**Reviewer #2**

**General comments:**

This study successfully implements interactive chemistry within E3SM, which consists of the chemUCI mechanism in the troposphere and the Linoz v3 module in the stratosphere. The new module performs well in standard atmospheric trace gas metrics while maintaining or improving the existing climate simulation. It holds significant potential to enhance E3SM's capability in studying chemistry-climate feedbacks, representing a notable advancement in Earth system modeling.

However, as a model development study, this manuscript lacks detailed descriptions of the introduced module. Additionally, some explanations can be misleading, and certain analyses of the simulation results are not sufficiently convincing. I suggest the manuscript be evaluated again after substantial improvements are made.

We appreciate the reviewer’s insightful and constructive comments, which will improve the overall quality of our revised manuscript. In response to the suggestions, we have made comprehensive revisions to the manuscript, including providing detailed descriptions, clarifying interpretations, and refining analyses. Our responses are outlined below in blue font.

**Major Points:**

* The author mentions in the title and abstract that E3SM-chem is based on E3SM version 3, but in the Introduction (Line 115) and Section 2.2, it is stated that the E3SM experiments use version 2. Can the authors confirm that, apart from the interactive atmospheric chemistry added in this study, all other features are consistent between E3SM-chem (version 3) and E3SM (version 2)? (Line 116 notes that version 3 has many new features compared to version 2.) This issue could significantly impact results presented in this manuscript, and the authors need to provide further clarification.

While the interactive chemistry capability has been developed for E3SM version 3, the base model utilized in the present study is version 2. This reflects the typical progression of model development, where new features intended for the next version are integrated into the framework of the preceding version. Consequently, all features other than the interactive chemistry remain consistent between E3SM-chem (version 2 with the new interactive chemistry) and the standard E3SM (version 2), which allows us to isolate the impact of the new interactive chemistry scheme on the simulated climate.

To avoid confusion, we deleted this sentence on Lines 116-117 “because the E3SMv3 atmospheric model has many other new features (Xie et al., 2025) and has not yet been officially released.”

* The interactive feature should be a key innovation of this study, yet Section 2 lacks a detailed description of its mechanisms. For instance, how meteorological fields influence chemistry and how chemistry, in turn, affects meteorological fields remain unclear. The authors briefly illustrate the impact of temperature (T) on ozone (O3) in Lines 135–138 and mention that E3SM-chem calculates radiative forcing, which partially demonstrates the interactive nature. However, it is not specified whether E3SM version 2 performs the same or similar calculations. A more thorough discussion is recommended.

E3SM and E3SM-chem share identical links between chemical tracers and other components of the model. Specifically, tracers are advected by winds, greenhouse gas concentrations are incorporated into radiation calculations, and chemical reaction rates are influenced by temperature and pressure. These configurations are standard in chemistry-climate simulations and are therefore not a focal point of our manuscript. The primary distinction between the two models lies in their treatment of chemical mechanisms in the troposphere: E3SM-chem includes interactively calculated chemical species, whereas E3SM predominantly relies on prescribed tracer datasets.

To clarify this point, we have added the following sentence on Line 149: "It is worthwhile pointing out that E3SM and E3SM-chem share identical links between chemical tracers and other components, such as the dynamic core and radiation."

* Line 152-162：The authors mention that the ISLET tracer transport method was already implemented in E3SM version 2. However, they do not clarify whether the ISLET method used in this study differs from the original version. If no modifications were made, this should not be presented as part of the methodological contributions of the current work. Further clarification is needed.

While the ISLET tracer transport method was implemented in E3SM version 2, its computational efficiency is fully realized in this study due to the incorporation of interactive chemistry, which significantly increases the number of advective tracers required. Without ISLET, the computational cost of chemUCI in E3SM would be substantially higher (due to 6 to 8 times slower tracer transport), making it impractical to adopt interactive chemistry as the default configuration for upcoming E3SM production simulations. This high computational expense of tracer transport is a key reason why only a limited number of CMIP6 models have delivered simulations with interactive chemistry, limiting our ability to understand chemistry-climate interactions.

Integrating ISLET with interactive chemistry is new to this model and provides a practical solution for enabling more affordable chemistry-climate simulations. Therefore, we believe it is appropriate to describe the ISLET method with our present study.

* The authors mention that ISLET allows for longer time steps, but they should specify how many times longer these steps are compared to CFL-limited ones. Additionally, does using longer time steps introduce computational errors? While the authors state in Line 160 that ISLET maintains "equivalent accuracy," a more thorough evaluation is necessary when proposing a model development. Furthermore, the authors claim that ISLET is computationally efficient, with the combined chemistry calculations adding at most 20% to computational costs. However, they should specify the hardware conditions under which these tests were conducted. The lack of precise numerical values raises questions—if a range is reported, what causes this variability? Does the additional computational cost scale with the number of CPU cores used? A more detailed explanation is warranted.

