Geosci. Model Dev. Discuss., 2, C533–C541, 2010
www.geosci-model-dev-discuss.net/2/C533/2010/

**Geoscientific
Model Development
Discussions**

Interactive comment on "Incremental testing of the community multiscale air quality (CMAQ) modeling system version 4.7" *by* K. M. Foley et al.

K. M. Foley et al.

appel.wyat@epa.gov

Received and published: 29 January 2010

We thank the referee for their constructive comments on the paper. We have addressed the questions and suggestions that were provided and we feel that these changes have improved the paper. Below we have listed our response to each of the specific points that were raised, using the page, line and Figure numbers from the paper published in GMDD. Reviewer comments are shown in italics; the author reponses are in regular text.

Major concerns

1. *The conclusion of the paper cannot be regarded as general since the West Coast States are not included in the analysis.*

We agree with the concern that the results of this study do not include an analysis of model performance for the West Coast States. Plans for the evaluation of the next CMAQ release include running the model on a continental domain. There is also an evaluation of CMAQv4.7 underway that will be based on model output for the entire US. (See http://aqmeii.jrc.ec.europa.eu/aqmeii2.htm for additional information.) We now mention this issue and the need for additional analysis of Western states in the paper.

2. *The improvements of CMAQ v4.7 on the predicted PM concentrations are generally small and might be statistically insignificant.*

The paradigm of statistical significance testing is not appropriate for this type of data. Since we are analyzing output from a deterministic model we cannot use probability theory or stochastics. In a model to model comparison there is no way to cast the two sets of model output as independent random samples from some underlying population. We cannot answer these two questions: "What is the population that we are sampling from?" and "How do we define a probability sampling scheme from this population?" As a result there is no meaningful null hypothesis in this case because there is no way that differences in the modeled concentrations are an artifact of inherent variability in a random sample. Since these are deterministic models, if we rerun the two simulations, we are guaranteed to get the same result every time.

Rather than statistical significance, what is of greater interest in this study is what is often called "practical significance". In other words, we need to identify what change in monthly average $PM_{2.5}$ would be considered "significant" or notable based on the context of how the model output will be used. This decision can be based on many things, such as experience with previous model changes, e.g. if the improvement in model bias is large compared to improvements seen in past model updates. We agree that some of the improvements in the model science have not translated into large improvements in model predictions. We have made an effort to identify, with appropriate language, areas where we believe the changes in model biases are important, i.e. the improvements have practical significance (e.g. improvements in predicted coarse

mode nitrate, pg 1264 line 27). We have also identified areas where the change in model performance is small but in the right direction (e.g. the impact of the cloud model improvements on sulfate, pg 1265 line 19-23).

Specific comments

1. *The paper reported that the modified gamma N_2O_5 decreases the bias in the simulated particulate nitrate concentrations in the eastern US in the winter. However, the absolute change in the nitrate concentrations is small due to relatively low nitrate concentrations in the eastern US in general. The evaluation of this new parameterization could have been extended to include West Coast States, such as California, where the nitrate concentrations during wintertime are more substantial.*
This is a good suggestion for future research. Although the impact of our modifications to γN_2O_5 on TNO_3 concentrations may be larger in California in terms of absolute magnitude, we expect that the relative effect will still be damped by the compensating effects of gas-phase chemistry as noted on lines 3 – 6 of page 1262.

2. *Using the "previous increment" as a label (such as Figure 2) without clearly stating what increment is included leads to confusion. This makes it difficult for readers to replicate the results discussed in the paper. Since most of the users of CMAQ are likely not aware of these incremental changes not documented in this study, why not use the unmodified CMAQ as the base case consistently throughout the text?*
In Figure 2, the "previous increment" is very similar to Increment A. This increment used the γN_2O_5 parameterization given in the body of the Davis et al. paper. In contrast, Increment A (and the released model version, CMAQv4.7) includes the equations in Appendix A of Davis et al. This change was made late in the incremental testing process in response to reviewer comments for the Davis et al. paper. We felt it would be unnecessarily confusing to explain this small difference in this paper since it has no real impact on the final conclusions of the evaluation of the SOA increment. In Figure 6, the "previous increment" is actually the HONO Increment. Reasons for not including this increment in Section 4 are given on page 1256. Rather than adding these

descriptions to the text (since the paper is already quite long) we now reference this open discussion so that interested readers will know exactly what is meant by "previous increment".

The incremental test was designed to show how the model changes build upon one another, leading to the final released version. In other words, the incremental test results have the nice property that summing all of the changes will produce the final model output. As suggested, another strategy could have been to use the unmodified CMAQ as the base case for each increment. This would also provide a systematic test of each model change but could not necessarily be used to explain the final model output that includes all the model changes together.

