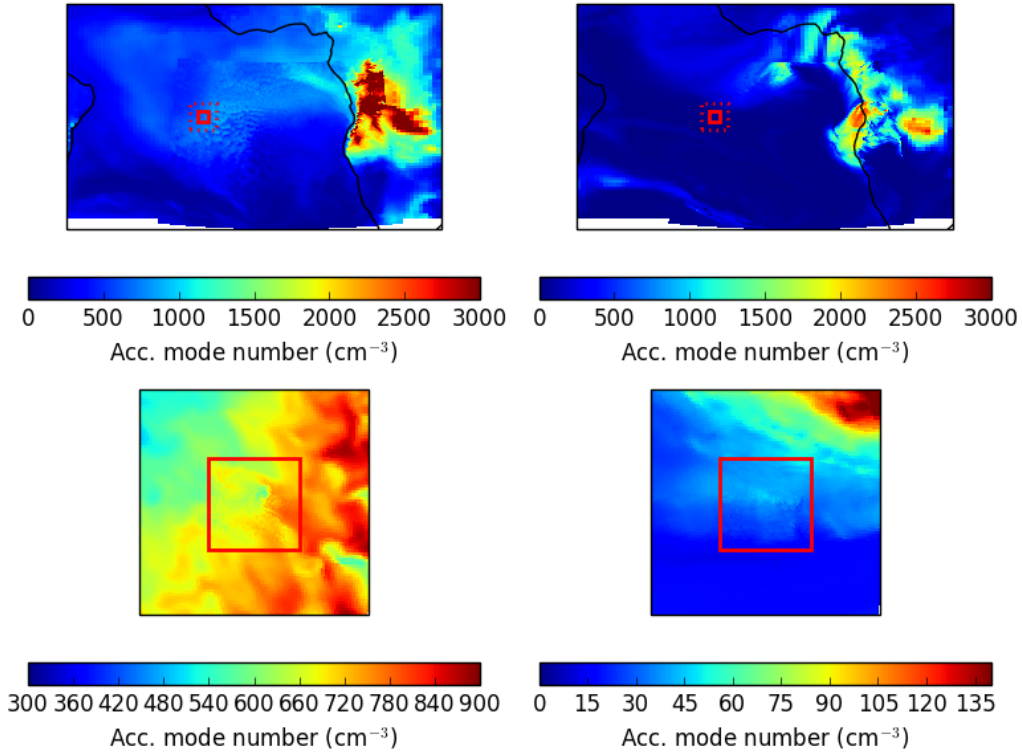


Figure 2. Simulated number concentration of accumulation-mode aerosol, (abbreviated as **Acc-mode** in the legend labels) in the boundary layer (averaged from the surface to 750m altitude) on the left and in the lower free troposphere (averaged from 2900 m to 3400 m altitude) on the right), in cm^{-3} , at **midday UTC on 19 August**. The 500 m regional domain is shown as a red square on all four plots; the top plots show the global and 7 km models with the domain of the bottom plots as a red dotted square, while the bottom show the 7 km model and the 500 m model.

of Wilson and Ballard (1999) and the double-moment GLOMAP aerosol microphysics scheme (Mann et al., 2010). GLOMAP stands for GLObal Model of Aerosol Processes. Because the aerosol microphysics is double-moment, cloud droplet number concentration can be represented to some extent in the cloud microphysics and radiation schemes, but as a diagnostic variable rather than a prognostic. We use ‘diagnostic’ here to indicate that the diagnostic droplet number concentration is calculated from the simulated aerosols, updraft speed, and temperature on each timestep without reference to the droplet concentration on the previous timestep, while the prognostic droplet concentration is retained in memory from one timestep to the next and advected by the simulated wind fields, though it may also be updated if the simulated aerosol concentration or updraft speed changes. The details are explained in the next Section.

The 500 m-resolution simulation uses CASIM double-moment cloud microphysics (Shipway and Hill, 2012) as well as GLOMAP aerosol microphysics. This simulation is driven by the 7 km-resolution model, and has 450 grid-boxes by 450 grid-boxes in latitude and longitude, and the same 140 vertical levels as the 7 km-resolution simulation. It is centred on -11°

Table 1. Summary of differences between model configurations used. Unless explicitly mentioned, affected by recent minor updates, or dictated by the name of the model configuration as described in the appropriate documentation paper, the code for global and regional simulations is identical. In the ‘Cloud microphysics’ row 1M signifies single-moment and 2M double-moment. In the ‘Aerosol activation’ row, ARG stands for the parameterization of Abdul-Razzak and Ghan (2000). This parameterization has two implementations, labelled DIAG when cloud droplet number concentration is diagnostic, and PROG when it is prognostic. The microphysics sub-step is marked with an asterisk because we wish to emphasise that the sub-step is not applied to condensation or evaporation, which are treated by the cloud parameterization and not by the microphysics scheme. However, it does apply to precipitation-related processes such as autoconversion and accretion, for example.

Configuration or parameterization	Global	7 km model	500 m model
Configuration name	GA7.1	RA1	RA1
Number of lat,lon grid cells	324x432	450x670	450x450
Timestep	20 minutes	2 minutes	20 seconds
Microphysics sub-step*	2 minutes	none	10 seconds
Number of vertical levels	85 (to 85km)	140 (to 40km)	140 (to 40km)
Code version	11.2	10.8	11.3
Cloud microphysics	1M (Wilson and Ballard, 1999)	1M (Wilson and Ballard, 1999)	2M (CASIM; Shipway and Hill (2012))
Aerosol microphysics	GLOMAP (Mann et al., 2010)	GLOMAP (Mann et al., 2010)	GLOMAP (Mann et al., 2010)
Aerosol activation	ARG (DIAG)	ARG (DIAG)	ARG (PROG)
Sub-grid cloud	PC2 (Wilson et al., 2008)	Smith (1990)	Smith (1990)

latitude, -14.5° longitude. Like the 7 km-resolution simulation, we use the RA1 configuration, but a more recent UM version, 11.3 is used, as this has the latest iteration of the CASIM microphysics. Other differences compared to version 10.8 are expected to have only minor effects. As in the 7 km and global models, the offline oxidants configuration of UKCA chemistry is used. This simulation is initialized from the global model on 18 August at 0000UTC. By 19 August, all of the air masses that advect into the domain from the boundaries will have been simulated by the 7 km-resolution model rather than the global model for at least two days, and therefore the resolved wet scavenging processes evident in Figure 2 and in the 4 km simulation of Gordon et al (2018) will have affected the aerosol concentrations, and the higher resolution will have had time to affect the winds. In the domain averages we present, we exclude the 20 grid-boxes nearest to the domain boundaries to remove the transition region between the 7 km and 500 m resolution simulations. The number 20 is arbitrary and chosen by eye, and corresponds to around 30 minutes of advection time for a wind speed of 5 ms^{-1} : enough time to produce some more resolved turbulence but not enough time for full mixing of the boundary layer. We are able to run 500 m-resolution simulations driven by the 65 km global model without the intermediate-resolution nest, but then we would need to exclude more grid boxes at the boundaries of the innermost simulation to allow the high resolution structure to spin up.

The boundary layer scheme we use in the 500 m model in all three models is based on that of Lock et al. (2000), which is blended with the Smagorinsky-type scheme from the Met Office Large Eddy Model (Brown, 1999) as grid spacing decreases,

as described by Boutle et al. (2014). The scheme is expected to be dominated by the Lock et al. (2000) scheme rather than the Smagorinsky-type scheme, **even at the 500 m spatial resolution**, while their contributions would be approximately equal at 250 m resolution.

Following the GA7.1 and RA1 configurations of the UM, the global model uses the PC2 subgrid cloud scheme of Wilson et al. (2008) while the regional models use the subgrid cloud scheme of Smith (1990). All three models employ the area adjustment approach of Boutle and Morcrette (2010), so the cloud fraction seen by the microphysics is the mean of that in three sub-layers of each vertical level in each grid cell, while that seen by the radiation code is the maximum.

In the global model and in the higher-resolution regional models, anthropogenic and natural aerosol emissions are taken from the CMIP5 database (Lamarque et al., 2010), except for biomass burning emissions, which are from the Fire Energetics and Emissions Research (FEER) inventory (Ichoku and Ellison, 2014) for August 2017. **In addition, sea spray and dust emissions are represented interactively using the parameterizations of Gong (2003) and Woodward (2001) respectively.** While the configuration used by Gordon et al. (2018) was based on GA6.1 and UM version 10.3 we now use GA7.1 (Walters et al., 2019) and UM version 11.2. The offline-oxidants configuration of United Kingdom Chemistry and Aerosols (UKCA) chemistry is used **together with dust from the CLASSIC aerosol scheme, as in the HadGEM3-GC3.1 climate model (Mulcahy et al., 2020).**

~~The model is~~ **Simulations at all resolutions** are fully coupled to the standard radiative transfer scheme in the UM, via the RADAER module for the aerosols (Bellouin et al., 2013). Direct and semi-direct effects of the absorbing aerosols on the cloud are therefore included in the simulations but will be fairly small in this period due to the relatively low aerosol concentrations, and are not the focus of this work.

Our setup can be viewed as an update of that documented by Gordon et al. (2018). However, as the 500 m-resolution model configuration is different to that published by Gordon et al. (2018) (for example, most aerosol-related settings are upgraded to GA7.1 from GA6.1), many of the tunings used in our previous simulations are no longer required. In the wet scavenging code, Gordon et al. (2018) changed a parameter designed to represent the fraction of the area of a grid box over which rain occurs from 30% to 100%. We reverse this change because while it is still more likely that entire 500 m grid-boxes are raining than entire global model grid boxes, we do not account for the evaporation of rain returning scavenged aerosols to the atmosphere. The 30% parameter can therefore be thought of as the fraction of rain which does not evaporate. We have not verified the accuracy of this assumption, which will clearly depend on the regime studied and should be revisited in future. We no longer tune the dry deposition velocity. The biomass burning emissions diameter (specifically, the number geometric mean diameter) is still 120 nm instead of the default for GLOMAP of 150 nm. This diameter is shown by Shinozuka et al. (2019) to give aerosol dry diameters in reasonable agreement (within 40%) with measurements from the parallel NASA ORACLES campaign (ObseRvations of Aerosols above CLouds and their intEractionS), although the diameter is still slightly overestimated compared to ORACLES. For example, in the lower free troposphere most affected by smoke, the simulated dry diameters are biased 37% high. The mass of organic carbon in the same location, which dominates the overall aerosol mass, is biased high by 8%.

4 Activation, microphysics and coupling

In the global and 7 km models, a diagnostic activation scheme based on the parameterization of Abdul-Razzak and Ghan (2000) is used to calculate the droplet number concentration (West et al., 2014). We refer to this as the ARG (DIAG) scheme later in the paper. On each timestep, the scheme calculates the droplet concentration at cloud base, and imposes it on grid boxes that are above cloud base and still in the same cloud. The droplet concentration does not depend on the concentration on previous timesteps. These diagnostic droplet concentrations are used to calculate autoconversion rates in the single-moment cloud microphysics scheme of Wilson and Ballard (1999). This activation scheme is referred to as the ‘UKCA activation scheme’ later in the paper.

In our 500 m-resolution simulations, aerosols activate in the CASIM code to form cloud droplets, also using the ‘ARG’ parameterization of Abdul-Razzak and Ghan (2000). Other examples of the use of this parameterization in models at or near the cloud-resolving scale are documented by Ghan et al. (2011), Table 3. The treatment, subsequently referred to as the ‘ARG (PROG)CASIM activation scheme’ where PROG refers to the prognostic droplet concentration, is called once per timestep in grid-boxes when a non-zero mass of water is condensing from the vapor phase. The updraft speed used in the ARG scheme is set equal to the grid-box mean updraft speed or 0.001 ms^{-1} , whichever is higher. We reduce this threshold from the original 0.1 ms^{-1} in this paper, which avoids an unphysical spike in the distribution of droplets (Supplementary Figure S1) and an underestimation of the frequency of low droplet concentrations. Instead of the tuned values used by Gordon et al. (2018), the hygroscopicities used are now those recommended by Petters and Kreidenweis (2007) including a kappa value for organic carbon of 0.2. If the number of droplets activated in a timestep exceeds the number of droplets already existing in that grid cell, the droplet concentration used in the microphysics and radiation schemes is updated to the new value. If, on the other hand, the cloud fraction in the grid box goes down, the droplet concentration is altered in proportion, as discussed later. Cloud droplet number concentrations in our 500 m-resolution simulation are also calculated diagnostically on each timestep by the UKCA ARG (DIAG) scheme (West et al., 2014), using the same procedure as in the global model, described above. These number concentrations are not used by the model’s microphysics or radiation scheme, and so in future simulations this parameterization could be switched off. However, for this study we leave it switched on in order to examine its scale invariance and to compare the predicted droplet concentrations with those from CASIM ARG (PROG).

In the CASIM microphysics sub-grid cloud parameterization, described by Smith (1990), condensation of water vapor onto cloud droplets is treated with ‘saturation adjustment’ so supersaturation is not prognostic and droplets are assumed to be in equilibrium at the end of each model timestep. We note that because we use a sub-grid cloud scheme, the grid-box mean equilibrium relative humidity in clouds may be below 100%. In the CASIM microphysics scheme, autoconversion and accretion are handled by the parameterization of Khairoutdinov and Kogan (2000). In our simulations the clouds are entirely warm phase; for a description of the representation of cold clouds in CASIM see Miltenberger et al. (2018). In addition to our reduction of the activation-updraft minimum minimum updraft speed used in the activation scheme, we also change the cloud droplet size distribution assumed by the bulk scheme from an exponential distribution to a gamma distribution. This change is explained further in the context of the evaluation in Section 6.7.

The coupling from the ~~UKCA~~**GLOMAP** aerosol microphysics code to the CASIM cloud microphysics proceeds simply by passing the aerosol mass and number in the soluble Aitken, accumulation and coarse modes, and the volume-weighted hygroscopicities **kappa values** of these modes, to the activation scheme in CASIM. **The kappa values are parameters which describe the hygroscopicity of an aerosol chemical component (Petters and Kreidenweis, 2007).** The CASIM microphysical process rates are coupled to ~~UKCA~~ **GLOMAP aerosols** following the procedure used in the default configuration of the UM with single-moment microphysics from Wilson and Ballard (1999). The autoconversion and accretion rates are summed and passed back to the aerosol microphysics code to determine the rate of removal of aerosols in droplets by rain, while the rain and snow rates are used to determine the rate of impaction scavenging of aerosol by precipitation. **Autoconversion and accretion** These processes also reduce the prognostic droplet number concentration. The liquid water content is used in the calculation of the rate of conversion of sulfur dioxide to aerosol-phase sulfate inside cloud droplets in the GLOMAP module of the code. The CASIM microphysics code has the capability to simulate ~~aerosol~~**cloud** microphysical processing **of aerosol** (Mildenberger et al., 2018), for example the reduction of aerosol number concentration when cloud droplets collide and coalesce, or the increase in aerosol number concentration when rain evaporates. However, there is no capability to track the composition of the aerosol inside hydrometeors during processing, nor to perform aqueous chemistry in the CASIM module. For simplicity and to save on computational expense, therefore, we do not keep track of aerosols in hydrometeors separately to aerosols outside them, and **we** keep the aqueous chemistry in the GLOMAP code. Aerosols that activate to form cloud droplets are only removed if they are wet scavenged (i.e. they form rain) and if this happens it is irreversible: the evaporation of rain does not return aerosols to the atmosphere. A coupled GLOMAP-CASIM double-moment model that includes **cloud**~~aerosol~~ microphysical and chemical processing **of aerosol** is deferred to future work.

20 5 Developments to the activation scheme

The procedure for aerosol activation adopted in both ~~CASIM and UKCA~~**ARG (DIAG) and ARG (PROG)** activation schemes is to activate aerosols on each timestep as if no cloud were present whenever there is a tendency for water mass to condense. In ~~UKCA~~**ARG (DIAG)**, the number of droplets that exist in the box before the activation scheme is run is not stored, so the new value of the droplet concentration is used regardless of the previous concentrations. In ~~CASIM~~**ARG (PROG)**, the number of droplets already in the grid-box is stored **for the double-moment CASIM microphysics**, and if the new number exceeds the old, the number of droplets is increased to the new value. The overwriting of old droplet concentrations by new droplet concentrations if they exceed the old is the procedure of Stevens et al. (1996); Lohmann (2002), and others, but Lohmann (2002) only activated aerosols at cloud base, and assumed cloud droplet concentrations were uniform in columns within a cloud, a procedure since followed by the ~~UKCA~~**ARG (DIAG)** activation scheme (West et al., 2014). The procedure in CASIM, as in ~~most models that employ the more widely used~~ **in models such as CAM5.0 that employ** the Morrison and Gettelman (2008) activation scheme or its successors, is to do activation **via ARG (PROG)** at all vertical levels within the cloud. **In CASIM,** the maximum supersaturation is diagnosed assuming no hydrometeors are present on each timestep, while above cloud base in CAM5.0, a fixed supersaturation of 0.3% is assumed (Wang et al., 2013). The use of the grid-scale mean updraft in the

CASIMARG (PROG) activation scheme, as in RAMs (Saleeby and Cotton, 2004; Thompson, 2016), differs from the scheme of Morrison and Gettelman (2008), which was written with climate model resolution in mind and therefore uses a turbulent sub-grid-scale updraft (Morrison et al., 2005) instead of the grid-scale mean, which is mostly unresolved at low resolution.

There are two possible mechanisms by which this double-moment scheme the ARG (PROG) scheme for the double-moment

5 CASIM microphysics may overestimate the impact of in-cloud activation, leading to overestimated cloud droplet concentration. First, the effect of existing cloud droplets on supersaturation is neglected by default, but activation is still repeated every timestep at all levels in the cloud. Second, provided a cloud does not evaporate, the cloud droplet number produced depends on the maximum updraft speed over the cloud’s lifetime rather than the mean updraft speed, though it is not clear that this mechanism leads to an overestimate, or is correct.

10 Existing cloud droplets certainly affect the supersaturation of water vapor in the cloud. In models like ours with saturation adjustment, there is an assumption that the concentrations of water vapor and liquid water reach equilibrium over one model timestep, which in our case is 20 s. In pre-existing clouds with sufficiently high liquid water content and droplet number, we may assume the sink of water vapor to unactivated aerosols will be negligible, and therefore we may write, after Squires (1952) and others, for example Politovich and Cooper (1988); Korolev and Mazin (2003), an equation for the time evolution of the
15 supersaturation s , using the notation of Ghan et al. (2011):

$$\frac{ds}{dt} = \alpha(T)w - \gamma^*GN\bar{r}s \quad (1)$$

Here w is updraft velocity, $\alpha(T)$ is a thermodynamic term that relates the updraft to the tendency for water vapor to condense as it cools, N is the droplet number, \bar{r} the number mean droplet radius, γ^* is another term which follows from the thermodynamics of rising moist air with assumptions detailed in Chapter 12 of Pruppacher and Klett (1997), and G is the growth coefficient,

20 which depends on the diffusivity of water vapor in air and on the thermal conductivity of the air. Our prescription here The prescription in Equation 1 is valid only for warm-phase clouds; Korolev and Mazin (2003) describe a more general approach for mixed phase clouds. We correct the diffusivity following the size-independent formulation of Fountoukis and Nenes (2005) except with an accommodation coefficient of 1 as recommended by Laaksonen et al. (2005). If (hypothetically) the system were not in equilibrium and w , T , N and \bar{r} were constant in time, the supersaturation $s(t)$ could then be approximated by (e.g.
25 Grabowski and Wang, 2013)

$$s(t) = s_{eq} + (s_0 - s_{eq}) \exp(-\gamma^*GN\bar{r}(t - t_0)) \quad (2)$$

where $s = s_0$ at $t = t_0$. In this equation, $(\gamma^*GN\bar{r})^{-1}$ may be interpreted as τ , the relaxation timescale for the supersaturation, and, in liquid clouds with sufficiently high water content, the supersaturation relevant to aerosol activation is given by the equilibrium or ‘quasi-steady’ value (Politovich and Cooper, 1988)

30 $s_{eq} = \alpha w \tau.$ (3)

Therefore the concentration of newly activated aerosol in each aerosol mode i , $N_{d,i,new}$, is related to the concentration of aerosol in that mode $N_{a,i}$ in the same way as in the activation parameterizations, by

$$N_{d,i,new} = \frac{1}{2} N_{a,i} \left(1 + \text{Erf} \left(\frac{2 \ln \frac{s_{eq}}{s_{c,i}}}{3\sqrt{2} \ln(\sigma_i)} \right) \right) \quad (4)$$

where σ_i is the mode width and

$$s_{c,i} = r_{a,i}^{-1.5} \sqrt{\frac{4A^3}{27B_i}} \quad (5)$$

is the critical supersaturation. The coefficient A is a function of temperature and B_i of the particle hygroscopicity, $r_{a,i}$ is the geometric mean radius of the i^{th} aerosol mode, and the equation assumes the aerosols are internally mixed (Pruppacher and Klett, 1997). The effect on the supersaturation of the condensation or evaporation of rain is currently neglected in these simulations. While the rain water mass is non-negligible compared to the cloud liquid water mass, above cloud base (where it matters) the product of the rain number and the radius is less than 1% of the product of the cloud number and cloud radius at least in the clouds we study here, so the effect on the relaxation time is negligible. At cloud base, the rain mass concentration can exceed the cloud droplet mass, but the small number of rain droplets still means the effect of rain on the relaxation time is negligible.

