

## 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.**

# Summary of Comments on Microsoft Word - 3\_reply\_to\_comments.docx

---

Page: 2

---

 Number: 1 Author: gromov Subject: Highlight Date: 09/01/2022 15:23:40

---

Please change the vertical axis in panel (c) so that it presents the area between -1% and +1% at a greater magnification (you can use axis breaks). The most interesting results are in this area! (I confirm that after magnifying this figure in vector manipulation software)

Please see the comment regarding RRMSE below. Following it:

- Panel (a) should be removed,
- Analysis should be shown for at least the boundary layer, free troposphere and outside-of-troposphere domains.

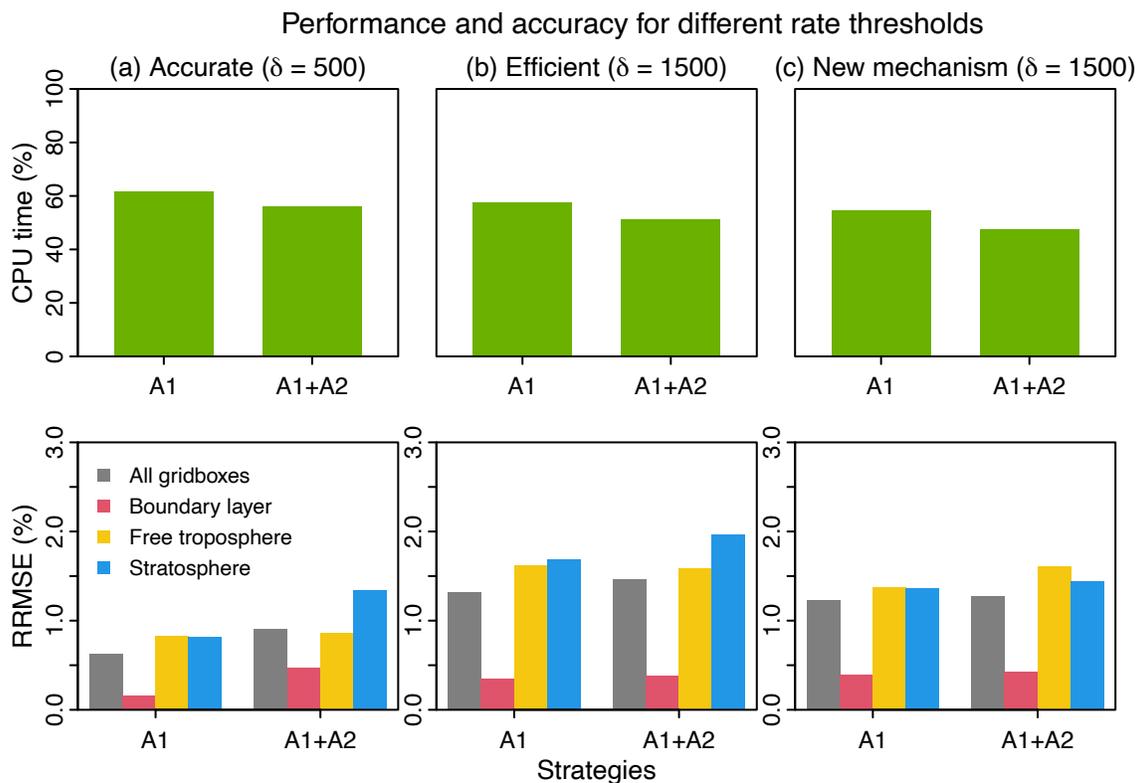
**Response.** Thanks for these suggestions. The axis breaks didn’t work well so we have a supplementary figure that specially presents the -1% and +1% area (Figure S6). We have removed panel (a) and show results for the boundary layer, free troposphere and stratosphere domain. We have updated Figure 6, added new Figure 7-8, S5-S8. We discuss these results in Line 310-355 (Section 3.2). Please check them for details.

Line 310-355. By isolating slow species (A1), we can reduce the chemical integration time by 38-43% with median errors of 0.6-1.3% among all species of all gridboxes. By further removing the slow reactions in each submechanism (A1+A2), we can reduce the CPU time by 44-49% and the median RRMSEs for the full atmospheric domain remain at 0.9-1.5%. The median RRMSEs are <0.4% in the boundary layer and 0.8-2.0% in the free troposphere and stratosphere. When using a higher threshold  $\delta = 1500$  molecules  $\text{cm}^{-3} \text{s}^{-1}$  to isolate slow species and removing the slow reactions, we can reduce the chemical integration time by 50%, and the median RRMSE is smaller than 2% for the full atmospheric domain (smaller than 0.5% in the boundary layer). The relative error on concentrations compared to the standard simulation is below 0.5% in tropical and mid-latitude regions for key species like O<sub>3</sub>, OH, sulfate and NO<sub>2</sub>, and could be higher (1-6%) in high latitudes where more chemical complexity reduction happens (Figure S4). Using a higher threshold of  $\delta (> 1500)$  only leads to marginal improvement in computer time but the RRMSE quickly increases.

