

## ***Interactive comment on “An aerosol activation metamodel of v1.2.0 of the pyrcel cloud parcel model: Development and offline assessment for use in an aerosol-climate model” by Daniel Rothenberg and Chien Wang***

### **Anonymous Referee #3**

Received and published: 21 December 2016

The authors show how their previously published methods for developing an emulator that approximates 1D parcel model output can be extended to include the parameter set used by a full aerosol model. Results of the derived emulator are compared against parcel model output for a large range in possible parameter values and for more realistic parameter value sets from model data. The emulator performance compares well to existing parameterizations with smaller biases in maximum supersaturation and activated particle number for the full range in parameter values, but compares less well when the parameter sets are taken from model data.

C1

My overall impression is that this new approach to parameterizing droplet activation shows promise for global modeling and could be of interest to those in the GCM cloud microphysics community. If indeed this is directed toward these interests (as suggested by the title/abstract) the manuscript could use more information on aspects of the scheme that are relevant for global modeling, notably computational expense, and also more of a focus on conditions and aerosol populations characteristics that are relevant to global modeling (i.e. CDNC > 10cm<sup>-3</sup>). I also have some questions about some of the choices made regarding initial parameter values and some of the simplifications made. My recommendation is for major revisions, including potentially to some aspects of the experiments and tests, in accordance with my comments and suggestions given below.

General comments:

1. One aspect of the emulators that is not mentioned in the text is the computational expense of this scheme relative to widely used parameterizations, such as the two used for comparison here, or the lookup table method. For use in a global model, the cost of computation would be considered along with the accuracy of the scheme. This is written about in Rothenberg and Wang (2016) and should be included here too, either by doing a quick assessment for this study, or by citing the previous work.
2. Since this is an extension of a previous paper but not a case where the code developed previously has been implemented in a global model, it is important to be clear and up front about what is new in the current study. On Pg 4 Lines 18-20 it is stated that the current study makes the Rothenberg and Wang (2016) emulator suitable for a global model, but I think it is not clear what that means at this point in the study. A bit of clarification would be helpful.
3. In general the paper seems unnecessarily long, especially given that it is an extension of previous work. I would recommend as a strategy for being more concise to find parts of the text that are common to this paper and Rothenberg and Wang (2016) and

C2

then just refer the reader to the latter paper for details. So this would include, for example, the description of the parcel model and the polynomial chaos expansion. Another area that could be cut down is the introduction and the descriptions of all the different parameterization methods. Since the goal of this paper is to prepare the emulator for a global model, it should suffice to list which GCMs use which activation schemes and the advantages/disadvantages of each.

4. I am somewhat uncomfortable with the exclusion of several of the aerosol modes and other parameters from the development of the emulator (Section 2.2) without having tested the sensitivity of the activation to these parameters, as is done for the remaining parameters in the following section anyhow. I would also note that no papers are cited to justify excluding these parameters. My recommendation is to either test all the parameters in the iterative calculations, or provide justification for excluding parameters from the literature. For example, it is reasonable to assume that nucleation mode particles are too small to contribute to CCN numbers, but it is not as obvious to me that a large population of nucleation mode particles could not have an impact on  $S_{max}$  by depleting available water vapor. In fact, there seems to be some evidence for this in the example given in Sect. 2.3? Maybe I am mis-reading this, but in any case please include a citation(s) to justify this exclusion or run the tests with this mode included. As a side note, it can be confusing that the three modes tested in the iterative calculations are “nucleation, accumulation and coarse” when the MARC modes are named “nucleation, Aitken and accumulation” and the nucleation mode has been excluded from the analysis. I think this requires some additional explanation.

5. Also with regard to the exclusion of parameters from the emulator, the use of a value of hygroscopicity equal to that of pure sulfate for the sulfate/BC mixture does not match up with my understanding of how the kappa parameter is usually applied. My understanding is that kappa is computed as a volume-weighted average of the component species regardless of the arrangement of the species within the particle. One could conceive of a situation in which this would not work, namely hygroscopic

C3

material within a surrounding shell of insoluble material, but I think this is considered rare. For the opposite, more common arrangement (hygroscopic with insoluble core) the volume-weighted kappa is appropriate, especially in this case since sulfate is highly soluble. So the water activity of a half:half BC:sulfate particle would be much less than the pure sulfate as is assumed here and I suspect the activation within the parcel model/emulator is sensitive to this for certain conditions. Please revisit this and see if you agree.

