Response to Reviewer 1

This work presents an update to the Abdul-Razzak and Ghan (ARG) parametrization for aerosol activation. ARG is a well-established and widely used parametrization in the regional and global modelling community, so this is a very useful development with a large application potential. The paper is clearly written and the method is documented in detail, facilitating the implementation by other modelling groups. The improved parameterization is then tested and evaluated in a state-of-the-art global aerosol-climate model, demonstrating a generally improved model performance for CDNC against satellite data. The paper is definitely worth of publication in GMD. Just a few questions and remarks, which the authors may consider before publication:

We sincerely thank the reviewer for their thoughtful and constructive comments on the manuscript. We have carefully considered each comment and have addressed them in detail below, making the necessary revisions and providing additional clarifications where appropriate. Reviewer comments are in **Blue**, our responses are in **Black**, and manuscript updates are in **Red**.

1. The global model simulations performed for this study cover only 1 year (2014), so it is not possible to quantify the variability of the results. Could you briefly comment on that, also based on your previous experience with the model?

We thank the reviewer for raising this concern. We chose 2014 because this is the endpoint year for the CMIP6 project, so it makes sense to define it as 'present day' and compare the results with existing literature (included in the revised draft, see reply to the minor comments). Our model setup is very similar to the UK Earth System Model version 1.1 (UKESM1.1) discussed in Mulcahy et al. (2023). In their work, similar biases in cloud top daytime droplet number concentrations were found when compared with the Grosvenor et al. (2018) dataset for 2003 to 2015. Hence, we are confident that simulating a longer period would result in changes similar to those shown in our manuscript. However, we acknowledge that using different models (Smith et al., 2020) or versions or simulation years might result in slightly different values compared to what we obtain here.

In the revised manuscript, on line L147 we add: "Climate models that participated in the Coupled Model Intercomparison Project Phase 6 (CMIP6) also reported radiative forcing due to aerosol-cloud interactions over a period ending in the year 2014 (Forster et al., 2021)."

In line 266 we add: "The magnitude of these biases and associated cloud radiative effects might also depend on choice of model (Smith et al., 2020), or version, or the simulation year."

2. Sect. 3.2: it would be good to complement the evaluation with the RMSE in addition to the NMB, as the latter is affected by the compensation of over/underestimation.

We thank the reviewer for this great suggestion. We agree that the Normalized Mean Bias (NMB) metric can be influenced by the compensation of biases. Like RMSE, it is also asymmetric in the sense that fractional overestimation is penalized more than fractional underestimation. Therefore, to address this concern, we have calculated RMSE and also Normalized Mean Absolute Error Factor (NMAEF). We find that the RMSE becomes worse with the updated ARG (numbers below). NMAEF, which is unbiased for fractional uncertainties, decreases with the updated ARG, indicating better model performance.

In the revised manuscript, on line 240 we add:

"We calculated several metrics to quantify the model performance. The Normalized Mean Bias (NMB) is useful in that it describes whether the model overestimates or underestimates observations overall, but positive and negative biases tend to cancel in the metric. Therefore, Root Mean Square Error (RMSE) is also commonly used to summarize the magnitude of the differences between the model and observations. However, both metrics penalize fractional overestimates more than fractional underestimates: for example, a model that overestimates observations by a factor two has a NMB of +100%, while one that underestimates observations by a factor two has a NMB of −50%. The model that overestimates has a larger absolute error but the same fractional error compared to the model that underestimates. Hence, following Gustafson and Yu (2012), we also compute the Normalized Mean Absolute Error Factor (NMAEF), which penalizes fractional overestimates and underestimates equally. It also differs from RMSE in that RMSE depends quadratically on errors while NMAEF depends on them linearly. We quote a simplified version of the NMAEF valid for positive model values and observations; the full version is given in Equation 5 of Gustafson and Yu (2012)."

In line 262 we add:

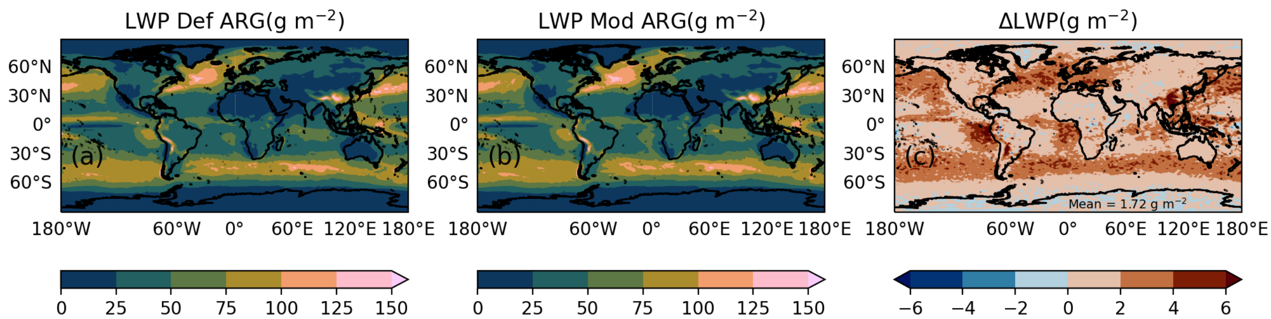
"The globally averaged RMSE increases from 64 cm^{-3} to 72 cm^{-3} when the ARG scheme is updated, while the global average NMAEF decreases from 0.48 to 0.39. Thus, the overall bias and fractional error decrease when we update the ARG parameterization, while the absolute error (measured by the RMSE) increases. Whether or not the performance of the model improves overall therefore depends on the priorities of the user. The magnitude of these biases and associated cloud radiative effects might also depend on choice of model (Smith et al., 2020), or version, or the simulation year."

Based on the comment of reviewer 2, we also design another sensitivity study to better understand the model performance. Please see below.

3. Would it be reasonable to also look at other variables, like liquid water path, cloud cover and precipitation? Do you expect changes in these variables with the improved ARG parametrization?

We thank the reviewer for this valuable suggestion. In Supplementary Figure S12, we show the changes in Liquid Water Path between the simulation with the updated ARG and the default ARG. The global average change is small yet systematic (3.4%). However, we find only minor changes in cloud cover and precipitation (not shown). In the revised manuscript, on line 292, we add the following:

"In Figure S12, we present the difference in liquid water path (LWP) between present-day simulations using the default ARG and the updated ARG. Our results indicate a small yet systematic increase of 1.72 g m^{-2} (3.4%) in LWP with the updated ARG, concentrated in regions where cloud cover is high. We find only minor differences in precipitation and cloud cover between these simulations (not shown)."



Minor suggestions Line 10: you could also provide the changes of CRE and aerosol forcing in relative terms.

We thank the reviewer for their suggestion. We have provided the percentage changes of CRE and aerosol forcing in the parentheses in the abstract.

Line 40: in the conditions used, can you be more specific?

We agree with the reviewer that this statement requires further clarification. The "conditions" mentioned in the paper refer to ammonium nitrate aerosols with a hygroscopicity of 0.7 (lognormally distributed with a number mode diameter of 100 nm) and a fixed updraft velocity of 0.5 m s^{-1}. We have also refined the sentence structure and ensured the accuracy of the reported values.

In the revised manuscript, line 40 is now modified to the following:

"They used ammonium nitrate aerosols of hygroscopicity 0.7 (lognormally distributed with number geometric mean diameter of 100 nm) and a fixed updraft velocity of 0.5 m s^{-1} to prepare that figure. Biases in the activation fraction in most schemes were small at widths of around 2.0 but increase below this. At a width of 1.6, the activation fraction in the ARG scheme is ~15% lower than the cloud parcel model, however at a width of 2.0 the fractions activated are within 10%."

Line 52: maybe add that targeting the accumulation mode is justified since this is the most relevant mode for activation?

We thank the reviewer for this suggestion. In the revised manuscript, on line 106, we add the following text:

"We vary only the mode width of the accumulation mode, as we find that it is more important to the determination of N_d than than the Aitken and coarse mode widths."

Line 63: citing the original Köhler theory would be appropriate in this context.

We thank the reviewer for this suggestion and have added the appropriate reference on line 66 of the revised manuscript.

Line 90: could you add a reference to the Latin hypercube sampling method?

We thank the reviewer for pointing this out. We have added a couple of references to the Latin hypercube sampling method on line 99 of the revised manuscript.

Table 2: in the list of parameters, Aitken should appear before accumulation.

We thank the reviewer for raising this point. The table is fixed following the reviewer's suggestion

Table 3: I would sort the columns in a way that the old and new values of a given parameter appear next to each other, otherwise it is quite hard to compare them.

We thank the reviewer for their suggestion. We have fixed the Table, such that new and old values of the parameters appear next to each other.

Figure 1 caption: please use the SI units for pressure (Pa) instead of atm throughout the paper.

We thank the reviewer for raising this point. We have changed the units throughout the paper.

