In this study, the authors applied a Lagrangian framework to the E3SM to evaluate the aerosol-cloud interactions against multiple observations and reanalysis. Additionally, the framework uses direct measurements of CCN instead of AOD, and sensitivity tests to the parameters in the autoconversion and accretion scheme were performed. The biases in the E3SM simulation and possible causes were discussed. Given that the Lagrangian framework can provide some additional information for model evaluation compared to the common Eulerian perspective and the manuscript overall is well structured, I think this manuscript is suitable for the publication in ACP after addressing the following concerns.

General Comments:

In Section 4.1, specifically in Figure 5, the authors investigated the relationship between CCN and rain rate. For different E3SM sensitivity tests, there are no obvious differences among different present-day simulations. Since aerosol wet removal can be done both in and below clouds and model can diagnose these processes, I suggest that the authors show the differences of in-cloud and below-cloud wet scavenging among different simulations in detail. It might also be useful to understand why $\frac{\Delta CCN}{\Delta R}$ in the ARM observations is larger than that in the E3SM simulations.

In Figure S7, for a given rain rate, ARM features higher efficiency in washing out aerosols than the model, especially for rain rates smaller than 10^{-3} mm/hr. Why?

Specific Comments:

L187: Please define N_d in the equation.

L233: A brief description of the Froude number and the threshold indicating the flow was not blocked is expected.

Figure 2: ARM shows larger variations than E3SM. Why?

Figure S4: I did not find that means and standard deviations are provided.

Figure 3: Are the values in the parenthesis standard deviations?

L265-L273: In this study, the authors define light rain smaller than the accumulated average threshold of 0.05 mm/hr and heavy rain larger than this threshold. However, according to previous studies (e.g., Chen et al., 2021; Wang et al, 2021), light rain is often defined as a daily average smaller than 10 or 20 mm/day. Following these criteria, heavy rain defined here should also be grouped into the light rain category. Therefore, I suggest that the authors change the terminology of light and heavy rain to rain below and above the accumulated average threshold of 0.05 mm/hr, respectively throughout

the text.

Section 4.5.2: For the first approach, the difference terms represent differences in cloud properties between the clean and polluted state based on measured values of the CCN at the ARM site. However, for $\frac{\tau_a^{PD}-\tau_a^{PI}}{\Delta\tau_a}$ in Eq. (5), how to derive it in observations? Is the term equal to 1? If yes, please clarify this.

For the second approach, did the derived ERFaci significantly differ from the value using the method stated in Ghan (2013)? In model simulations, ERFaci can be calculated as $\Delta(F_{clean} - F_{clean,clear})$, where Δ denotes the difference between PD and PI, and each term can be directly output by the model.

References:

Chen, D., Dai, A., & Hall, A. (2021), The convective-to-total precipitation ratio and the "drizzling" bias in climate models. *Journal of Geophysical Research: Atmospheres*, 126, e2020JD034198. https://doi.org/10.1029/2020JD034198.

Wang, Y., G. J. Zhang, S. Xie, W. Lin, G. C. Craig, Q. Tang, and H.-Y. Ma (2021), Effects of coupling a stochastic convective parameterization with the Zhang–McFarlane scheme on precipitation simulation in the DOE E3SMv1. 0 atmosphere model, *Geoscientific Model Development*, *14*(3), 1575-1593.