Response to referee comments on "A machine learning-guided adaptive algorithm to reduce the computational cost of atmospheric chemistry in Earth System models: application to GEOS-Chem versions 12.0.0 and v12.9.1"

We thank the editor for his careful reading of the manuscript and the valuable comments. This document is organized as follows: the editor's comments are in *italic*, our responses are in plain text, and all the revisions in the manuscript are shown in blue. **Boldface blue text** denotes text written in direct response to the Referee's comments. The line numbers in this document refer to the updated document with tracked changes.

There is still one last important point left to be addressed, however. You have introduced the use two RRMSE definitions now, with cutoff concentrations corresponding to 99% of the global- and domain-wise masses. Please remove the former one – this is the issue we have discussed earlier about using the same threshold concentration for the entire atmosphere, it is deceptive. Global threshold may not be used for separate atmospheric domains, full stop.

Response. Thanks. We now remove all text and figures using the global threshold.

Line 181. We use the Relative Root Mean Square Error (RRMSE) metric as given by Sandu et al. (1997) to characterize the error in our reduced mechanism:

$$RRMSE_i = \sqrt{\frac{1}{Q_i}\sum_{j=1}^{Q_i}\left(\frac{n_{i,j}^{\text{reduced}} - n_{i,j}^{\text{full}}}{n_{i,j}^{\text{full}}}\right)^2} \quad (7)$$

where $n_{i,j}^{\text{reduced}}$ and $n_{i,j}^{\text{full}}$ are the concentrations for species i and gridbox j in the reduced and full chemical mechanisms, and the sum is over the Q_i ordered gridboxes **that account for 99% of the total mass of species *i* in the boundary layer (surface to 2 km), free troposphere (2 km to tropopause), and stratosphere (the 99% thresholds are different in different atmospheric domains).**

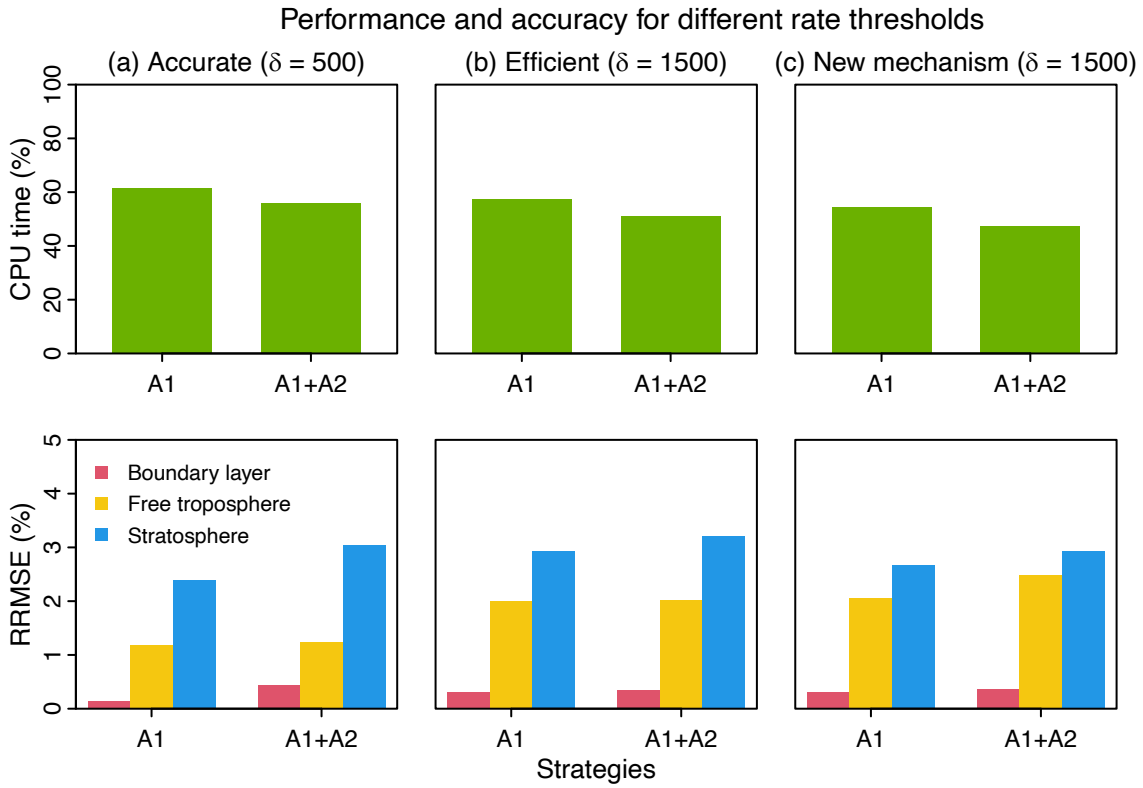

Figure 6. **Performance and accuracy of the adaptive chemical mechanism**. We test the performance of the adaptive method by (A1) removing slow species (P_i or $L_i > \delta$) and (A2) removing slow reactions (reaction rate < 10 molecules cm^{-3} s^{-1}). Results are shown on the last day of 3-year simulations. The unit of δ is molecules cm^{-3} s^{-1}. The performance is measured by the computing processor unit (CPU) time used by the chemical operator, and the accuracy is measured by the median relative root mean square error (RRMSE) for species concentrations using the full chemical mechanism in the boundary layer (0-2 km altitude), free troposphere (2 km to tropopause), and stratosphere. For (a) and (b), we use δ as 500 and 1500 molecules cm^{-3} s^{-1} in GEOS-Chem 12.0.0 that has 228 species and 724 reactions. For (c), we port the algorithm to GEOS-Chem 12.9.1 that has 262 species and 850 reactions. The number of blocks (N) is 13 and the number of chemical regimes is 21 (20 submechanisms (M=20) and one full mechanism).

