Dr. Sebastian D. Eastham
Principal Research Scientist
Joint Program on the Science and Policy of Global Change
Associate Director of the Lab. for Aviation and the Env.

Massachusetts Institute of Technology
77 Mass. Ave. Office E19-439F, Cambridge MA 02139, USA
http://globalchange.mit.edu | http://lae.mit.edu
seastham@mit.edu | (617) 253-2170

Editorial Office

Geoscientific Model Development

October 23rd, 2022

Dear Dr. O'Connor,

Re: submission of "Implementation and evaluation of the GEOS-Chem chemistry module version 13.1.2 within the Community Earth System Model v2.1" to *Geoscientific Model Development*

Thank you for arranging this review of our work. We have used the time since we received our reviews to address the reviewers' comments, and the manuscript has been improved as a consequence. We believe that we have now addressed all comments where possible.

Please find below our responses (in **bold**) to each of the review comments (in *italics*) and our revised manuscript enclosed.

Referee #1

The resubmitted draft describing the new implementation of GEOS-Chem within CESM is much improved over the initial submission. I recommend it for publication after my minor comments and technical edits listed below are considered. The line numbers refer to the marked-up version that followed after the authors' comments to the reviewers and not line numbers in the submitted draft as I did not have the new draft when I started my review and was unfortunately without internet at that time.

We thank the reviewer for their time and effort and for their helpful comments. We have done our utmost to respond to each comment in turn (below).

Line 55: The authors need to check references throughout the manuscript are in alphabetical or chronological order. Often there is a mix. For example, "Community Earth System Model (CESM) (Hurrell et al., 2013; Tilmes et al., 2016; Lamarque et al., 2012; Emmons et al., 2020)"

References have now been changed to chronological ordering throughout.

Line 64 : extra space at the end of sentence before the period.

This has now been corrected.

Line 82: GEOS FP does not have a hyphen (see https://gmao.gsfc.nasa.gov/GMAO_products/NRT_products.php) and this should be corrected throughout the manuscript.

GEOS FP is now written without a hyphen throughout the manuscript.

Line 83-85: GEOS FP and MERRA-2 should have references to Lucchesi 2018 and Gelaro et al. 2017 (https://doi.org/10.1175/JCLI-D-16-0758.1), respectively.

These references have now been added.

Also, is there an extra space after output and before provides? ("output provides data")

The extra space has been removed.

Line 87: Should ESM be redefined, or is it assumed folks know it from the CESM definition?

ESM is now defined as Earth system model on first use (line 48).

Line 99: The Keller et al. 2017 reference is incomplete. My assumption is it is redundant to the Hu et al reference so the Keller et al. 2017 should be removed and only the Keller et al 2021 reference should be included.

The reference in question now includes only Keller et al. (2021).

Line 101: Goddard Earth Observing System (GEOS) (Bey et al., 2001; Eastham et al., 2018) is redundant (to text above as GEOS already defined) and confusing (as readers could mistake Bey and Eastham references for GEOS meteorology references). I suggest removing references and "Goddard Earth Observing System".

The references have been removed, as has the re-definition of GEOS.

Line 105: CESM already defined and references provided.

The re-definition and references have been removed.

Line 107: "observed meteorology including from GEOS", I suggest changing observed to analyzed, unless you can actually nudge towards observations.

This has been changed to "analyzed" (line 107).

Line 119: Earth system models is written out instead of using "ESM"

We now write "ESM" (line 118).

Line 141: "Sections 4 and 5 present a one-year simulation (not including spin up)". I know I suggested to make it clear that there was a spin-up vs the year used for evaluation in my initial review but I think this is confusing as it is written, such that one might wonder if no spin-up was included. Instead, maybe something like "Results from the one-year simulation (following the appropriate spin-up) performed for each model configuration 1) CESM with GEOS-Chem, 2) CESM with CAM-chem, and 3) the stand-alone GEOS-Chem CTM are presented with includes model intercomparison in Section 4 and evaluation against surface and satellite measurements in Section 5. Final discussion and conclusions are in Section 6.". Note, I restructured the sentence and included a final sentence for Section 6.

We agree and have adopted the proposed changes in full (lines 139-142).

Line 182: Should HEMCO have a Keller et al. reference? Line 301 redefines HEMCO and provides the Keller et al. 2014 reference.

