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 referees for their careful reading of the manuscript and the valuable comments. This document is organized as follows: the Referee'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 manuscript with tracked changes.

Referee 1

The computational cost of atmospheric chemistry with atmospheric chemistry transport or an Earth System Models is large and methods to reduce these costs are so useful. The authors have over the years presented a series of papers (Santillana et al., 2010; Santillana et al., 2016; Shen et al., 2020) that attempt to reduce this computational burden by means of (simplifying here) separating the chemistry into fast species for which the differential equations need to be explicitly solved and slow species which can be solved analytically. The complexity of the approach has increased over the years and this paper represents the current incarnation of the methodology.

These technical advances in the numerical methods for atmospheric chemistry transport model (and more generally for geophysical models) are not seen as being sexy science, however, they are essential if these models are to be useful for the wider community. I am thus supportive of this paper and would suggest publication after some clarifications and corrections suggested below.

Major comments.

***The figures from the supplementary material should be included in the main text of the paper.** For me, it is hard to understand some of the more complex mathematical aspects of the paper without a diagram to support it (e.g. S1 etc). I don't really see why all of the figures from the supplementary material can't be included in the main text. They are not particularly repetitive and I think it would be useful to the reader to see them all.*

Response. Thanks. We have moved Text S1, Figure S1, S2 and S6 to the main text.

***It would be useful at the end of the introduction to provide a context for what is coming up in the rest of the paper**. The lines around 185 provide this commentary about what has been done in the past, the problems associated with those and the method of addressing those which is discussed in the rest of the paper. This would help to contextualize Section 2 which seems like a rather remote set of definitions at the moment. In a few places, the paper feels like it is rather disjointed with one section not necessarily rolling into the next with much cohesion. Perhaps a re-read and a re-think of some of the structure would be beneficial for much of the paper.*

Response. Thanks. We have added a new paragraph at the end of the introduction to discuss what we have done in previous work and what to do next.

Line 69. In this work, we continue developing the adaptive method described by Shen et al. (2020). This method pre-assembles a small number of subsets of the full chemical mechanism representing the range of conditions in the troposphere and stratosphere, and selects the most appropriate submechanism to use in the model locally and on the fly. The submechanisms are constructed by first splitting the full mechanism's atmospheric species into *N* different blocks based on similarity of chemical behaviors, using a machine learning clustering method. We then define the submechanisms as different assemblages of blocks, select *M* of these assemblages to encompass the majority of chemical conditions in the atmosphere, and build them into the model. The choice of submechanism in the model is then made locally by computing chemical production and loss rates of the mechanism species and deciding which need to be part of the coupled chemical computation ('fast' species) and which can be tracked independently ('slow' species). A major development here is to define chemically coherent blocks that allow the method to easily accommodate changes in the chemical mechanism and to be readily applied to different mechanisms. We further improve the performance of the method by reducing the number of reactions as well as the number of species in the submechanisms.

The available code is included in a RAR format. This doesn't seem to decompress on my Mac. Can we have the data in a standard zip format?

Response. It can be decompressed after you remove .rar. I have clarified this in README in the same directory. Thanks.

Minor Comments.

Line 60. This sounds like the end of an abstract rather than an introduction. At this point, the methodology hasn't been explained or tested so how can it be called chemically coherent, accurate or

Response. Thanks. We now rewrite this sentence.

Line 77. A major development here is to define chemically coherent blocks that allow the method to easily accommodate changes in the chemical mechanism and to be readily applied to different mechanisms. We further improve the performance of the method by reducing the number of reactions as well as the number of species in the submechanisms.

Line 87. Can the value of 10 molecules cm-3 s-1 be put into some context? Why was this chosen? Which reactions fall into this category?

Response. Thanks. Now we say
Line 192. In each submechanism, **if a reaction is slower than 10 molecules cm^{-3} s^{-1} over all gridboxes that select this submechanism, this reaction is considered as unimportant in contributing to the threshold δ and is removed from the submechanism; but this reaction will be kept if it is faster than 10 molecules cm^{-3} s^{-1} in any of these gridboxes selecting this submechanism. The threshold we used to separate fast and slow reactions is slightly larger than 0 molecules cm^{-3} s^{-1} because of numerical precisions (unimportant reactions may still have a reaction rate > 0 molecules cm^{-3} s^{-1} in the numerical chemical solver in some timesteps)**. About 40-60% reactions can be removed using this strategy. For example, reactions of short-lived volatile organic compounds (VOCs) are removed in stratospheric gridboxes, and daytime photochemical reactions are removed in nightime gridboxes.

Line 99. Giving the number of species (228?) in the reaction mechanism would be useful for contextualizing the 3400 other numbers?

Response. Thanks. Now we say
Line 216. For example, in the reaction A+B-→C, there are 2 pairs (A-C and B-C) of reactants-products, 3 vertices (A, B, and C) and 2 edges (A-C and B-C).

Line 104. "$T_{i,j}$ is the is the number of reactions that include both species i and j." Is that tas a reactant, products or either?

Response. Now we say.