Figure 7 shows the evolution of the RRMSE over an 8-year period for all 228 individual species in the mechanism, using a  $\delta$  of 1500 molecules  $\text{cm}^{-3} \text{s}^{-1}$  to isolate slow species and also removing the slow reactions. The results are shown in different atmospheric domains including the boundary layer, free troposphere, and stratosphere. There is no significant growth in error over the 8-year period. The RRMSEs for key species including ozone, OH, sulfate and NO<sub>2</sub> are smaller than 0.5% and are within  $\pm 10\%$  for >95%

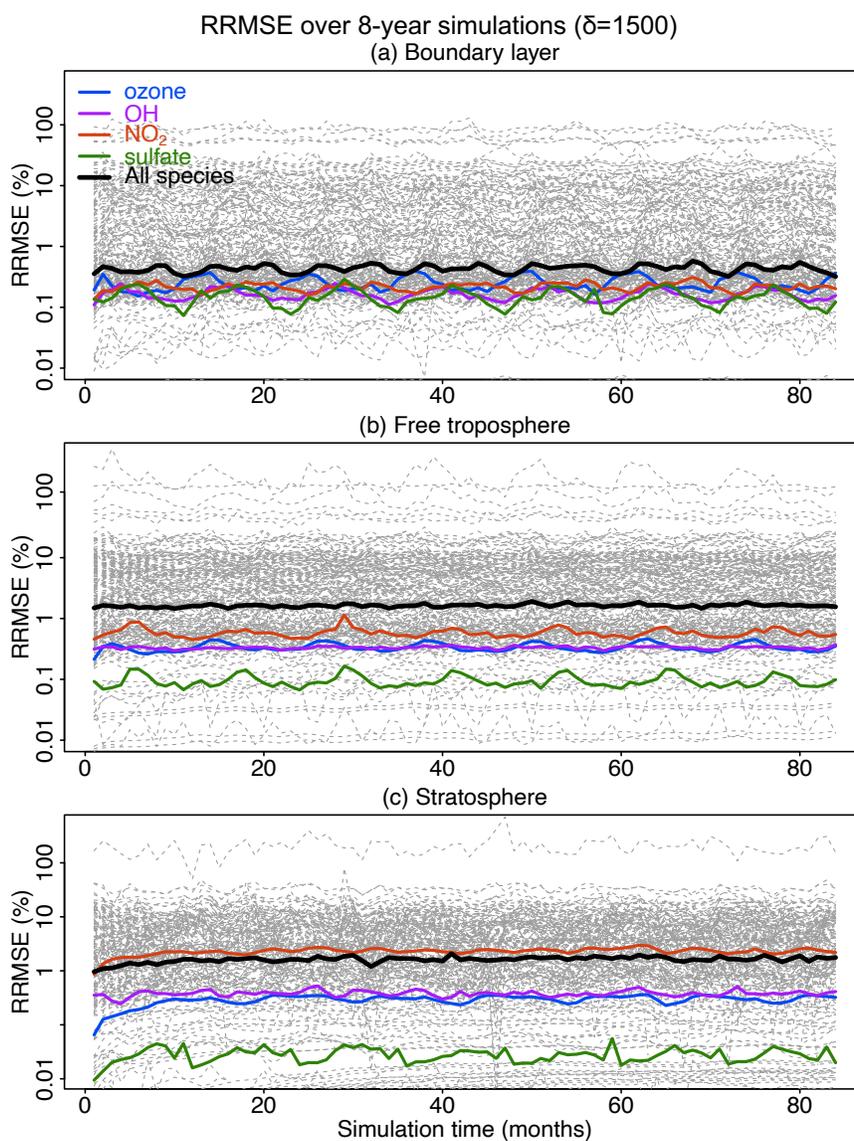
of the other species in the boundary layer (Figure 7a). The median RRMSEs are higher (1-2%) in the free troposphere and stratosphere where most of the reduction of chemical complexity occurs (Figure 7b-c). The median RRMSE in the stratosphere increases slightly in the first 20-30 months and then stabilizes, reflecting the long time scale for chemical aging to abate the sensitivity to initial conditions. We also calculate the RRMSEs for each species in the three atmospheric domains using three different thresholds so that we can account for 99% mass in each domain (Figure S5). This means more gridboxes with lower species concentrations will be accounted for in the calculation so the RRMSE are slightly higher (0.4% in the boundary layer and 2-3% in the free troposphere and stratosphere). Table S1 lists the species with 10% highest RRMSE in each of the three atmospheric domains, dominated by secondary VOCs and iodine radicals in the troposphere, and VOC species in the stratosphere. None of these species play a central role in the chemistry for the corresponding atmospheric domains.

Figure 8 shows the relative differences of global atmospheric masses over the 8-year simulation in the boundary layer, free troposphere, and stratosphere. The relative differences are within 10% for >99% of the species in troposphere and for >95% of the species in the stratosphere. Species with the largest errors are inorganic halogens and VOC species (more details can be found in the boxplots in Figure S6 and S7). Table S1 lists the species with 10% highest relative bias in atmospheric masses; all have minor importance in atmospheric chemistry. OH has a bias <0.2% in the troposphere and <0.01% in the stratosphere. Other key species like ozone and sulfate have a relative difference <0.5% in the troposphere and <0.1% in the stratosphere (Figure S8). The relative difference for NO<sub>2</sub> in the stratosphere changes slightly from 0% to -0.6% in the first 30 months and then stabilizes at -0.6% (Figure S8).