In well-established clouds with high liquid water content, Korolev (1995); Ming et al. (2007); Ghan et al. (2011) suggested that using Equation 3 should be a better approximation for the supersaturation than the maximum supersaturation s_{max} generated by activation parameterizations such as ARG. Dearden (2009) tested this approximation in large eddy simulations, and found the maximum supersaturation s_{eq} calculated assuming equilibrium with existing droplets to be a much better approximation than that derived using a precursor to the ARG parameterization (Twomey, 1959), which is valid in the approximately same conditions as ARG: at cloud base. In WRF-chem, Yang et al. (2015) found that implementing the suggestion improved simulated wet scavenging. We use the quasi-steady-state equation only when it produces a lower supersaturation than the ARG parameterization. The more detailed microphysics scheme of Phillips et al. (2007) also uses a diagnostic parameterization similar to that of ARG at cloud base and a different approach above. However, inside clouds Phillips et al. (2007) represent supersaturation prognostically, without saturation adjustment, in contrast to our cruder quasi-steady approximation.

In a general circulation or mesoscale model with a sub-grid cloud fraction scheme, partially cloudy grid boxes must be accounted for. By default in ARG (PROG) within CASIM, the grid-box mean change in cloud droplet number concentration on each timestep is

$$\frac{\Delta \bar{N}_{d,t+1}}{\Delta t} = \frac{F_{t+1} \times \left(\max \left((\tilde{N}_{d,t+1} - \tilde{N}_{d,t} F_t / F_{t+1}), 0 \right) \right)}{\Delta t} \quad (6)$$

where Δt is the model timestep, 20 s in our tests, $\tilde{N}_{d,t+1}$ denotes the newly calculated in-cloud (denoted by tilde) cloud droplet number concentration, $\tilde{N}_{d,t}$ that calculated on the previous timestep, $\bar{N}_{d,t+1}$ denotes the newly calculated grid-box mean (denoted by overbar) cloud droplet number concentration, and F_{t+1} is the current cloud fraction. The \tilde{N}_d is calculated by running the ARG parameterization over the whole grid box on the assumption that it is completely cloudy so the real grid-box

mean cloud droplet number concentration is this value multiplied by the cloud fraction. In some models Δt is set to an activation timescale rather than the model timestep (20 minutes in the case of Morrison and Gettelman (2008)). The multiplication by cloud fraction is done after taking the maximum rather than before to handle the case where the cloud fraction increases but the cloud droplet number decreases. The activation scheme is not run when the cloud fraction decreases. When a cloud evaporates, there is a homogeneous mixing assumption: the cloud is assumed to evaporate uniformly across the grid box so that all of the cloud droplets get smaller as the liquid water content decreases, and they are only removed when the cloud fraction or mass of liquid water reach thresholds close to zero (10^{-10} kg kg $^{-1}$ for liquid water content and 10^{-12} for cloud fraction).

In our improved activation scheme, we assume that in each partially cloudy grid box, the total number of droplets that will either be activated, or remain activated, on a new timestep, denoted by $t + 1$, is the sum of those activated inside the old cloud and those activated in any new cloud that forms. Therefore we replace $\tilde{N}_{d,t+1}$ in Equation 6 by

$$\frac{\tilde{N}_{d,cloud,t+1}F_t + (F_{t+1} - F_t)\tilde{N}_{d,ARG,t+1}}{F_{t+1}}. \quad (7)$$

The first term in this Equation 7 represents aerosols that will be activated at equilibrium in the existing cloud. This term comprises both aerosols that are already contained in large cloud droplets and aerosols that may be newly activated in the existing cloud due to increasing updraft speeds and supersaturations ('secondary activation'). The second term in Equation 7 represents aerosols that will be activated in any additional new cloud that forms in the grid box. We thus obtain

$$\frac{\Delta \bar{N}_{d,t+1}}{\Delta t} = \frac{F_{t+1} \times \left(\max \left(\frac{\tilde{N}_{d,cloud,t+1}F_t + (F_{t+1} - F_t)\tilde{N}_{d,ARG,t+1}}{F_{t+1}} - \frac{\tilde{N}_{d,t}F_t}{F_{t+1}}, 0 \right) \right)}{\Delta t} \quad (8)$$

where F_t is the cloud fraction calculated at the end of the previous timestep and then advected, $\tilde{N}_{d,cloud,t+1}$ is the in-cloud droplet number calculated using Eq. 3 for the supersaturation and $\tilde{N}_{d,ARG}$ is the result of the ARG parameterization (or a similar treatment such as that of Twomey (1959) or Nenes and Seinfeld (2003)). In our model, F_t is available even when the sub-grid cloud scheme is diagnostic rather than prognostic, and the equation is not used if $F_{t+1} - F_t < 0$: if this is the case we use instead

$$\frac{\Delta \bar{N}_{d,t+1}}{\Delta t} = \frac{(F_{t+1} - F_t)\tilde{N}_{d,t}}{\Delta t}. \quad (9)$$

If the cloud fraction is below 5% or the in-cloud supersaturation calculated from the cloud droplets via Eq. 3 is higher than that calculated with the ARG parameterization, we revert to Equation 6, using only the ARG scheme. The equations for the various conditions are summarised in Table 2.

With all of the double-moment approaches we are aware of, over the lifetime of a cloud, the in-cloud droplet number is more likely to increase than it is to decrease, because the droplet number in each gridbox is overwritten if the new droplet number exceeds the old. Therefore, the fraction of activated aerosol will end up corresponding to the highest updraft speed seen during that lifetime. Morrison et al. (2005) divided the number of additional new droplets activated by 2 to help compensate for this. This 'ratcheting' mechanism is not necessarily unrealistic, since turbulence-induced upward fluctuations in updraft speed in clouds may well activate more droplets, while downward fluctuations around a positive mean updraft are unlikely

Table 2. Equations for activation to apply in different possible situations as described in the text

Condition	Equation	Application
$F_{t+1} - F_t > 0$, $F_{t+1} > 0.05$, and $S_{cloud} < S_{arg}$	Eq. 8	Activation inside existing clouds
$F_{t+1} - F_t > 0$, $F_{t+1} < 0.05$ or $S_{cloud} > S_{arg}$	Eq. 6	Activation in newly forming clouds
$F_{t+1} - F_t < 0$	Eq. 9	Evaporating clouds

to lead to large existing droplets completely evaporating. However, the ratcheting means the double-moment scheme should produce more droplets on average than a single-moment scheme **provided the number concentration of droplets is diagnosed from the aerosol concentration in the single-moment scheme**. For example, if the **ARG (DIAG) UKCA** activation scheme were fed by the same updraft speeds and the same aerosol as a double-moment scheme, fewer droplets on average would be produced. The droplet concentrations shown **later**, in Figure 8, suggest that indeed the **CASIMARG (PROG)** scheme does produce more droplets, but the comparison is complicated by the **UKCAARG (DIAG)** procedure of setting vertically constant droplet concentrations above cloud base.

In the version of CASIM we are using, aerosol processing is not included, so aerosols are not removed from the gas phase or tracked separately in the cloud phase. Therefore, in our improved scheme, once a cloud forms and the supersaturation used in the activation scheme is reduced due to the existing droplets, the number of aerosols activated on subsequent timesteps within the cloud will be smaller than the number activated the first time the cloud formed. Because the droplet number is only updated if it is higher than the previous droplet number, this will not generally cause the droplet number to decrease, but it will artificially hinder secondary activation from increasing the droplet number. In other words, while our improvements are designed to reduce secondary activation **above cloud base**, because of the lack of processing we may have reduced it too much. This potential bias could be avoided in future work when aerosol processing is re-introduced to the model.

Aerosol activation is affected by the degree to which updraft speeds in clouds are resolved by a model. If the grid-box mean updraft speed is used in the activation scheme and not all updrafts are resolved, one might expect an underestimate in the concentrations of activated aerosols to result. This bias may counteract the effect of overestimating in-cloud activation discussed earlier. The result of the study of Malavelle et al. (2014) is that the ratio of the variance of simulated updrafts to the variance of real updrafts is given by

$$R = 1 - \frac{\left(\frac{\Delta x}{Z_{ml}}\right)^{E1} + a \left(\frac{\Delta x}{Z_{ml}}\right)^{E2}}{\left(\frac{\Delta x}{Z_{ml}}\right)^{E1} + b \left(\frac{\Delta x}{Z_{ml}}\right)^{E2} + c} \quad (10)$$

for grid resolution Δx , a parameter Z_{ml} proportional to the boundary layer height as described in Supplementary Table S4, and fitted constants $E1 = 2.59$, $E2 = 1.34$, $a = 7.95$, $b = 8.00$ and $c = 1.05$. An additional correction factor f is required to

account for the difference between effective and actual grid resolution. This factor varies from model to model but for the UM it is a factor four correction to the variance (Malavelle et al., 2014). The constant of proportionality between Z_{ml} and the boundary layer height depends on the cloud regime. The updraft velocity in the activation scheme can then be set to

$$w_{act} = w \sqrt{\frac{f}{R}} \quad (11)$$

5 to correct for unresolved fluctuations in the velocity distribution, following Equation 21 in Malavelle et al. (2014).

In the cloudy part of the model domain, the mixed layer height is around 2000 m, so the formulation, accounting for effective grid resolution, suggests a scaling factor of approximately 3 (see Supplementary Table S4). The comparison of the PDF of updraft speeds between the model and observations suggests a factor that varies from 1.5 to 4.2 depending on altitude and liquid water content. We give more weight to results at cloud base and to the data binned by liquid water content in Supplementary Figure S6, so we use a factor of 2. We do not implement the Malavelle et al. (2014) scheme in full as it appears to exaggerate the scaling factor required for the model resolutions and boundary layer we are studying. Moreover, we note that scaling all our updrafts up by a factor two doubles the mean updraft speed used in the activation scheme as well as the standard deviation. For the cloud we simulate, at cloud base and at cloud top this leads to poorer agreement of the mean updraft speed with observations, while in the middle of the cloud the agreement improves. The potential for such a scaling to introduce bias requires further study.

6 Evaluation of the double-moment microphysics

In this section we evaluate the configuration of the model with double-moment microphysics without the additional developments described in the previous section, to establish a baseline for testing these additional developments. All of our evaluation is performed with a snapshot of the model at 1200 UTC on 19 August. The aircraft took off at 1001 UTC, entered the domain of the 500 m model at 1103 UTC, left it at 1310 UTC and landed at 1344 UTC. We assume that changes in meteorological or aerosol variables in the region between 1103 and 1310 UTC are small.

6.1 Satellite liquid water path and cloud droplet number

Figure 3 shows the liquid water path and cloud droplet number in the simulation domain compared to the MODIS instrument on the TERRA satellite (Platnick et al., 2015). The calculation of droplet number concentration from MODIS effective radius, cloud top height and optical depth is described by Gordon et al. (2018) and follows Boers et al. (2006) and Grosvenor and Wood (2014). The cloud feature is reproduced by the model, but it is substantially smaller than that observed by the satellite and it is displaced to the south-east. The south-easterly displacement suggests the wind speed in the boundary layer in the 7 km model is likely a little slower than in reality. The liquid water path seems to match the satellite well in the 7 km simulation. In the 500 m simulation, the in-cloud liquid water path is also realistic, but the relative frequency of intermediate liquid water paths between 100 and 300 gm^{-2} is underestimated compared to the satellite. We were able to increase the cloud cover in the 500 m-resolution model by tuning the critical minimum relative humidity in the sub-grid cloud scheme, as shown in Supplementary

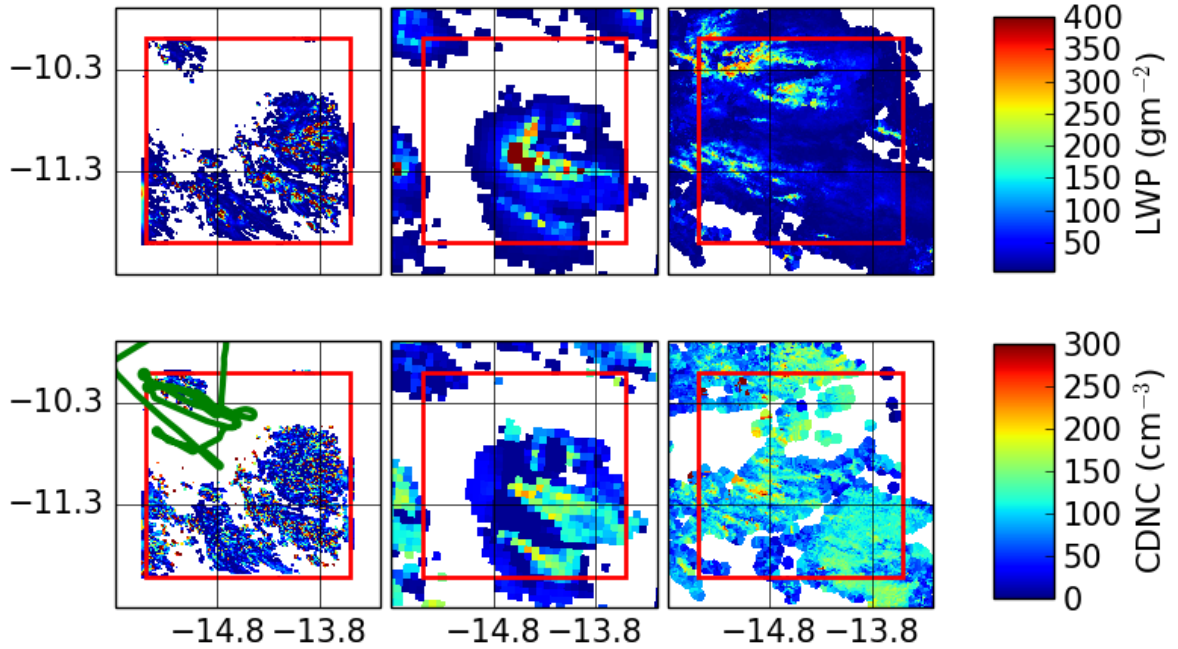


Figure 3. Liquid water path (LWP, top) and cloud droplet number concentration (CDNC, below) compared to MODIS on 19 August at midday. Instantaneous output from the 500 m-resolution simulation is shown on the left, the 7 km-resolution simulation in the centre, and MODIS TERRA on the right. The sizes of isolated MODIS pixels are increased during regridding, for ease of viewing. There is not always a valid MODIS N_d CDNC retrieval where there is a valid liquid water path, so the cloud extent can only be inferred from the top-right figure. The green trace on the bottom-left figure shows the path of the aircraft.

Figure S2. However, to maintain consistency with the official RA1 UM configuration (Bush et al., 2019) we retain the default values of these parameters in the simulations that follow.

The retrieval of cloud droplet number does not yield data in all cloudy pixels due to confusion with overlying cirrus cloud which prevents the retrieval of cloud top temperature. However, from the valid pixels, the cloud droplet concentration in the 7 km simulation is realistic. In the 500 m simulation, a large number of isolated high peaks in the droplet concentration are seen, while there are no such high peaks in the measurements. In the simulation, these peaks correspond to high updraft speeds at some point in the history of the cloud. It is possible that these peaks correspond to the few grid cells where our saturation adjustment assumption has led to excessive release of latent heat, temporarily strengthening the updrafts. In practice, this variability may not bias the area average: Supplementary Figure S3 shows the same figure, except with the central subfigures replaced by the droplet concentration and water path in the 500 m simulation regridded to 5 km resolution, which is not dissimilar to the apparent resolution of the relatively sparsely sampled MODIS cloud droplet number concentration. The simulations in these subfigures show better agreement with the observations.

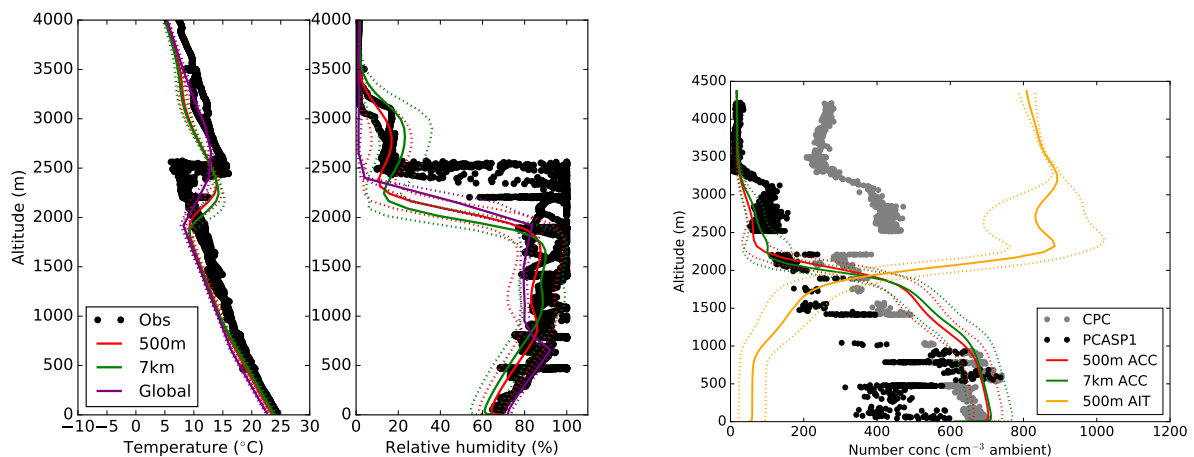


Figure 4. FAAM aircraft observations of temperature (left), relative humidity (centre) and out-of-cloud Aitken (AIT) and accumulation (ACC)-mode aerosol number concentration (right) with the corresponding model means over the 500 m domain. all at or around midday on 19 August 2017. The aircraft data are used whenever they are in the domain of the 500 m-resolution simulation. Variability across the model domain is assumed here to be small, but, for accumulation-mode aerosol, it may be inferred from Figure 2. Dotted lines indicate one standard deviation. Aerosol data inside clouds are removed as described in Section 2.

6.2 Temperature and humidity

Figure 4 shows the aircraft measurements of temperature and relative humidity compared to the global and 7 km-resolution simulations in the left and central sub-figures. These measurements are supported by Supplementary Figure S4, which shows the corresponding data from radiosondes at Ascension Island compared to the regional and global simulations. The simulations

5 clearly underestimate the boundary layer height, by around 400 m, with the 7 km simulation performing slightly worse than the 500 m simulation. The grid spacing at the top of the boundary layer is approximately 80 m. This underestimate will affect our comparison of simulated cloud properties, especially close to cloud top, to observations, as we discuss later. The underestimate is partly due to an underestimate of the boundary layer height by the ERA-interim meteorology used to nudge the global simulation, and it is exacerbated by the 7 km-resolution simulation. We verified that the bias persists in the more up-to-date

10 ERA5 reanalysis. We were able to reduce the underestimate by about 100 m by tuning entrainment parameters in the global model, but as this gain is small relative to the overall 400 m bias, we retain the default configuration in the following analysis. We also do not try to correct for this discrepancy in the evaluation by comparing lower altitudes in the model to higher altitudes in observations, as making such a correction would lead to further complications from the different temperatures at the different altitudes. The radiosondes from Ascension Island, which lies just outside the domain of the 500 m -resolution

15 simulation, indicate that the boundary layer height there is lower than the boundary layer height inside the 500 m simulation domain, by approximately 500 m in the observations and 300 m in the 7 km-resolution model.

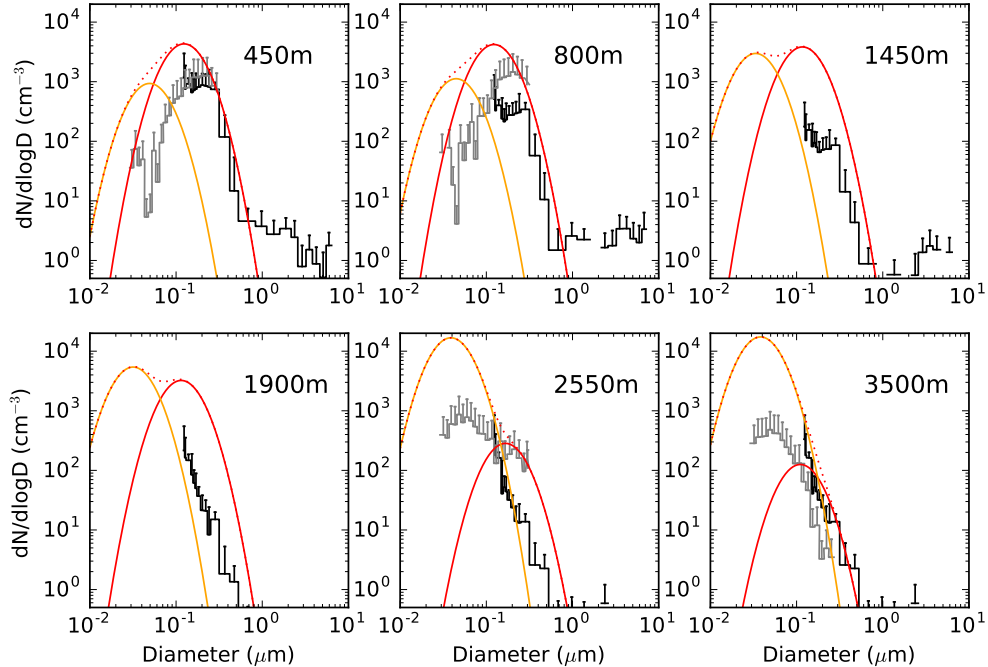


Figure 5. The dry size distributions (out of cloud) observed by the aircraft from a combination of PCASP (black) and SMPS (grey) instruments are shown at different altitudes, compared to the model Aitken (orange) and accumulation mode (red) number concentrations, with the total shown as a red dotted line. The simulated coarse mode is not shown. All data shown are out-of-cloud at standard temperature and pressure. The simulated data are a mean over the domain at 1200 UTC. Only aircraft data from straight-and-level legs at the specified altitude are shown.