We now include Keller et al. (2014) on line 181.

Line 202: "OH, ozone and NO3". OH and NO3 should be defined. OH is defined later on line 486.

Both OH and NO3 are now defined on line 201.

Line 206,229: The unit has a period after gram instead of a space.

The erroneous period has now been replaced with a space.

Line 249: MAM4 already defined so not necessary to do here.

We now simply state on line 249 that CESM uses MAM4 with a minor edit to make it clear that the 4 in MAM4 refers to the number of modes (lines 248-250).

Line 315: NOx is defined as NO+NO2 but then NO is used a few lines later without being first defined as nitric oxide.

NO and NO_2 are now defined as nitric oxide and nitrogen dioxide on first use (lines 312-313).

Line 322: Sulfur dioxide not defined first before SO2.

SO_2 now defined as sulfur dioxide on first use (line 321).

Line 331: Should this say "since iodine chemistry is not explicitly modeled" or do the authors mean only "iodine"

We now state that "iodine species and chemistry" are not explicitly modeled (line 330).

Line 346: "(methane, N2O, and chlorofluorocarbons)". When are these used again and if so should they be defined? N2O is not defined yet. CFCs defined at line 478.

All three are now defined here (lines 345-346). Although we do not discuss methane or N_2O further, we felt it was helpful to include the chemical formulae since both (especially N_2O) are so frequently discussed in such terms.

Line 377: The previous paragraph also begins with "Although". I suggest changing one of these to an alternative word, like "While" (or "Whilst" if you prefer the British English)

We chose to go with "While" (although I like the sound of "Whilst", I must bite my tongue and recognize that I am in a US institution and have otherwise used American English).

Line 387: "Then, we perform a comparison of its output to that generated by two other model configurations (Section 4)". I suggest adding something to this sentence so it reads to the effect of "the two conventional configurations for CESM-chem and GEOS-chem". Otherwise, the "other model configurations" is ambiguous as the next sentences states "by comparing these results to C-CC", it is unclear if the "results" is just the C-GC or C-GC plus these two other model configurations.

We have adopted this change as suggested in full (lines 386-388).

Line 398: I suggest changing "data" to "output" as this is from model simulations, and not observations.

We now refer to "model output" instead (line 397).

Does "This section" in Line 399 refer to Section 3 or Section 5?

The reference was to Section 3, which is now stated explicitly on line 398.

Line 402: I suggest changing "All simulations cover January 1st to December 31st 2016, with an additional 6 months (S-GC) or 1 year (C-GC/C-CC) of spin up" to "Following a model spin-up period (6 months for S-GC and 1 year for C-GC and C-CC), the 1-year period of January 1 to December 31, 2016 was suitable for multi-model evaluation."

We have changed the lines in question (401-402) to read "Following a model spin-up period (6 months for S-GC, 1 year for C-GC and C-CC), the one-year period of January 1st to December 31st 2016 is simulated and used for multi-model evaluation".

Line 415: could state this is the stratopause

We have adopted this change (line 415).

Line 423: Are the truncated levels the first 56 layers of the GEOS grid and then are they somehow regridded to match the 56 hybrid pressure levels of CESM? Can the authors clarify this?

We now clarify (lines 423-424) that the "upper 16 layers from MERRA-2 are removed, leaving a truncated 56-layer vertical grid which is used unmodified by CAM6".

Line 438: CEDS should be defined and reference provided, especially as a new version of CEDS was recently released.

CEDS is defined previously on line 320 and includes the relevant reference. We now repeat the reference on line 439.

Line 453: Is there a reference for AeroCom?

Thank you for catching this oversight! We now reference the relevant source (line 455).

Line 506: What is the pressure range in Figure 2? It is hard to know from the y-axis labels. Some of the other vertical plots have the top of the y-axis given in the Figure caption. Check this is given in all captions.

The range is 300 hPa to 1.65 hPa (the model top for C-CC and C-GC). This is now specified in the caption (line 510). A similar specification is also now provided for all other figures.

Line 515: Remove "hydroxyl radical" as OH already defined.

This redundant definition has been removed.

Line 531: Should "ozone" be clarified to be tropospheric ozone or UTLS or which levels did the Park et al study look at?