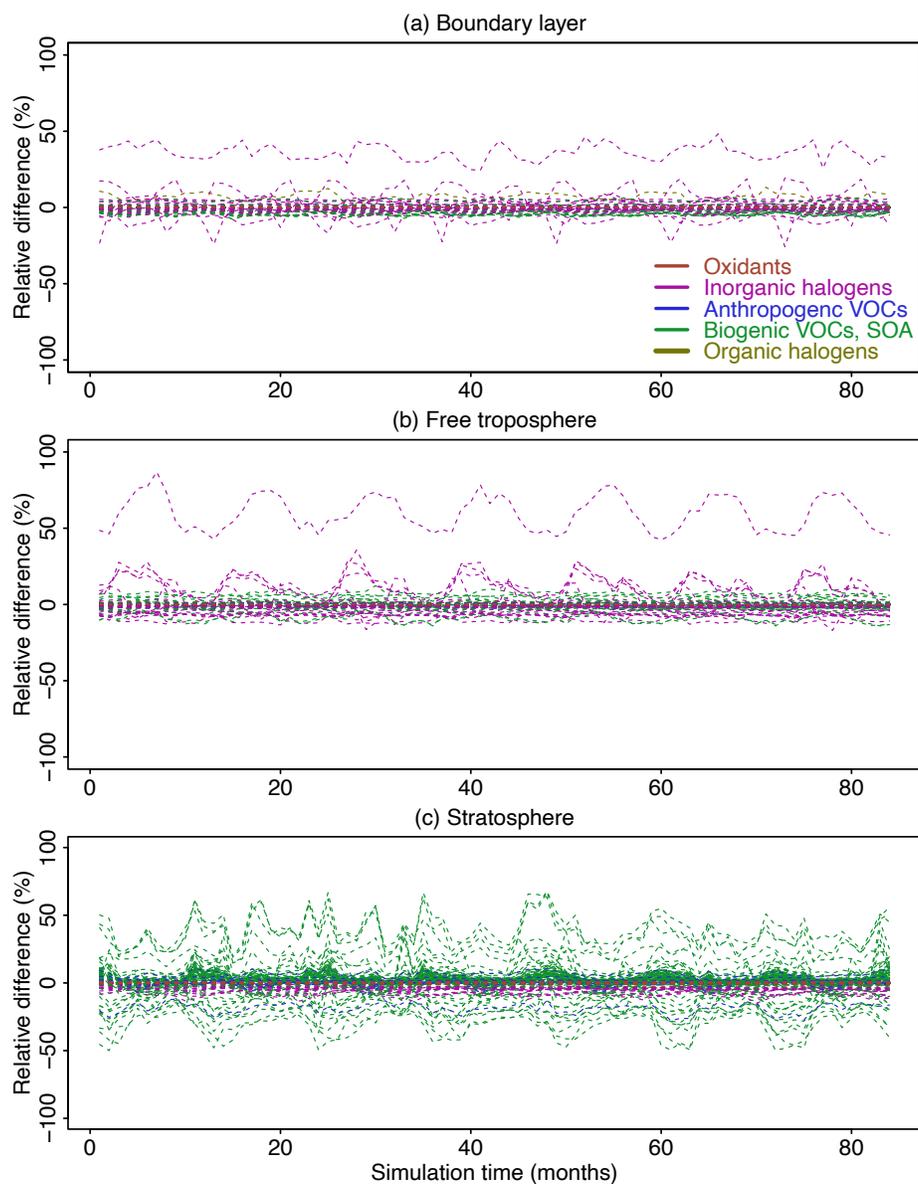
**Figure 6. Performance and accuracy of the adaptive chemical mechanism.** We test the performance of the adaptive method by (A1) removing slow species ( $P_i$  or  $L_i > \delta$ ) and (A2) removing slow reactions (reaction rate  $< 10$  molecules  $\text{cm}^{-3} \text{s}^{-1}$ ). Results are shown on the last day of 3-year simulations. The unit of  $\delta$  is molecules  $\text{cm}^{-3} \text{s}^{-1}$ . 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 for all gridboxes, in the boundary layer (0-2 km altitude), free troposphere (2 km to tropopause), and stratosphere. For (a) and (b), we use  $e \delta$  as 500 and 1500 molecules  $\text{cm}^{-3} \text{s}^{-1}$  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).



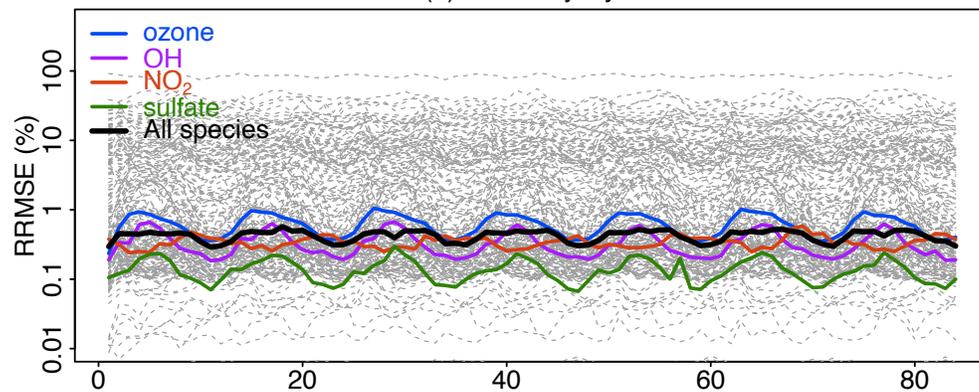
**Figure 7.** Accuracy of the adaptive reduced chemistry mechanism algorithm over an 8-year GEOS-Chem simulation using a threshold  $\delta$  of 1500 molecules  $\text{cm}^{-3} \text{s}^{-1}$  to separate fast and slow species. We show the RRMSE in the (a) boundary layer, (b) free troposphere, and (c) stratosphere. Results are also shown for the median RRMSE across all species in the mechanism and more specifically the RRMSE for ozone, OH,  $\text{NO}_2$ , and sulfate.