Fig. 6 shows RRMSE, however it is not clear, which one, please make sure that the domain-wise cutoff is being used. Your new Figure S5 (with domain-wise cutoffs) presents excellent results, so please use it the manuscript instead of Figure 7. The latter, together with the "global" threshold RRMSE use (I refer to lines 185-189 of the manuscript v.6) must be removed.

Response. Thanks. We have removed the old Figure 7 and replaced it with Figure S5. All texts using the global threshold are removed now.

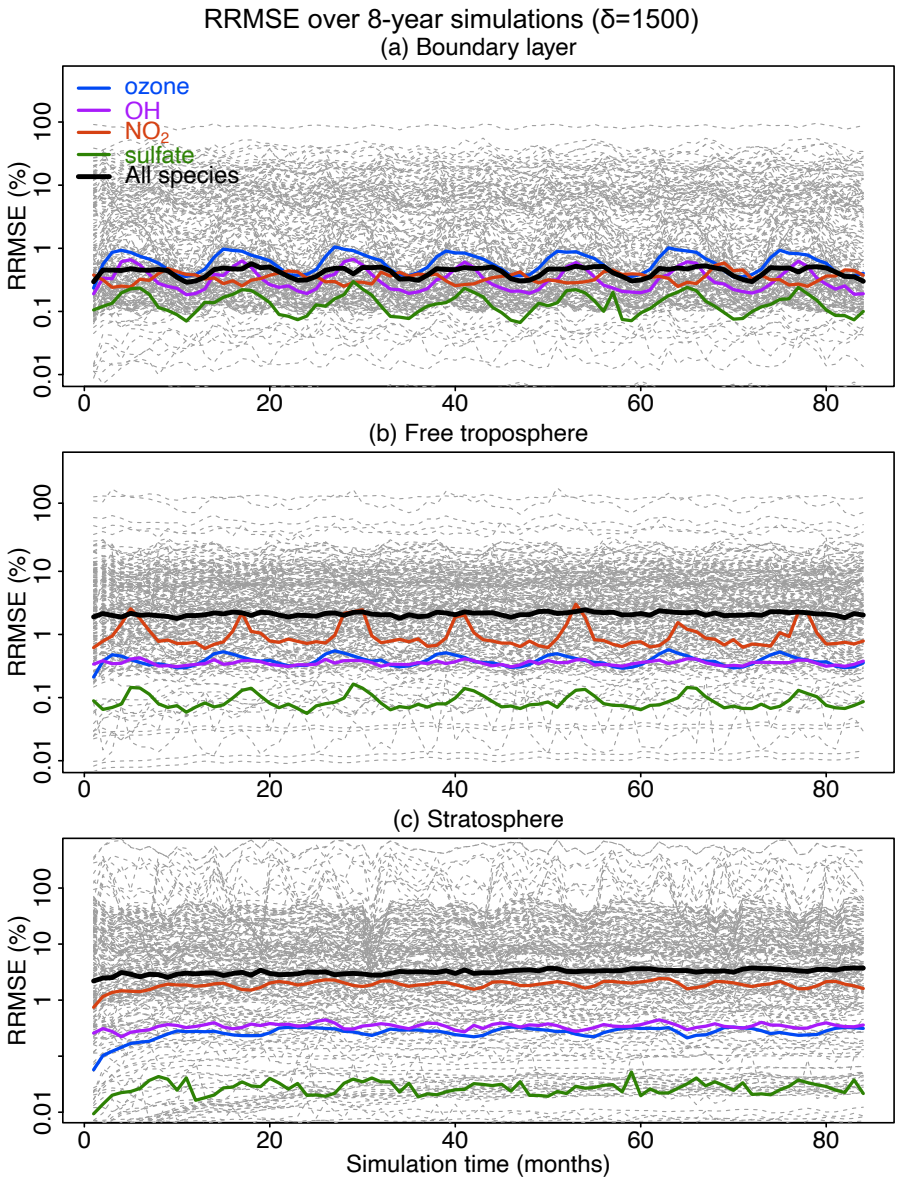

Figure 7. Accuracy of the adaptive reduced chemistry mechanism algorithm over an 8-year GEOS-Chem simulation using a threshold δ of 1500 molecules cm^{-3} s^{-1} to separate fast and slow species. We show the RRMSE in the (a) boundary layer, (b) free troposphere, and (c) stratosphere. Results are also shown for the median RRMSE across all species in the mechanism and more specifically the RRMSE for ozone, OH, NO_2, and sulfate.