We thank the reviewers for their thoughtful comments on the manuscript and have addressed them in detail below. Reviewer comments are in *blue italics*, and our responses are in black normal text.

Response to Reviewer #2

Table 1: Please provide a long name for each parameter, e.g., temperature, pressure, etc.

Updated.

Figure 2: Is there a reason why the performance of the Twomey scheme is not shown? Also, please state in the figure (or at least in the figure label) that these are the results for the physically naïve emulators.

The main purpose of Figure 2 was to show the performance of the naïve emulators, with the ARG scheme only shown as a reference. The Twomey scheme skill is very bad, with orders of magnitude higher error and worse R^2, as stated on Line 215. As such, we chose not to include the Twomey performance, as it is would distract from the main purpose of the figure. We have updated lines 218-219 to more specifically address this:

"The Twomey scheme is not shown in Figure 2, as it performs relatively quite poorly (MSE = 0.29, R^2=0.03) and is not a particularly useful benchmark as compared to the relatively skillful ARG parameterization."

Section 4: it would be useful to show the machine learning statistics for both the training data and the test data to demonstrate that the models are not suffering from overfitting.

The reviewer raises a good point, that predictive skill on the training set is not necessarily indicative of skill on the test set predictions. We are more explicit in the manuscript on lines 207-210 that all evaluation is done on the test set to more accurately represent the lack of strong overfitting in our models:

"We evaluate the skill of these emulators in reproducing the activation fraction prediction within the test set, as described in Section 3. As machine learning predictive skill on the training set is not always an indicator of predictive skill on the test set, we discuss only test set performance here as a more strict evaluation criteria."

On line 260, I think it should say: "tend to perform. . ."

Updated, thank you.

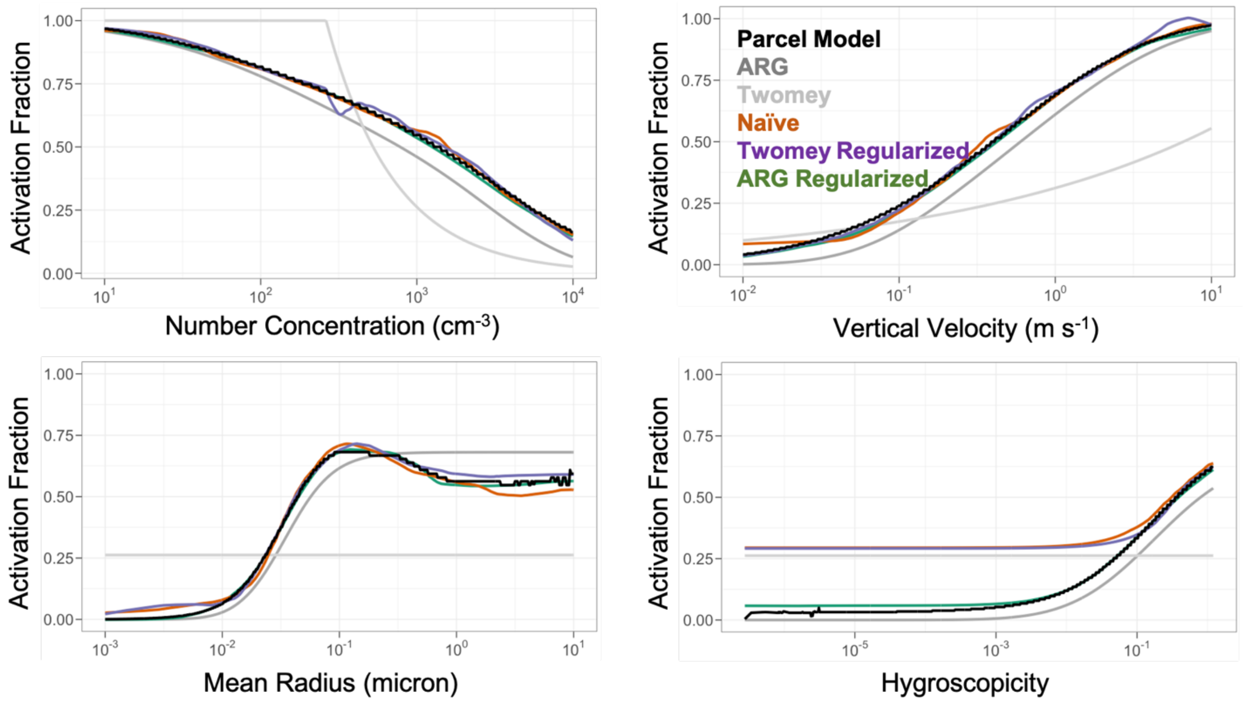
Section 4.4: The weak performance of the naïve and Twomey regularized emulator under low hygroscopicity regimes seems somewhat surprising. Doesn't this imply that the training data did not include enough training data capturing a low hygroscopicity environment? Based on Table 1, the hygroscopicity range of the training data spans 0 to 1.2, so the training data should capture this range at least to some extent. Or does this result suggest that the hygroscopicity value should be log-transformed to give higher weights to the lower bound? Given that low hygroscopicity values are not uncommon in the real atmosphere, this should be addressed a bit more convincingly in the revised version of the manuscript.

The weak performance of the naïve and Twomey regularized emulators indeed might be due to poor representation in the training data. We discuss this further on lines 321-325, however in the original text it was not apparent that this discussion was specifically toward low hygroscopicity values. We have updated the text:

“Though the specific issue of the poor performance of the Twomey regularized and naïve emulators in this low hygroscopicity range could potentially be somewhat resolved with additional model training data and other training optimization techniques (e.g., transfer learning on a subsample of the data, optimizing in log space, etc.), initial tests suggest that none of these issues completely solve the performance issues.”

Figure 7: is it possible to also show the performance of the regular parameterizations (without any machine learning)? This would help demonstrate the value of adding the machine learning correction to these parameterizations.

To maintain an easy-to-read figure, we have included only the machine learning-based parameterizations, as the sensitivities of the Twomey and ARG schemes have been described in detail in previous work. We have included a figure with the original parameterizations below.


Parcel Model
ARG
Twomey
Naïve
Twomey Regularized
ARG Regularized
Activation Fraction
10^{1}
10^{2}
10^{3}
10^{4}
Number Concentration (cm^{-3})
10^{-2}
10^{-1}
10^{0}
10^{1}
Vertical Velocity (m s^{-1})
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}
Mean Radius (micron)
10^{-5}
10^{-3}
10^{-1}
Hygroscopicity