Relative differences of atmospheric masses over 8-year simulations ( $\delta=1500$ )

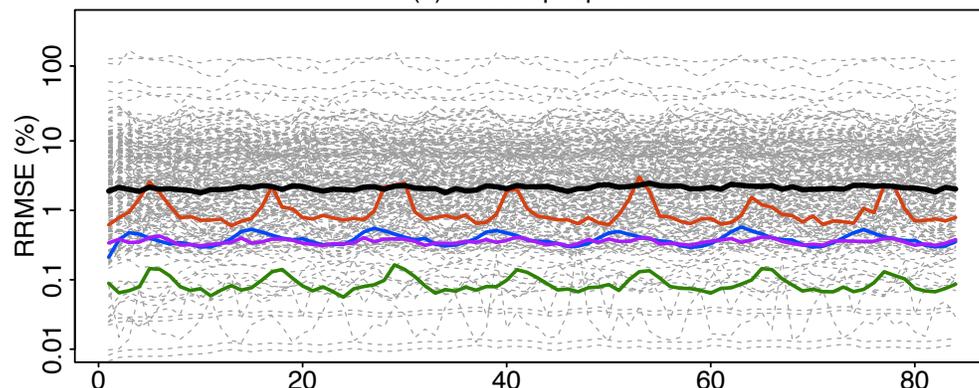


**Figure 8.** Relative difference of atmospheric masses in the adaptive reduced chemistry mechanism algorithm over an 8-year GEOS-Chem simulation using a threshold  $\delta$  of  $1500 \text{ molecules cm}^{-3} \text{ s}^{-1}$  to separate fast and slow species. We show the RRMSE in the (a) boundary layer, (b) free troposphere, and (c) stratosphere. Different colours denote different species categories (more details can be found in Table 1). Figure S8 presents more detailed results for the species with RRMSEs in the -1.5% to 1.5% range.

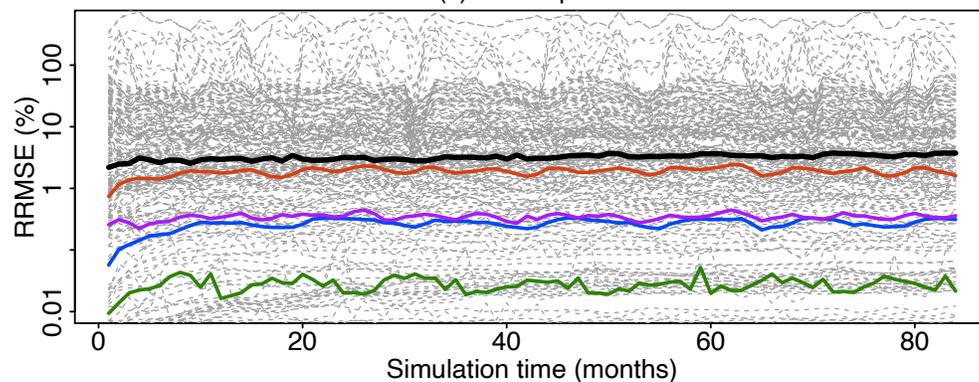
RRMSE over 8-year simulations ( $\delta=1500$ )  
(a) Boundary layer



