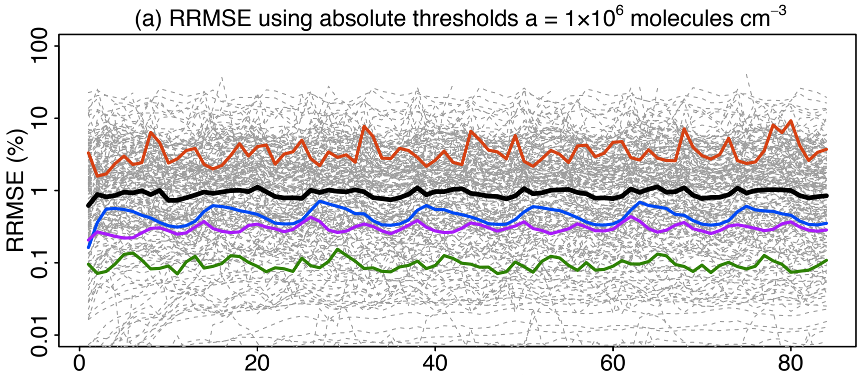
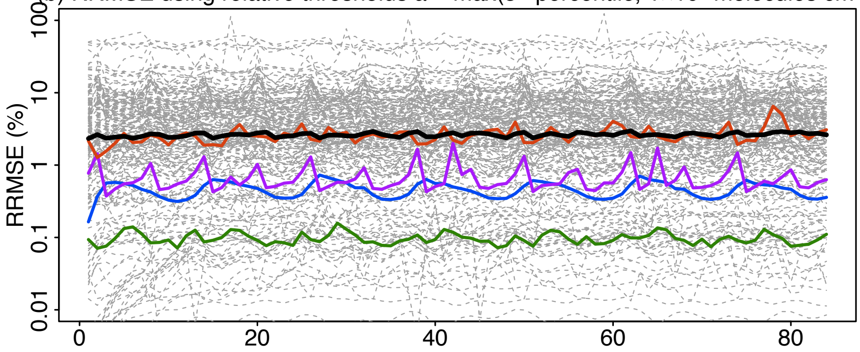
Accuracy over 8-year simulations using different error definitions (δ=1500)

(a) RRMSE using absolute thresholds a = 1×10^6 molecules cm^{-3}



(b) RRMSE using relative thresholds a = max(5th percentile, 1×10^4 molecules cm^{-3})



(c) Relative difference of global atmospheric masses

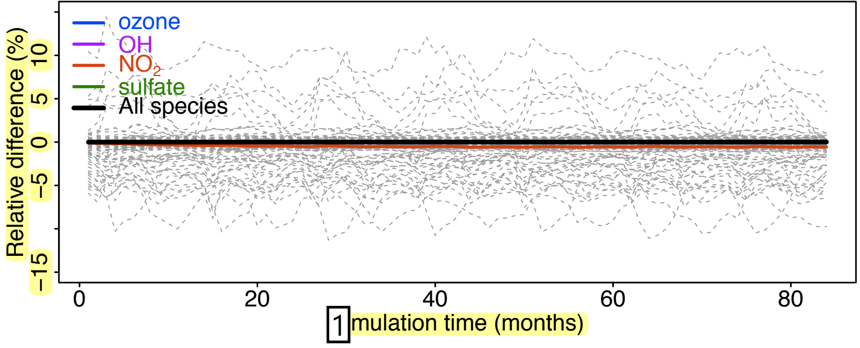


Figure 8. Accuracy of the adaptive reduced chemistry mechanism algorithm over an 8-year GEOS-Chem simulation using a threshold δ of 1500 molecules cm^{-3} s^{-1} to separate fast and slow species. (a) Same as Figure 7c but for the 8-year simulation. Here we use an absolute threshold of species concentration (*a* = 1×10^6 molecules cm^{-3}) to define the RRMSE (See Section 2.5). (b) Same as Figure 8a but using a relative threshold (*a* = max (5th percentile, 1×10^4 molecules cm^{-3})) to define the RRMSE (See Section 2.5). (c) Relative difference of global atmospheric masses for each species. Dashed lines show results for all 228 species in the mechanism. 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.