As well as the error in the boundary layer height, the temperature is generally underestimated by the simulations by around 1.5°C above the boundary layer, although the simulations do reproduce the hint of a secondary inversion at around 3.5 km altitude reasonably well. The relative humidity is slightly underestimated in most of the boundary layer compared to the aircraft data. Unlike the humidity derived from the aircraft, the radiosondes record lower humidities than the model. This discrepancy may be due to imperfect matching between the clouds in the model grid boxes and the observations as we also saw in Figure 3, or due to effects from Ascension Island, which is not included in the model. Above cloud, all simulations produce an elevated relative humidity in the moderately polluted aerosol layer, in good agreement with the aircraft measurements in Figure 4.

6.3 Aerosol number concentration and size distribution

The simulated vertical profiles of aerosol number concentration are compared to observations on the right of Figure 4. The observed accumulation-mode aerosol number concentration in the boundary layer varies from 350 to 700 cm⁻³ below 1000 m

altitude, and 200 to 400 cm⁻³ between 1000 and 2000 m. The domain mean in the simulation, around 700 cm⁻³ below 1000 m and 600 cm⁻³ above, is biased high compared to the observations. The Aitken mode number concentration above the boundary layer is over-predicted by the model by a larger amount, more than a factor of two, suggesting that new particle formation in the free and upper troposphere may be overestimated. The overestimate could be in one or both of the regional models, or the global model, as these particles may advect in through the boundaries of the regions. The overestimate is due to excessive new particle formation in the upper troposphere in the global simulation. The full vertical profile of Aitken-mode particle concentrations in the global model is shown in Supplementary Figure S5. The bias is also present in the evaluation of Mulcahy et al. (2020). The excessive new particle formation is itself due at least in part to an overestimate of sulfur dioxide concentrations in the model (not shown). The overestimation of the Aitken mode number concentration is likely responsible in part for overestimates in the cloud droplet concentration at the top of the simulated clouds (around 2215 m), as discussed in Section 6.6. rather than biomass burning since the geometric mean diameter of the smoke emissions is 120 nm, which is already in the accumulation mode, and there is no mechanism in the model for the smoke particles to shrink. As this study is focused on the boundary layer, and the concentrations in the Aitken mode there seem realistically similar to the difference between the PCASP concentration and the CPC concentration, we do not investigate further. However, It will be important to study the scale invariance of nucleation-mode microphysics at high resolution, and the biases in its parent model, which also affect the UK contributions to the CMIP6 experiments, in future work.

The observed aerosol size distribution is shown in Figure 5. As suggested by the number concentrations in Figure 4, the Aitken mode concentration is consistently overestimated, but the accumulation mode dry diameter is simulated well, generally within 30% of observations, though it appears to be overestimated by a larger amount at the 1900 m level, the altitude of the cloud layer (note that cloudy grid boxes are excluded from the average). The limited amount of aerosol data available at the level of the clouds precludes a more quantitative comparison at this altitude. This overestimation may indicate too much aqueous sulphate production, or too little rainout of the larger particles. The poor agreement of the simulation and observations at 2550 m is likely to be due at least in part to the underestimation of the boundary layer height by the model. The overestimation of the Aitken mode number concentration is unlikely to lead to substantial biases in the cloud droplet number because, while too high, the concentrations are still much lower than those of the accumulation mode in the boundary layer.

6.4 Liquid water content

The aircraft targeted areas of strong radar reflectivity, and so its sampling of liquid water content is biased towards thicker clouds. This bias must be addressed as we compare the observations with the whole range of model grid cells in our domain, as discussed by Field and Furtado (2016). The distribution of liquid water content at the five altitudes sampled by the aircraft is compared to the distribution at the same altitudes of the cloudy grid-boxes of the 500 m- resolution model in Figure 6. Unsurprisingly, the observed liquid water contents are skewed towards high values while the simulated liquid water contents are mostly much lower. To compare vertical velocities and cloud droplet number concentrations more fairly between simulations and observations, for some of our analysis we split the samples at a liquid water content of 0.15 gkg⁻¹. The difference in the distribution of liquid water content between the model and the observations within each of these bins is then much smaller

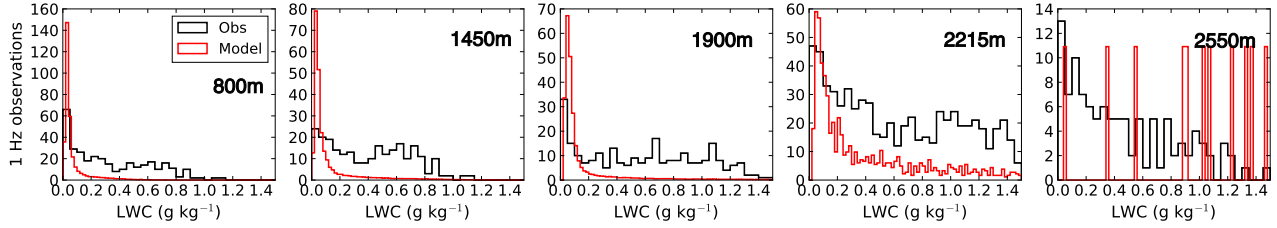


Figure 6. Frequencies of observed in-cloud liquid water content from the cloud droplet probe compared to the liquid water contents in cloudy grid cells in the 500 m-resolution simulation at the five altitudes marked. We include all model grid cells at the specified altitude within the simulation domain with a liquid water content greater than 0.01 g kg^{-1} at the instant of 1200 UTC on 19 August in the histogram, except for the grid cells within 10 km of the domain boundary. The histograms of the simulated liquid water contents are scaled to contain the same number of entries as the histograms of the observations (so, for example, there are 11 model grid boxes plotted at 2550 m altitude).

than the difference over the full range of liquid water content (Figure 6). In some of the subsequent evaluation, we focus on observations and simulation data made at liquid water contents above 0.15 g kg^{-1} , as this is where most of the observations lie. The histograms of vertical velocities in two bins of liquid water content shown in Supplementary Figure S7 show that, as expected, high liquid water content is associated with updrafts while low liquid water content is more likely in downdrafts.

5 6.5 Updrafts

We compare simulated grid-box mean updraft speeds in our 7 km and 500 m-resolutions models to observations in Figure 7. We include all grid boxes where the overall cloud liquid water content is greater than 0.01 g kg^{-1} irrespective of the cloud fraction. Clearly the width of the distribution is underestimated substantially, even in the 500 m simulation. An underestimate is expected, because the 500 m model grid boxes does not resolve all the updrafts. On the other hand, the aircraft records vertical velocity at 32 Hz, corresponding to approximately one measurement for every 4 m it traverses, and therefore it should represent all of the turbulence that would be captured by a large eddy simulation. In our model, we therefore rely on parameterized boundary layer mixing. In the 7 km simulations, almost no variability in updraft speed in these clouds is resolved.

Model resolution explains only part of the underestimated variability in vertical velocity. The other reason the updraft widths are too narrow is sampling biases: the sampling of clouds with high radar reflectivity discussed in the previous section. We compare the moments of the simulated and observed updraft distribution in bins of cloud liquid water content, in Supplementary Figure S6. The distributions and their moments when split into two bins are shown in Supplementary Figure S7 and Supplementary Table S2. When split by liquid water content, the width of the vertical velocity distributions in the model matches observations better, although it is still underestimated. Before splitting by liquid water content, the mean vertical velocity at each altitude was generally within 0.1 ms^{-1} of zero, but after the split this is no longer the case, as indicated in Supplementary Table S2. We note that the two variables are not independent: large positive liquid water contents are correlated to large positive vertical velocities because high vertical velocities imply high supersaturations. The reader is also reminded that the 2215 m

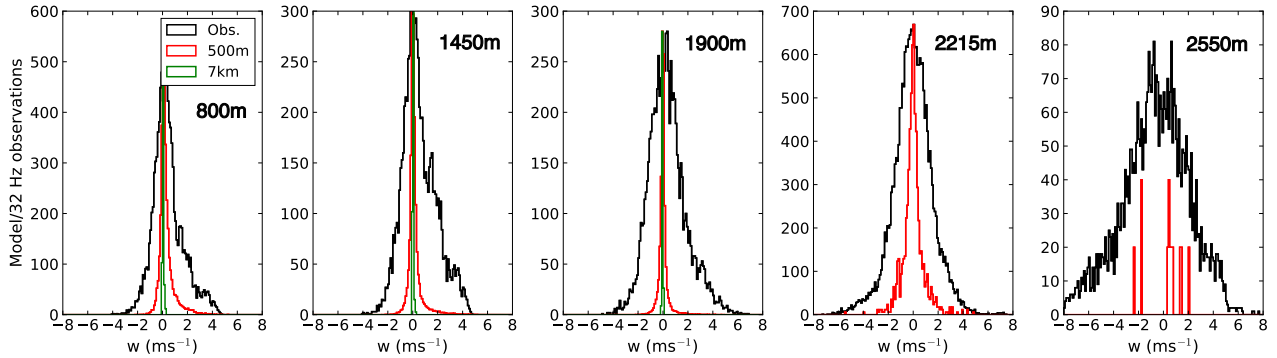


Figure 7. Updraft speed in clouds in 32 Hz CLARIFY observations and in the model (before changes to the activation scheme), at altitudes of 800, 1450, 1900, 2215 m and 2550 m from left to right. **Model data are from instantaneous simulation output of the 500 m and the 7 km-resolution simulations, in both cases across the domain of the 500 m simulation, at 1200 UTC on 19 August 2017.** Data are selected as being in-cloud if the liquid water content exceeds 0.01 g kg^{-1} in both the model and in the observations. The distribution of updraft speeds in the 7 km-resolution driving model is added, showing that at 7 km horizontal resolution very little of the updraft distribution is resolved and a sub-grid turbulent vertical velocity, or a PDF of vertical velocities, should be used in the activation scheme instead. The histograms of simulated vertical velocity are scaled so their maxima match the maximum of the histogram of the observed vertical velocity, except at 2550 m altitude where the simulated velocity histogram is scaled by a factor 20, for clarity.

altitude level is probably more representative of cloud top in the model than the 2550 m level, due to the underestimated height of the boundary layer.

If we approximately correct for the sampling bias associated with the path of the aircraft by only considering vertical velocities where liquid water content exceeds 0.15 g kg^{-1} and also smooth the observed vertical velocity distribution, to show the mean observed vertical velocity over approximately 500 m, we would expect good agreement of the model with observations. When we do this, as shown in Supplementary Figure S8, we do see substantially improved agreement of the model with observations at cloud base and cloud top, but not in the middle of the cloud: the model still underestimates the variability in updraft speed. The residual bias may be a sampling artefact we did not successfully remove, or it may indicate that there is not enough simulated convection.

6.6 Droplet concentrations

The simulated in-cloud droplet number concentration is calculated by first dividing the grid-box mean droplet number concentration by the cloud fraction, then removing any grid cells where the in-cloud mean liquid water content is below 0.01 g kg^{-1} or where the cloud fraction is below 0.05. With the **CASIMARG (PROG)** activation scheme, Table 3 shows that the mean droplet concentration is overestimated by a factor of two to three depending on altitude in the clouds with higher liquid water content, and by a larger factor in clouds with low liquid water content. The overestimate is likely to be due in part to the substantial overestimate in aerosol concentrations in the accumulation mode, and in part due to the activation scheme and warm rain

representation (discussed later). Figure 8 shows the spatial distributions of the simulated and observed droplet concentrations (using the whole sample, not split by liquid water content). We do not plot droplet concentrations at 2550 m altitude due to the low sample size. The figure shows the variability (the width of the distribution) is also overestimated, although the frequency of very low droplet concentrations is underestimated. Table 3 shows the overestimate in mean droplet concentration is more severe than the figure suggests, because the aircraft targeted areas of strong radar reflectivity and therefore preferentially sampled high liquid water contents and high droplet concentrations. At 2215 m altitude, it seems likely that the overestimated mean and width of the droplet concentration spatial distribution is at least partly due to activation of the Aitken mode, given that the number concentration of the Aitken mode is substantially overestimated as discussed in Section 6.3.

The UKCA~~ARG~~ (DIAG) activation scheme in the same simulation also overestimates the variability in droplet concentration, but the mean is closer to observations. Most likely, UKCA~~ARG~~ (DIAG) produces fewer droplets than CASIM~~ARG~~ (PROG) because the CASIM~~ARG~~ (PROG) droplet number depends on the highest updraft in the history of the cloud while the UKCA~~ARG~~ (DIAG) droplet number depends on the updraft at the particular instant the droplet number is diagnosed, as discussed in Section 5. In UKCA~~ARG~~ (DIAG) at 7 km resolution, by contrast, the variability is underestimated and the mean droplet concentration is underestimated, ~~more severely, by up to a factor~~ often around two three. This UKCA~~ARG~~ (DIAG) activation scheme (so far only used at convection-permitting resolution by Gordon et al. (2018)), also produces a spike in the very lowest bin. This spike is the result of a lack of scale-awareness in the scheme as it is currently coded. It corresponds to the minimum droplet concentration being assigned, 5 cm^{-3} , because the characteristic updraft speed is out of the range allowed in the code (a more detailed discussion is given in the supplement). Grid-boxes with updrafts out-of-range are seen in the 7 km-resolution simulations as well as the 500 m simulations. We describe this effect in more detail, and propose a fix to this specific problem, in Supplementary Section 13.

~~As well as the bias mentioned above, the UKCA diagnostic activation scheme suffers from a lack of scale invariance at high resolution in the sense that~~ The lack of scale invariance in the ARG (DIAG) diagnostic activation scheme also means the 7 km simulation does not yield the same distribution of cloud droplet concentrations as the 500 m simulation. This is unsurprising as ~~at 500 m the updraft PDF can often have a substantially non-zero mean in any grid box, while at 7 km it is close to zero, as in the climate model.~~ At 1900 m altitude the simulations agree on the domain-mean droplet concentration: it is 86 cm^{-3} in the 500 m simulation and 92 cm^{-3} in the 7 km simulation, while at 800 m altitude they do not: the droplet concentration is 181 cm^{-3} in the 500 m simulation and 117 cm^{-3} in the 7 km simulation. The reduced variability in the 7 km simulation compared to the 500 m simulation is expected, but further work is needed to ensure the means are consistent. However, because the diagnostic droplet concentrations in the 7 km model do not feed through the lateral boundaries of the 500 m simulations, biased droplet concentrations in our 7 km model should not substantially affect the 500 m model.

The spike at 5 cm^{-3} is not observed in the global climate model and, partly because of this, there is no substantial low bias in the mean droplet concentration. Other reasons for the lack of low bias in the climate model could be the lack of resolved wet scavenging of aerosols. A global evaluation of droplet number in the climate model is presented by Mulcahy et al. (2018). Our main conclusion here is that the UKCA activation scheme should not be used as-is in cloud-resolving simulations, even though the domain mean droplet numbers it produces are realistic in this case study. Indeed, this UKCA activation scheme

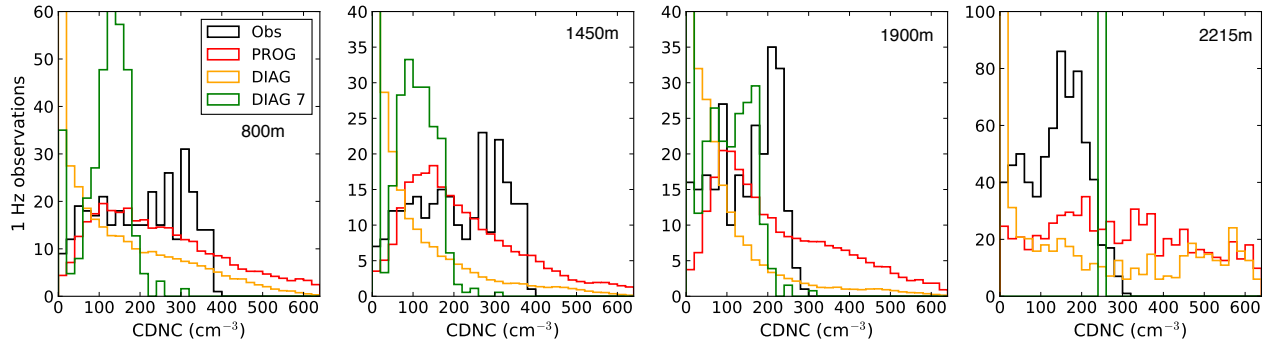


Figure 8. Cloud droplet number concentration in CLARIFY observations at 1 Hz, in the 500 m resolution model as predicted by **CASIM-ARG** (**PROG**), labelled ‘**PROG**’ in **CASIM** and by **UKCA-ARG** (**DIAG**) with its updraft PDF, labelled ‘**DIAG**’, and in the 7 km resolution model as predicted by **UKCA-ARG** (**DIAG**), labelled ‘**DIAG 7**’ at altitudes of 800, 1450, 1900, and 2215 m from left to right. The histograms of simulated cloud droplet number concentration are scaled to contain the same number of entries as the 1 Hz observations. The histograms show droplet numbers over the full range of in-cloud liquid water contents. The observations are compared to gridboxes sampled from the whole model domain, except for the 20 gridboxes nearest the boundaries. Values are grid-box means rather than in-cloud means. Strictly, the aircraft traverses 500 m in about 4 seconds, so 1 Hz is too high a sampling frequency for comparison to the model. However, down-sampling to 0.25 Hz reduces the data sample without changing the shape of the distribution.

must be used with care even at the relatively coarse 7 km resolution (perhaps with an increased minimum width for the updraft distribution).

6.7 Cloud and rain particle size distributions

We compare the cloud and rain size distributions between the model and the observations in Figure 9. We find the model
 5 represents large cloud droplets well, but overestimates the number of small droplets and underestimates the number of large rain drops. The cloud droplet size distributions are broadly similar for cloud liquid water contents below and above 0.15 g kg^{-1} , as shown in the Supplement as Figure S10. The simulated surface rainfall amount is compared to observations from the GPM satellite dataset (Huffman et al., 2014) in Supplementary Figure S11. Despite the poor representation of the rain size distribution
 10 observed by the aircraft, the total surface rainfall rate simulated by the model is in generally good agreement with the satellite data. Simulated vertical profiles of rain mass and number concentration are evaluated in Figure 12, which is shown later as it includes model developments discussed in the following sections **Section 5**. This evaluation confirms that, compared to the aircraft data, the rain mass is generally well simulated but the number concentration of small drops is overestimated and that of very large drops underestimated. We note that the autoconversion and accretion parameterizations from Khairoutdinov and Kogan (2000), or the parameters of the gamma size distribution, could in principle be tuned to attempt to address this bias, but
 15 we did not attempt to do so.

Table 3. Cloud droplet number concentrations in all clouds, and in two bins of liquid water content. The 500m-resolution, 7 km-resolution, and global models are shown, with the ARG (DIAG) and ARG (PROG) activation schemes where they are run. In both model and observations, the threshold liquid water content to define a cloud is 0.01 g kg^{-1} . All data is shown above the first horizontal line, the model grid cells and observations with in-cloud liquid water content above 0.15 g kg^{-1} are shown above the second horizontal line, and those with liquid water content below 0.15 g kg^{-1} are shown below. The global model, labelled ‘Global DIAG’ is not separated by liquid water content as the number of grid-boxes in the domain is too small.