(b) Free troposphere

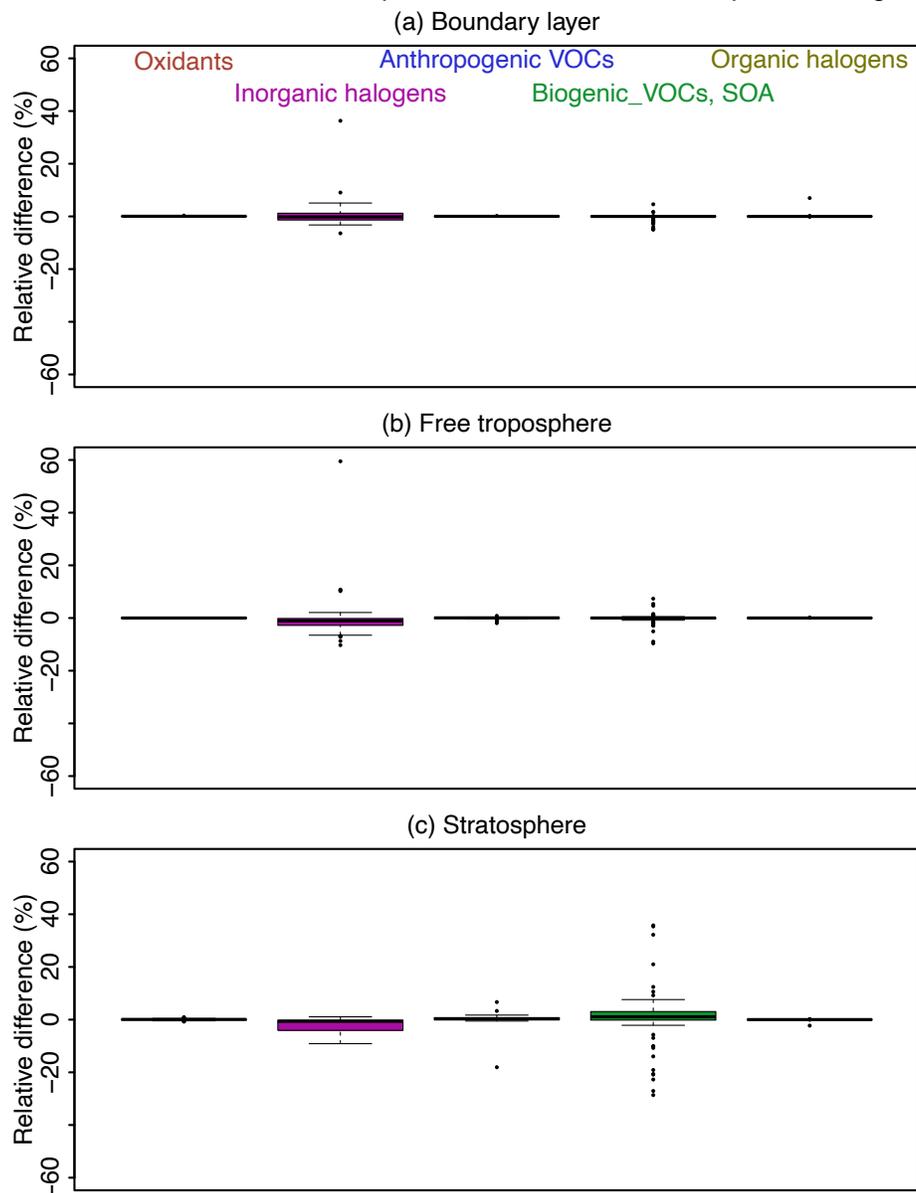


(c) Stratosphere



**Fig. S5.** Same as Figure 7 but we calculate the RRMSE by accounting for gridboxes that can comprise 99% mass in each atmospheric domain (the 99% thresholds are different in different domains in this case), including the (a) boundary layer, (b) free troposphere, and (c) stratosphere.

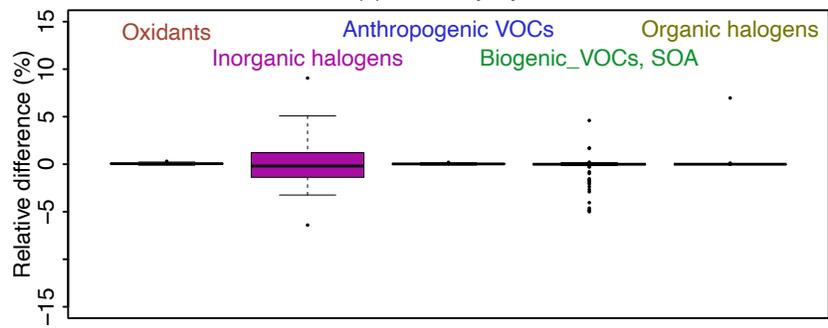
Relative differences of atmospheric masses for different species categories



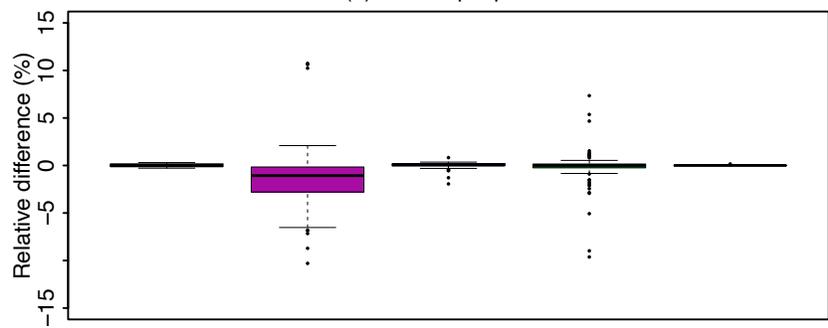
**Fig. S6.** The relative differences of atmospheric mass (averaged in the 8-year simulations) in the adaptive reduced chemistry mechanism for different species categories in the (a) boundary layer, (b) free troposphere, and (c) stratosphere. The top and bottom of each box are the 25th and 75th percentile, and the centerline is the 50th percentile. We use a threshold  $\delta$  of  $1500 \text{ molecules cm}^{-3} \text{ s}^{-1}$  to separate fast and slow species here.