6. With regard to the choice of parameter space to cover - I understand the idea to cover a large range of parameter values for the sake of completeness. However, I find it curious that the lower bound for the geometric mean sizes is 1 nm for all modes after using the small size of the nucleation mode particles to justify excluding them altogether from the analysis. Perhaps since the values are sampled from a uniform distribution these small values do not come up very often and so the lower bound does not matter very much? And in this case does the sampling overemphasize higher values of, for example, ACC sulfate number concentration which ranges from roughly 0 to 10,000 /cm<sup>3</sup> and therefore could be expected to be > 5000/cm<sup>3</sup> about half the time? I think my major question here is: do you have a sense for how sensitive the emulator is to the parameter set used to train it? Could making reasonable changes to the range or sampling of parameters make a big difference in emulator performance? If so, it might be worth noting that many GCMs set a minimum CDNC, often 10/cm<sup>3</sup> or greater, and this might direct future emulator training efforts.

Minor comments:

Pg 2, Lines 14-18: This paragraph begins by describing ESMs and moves on to describing parcel models without a clear transition. I suggest starting a new paragraph beginning with Line 14.

Pg 2, Line 19: My understanding is that the barrier to including a parcel model within a GCM is solely related to computational expense. Otherwise there would be no need to

C4

parameterize or emulate the parcel model.

Pg. 4, Line 33: I suggest noting, as you do, that MARC can be used instead of MAM3 or MAM4 in CAM5 (I am not sure it replaced MAM3?), but remove the wording that MARC “extends” CESM. Is there a reference for the implementation of MARC in CAM5?

Pg 5, Lines 23-25, 28-30: I may be just confused about what MARC is – does it also include cloud microphysics?

Pg 5, Lines 32-33: I wonder what impact this assumption has on the results especially since  $V$  is lower than the set limit 50% of the time. Is there a reference that could be used to justify making this assumption or maybe some tests of your own?

Pg 6, Line 11: Please note and cite the emissions datasets that are used, particularly for dust and sea-salt.

Pg 6, Line 33: I think  $\mu$ -sub-g needs to be defined.

Pg 7, Lines 1-7: This paragraph could probably be deleted in the interest of space.

Pg 7, Lines 18-19: For practical purposes this is roughly true, but in theory organic carbon aerosol are not all insoluble or non-hygroscopic and even insoluble particles can activate droplets at high enough SS if they are large.

Pg 7, Lines 21-22: Is this assuming a fixed size range for the nucleation mode? Later on it is shown that the accumulation mode mean size covers a huge range and I wonder if it is conceivable that a variable nucleation mode mean size could vary up into sizes that are more relevant for CCN?

Pg 8, Lines 26-30: I have difficulty interpreting the reductions in  $S_{max}$  by including the nucleation or coarse modes as small or large effects. Is it possible to report the changes in activated droplet number instead? This might be easier to understand for other readers as well. I would add that the reductions in  $S_{max}$  are also due to the decrease in available water vapor in the case with more condensation.

C5

Pg 8, Lines 31-32: Could you provide more details about these 50,000 parameter sets? Were they sampled randomly for a range in altitudes?

Pg 9, Lines 11-13: I have to admit that I do not know the relative impacts of the reduction in parcel buoyancy vs. the reduction in water vapor for suppressing the  $S_{max}$ . Is this something you have looked into? Might be worth mentioning both mechanisms here in any case.

Pg 12, Lines 13-15: Is this speculation or is this something that was observed from the MARC output? Several major biomass burning areas (large sources of BC) are coastal or near-coastal such as in Australia and Indonesia and large indirect effects from biomass burning aerosols have been deduced from modeling off the west coast of Africa and South America. I would expect a mix of BC, sulfate and sea-salt in these places.

Pg. 12, Line 25: I would define “parcel model behavior” or use a more specific phrase such as “parcel model  $S_{max}$ ”

Pg 13, Line 19: This sentence was difficult to understand, perhaps just changing the word choice for “case” would help clarify.

Pg 14, Lines 8-13: I would argue that the bias for simulations with  $CDNC > 10/cm^3$  are of more interest to the global modeling community – maybe this discussion could be removed from the text.

Pg 15, Lines 32-33: I was surprised to read that the aerosol number concentrations ever went below the thresholds used to train the emulators which were  $1e-5/cm^3$  for all except ACC mode sulfate ( $1e-3/cm^3$ ). Is the model often predicting values this small? It is possible that I misread this.

Figure 5-8: Please include units for figures 5-8 (% and  $/cm^3$ )

Table 2: Are the hygroscopicity range values rounded off?  $\kappa$  should not be above the value for sulfate and could drop well below 0.1, although maybe sulfate usually

C6

dominates the internal mixture volume.

---

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-228, 2016.