Dear Editor,

We appreciate for your time revieing this paper and providing constructive comments and suggestions. Here are our responses point by point in blue.

Comments to the author:

1. Regarding the comments/replies to Reviewer #2:

I see none of them implemented in the most recent version (gmd-2022-82-manuscript-version6.pdf). Some are included in the author-tracked version, however not all, e.g. mentioning the additional analysis (and supplemental figures). Regarding the latter, I suggest *not* to it because 1) maximum annual biases are already a strong measure of model deviations, and 2) there is no sense comparing biases for the first and last day of the simulations for species whose lifetimes are much shorter than a year (read their budgets are exchanged by emissions/sinks several times a year). Otherwise, explicate *how* do you expect year-long biases accumulation and propagation to occur. Should you decide to keep the supplement material, please follow the rules of its preparation (title page, etc.).

A copy of the tracked version (without the tracking feature) was saved, but it wasn't used for submission due to human error. We apologize for this mistake.

The intension of showing the maximum absolute bias for O_3 and $PM_{2.5}$ on the first and last days of the simulations, was an attempt to resolve the 2nd reviewer's concern regarding (CMAQ) model degradation (does error worsen over time). We agree with your assessment about the lifespan of simulated pollutants. We will not include the supplemental figures and we will add the following text at the end of Section 3.3 in the manuscript to reassure the stability of our simulations (for reviewer 2):

"No error accumulation due to the non-systematic changes in model inputs (changing precision introduces both positive and negative changes in a spatially and temporally random manner) can occur over the course of the annual simulation for chemical species of interest such as O_3 and $PM_{2.5}$. Their lifetimes are much shorter than a year, i.e. their simulated budgets within the continental-scale modeling domain are repeatedly exchanged through transport, emissions, and chemical and physical sinks. All simulations (*orig*, *A05*, *A04*, and *A03*) are numerically stable (no compounding error over time)."

2. Regarding the changes to the model run time (Sect. 2.3):

Since it is not possible to disentangle the influence of external factors that have affected the run-time of the simulations, presentation of Fig. 3 is impractical. Some computer systems/codes provide alternative measures of simulation effort (e.g. cpu-time, CPU cycles or whatever metric scaling with the number of operations performed) which are independent of wall-time used. Should you possess such from your simulations, please use them in Fig. 3 and in the related discussion instead. Otherwise, Fig. 3 should be removed.

We do agree with the editor's view "Since it is not possible to disentangle the influence of external factors that have affected the run-time of the simulations, presentation of Fig. 3 is impractical.". Instead, we believe execution time is a vital part of our study. We have provided potential explanations for the behaviour execution time which was illustrated in Figure 3. Even though these explanations might not determine the influence of external factors that have impacted simulation run-time, it will be beneficial for the community to understand that such behavior exists (being transparent and maintaining integrity).

Indeed, for n = 3, the execution time was smaller than the original case almost everyday and for the other two cases, the execution time was smaller for most days. We don't believe this is a product of random behaviour. Additionally, CMAQ is an MPI-based model (reflected in the first sentence in Section 1 in the revised manuscript). The model elapsed time (execution time depicted in Figure 3) was reported by an MPI function called, 'MPI_WTIME' (reflected in the first sentence in Section 3.2 in the revised manuscript) which utilizes a native high-resolution timer, within the model itself. We also used the Linux 'time' command which summarized real-time, user CPU time, and system CPU time spent by executing the mpirun command. The user CPU time was almost identical to the reported elapsed model time. This is a method to validate the legitimacy of the elapsed time we provided in Figure 3.

3. Presentation issues:

The values in Tables 3–6 are anything but a comprehensible way to present the changes between the orig and altered-precision simulations; the ranges between the latter have no practical meaning (or please elucidate what useful information you derive from them, I see only sense in minimum and maximum difference, but not the difference of the latter). Since you investigate the impact of altering the precision on your results, you need to present the *differences* in biases, MNB, etc. of the sensitivity simulations versus the orig case. These in turn may even provide you with a possibility of a rough extrapolation of how the accuracy changes depending on the number of reduced digits.

Thus, I advise that the bias, MNB, r, RMSE and minima/maxima from the orig simulation *only* should be presented in one table for all species (PM2.5, O3, NH3). *Changes* to the bias, RMSE and minima/maxima values should be shown in bar chart plots (one plot per characteristic, i.e. four plots maximum, please use overlapping axes for different species). Ranges should be removed. So should be the changes to r (they bear no meaning on these scales) and NMB values (they are equivalent to biases).

We have removed Tables 3 – 6 and we created Table 3 to show bulk statistical metrics of bias, NMB, r, and RMSE for the *orig* simulation so a reader can gauge the accuracy of CMAQ. To follow the editor's suggestion, we have created Figure 4 to show the absolute change in bias, RMSE, and minima/maxima with respect to $PM_{2.5}$, O_3, and NH_3. The manuscript has been revised to reflect these changes. Information about ranges, change of r, and NMB are not included.

Figures 6–8: please re-render with distinct hue colour scale (i.e. "rainbow", or any other multi-colour) and lesser number of levels – right now it is not possible to read the actual value from the colour in the plot. Same applies to Figures S1–S3, should you decide to keep these (I strongly discourage that, see the comment point 1. above).

Thank you for the suggestion. For Figures 6 - 8 (Figures 7 – 9 in the revised manuscript), we followed the guidelines < https://www.geoscientific-model-development.net/submission.html#figurestables > to make our images vision impairment friendly. We chose the color map titled, 'viridis'. Therefore, we would like to keep them as they are.

Figure 5: please scale vertical axes according to the maxima of values presented in a given plot, or at least according to the maximum for a given species. Current scales have little sense w.r.t. to the data shown.

Figure 5 (Figure 6 in the revised manuscript) has been re-scaled.

4. Regarding the comment of Reviewer #1, Point 11.:

I very much agree with the Reviewer #1 that presenting the estimate of the changes to the *cumulative* deposition is beneficial for the study. I reckon that obtaining such is possible (you just need to integrate the hourly rates available). Therefore, please replace the analysis of hourly estimates with the integral ones (annual). Presentation of the latter should be improved similar to that described in p. 3 above. That is, integrals for orig simulation should be presented in a separate table or added to that including biases, RMSE, etc., and differences should be shown in plots.

We have calculated the sum of absolute differences (annual for all hours and cells over the US over land) for the dry deposition of ozone and for the wet deposition of sodium, ammonium, chlorine, nitrate, and sulphate throughout 2016. The values have been plotted as bar plots and are provided in Figure 10.

Beyond these four responses, we have indicated that the high bias in $PM_{2.5}$ was associated with the Pioneer wildfire in Idaho from July to September of 2016 (line 235 – 237) as well as we have updated the conclusion with an explicit recommendation: “These results show the optimal benefit of altering CMAQ input data by keeping three significant digits then subsequently keeping four significant digits for CMAQ output data.”