Line 285. Where $T_{i,j}$ is the number of reactions that include both species i and j (one is the reactant and the other is the product),

Figure S1 should appear here to help explain the definition of Di,j. It would be useful to give an explanation for the numbers which are obtained given the chemical mechanism linking Toluene, xylene and Glyxoyl.

Response. We now move Fig. S1 back to the main text. We have made it clear about how we calculate the numbers in the figure caption.

The methods used to calculate "distance" appears to refer to the mechanism without any chemistry occurring. In the case of A+B-->C and A+D-->C, the "distance" between A and C doesn't care about the 2 rate constants or the concentration of the reactants. It would appear that chemically minor routes or channels are given the same weight as the dominant routes or channels. Presumably, the ideal way of working out the distance between species would be to explore the mechanism with some chemistry occurring and the weigh the distances between vertices by the flux between species or something like that? This obviously has a significant downside of being much more complex to implements and contextual (the distances would change with the chemical environment). It might be useful to describe this as being an optimal approach, but the approach taken is a simplification of this. Otherwise, it is rather hard to understand why the fluxes have not been used to represent the distance between species?

Response. Thanks for raising this good point. We indeed considered weighting the distances using the logarithms of global average chemical reaction rates. However, we don't find this will make a big difference in our results but it makes the method very complex. Now we say this in text.

Line 299. We also tried weighting the species distances using the logarithms of their global mean reactions rates but this does not have significant effects on our final results.

Line 116. It might be worth including the definition of "long-lived" and fast here.

Response. Thanks. This sentence is removed in the revision and we have defined "long-lived" and fast elsewhere.

Line 130. It would be worth pulling in the description of f from the supplementary information.

Response. Now we move Text S1 back into the manuscript.

Line 134. Bring in the figure S2 into the main body.

Response. Done, thanks.

Line 144. By only considering values greater than 1e6 cm-3 from your statistical analysis you are not excluding many values for typically high concentration species (O3 etc) but you will be excluding a significant fraction of the values for OH shown in Figure S6. It is not appropriate to do this for OH. A significant fraction the grid boxes will have OH concentrations less than 1e6 (the canonical global mean value). I appreciate that there might be increased errors at lower concentrations but if a large number of the global grid boxes have lower concentrations than this, this metrics is being rather selective in the grid boxes that it is using the analyse the *model. From Figure S8 it would appear that this excludes evaluation of the model error in large chunks of the surface of the globe for OH?*

Response. We now test also other lower thresholds (10^5). See Figure S5-S6 for more details.

