Review of

Overview and evaluation of the Community Multiscale Air Quality (CMAQ) model version 5.1

By *K. Wyat Appel, Sergey L. Napelenok, Kristen M. Foley, Havala O. T. Pye, Christian Hogrefe, Deborah J. Luecken, Jesse O. Bash, Shawn J. Roselle, Jonathan E. Pleim, Hosein Foroutan, William T. Hutzell, George A. Pouliot, Golam Sarwar, Kathleen M. Fahey, Brett Gantt, Robert C. Gilliam, Daiwen Kang, Rohit Mathur, Donna B. Schwede, Tanya L. Spero, David C. Wong, and Jeffrey O. Young*

I bring here 3 points for the final review. Point 1 strikes me as critical and needs to be corrected because it leads to a (in my opinion) biased and unjustified piece of conclusion regarding the relative importance of the update in emissions and the scientific updates.

Point 2 brings back to consideration a remark from the initial review that I think has been too overlooked by the authors,

Point 3 is a request for change in the color scale of a Figure so that the albedo is between 0 and 1

I nonetheless wish to thank the authors for the considerable work in the Review process, even though I still think that more written information about the model design could have been brough in this new CMAQ reference article.

Below, in green the Authors' text (either answers to my initial review or text from the manuscript), in black my text, in blue statements from my initial review.

Point 1

"Obviously it was not made clear in the manuscript that the overall impact from the emission platform change was small. Hopefully this is now made clear in the text. In addition, a figure showing the impact of the emissions platform change on ozone and PM2.5 in January and July has been added to the text to quantify to the reader the impact from the emissions platform change."

The following statement is introduced in the revised version (it would be easier to find if the Authors had indicated explicitly where they had made such an adjustement):
"*However, based on sensitivity simulations performed for January and July 2011 where the only difference was the emissions platform used, the differences in O3 30 and PM2.5 between those two simulations used were generally small and isolated, suggesting there is minimal impact to the comparison between the v5.0.2 and v5.1 simulations from the change in the emissions platform used. Figure S1 shows the impact on winter (January) and summer (July) O3 and PM2.5 between simulations using the different emission platforms.*"

I have several remarks on the Author's response and the corresponding changes that have been performed :
- The Figure has not been added "in the text" but as a supplement S1
- The statement that "the differences between those two simulations were generally small and isolated" seem to me as very strange : if one looks at Fig. S1a along with Fig. 6a of the revised manuscript, it is evident that **about 100% of the PM25 difference between v5.0.2 and v5.1 is due to the change in**

emissions ! Comparison with Figs. 5c, 4c, 2c and 1c reveal that all other causes of change in wintertime PM25 are about 1 to 2 order of magnitudes smaller than the changes in the emission dataset.

Unless the authors demonstrate in a convincing way that this remark is due to me misunderstanding the figures, which is possible, I recommend that :
- Fig. S1 is moved into the main manuscript since it reveals effects 1 order of magnitude larger than the figures shown in the main manuscript for wintertime PM2.5
- The statement that there is "*minimal impact to the comparison between the v5.0.2 and v5.1 simulations from the change in the emissions platform used.*" strikes me as very biased at least regarding PM2.5 and I recommend that it is replaced by a more realistic discussion.

This failure to analyze in a realistic way the effect of emission changes on PM25 leads, in my opinion, to a biased conclusion : "the scientific updates in v5.1 resulted in relatively dramatic improvements in model performance for PM2.5 in winter and summer" while, as discussed above, comparison of Fig. S1 with Figs. 5c, 4c, 2c and 1c reveal that all other causes of change in wintertime PM25 are about 1 to 2 order of magnitudes smaller than the changes due to the upset in the emission dataset, so the effect of scientific updates alone seem to be at best marginal compared to the effects of the update of dataset. As wintertime PM25 concentrations are usually a major worry for air quality modellers (due to usually strong emissions and stable atmospheric conditions), **I consider critical that this statement is replaced by a statement telling explicitly that scientific updates brought relatively small changes to wintertime PM2.5 when compared to the emissions changes, not allowing the authors to quantify the changes in model performance regarding wintertime aerosols**. I think that this is a caveat of the study that needs to be acknowledged. The only effect of the scientific updates that seem very significant to me is the effect opf the new aerosol processes in summertime as shown in Fig. 2d.