Relative differences of atmospheric masses for different species categories

(a) Boundary layer



(b) Free troposphere



(c) Stratosphere

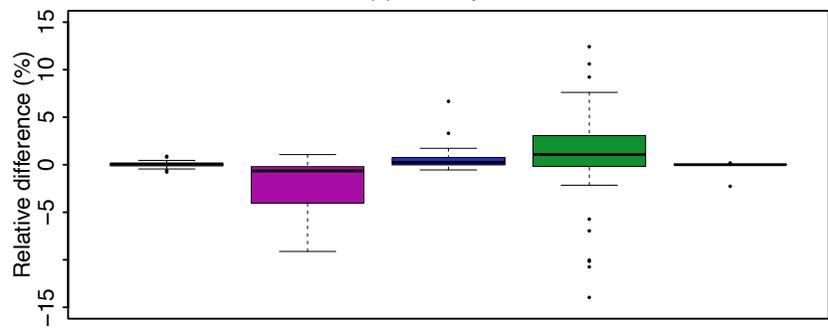
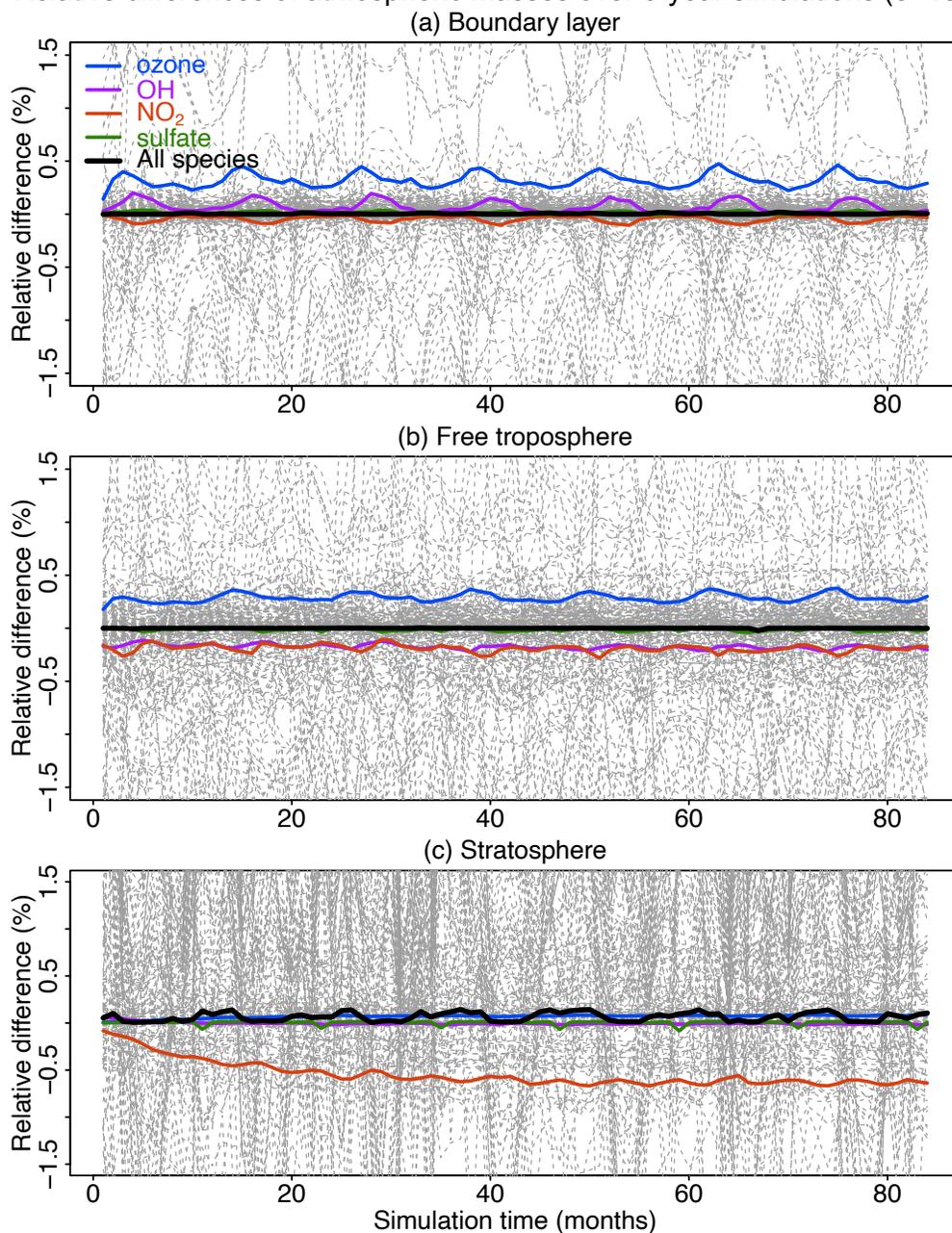


Fig. S7. Same as Fig. S6 but we present more details for the data from -15% to 15%.

Relative differences of atmospheric masses over 8-year simulations ( $\delta=1500$ )



**Fig. S8.** Same as Figure 8 but for more details in the range between -1.5% and 1.5%. Results are also shown for the median relative difference across all species in the mechanism and more specifically for ozone, OH, NO<sub>2</sub>, and sulfate.

## Page: 3

---

**T** Number: 1 Author: gromov Subject: Highlight Date: 04/01/2022 23:44:01

---

Exactly this statement has to be present in the manuscript. Because it explains how you came to this number!

**Response.** Thanks. We now say

Line 102. In each submechanism, if a reaction is slower than  $10 \text{ molecules cm}^{-3} \text{ s}^{-1}$  for all gridboxes that select this submechanism, then the reaction is considered negligible and removed from the submechanism. The logic is that such a slow reaction will not contribute significantly to the total species production/loss rate threshold  $\delta = 500\text{-}1500 \text{ molecules cm}^{-3} \text{ s}^{-1}$ . About 40-60% reactions can be removed using this strategy without incurring significant error. For example, reactions of short-lived volatile organic compounds (VOCs) are removed in stratospheric gridboxes, and daytime photochemical reactions are removed in nighttime gridboxes. **Tests indicate that increasing the reaction rate threshold to  $100 \text{ molecules cm}^{-3} \text{ s}^{-1}$  incurs significant error.**

---

**T** Number: 2 Author: gromov Subject: Highlight Date: 09/01/2022 20:51:35

---

Same as above, this statement has to be in the manuscript, augmented with the one mentioning that decreasing distances allows further optimisation without increasing the errors (what you formulate as "did not significantly change the results" highlighted below).

**Response.** Thanks. We now say

Line 144. Second, for each species  $i$ , we decrease its distance with the 5 species that have highest similarity with it by 50% and this scaling is applied only once for each species pair. The logic is that the number of species with similar chemical characteristics is usually around 5 and decreasing the distances among them by 50% can increase the probability of these species to be in the same chemical blocks after the optimization process. **We carried out a number of tests by perturbing the parameters used here and examine if the optimized chemical blocks are chemically logical, and results show that using 10 highest-similarity species instead of 5 or decreasing distances by 30%-70% instead of 50% did not significantly change the results.**

**T** Number: 3 Author: gromov Subject: Highlight Date: 09/01/2022 19:56:03

---

I believe you do may not correctly understand the meaning of the term "intuition" (look up the definition, e.g. at <https://dictionary.cambridge.org/dictionary/english/intuition> ). If your feelings determine which mechanism is correct, then your are submitting this study to a wrong journal -- GMD is dealing with natural sciences, foremost mathematics, physics and chemistry that are based on facts and logic.

Having said this, I admit that intuition is a subjective criterion and may precede an inference in physical sciences, however it may not be used as an argument in contrast to a strict mathematical/logical proof.

