Referee #2

Dear Referee, #2, thank you so much for reviewing this paper and providing constructive review and comments. Here are our responses to your comments and suggestion point by points (in blue).

In this paper, the authors assess the degradation in accuracy of the CMAQ model when they decrease input/output precision to save disk space. They do a number of tests to see how the model results change when precision is degraded. For the most part, they find very little change.

We don't think the word degradation or degraded is right choice of word to describe our approach. Truncate (to 'n' significant digits) is more applicable. Certainly, by altering the data precision does not change the quality of the data or make it inferior. The accuracy of CMAQ is essentially unchanged. Our work examined a given number value in 32-bits precision how significant the last few digits areto a numerical model and if they are removed, how does it impact on data compression.

When I first got this review request, I was concerned that there would be very little to the paper. Reading a bit in, my though changed to it could provide a bit of useful information. At the end, my real concern is that they really don't get in to the most interesting results.

Let us use this opportunity to reiterate the interesting result that we offered in this work. Altering the precision of the data did not impact the model performance significantly, statistically speaking, but increases the data compression efficiency substantially by using ordinary compression tools. This is useful to most researchers who need to retain large quantities of data for a long period of time. Definitely this approach reduces storage costs significantly.

First, fundamentally what they find is that the model results change very little, on average, when they reduce emissions and stored variable precision. This is a good result, but actually, it would be of great concern if that was not the case, and would indicate that what we know about air quality models is wrong. Why do I say this? There have been lots of studies of how changes in emissions impact results, and changing the inputs by 0.1% (less on average) should have virtually no impact. Good to see that is the case. It is also expected that the change in stored variable precision would also have little impact on average. Thus, their results are very much in line with expectations.

The authors share your thoughts regarding our results. It is known that for a numerical model, when an input is perturbed, the model will produce different answers. Furthermore, researchers have done numerous emission reduction studies and the reduction amounts are in 10%, 20%, or higher range. In our study, the change is about 0.1% or less as the referee has indicated which was contributed by removing the last few digits with the 32-bits data precision. We were not aware of any studies interested in this small amount of change in emissions. We agree with the referee's comment "very much in line with expectations", however expectation is expectation. We validated such expectation with a sound scientific approach and that is what this paper is about.

An interesting twist is that they also look at run time and find little difference, though there are some spikes. Those spikes are insufficiently explored.

We shared the referee's same concern "Those spikes are insufficiently explored". We have rerun those dates with spikes and the spikes was not observed. In other words, those spikes could be considered artifacts of running jobs on a non-dedicated system and are not repeatable.

One of their most interesting results is how changes in emissions precision had a radical result on the maximum. That result begs for more analysis.

For additional analysis, we have added the following paragraph and three new figures in the supplemental document.

"Figures 6–8 show the annual maximum absolute bias of the entire domain with respect to PM2.5, O3, and NH3. We have done additional analysis by examining daily maximum absolute bias with respect to those three variables (Fig. S1–S3). By displaying the daily maximum absolute bias of the first and last day from an annual simulation: 1. regardless of simulation cases, A05, AO4, or AO3, the bias range is quite similar for all three variables, 2. daily bias coverage scale is much small compared with the equivalent annual maximum absolute bias (Fig 6-8), 3. bias from the first day did not accumulate through the entire simulation. For A03 case, NH3 daily maximum absolute bias on the first day showed some large values but the last day did not. This echoes the third point quite well. Overall, this ensures the integrity of the model performance when an altered-precision emissions were used."

A final comment is that not only did they not go far enough on both the run-time results and why the precision can have such large impacts, if even on isolated values, but answer the question, at what point do results really start to degrade?

Regarding the run time results, we concluded that run time 'spikes' were artifacts caused by simultaneously running tasks on a non-dedicated system. We determined this by rerunning the A05, A04, and A03 simulation for certain days in which a spike occurred. Upon re-running simulations, the spikes were gone or reduced in magnitude. For the FX03 case, the run time is lower than the original case most of the time; a potential explanation is that a change in emission input could alter the pathway in aerosol dynamics calculation or reduce the number of iterations in the chemistry solver.

An additional analysis was performed (please see above response). It ensures the model did not degrade when altered-precision emission was used.

Detailed comments:

Line 54: What is meant by the chemical transport model within CMAQ. Is not CMAQ a chemical transport model?

Thanks for pointing that out. CMAQ indeed is a chemical transport. To clarify that, we have revised "The chemical transport model within CMAQ" to "CMAQ model".

Ine 61: "could"? Seems a bit weak.

The phrase "management approaches could be justified" has been revised as "management is justifiable".

Line 135: remove the comma after called.

Done.

In Table 1, add a footnote to state what AOX are.Table 2 does not do its job well. It does not say what the FX cases are.

AOX are defined on line 100 of the original version of the manuscript and on line 109 of the revised version so I don't think a footnote is necessary. FX cases are defined on line 104 of the original version of the manuscript and on line 114 of the revised version.

Line 183: "Resultantly"? Awkward.

The word "Resultantly" has been removed.

Given the small changes, most of the tables and figures are not needed (certainly the figures are not needed as there are no real observable changes).

Despite the small differences in tabular values and geospatial images, we believe the tables and figures strongly support our conclusion. We will leave the tables and the figures in the manuscript.