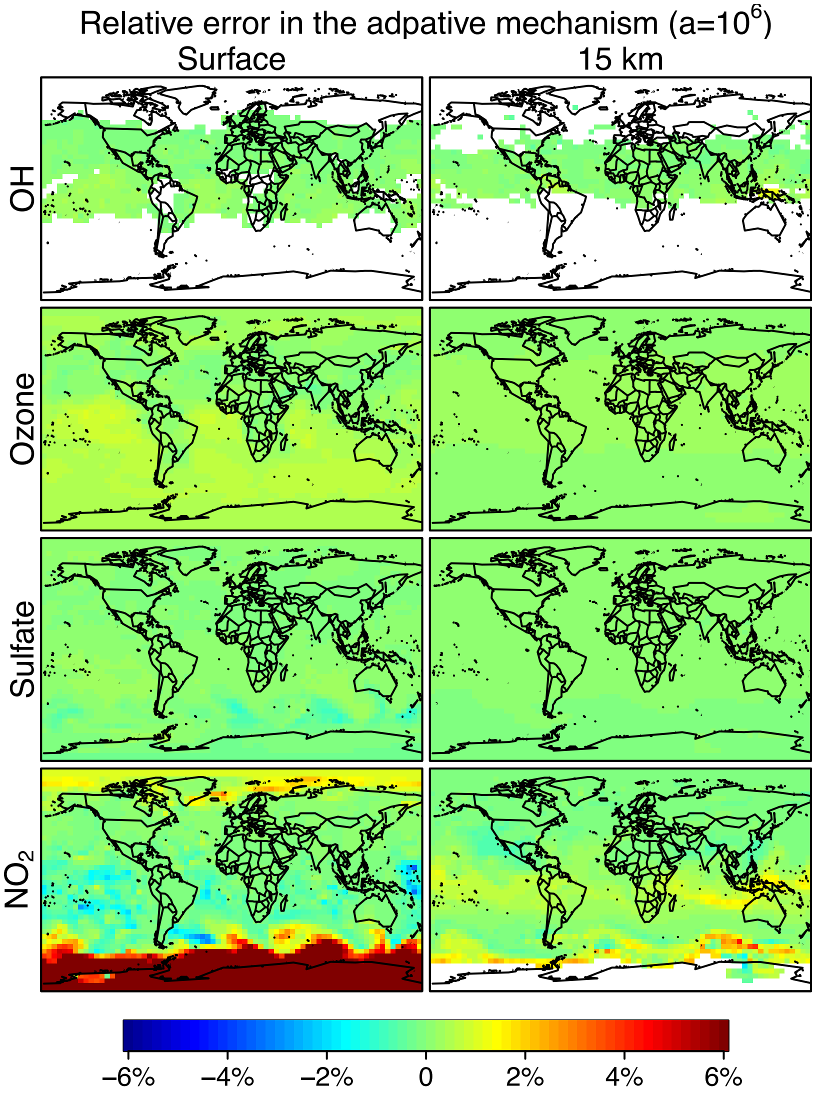

Fig. S5. Relative error from the adaptive mechanism reduction method after three years of simulation in the GEOS-Chem global 3-D model for tropospheric-stratospheric chemistry. The figure shows relative differences of 24-h average OH, ozone, sulfate and NO_2 concentrations relative to the full-chemistry simulation on the last day of the three-year simulation. The calculation removes slow (P and L < 1500 molecules cm^{-3} s^{-1}) and slow reactions (rate < 10 molecules cm^{-3} s^{-1}). The number of blocks (*N*) is 13 and the number of chemical regimes (*M*) is 20. We only consider gridboxes with species concentrations >= 10^6 molecules cm^{-3}; otherwise the gridboxes are shown as white.

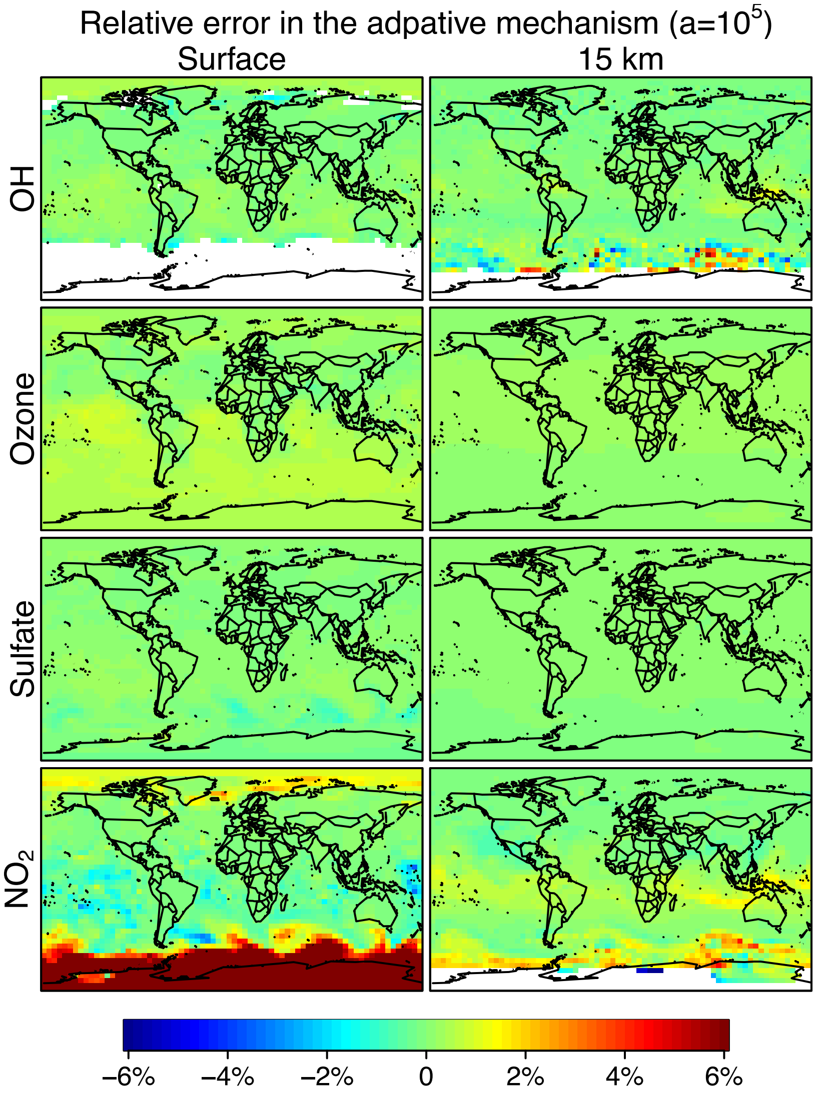

Fig. S6. Same as Fig. S5 except that we only consider gridboxes with species concentrations >= 10^5 molecules cm^{-3}.

It's also not clear why the value is set at 1e7 for NO2 in Figure S6? Is this alternative cut off used in Figure S8? Presumably, this is the reason why the polar regions are excluded from Figure S6 for NO2?

Response. Now we use the same thresholds for all species and we also test lower thresholds ($a=10^5$). See Figure S5 and S6 for more details

Line 158. Can Figure 1 include information about the definition of slow and long-lived for reader clarity otherwise they are having to flick back to find the definitions used?

Response. Done, thanks.

Line 182. "different blocks based on similarity of chemical behaviours using a machine learning clustering method." I think this is described in the supplementary material. This should be brought into the main text of the paper or a reference to where this is described included in the text.