Park et al. (2021) were focused on the surface and free troposphere, so this has been modified to state "tropospheric ozone" (line 532).

Line 546-557: I am struggling with the structure of these two paragraphs. The bigger differences are in panels b and i, yet this first paragraph focuses on (h) and then the second on (i). I suggest reconsidering the flow of the paragraphs to something like:

There is a clear link between the ozone distributions and water vapor. Outside of the tropics and below the tropopause (Figure 3i), water vapor concentrations are up to 30 % greater in C-GC than in S-GC…<continue as written>...(an indirect sink for ozone). While ozone concentrations are uniformly lower for C-GC than for C-CC (Figure 3c), water vapor concentrations are uniformly greater for C-GC than for C-CC (Figure 3d). This is not surprising since the representation of moist physics in the two models is identical. However…<continue as written>...HNO3. Differences in ozone related to tropospheric NOy and halogens will be explained in detail in Section 4.4.

We have attempted to restructure this section following the reviewer's advice. (lines 548 to 577).

Lines 565-569: Is it differences in moist processes or transport?

Since both could be relevant, we now say "since the representations of transport and tropospheric moist physics in the two models are identical" (lines 558-559).

Line 576-577: I am struggling to see this link.

The original text read "…in the Northern mid- and upper latitudes below 900 hPa, OH concentrations are 10-20% greater in C-GC than in S-GC. This reflects the greater water vapor concentrations and roughly equal ozone concentrations between the two models". We now clarify that we are referring to Figure 3e on line 567.

Line 590-595: What do the authors mean by "does not show the same hemispheric asymmetry in absolute terms"? The difference over the southern hemisphere is about the same, it is really the northern hemisphere that is different. Is the emission of bromine from sea salt limited to either hemisphere or could it be driving the differences seen in the arctic which looks bluer to me in panel c than panel b. In the northern hemisphere, the C-GC and S-GC are fairly close (2.2 ppbv difference). Are the anthropogenic emissions driving the differences in the northern hemisphere or the southern hemisphere?

We have rephrased this to instead say that "The comparison between C-GC and C-CC (panel c) shows a similar difference in Southern Hemispheric ozone over oceans, but the relative difference now also extends to Northern Hemispheric oceans." (lines 591-592). After a brief discussion of the role of bromine and biogenic emissions (lines 593-597), we now also state that differences in the response to anthropogenic emissions may play a role in the different Northern Hemispheric ozone concentraitons (lines 597-598).

Line 602-604: This hot spot is also seen in both C-CC v C-GC plots. So is it almost more remarkable that January S-GC v C-GC does not have it? I wanted to link it to the hot spot over the amazon, but I got my sign wrong. Is there a clear feature in the C-GC minus C-CC in Figure 4 that is consistent all year round that could be linked to this hot *spot in the bottom panel of Figure 5 that could partially explain July top panel? (if not, then this falls under the need for future work).*

Based on additional review, we agree that the July hot spot is present in both C-CC and S-GC, and now state as much in the text (line 606). However, we believe that the feature in the January C-CC vs C-GC comparison is distinct, and not necessarily related. This feature is now commented on at the end of the same paragraph (lines 607-608).

Line 626: reference Figure 7 when claiming the differences are greatest over the oceans, as this cannot be determined in Figure 6.

We now do so (line 630).

Line 626-628: The link to OH differences, is that still for below 800 hPa or throughout the troposphere? I do not see a clear link between Figure 3f and Figure 6c regarding the higher sulfate aerosol mass in the southern latitudes, but I can see patterns that match the OH zonal plots elsewhere in the troposphere.

We agree that the link between Figures 3f and 6c is not especially clear, and that OH alone is unlikely to explain this discrepancy. However, we had overlooked the possibility of differences due to in-cloud sulfur processing, which is handled differently in the C-GC/S-GC compared to C-CC. We now mention this possibility on line 633 and refer the reader to the original description in Section 2.3.1.

Line 629: Can you quantify the "more closely follows"? I think this is the differences of +/- 25% quoted at the end of the paragraph but I think it should be quoted earlier. Can you give a reason by the C-GC and S-GC differ so much? Is it the difference in the bulk representation vs summing across all aerosol size bins? Oh, based on my latter reading of line 654 saying it is the convective scavenging, I think you may need to check here in line 630 if this should be broken into two sentences or at least properly reference (b) and (c) as I missed here the authors were referring to the 'differences in the representation of convective scavenging" was now discussing panel (b) and not still panel (c).