Please refrain from using this term. Also -- I repeat -- please give a clear definition (in the beginning, i.e. in the Introduction) to the "chemical coherence" term that you are using throughout the manuscript. "Coherence" in your case, I recon, means "having its parts related in an organized and reasonable way" (see definition at, e.g., <https://dictionary.cambridge.org/dictionary/english/coherent> ), so please explicate in which parts the mechanisms called "chemically coherent" are related in organised and reasonable way. In a way, this was somewhat satisfied in the sentence that you've edited (I refer to ll. 58-60 of the manuscript version 5 as compared to ver. 4). Please refrain from using "chemical coherence" in the abstract at all.

**Response.** Thanks for pointing this out. We have removed 'coherent' everywhere in the paper. In some places, we use 'chemically logical' and define it in the introduction.

[Line 37. chemically logical \(i.e., retaining connections between species involved in the same or similar reactions\)](#)

**T** Number: 4 Author: gromov Subject: Highlight Date: 04/01/2022 23:46:17

---

**Response.** Done.

**T** Number: 5 Author: gromov Subject: Highlight Date: 09/01/2022 20:37:47

---

I see improvements in Figs. 7 & 8, however I still do not see any relevance in what is shown S4 and S8. What if the most important species turn out to be those in >90% quantile part with RRMSE of 10% and higher? Which are those species that have highest RRMSE?

**Response.** Now we give a list of these species that have a mass bias or RRMSE in the >90% quantile part in table S1 and also discuss the results in main text.

[Line 346. Table S1 lists the species with 10% highest RRMSE in each of the three atmospheric domains, dominated by secondary VOCs and iodine radicals in the troposphere, and VOC species in the stratosphere. None of these species play a central role in the chemistry for the corresponding atmospheric domains.](#)

[Line 354. Table S1 lists the species with 10% highest relative bias in atmospheric masses; all have minor importance in atmospheric chemistry.](#)

[Table S1. Top 10% Species with highest RRMSE and relative mass bias in the boundary layer, free troposphere and stratosphere.](#)

	RRMSE	Relative mass bias
Boundary layer	I2O4, I2O2, I2O3, IBr, INO, ICl, IONO, OIO, PRPN, HOI, BrSALC, BrSALA, ISNOHOO, IONO2, ISNOOB, I2, PRN1, MAN2, MACRN, MAOPO2, IO	I2O2, I2O3, INO, TRO2, N, XRO2, Cl2O2, IONO, OIO, PRPN, Br2, BrSALC, BrSALA, ISNOHOO, IONO2, ISNOOB, I2, Cl2, PRN1, MAN2,

		MACRN, MAOPO2
Free troposphere	ISN1OG, I2O4, I2O2, I2O3, INO, ICI, OIO, PRPN, BrSALA, MAOP, ISNOHOO, ISNOOB, I2, PRN1, MAN2, ISNOOA, MACRN, MAOPO2, I, IO, NMAO3, INO2	AERI, ISN1OG, I2O4, I2O2, I2O3, INO, HI, IONO, OIO, HOI, BrSALC, BrSALA, ISNOHOO, ISNOOB, I2, CI2, PRN1, MAN2, MACRN, I, IO, NMAO3
Stratosphere	ISN1OA, ISN1OG, LVOC, PMNN, MRP, IPMN, MACRNO2, MONITS, GAOO, MVKN, MGLYOO, GLYX, MGLOO, MAN2, MACRN, HCOOH, KO2, MGLY, RIO2, INO2, MRO2	ISN1OA, ISN1OG, DHDC, PRPN, DHPCARP, ISNOHOO, ISNOOB, INPN, I2, PRN1, PROPNN, MAN2, ISNOOA, MACRN, MAOPO2, OLND, OLNN, KO2, NMAO3, ISN1, RIO2, INO2

 Number: 6 Author: gromov Subject: Highlight Date: 09/01/2022 14:49:15

My comment was NOT about N2O, it was about OH. This answer is irrelevant.

**Response.** Yes, you are right that OH is so critical in atmospheric chemistry and a small bias in OH may have large effect on certain species. But the bias of OH is very small in our case (<0.2% in the in the troposphere and <0.01% in the stratosphere) (Figure S7) and there is no significant sign of increasing errors over our 8-year period (see Figure 7-8, S5-S7 for more details). We have removed all our old statements claiming that our algorithm can be applied in earth system models and make it clear that our error is stable in an 8-year period (we never over-state the performance of our method now). Now we say this in the text. [Line 354. OH has a bias <0.2% in the troposphere and <0.01% in the stratosphere.](#)

## Page: 4

 Number: 1 Author: gromov Subject: Highlight Date: 09/01/2022 20:42:26

Relative differences for global atmospheric masses do not make any sense because species are not evenly distributed spatially and their concentrations change within orders of magnitude, also following the air density changes with height. What does an error of 10% for global mass tell me? Is it in the stratosphere, which has about 10% of all global mass, or at the surface? If the former, you have 200% error in the stratosphere?

See also the comments above regarding the calculation of RRMSE and presentation of errors for all species.

**Response.** Thanks for this good point. Now we show the error of boundary layer, free troposphere and stratosphere. Please see Figure 7-8, S5-S8 for more details. The related discussion can be found in Line 336-356.

---

**T** Number: 1 Author: gromov Subject: Highlight Date: 09/01/2022 20:30:13

---

I am still not fine with the averaging used in Eq (7) (instead of selection of maximum error across the gridboxes as in the cited literature). But, OK, then the only fairly acceptable form of threshold (a) for RRMSE you may define is the relative one (i.e. 5th percentile that you use, I do not advise to put a lower limit at all; even if, then  $\leq 1$  molecule  $\text{cm}^{-3}$ , as in classical approaches you refer to).