The tracer transport time step (1800 s) is 6x longer than the dycore time step (300 s). The accuracy of the ISLET method was extensively evaluated by Bradley et al. (2022). The tracer transport computational cost scales well with CPU cores as shown in Figure 3 of Golaz et al. (2022). Furthermore, this scaling behavior doesn’t change with more tracers (see Figure 15 of Bradley, 2024).

We revised the sentence on Lines 157-159 to “This approach allows tracer transport in these simulations to operate with a significantly longer (6x; 1800 s) time step than the CFL-limited dynamics grid time step (300 s) while maintaining accuracy (Bradley et al., 2022). The ISLET tracer transport scales well with CPU cores as shown in Figure 3 of Golaz et al. (2022). Furthermore, it scales linearly with increasing tracer numbers as demonstrated in Bradley (2024).”

Thanks for the suggestion of adding hardware information for the simulation cost. We added these details of hardware in the following paragraph, “The computational performance tests were carried out on 30 nodes of the Chrysalis cluster at Argonne National Laboratory. Chrysalis consists of 512 compute nodes. Each node is equipped with two AMD Epyc 7532 “Rome” 2.4 GHz processors. Each processor has 32 cores and hence a total of 64 cores per node. Each node has 256 GB 16-channel DDR4 3200 MHz memory. The nodes are connected using Mellanox HDR200 InfiniBand and use the fat tree topology. On average, a one-year E3SM-chem simulation test was completed in 4,200 seconds (wallclock time), with 350 seconds spent on tracer advection and 45 seconds on chemistry solvers.”

Golaz, J.-C., Van Roekel, L. P., Zheng, X., Roberts, A. F., Wolfe, J. D., Lin, W., et al. (2022). The DOE E3SM Model version 2: Overview of the physical model and initial model evaluation. Journal of Advances in Modeling Earth Systems, 14, e2022MS003156. <https://doi.org/10.1029/2022MS003156>

Bradley, A. M. (2024). Stabilized bases for high-order, interpolation semi-Lagrangian, element-based tracer transport, Journal of Computational Physics, 508, 113034, https://doi.org/10.1016/j.jcp.2024.113034.

* Figure 6 only presents the zonal mean SCO from E3SM-Chem but the corresponding E3SM results for comparison are not included. Moreover, based on Figure 6d, it appears that E3SM's results seems exhibit better overall agreement with observations than those from E3SM-Chem. This weakens the comparative analysis, and including the E3SM reference would strengthen the evaluation of model performance.

E3SM-chem has a different chemistry module in the stratosphere as discussed in the paper. We have now added the monthly zonal mean SCO from the previous E3SM and dropped the difference plot (c) since the panel (d) shows the annual mean. We also updated to use the latest observational data version. The monthly patterns for E3SM and E3SM-chem are alike except for the mean tropical biases. Both models are biased low at high northern latitudes.

* Figure 10-13: The authors should provide a more detailed discussion of the results presented in these figures rather than summarizing them briefly. Section 4 offers insufficient analysis of the global climate results, merely concluding that E3SM-Chem produces simulations similar to E3SM. While I agree with the authors’ statement in Line 356 that introducing new features does not necessarily require improved model performance, the significant enhancements in O3 and H2O simulations in E3SM-Chem should logically lead to some climate differences (at least regionally), such as in radiative forcing. To better show these differences, the authors could consider calculating the difference between panels (c) and (d) in Figures 10–13, which may reveal more pronounced discrepancies. Additionally, evaluating these differences would strengthen the completeness of this study.

This suggestion is interesting, but it would greatly expand the paper and not produce conclusive results. Looking for regional differences would reflect more about the internal variability than the forced response from interactive chemistry. To find the forced climate differences (due to tropospheric O3 and stratospheric H2O) would require a major Detection and Attribution Model Intercomparison Project (DAMIP) exercise and - while it is interesting as the reviewer notes - it is outside the scope. The figures as is show that, overall, the model climate discrepancies are not largely changed. That is the point here.

We added these discussions of Figures 10-13 on Line 372. “The global mean biases are very similar for TOA net radiation (E3SM-chem=0.38 W/m2 vs E3SM=-0.68 W/m2), 2-m air temperature over land (0.28 oC vs. 0.16 oC), and 850-hPa zonal wind (0.26 m/s vs. 0.24 m/s); and the same for precipitation (0.32 mm/day). Some small regional changes are observed, including reduced TOA net radiation over the tropical western Pacific, increased precipitation in northern South America, and variations in 850-hPa zonal wind over the central Pacific.”