Response. Thanks, we have moved this part back to main text.

Line 185. It would be very useful to have had the contextual information given here much earlier in the paper so the reader can understand what didn't work in the past, what the proposed solution is and how this will be implemented.

Response. We now move this to the last paragraph of the introduction part. Please check there for more details.

Line 209. I'm not sure I understand the comment that iodine reservoirs are inert? Many of them are highly photolabile. This doesn't seem to make sense?

Response. I think the more active iodine reservoirs are in Block 3. The species in Block 4 are relatively inert and they are considered as fast in only <5% of gridboxes.

Line 243. When talking about the error this is the RRME? This should probably be clarified. It would also be useful to discuss the implications of the >1e6 cm-3 limit here on the statistics. Is the NO2 value with the 1e7 cm-3 limit as indicated in the supplementary material?

Response. We use a different threshold NO_2 just for visualization in the old manuscript and we use 10^6 molecules cm^{-3} s^{-1} for all species now. See Figure S5-S6 for more details.

And we have clarified that whether we use the RRMSE error or the relative difference compared to the standard simulation in the text.

Now we say this in text.

Line 721. Computing the RRMSE for all species with concentrations higher than $a=1\times10^5$ molecules cm^{-3} (instead of $1x10^6$ molecules cm^{-3}) shows similar results except that the magnitude of the error is higher because the relative difference is expected to be higher at low species concentrations (Fig. S6, S8, S9).

Line 245. Being able to update the chemical mechanism is an important aspect of maintaining the viability of the model in the long term. If there was a significant change to the mechanism the whole process of running the model without the chemical mechanism splitting would need to be done again? The method outlined here is for "on the fly" updates and relies upon the changes being small and the chemical intuition of the person doing the update. It would be useful to explain that the approach described here is for "patching" the mechanism etc. rather than updating the whole mechanism

Response. Thanks for pointing this out. Now we make it clear how to update our mechanism.

Line 725. Chemical mechanisms in models are frequently updated, including addition and removal of species. Because the species blocks are chemically coherent, our algorithm can accommodate mechanism updates without requiring reconstruction of the submechanisms. New species simply need to be added to the appropriate blocks. Figure S10 shows the diagram for adding new species into the mechanism. Attribution of a species to a given block can be easily determined by its chemical family and the percentage of gridboxes that treat this species as fast when averaged globally. In order not to compromise the computational efficiency, the basic rule is to not mix faster species with slower ones.

It's not clear where new "biogenic VOC" degradation products would go (blocks 7,8 or 9 etc). If an exact mechanism for working out the placement hasn't been found and new species were randomly allocated to 7,8 or 9 (or another mechanism) it would be useful to have that documented.

Response. Thanks. Now we say this in text.

Line 729. In order not to compromise the computational efficiency, the basic rule is to not mix faster species with slower ones. For example, biogenic VOC species and their products could go to Block 8-9 if the percentage of gridboxes that treat them as fast is >1% or Block 10-11 if the percentage is <1%. Our algorithm is robust to misplacements of new species, which may affect computational performance but will not enlarge the error.

Line 272. I don't think that it has been demonstrated that it can be ported to different atmospheric models "easily". Relatively minor changes to the chemical mechanism within one model were shown to be able to be incorporated without re-running the whole tuning procedure but I don't think it can be demonstrated that it is easy to move into a different model (CESM etc).

Response. Thanks, and we have deleted this sentence.

Figure 3. Could be greyed out area showing the number of grid boxes in that category be split into some subcategories: marine boundary layer, continental boundary layer, free troposphere, stratosphere etc to provide some additional information?

Response. Done.

