Responses to Reviewers

Anonymous Referee #1

Received and published: 1 October 2020

General Comments:

The manuscript 'A comprehensive 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' written by Kai Wang presented the comprehensive comparison of offline (i.e., traditional) CMAQ and two-way coupled CMAQ over the CONUS. To promote our understating on the interaction between meteorology and air quality, the approach of two-way coupled modeling is necessary, and this manuscript can contribute to this purpose. The authors claimed that long-term simulations on both (two-way and offline) models over the CONUS is important point in this study, because the previous studies have been limited in many aspects (different chemical options, difference meteorological options, or, limited in time to focus on episode analysis). Although I would like to recognize the importance of this study, the evaluation is not well conducted in depth to make the best use of this long-term simulation. Please consider to address the following one major point, and also check the minor comments to improve the manuscript.

Reply: We thank the reviewer's constructive comments and also for recognizing the importance of this study. We have carefully revised our paper to fully address the reviewer comments. Please find our point-by-point responses below.

Major point:

I would like to disagree the evaluation framework of long-term simulations conducted in this study. The authors stated that "more robust assessments" through five-year simulations; however, the evaluation is only conducted by averaging the five-year dataset. This does not take advantage long-term simulations, and does not provide deep understanding of two-way coupled and offline models comparison. In addition to the averaged field of climatological type data, the comparison should be furthermore focused on trends in five years (if detected from observed facts) or year-to-year variations of both meteorology and air quality. Based on this extended evaluation, it could be finally proved the importance of two-way coupled model. Without such kind of evaluations, this study of long-term simulations will be less important.

Reply: In order to address the reviewer's concern, we have performed the annual trend analyses and added a new figure (i.e., Figure 3) for major meteorological variables and air pollutant species. Additional analyses are also added in Sections 3.1.1, 3.2.1, and 3.2.2 in the revision and attached below (in red color) as well:

"Figure 3 shows the bar charts of annual trends for T2, RH2, WS10, and precipitation in 2008-2012. Two-way WRF-CMAQ predicts the annual average T2 very well with MBs < 0.25 °C in all years. The simulation can also capture the increasing trend of T2 from 2008 to 2012 observed by NCDC. RH2 is consistently overpredicted by the two-way WRF-CMAQ in all years despite relatively low biases (MBs < 3%). Both observations and simulations show the lowest RH2 in 2012 and the highest in 2009. As also shown in Figure 1, the model tends to systematically overpredict both WS10 and precipitation throughout all years as well. There are no clear trends (i.e., increasing or decreasing) for WS10 and

precipitation between 2008 to 2012 from either observations or simulations. However two-way WRF-CMAQ is able to capture the lowest wind speed and precipitation both in 2012 and the highest wind speed in 2008 from observations. In general, the model performs very well in reproducing the year-to-year variation for the major meteorological variables between 2008 to 2012."

"Figure 3 also shows the bar charts of annual trends for max 8-h O3 from two-way WRF-CMAQ against AQS and CASTNET observations in 2008-2012. Two-way WRF-CMAQ systematically overpredicts O3 especially against AQS data with MBs typically > 4.0 ppb. The potential reasons for model biases have been discussed earlier in this section. There are no obvious decreasing or increasing trends for max 8-h O3 from AQS or CASTNET observations. However, the model can generally capture the high O3 mixing ratios in 2008 and 2010 and the low O3 mixing rations in 2009 from both AQS and CASTNET. The similar down and up trends between 2008 to 2010 for O3 (i.e., decreasing from 2008 to 2009 and increasing from 2009 to 2010) from AQS observations were also found by Yahya et al. (2016), but not captured by their simulations. Zhang and Wang (2016) was able to reproduce the similar trend over the southeastern U.S. between 2008 to 2010 using their models and attributed the abnormal high 2010 O3 mixing ratios to the extreme dry and warm weather conditions during fall 2010."

"Figure 3 shows the bar charts of annual averaged observations and simulations for PM2.5 over the CSN and IMPROVE sites. Overall, the model performs well for PM2.5 for most of years and better over CSN than IMPROVE sites with general underpredictions in most years. The observations for both CSN and IMPROVE show a general decreasing trend (except for 2010 over CSN) especially over IMPROVE sites. Two-way WRF-CMAQ is able to reproduce the declining trend well particularly over IMPROVE sites and again demonstrate its capability in accurately simulating the year-to-year variations of not only meteorology but air quality."

Overall, our simulations can either capture the decreasing trend for some variables (e.g., T2 and $PM_{2.5}$) or reproduce the year-to-year variation for most of rest variables well, which provide great fidelity in applying this version of two-way coupled WRF-CMAQ model for the future studies.

In addition, we also significantly revised majority of figures and tables (i.e., moving the old Figures 1-7 and Tables 1-2 into supplementary materials and creating the new Figures 1-13 and Tables 1-4) in our revision by adding the seasonal analyses as suggested by the other reviewer which we believe should provide even deeper understanding of the results and further address reviewer's concern here.

Minor points:

1. L26: "modes" is typo of "models"?

Reply: It's indeed modes. Two different coupling modes for the same version of CMAQ model.