* Line 339: The text states that E3SM-Chem exhibits a 30 DU reduction compared to E3SM in high-latitude Southern Hemisphere regions, yet Figure 6d does not reflect such a substantial discrepancy (the maximum difference appears to be no more than 10 DU). The authors should clarify this inconsistency.

The 30 DU reduction in E3SM-chem compared to E3SM, as shown in Figure 8a, refers specifically to the Southern Hemisphere (SH) polar region, which is implied in the SH minimum. In contrast, the results in Figure 6d are limited to latitudes up to 60°S. The difference arises because these two figures represent different geographical areas.

* Line 340: If, as stated in Line 339, the background SCO values are reduced by 30 DU in E3SM-Chem, the ~30 DU difference between E3SM and E3SM-Chem in Figure 8a suggests that the improved ozone hole simulation may primarily result from this background reduction rather than the proposed radiative cooling by water vapor. The authors should provide more robust evidence to substantiate their interpretation and clarify the relative contributions.

It is also our understanding that the primary driver of the E3SM-chem ozone hole improvement is the reduction in background stratospheric column ozone (SCO), with radiative cooling by water vapor playing a secondary role. Thanks for pointing out this potential confusion. We clarified this on Line 342 and it now reads “A secondary reason for this improvement may be linked to changes in the strength of the Antarctic circumpolar stratospheric vortex…”

**Minor Points:**

* It is recommended to include a table listing the differences between E3SM-chem and E3SM. This would help reader to compare and more clearly demonstrate the contributions of this study to model development.

We added Table SY in the revised manuscript to summarize the differences between E3SM-chem and E3SM.

* Line 135: The phrasing "O3 depends on O3 and column O3" is somewhat confusing. Isn’t column O3 also part of O3? Additionally, since the focus and innovation of this paper lie in model development, it is suggested to explicitly describe the calculation methods for O3 net production in relation to O3, T, CH4, NOy, and H2O. The calculation method for Linoz v2 should also be provided to allow readers to compare.

This information has been well published before; however, we have added some of this to the summary table suggested above. The troublesome description has been corrected to:

“The O3 chemical tendency (net production minus loss) in E3SM-chem depends on the local concentration of O3, CH4, NOy, N2O and H2O as well as the overhead column of O3 and T that determines the photolysis rates.”

Linoz calculates the net production of each tracer as a function of all tracers including itself as well as temperature (T), and photolysis (parameterized in terms of overhead ozone). Specifically for O3, Linoz v3 derives net production based on O3, T, column O3, CH4, NOy, and H2O.

The details of Linoz v3 are documented in Hsu & Prather (2010), while Linoz v2 is described in McLinden et al. (2000) and Hsu & Prather (2009). These references were cited in the revised manuscript for calculation details in Section 2.1: “The calculation details are provided in McLinden et al. (2000) and Hsu & Prather (2009) for Linoz v2, and in Hsu & Prather (2010) for Linoz v3.”

Hsu, J., & Prather, M. J. (2009). Stratospheric variability and tropospheric ozone. Journal of Geophysical Research: Atmospheres, 114(D6), D06102. <https://doi.org/10.1029/2008JD010942>

Hsu, J., & Prather, M. J. (2010). Global long-lived chemical modes excited in a 3-D chemistry transport model: Stratospheric N2O, NO, O3 and CH4 chemistry. Geophysical Research Letters, 37(7). <https://doi.org/10.1029/2009GL042243>

McLinden, C. A., S. C. Olsen, B. J. Hannegan, O. Wild, and M. J. Prather (2000), Stratosphere ozone in 3-D models: A simple chemistry and the cross-tropopause flux, J. Geophys. Res., 105(D11), 14,653–14,665.

* Line 136-137: Is O3 radiative forcing also calculated in E3SM? Besides, It would be helpful to specify the value of the radiation timestep.

Based on the context of Lines 136-137, we assume the question is whether O3 numbers are used in the radiative transfer calculation in E3SM. The answer is yes.

We added the radiation timestep of 1 hour on Line 137. Thanks.

* Line 140/Line 152: Does tracer transport scale linearly with the number of advected tracers? For example, would 56 tracers double the computational time for the transport process? This could provide useful insights for future work.

Yes, the tracer transport scales linearly with the advected tracer number. We added “...it scales linearly with increasing tracer numbers as demonstrated in Bradley (2024).” in the response to one of the major points above.