Altitude	Obs (cm^{-3})	500 m PROG (cm^{-3})	500 m (cm^{-3}) DIAG	7km DIAG (cm^{-3})	Global DIAG (cm^{-3})
800	200	253	144	118	180
1450	206	229	88	92	180
1900	144	227	79	92	145
2215	138	301	192	250	180
2550	101	370	364	-	-
800	249	512	280	180	-
1450	247	416	164	139	-
1900	173	421	131	127	-
2215	158	383	234	250	-
2550	126	388	338	-	-
800	117	220	127	116	-
1450	89	183	69	86	-
1900	45	174	65	78	-
2215	45	156	118	-	-
2550	14	182	647	-	-

In the simulations shown in Figure 9, we have used a gamma size distribution instead of the default exponential cloud droplet size distribution assumed by CASIM and published by Grosvenor et al. (2017); Miltenberger et al. (2018) and others. We find (not shown) that once both the moments for droplet mass and number concentration are predicted by the model, adjusting the prescribed size distribution shape (e.g. changing from an exponential to a gamma distribution) without changing these moments does not strongly affect the important results of the microphysics parameterization, such as rain formation rates. However, in Section 5 we the next section we wish to calculate the number mean droplet radius in order to determine supersaturation, and this is sensitive to the shape. For example, an exponential size distribution has a number mean radius equal to one-third of the effective radius and 0.55 times the volume mean radius, while the more realistic gamma size distribution with $\mu = 5$ has a number mean radius equal to 75% of the effective radius and the effective radius is larger than the volume mean radius by 15%. This 15% is still larger than the 8% found in observations (Freud et al., 2011), but in much better agreement than the 66% factor by which the effective radius of the exponential distribution is greater than its volume mean radius. The gamma

size distribution is given by

$$N(D) = \frac{N_d \lambda^{\mu+1} D^\mu e^{-\lambda D}}{\Gamma(\mu+1)} \quad (12)$$

for drop diameter D , shape parameter λ , and droplet concentration N_d . The size distributions shown in Figure 9 and used in our subsequent simulations have $\mu = 5$. The use of $\mu = 5$ is inspired by the relationship of Martin et al. (1994) used by Morrison and Gettelman (2008), which yields $\mu = 5$ for a cloud droplet number concentration of 240 cm^{-3} , close to the concentrations observed here. The exponential size distribution, the size distribution used by Morrison and Gettelman (2008), and the gamma distribution with $\mu = 5$ are compared in Supplementary Figure S9 and the number mean droplet radii that result are tabulated in Supplementary Table S3. The gamma size distribution we use overestimates the number of small droplets compared to observations, but it captures the number of larger cloud droplets very well.

Simulations by Pinsky and Khain (2002) and others showed that in principle a bimodal droplet spectrum could arise as a result of secondary droplet nucleation. We have some indications, from Figure 7, that updrafts may accelerate in our clouds, leading to secondary activation, as strong updrafts, above 4 ms^{-1} , are more frequently observed at higher altitudes. It is then tempting to claim the droplet spectrum in Figure 9 is a bimodal distribution with peaks at $7 \mu\text{m}$ and about $20 \mu\text{m}$. When we filter the observed size distribution to exclude low liquid water contents in which secondary activation is not expected we find (on the right of Supplementary Figure S10) that at high altitudes in the cloud the slightly bimodal distribution remains. However, it could easily be an instrument artefact, and it could probably also arise from different aerosol types, entrainment, or collision-coalescence processes. Thus it is very speculative to suggest secondary activation is actually taking place. In our simulations, by contrast, there is no mechanism for secondary activation to lead to a bimodal size spectrum: the new droplets must fit into the existing size spectrum definition.

7 ~~Developments to the activation scheme II: sub-grid vertical velocity for cloud-resolving resolutions~~

~~An overestimate in droplet number must be balanced to some extent by underestimates, since it has been shown that the default activation scheme performs realistically (Grosvenor et al., 2017; Miltenberger et al., 2018; Stevens et al., 2018). The likely explanation is an underestimate of the turbulence by cloud-resolving simulations which still have much lower resolution than large eddy simulations. The updraft speed distribution is compared to CLARIFY measurements in Figure 7. It is clear that the width of the distribution is significantly lower in the model, even when the bias introduced by the aircraft sampling high radar reflectivity is approximately accounted for.~~

8 Results of improvements to **the ARG (PROG) activation scheme**

For our 500 m simulations, Figure 10 shows that when we correct the **ARG (PROG) activation scheme in two places in our model to account for the effect on supersaturation of existing droplets, and the effect of unresolved vertical velocities**, we obtain relatively similar cloud droplet numbers to those we started with, with a slight improvement to the agreement of the

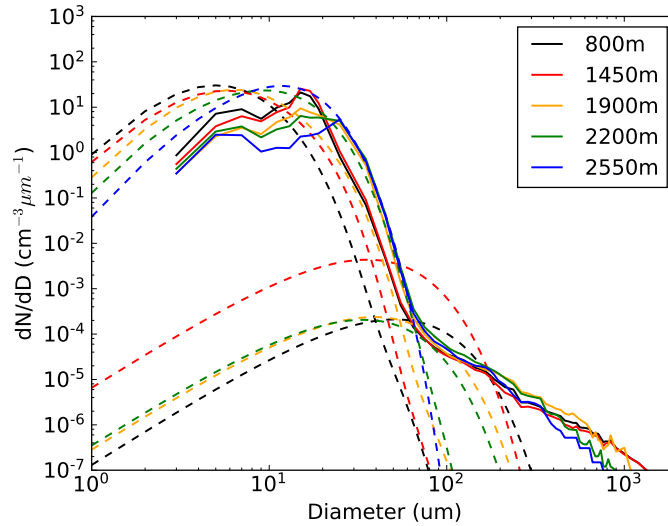


Figure 9. Evaluation of cloud and rain droplet size distributions (in microns) at the altitudes sampled by the aircraft. **The simulations are from the 500 m-resolution model (the only simulation with double-moment cloud microphysics).** The observed size distributions are shown by solid lines and the simulated distributions by dotted lines. The observation data are combined from the cloud droplet probe (CDP), two-dimensional stereo probe (2DS) and cloud imaging probe (CIP). The CDP data and simulated cloud droplet concentrations are filtered to be in-cloud, i.e. to include only samples with $LWC > 0.01 \text{ g kg}^{-1}$, while the CIP data and simulated rain are not.

model with observations. Overall, it looks like the factor 2 correction to the updraft (as suggested by observations at cloud base) produces cloud droplet number in good agreement with the default model for this resolution. However, as well as being resolution-dependent, this finding may be specific to the type of cloud we study, and will likely be influenced by other sources of bias - for example, the assumption that the aerosols are internally mixed, the assumptions about hygroscopicity, or errors in

5 simulating the aerosol size distribution.

The mean maximum supersaturation from the ARG parameterization is around 0.25% at cloud base in the default simulation and around 0.35% when the updraft speed is multiplied by two, in line with expectations (Ghan et al., 2011). The maximum diagnosed supersaturations in new clouds, calculated by the ARG parameterization, and in existing clouds calculated with the quasi-steady-state assumption, are shown as in-cloud domain mean vertical profiles in Figure 10 and as in-cloud histograms in

10 Supplementary Figure S13. They demonstrate that in-cloud activation is clearly not negligible even in this relatively shallow cloud: the mean in-cloud maximum supersaturation can exceed 0.1%. However, in the cloud we study, it is lower than the 0.3% prescribed for convective clouds above cloud base by Wang et al. (2013). The standard deviation of the maximum supersaturations (not shown, for clarity) is also around 0.1% for the case when the updraft speed is multiplied by two, and 0.05% when the updraft is not corrected. The mean in-cloud value is similar to the 0.1% found in a smaller cumulus cloud

15 by Politovich and Cooper (1988). Very occasionally, the relaxation time can exceed the model timestep of 20 s, as shown in the

Table 4. Standard deviation of in-cloud updraft speed in observations and in simulations at three different resolutions, on 19 August 2017 at 1200 UTC. The simulations with 1.5 km and 3.0 km resolution were driven by the same 7 km-resolution simulation as the 500 m -resolution simulation, but, to save CPU, they were run with 224 grid cells in the vertical and horizontal (compared to 450 for the 500 m -resolution simulation). All three simulation domains are centered on 11°S, 14.5°W.

Altitude (m)	σ_w obs. ms^{-1}	σ_w 500 m ms^{-1}	σ_w 1.5 km ms^{-1}	σ_w 3 km ms^{-1}
800	1.23	0.61	0.57	0.40
1450	1.39	0.58	0.42	0.32
1900	1.52	0.44	0.20	0.17
2215	1.60	0.86	0.47	0.25
2550	2.68	1.32	-	1.12

histograms in Supplementary Figure S13, and in this case a bias will result, but the number of grid boxes in which this happens is not significant.

When we account for the effect of existing cloud droplets on supersaturation in the activation scheme, the instantaneous concentration of new droplets activated is always lower above cloud base than when we do not. Increasing the updraft speed increases the activation at cloud base, leading to more activation at cloud base, but, for a factor of two increase in updraft, the number of droplets activated inside the cloud remains lower than in the original model. This difference in vertical profile propagates to the prognostic concentration of cloud droplets, and the consequence is a lower cloud droplet concentration where it is most important for radiative transfer - where the domain-mean liquid water content is highest close to cloud top— compared to the droplet concentration at cloud base.

The spatial distributions of cloud droplet number concentration, shown in Figure 11, match the aircraft slightly better after the constraint our modifications to the activation scheme, as the PDF of the cloud droplet number concentration assembled from the model grid cells narrows slightly. However, the width of the PDF is still biased to be too wide and sometimes too concentrated at low droplet number concentrations.

Table 4 shows a comparison of the simulated and observed width of the spatial distribution of the in-cloud updraft speed before scaling. The factor two correction is clearly appropriate, except at 1900 m altitude where it is too small. By contrast the correction of a factor of three suggested by the parameterization of Malavelle et al. (2014) would be too large. We also ran simulations with 1.5 km and 3 km resolution, where the Malavelle et al. (2014) scheme would suggest a scaling factor of about 8 is needed, and we find similarly that for most altitudes this would increase the width of the updraft PDF substantially beyond that which is observed (Table 4 and Supplementary Figure S12). We show the spatial distributions of updraft speeds across the simulation domains in Table 4. Further work is needed to understand the reasons for this overestimated correction factor fully, but the underlying premise that the updraft may be scaled up to account for its unresolved fraction may still be appropriate.

Our improved activation scheme gives a modest improvement to model performance in the cloud we chose, but we have not demonstrated it would work in any cloud. In a thin cloud with high cloud fraction but low liquid water content, the relaxation time might reach values comparable to the model timestep, which will most likely lead to biases. To avoid unphysical

Table 5. Mean in-cloud cloud droplet number concentration in the three simulations in Figures 10 and 11 (original, supersaturation corrected in clouds, and supersaturation corrected and updraft speed increased) and in the observations from the FAAM aircraft cloud droplet probe. Model grid-boxes and CDP observations where the liquid water content exceeds 0.15 g kg^{-1} are shown above the horizontal line, those with lower liquid water contents below.

Altitude (m)	Obs. cm^{-3}	Default cm^{-3}	(in-cloud SS) cm^{-3}	(in-cl SS, w x 2) cm^{-3}
800	249	512	316	445
1450	247	416	223	326
1900	173	421	228	329
2215	158	383	244	310
2550	126	388	277	329
800	117	220	152	231
1450	89	183	109	172
1900	45	174	98	147
2215	45	156	101	154
2550	14	182	160	236

results, the quasi-steady-state scheme could be switched off if the relaxation time is equal to, or exceeds, the timestep, and the ARG scheme used by itself instead. In our simulation, the relaxation time is either well below the timestep, or the quasi-steady-state supersaturation is higher than the ARG supersaturation, and in this case the quasi-steady-state supersaturation is not used. However, there is no guarantee that these conditions would always be satisfied with different timestep lengths or in different cloud types.

The effect on cloud microphysics of the changes to the activation scheme is detailed in Figure 12. Reducing the cloud droplet number concentration by accounting for the effect of existing clouds on supersaturation increases rain number and mass concentration, unsurprisingly, and increasing the updraft speed in the activation scheme reduces rain formation **via autoconversion and accretion** in the very highest clouds. However, the impact of the results on microphysics is small: in particular, there is no substantial impact on cloud liquid water content in these simulations, and no effect on the altitude at which rain forms.

9 Discussion and conclusions

We have coupled the **UKCA GLOMAP** and CASIM two-moment aerosol and two-moment cloud microphysics components within the Unified Model and tested the resulting model at **500 m horizontal cloud-resolving** resolution against CLARIFY campaign data. **The new configuration is intended to lead to improved simulations of aerosol-cloud interactions. We made some additional developments to the model, summarised in Table 6.**

~~The model emits~~**In the case study we simulated, smoke and marine aerosols are emitted and propagated** to the neighbourhood of Ascension Island, where they interact with clouds and are scavenged by rain. ~~Our evaluation is ambitious as it relies on the~~

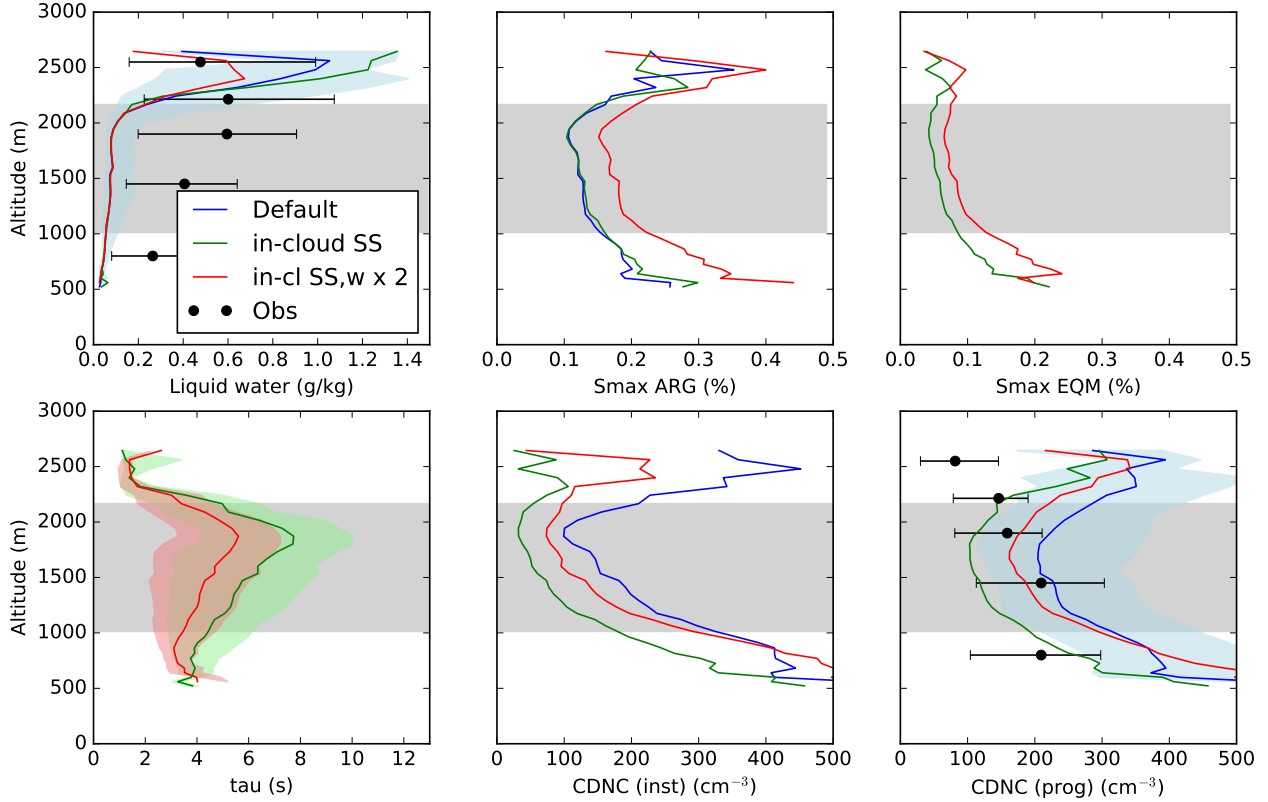


Figure 10. Liquid water content, maximum supersaturation ‘Smax’ calculated via the ARG parameterization and the in-cloud equilibrium ‘EQM’, supersaturation relaxation time ‘tau’, instantaneous concentration of new droplets activated ‘CDNC (inst)’ and prognostic droplet number concentration ‘CDNC (prog)’, at 1200 on 19 August. Each plot shows the old version of the model, labelled ‘Default’, the case where the reduction of supersaturation in clouds is accounted for, ‘in-cloud SS’, and the cases where this reduction is accounted for and the updraft speed in the activation scheme is increased by a factor 2, ‘in-cl SS, w x 2’. In-cloud medians are shown (so the liquid water content is above 0.01 g kg^{-1} and we divide by the cloud fraction), and, except in the case of liquid water content and prognostic cloud drop number (the top-left and bottom-right) the plots also only show means over grid cells where a positive water mass is condensing. The observed median liquid water content and droplet concentrations are shown as black dots on the first and last subfigures. Where shown, error bars or shading indicate the interquartile range either of the sampled observations of the model grid cells. The plot of the relaxation time ‘tau’ only includes grid cells where the relaxation time is used - i.e. where the cloud fraction exceeds 0.05 and the equilibrium supersaturation in the cloudy part of the grid cell is lower than the ARG supersaturation. We also show shaded in grey the vertical band containing 95% of the total liquid water content, because otherwise the eye is drawn to the higher altitudes. Above the shaded region, there are only a few cloudy grid cells, which do not have a strong impact on the domain-mean radiative properties, because the tops of most of the clouds in the domain are at or slightly below the top of the shaded region, not above it.

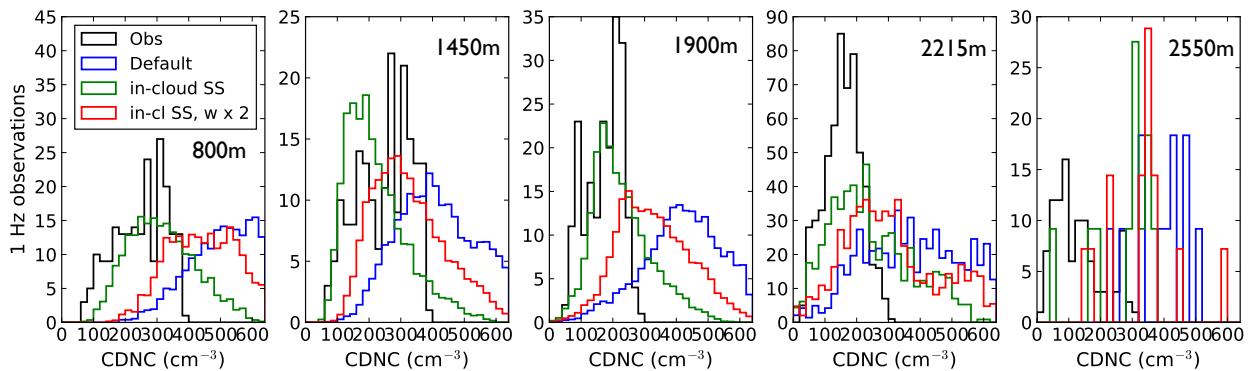


Figure 11. Comparison of in-cloud cloud droplet number distributions in the five straight-and-level runs at altitudes of 800, 1450, 1900, 2215 and 2550 m from left to right. Each plot shows the old version of the model, labelled ‘Default’, the case where the reduction of supersaturation in clouds is accounted for, ‘in-cloud SS’, and the cases where this reduction is accounted for and the updraft speed in the activation scheme is increased by a factor 2, ‘in-cl SS, $w \times 2$ ’. Only model grid-boxes and CDP observations where the liquid water content exceeds 0.15 g kg^{-1} are shown.

model accurately representing this complete cycle. While far from perfect, the model performs well. Further detailed evaluation in different cloud regimes is needed to stress-test the aerosol microphysics at high space and time resolution, as we observed some substantial deviations from observed Aitken-mode number concentrations in the free troposphere, and a high bias in cloud droplet number concentration. The model developments we have made are summarised in Table 6. The case study we focused on, The CLARIFY flight on 19 August 2017 has proved a useful test-bed for model evaluation. The five level in-cloud aircraft flight segments allowed large data samples to be obtained with state-of-the-art instrumentation. Additional profiles and saw-tooth segments are also available for analysis. The cloud deck, which is just under 200 km across, is substantial but self-contained, allowing plenty of data to be obtained both in and out of cloud. The case can be simulated reasonably well with our numerical weather prediction model. While other cloud cases have been measured in more detail, this is one of the best cases from the south-east Atlantic stratocumulus-to-cumulus transition zone and would be suitable for future detailed studies of aerosol-cloud interactions.