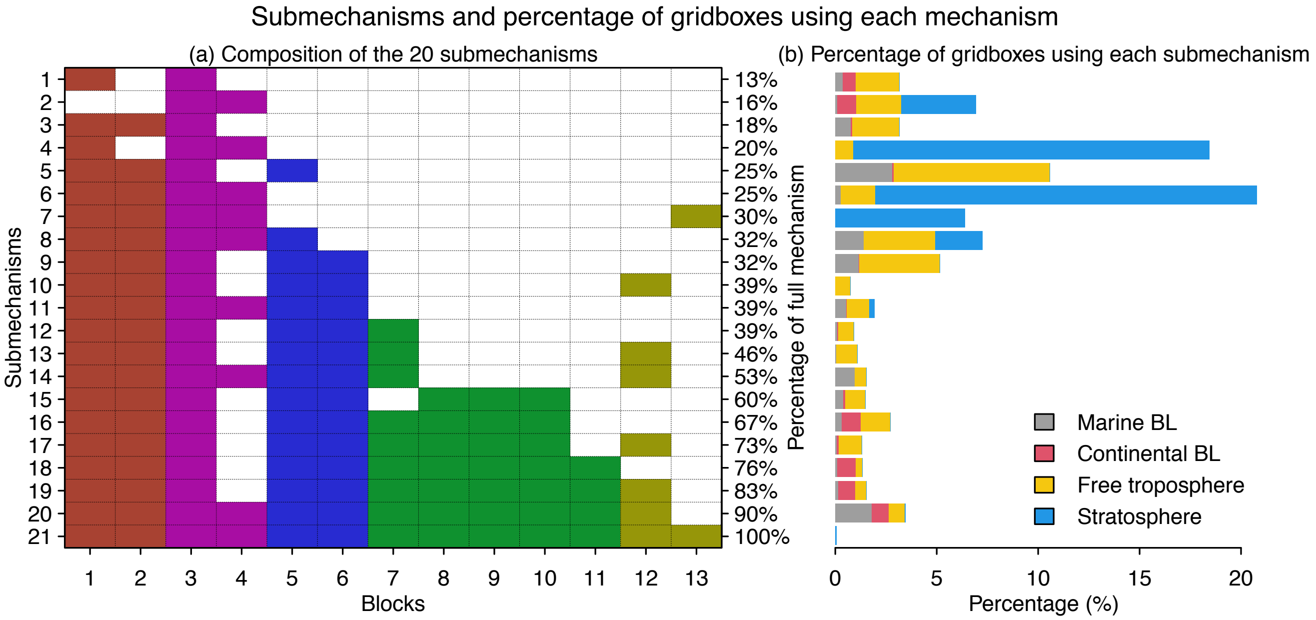

Figure 5. **Submechanisms and percentage of gridboxes using each mechanism**. Panel (a) shows the composition of the 20 submechanisms and full mechanism (the 21st one) as well as the percentage of species from the full mechanism that are treated as fast in each of them. Colors denote species block types as defined in Figure 4. Panel (b) shows the percentage of gridboxes using each submechanism in the marine boundary layer (BL), continental BL, free troposphere, and stratosphere.

Figure 4. This is the performance over what timescale? All of the 1-year timesteps for all grid boxes? It's not clear that the bars indicate the simulation speed up and the symbols represent the accuracy.

Response. They are based on the last day of 3-year simulations and we have made this clear in the figure caption.

Supplementary material. Much if not all of this should be in the main body of the paper.

Response. Thanks. We have moved Figure S1, S2 and S6 to the main text.

Figure S4. Is this mislabelled as "anthropogenic blocks", think it should be biogenic blocks? How is the decision made about fast/slow when there are multiple blocks? This could be explained.

Response. Done, thanks.

Figure S5. I'm not sure that the figure actually shows the "mechanism complexity needed"? It shows the sub-mechanism at each location. I found the %fast labelling slightly confusing as it wasn't that obvious whether the colour scale was describing the Regime or the %fast? Perhaps the %fast could just be removed for simplicity?

Response. Now we say this the caption.

Fig. S3. Chemical mechanism complexity used in the adaptive chemical mechanism in different regions of the atmosphere.

We prefer to keeping the "%fast" because it can better visualize the complexity of each chemical mechanism.

Figure S7. How do these curves look if the <1e6 cm-3 restriction in calculating the RRMS is removed?

Response. The relative error will become very large for species with low concentrations. So here we prefer to show the curve only gridboxes with enough high species concentrations. But we do include a new figure in the supplementary by considering all gridboxes with species concentrations >1e5 cm^{-3}.

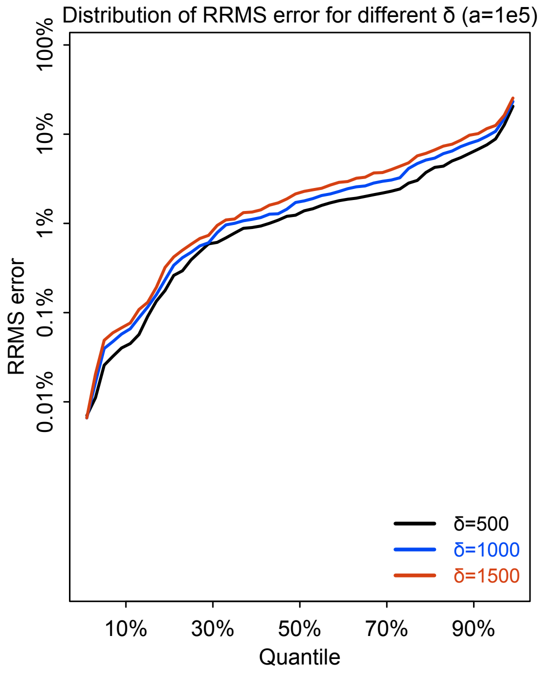

Fig. S9. Same as Fig. S4 except that we consider gridboxes with species concentrations >= 10^5 molecules cm^{-3}.

Figure S8. Can the location where the RRMS has not been calculated due to the 1e6 cm-3 restriction be indicated (grey the area?).

Response. Done, please check Fig. S5 and S6 for more details.

Figure S9. What is the H / L notation indicating?

Response. Now we make this clear in the caption.

Fig. S10. ‘H’ (‘L’) means higher (lower) percentage of gridboxes that consider this species as fast.