Major comments

One of my main concerns is the superficial and nature of discussion you offer, often having little physical context. Such instance are:

Response to Reviewer 1 comment on Line 87: Your response still does not offer any context regarding the

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

selection criteria. Why 10 molecules and not 100 or 1000? How much larger is “slightly larger than 0 molecules cm-3 s-1” you mention? What “because of numerical precisions” means here at all? I cannot accept this response as scientific.

Response. We examined different thresholds in some short-term experiments in the early stage of this project. We found that the error quickly increases if the threshold is above 10 molecules cm^{-3} s^{-1}. You are right that our previous discussion was superficial, and we provide more details here.

Line 143. In each submechanism, if a reaction is slower than 10 molecules cm^{-3} 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 δ =500-1500 molecules cm^{-3} s^{-1} for the species to be included in the Jacobian matrix.**

Similarly,
Reviewer 1 on Line 243, Reviewer 2 on Line 297 (and generally comments/replies regarding the choice of thresholds): You select a new threshold without explicating why – an approach “we use now lower value and it does (or does not) change things” is vaguely related to the nature of the selection, it does not help understanding which factors are important or not. Where are the important thresholds for distances? Why testing for 10 species is now sufficient as compared to 5 species before? It appears to be an arbitrary choice.

Response. Thanks for this good point. This was done with repeated tests of different thresholds and then we determine if the optimized chemical blocks are consistent with our chemical intuition. We rewrite it here.

Line 208. 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. We found in tests that using 10 highest-similarity species instead of 5 and decreasing distances by 30%-70% instead of 50% did not significantly change the results. The logic is that the number of species with similar chemical characteristics is usually around 5-10 and decreasing the distances among them by 30-70% can increase the probability of these species to be in the same chemical blocks after the optimization process.

Furthermore,
Reviewer 1 on Figure S6, Reviewer 2 major comment: Showing the average difference for the species is alike reporting the average body temperature for all patients in the hospital. Not only that does not help understanding the biases for the species of importance, how would you judge/set the input (e.g. weight) of every species to this average?
Response. Thanks. Now we show the error of all species in Figure 6, 8, S4, S7 and S8. Please check them for more details.

A large bias in simulating N2O in troposphere (where it is mostly inert) will hardly affect CH4 lifetime there, whilst a small bias in OH will do so (especially on the long-term, I will return to that below).
Response. We don’t simulate the evolution of long-lived species in our standard chemistry model, following standard practice of simulations focused on tropospheric chemistry – we now clarify that this is our focus. We have also shown that the error is stable in an 8-year simulation.

Page: 3

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!

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).

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.

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

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?

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.

Line 113. In the simulations presented here, methane, N_2O, and other long-lived halocarbons have fixed concentrations in surface air (Eastham et al., 2014; Murray, 2016) so that the longest resolved chemical modes are less than a year.

We also removed all text claiming that our algorithm can be applied in earth system models because we haven't tested it. We have also changed the title,
Title. A machine learning-guided adaptive algorithm to reduce the computational cost of integrating kinetics in global atmospheric chemistry models: application to GEOS-Chem versions 12.0.0 and 12.9.1

Fig. S7 and its vertical scale are deceptive – behind larger relative differences for some species (not indicated which) one does cannot check for growing biases (I spot one in the right panel for NO2 nonetheless)
Response. Thanks for pointing this out. Now we show all errors (in log scale) of all species in figures. Please check updated Fig. 7, 8 and S7.