Our evaluation of meteorology and of aerosol and cloud microphysics is ambitious as it relies on our hierarchy of simulations accurately representing the complete cycle of aerosol emissions, transport and deposition at the synoptic scale, as well as cloud microphysical processes. We highlighted some areas where the models perform well but also some shortcomings. The 500 m-resolution model simulates the boundary layer clouds realistically, although the top of the boundary layer is low and the cloud we focus on is not quite simulated in the right place. Unfortunately, the low boundary layer top means the matching of the simulated cloud to observations at the same altitudes leads to an imperfect comparison of other cloud properties such as droplet concentration and updraft speed. However, accumulation-mode aerosol, cloud and rain number concentrations are all realistic: simulated mean accumulation-mode aerosol and cloud droplet number concentrations are generally within a factor two of observations, although larger discrepancies exist in certain cloud regimes, for example, droplet number concentration is

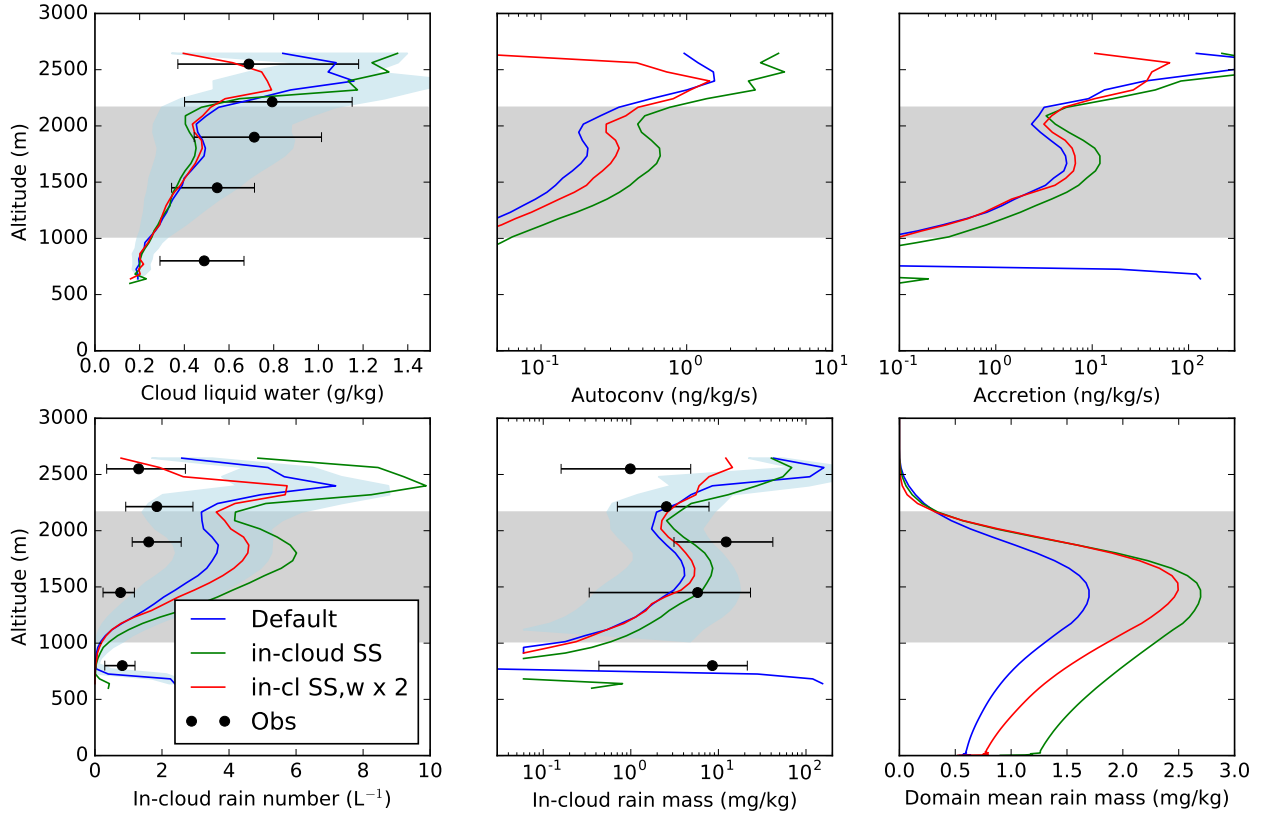


Figure 12. Rain and the rain formation process for observations (black dots) and model grid cells (coloured lines) where the liquid water content exceeds 0.15 g kg^{-1} , in the three simulations shown in Figures 10 and 11 (original, supersaturation corrected in clouds, and supersaturation corrected and updraft speed increased). In CASIM as in many microphysics schemes, rain formation begins with an autoconversion rate, labelled ‘Autoconv’. Rain droplets then accumulate liquid water by accretion. All plots except the last show in-cloud medians where we only include grid-boxes with cloud fraction greater 0.05, and the grid-box value is divided by the cloud fraction. Where shown, error bars and blue shading represents the interquartile range, while the grey shading shows the region where 95% of the domain mean liquid water content resides. The last plot shows the domain mean rain mass mixing ratio with no liquid water content threshold applied.

Table 6. Summary of model developments and adjustments documented in this manuscript

Section	Development
Section 4	GLOMAP-CASIM coupling via activation
Section 4	GLOMAP-CASIM coupling via scavenging/chemistry
Section 6.7	Adjustment of μ in cloud droplet size distribution
Section 4	Reduction of CASIM threshold updraft for ARG (PROG) activation
Section 5	In-cloud droplet activation in ARG (PROG) activation scheme
Section 5	Sub-grid updraft correction in ARG (PROG) activation scheme (only implemented as a case-specific tuning factor)

more severely overestimated in clouds with low liquid water contents. Rain mass is also simulated realistically (within a factor two) but rain number concentration is overestimated by around a factor three. Further work on the global model aerosol code is needed to address the substantial overprediction of Aitken-mode aerosol number concentrations. Cloud liquid water content and rain mass compare reasonably well to satellite measurements. We found the ARG (DIAG) aerosol activation in UKCA without CASIM microphysics breaks down requires a fix to work once high fractions of updraft speeds are partially or fully resolved, and for now CASIM is preferred to activate aerosols in simulations with convection-permitting resolutions of 7 km or finer. However, in the climate model, the same activation scheme has no such problems.

We examined tested two possible improvements to the ARG (PROG) aerosol activation scheme in CASIM: the correction to the supersaturation for existing cloud, and the offline application of the Malavelle et al. (2014) correction to the updraft speeds. A gamma distribution with $\mu = 5$ was found to be needed for the first of these improvements to work, instead of the exponential distribution usually used for the CASIM cloud droplet size distribution, and so all of our results are presented with the gamma distribution. Further work is needed to ensure the Malavelle et al. (2014) correction scales with resolution correctly before it can be implemented online. Conversely, the CASIM ARG (PROG) activation scheme will break at coarser model resolution, as it is not possible to apply the Malavelle et al. (2014) correction unless some fraction of the updrafts are resolved. To get around this a sub-grid vertical velocity of the form suggested by Morrison and Gettelman (2008), or a PDF as in UKCA ARG (DIAG), could be implemented into CASIM and switched on for low grid resolutions (probably coarser than around 3 km, depending on the cloud type) so that aerosols will still activate when the grid-box mean updraft speed is zero. We also recommend reducing the minimum updraft speed in the CASIM ARG (PROG) activation scheme from 0.1 ms^{-1} to 0.001 ms^{-1} to avoid unphysical spikes in the distribution of cloud droplet number.

We have now made the model more physically sound, by reducing incorrect in-cloud activation and using vertical velocities to activate aerosol that are closer to the real vertical velocities observed, and without introducing any new tuning parameters or computational expense. For high-resolution simulations, we believe the approach to secondary activation we adopted following Korolev (1995); Ming et al. (2007); Ghan et al. (2011); Yang et al. (2015) is relatively generalizable and could readily be applied to other models that represent sub-grid cloud fraction, for example WRF-chem with Morrison and Gettelman (2008)

or Gettelman and Morrison (2015) microphysics, and other activation schemes such as that of Nenes and Seinfeld (2003). We note that the Morrison and Gettelman (2008) microphysics uses an updraft speed that is already adjusted for sub-grid turbulence and so has no need of the Malavelle et al. (2014) correction. Implementing only the correction to the supersaturation, and no correction to the updrafts, would lead to lower simulated droplet concentrations.

- 5 There remain arguments for either switching off droplet nucleation above cloud base altogether, assuming the cloud droplet concentration is uniform in vertical columns above cloud base, or assuming the supersaturation is constant above cloud base as done in convective clouds in CAM5 by Wang et al. (2013). In-cloud activation leads to a broadening of the cloud droplet size distribution above cloud base, but with our approach, any new droplets activated above cloud base must follow the same size distribution as existing droplets (Khain et al., 2015). However, we have still improved on the original method where the
- 10 importance of activation above cloud base was substantially exaggerated.

- The improvements to the CASIM **microphysics and its ARG (PROG) activation scheme** may lead to more reliable simulations of aerosol activation and aerosol-cloud interactions that depend more on the concentration of fresh aerosols at cloud base and less on the more processed aerosols found higher in the cloud than previous simulations using similar models, while still allowing activation above cloud base. However, we emphasise that we have not tested the improved model in deep or cold
- 15 clouds, and we do not suggest it could replace more sophisticated activation parameterizations, some of which already calculate in-cloud supersaturation prognostically (e.g. Fan et al., 2018). In particular, the explicit, prognostic calculation is likely to be needed for calculations of convective invigoration since it is likely that latent heat release and cloud evaporation rates will be erroneous in bulk models with saturation adjustment (e.g. Grabowski, 2007; Hill et al., 2008; Lebo et al., 2012). However, more approximate calculations, used in many if not most convection-permitting studies, may benefit from the improvements to
- 20 the procedure for aerosol activation we tested, with no additional computational cost.

10 Author contributions

- HG, KSC, and PRF developed the concepts and ideas for the direction of the paper. HG, PRF, DPG, AAH and JW coupled the GLOMAP and CASIM aerosol and cloud microphysics codes. SJA, PB, KB, IC, ZC, JT, and HW collected and calibrated the measurement data. HG carried out and analysed the model runs, performed the model evaluation with input from PRF, SJA
- 25 and PB, and wrote the paper with input and comments from KSC, PRF, SJA, PB, ZC, DPG, and AAH.

11 Acknowledgements

- This research was funded by the Natural Environment Research Council (NERC) CLARIFY project NE/L013479/1. We thank the Facility for Airborne Atmospheric Measurements, who operate the aircraft and some of the instrumentation used for the CLARIFY campaign, and everyone involved with the campaign. Radiosonde data was obtained from the Atmospheric Radiation Measurement (ARM) Climate Research Facility, a U.S. Department of Energy Office of Science user facility sponsored
- 30 by the Office of Biological and Environmental Research. The Terra/MODIS cloud optical depth and effective radius datasets

were acquired from the Level-1 and Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC), located in the Goddard Space Flight Center in Greenbelt, Maryland (<https://ladsweb.nascom.nasa.gov/>). We acknowledge use of the Monsoon2 system, a collaborative facility supplied under the Joint Weather and Climate Research Programme, a strategic partnership between the UK Met Office and NERC.

5 12 Data availability

Data from the FAAM aircraft are available on the CEDA repository <http://archive.ceda.ac.uk/>. Simulation data are stored at the Met Office and available from the corresponding author on reasonable request. The Met Office Unified Model is available for use under licence. A number of research organisations and national meteorological services use the UM in collaboration with the Met Office to undertake basic atmospheric process research, produce forecasts, develop the UM code, and build and
10 evaluate Earth system models. For further information on how to apply for a licence, see

<http://www.metoffice.gov.uk/research/modelling-systems/unified-model>

(last access: 1 June 2020).

13 Competing interests

The authors declare that they have no conflict of interest.

References

- Abdul-Razzak, H. and Ghan, S. J.: A parameterization of aerosol activation: 2. Multiple aerosol types, *Journal of Geophysical Research: Atmospheres*, 105, 6837–6844, doi:10.1029/1999JD901161, <http://dx.doi.org/10.1029/1999JD901161>, 2000.
- Abel, S. J. and Boutle, I. A.: An improved representation of the raindrop size distribution for single-moment microphysics schemes, *Quarterly Journal of the Royal Meteorological Society*, 138, 2151–2162, doi:10.1002/qj.1949, <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.1949>, 2012.
- Arnason, G. and Brown, P. S.: Growth of cloud droplets by condensation: A problem in computational stability, *Journal of the Atmospheric Sciences*, 28, 72–77, doi:10.1175/1520-0469(1971)028<0072:GOCDBC>2.0.CO;2, [https://doi.org/10.1175/1520-0469\(1971\)028<0072:GOCDBC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1971)028<0072:GOCDBC>2.0.CO;2), 1971.
- 10 Barrett, P. A., Blyth, A., Brown, P. R. A., and Abel, S. J.: The structure of turbulence and mixed-phase cloud microphysics in a highly supercooled altocumulus cloud, *Atmospheric Chemistry and Physics Discussions*, 2019, 1–36, doi:10.5194/acp-2019-749, <https://www.atmos-chem-phys-discuss.net/acp-2019-749/>, 2019.
- Bellouin, N., Mann, G. W., Woodhouse, M. T., Johnson, C., Carslaw, K. S., and Dalvi, M.: Impact of the modal aerosol scheme GLOMAP-mode on aerosol forcing in the Hadley Centre Global Environmental Model, *Atmospheric Chemistry and Physics*, 13, 3027–3044, doi:10.5194/acp-13-3027-2013, <https://www.atmos-chem-phys.net/13/3027/2013/>, 2013.
- 15 Boers, R., Acarreta, J. R., and Gras, J. L.: Satellite monitoring of the first indirect aerosol effect: Retrieval of the droplet concentration of water clouds, *Journal of Geophysical Research: Atmospheres*, 111, n/a–n/a, doi:10.1029/2005JD006838, <http://dx.doi.org/10.1029/2005JD006838>, d22208, 2006.
- Boutle, I., Price, J., Kudzotsa, I., Kokkola, H., and Romakkaniemi, S.: Aerosol–fog interaction and the transition to well-mixed radiation fog, *Atmospheric Chemistry and Physics*, 18, 7827–7840, doi:10.5194/acp-18-7827-2018, <https://www.atmos-chem-phys.net/18/7827/2018/>, 2018.
- 20 Boutle, I. A. and Morcrette, C. J.: Parametrization of area cloud fraction, *Atmospheric Science Letters*, 11, 283–289, doi:10.1002/asl.293, <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/asl.293>, 2010.
- Boutle, I. A., Eyre, J. E. J., and Lock, A. P.: Seamless stratocumulus simulation across the turbulent gray zone, *Monthly Weather Review*, 142, 1655–1668, doi:10.1175/MWR-D-13-00229.1, <https://doi.org/10.1175/MWR-D-13-00229.1>, 2014.
- 25 Brown, A. R.: The sensitivity of large-eddy simulations of shallow cumulus convection to resolution and subgrid model, *Quarterly Journal of the Royal Meteorological Society*, 125, 469–482, doi:10.1002/qj.49712555405, <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49712555405>, 1999.
- Bush, M., Allen, T., Bain, C., Boutle, I., Edwards, J., Finnenkoetter, A., Franklin, C., Hanley, K., Lean, H., Lock, A., Manners, J., Mittermaier, M., Morcrette, C., North, R., Petch, J., Short, C., Vosper, S., Walters, D., Webster, S., Weeks, M., Wilkinson, J., Wood, N., and Zerroukat, M.: The first Met Office Unified Model/JULES Regional Atmosphere and Land configuration, RAL1, *Geoscientific Model Development Discussions*, 2019, 1–47, doi:10.5194/gmd-2019-130, <https://www.geosci-model-dev-discuss.net/gmd-2019-130/>, 2019.
- 30 Clark, T. L.: Numerical Modeling of the Dynamics and Microphysics of Warm Cumulus Convection, *Journal of the Atmospheric Sciences*, 30, 857–878, doi:10.1175/1520-0469(1973)030<0857:NMOTDA>2.0.CO;2, [https://doi.org/10.1175/1520-0469\(1973\)030<0857:NMOTDA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1973)030<0857:NMOTDA>2.0.CO;2), 1973.
- 35

- Costantino, L. and Bréon, F.-M.: Aerosol indirect effect on warm clouds over South-East Atlantic, from co-located MODIS and CALIPSO observations, *Atmospheric Chemistry and Physics*, 13, 69–88, doi:10.5194/acp-13-69-2013, <https://www.atmos-chem-phys.net/13/69/2013/>, 2013.
- Dearden, C.: Investigating the simulation of cloud microphysical processes in numerical models using a one-dimensional dynamical framework, *Atmospheric Science Letters*, 10, 207–214, doi:10.1002/asl.239, <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/asl.239>, 2009.
- Fan, J., Leung, L. R., Li, Z., Morrison, H., Chen, H., Zhou, Y., Qian, Y., and Wang, Y.: Aerosol impacts on clouds and precipitation in eastern China: Results from bin and bulk microphysics, *Journal of Geophysical Research: Atmospheres*, 117, doi:10.1029/2011JD016537, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JD016537>, 2012.
- 10 Fan, J., Rosenfeld, D., Zhang, Y., Giangrande, S. E., Li, Z., Machado, L. A. T., Martin, S. T., Yang, Y., Wang, J., Artaxo, P., Barbosa, H. M. J., Braga, R. C., Comstock, J. M., Feng, Z., Gao, W., Gomes, H. B., Mei, F., Pöhlker, C., Pöhlker, M. L., Pöschl, U., and de Souza, R. A. F.: Substantial convection and precipitation enhancements by ultrafine aerosol particles, *Science*, 359, 411–418, doi:10.1126/science.aan8461, <https://science.sciencemag.org/content/359/6374/411>, 2018.
- Field, P. R. and Furtado, K.: How biased is aircraft cloud sampling?, *Journal of Atmospheric and Oceanic Technology*, 33, 185–189, doi:10.1175/JTECH-D-15-0148.1, <https://doi.org/10.1175/JTECH-D-15-0148.1>, 2016.
- 15 Fountoukis, C. and Nenes, A.: Continued development of a cloud droplet formation parameterization for global climate models, *Journal of Geophysical Research: Atmospheres*, 110, doi:10.1029/2004JD005591, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JD005591>, 2005.
- Freud, E., Rosenfeld, D., and Kulkarni, J. R.: Resolving both entrainment-mixing and number of activated CCN in deep convective clouds, *Atmospheric Chemistry and Physics*, 11, 12 887–12 900, doi:10.5194/acp-11-12887-2011, <https://www.atmos-chem-phys.net/11/12887/2011/>, 2011.
- 20 Fridlind, A. M., Ackerman, A. S., Jensen, E. J., Heymsfield, A. J., Poellot, M. R., Stevens, D. E., Wang, D., Miloshevich, L. M., Baumgardner, D., Lawson, R. P., Wilson, J. C., Flagan, R. C., Seinfeld, J. H., Jonsson, H. H., VanReken, T. M., Varutbangkul, V., and Rissman, T. A.: Evidence for the predominance of mid-tropospheric aerosols as subtropical anvil cloud nuclei, *Science*, 304, 718–722, doi:10.1126/science.1094947, <https://science.sciencemag.org/content/304/5671/718>, 2004.
- 25 Furtado, K., Field, P. R., Luo, Y., Liu, X., Guo, Z., Zhou, T., Shipway, B. J., Hill, A. A., and Wilkinson, J. M.: Cloud microphysical factors affecting simulations of deep convection during the presummer rainy season in Southern China, *Journal of Geophysical Research: Atmospheres*, 123, 10,477–10,505, doi:10.1029/2017JD028192, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2017JD028192>, 2018.
- 30 Gettelman, A. and Morrison, H.: Advanced two-Moment bulk microphysics for global models. Part I: off-Line tests and comparison with other schemes, *Journal of Climate*, 28, 1268–1287, doi:10.1175/JCLI-D-14-00102.1, <https://doi.org/10.1175/JCLI-D-14-00102.1>, 2015.
- Ghan, S. J., Abdul-Razzak, H., Nenes, A., Ming, Y., Liu, X., Ovchinnikov, M., Shipway, B., Meskhidze, N., Xu, J., and Shi, X.: Droplet nucleation: physically-based parameterizations and comparative evaluation, *Journal of Advances in Modeling Earth Systems*, 3, n/a–n/a, doi:10.1029/2011MS000074, <http://dx.doi.org/10.1029/2011MS000074>, m10001, 2011.
- 35 Gong, S. L.: A parameterization of sea-salt aerosol source function for sub- and super-micron particles, *Global Biogeochemical Cycles*, 17, doi:10.1029/2003GB002079, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003GB002079>, 2003.

