Response to Reviewer #1

Reviewer comments in black and our response in blue text.

Summary

The authors present a new GPU-enabled aerosol module MAM4xx for the Earth system model EAMxx, derived from the four-mode module MAM4. They document their development steps: refactoring the Fortran code, porting the code to C++/Kokkos, and integrating the code into EAMxx. They evaluate the performance of EAMxx-MAM4xx with a global simulation at 12-kilometer resolution over 33 days. The paper is well written, well structured, and provides valuable insights into the development process and model performance. I have three general comments and a few specific comments.

We thank the reviewer for the comments on the paper. The comments are repeated here with our responses denoted by blue text.

General comments

In the introduction, the authors should mention recent developments in kilometer- scale Earth system modeling. Segura et al. (2025) presented Earth system simulations with ICON and IFS-FESOM at 10-kilometer resolution in the atmosphere over 30 years. Klocke et al. (2025) presented simulations with ICON over short periods at 1-kilometer resolution. Weiss et al. (2025) presented a simulation with ICON coupled to the aerosol module HAM-lite at 5-kilometer resolution.

- https://gmd.copernicus.org/articles/18/7735/2025
- https://dl.acm.org/doi/10.1145/3712285.3771789
- https://gmd.copernicus.org/articles/18/3877/2025

The introduction has been modified to include text related to the three references as follows: "More recently, Segura et al. (2025) describe 30-year Earth system simulations using the ICON and IFS-FESOM models with a 10-km grid spacing. ICON has also been used over shorter periods using a 1 km grid spacing (Klocke et al. 2025) and coupled with the HAM-lite aerosol model with a 5 km grid spacing (Weiss et al. 2025)."

In the methods, the authors could include a figure that shows the interface of EAMxx and MAM4xx and the time stepping. That would help the reader to understand the ordering and handling of processes and the computational costs.

Included a new figure (now Figure 3) to show a flow chart of how modules within EAMxx are called, including where sub-cycling is performed. The figure illustrates how sub-cycling is one (but not all) of the contributors to the larger computational burden of aerosol-cloud interactions.

In the evaluation, the authors could include a global overview of the simulation. Quantities like global burdens, lifetimes, and optical depths would help the reader to get a basic understanding of how well the global aerosol cycle is represented.

Included a new figure (Figure 16) at end of Section 4.4 to show global distribution of observed (MODIS) and simulated AOD during the simulation period. Also added a figure (Figure 17) summarizing the simulated global aerosol burdens. Some additional discussion has been added to describe the results in the figures and put them into context with other studies.

Specific comments

In the introduction, the authors could mention that GPU nodes are more energy efficient than CPU nodes.

Done. Added "GPU chips are more energy efficient than CPU chips when running a massive batch of calculations in parallel, such as atmospheric model simulations." after SCREAM is introduced.

In line 117-118, are there plans to account for ice-nucleating particles as well?

INP are already included for the SPA treatment but were not mentioned on line 117. The sentence has been modified to include INP. For MAM4xx, both CCN and INP are affected by prognostic aerosols.

In line 220-221, how are the loops over columns parallelized in the new code?

This text refers to removing message passing interface (MPI) related code used for parallelism on CPU clusters to what Kokkos expects. As described in Donahue et al. (2024), a hierarchical parallelism is used in EAMxx with three levels: columns, teams, and threads. We added a sentence to this paragraph pointing to this paper for more information for interested readers. Other details can be found in https:// https://docs.e3sm.org/scream/developer/kokkos_ekat/

In line 283-289, how large are the conservation errors in the new code?

This is an excellent question. MAM4xx was ported verbatim from and tested against the Fortran MAM4. Many parts of MAM4xx were unit tested for mass conservation during the port. In addition, the dynamic transport in EAMxx is very conservative. However, we have not yet performed mass conservation tests for MAM4xx within EAMxx. There exists software infrastructure to track and enforce conservation within EAMxx, but it is not enabled or tested yet by the E3SM team.

In section 3.4, how do the computational costs scale with resolution? If not to much work, the authors could perform two additional tests at resolutions of 6 and 3 kilometers and evaluate how the total and relative costs change.

Performing higher resolution simulations are computationally more expensive and we have nearly run out of our allocation. An E3SM project is extending the evaluation of MAM4xx to include aerosol-cloud interactions. They are also investigating the computational issues associated with aerosol-cloud interaction calculations shown in Fig. 2. We will pass along this suggestion to that team to also include resolution dependency issues.

In line 316-317, is the vertical mixing of aerosols still physically consistent with the vertical mixing of other species like water?

Yes. While the vertical mixing is handled outside of SHOC, which handles the vertical mixing of meteorological variables, the treatment is similar. This is the method employed in E3SM v2 (Fortran version) and was retained in the initial port to be consistent. Plans are underway to integrate aerosol mixing into SHOC.

In line 319-320, could you provide possible explanations for why the relative computational costs depend on the hardware?

GPU chips vary among different supercomputers, not only by vendor but by version.

In line 339-341, how do the emissions of sea salt and dust change with resolution?

The online emissions of sea salt and dust depend on simulated surface wind speed and soil moisture for dust. Changing the resolution will affect local meteorology and subsequently alter local emissions rates. The accumulated vertical flux over a region from a high-resolution simulation may or may not be the same as those from a coarse simulation.

In line 440-441, how flexible is the aerosol module? How much work would it be to include those species?

These species have already been included in a research Fortran version of MAM4 (Wu et al. 2022) using the MOSAIC aerosol microphysics treatment. It has not been included in official releases of E3SM due to its additional computational cost. It is technically feasible to modify MAM4xx to include those species (e.g., adding additional tracers is relatively straightforward) but would require effort in converting the code to C++, integrating it with MAM4xx, and testing the results.

Wu, M., Wang, H., Easter, R.C., Lu, Z., Liu, X., Singh, B., Ma, P.-L., Tang, Q., Zaveri, R.A., Ke, Z., Zhang, R., Emmons, L.K., Tilmes, S., Dibb, J.E., Zheng X., Xie, S., and Lueng, R.L.: Development and evaluation of E3SM-MOSAIC: Spatial distributions and radiative effects of nitrate aerosol. *J. Adv. Model Earth Sys*., 14, e2022MS003157. DOI:10.1029/2022MS003157.