We have moved the quantification (+/-25%) to immediately follow the statement that C-GC more closely follows C-CC, and also direct the reader to the panel c. The sentence in question has been split into two as suggested, and panel b is now indicated in the discussion of the differences between C-GC and S-GC (lines 634-637).

Line 644: In the previous paragraph there was a link to DMS for these greater concentrations and OH. Does that not apply here?

It does apply, but we have not been able to definitively identify the root causes of differences. Since we already mention several possible causes for differences in the prior paragraph, we chose not to repeat those explanations here.

Line 655-657: it is not clear to me why POM would be greater in C-CC if the emissions of POM in C-CC are 29% lower.

This was an oversight, and we thank the reviewer for catching it. We now state (lines 661-663) "[a] lthough emissions of POM in C-CC are 29% lower they occur as accumulation-mode rather than primary organic mode aerosol, which may extend their lifetime".

Line 674: except the two points just above 50% (as mentioned in line 677). May be best to change your upper bound to match or give the range (looks like about -25% to 55%).

The range has been updated to read "-26 to +55%" (line 683).

Line 688-691: Is the difference at 10 hPa really significant? At 200 hPa, the 60 and 63% are also only 3% different and read like they are about the same, but the striking difference is the much greater 78% in C-CC at 200 hPa. Is it fair to open the paragraph with "partitioning between NOx and HNO3 differs significantly between the three models"? Is there a point where the difference becomes significant?

We agree that the differences in partitioning are predominantly between the GC and CC-based models. 10 hPa is highlighted because at this pressure the total NO_y is consistent between C-CC and C-GC, but the

partitioning is more consistent between S-GC and C-GC. We have re-written the paragraph opening to highlight this (lines 697-699).

Line 693: Can this statement on sulfate aerosol size distribution be linked to Figure 6 at all?

Our intent in this statement ("the CESM-driven simulation includes a more detailed representation of the sulfate aerosol size distribution") is to highlight that the explicit size distribution representation may affect the rates of hydrolysis. Unfortunately Figure 6 shows only total sulfate aerosol mass, so we could not draw a clear link to this statement. However, we do now draw the reader's attention back to Figure 6 on lines 703-704.

Line 704: At 500 hPa, I cannot see any NOx in the figures. Given the previous sentences say that from 200-900 hPa HNO3 and PAN are dominant and the following sentence says NOx once again becomes significant contributor, is there much NOx contributing at 500 hPa that it is worth listing it here in the 78%, 85% and 97%? Or is this to be consistent with what is in line 702?

We include NO_x for consistency with the prior statement (previously line 702, now line 710) as surmised by the reviewer. To try and make this clearer we now say "the same combination of species (NO_x, HNO_3, PAN)" (lines 712-713).

Line 706-707: This sentence is confusing as it opens talking about lower altitudes but then ends referencing 200 to 300 hPa and then the new text again refers to lower altitudes and is possibly redundant. Can the authors revisit these last two sentences to make sure the message is clear.

On review these final two sentences seem redundant, and as such have been removed.

Line 710: Lightning NOx emissions are not identical in Table 2 for C-GC and C-CC, but they are calculated using the same online parameterization. This should be clearer in this sentence.

We now state instead that the same parameterization is used (line 720).

Line 715: Make sure there are not two Table 3 references. May be an error in the marked-up version only.

The redundant reference has been removed.

Line 717-718: Is the percentage of NIT about the same between C-GC and S-GC?

Similar but not identical. We now state (lines 728-729) that "[t]he ratio of nitrate in aerosol compare to in gaseous HNO3 is similar at low altitudes (below 900 hPa) between C-GC and S-GC, with nitrate mixing ratios being lower than HNO3 at 900 hPa but greater than HNO3 at the surface".

Line 723: Is there a reference you can put for the Neu scheme? It is a bit confusing given the Neu and Prather reference after MAM.

We now cite Neu and Prather (2012) immediately after "Neu scheme" instead of after "MAM4" (line 734). We also state on the first usage of "Neu scheme" that this is the name most used for the scheme described by Neu and Prather (2012) (line 194).