Point 2

C1 p. 5, l. 31: Is this time interval valid for all the domain? Days in July should last much longer than 12 hours at least in the north of the domain, and the daytime interval must be very different from the west to the east of the simulation domain (about 5000 km, which is about 4 hours time lag in the solar time). Using points from 11:45 to 23:45 UTC from west to east would result to using data points from mid-morning to the sunset at the eastern part of the domain, and from dawn to mid-afternoon in the west of the domain, which is critical as cloudiness often has a strong diurnal cycle.

I recommend that all the available daytime data points shall be used for this comparison.

Response: All available data from the satellite product are used in the average in Figure 1a (Note this Figure has been moved and is now referred to as FigureXa). The figure in the Supplemental Information section S1 shows the number of daytime hours (11:45UTC – 23:45UTC) with available GOES cloud albedo data during July 1 – July 31, 2011 for the modeling domain. Regions in the eastern half of the US have a larger number of available satellite observations (on the order of 390hrs) compared to the western coast which has < 340hrs. Since the reference to the time window of 11:45UTC-23:45UTC caused unnecessary confusion we have removed this from the main text. We now point readers to the Supplemental Information for further description of the hours of available satellite data:
"The satellite data are available at 15 minutes prior to the top of the hour during daytime hours and were matched to model output at the top of the hour (see section S.1 in the supplemental material for further information)."

The figure that is shown in the Supplementary material seems to only confirm what I was stating, that the time window from 11:45UTC to 23:45 UTC is arbitrary and may produce problems biases : on that figure in the Supplement showing the number of available daypoints, a strong east-west gradient is visible, and while more than 400 daytime points are available for the center-east of the US, less than 340 are available for California, suggesting that late-afternoon points are missing over California and generally the west of the domain. Less daylight time on the west coast than east-coast would actually be a very surprising result...

In San Francisco in summertime, the sunset is about 20:30 local time (4:30 AM UTC), so 23:45UTC is 15:45 local time, what one would call mid-afternoon, almost 5 hours before sunset, so are 4 to 5 hours of valid data in the afternoon/evening discarded over the western US ? Why impose an arbitrary time window and not just use all available GOES data ?

While I do not consider this point critical, I consider that the figure shown in the Supplement only reveals that what I feared in my review is actually what happens, so instead of bypassing this remark **I would like to maintain the recommendation that all the daytime data points are used for the comparison, not just between 11:45 and 23:45 UTC**.

Point 3

C2 p. 5, l. 32: the description of Fig. 1 does not fit that in the caption of Fig. 1 (the latter one seems to be more relevant). The average cloud albedo seems not to be shown.
This should be clarified.

Response: The reviewer is correct. The wrong Figure 1 was included with the original submission of the manuscript. In the revised version of the manuscript this Figure (now called Figure 4) does show the average cloud albedo, consistent with the description in the text. The Figure caption has also been changed to say:
"Figure 4. The average cloud albedo during daytime hours in July 2011 derived from (a) the GOES satellite product (b) WRF3.7 (c) CMAQv5.1 with photolysis/cloud model treatment from v5.0.2 and WRF3.7 inputs (CMAQv5.1_RetroPhot) (d) CMAQv5.1 using WRF3.7 inputs (CMAQv5.1_Base)."

There is a problem in Fig. 4 : the albedo ranges between 0 and 40, it should be a value between 0 and 1. Actually, I had ot go through the Supplement to realize that the albedo is probably present as percentage values which is, I think, very uncommoon. I recommend that the albedo is brought to its classical dimensionless form between 0 and 1 in the Figure.