Figs. S8 and S9: the captions refer to "annual averages" and "January 2014", but the plot show DJF and JJA, respectively (the main text too). Please Check.

We thank the reviewer for bringing this to our attention. We confirm that Figure 3 in the main text shows the annual averages. In the supplement, we have updated the captions for Figures S9 and S10 to clarify that Figure S9 represents the DJF average, while Figure S10 corresponds to the JJA average.

Line 253: you could mention the IPCC-AR6 ranges and the recent Bellouin et al. (2020, doi:10.1029/2019RG000660) results here, to support the statement that your RFs are reasonable.

We thank the reviewer for this suggestion. We have mentioned the estimated radiative forcings from the sixth assessment report and the Bellouin et al. (2020) to better support our results.

In the revised manuscript, on line 311 we add the following:

"Bellouin et al. (2020) estimates the present-day aerosol effective radiative forcing to range from -1.60 to -0.65 Wm^{-2}, with a 16%–84% uncertainty range. The Sixth Assessment Report (AR6) of the Working Group I (WGI) of the Intergovernmental Panel on Climate Change (IPCC) estimates the direct radiative forcing at -0.3 (-0.6 to 0.0) Wm^{-2}, and the cloud radiative forcing at -1.0 (-1.7 to -0.3) Wm^{-2} (Forster et al., 2021) over the industrial era."

Line 269: different accumulation mode - > different aerosol accumulation mode.

Fixed (see Line 334).

References

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K. S., Christensen, M., Daniau, A.-L., Dufresne, J.-L., Feingold, G., Fiedler, S., Forster, P., Gettelman, A., Haywood, J. M., Lohmann, U., Malavelle, F., Mauritsen, T., McCoy, D. T., Myhre, G., Mülmenstädt, J., Neubauer, D., Possner, A., Rugenstein, M., Sato, Y., Schulz, M., Schwartz, S. E., Sourdeval, O., Storelvmo, T., Toll, V., Winker, D., and Stevens, B. (2020). Bounding global aerosol radiative forcing of climate change. *Reviews of Geophysics*, 58(1):e2019RG000660. e2019RG000660 10.1029/2019RG000660.

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D. J., Mauritsen, T., Palmer, M. D., Watanabe, M., Wild, M., and Zhang, X. (2021). The earth's energy budget, climate feedbacks, and climate sensitivity. In Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., editors, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 923–1054. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., Alexandrov, M. D., Bennartz, R., Boers, R., Cairns, B., Chiu, J. C., Christensen, M., Deneke, H., Diamond, M., Feingold, G., Fridlind, A., Hünerbein, A., Knist, C., Kollias, P., Marshak, A., McCoy, D., Merk, D., Painemal, D., Rausch, J., Rosenfeld, D., Russchenberg, H., Seifert, P., Sinclair, K., Stier, P., van Diedenhoven, B., Wendisch, M., Werner, F., Wood, R., Zhang, Z., and Quaas, J. (2018). Remote sensing of droplet number concentration in warm clouds: A review of the current state of knowledge and perspectives. *Reviews of Geophysics*, 56(2):409–453.

Gustafson, W. I. and Yu, S. (2012). Generalized approach for using unbiased symmetric metrics with negative values: normalized mean bias factor and normalized mean absolute error factor. *Atmospheric Science Letters*, 13.

Mulcahy, J. P., Jones, C. G., Rumbold, S. T., Kuhlbrodt, T., Dittus, A. J., Blockley, E. W., Yool, A., Walton, J., Hardacre, C., Andrews, T., Bodas-Salcedo, A., Stringer, M., de Mora, L., Harris, P., Hill, R., Kelley, D., Robertson, E., and Tang, Y. (2023). Ukesm1.1: development and evaluation of an updated configuration of the uk earth system model. *Geoscientific Model Development*, 16(6):1569–1600.

Smith, C. J., Kramer, R. J., Myhre, G., Alterskjær, K., Collins, W., Sima, A., Boucher, O., Dufresne, J.-L., Nabat, P., Michou, M., Yukimoto, S., Cole, J., Paynter, D., Shiogama, H., O'Connor, F. M., Robertson, E., Wiltshire, A., Andrews, T., Hannay, C., Miller, R., Nazarenko, L., Kirkevåg, A., Olivié, D., Fiedler, S., Lewinschal, A., Mackallah, C., Dix, M., Pincus, R., and Forster, P. M. (2020). Effective radiative forcing and adjustments in cmip6 models. *Atmospheric Chemistry and Physics*, 20(16):9591–9618.