The stated errors that "also remain at 0%" make no sense from the numerical point of view – does this mean that you compute numerically identical result (read copy it)? So how large are deviations really? Do you mean below 0.1%? The metric "+-10% for >99% of the other species" is also not reasonable – what if the remaining 1% of species has other important (in addition to key ones) species, e.g. reactive compounds?
Response. Sorry, our old narrative wasn't precise, and we have removed all these sentences. We now show all errors in Figure 7, 8 and the distribution of errors in Figure S4.
Line 376. [1]The relative differences of global atmospheric masses are within 10% for >99% species (Figure 8c) and show no sign of increasing trends. The relative difference for NO_2 increases slightly from 0% to 0.4% in the first 30 months and then stabilizes at 0.4%. Key species like OH, ozone and sulfate have a relative difference smaller than 0.01% throughout this simulation period.

*I give a few hints to help adequately re-address these comments. Classical ODE integrators employ absolute and relative tolerances for estimating the errors about the solution that is obtained. This is done to constrain errors when relative scales of changes for the spectrum of variables integrated vary many orders of magnitude, and is shown to be valid using through the strict mathematical analysis (see, e.g., [1], Chapter on ODE integration). Whilst I do not require such here, **a minimum sound mathematical analysis of the order of the errors in your method should be present in the revised version**.*
Response. Thanks, but we are afraid that a simplified ODE assumption may not be applied here. We mentioned in text that the Rosenbrock chemical solver is a 4th-order implicit method and the linear approximation for slow species is a 1st-order explicit method. But we feel it is difficult to further analyze the order of the errors mathematically because of operator splitting with other processes including deposition and transport.

A shortcoming of your current method, analysis and presentation of the results is also in using only absolute thresholds. Ozone and OH differ by orders of magnitude, yet you use the same absolute threshold values for these two species in the analysis. Presenting errors at the surface and at 15 km altitude (Figs. S6-7) has little sense for OH, whose tropospheric sink is 80% in the free troposphere. Which atmospheric abundances are compared in Fig. S7, are these gridboxes or domain averages? These are only a few questions, I could go on… but they already show how inconclusive and not-well-thought-through are the analysis and discussion you present.
Response. Thanks for these good points. Here we report the error using both absolute and relative thresholds. See Figure 8 for more details. And we have a new paragraph discussing this.

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.

Line 272.

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

Absolute threshold $a = 1\times10^6$ molecules cm^{-3}

Relative threshold a = max (5th percentile, 1×10^4 molecules cm^{-3})

Line 372. Our algorithm also shows no sign of increasing errors over an 8-year simulation. Figure 8a-b displays the RRMSE over this period by taking account of all gridboxes with species concentrations above an absolute threshold of 1×10^6 molecules cm^{-3} or a relative threshold of the 5th percentile of species concentrations across all gridboxes (see Section 2.5 for more details). In both cases, the median and the maximum RRMSE remain constant over this simulation period. The relative differences of global atmospheric masses are within 10% for >99% species (Figure 8c) and show no sign of increasing trends. The relative difference for NO_2 increases slightly from 0% to 0.4% in the first 30 months and then stabilizes at 0.4%. Key species like OH, ozone and sulfate have a relative difference smaller than 0.01% throughout this simulation period.

To recap, you need to substantially improve the analysis of errors, presentation and discussion in order to confirm that your method is accurate and convergent (w.r.t. the reference) and support your statements, especially the ones in the conclusion section. Please add a definition or explicate what exactly you imply by "chemical coherence" (better in the introduction).

Response. Thanks for pointing this out. We use chemical initiation instead.

Line 69. A major development here is to enforce that chemically connected species be grouped in the same blocks, so that the blocks are consistent with chemical intuition and can be logically modified and extended as the mechanism changes.

As of now, you show no evidence of "no error growth over multi-year global simulations" because many atmospheric compounds have lifetimes longer than 3 years (read they need to equilibrate to be consistently compared to the reference). Fig. S7 adds dubiety here, too. I suggest that you simulate at least a 30 years long period in order to corroborate that no substantial errors in CH_4, NOx and other reactive N reservoirs and hence O_3 are building up in the optimised mechanism.