Line 736: What is the scheme used in S-GC? It is not referenced earlier in this paragraph. Can the authors remind the readers?

We now cite Luo et al. (2001) again on line 747.

Line 747-749: This kind of introduction of halogens (referring to both bromine and chlorine) may be better suited for the start of Section 4.4 (Line 660) instead of in the start of the bromine section.

We have moved this text accordingly (now lines 666-668).

We respond to the following three reviewer comments collectively:

Line 759: It looks to me like at 1000 hPa, the difference is exactly if not slightly less than 100%. Is it correct to say "exceeding … by 100%"? Add reference to Figure 12b in the sentence. If Figure 12 is not referenced again, this paragraph only focuses on the result of surface concentrations and this could likely be summed up in a table instead. Line 762: If Figure 12 is kept, reference Fig 12c in this sentence.
Line 765: This sentence is misleading as we do not have information on CH3Br and OH in Figure 13. If this is a statement of chemistry knowledge, about Bry increasing with height due to this reaction, it should go with Figure 12. I encourage the authors to reconsider the reference to Figure 13 for the description of Bry profiles (Line 765-771). Possibly this was a Figure reference that was not updated correctly (line 764) and was intended to go with Figure 12 all along.

We agree both that Figure 12 was under-utilized and that the initial discussion of CH_3Br belonged with Figure 12 rather than Figure 13. We have extended the discussion of Figure 12 (lines 767-773), more clearly referencing the role of OH. We believe that Figure 12 is also more appropriate for the discussion of total Br_y than Figure 13, and have moved the discussion of variation with altitude to this location.

Line 800: Could refer back to Figure 12b to quantify this large difference.

We now do so (lines 810-811).

Line 804: In the marked-up copy there is a double "mid-tropospheric mid-tropospheric" and then at the end of the sentence another "mid troposphere". Can the authors reduce it to just one?

Done (line 815).

Line 812: Could remove "vertical" before "profiles". Cly has been defined a few times. I am not sure if you need to redefine it here. The authors did not redefine Bry in line 750.

We now simply state "[a]nnually-averaged profiles of Cl_y" (lines 821-822).

Line 815: Add C-GC to this sentence "total Cly for C-GC follows…". Is this true, given there are swings in the % difference with values over 100 % near the surface, then negative, then back to over 100%? The distribution is about the same from 100 to 2 hPa (Figure 14b).

We agree that "the same" was too strong. We now instead state that "Comparing C-GC to S-GC (Figure 14b), differences in total Cl_y follow a similar pattern to Br_y (Figure 12b) up to 10 hPa" (line 825).

Also, I do not agree with the "above this pressure, the vertical distribution of C-GC is closer to C-CC", and this statement does not agree with what is stated in Lines 818-819. The authors need to revisit the discussion of Figure 14. Would it flow better if the authors followed a similar description as for Figure 12 (i.e., first the extreme differences at the surface and then the rest of the profile)?

This statement was inaccurate – our intent had been to point out that the relative difference in Cl_y above 10 hPa was stable between C-GC and C-CC, but increasing when comparing C-GC to S-GC. We have followed the reviewer's advice and restructured this paragraph, doing away with the misleading initial summary and instead following the same structure as was used when describing Figure 12 (lines 825-830).

Line 831: Can the authors quote altitude range here for UTLS HCl (e.g., 400 to 100 hPa)? I am trying to estimate by eye comparing panel c to b and I do not know how far up to the range should be. Does it go over 50 hPa, given the sentence that follows picks this height out for further discussion?

We now state "specifically, between 200 and 50 hPa" (line 843). Although the comment is true at lower altitudes also, this range is relatively clear in Figure 15.

Line 832: While this statement is true, my mind went to Figure 14 looking for the percent difference, not from Figure 15. It would likely be easier on the reader if the 15% came off Figure 14, even if you reference Figure 14a which clearly shows C-CC is greater than C-GC and SGC. I wanted to get this number off panel b or c, and that is less straightforward. As for the difference in HCl, this is not shown in a figure like panels b and c of Figure 14.

We now reference Figure 14a in this location (now line 844).

Line 840: Add C-GC to this first sentence. Can the authors link these values to the percent differences shown in Figure 14b and c?

