

## ***Interactive comment on “A Comparative Study of Two-way and Offline Coupled WRF v3.4 and CMAQ v5.0.2 over the Contiguous U.S.: Performance Evaluation and Impacts of Chemistry-Meteorology Feedbacks on Air Quality” by Kai Wang et al.***

**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

C1

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).

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.

Minor comments: Sections 3.1.1 and 3.1.2: It would be much more helpful to at break these comparisons up into summer vs. winter as some biases could be cancelling one another out.

Figure 4: The colors used in the top panel are very hard to distinguish.

3.2.1. Annual average ozone is not really a useful diagnostic, I think showing summer only would be very beneficial.

3.2.2. Again, a seasonal analysis here would be more appropriate (i.e., winter is dominated by NO<sub>3</sub>, summer with OA (and SO<sub>4</sub>)).

Figure 5a-b: You could shift the color limits by 20 ppb.

Figure 8. Why not just use more colors instead of the varying dot sizes – hard to distinguish.

Figure 10. Looks to be some weird striping for O<sub>3</sub>.

C2