Response. A 30-year simulation is not necessary for tropospheric oxidant/aerosol chemistry, which we now clarify as our focus. Here we conduct an 8-year simulation and we find the error is stable over this time period. Please check Figure 8 for more details.

Line 372. Our algorithm also shows no sign of increasing errors over an 8-year simulation. Figure 8a-b displays the RRMSE over this period by taking account of all gridboxes with species concentrations above an absolute threshold of 1×10^6 molecules cm^{-3} or a relative threshold of the 5th percentile of species concentrations across all gridboxes (see Section 2.5 for more details). In both cases, the median and the maximum RRMSE remain constant over this simulation period.

Note that an error of a few percent in oxidising radicals may lead to largely diverging results in a few decades. For instance, using a simple spreadsheet integration you may infer that a longer by 1% (as compared to assumed average of 9 years) atmospheric CH_4 lifetime would yield CH_4 abundance larger by already 3% (above 50 ppb, or growth rate in last 5 years!) in just recent 30 years, assuming that emissions are the same in both cases. So can your optimised method really simulate 3 decades of CH_4 evolution with an accuracy of 1%?

Page: 5

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 <=1 molecule 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/cm3, 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 of 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.

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

A typo? What is chemical initiation???

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

See the comment regarding the intuition above

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 CH4, a rather large number of trace gases are being oxidised by the HOx and NOx. An 8-yr simulation, I tend to agree here, will likely be sufficient to have reactive N-bearing reservoirs equilibrated.

Response. We don’t simulate the evolution of long-lived species in our standard chemistry model. We now make it clear in text.

Line 114. In the simulations presented here, methane, N_2O, and other long-lived halocarbons have fixed concentrations in surface air (Eastham et al., 2014; Murray, 2016) so that the longest resolved chemical modes are less than a year.

Minor comments

Reviewer 1 on Line 212: What are the objective criteria, if you state “This is done objectively here…”.
Response. We removed ‘objectively’ and now explain in detail how we construct coherent blocks in the following sentences.

How can “two species … appear in the same reaction more times”? Present an example.
Response. “appear in the same reaction” was confusing. We have replaced by “appear together in multiple reactions’. For example, NO and NO2, OH and HO2 may appear in a lot of reactions related to VOC oxidation.
Line 157. In general, two species should have shorter distances if they appear together in multiple reactions (e.g. NO and NO_2, HO and HO_2) or have similar products in the mechanism.

Reviewer 1 on Line 285: Specify exactly which index denotes product and which denotes educt, i.e. “… include both educt i and product j species”.
Response. Sorry, we didn’t make it clear. Either species i or j could be the reactant.
Line 164. Where T_(i, j) is the number of reactions that include both species i and j (with one as reactant and the other as product, **either *i* or *j* could be the reactant**).

Reviewer 2 on Line 465: Copy-pasted responses do not respect reviewers and editors time invested in this manuscript. I consider it a mauvais ton, especially as this happens not the first time in this submission.
Response. Sorry, I did this because reviewers have similar questions. It is not a good manner.

Reviewer 1 on Line 209: Do you “think” that there is more active iodine in the Block 3, or do you “believe” in that, or do you postulate that? Can you actually look into the model results to confirm that?
Response. We indeed looked into the model results. Please see the details here.

Percentage of gridboxes in the global tropospheric+stratospheric domain that treat this species as fast.
Block 3: BrNO2(16.5%), IONO(9.5%), OIO(28.9%), ClOO(78.6%), OClO(20.6%), BrCl(25.5%), HOI(35.1%), Br2(28.2%), IONO2(24.7%), BrNO3(40.2%), I(51.1%), IO(50.7%), HOBr(48.6%), HOCl(34.4%), ClNO3(30.8%), BrO(51.3%), HCl(46%), HBr(26.9%), Cl(79%), Br(51.4%), ClO(49.1%)

Block 4: AERI(0.2%), ISALA(0.1%), ISALC(0.2%), I2O4(1.2%), I2O2(4.7%), I2O3(0.9%), IBr(0.7%), INO(0.9%), HI(1.2%), ICl(4.4%), Cl2O2(2.2%), ClNO2(2%), BrSALC(1.9%), BrSALA(2.2%), I2(2.3%), Cl2(2.5%)

Page: 6

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

Sorry, you still haven't made if 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?