Referee 2

This paper present some clearly very well thought out and well executed methodologies that are shown to be effective at minimising computational costs for solving a complex chemical mechanism in the GEOS-Chem model. These tool are likely to have substantial benefits for the air quality and Earth System modelling communities. Unfortunately, the writing of the paper is confusing at times, understandable given the complexity of the subject matter but it does distracts from messages it is trying to convey. With a few small changes to the structure and presentation, this will be an excellent paper well suited for publication in GMD.

Major Comments

The biggest issue I find with the paper is the it goes into the nitty gritty of how it partitions different species/reactions into different categories (sections 2.2-2.5) before it explains the overall design philosophy and what these different categories are used for (in section 3.1). The result is that on first reading, I got very confused reading through section 2 and was only able to make sense of it on second reading. This could be improved substatially if either Section 2.1 were expanded upon to give an overview of the whole design philosophy and what each of the the different subcategories (fast, slow species, unimportant reactions etc.) are going to be used for before they are described in detail, or a new Section 2.2. is added giving an overview of all of the developments. Some of this description is in the first paragraph of Section 3.1. Also section 3.1 repeats a lot of the text currently in section 2.1 about the model description and is inapropriate here, therefore section 3.1 should also be editted to avoid repition. In short, Section 2 should contain all of the descriptions and definitions for the model and the adaptive algorithm for the chemical operator. Section 3 should focus on testing, evaluating and optimising the algorhithm in the 3D model. If these two sections were more clearly defined, the paper would be much easier to follow.

Response. Thanks. We have made the following changes to improve the presentation quality.

1. We now give an overview of the whole design philosophy in the last paragraph of the introduction part (as also suggested by the first referee).
2. We have moved part of the text in the first paragraph of Section 3.1 to Section 2.1 and 2.2.
3. We have moved Text S1 back to Section 2 to give a better description of our method.
4. We have moved Figure S1, S2 and S6 to the main text.

There are also some inconsistancies in language and definitions - I found the use of "slow species" and "slow reactions" (each of which use different threshold definitions) particularly confusing. It would help to be consistent and use "unimportant reactions" for those <10 molecules cm-3 s-1. I have specific comments below to help with this.

Response. Thanks. We now use ‘slow reactions’ throughout the paper because it is more accurate than ‘unimportant reactions’.

In terms of the statistical analysis (Section 2.6) it is important to have a measure of bias as well as error - I would generally be more concerned about a species that has an error of 1% and bias of ~1% (which would imply a consistant error in one direction), than one which has an error of 1% and a bias of ~0% (which would implpy more random error). Looking at Figure S8, the errors in surface Ozone seem to be biased in one direction. That is potentially concerning if the bias is enough to significantly affect tropospheric ozone burden - as ozone is a radiatively important species this could have consequences in an Earth System model. Please also include a normalised measure of bias in Section 2.6.

Response. Thanks for this good point. We now define a relative abundance metric in the paper.

Line 465. A second metric to evaluate our adaptive chemical mechanism is the relative difference of atmospheric abundances for all species compared to the standard simulation. This tests for accumulating bias over long simulation periods.

The tropospheric ozone doesn’t show accumulated bias in one direction. As seen from Figure 7, its bias has some seasonality. And now we have a new figure to show the relative difference in atmospheric abundances of all species.

Line 719. The median relative difference in atmospheric abundances among all species remains at 0% over this 3-year period; the relative differences for key species like ozone, OH, sulfate and NO2 also remain at 0% and are within ±10% for >99% of the other species (Figure S7).

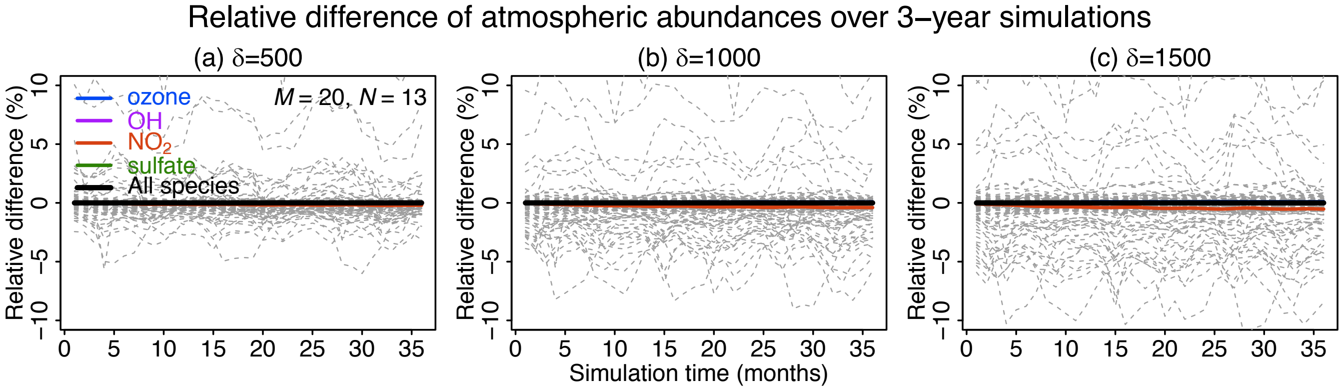

Fig. S7. Relative difference of total atmospheric abundances using adaptive reduced chemistry mechanism algorithm over a three-year GEOS-Chem simulation. Results are shown for the median relative difference across all species in the mechanism and more specifically for ozone, OH, NO_2, and sulfate. The grey dashed lines are for all 228 species.

