Answer to Review #3

Thank you for your valuable comments. The manuscript is revised accordingly, please see revised manuscript for detail. Here are the answers for your comments.

This paper compares the meteorological and chemical performance of two different ways to run CMAQ: the conventional WRF-CMAQ version and the latest GFSv16-CMAQ. Focus is given on August 2019 when the FIREX-AQ provided 3D observations for evaluation.
This paper is a revision and has been revised following comments from three reviewers. The authors have addressed the raised comments and I find the paper suited for publications. Below some minor comments that could help strengthen the paper.
1) Some improvements in the English language would be beneficial (e.g. the use of articles).
➢ Thank you for your comments. We made some literature revisions. Please see revised manuscript for detail.

2) I agree with the previous reviewers that one month is a fairly short time period and this limitation should be better acknowledged in the Introduction.
➢ We add the acknowledgement for this study period in the introduction. "This study focuses on the period of summer 2019, and Campbell et al. (2022) evaluated the GFS-CMAQ for longer periods."

3) Abstract: I suggest also stating the a comparison to operational surface met and chemistry observations is done
➢ Added the statement.

4) Section 2.1: Please also state what the temporal frequency of the driving met fields is
➢ Added "hourly".

5) Section 2.2: What is meant by the GFS levels were "collapsed"? Would interpolated be a better term?
➢ Changed.

6) Section 2.3: State how VOC emissions were derived for GBBEPx
➢ Added

7) Section 3.2: I am really puzzled by the sudden degradation of WRF 10m wind speed in the Western U.S. at the end of August. Can the authors elaborate on this?
*Also related to this. Did the authors look whether WRF performance drifts over time? I would assume that despite nudging, an over one month long simulation might show a drift.*
➢ The WRF degrading issues are related to storm weather, tropical storm Erin near the east coast and tropical storm Ivo near west U.S., later August, 2019 (see Page 9, line 7-14). WRF's setting and configuration in this study are mainly for air quality study, not tuned for predicting storm weather.

8) Page 9, Line 8: Did the authors confirm the higher NOx? If so, add at least ("Not shown")
➢ Added the "not shown". The GFS-CMAQ's higher surface NOx at night near source regions can be confirmed (see plots below). The NOx difference is almost opposite to the PBL difference (Figure 2).

Monthly Mean NOx Difference (GFS/CMAQ - WRF/CMAQ) at 18UTC

Monthly Mean NOx Difference (GFS/CMAQ - WRF/CMAQ) at 06UTC
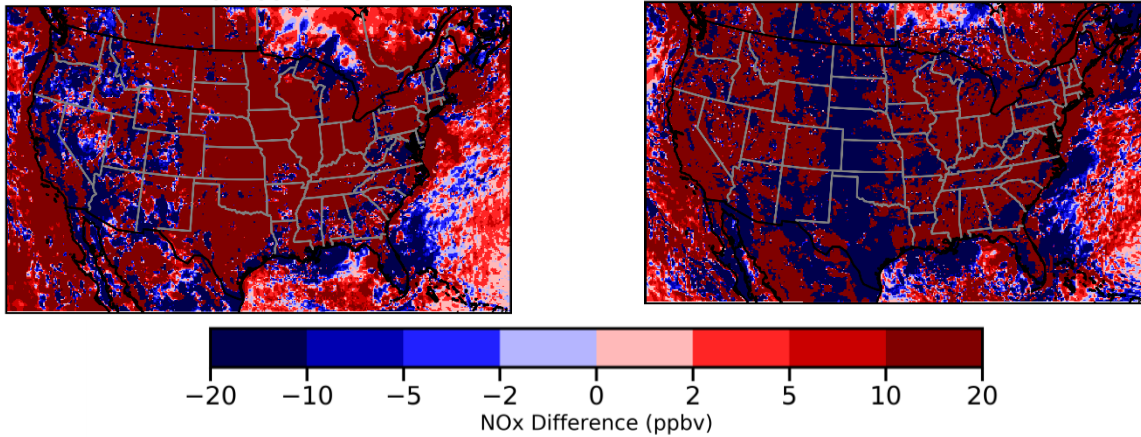
9: Summary, Page 19: I see that the speed and less computing resources is a very valid argument. However, as for nudging I wonder whether the authors can also comment on what would be if WRF is run with observation nudging or meteorological data assimilation. Or at least acknowledge that this is also a possibility.

➢ Added the acknowledgement "Nudging toward observations or including data assimilation may yield different results for the WRF run" on page 18.

Again. Thank you for your comments.