2 Method description

Here we describe the adaptive method as applied in the GEOS-Chem global model, although it is applicable to any model. We begin with a brief description of the model as relevant to the presentation.

2.1 GEOS-Chem model

We use the GEOS-Chem version 12.0.0 global 3-D model for tropospheric and stratospheric chemistry (https://doi.org/10.5281/zenodo.1343547) with 12 CPUs in a shared-memory Open Message Passing (Open-MP) parallel environment. For development and testing purposes, we choose a horizontal resolution of 4°×5° and 72 pressure levels extending from surface to 0.01 hPa and drive the model with NASA MERRA2 assimilated meteorological data. The full mechanism for oxidant-aerosol chemistry in the model has 228 species and 724 reactions, including coupled gas-phase and aerosol chemistry for the troposphere and stratosphere (Sherwen et al., 2016; Eastham et al., 2014). The chemical operator uses a 4th-order Rosenbrock implicit method, implemented through the Kinetic Pre-Processor (KPP) (Sander and Sandu, 1996), to solve for the chemical evolution of species concentrations, involving iterative calculations and inversion of the Jacobian matrix that stores the sensitivity of species tendencies (production minus loss rates) to concentrations. In the simulations presented here, methane, N_2O, and other long-lived halocarbons have fixed concentrations in surface air (Eastham et al., 2014; Murray, 2016) so that the longest resolved chemical modes are less than a year.

As part of this study, we test the portability of our adaptive algorithm by moving it from GEOS-Chem version 12.0.0 to GEOS-Chem version 12.9.1 (https://doi.org/10.5281/zenodo.3950473). This new version [1]f has a thoroughly updated mechanism of 262 species and 850 reactions, including improved organic nitrate chemistry **(**Fisher et al., 2018**)**, isoprene chemistry **(**Bates and Jacob, 2019**)**, and halogen chemistry **(**Wang et al., 2019**)**. From version 12.0.0 to 12.9.1, we need to remove 49 old species and add 83 new species.

2.2 Separation of fast and slow species and reactions

Coupling between species in the Rosenbrock chemical solver is needed only for species with sufficiently fast production or loss rates (fast species), and similarly reactions need to be considered only if they are sufficiently fast. We separate the atmospheric species as fast or slow based on their production and loss rates relative to a threshold δ: fast if either $P_i(\boldsymbol{n}) \geq \delta$ or $L_i(\boldsymbol{n}) \geq \delta$, slow if $P_i(\boldsymbol{n}) < \delta$ and $L_i(\boldsymbol{n}) < \delta$ (P_i and L_i refer to the production and loss rates of the i^{th} species, δ is a threshold, and $\boldsymbol{n}$ is a vector of concentrations of all species). To get a sense of a relevant threshold, consider the hydroxyl radical (OH) which is central to driving oxidant-aerosol chemistry. OH has a daytime concentration of the order of 10^6 molecules cm^{-3} and a lifetime of 1s, so its production and loss rates are of the order of 10^6 molecules cm^{-3} s^{-1}. Species with production and loss rates smaller than 10^2-10^3 molecules cm^{-3} s^{-1} are unlikely to have fast influence on other species in the mechanism

Summary of Comments on Microsoft Word - 1_Manuscript_clean.docx

Page: 3

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

The explicit solution by Eq. 3 does not strictly conserve mass (Shen et al., 2020), and Shen et al. (2020) previously found that this is a problem for halogen species in the stratosphere due to the long lifetime of the collective halogen families and the alternance of the component species as fast and slow over day and night. To avoid this problem, we treat all 37 reactive inorganic halogen species as fast in the stratosphere. Thus, among the *N* blocks, 2 are allocated to the reactive inorganic halogen species, and *N*-2 are allocated to the other species. The transitions of species between the 2 inorganic halogens blocks and the other *N*-2 blocks are not accepted in the optimization process.