Finally, there are a number of figures in the supplement which would have been useful in the main paper and I did not understand why they were not in the main paper. I would recomend moving at least figures S1, S2, S5, S6, S7 and S8 into the main paper.

Response. Thanks. After considering both referee’s suggestions, we now move Figure S1, S2 and S6 back to the main text.

Specific comments

ln 26. Ambigous "it" in: "because it exerts strong forcing and feedbacks...". change to "because chemical and aerosol species extert strong forcings and feedbacks..."

Response. Done, thanks.

*ln 49. Change "guide us build" to "guide us **to** build".*

Response. Done, thanks.

Ln 88. here and elsewhere, please be consistent and use the term "unimportant" instead of "slow" to describe those reactions that are removed with fluxes <10 molecules cm-3 s-1, otherwise it is easily confused with the "slow species".

Response. Thanks. But we think ‘slow reactions’ are more accurate than ‘unimportant reactions’, so we use it throughout the paper.

Section 2.3. If I am to understand this correctly, when species are defined as being "slow" and/or "long-lived", they are not "removed" from the mechanism per se*, rather they are solved using the analytical approach instead of the 4-th order Rosenbrock solver. Given this, I think you need to make clear here that when you say "coupled system", you mean all of the species and reactions which are solved using the Rosenbrock solver. All of the species that are "removed" or "excluded" from the coupled system, as you say later, are instead solved using the simple analytical approach.*

Response. Thanks for correcting this. Now we say

Line 203. We solve the fast species in their submechanism using the standard Rosenbrock solver. For the slow or long-lived species, we approximate the evolution of concentrations using an explicit analytical solution that assumes first-order loss (Santillana et al., 2010).

Line 209. **As such, we still update the concentrations of all species but in a more efficient way**.

Line 99. Please can you define "distance" qualitively here before the quantitative definition.

Response. Now we say

Line 212. We construct coherent subsets ('blocks') of the species in the mechanism species based on their linkages through the mechanism reactions. This is done objectively by defining the species distances in the mechanism using graph theory. **In general, two species should have shorter distances if they appear in the same reaction more times and have similar products in the mechanism**.

Line 112. The term "distance" is overloaded in the text to mean multiple different things. You have defined a new term "Euclidian distance" |Di-Dj|, which is a scaler, rather than the previous "distance" Di which is a vector. This Euclidian distance is then modified to cluster species together (I would call this something like "clustered distance"). Please make clear that it is this "clustered distance" that is stored in the matrix and used to calculate the cost function Z (eq 4).

Response. Thanks, we prefer to name it as 'modified distances'. Now we say this in the text.

Line 298. We store these modified distances of all pairs in a 228x228 matrix.

*Line 113. When applying the 50% factor, is this applied iteratively? i.e. if two species are each others closest pair, will their Euclidian distance be reduced by a factor 0.5*0.5=0.25, as iterating through each species, or is the 50% factor only applied once? Please be clear which approach you are using.*

Response. This scaling will be applied only once.

Line 297. Second, for each species *i*, we will decrease its distance with the 5 species that have highest similarity with it by 50% and **this scaling is applied once for one pair of species**.

The 50% factor and 5 closest neighbours both seem quite arbitrary, but I think I can see how this approach will cause clustering of species into close families. What was the reason for the use of these two values, and did you test other values?

Response. We indeed tested other numbers and the results are in general consistent.

Line 297. Using 10 highest-similarity species instead of 5 and decreasing distances by 30% or 70% does not change the results.

Line 135. Would benefit from Figure S2 being in the main text here. Looking at Figure S2, I think you can make a fair cost-benefit argument that M=20, N=13 is well optimised as it is on the bottom-left of the 30% contour. It is clear from the contour lines that you get diminishing returns of the fraction of species if you were to increase M and N, hence selecting on the 30% contour seems reasonable. By selecting the bottom-left part of the 30% contour, you are minising both N and M.

Response. Thanks, we now move Figure S2 back to the main text. Now we say this in text.

Line 528. In order to make the code manageable, we choose to use $M = 20$ resulting in an optimal value $N = 13$ at which only 30% of the species need to be treated as fast in the global tropospheric and stratospheric domain (Figure 3).

Basically, you can be more rigorous in the justification for why you used M=20, N=13 than simply saying you "choose" them.
Response. Now we say this in the main text.
Line 528. In order to make the code manageable, we choose to use $M = 20$ and an optimal value $N = 13$ at which only 30% of the species need to be treated as fast in the global tropospheric and stratospheric domain.

Line 137-139. Please move this text to the top of section 2.5, as this describes the training set used to derive the submechanisms and blocks.
Response. Done, thanks.

