Responses to Reviewers

Anonymous Referee #2

Received and published: 28 October 2020

Wang et al. present a study which evaluates the performance of the of Two-way and Offline Coupled WRF v3.4 and CMAQ v5.0.2 over the Contiguous U.S for an extended time period (5 years). Previous works had experimental design deficiencies (e.g., differing physics, chemistry) that his work addresses. The importance of chemical meteorological
feedbacks are increasingly being recognized as essential for the prediction of both weather and atmospheric chemistry, and this work adds well to that body of work. Outside of a major comment w/ regards to the experimental design (cycling between 5-day periods), my main critique of the manuscript is the heavy reliance on the use of 5-year averages to discuss model performance and comparisons. This is also somewhat related to my major comment about cycling. I think it would benefit the community to examine and discuss seasonal spatial patterns (and thus reasons for model deficiencies), periods of peak aerosol and/or high ozone days (not just number of exceedance, but more details in how the model performs/evolves).

Reply: We thank the reviewer for the general positive comments and recognizing the importance of this study. We have carefully revised our paper to fully address the reviewer's comments. Some major revisions have been performed including:

1) Old Figures 1-7 and Tables 1-2 are moved into supplementary materials. New Figures 1-13 and Tables 1-4 have been added by including the seasonal results.
2) A new figure (Figure 3) with the annual trend results has been added and corresponding analyses have been added in Sections 3.1.1, 3.2.1, and 3.2.2.
3) All texts in major Sections 3 and 4 have been updated with the seasonal results.

Please see our point-by-point detailed responses below.

Major comment: Line 185: Are any fields cycled between consecutive 5-day simulations besides chemistry? (e.g., land surface fields?) I think this needs to be discussed in detail how it relates to the experiments. If they are reinitialized every 5 days, should the first day or two be considered in the comparisons? The deviation between the two simulations would likely increase as lead time increases. Here is really comparing 5 years of 5-day forecasts.

Reply: The 5-day reinitialization are actually only applied to the meteorological fields (land use or land surface fields are assumed to be constant) and the chemistry fields are continuously simulated without any reinitialization. The same approach has been applied to both two-way WRF-CMAQ and WRF-only (providing the meteorological fields for offline CMAQ simulations) simulations. So any deviation for meteorology fields between two simulations are really more determined by the feedback processes. The reinitialization approach used in this work is very common practice for the air quality/chemistry transport models to ensure the accurate simulations of meteorological fields. The reinstallation may lead to some initial shock at very beginning of each 5-day, but wouldn't make significant impacts on simulation results based on our previous studies (see Wang et al., 2021).

We have added the following statements in revision (L190-196) to further clarify this issue.

"Two-way coupled WRF-CMAQ simulations are reinitialized every 5 days for meteorology fields only. We have conducted sensitivity simulations in the past (Wang et al., 2021) and found that a 5-day reinitialization frequency is more suitable to improve the overall simulation quality to make meteorology simulations as accurate as possible while preserving the two-way chemistry-meteorology feedbacks. The WRF-only simulations that are used to drive the offline CMAQ simulations apply the same reinitialization method to make sure any deviation between two simulations are more determined by the feedback processes"

Minor comments: Sections 3.1.1 and 3.1.2: It would be much more helpful to at least break these comparisons up into summer vs. winter as some biases could be cancelling one another out.
Reply: As suggested by the reviewer, we have broken up our previous 5-year annual average evaluation/analyses into summer vs. winter comparison for Sections 3 and 4. We have replaced most of previous figures and all tables in the original submission into seasonal comparison (i.e., summer vs winter). New analyses have also been added in the revision to reflect the changes of those figures and tables (see the track-mode revision on those changes).

In general, the new seasonal results show general consistent performance when comparing with old annual results (very limited cases of cancel-out of model biases in different seasons occur for example for T2, which has been explicitly pointed out in the revision; see L260-263 in the track-mode revision) and thus won't affect our previous conclusions based on the annual performance analyses. The seasonal results indeed shed more lights for some of our previous analyses and speculations. For example, the new seasonal T2 performance now can well support the O3 monthly and diurnal performance now. The speculation of model biases caused by biogenic emissions from BEIS on column HCHO in the original version has been eliminated after checking the season results (see L584-585 for changes in track-mode).

Figure 4: The colors used in the top panel are very hard to distinguish.
Reply: All the SWCF figures (now Figure 8 in revision) are updated with the new color scheme. Please note in both old and new plots, we intentionally use the cold-color only schemes to better represent SWCFs which only contain negative values.

3.2.1. Annual average ozone is not really a useful diagnostic, I think showing summer only would be very beneficial.
Reply: As suggested, we have replaced the annual average ozone with summer only and updated the corresponding texts in Section 3.2.1.

3.2.2. Again, a seasonal analysis here would be more appropriate (i.e., winter is dominated by NO3, summer with OA (and SO4)).

Reply: Both the spatial overlay plots for PM2.5 and scatter plots for PM compositions have been broken up into summer vs winter comparison and the corresponding texts are updated in Section 3.2.2.

Figure 5a-b: You could shift the color limits by 20 ppb.
Reply: Both figures (now Figure 9a-b in revision) have been replaced by summer ones and the scales have been adjusted to from 30 to 60 ppb as suggested.

Figure 8. Why not just use more colors instead of the varying dot sizes – hard to distinguish.
Reply: The figure (now Figure 14 in revision) has been updated by using different colors instead of varying dot sizes.

Figure 10. Looks to be some weird striping for O3.
Reply: The striping seems to be caused by the WRF-CMAQ interface to deal with feedback interactions among multiple CPUs while conducting parallel computing, which hardly can affect any conclusions.

Reference:

Wang, K., Y. Zhang, and K. Yahya (2021), Decadal application of WRF/Chem over the continental U.S.: Simulation design, sensitivity simulations, and climatological model evaluation, Atmospheric Environment, 118331.