2.5 Error analysis

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

$$\boldsymbol{RRMSE_i} = \sqrt{\frac{\mathbf{1}}{\boldsymbol{Q_i}}\sum_{\boldsymbol{j=1}}^{\boldsymbol{Q_i}}\left(\frac{\boldsymbol{n_{i,j}^{\text{reduced}}} - \boldsymbol{n_{i,j}^{\text{full}}}}{\boldsymbol{n_{i,j}^{\text{full}}}}\right)^{\mathbf{2}}} \quad (\textbf{where } n_{i,j}^{\text{full}} \geq \boldsymbol{a}) \qquad (\mathbf{7})$$

Absolute threshold $a = 1\times10^6$ molecules cm^{-3}

Relative threshold a = max (5th percentile, 1×10^4 molecules cm^{-3})

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, the sum is over the gridboxes where $n_{i,j}^{\text{full}}$ is greater than a threshold *a*, and Q_i is the number of such gridboxes. Here we use an absolute threshold of $a = 1\times10^6$ molecules cm^{-3} as in Eller et al. (2009) and Santillana et al. (2010), and will also show results with $a = 1\times10^5$ molecules cm^{-3} and with a relative threshold of *a* as the 5th percentile of concentrations of the species over all gridboxes.

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

3 The adaptive algorithm for the chemical operator

3.1 Potential for local simplifications of atmospheric chemistry mechanisms

Figure 2 displays the potential for local simplification of the full mechanism over the global domain, based on local chemical production and loss rates for the 228 species simulated by GEOS-Chem. Using a threshold δ of 500 molecules $cm^{-3}s^{-1}$ for production and loss rates to define the fast and slow species (see Section 2.2 for the selection of this threshold), a given

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.

Keller, C. A. and Evans, M. J.: Application of random forest regression to the calculation of gas-phase chemistry within the GEOS-Chem chemistry model v10, Geosci. Model Dev., 12, 1209–1225, https://doi.org/10.5194/gmd-12-1209-2019, 2019.

Kirkpatrick, S., Gelatt, C.D. and Vecchi, M.P.: Optimization by simulated annealing, Science, 220 (4598), 671-680, 1983.

Murray, L.: Lightning NOx and Impacts on Air Quality, Current Pollution Reports, 2, 115–133, https://doi.org/10.1007/s40726-016-0031-7, 2016.

National Resarch Council: A National Strategy for Advancing Climate Modeling, National Academies Press, Washington DC, 2012.

Rastigeyev, Y., Brenner, M.P., Jacob, D.J.: Spatial reduction algorithm for atmospheric chemical transport models. Proc. Natl. Acad. Sci. USA, 104, 13875-13880, 2007.

Sportisse, B., Djouad, R.: Reduction of chemical kinetics in air pollution modeling, J. Comput. Phys., 164, 354-376, 2000.

andu, A., Verwer, J. G., Blom, J. G., Spee, E. J., Carmichael, G. R., and Potra, F. A.: Benchmarking stiff ode solvers for atmospheric chemistry problems II: Rosenbrock solvers, Atmos. Environ., 31, 3459–3472, 1997

Santillana, M., Le Sager, P., Jacob, D.J., and Brenner, M.P.: An adaptive reduction algorithm for efficient chemical calculations in global atmospheric chemistry models, Atmos. Environ., 44(35), 4426-4431, 2010.