* Line 143: The "e90 tracer method" should be described upon its first mention.

Agreed. We revised Line 143 with the e90 tracer description “e90 is a synthetic tracer introduced with a uniform surface emission rate, calibrated to achieve a global mean mole fraction of 100 parts per billion (ppb) relative to the dry air mass. The tracer undergoes atmospheric decay with an e-folding time scale of 90 days.”

* Line147: How is the specific value of e90 obtained? The calculation method should be provided here.

As noted in the response to the previous question, the e90 emission is calibrated to give a global mean e90 of 100 ppb. Since different tropopause definitions are expected to yield comparable tropospheric air masses, the e90 tropopause threshold is then obtained by calibrating “against the thermal lapse rate tropopause to ensure a similar tropospheric air mass with both methods”, as described on Lines 145-146.

* Line 150: "ChemDyg" appears only here and is not applied in subsequent analyses—or if it is used, it is not clearly indicated.

Good point. The sentence was revised to indicate which figures were generated by ChemDyg: “In this study, ChemDyg is utilized for the analyses presented in Figures 1, 3, 4, and 5 <<Double check the revised figure numbers>>.” The ChemDyg paper is now published, and we updated the reference to Lee et al. (2025) accordingly.

Lee, H.-H., Q. Tang, M. J. Prather, and J. Xie (2025), A Comprehensive Chemistry Evaluation and

Diagnostics Package for E3SM – ChemDyg Version 1.1.0, Environ. Model. Softw., 191, 106498,

doi:10.1016/j.envsoft.2025.106498.

* Line327: The authors present only the overall bias of E3SM-chem relative to observations but do not provide the corresponding values for E3SM, merely stating that it is "slightly worse than E3SM." It is recommended to include the specific numerical values for comparison.

Revised to “... slightly worse than E3SM (a high bias of 0 to 6 DU (~1.5%))”.

* Line 328: While Figure 6d shows that the bias between E3Sm-chem’s SCO and observations’ is indeed smaller in the tropics, E3SM-chem appears to have a larger bias than E3SM at other latitudes outside the tropics. It’s hard to say the E3SM-chem improves the simulations of SCO.

Added this sentence on Line 329. “Outside the tropics, E3SM-chem exhibits a bias with an absolute magnitude slightly larger than that of E3SM, except for a nearly identical bias observed at 50–60°S.”

* Line 333：The authors do not specify the data years for E3SM, E3SM-chem, and ERA5. Figure 7 notes that the model years are 1985–2014, whereas ERA5 covers 1979–2019. It would be more convincing to align the comparison periods, such as using 1985–2014 for all these three.

We updated the ERA5 panel in Figure 7 to match the same time period as the models (1985–2014) and added the corresponding years on Line 334. The ERA5 climatology exhibits only minor differences (4.49 ppm vs. 4.42 ppm) when averaged over 1979–2019 versus 1985–2014, and these adjustments do not affect our conclusions.

* Line 347-349: Why is there a difference in the polar stratospheric cloud chemistry temperature threshold between E3SM-chem and E3SM? The authors should explain the reason for this difference.

The chlorine-catalyzed ozone-hole chemistry in E3SM-chem and E3SM is parameterized following the method of Cariolle et al. (1990), which depends on the polar stratospheric clouds (PSC) chemistry temperature threshold. E3SM-chem and E3SM simulations display differences in their temperature fields in Antarctica. Hence, we recalibrated the PSC chemistry temperature threshold to better match observations. This is standard with the Cariolle and other polar chemistries because of the threshold behavior of chlorine-catalyzed ozone destruction.

The revised text reads “The SH O3 is very sensitive to the temperature and dynamics over Antarctica, and these have changed in part to the new abundances of O3 and stratospheric H2O with E3SM-chem. Therefore, we retune the polar stratospheric clouds chemistry temperature threshold to 198.0 K in E3SM-chem vs. 197.5 K in E3SM (Tang et al., 2021) to better match the ozone hole observations.”

* Figure 10: The units should be standardized to W m-2. (consistent with Figure 9). In Figure 10, even panels (a) and (b) use different unit notations. The authors should carefully review the manuscript for similar issues.

Thanks. Good catch. The inconsistent unit format is due to the diagnostic tool updates. We replotted Figures 10 and 13 to use the same format.

* Line 378: This statement contradicts the title. The authors should clarify which version of E3SM this work is based on.

Please see our response above. This study implements interactive chemistry on top of E3SMv2, which will be included as one of the new features of E3SMv3. We further clarified this by revising Line 378 to “We successfully implement interactive chemistry on top of E3SMv2 (as a new feature in the upcoming E3SMv3 release)...”