2. L178-180 (and abstract): Are this chemical ICON/BCON considered year-to-year

variation simulated by CESMv1.2.2/CAM5? Did this model perform well compared to other model(s)? If this model had superiority, please note how this model is important. Without any specific reasons, I feel it is no need to mention this model in the abstract.
Reply: Yes. The ICONs/BCONs simulated by CESMv1.2.2/CAM5 are year specific. It's an online coupled global model with many improvements in terms of chemistry and aerosol treatments. The simulations have been comprehensively evaluated against surface, remoting sensing including satellite data, and reanalysis data for major meteorological and chemical variables over Europe, Asia, North America, and the globe. The results show generally satisfactory performance and are also compared with existing global model results such as CESM/CMIP5. More details and model evaluation can be found in He and Zhang (2014) and Glotfelty et al. (2017) as already cited in the original submission.

We have added the following statement in the revision (L180-185 in the track-mode version) to cover the above points:
"The chemical ICONs/BCONs generated from CESM simulations consider the year-to-year variation. The CESM simulations have been comprehensively evaluated against surface, remoting sensing including satellite data, and reanalysis data for major meteorological and chemical variables over Europe, Asia, North America, and the globe. The results are also compared with other existing global model results and show generally satisfactory/superior performance."

3. L183-L185: In my best knowledge, inline dust scheme implemented in CMAQ version 5.0.2 is not the scheme reported by Foroutan et al. (2017) (see, also https://www.airqualitymodeling.org/index.php/CMAQv5.0.2_Wind_blown_dust_updates). In addition, this statement contradicts to the discussion in its evaluation (L427-429). Please address this issue. If the authors implemented the scheme by Foroutan et al. (2017) in this study, exact explanation is required because this is model development paper.
Reply: We thank reviewer for catching this issue. We indeed used the default dust scheme in CMAQv5.0.2, which should be based on Zender et al. (2003) instead of Foroutan et al. (2017). We have fixed this inaccurate citation by updating the reference and now it should be consistent with L427-429 (now L511-513 in track-mode revision).

4. L212: "PM10" will include PM2.5, hence the expression of "coarse particulate matter" is not appropriate. Or, did the authors calculate PM10-PM2.5 to represent coarsemode particulate matter?
Reply: The reviewer is right that PM10 evaluated in this work includes PM2.5. So we have rewritten the definition as "particulate matter with diameters of 10 μm or less" at L212 (L222 in track-mode revision). To be consistent, we also redefine PM2.5 as "particulate matter with diameters of 2.5 μm or less" at L208 (L218 in track-mode revision).

5. L222 (and related to Section 3.2.3): "paired with the satellite retrievals" means the deficit grid points in satellite observation are applied for model results? Please clarify. I guess that some satellite products provide averaging kernel, but how did the author apply averaging kernel for better comparison between model and satellite

measurements? The detail seems to be dropped here. Please specify.
Reply: Only those grid points with valid satellite observations are considered when paring model results with observations. As noted by the reviewer, averaging kernels (AKs) for some satellite products such as NO_2 are only available for level 2 data, however all analyses in this work are based on level 3 data. Also one of previous studies by Schaub et al. (2006) found that both satellite retrievals with or without applying AKs from Global Ozone Monitoring Experiment (GOME), which uses the similar retrieval methods as SCIMACHY used in this study, show generally good agreements with ground-based measured NO_2 columns.

In short, for the current study, no AKs are applied which may introduce some uncertainties, but won't affect our conclusion. We have acknowledged this issue and also further clarified the data pairing in the revision (L233-236) as below.

"**Note that only those grid points with valid satellite observations are considered when paring model results with observations and the averaging kernels are not considered when analyzing the column CO and NO_2 results, which may introduce some uncertainties**."

6. L443-446: The review paper by Emery et al. (2017) (used for ozone evaluation in this study) also presented the model performance goal/criteria for aerosols. Why these criteria is not used?
Reply: Thanks for bringing up the criteria proposed in Emery et al. (2017) for aerosols. We have cited Emery et al. (2017) here and removed the sentence "There are no universally accepted performance criteria for aerosols." from the revision. Actually the criteria used in this study are consistent with those recommended in Emery et al. (2017).

7. Figure 10: For gas species, differences are seen along latitude (approx. each 3-4 deg.) over western U.S.A. and Mexico. What is this difference?
Reply: It 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.

8. Author contribution (L720-723): The contributions of all authors are not explicitly described here. Is it accepted in this journal style? (see, https://www.geoscientificmodel-development.net/submission.html#manuscriptcomposition)
Reply: Thanks for bring up this point. We have revised the contribution part by explicitly stating contributions of all authors as follow (L824-827 in track-mode):
"YZ and MB defined the scope of the manuscript. YZ and KW designed the study and all the simulations. SY and DW developed the two-way coupled WRF-CMAQ code. KW conducted all the simulations and performed the analyses. KW prepared drafted the manuscript. YZ, SY, DW, JP, RM, JK, and MB reviewed and edited the manuscript."

References:

Glotfelty, T., J. He, and Y. Zhang (2017), Impact of future climate policy scenarios on air quality and aerosol-cloud interactions using an advanced version of CESM/CAM5: Part I. model evaluation for the current decadal simulations, Atmospheric Environment, 152, 222-239.

He, J., and Y. Zhang (2014), Improvement and further development in CESM/CAM5: Gasphase chemistry and inorganic aerosol treatments. Atmos. Chem. Phys. 14, 9171-9200. http://dx.doi.org/10.5194/acp-14-9171-2014.

Schaub, D., K. F. Boersma, J. W. Kaiser, A. K.Weiss, D. Folini, H. J. Eskes, and B. Buchmann (2006), Comparison of GOME tropospheric NO_2 columns with NO_2 profiles deduced from ground-based in situ measurements, Atmos. Chem. Phys., 6, 3211–3229.