Shen, L., Jacob, D. J., Santillana, M., Wang, X., and Chen, W.: An adaptive method for speeding up the numerical integration of chemical mechanisms in atmospheric chemistry models: application to GEOS-Chem version 12.0.0, Geosci. Model Dev., 13, 2475–2486, https://doi.org/10.5194/gmd-13-2475-2020, 2020.

Sherwen, T., Schmidt, J. A., Evans, M. J., Carpenter, L. J., Großmann, K., Eastham, S. D., Jacob, D. J., Dix, B., Koenig, T. K., Sinreich, R., Ortega, I., Volkamer, R., Saiz-Lopez, A., Prados-Roman, C., Mahajan, A. S., and Ordóñez, C.: Global impacts of tropospheric halogens (Cl, Br, I) on oxidants and composition in GEOS-Chem, Atmos. Chem. Phys., 16, 12239–12271, http://sci-hub.tw/10.5194/acp-16-12239-2016, 2016.

Wang, X., Jacob, D. J., Eastham, S. D., Sulprizio, M. P., Zhu, L., Chen, Q., Alexander, B., Sherwen, T., Evans, M. J., Lee, B. H., Haskins, J. D., Lopez-Hilfiker, F. D., Thornton, J. A., Huey, G. L., and Liao, H.: The role of chlorine in global tropospheric chemistry, Atmos. Chem. Phys., 19, 3981–4003, https://doi.org/10.5194/acp-19-3981-2019, 2019.

Young T. R. and Boris J. P.: A numerical technique for solving stiff ordinary differential equations associated with the chemical kinetics of reactive flow problems, J. Phys. Chem., 81, 2424–2427, 1977.