- Gordon, H., Field, P. R., Abel, S. J., Dalvi, M., Grosvenor, D. P., Hill, A. A., Johnson, B. T., Miltenberger, A. K., Yoshioka, M., and Carslaw, K. S.: Large simulated radiative effects of smoke in the south-east Atlantic, *Atmospheric Chemistry and Physics*, 18, 15 261–15 289, doi:10.5194/acp-18-15261-2018, <https://www.atmos-chem-phys.net/18/15261/2018/>, 2018.
- Grabowski, W. W.: Representation of turbulent mixing and buoyancy reversal in bulk cloud models, *Journal of the Atmospheric Sciences*, 5 64, 3666–3680, doi:10.1175/JAS4047.1, <https://doi.org/10.1175/JAS4047.1>, 2007.
- Grabowski, W. W. and Jarecka, D.: Modeling condensation in shallow nonprecipitating convection, *Journal of the Atmospheric Sciences*, 72, 4661–4679, doi:10.1175/JAS-D-15-0091.1, <https://doi.org/10.1175/JAS-D-15-0091.1>, 2015.
- Grabowski, W. W. and Morrison, H.: Modeling condensation in deep convection, *Journal of the Atmospheric Sciences*, 74, 2247–2267, doi:10.1175/JAS-D-16-0255.1, <https://doi.org/10.1175/JAS-D-16-0255.1>, 2017.
- 10 Grabowski, W. W. and Wang, L.-P.: Growth of cloud droplets in a turbulent environment, *Annual Review of Fluid Mechanics*, 45, 293–324, doi:10.1146/annurev-fluid-011212-140750, <https://doi.org/10.1146/annurev-fluid-011212-140750>, 2013.
- Grosvenor, D. P. and Wood, R.: The effect of solar zenith angle on MODIS cloud optical and microphysical retrievals within marine liquid water clouds, *Atmospheric Chemistry and Physics*, 14, 7291–7321, doi:10.5194/acp-14-7291-2014, <https://www.atmos-chem-phys.net/14/7291/2014/>, 2014.
- 15 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 closed-cell stratocumulus: insight from the modelling of a case study, *Atmospheric Chemistry and Physics*, 17, 5155–5183, doi:10.5194/acp-17-5155-2017, <https://www.atmos-chem-phys.net/17/5155/2017/>, 2017.
- Guichard, F. and Couvreux, F.: A short review of numerical cloud-resolving models, *Tellus A: Dynamic Meteorology and Oceanography*, 69, 1373 578, doi:10.1080/16000870.2017.1373578, <https://doi.org/10.1080/16000870.2017.1373578>, 2017.
- 20 Hall, W. D.: A detailed microphysical model within a two-dimensional dynamic framework: model description and preliminary results, *Journal of the Atmospheric Sciences*, 37, 2486–2507, doi:10.1175/1520-0469(1980)037<2486:ADMMWA>2.0.CO;2, [https://doi.org/10.1175/1520-0469\(1980\)037<2486:ADMMWA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1980)037<2486:ADMMWA>2.0.CO;2), 1980.
- Heymsfield, A. J., Dye, J. E., and Biter, C. J.: Overestimates of entrainment from wetting of aircraft temperature sensors in cloud, *Journal of Applied Meteorology*, 18, 92–95, doi:10.1175/1520-0450(1979)018<0092:OOEFWO>2.0.CO;2, [https://doi.org/10.1175/1520-0450\(1979\)018<0092:OOEFWO>2.0.CO;2](https://doi.org/10.1175/1520-0450(1979)018<0092:OOEFWO>2.0.CO;2), 1979.
- 25 Heymsfield, A. J., Bansemer, A., Heymsfield, G., and Fierro, A. O.: Microphysics of maritime tropical convective updrafts at temperatures from -20° to -60° , *Journal of the Atmospheric Sciences*, 66, 3530–3562, doi:10.1175/2009JAS3107.1, <https://doi.org/10.1175/2009JAS3107.1>, 2009.
- Hill, A. A., Dobbie, S., and Yin, Y.: The impact of aerosols on non-precipitating marine stratocumulus. I: Model description and prediction 30 of the indirect effect, *Quarterly Journal of the Royal Meteorological Society*, 134, 1143–1154, 2008.
- Hill, A. A., Shipway, B. J., and Boutle, I. A.: How sensitive are aerosol-precipitation interactions to the warm rain representation?, *Journal of Advances in Modeling Earth Systems*, 7, 987–1004, doi:10.1002/2014MS000422, <http://dx.doi.org/10.1002/2014MS000422>, 2015.
- Honnert, R., Masson, V., and Couvreux, F.: A diagnostic for evaluating the representation of turbulence in atmospheric models at the kilometric scale, *Journal of the Atmospheric Sciences*, 68, 3112–3131, doi:10.1175/JAS-D-11-061.1, <https://doi.org/10.1175/JAS-D-11-061.1>, 35 2011.
- Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., and Xie, P.: Integrated Multi-satellitE Retrievals for GPM (IMERG), version 4.4, NASA’s Precipitation Processing Center, doi:10.5067/GPM/IMERG/3B-HH/05, 2014.

- Ichoku, C. and Ellison, L.: Global top-down smoke-aerosol emissions estimation using satellite fire radiative power measurements, *Atmospheric Chemistry and Physics*, 14, 6643–6667, doi:10.5194/acp-14-6643-2014, <https://www.atmos-chem-phys.net/14/6643/2014/>, 2014.
- Khain, A. and Lynn, B.: Simulation of a supercell storm in clean and dirty atmosphere using weather research and forecast model with spectral bin microphysics, *Journal of Geophysical Research: Atmospheres*, 114, doi:10.1029/2009JD011827, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JD011827>, 2009.
- Khain, A. P., Phillips, V., Benmoshe, N., and Pokrovsky, A.: The Role of Small Soluble Aerosols in the Microphysics of Deep Maritime Clouds, *Journal of the Atmospheric Sciences*, 69, 2787–2807, doi:10.1175/2011JAS3649.1, <https://doi.org/10.1175/2011JAS3649.1>, 2012.
- Khain, A. P., Beheng, K. D., Heymsfield, A., Korolev, A., Krichak, S. O., Levin, Z., Pinsky, M., Phillips, V., Prabhakaran, T., Teller, A., van den Heever, S. C., and Yano, J.-I.: Representation of microphysical processes in cloud-resolving models: Spectral (bin) microphysics versus bulk parameterization, *Reviews of Geophysics*, 53, 247–322, doi:10.1002/2014RG000468, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014RG000468>, 2015.
- Khairoutdinov, M. and Kogan, Y.: A New cloud physics parameterization in a Large-Eddy Simulation model of marine stratocumulus, *Monthly Weather Review*, 128, 229–243, doi:10.1175/1520-0493(2000)128<0229:ANCPPI>2.0.CO;2, [https://doi.org/10.1175/1520-0493\(2000\)128<0229:ANCPPI>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<0229:ANCPPI>2.0.CO;2), 2000.
- Kogan, Y. L. and Martin, W. J.: Parameterization of bulk condensation in numerical cloud models, *Journal of the Atmospheric Sciences*, 51, 1728–1739, doi:10.1175/1520-0469(1994)051<1728:POBCIN>2.0.CO;2, [https://doi.org/10.1175/1520-0469\(1994\)051<1728:POBCIN>2.0.CO;2](https://doi.org/10.1175/1520-0469(1994)051<1728:POBCIN>2.0.CO;2), 1994.
- Korolev, A., Strapp, J. W., Isaac, G. A., and Emery, E.: Improved airborne hot-wire measurements of ice water content in clouds, *Journal of Atmospheric and Oceanic Technology*, 30, 2121–2131, doi:10.1175/JTECH-D-13-00007.1, <https://doi.org/10.1175/JTECH-D-13-00007.1>, 2013.
- Korolev, A. V.: The influence of supersaturation fluctuations on droplet size spectra formation, *Journal of the Atmospheric Sciences*, 52, 3620–3634, doi:10.1175/1520-0469(1995)052<3620:TIOSFO>2.0.CO;2, [https://doi.org/10.1175/1520-0469\(1995\)052<3620:TIOSFO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1995)052<3620:TIOSFO>2.0.CO;2), 1995.
- Korolev, A. V. and Mazin, I. P.: Supersaturation of water vapor in clouds, *Journal of the Atmospheric Sciences*, 60, 2957–2974, doi:10.1175/1520-0469(2003)060<2957:SOWVIC>2.0.CO;2, [https://doi.org/10.1175/1520-0469\(2003\)060<2957:SOWVIC>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<2957:SOWVIC>2.0.CO;2), 2003.
- Laaksonen, A., Vesala, T., Kulmala, M., Winkler, P. M., and Wagner, P. E.: Commentary on cloud modelling and the mass accommodation coefficient of water, *Atmospheric Chemistry and Physics*, 5, 461–464, doi:10.5194/acp-5-461-2005, <https://www.atmos-chem-phys.net/5/461/2005/>, 2005.
- Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R., Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K., and van Vuuren, D. P.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, *Atmospheric Chemistry and Physics*, 10, 7017–7039, doi:10.5194/acp-10-7017-2010, <https://www.atmos-chem-phys.net/10/7017/2010/>, 2010.
- Lebo, Z. J., Morrison, H., and Seinfeld, J. H.: Are simulated aerosol-induced effects on deep convective clouds strongly dependent on saturation adjustment?, *Atmospheric Chemistry and Physics*, 12, 9941–9964, doi:10.5194/acp-12-9941-2012, <https://www.atmos-chem-phys.net/12/9941/2012/>, 2012.

- Lock, A. P., Brown, A. R., Bush, M. R., Martin, G. M., and Smith, R. N. B.: A New Boundary Layer Mixing Scheme. Part I: Scheme Description and Single-Column Model Tests, *Monthly Weather Review*, 128, 3187–3199, doi:10.1175/1520-0493(2000)128<3187:ANBLMS>2.0.CO;2, [https://doi.org/10.1175/1520-0493\(2000\)128<3187:ANBLMS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<3187:ANBLMS>2.0.CO;2), 2000.
- Lohmann, U.: Possible aerosol effects on ice clouds via contact nucleation, *Journal of the Atmospheric Sciences*, 59, 647–656, doi:10.1175/1520-0469(2001)059<0647:PAEOIC>2.0.CO;2, [https://doi.org/10.1175/1520-0469\(2001\)059<0647:PAEOIC>2.0.CO;2](https://doi.org/10.1175/1520-0469(2001)059<0647:PAEOIC>2.0.CO;2), 2002.
- Lu, Z., Liu, X., Zhang, Z., Zhao, C., Meyer, K., Rajapakshe, C., Wu, C., Yang, Z., and Penner, J. E.: Biomass smoke from southern Africa can significantly enhance the brightness of stratocumulus over the southeastern Atlantic Ocean, *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.1713703115, <http://www.pnas.org/content/early/2018/02/27/1713703115>, 2018.
- 10 Malavelle, F. F., Haywood, J. M., Field, P. R., Hill, A. A., Abel, S. J., Lock, A. P., Shipway, B. J., and McBeath, K.: A method to represent subgrid-scale updraft velocity in kilometer-scale models: Implication for aerosol activation, *Journal of Geophysical Research: Atmospheres*, 119, 4149–4173, doi:10.1002/2013JD021218, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JD021218>, 2014.
- Mann, G. W., Carslaw, K. S., Spracklen, D. V., Ridley, D. A., Manktelow, P. T., Chipperfield, M. P., Pickering, S. J., and Johnson, C. E.: Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition-climate model, *Geoscientific Model Development*, 3, 519–551, doi:10.5194/gmd-3-519-2010, <http://www.geosci-model-dev.net/3/519/2010/gmd-3-519-2010.html>, 2010.
- 15 Martin, G. M., Johnson, D. W., and Spice, A.: The measurement and parameterization of effective radius of droplets in warm stratocumulus clouds, *Journal of the Atmospheric Sciences*, 51, 1823–1842, doi:10.1175/1520-0469(1994)051<1823:TMAPOE>2.0.CO;2, [https://doi.org/10.1175/1520-0469\(1994\)051<1823:TMAPOE>2.0.CO;2](https://doi.org/10.1175/1520-0469(1994)051<1823:TMAPOE>2.0.CO;2), 1994.
- 20 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, *Atmospheric Chemistry and Physics*, 18, 3119–3145, doi:10.5194/acp-18-3119-2018, <https://www.atmos-chem-phys.net/18/3119/2018/>, 2018.
- Ming, Y., Ramaswamy, V., Donner, L. J., Phillips, V. T. J., Klein, S. A., Ginoux, P. A., and Horowitz, L. W.: Modeling the Interactions between Aerosols and Liquid Water Clouds with a Self-Consistent Cloud Scheme in a General Circulation Model, *Journal of the Atmospheric Sciences*, 64, 1189–1209, doi:10.1175/JAS3874.1, <https://doi.org/10.1175/JAS3874.1>, 2007.
- 25 Morrison, H. and Gettelman, A.: A new two-moment bulk stratiform cloud microphysics scheme in the Community Atmosphere Model, Version 3 (CAM3). Part I: description and numerical tests, *Journal of Climate*, 21, 3642–3659, doi:10.1175/2008JCLI2105.1, <https://doi.org/10.1175/2008JCLI2105.1>, 2008.
- Morrison, H. and Grabowski, W. W.: Modeling supersaturation and subgrid-scale mixing with two-moment bulk warm microphysics, *Journal of the Atmospheric Sciences*, 65, 792–812, doi:10.1175/2007JAS2374.1, <https://doi.org/10.1175/2007JAS2374.1>, 2008.
- 30 Morrison, H., Curry, J. A., and Khvorostyanov, V. I.: A New Double-Moment Microphysics Parameterization for Application in Cloud and Climate Models. Part I: Description, *Journal of the Atmospheric Sciences*, 62, 1665–1677, doi:10.1175/JAS3446.1, <https://doi.org/10.1175/JAS3446.1>, 2005.
- Morrison, H., Thompson, G., and Tatarskii, V.: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: comparison of one- and two-moment schemes, *Monthly Weather Review*, 137, 991–1007, doi:10.1175/2008MWR2556.1, <https://doi.org/10.1175/2008MWR2556.1>, 2009.
- 35 Mulcahy, J. P., Jones, C., Sellar, A., Johnson, B., Boutle, I. A., Jones, A., Andrews, T., Rumbold, S. T., Mollard, J., Bellouin, N., Johnson, C. E., Williams, K. D., Grosvenor, D. P., and McCoy, D. T.: Improved aerosol processes and effective radiative forcing in HadGEM3 and

- UKESM1, *Journal of Advances in Modeling Earth Systems*, 10, 2786–2805, doi:10.1029/2018MS001464, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001464>, 2018.
- Mulcahy, J. P., Johnson, C., Jones, C. G., Povey, A. C., Scott, C. E., Sellar, A., Turnock, S. T., Woodhouse, M. T., Abraham, N. L., Andrews, M. B., Bellouin, N., Browse, J., Carslaw, K. S., Dalvi, M., Folberth, G. A., Glover, M., Grosvenor, D., Hardacre, C., Hill, R., Johnson, B., Jones, A., Kipling, Z., Mann, G., Mollard, J., O'Connor, F. M., Palmieri, J., Reddington, C., Rumbold, S. T., Richardson, M., Schutgens, N. A. J., Stier, P., Stringer, M., Tang, Y., Walton, J., Woodward, S., and Yool, A.: Description and evaluation of aerosol in UKESM1 and HadGEM3-GC3.1 CMIP6 historical simulations, *Geoscientific Model Development Discussions*, 2020, 1–59, doi:10.5194/gmd-2019-357, <https://www.geosci-model-dev-discuss.net/gmd-2019-357/>, 2020.
- Nenes, A. and Seinfeld, J. H.: Parameterization of cloud droplet formation in global climate models, *Journal of Geophysical Research: Atmospheres*, 108, n/a–n/a, doi:10.1029/2002JD002911, <http://dx.doi.org/10.1029/2002JD002911>, 4415, 2003.
- Peers, F., Francis, P., Fox, C., Abel, S. J., Szpek, K., Cotterell, M. I., Davies, N. W., Langridge, J. M., Meyer, K. G., Platnick, S. E., and Haywood, J. M.: Observation of absorbing aerosols above clouds over the south-east Atlantic Ocean from the geostationary satellite SEVIRI – Part 1: Method description and sensitivity, *Atmospheric Chemistry and Physics*, 19, 9595–9611, doi:10.5194/acp-19-9595-2019, <https://www.atmos-chem-phys.net/19/9595/2019/>, 2019.
- Petersen, G. N. and Renfrew, I. A.: Aircraft-based observations of air–sea fluxes over Denmark Strait and the Irminger Sea during high wind speed conditions, *Quarterly Journal of the Royal Meteorological Society*, 135, 2030–2045, doi:10.1002/qj.355, <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.355>, 2009.
- Petters, M. D. and Kreidenweis, S. M.: A single parameter representation of hygroscopic growth and cloud condensation nucleus activity, *Atmospheric Chemistry and Physics*, 7, 1961–1971, doi:10.5194/acp-7-1961-2007, <https://www.atmos-chem-phys.net/7/1961/2007/>, 2007.
- Phillips, V. T. J., Donner, L. J., and Garner, S. T.: Nucleation processes in deep convection simulated by a cloud-system-resolving model with double-moment bulk microphysics, *Journal of the Atmospheric Sciences*, 64, 738–761, doi:10.1175/JAS3869.1, <https://doi.org/10.1175/JAS3869.1>, 2007.
- Pinsky, M. B. and Khain, A. P.: Effects of in-cloud nucleation and turbulence on droplet spectrum formation in cumulus clouds, *Quarterly Journal of the Royal Meteorological Society*, 128, 501–533, doi:10.1256/003590002321042072, <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/003590002321042072>, 2002.
- Planche, C., Mann, G. W., Carslaw, K. S., Dalvi, M., Marsham, J. H., and Field, P. R.: Spatial and temporal CCN variations in convection-permitting aerosol microphysics simulations in an idealised marine tropical domain, *Atmospheric Chemistry and Physics*, 17, 3371–3384, doi:10.5194/acp-17-3371-2017, <https://www.atmos-chem-phys.net/17/3371/2017/>, 2017.
- Platnick, S., Ackerman, S., King, M., et al.: MODIS Atmosphere L2 Cloud Product (06_L2), NASA MODIS Adaptive Processing System, Goddard Space Flight Center, USA, doi:http://dx.doi.org/10.5067/MODIS/MYD06_L2.006, 2015.
- Politovich, M. K. and Cooper, W. A.: Variability of the supersaturation in cumulus clouds, *Journal of the Atmospheric Sciences*, 45, 1651–1664, doi:10.1175/1520-0469(1988)045<1651:VOTSIC>2.0.CO;2, [https://doi.org/10.1175/1520-0469\(1988\)045<1651:VOTSIC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<1651:VOTSIC>2.0.CO;2), 1988.
- Possner, A., Zubler, E., Lohmann, U., and Schär, C.: The resolution dependence of cloud effects and ship-induced aerosol-cloud interactions in marine stratocumulus, *Journal of Geophysical Research: Atmospheres*, 121, 4810–4829, doi:10.1002/2015JD024685, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD024685>, 2016.
- Pruppacher, H. R. and Klett, J. D.: *Microphysics of Clouds and Precipitation*, Springer, 1997.

- Rosenberg, P. D., Dean, A. R., Williams, P. I., Dorsey, J. R., Minikin, A., Pickering, M. A., and Petzold, A.: Particle sizing calibration with refractive index correction for light scattering optical particle counters and impacts upon PCASP and CDP data collected during the Fennec campaign, *Atmospheric Measurement Techniques*, 5, 1147–1163, doi:10.5194/amt-5-1147-2012, <https://www.atmos-meas-tech.net/5/1147/2012/>, 2012.
- 5 Saleeby, S. M. and Cotton, W. R.: A large-droplet mode and prognostic number concentration of cloud droplets in the Colorado State University Regional Atmospheric Modeling System (RAMS). Part I: module descriptions and supercell test simulations, *Journal of Applied Meteorology*, 43, 182–195, doi:10.1175/1520-0450(2004)043<0182:ALMAPN>2.0.CO;2, [https://doi.org/10.1175/1520-0450\(2004\)043<0182:ALMAPN>2.0.CO;2](https://doi.org/10.1175/1520-0450(2004)043<0182:ALMAPN>2.0.CO;2), 2004.
- Segal, Y., Pinsky, M., Khain, A., and Erlick, C.: Thermodynamic factors influencing bimodal spectrum formation in cumulus clouds, *Atmospheric Research*, 66, 43 – 64, doi:[https://doi.org/10.1016/S0169-8095\(02\)00172-2](https://doi.org/10.1016/S0169-8095(02)00172-2), <http://www.sciencedirect.com/science/article/pii/S0169809502001722>, 2003.
- 10 Seiki, T., Kodama, C., Noda, A. T., and Satoh, M.: Improvement in Global Cloud-System-Resolving Simulations by Using a Double-Moment Bulk Cloud Microphysics Scheme, *Journal of Climate*, 28, 2405–2419, doi:10.1175/JCLI-D-14-00241.1, <https://doi.org/10.1175/JCLI-D-14-00241.1>, 2015.
- 15 Shinozuka, Y., Saide, P. E., Ferrada, G. A., Burton, S. P., Ferrare, R., Doherty, S. J., Gordon, H., Longo, K., Mallet, M., Feng, Y., Wang, Q., Cheng, Y., Dobracki, A., Freitag, S., Howell, S. G., LeBlanc, S., Flynn, C., Segal-Rosenhaimer, M., Pistone, K., Podolske, J. R., Stith, E. J., Bennett, J. R., Carmichael, G. R., da Silva, A., Govindaraju, R., Leung, R., Zhang, Y., Pfister, L., Ryoo, J.-M., Redemann, J., Wood, R., and Zuidema, P.: Modeling the smoky troposphere of the southeast Atlantic: a comparison to ORACLES airborne observations from September of 2016, *Atmospheric Chemistry and Physics Discussions*, 2019, 1–77, doi:10.5194/acp-2019-678, <https://www.atmos-chem-phys-discuss.net/acp-2019-678/>, 2019.
- 20 Shipway, B. J. and Hill, A. A.: Diagnosis of systematic differences between multiple parametrizations of warm rain microphysics using a kinematic framework, *Quarterly Journal of the Royal Meteorological Society*, 138, 2196–2211, doi:10.1002/qj.1913, <http://dx.doi.org/10.1002/qj.1913>, 2012.
- Smith, R. N. B.: A scheme for predicting layer clouds and their water content in a general circulation model, *Quarterly Journal of the Royal Meteorological Society*, 116, 435–460, doi:10.1002/qj.49711649210, <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49711649210>, 1990.
- 25 Squires, P.: The Growth of Cloud Drops by Condensation. I. General Characteristics, *Australian Journal of Scientific Research A Physical Sciences*, 5, 59, doi:10.1071/PH520059, 1952.
- Stevens, B., Feingold, G., Cotton, W. R., and Walko, R. L.: Elements of the Microphysical Structure of Numerically Simulated Nonprecipitating Stratocumulus, doi:10.1175/1520-0469(1996)053<0980:EOTMSO>2.0.CO;2, [https://doi.org/10.1175/1520-0469\(1996\)053<0980:EOTMSO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1996)053<0980:EOTMSO>2.0.CO;2), 1996.
- 30 Stevens, B., Bony, S., Brogniez, H., Hentgen, L., Hohenegger, C., Kiemle, C., L’Ecuyer, T. S., Naumann, A. K., Schulz, H., Siebesma, P. A., Vial, J., Winker, D. M., and Zuidema, P.: Sugar, gravel, fish and flowers: Mesoscale cloud patterns in the trade winds, *Quarterly Journal of the Royal Meteorological Society*, n/a, doi:10.1002/qj.3662, <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3662>, 2019.
- 35 Stevens, R. G., Loewe, K., Dearden, C., Dimitrellos, A., Possner, A., Eirund, G. K., Raatikainen, T., Hill, A. A., Shipway, B. J., Wilkinson, J., Romakkaniemi, S., Tonttila, J., Laaksonen, A., Korhonen, H., Connolly, P., Lohmann, U., Hoose, C., Ekman, A. M. L., Carslaw, K. S., and Field, P. R.: A model intercomparison of CCN-limited tenuous clouds in the high Arctic, *Atmospheric Chemistry and Physics*, 18, 11 041–11 071, doi:10.5194/acp-18-11041-2018, <https://www.atmos-chem-phys.net/18/11041/2018/>, 2018.