Line 196. Please again change "slow reactions" to "unimportant reactions". Please clarrify - are the unimportant reactions removed from each of the submechanisms using the original training data, or are they removed on the fly depending on the concentrations of species in each grid cell at each timestep?
Response. Thanks. We prefer to name it 'slow reactions' and we have made it clear that the slow reactions removed in each submechanism are pre-defined and the same.
Line 524. The slow reactions removed in each submechanism are pre-defined (see Section 2.2 for more details).

Removing the reactions from each of the submechanisms in advance seems like the more efficient approach to me. However, there is a risk that the approach becomes inconsistant if used in different time periods with different chemical conditions to the training data. For example, there will be risk in using submechanisms derived with present day training data in preindustrial conditions. This is relevant for application in Earth System models.
Response. Thanks for raising this point and we now discuss it in the text.
Line 200. Here we remove these slow reactions in each submechanism based on present-day atmospheric chemistry environment and it should be re-evaluated if this method is applied in other periods (e.g. pre-industrial times) when the atmospheric conditions could be very different from our present-day one.

line 229. The full mechanism is by definition not a submechanism. Say that it is the 21st "chemical regime".
Response. Done, thanks.

Line 236. slow reactions -> unimportant reactions
Response. Now we use 'slow reactions' throughout this paper.

Figure 1. The line in panel a shows the fraction of species removed from the coupled mechanism. *Call the slow reactions unimportant reactions.*
Response. Now we use 'slow reactions' throughout this paper.

Figure 3. There are 21 chemical regimes, made up of M=20 submechanisms plus the whole mechanism.
Response. Now we use 'slow reactions' throughout this paper.

Figure 4. Number of chemical regimes is 21 (M+1), not 20.
Response. Done.

Supplementary material

line 12. think Z_2 *should be called* f *or* f(M,N) *instead. Use a different term to* D *for the gridcells, because* D *is already used for the distance vectors.*
Response. Fixed and we now use P instead of D.

line 18-20. Unclear how this algorithm works, don't know what is meant by "temperature" here.
Response. Now we add a reference to describe how this temperature parameter works.

figure S5. presumably chemical regimes should go up to 21, not 20? Also, its the 21st regime that has 100% of the mechanism, regime 20 has 90% according to Figure 3.
Response. We have updated this figure. Please check Figure S3 for more details.

Figure S6. Please include panels showing biases for the key species. If any have consistant/growing biases, that should be commented on (especially concerning for ozone). What value of δ did you use here?
Response. Thanks for this good point. We now define a relative abundance metric in the paper.

Line 465. A second metric to evaluate our adaptive chemical mechanism is the relative difference of atmospheric abundances for all species compared to the standard simulation. This tests for accumulating bias over long simulation periods.

And now we have a new figure to show the relative difference in atmospheric abundances of all species.
Line 719. The median relative difference in atmospheric abundances among all species remains at 0% over this 3-year period; the relative differences for key species like ozone, OH, sulfate and NO2 also remain at 0% and are within ±10% for >99% of the other species (Figure S7).

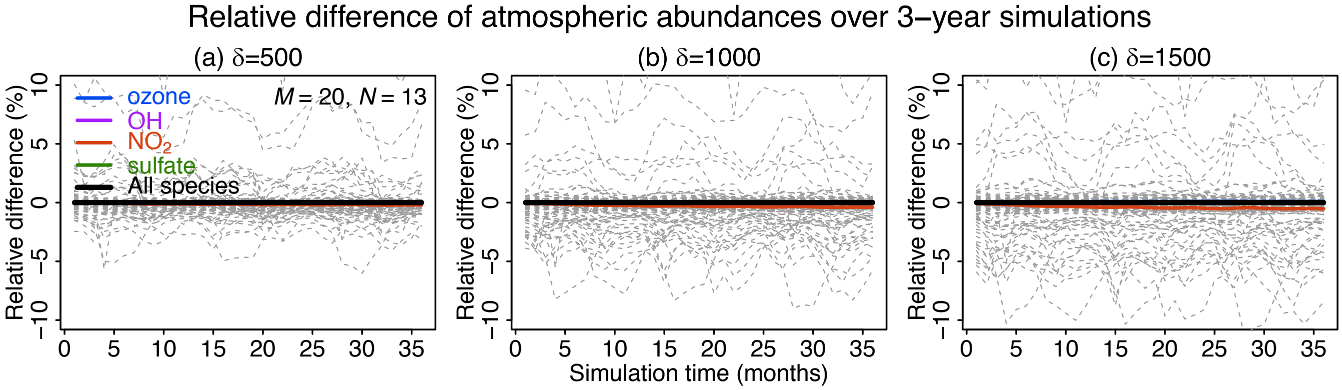

Fig. S7. Relative difference of total atmospheric abundances using adaptive reduced chemistry mechanism algorithm over a three-year GEOS-Chem simulation. Results are shown for the median relative difference across all species in the mechanism and more specifically for ozone, OH, NO_2, and sulfate. The grey dashed lines are for all 228 species.