Page: 13

Number: 1 Author: gromov Subject: Highlight Date: 09/01/2022 14:58:00

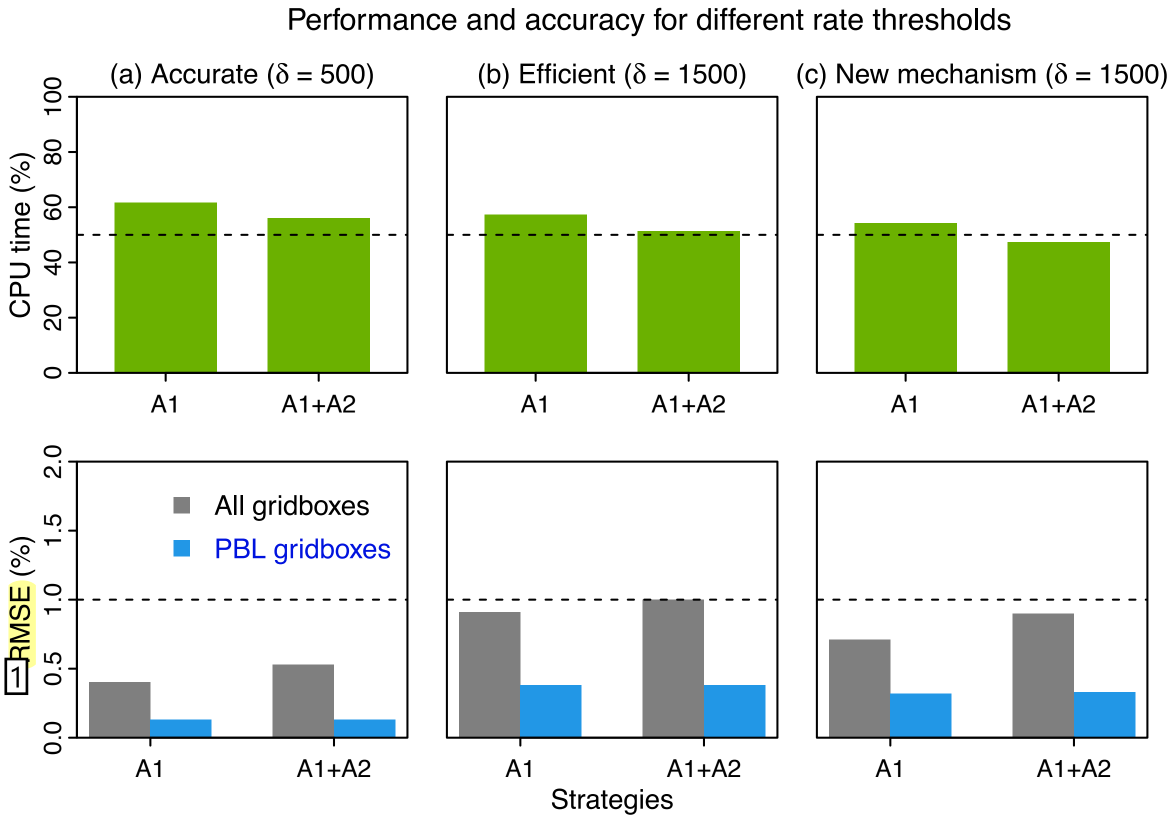


Figure 6. **Performance and accuracy of the adaptive chemical mechanism**. We test the performance of the adaptive method by (A1) removing slow species (P_i or $L_i > \delta$) and (A2) removing slow reactions (reaction rate < 10 molecules cm 2 $^{-1}$). Results are shown on the last day of 3-year simulations. 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. For (a) and (b), we use e 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).

Page: 20

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?

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

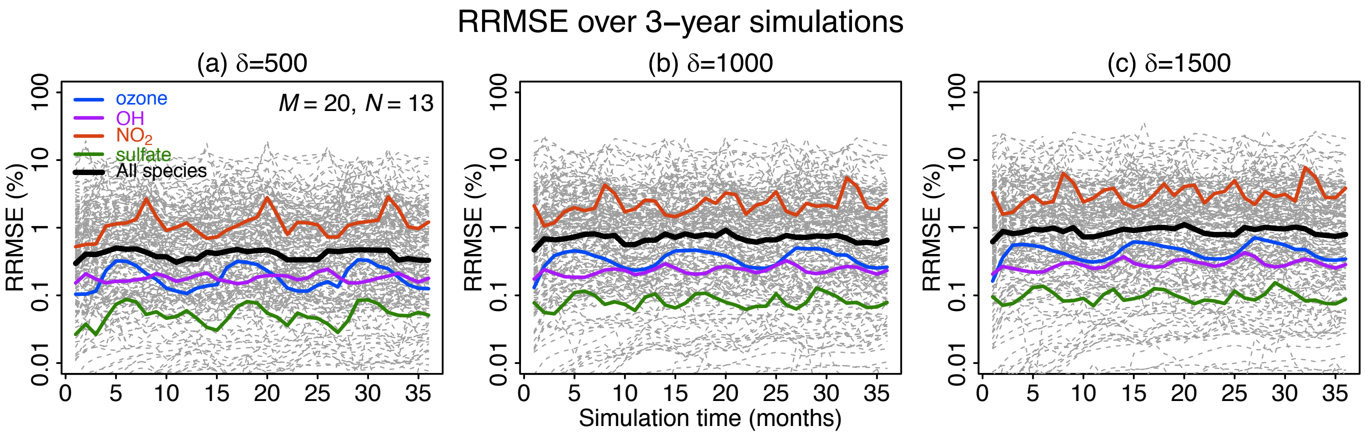


Figure 7. Accuracy of the adaptive reduced chemistry mechanism algorithm over a three-year GEOS-Chem simulation (see text). The accuracy is measured by the Relative Root Mean Square Error (RRMSE, see 1q. 5) error on simulated concentrations relative to a simulation including the full chemical mechanism. Dashed lines show results for all 228 species in the full mechanism, and solid lines show the median RRMSE as well as the RRMSEs for ozone, OH, NO_2, and sulfate. The three panels show the effect of using different thresholds δ ranging from 500 to 1500 molecules cm^{-3} s^{-1} to separate fast and slow species.

Page: 21

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

Eq. 7 ?