- Thompson, G.: Advances in a microphysics parameterization to predict supercooled liquid water and application to aircraft icing, PhD thesis, 2016.
- Twomey, S.: The nuclei of natural cloud formation part II: The supersaturation in natural clouds and the variation of cloud droplet concentration, *Geofisica pura e applicata*, 43, 243–249, doi:10.1007/BF01993560, <https://doi.org/10.1007/BF01993560>, 1959.
- 5 Vance, A. K., Abel, S. J., Cotton, R. J., and Woolley, A. M.: Performance of WVSS-II hygrometers on the FAAM research aircraft, *Atmospheric Measurement Techniques*, 8, 1617–1625, doi:10.5194/amt-8-1617-2015, <https://www.atmos-meas-tech.net/8/1617/2015/>, 2015.
- Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Carslaw, K., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Jones, C., Mann, G., Milton, S.,
- 10 Rumbold, H., Sellar, A., Ujiie, M., Whitall, M., Williams, K., and Zerroukat, M.: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, *Geoscientific Model Development*, 12, 1909–1963, doi:10.5194/gmd-12-1909-2019, <https://www.geosci-model-dev.net/12/1909/2019/>, 2019.
- Wang, H., Easter, R. C., Rasch, P. J., Wang, M., Liu, X., Ghan, S. J., Qian, Y., Yoon, J.-H., Ma, P.-L., and Vinoj, V.: Sensitivity of remote aerosol distributions to representation of cloud–aerosol interactions in a global climate model, *Geoscientific Model Development*, 6,
- 15 765–782, doi:10.5194/gmd-6-765-2013, <https://www.geosci-model-dev.net/6/765/2013/>, 2013.
- West, R. E. L., Stier, P., Jones, A., Johnson, C. E., Mann, G. W., Bellouin, N., Partridge, D. G., and Kipling, Z.: The importance of vertical velocity variability for estimates of the indirect aerosol effects, *Atmospheric Chemistry and Physics*, 14, 6369–6393, doi:10.5194/acp-14-6369-2014, <https://www.atmos-chem-phys.net/14/6369/2014/>, 2014.
- Wilson, D. R. and Ballard, S. P.: A microphysically based precipitation scheme for the UK Meteorological Office Unified Model, *Quarterly Journal of the Royal Meteorological Society*, 125, 1607–1636, doi:10.1002/qj.49712555707, <http://dx.doi.org/10.1002/qj.49712555707>, 1999.
- 20 Wilson, D. R., Bushell, A. C., Kerr-Munslow, A. M., Price, J. D., and Morcrette, C. J.: PC2: A prognostic cloud fraction and condensation scheme. I: Scheme description, *Quarterly Journal of the Royal Meteorological Society*, 134, 2093–2107, doi:10.1002/qj.333, <http://dx.doi.org/10.1002/qj.333>, 2008.
- 25 Woodward, S.: Modeling the atmospheric life cycle and radiative impact of mineral dust in the Hadley Centre climate model, *Journal of Geophysical Research: Atmospheres*, 106, 18 155–18 166, doi:10.1029/2000JD900795, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000JD900795>, 2001.
- Wyngaard, J. C.: Toward Numerical Modeling in the “Terra Incognita”, *Journal of the Atmospheric Sciences*, 61, 1816–1826, doi:10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2, [https://doi.org/10.1175/1520-0469\(2004\)061<1816:TNMITT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2), 2004.
- 30 CO;2, 2004.
- Yang, Q., Easter, R. C., Campuzano-Jost, P., Jimenez, J. L., Fast, J. D., Ghan, S. J., Wang, H., Berg, L. K., Barth, M. C., Liu, Y., Shrivastava, M. B., Singh, B., Morrison, H., Fan, J., Ziegler, C. L., Bela, M., Apel, E., Diskin, G. S., Mikoviny, T., and Wisthaler, A.: Aerosol transport and wet scavenging in deep convective clouds: A case study and model evaluation using a multiple passive tracer analysis approach, *Journal of Geophysical Research: Atmospheres*, 120, 8448–8468, doi:10.1002/2015JD023647, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD023647>, 2015.
- 35 Zuidema, P., Redemann, J., Haywood, J., Wood, R., Piketh, S., Hipondoka, M., and Formenti, P.: Smoke and Clouds above the Southeast Atlantic: Upcoming Field Campaigns Probe Absorbing Aerosol’s Impact on Climate, *Bulletin of the American Meteorological Society*, 97, 1131–1135, doi:10.1175/BAMS-D-15-00082.1, <https://doi.org/10.1175/BAMS-D-15-00082.1>, 2016.

Supplementary materials for: Improving aerosol activation in the double-moment Unified Model with CLARIFY measurements

Hamish Gordon^{1,2}, Paul R. Field^{1,3}, Steven J. Abel³, Paul Barrett³, Keith Bower⁴, Ian Crawford⁴, Zhiqiang Cui¹, Daniel P. Grosvenor¹, Adrian A. Hill³, Jonathan Taylor⁴, Jonathan Wilkinson³, Huihui Wu⁴, and Ken S. Carslaw¹

- 5 1. School of Earth and Environment, University of Leeds, LS2 9JT, United Kingdom
2. Engineering Research Accelerator, Carnegie Mellon University, Forbes Avenue, Pittsburgh 15213, United States
3. Met Office, Fitzroy Road, Exeter, EX1 3PB, United Kingdom
4. Department of Earth and Environmental Sciences, University of Manchester, M13 9PL, United Kingdom

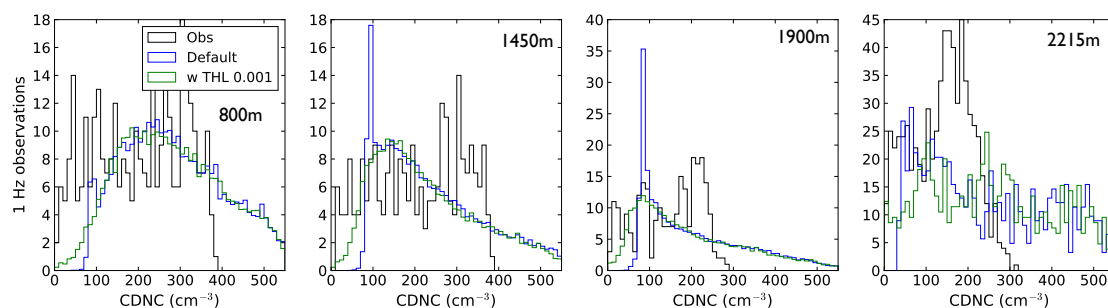


Figure S1. Change in droplet concentration (CDNC) when the ‘Default’ updraft PDF threshold ‘THL’ of 0.1m/s is reduced to 0.001m/s at 500 m resolution.

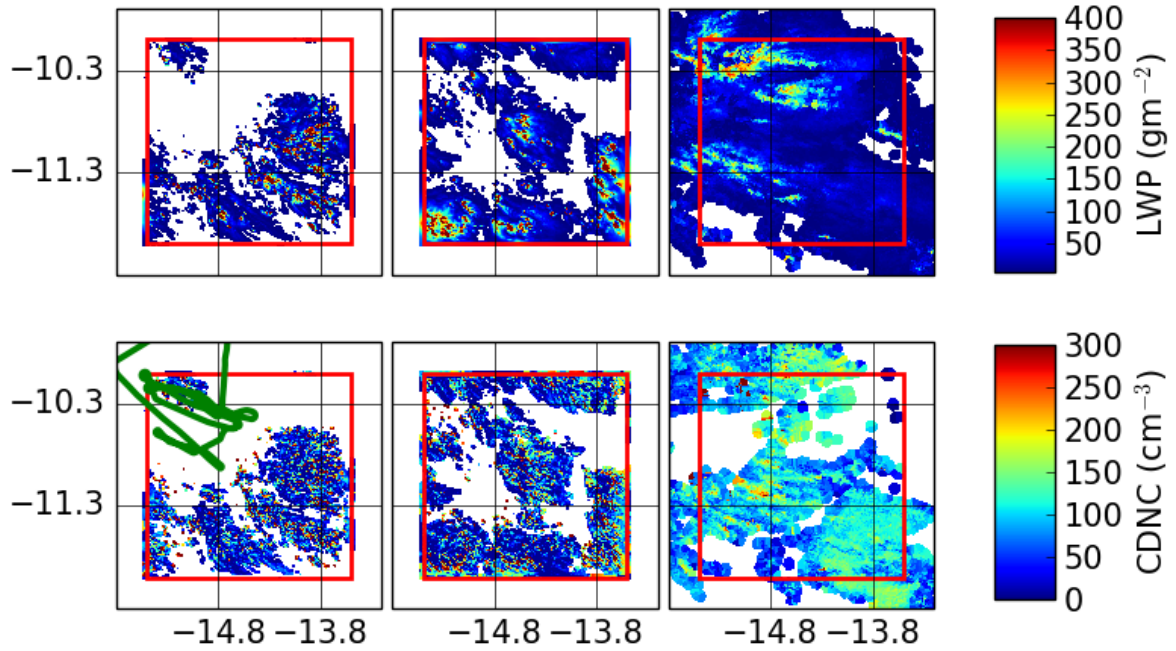


Figure S2. As Figure 3, except that the central subfigures now show a new 500 m simulation with the critical relative humidity in the sub-grid cloud scheme lowered by 0.1 (10%) to achieve more cloud cover.

Table S1. Mean in-cloud liquid water content in observations from the cloud droplet probe, compared to the ‘Default’ version of the model. In both model and observations, the threshold liquid water content is 0.01 g/kg to define a cloud.

Altitude	Obs (g/kg)	Default (g/kg)
800	0.34	0.07
1450	0.42	0.13
1900	0.58	0.18
2215	0.68	0.58
2550	0.70	0.98

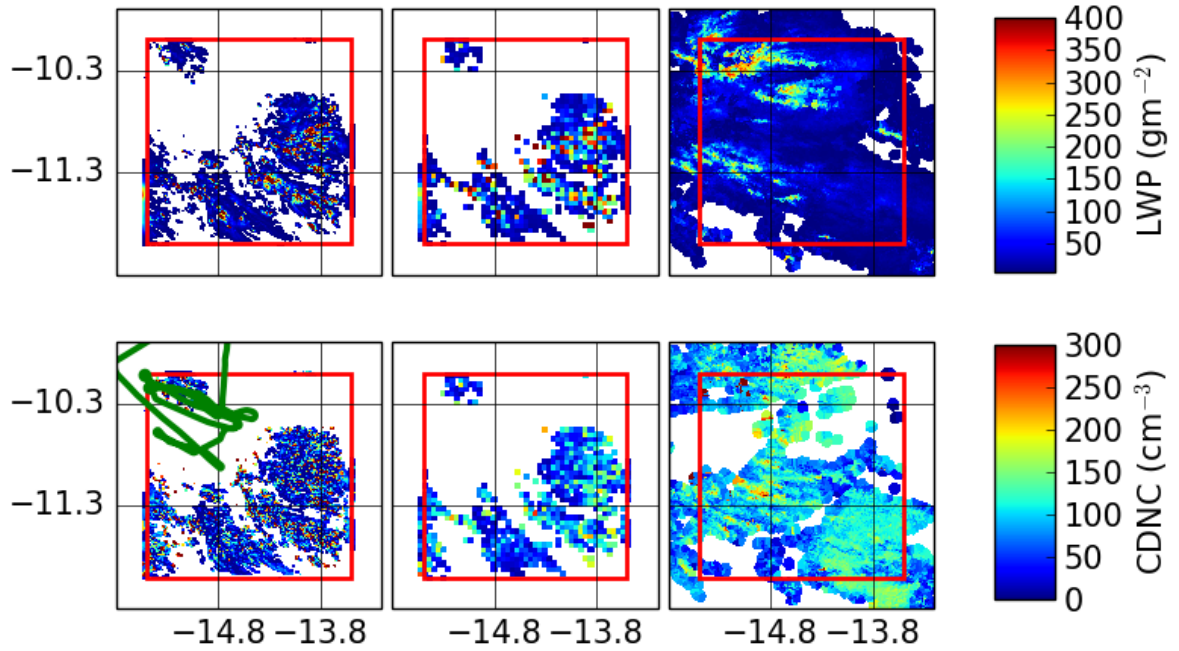


Figure S3. As Figure 3, except that the central subfigures show the 500 m simulation regridDED to 5 km resolution, for a potentially fairer comparison with MODIS

Table S2. Mean and standard deviation of in-cloud updraft speed in observations and in simulations, split by liquid water content (LWC) with $LWC > 0.15 \text{ gkg}^{-1}$ above the horizontal line and lower LWC below.

Altitude (m)	\bar{w} obs. ms^{-1}	\bar{w} model ms^{-1}	σ_w obs. ms^{-1}	σ_w 500 m ms^{-1}
800	0.69	0.98	1.32	0.90
1450	0.75	0.51	1.40	0.92
1900	0.72	0.17	1.47	0.73
2215	0.04	0.07	1.63	1.00
2550	-0.61	0.17	2.80	1.37
800	-0.006	0.12	0.88	0.34
1450	-0.12	-0.012	1.10	0.25
1900	-0.88	-0.04	0.92	0.22
2215	-0.42	-0.15	1.39	0.49
2550	-0.90	0.52	2.35	-

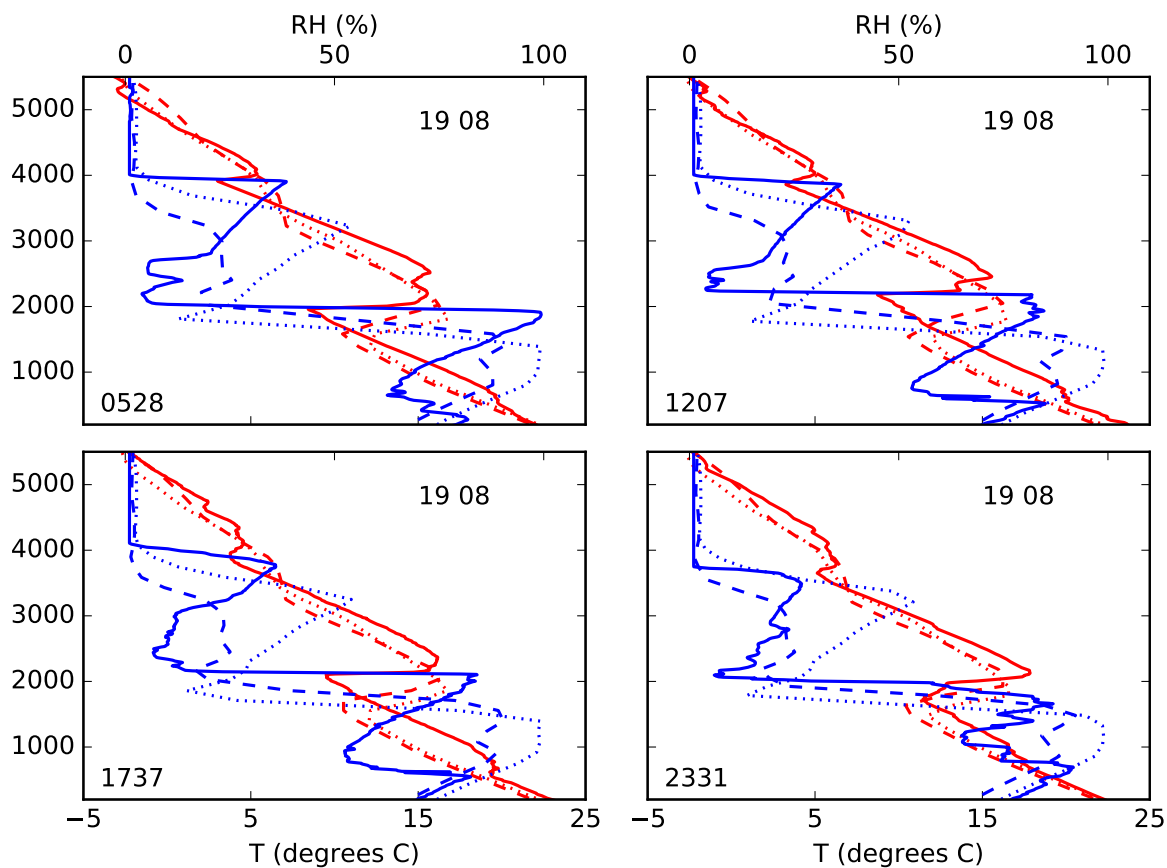


Figure S4. Driving model evaluation against Ascension Island radiosondes on 19 August. The global model is shown with the dashed line and the 7 km regional model with the dotted line. To reduce computational expense, the domain we simulate at 500 m resolution does not include Ascension Island.

Table S3. Change to simulated number mean droplet radius when the exponential size distribution ‘exp’ is replaced offline by the size distribution of Morrison and Gettelman (2008), labelled MG08, based on the same moments and when it is replaced by a gamma distribution with $\mu = 5$, labelled ‘ $\mu = 5$ ’.

Altitude	\bar{r} (exp) (μm)	\bar{r} (MG08) (μm)	\bar{r} ($\mu = 5$) (μm)
800	2.10	3.44	3.29
1450	2.40	3.89	3.76
1900	2.68	4.39	4.21
2215	3.54	5.60	5.56
2550	4.20	6.06	6.59

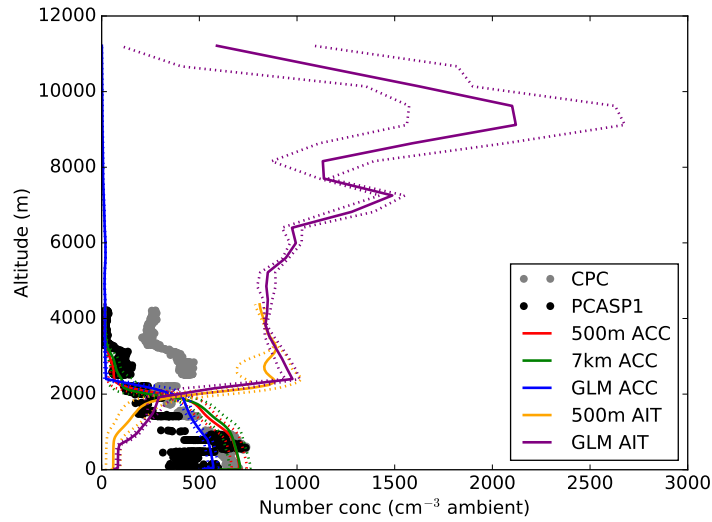


Figure S5. Simulated Aitken (AIT) and accumulation-mode (ACC) aerosol concentrations at ambient temperature and pressure, as in Figure 4, except including the global model (GLM) and extended to high altitudes. The comparison to the CLARIFY PCASP and CPC observations at low altitudes is retained.