We now specify C-GC (line 851). We chose not to reference Figure 14 or 15, as the contribution of Cl atoms to total Cl_y is negligible.

Line 849: MOPITT is on TERRA.

Thank you for catching this! We now state “Terra MOPITT” (line 860).

Line 902-903: Does this sentence refer to Figure 4 or another figure (possibly moved to the supplemental) that compares model to observations somehow spatially so the readers could infer the biases over the Mediterranean Sea and Northern Europe. And if there is a high bias, how does that line up with the sentence that follows on 907 that the models are biased low.

We realized that this wording was confusing; we had intended only to say that geographical patterns were consistent. When we stated that ozone was “high” over the Mediterranean and “low” over Northern Europe, this referred to value rather than bias. We have removed the sentence in question as it did not add to the discussion. Instead, we now refer readers to Figure 4, which shows the surface ozone in C-GC and the differences when comparing to the other two configurations (lines 914-916).

Line 909: These numbers could go nicely in Table 9 next to the correlation values. Might help the reader to follow that the greatest bias is found for the C-GC as quoted in line 916.

We agree and have moved these numbers into Table 9.

Line 930-931: Can the authors provide panel references in the text here. I expect this plot is produced by a CAM-Chem post-processing script so it is not something to change as folks familiar with these plots are familiar with this layout, but it is counter-intuitive to me to have the plots nearer the surface at the top and the stratospheric levels in the bottom rows. This figure is very busy; if the 1, 2, 3, 4 key in the bottom left of each plot could be removed except for say in the bottom panels (j, k, l) that would help where there is overlap with some of the plots.

Unfortunately we have limited control over the plot, as it is (as the reviewer correctly surmised) a standard output from CAM-chem post-processing software. However, we now explicitly state which panel we are referring to in each location (lines 937-947).

Line 936: Can you refer back to Figure 2 where the authors present the zonal mean ozone. The values at that height all appear to be about the same across the models so it does not surprise me that all three models appear to perform similarly.

We now do so (lines 956-957).

Line 947-948: Is this again a familiar comparison for CAM-Chem users to a climatology for 2004-2010? I suggest repeating this reason for using this other reference period if it is the case.

We now state that this is a standard output from CAM-chem processing software (line 958).

Also, could merge these first two sentences together because Figure 20 shows more than just TOC from OMI/MLS. “Figure 20 shows …<continue as written> …(Ziemke et al., 2011) (panel a) compared to results from C-GC, C-CC, and S-GC (panels b, d, and f, respectively)”.

These sentences have been merged to ensure that it is clear to the reader what is being shown (lines 956-958).

Line 956: Out of curiosity, is the difference in S-GC driven also by the spring Antarctic? Or is there something fundamentally different about the seasonality of the S-GC vs C-GC runs that leads to the 4.5 DU mean bias difference (7.8 vs 3.3)?

We are not sure of the drivers behind the differences, but suspect that behavior at the South Pole may indeed be relevant. To this end we now refer the reader to the SI, where we explore the differences in Antarctic ozone depletion between C-GC and S-GC (lines 966-967).

Line 958, Figure 20: The difference color bar labels are a bit odd. Any way to force it so the labels for white are -20 to 20 DU? It does not look like you need all the colors in the color bar. Similarly, Figure 21 difference color bar labels could be improved.

Unfortunately we are again constrained by the CAM-chem postprocessing software, which automatically picks the color bar and range. As such, we have left this as-is.

Line 964: Why only April and not the whole year as in Figure 20? Is it possible to have the mean values in the top right title label like in Figure 20 instead of the units mol/cm2?

Again we are showing a standard output in this case which does not generate an estimate of the mean bias. Although analyses are produced for four different months (January, April, July, and October) we concluded that it would be unwieldy to show all possibly combinations and therefore only show a single month.

Line 1053: Is there a reference for MUSICA (publication or website) which could be included here?

Yes – we now cite Pfister et al. (2020) on line 1064.

Referee #2

I am content that the authors have attempted to address my comments, and have also clarified their aims for the paper in their responses.

We thank the reviewer for their time and effort.

We would again like to thank the reviewers for their time and insight, and believe that their input during this review process has improved the paper substantially. Thank you again for considering our manuscript for publication in *Geoscientific Model Development.*

Regards,

Sebastian Eastham