Therefore: please remove any results for "absolute" thresholds from entire manuscript - this will affect Figs. 6 to 8 and S4 to S8, if I did not miss anything. Choosing an absolute threshold value at this magnitude is not appropriate - the obtained RRMSE estimates are deceptive (e.g. at the value of  $10^6$  molec/ $\text{cm}^3$ , only half of gridboxes with OH in your results will be accounted for on average). There are other species that have similar or even lower than OH concentrations in the atmosphere. Using fixed a may also create very large RRMSE values with large abundances, have you thought of this?

Furthermore, in the upper troposphere and stratosphere air density (and respective concentrations of trace gases with similar mixing ratios) is orders of magnitude lower. Therefore, please present the error analysis in Figure 8 for at least three atmospheric domains, viz. boundary layer (where most of "pollution chemistry" occurs), free troposphere (where, e.g., a major part of tropospheric OH is) and outside the troposphere. This is adamant for such analysis. For Fig. 7 it may be enough to show boundary layer results. I hope that you are aware of the fact that the value of (a) should be chosen from the statistic on concentrations pertaining to the given domain, not all simulated gridboxes.

**Response.** Thanks for this good point. Now we have re-defined the RRMSE and we show the errors in the boundary layer, free troposphere and stratosphere.

Line 215. 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 global mass of species  $i$ . We calculate separate RRMSEs for the boundary layer (surface to 2 km), free troposphere (2 km to tropopause), and stratosphere with the same 99% threshold in each atmospheric domain. As a test, we also calculate the RRMSEs over the gridboxes that can account for 99% of the mass in each atmospheric domain (the 99% thresholds are different in different domains in this case).

---

**T** Number: 2 Author: gromov Subject: Highlight Date: 09/01/2022 15:05:52

---

A typo? What is chemical initiation???

**Response.** It is a typo. Fixed.

---

**T** Number: 3 Author: gromov Subject: Highlight Date: 04/01/2022 19:09:01

---

See the comment regarding the intuition above

**Response.** Thanks. We now use 'chemically logical' and define it in the introduction.

Line 37. chemically logical (i.e., retaining connections between species involved in the same or similar reactions)

---

**T** Number: 4 Author: gromov Subject: Highlight Date: 09/01/2022 20:21:01

---

This is a very vague reply statement I cannot agree with unless you present solid arguments. Systematic biases are never welcome when estimates regarding any reactive (oxidant) species are made, e.g. not only CH<sub>4</sub>, a rather large number of trace gases are being oxidised by the HO<sub>x</sub> and NO<sub>x</sub>. An 8-yr simulation, I tend to agree here, will likely be sufficient to have reactive N-bearing reservoirs equilibrated.

**Response.** In our last replay statement, we say ‘A 30-year simulation is not necessary for tropospheric oxidant/aerosol chemistry’. As seen from Figure S6, the mass bias for atmospheric oxidants is very small. For OH, the bias is <0.2% in the troposphere and <0.01% in the stratosphere. Even though this cannot rule out the possibility of increasing errors over a 30-year period, our 8-year simulation results imply that our algorithm works well in a timeframe that is suitable for many tropospheric atmospheric chemistry studies (e.g. air quality studies). In text, we make it clear that the error is stable in an 8-year simulation and we don’t over-state our algorithm’s performance. We hope these can ease your concern here.

Line 354. OH has a bias <0.2% in the troposphere and <0.01% in the stratosphere.

Line 449. It is stable (no error growth over time) for 8-year simulations.

## Page: 6

---

**T** Number: 1 Author: gromov Subject: Highlight Date: 09/01/2022 16:09:15

---

Sorry, you still haven't made it clear. "Either ... could be" or "both ... are" reactants? Reactants are those appearing on the LHS of the chemical equation. Products are on the RHS. Do indices refer to reactants only?

**Response.** Now we make it clear in text.

Line 128. Where  $T_{i,j}$  is the number of reactions that include both species  $i$  and  $j$  (with  $i$  as reactant and  $j$  as product, or  $i$  as product and  $j$  as reactant)...

# Summary of Comments on Microsoft Word - 1\_Manuscript\_clean.docx

---

## Page: 3

---

**T** Number: 1 Author: gromov Subject: Highlight Date: 09/01/2022 14:39:21  
of GEOS-Chem?

---

**Response.** It is a typo. Fixed now.

## Page: 7

---

**T** Number: 1 Author: gromov Subject: Highlight Date: 09/01/2022 15:02:19

---

Check the reference (w.r.t. that given in the References section), this metric was introduced in

A. Sandu, J.G. Verwer, M. Van Loon, G.R. Carmichael, F.A. Potra, D. Dabdub, J.H. Seinfeld. Benchmarking stiff ode solvers for atmospheric chemistry problems-I. implicit vs explicit. Atmospheric Environment, Volume 31, Issue 19, 1997, [https://doi.org/10.1016/S1352-2310\(97\)00059-9](https://doi.org/10.1016/S1352-2310(97)00059-9).

**Response.** We have corrected it. Now we cite the Sandu's part I paper instead of part II.

## Page: 20

---

**T** Number: 1 Author: gromov Subject: Highlight Date: 09/01/2022 20:02:17

---

Do you mean RRMSE?

Is this also calculated with  $a=10^6$ ?

**T** Number: 2 Author: gromov Subject: Highlight Date: 09/01/2022 20:14:14

---

**Response.** We now use a different way to define RRMSE and have fixed the typo.

## Page: 21

---

**T** Number: 1 Author: gromov Subject: Highlight Date: 09/01/2022 20:14:53

---

Eq. 7 ?

**Response.** It is a typo. Now fixed.