Table S4. Values of the velocity scaling factor $\sqrt{f/R}$ for a horizontal resolution of 500 m, a factor $f = 4$ correction to the vertical velocity variance suggested by Malavelle et al. (2014), and two boundary layer heights representative of the CLARIFY case study.

Boundary layer type	Z_{mt}	$\sqrt{f/R} (Z = 1800\text{m})$	$\sqrt{f/R} (Z = 2200\text{m})$
II	$0.5Z$	4.27	3.85
III	$1.3Z$	2.81	2.64
IV	$0.5Z$	4.27	3.85
V (sc)	$0.5Z$	4.27	3.85
V (cbl)	Z	3.09	2.87
VI	Z	3.09	2.87
VII	$0.5Z$	4.27	3.85

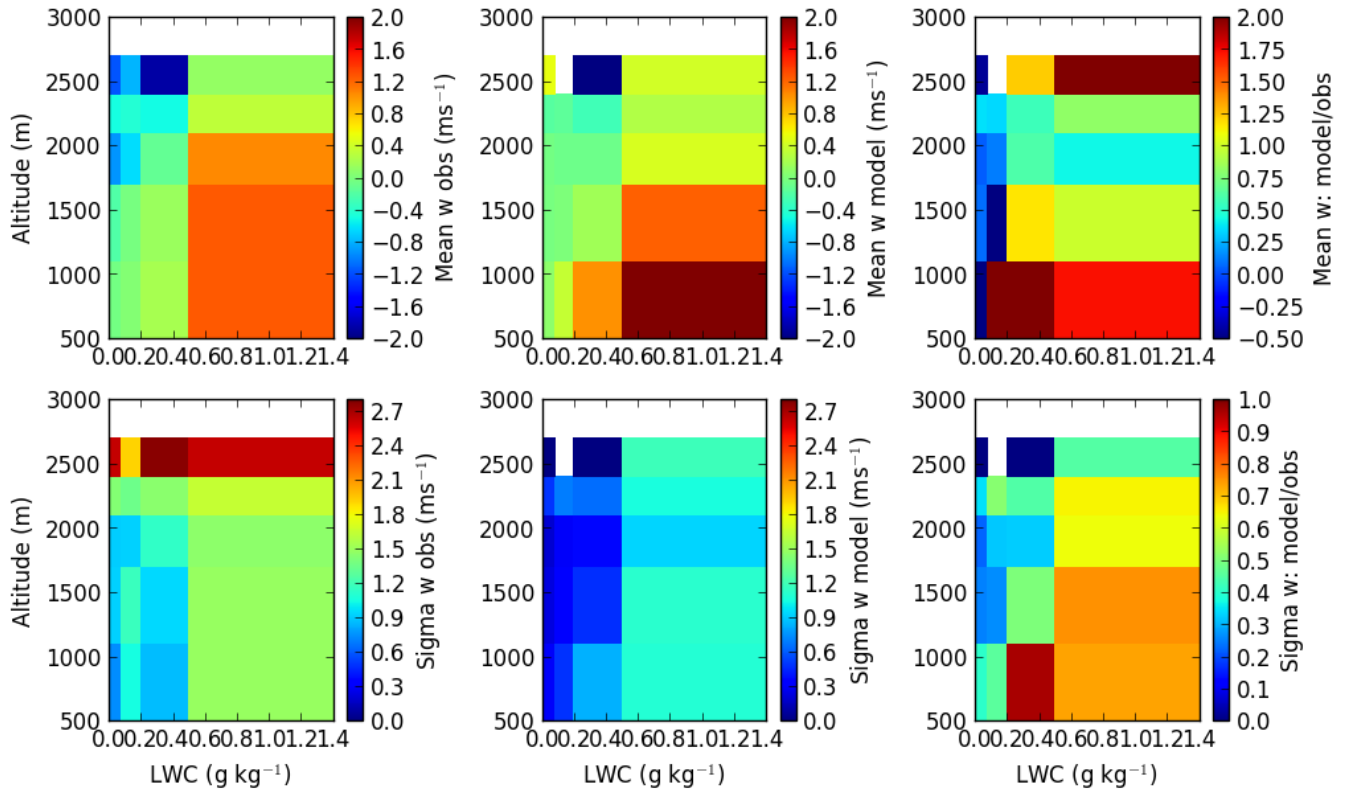


Figure S6. Mean and standard deviation of in-cloud updraft speed (vertical velocity w) in observations and in simulations, at the usual flight altitudes and in bins of liquid water content with boundaries at 0.08, 0.2, and 0.5 g kg^{-1}

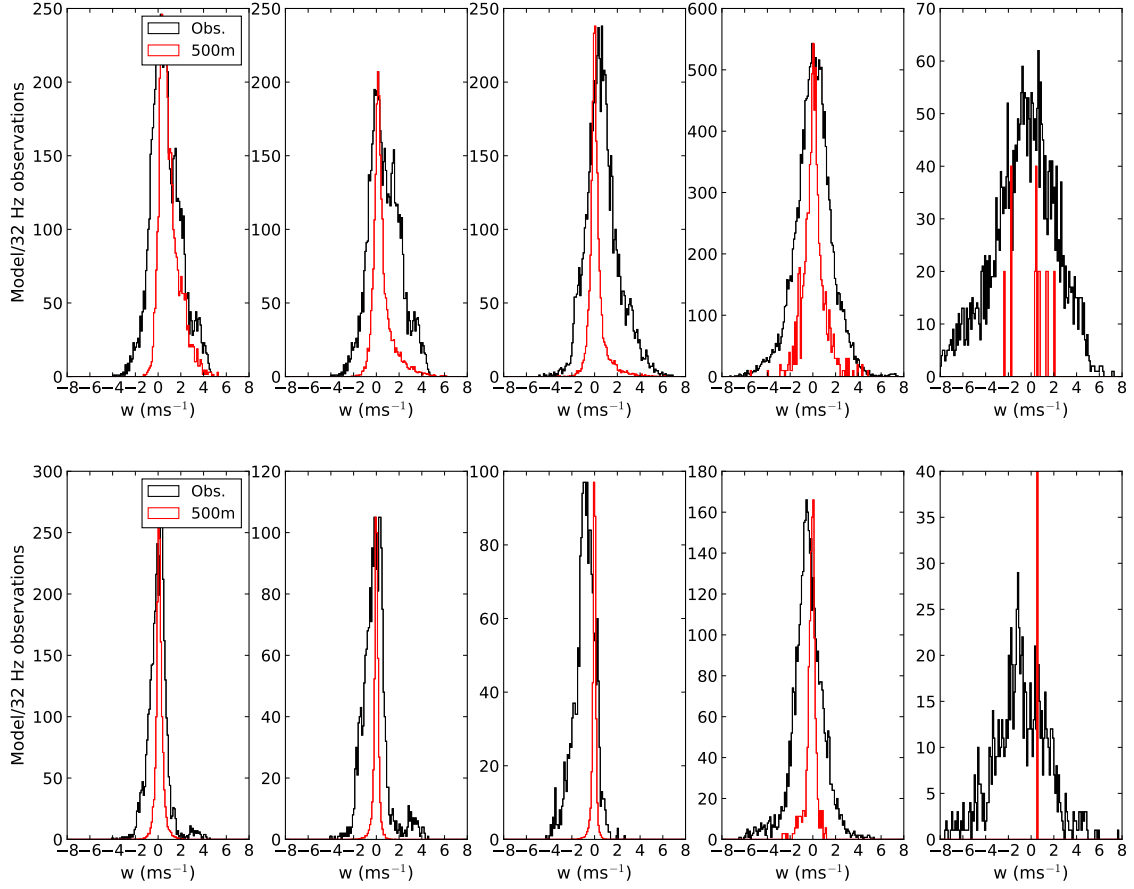


Figure S7. ~~Updraft speeds~~ **Spatial variability in vertical velocities w** in model and observations, top for in-cloud $\text{LWC} > 0.15 \text{ g kg}^{-1}$ and bottom for $\text{LWC} < 0.15 \text{ g kg}^{-1}$

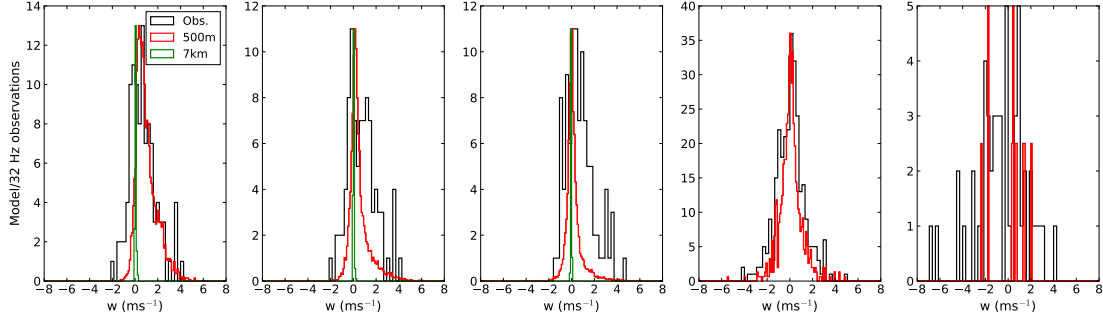


Figure S8. Spatial variability in simulated and observed vertical velocities w . Simulations compared to observed updrafts that have been smoothed out to average over intervals of 500m, for comparison with the model at this horizontal resolution. Only data for which for in-cloud liquid water content exceeds 0.15 g kg^{-1} are shown.

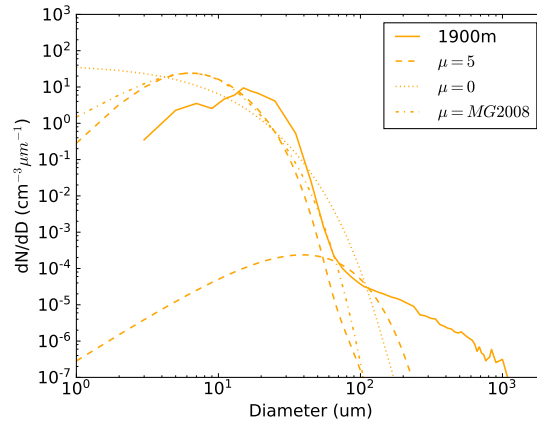


Figure S9. Particle size distributions at 1900 m altitude: observations as in Figure 9, the gamma size distribution with $\mu = 5$ used in this paper (dashed line), the exponential size distribution that was previously in the CASIM code (dotted, labelled $\mu = 0$), and the Morrison and Gettelman (2008) size distribution (dash-dot) that would be derived from the same moments.

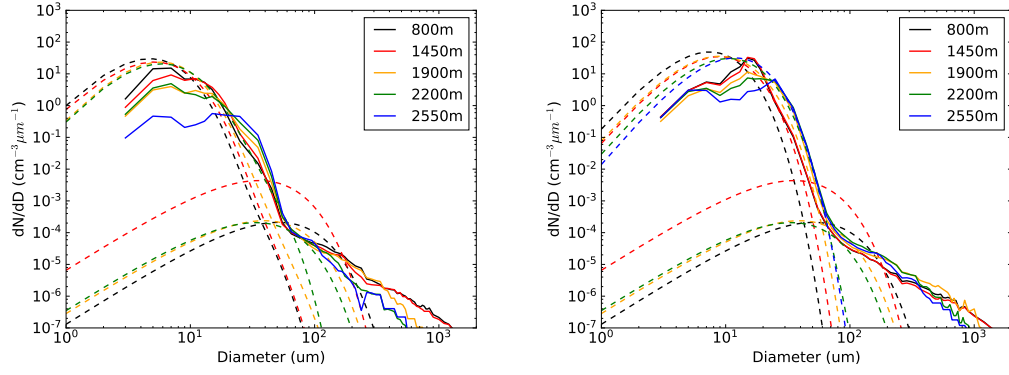


Figure S10. Particle size distributions as Figure 9 except with $\text{LWC} < 0.15 \text{ gkg}^{-1}$ (left) and $\text{LWC} > 0.15 \text{ gkg}^{-1}$ (right).

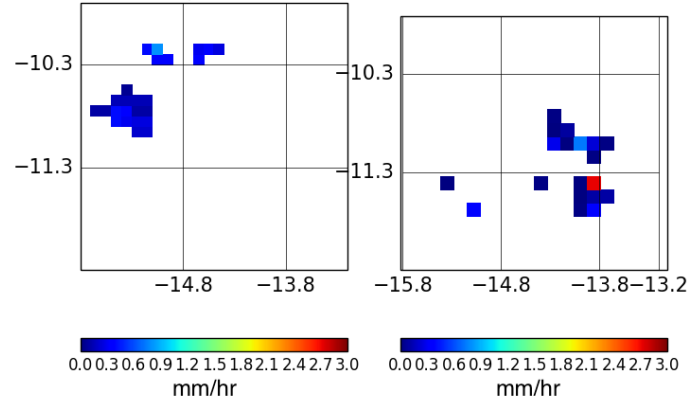


Figure S11. GPM observations of rain (left) compared to simulated rain (right) in the default version of the model.

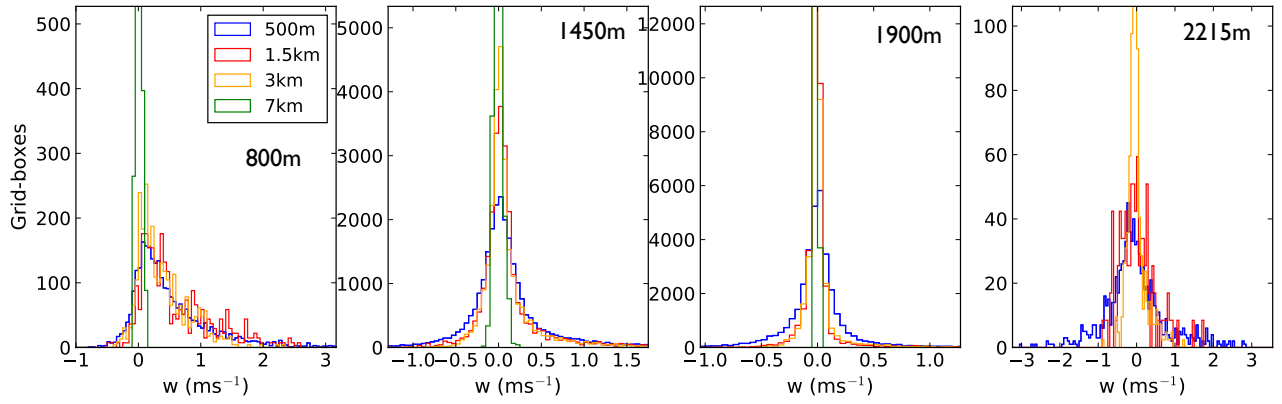


Figure S12. Spatial variability in vertical velocities w as PDFs Vertical velocities in simulations at 500 m, 1.5 km and 3 km resolution at the altitudes of the four lower straight and level flight legs, where the sample of in-cloud grid boxes is large enough. No threshold liquid water content is applied, and the x axis range is reduced compared to Figure 6 for clarity.

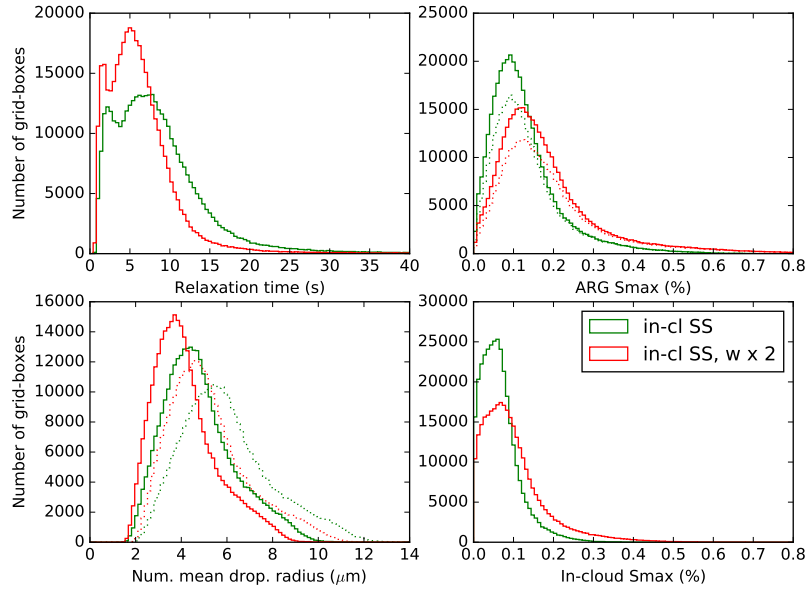


Figure S13. Histograms of grid-box mean relaxation time where relaxation time is used (in grid-boxes where in-cloud quasi-steady supersaturation is used to activate new droplets), ARG maximum supersaturation ‘ARG Smax’ (on all grid boxes where activation is taking place, and, with dotted lines, on grid boxes where relaxation time is calculated), number mean droplet radius where relaxation time is used, and in-cloud quasi-steady supersaturation where it is used. Dotted lines on the plot of number mean droplet radius show effective radius, which is 25% higher.

The strange ARG (DIAG) distribution of cloud droplet concentrations

At cloud base, the diagnostic ARG (DIAG) cloud droplet number concentration calculation samples a Gaussian PDF with mean equal to the gridbox mean updraft and width equal to the larger of 0.01 and $\sqrt{2/3TKE}$ (Boutle et al., 2018; Mulcahy et al., 2018), where TKE is the unresolved turbulent kinetic energy. A histogram showing the correlation of the droplet number concentration and the vertical velocity for the lowest 900 m in the model (roughly corresponding to cloud base altitudes), is shown in Supplementary Figure S14. The droplet concentrations and updraft speeds are calculated for 20 bins of the PDF and then the weighted mean is taken:

$$N_d = \frac{\sum N_i p(w_i)}{\sum p(w_i)} \quad (S1)$$

Here $p(w_i)$ is the probability a sub-grid updraft takes the value w_i . So far so good, but this procedure is coded such that the minimum updraft ever sampled from the PDF is zero and the maximum number ever sampled is four times the width, independently of the mean of the PDF. The maximum value of $\sqrt{2/3TKE}$ is around 1.2 at cloud base but the mean is around 0.15 ms^{-1} , and so, for example, if the mean updraft speed is less than -0.45 ms^{-1} or greater than 1.05 ms^{-1} (both frequent occurrences in the 500 m resolution simulations) and the width is 0.15, the integral of the PDF between 0 and 0.6 will be essentially zero, as this range is always over 3σ from the mean. The cloud droplet concentration that results will be poorly defined, most likely zero, because one is dividing by zero in the denominator (and multiplying by zero in the numerator) of the weighted mean. The droplet concentration is then reset to its hard-coded minimum of 5 cm^{-3} . The characteristic updraft speed, which is the updraft speed sampled from the PDF that would give the same cloud droplet concentration as the actual weighted mean concentration, will be zero if the real updraft speed is large and negative, or around four times the PDF width if the real updraft speed is large and positive.

We fixed this bug in the ARG (DIAG) activation scheme (from UM version 11.7) by redefining the minimum updraft in the PDF that is sampled as

$$w_{min} = \max(0, \bar{w} - 2\sigma_w) \quad (S2)$$

instead of zero, and the maximum as

$$w_{max} = \max(4\sigma_w, \bar{w} + 2\sigma_w) \quad (S3)$$

instead of $4\sigma_w$, to preserve existing behaviour as far as possible but still ensure that the PDF is still sampled in strong updrafts.

~~We conclude that the single-moment UKCA activation scheme should not be used at resolutions where the grid-box mean updraft speed may exceed four times its distribution's width, or that the minimum width should be artificially increased in these situations.~~

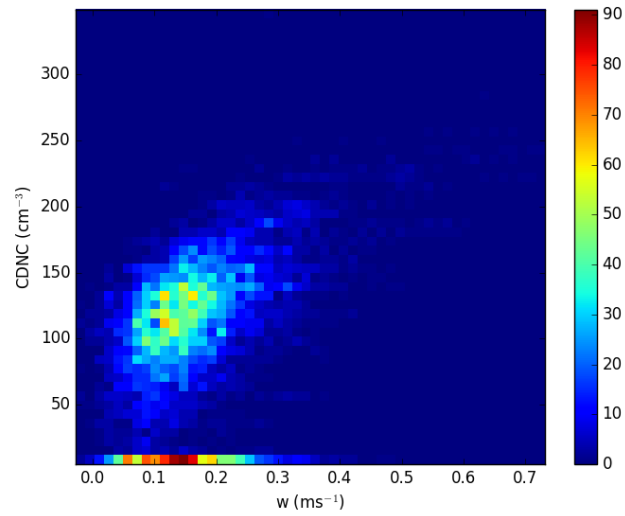


Figure S14. Correlation of updraft with cloud droplet number calculated with the UKCA activation scheme at low altitudes (below 900 m) for the 7 km-resolution simulation. At or close to cloud base, the updrafts can still be significant compared to the minimum updraft width, and consequently the PDF is not always fully sampled. All times are used (not just midday on 19 August) to increase the statistics available.