3. *The paper claims that seasonal SOA pattern predicted by the updated SOA module in CMAQ 4.7 is "in better agreement than v4.6 with observational estimates of SOA" (line 4, p 1263). This sentence is not supported with statistics or time series. In addition, Figures 2 and 3 are comparing "CMAQ increment B" with "previous increments", which is not CMAQ v4.6, and cannot be used to support the claim. Similarly, "Updates to the SOA module also improve diurnal patterns. . ." (line 7, p 1263) is not well supported. Figure 3 shows a regional difference but it is insufficient to support the conclusion without comparing with observations.*
On page 1263 we state that the agreement is better "qualitatively". We are simply pointing out that the observed data show higher SOA concentration in the summer than in the winter and the simulations now reflect this seasonal change. We are not trying to make any stronger statement about the model performance. Also the "previous increment" in this case is the same as the base model (CMAQ v4.6) with a change to the emissions inputs and a change to the parameterization for γN_2O_5. These changes have little to no impact on the SOA predictions and therefore a comparison between increment B and CMAQ v4.6 is appropriate in this case. Finally, the statement about the improved diurnal patterns is based on Figure 4 as stated in the text (page 1263, line 10). Figure 4 shows the change in the median of the daily TC amplitudes across

the domain. There are only two grid cells where data are available to evaluate this change and improved agreement with the observations is noted in both of those cells (see page 1263, lines 12-19).

4. *Figure 7: Presumably, the reason that authors limit the scales in the difference plots (third column) to 0.75 μg/m^3 is that there are some locations with much larger difference. What are the maximum differences and where are the locations where these maximum differences occur? Why are there significant decreases in the predicted concentrations in the southern part of Louisiana?*
The largest differences in both January and August are 1.7 μg/m^3. In January, there are only 24 grid cells with differences greater than .75 μg/m^3. Most of these are in LA with a few in VA and Canada. In August there are about 200 grid cells with differences greater than .75 μg/m^3 scattered throughout a few states (see Figure 1 included below). The largest difference of 1.7 μg/m^3 occurs in a single grid cell in Mexico. There is no clear pattern in where these largest differences occur.

The decrease noted in predicted SO_4 over southern LA is most likely caused by the modifications made to the sub-grid non-precipitating (NP) clouds. Section 2.3 describes the changes made to the NP clouds that could have impacted the formation of these clouds. We did not expect that the cloud model modifications would have a uniform change across the modeling domain. Some of the cloud model changes reduced sulfate concentrations (e.g. limiting the formation of NP clouds), while other changes increased concentrations (changing the cloud process integration to coincide with the model synchronization time-step, allowing more cycling of pollutants through clouds that increased SO_4 through aqueous production).

5. *The in-line photolysis option is useful but it should be used with care. In general, using the in-line option tends to decrease in the surface photolysis rates but could lead to increased photolysis rates in higher elevations. It will be more informative for the authors to show vertical profiles of predicted actinic flux and photolysis rates to better illustrate the effect of the in-line calculation. Some sort of evaluation of the accuracy of*

the in-line model in terms of both surface and vertical profile is also necessary.
The in-line photolysis option is included in the CMAQv4.7 release as a research/beta option for testing and experimentation. It is categorized as "beta" because additional work is needed before this option becomes the default option in the model release. We've included two vertical cross-section plots of the differences in monthly average photolysis rates for NO_2 and $O_3 \rightarrow O(^1D)$ between the in-line model and the values interpolated from JPROC (Figure 2 below). Both plots show, in general, that the largest differences are in lower model layers and in particular over higher terrain area such as the Rocky and Appalachian Mountains. Differences in NO_2 photolysis rates are much lower in the upper troposphere while the $O_3 \rightarrow O(^1D)$ rates are 10-20% lower with the in-line model. We will repeat this analysis once the inline model is updated (by incorporating temperature effect into the absorption cross section / quantum yield data and incorporating satellite O_3 column data in the radiative transfer calculations to give a better estimate of the stratospheric O_3 component). We agree that an evaluation of photolysis rates is needed to assess the accuracy of the in-line photolysis module, and have already initiated work in that direction; the results from these on-going assessments will be reported in a subsequent manuscript.

6. *The authors state that CMAQ v4.7 slightly improves the PM results in general. However, most of these differences are on the order of 0.1 μg/m^3. So, are the CMAQ v4.7 model predictions statistically different from the previous version? Compensating errors in other model parameters, such as minimum vertical diffusivities, or meteorology inputs can easily lead to much larger differences and reverse the conclusion in this paper. It is not to say that these changes are not necessary but the authors should be careful in drawing their conclusion regarding whether the improvement in model science really leading to improvement in model predictions. In addition, as mentioned in comment 1, neglecting the West Coast is a significant omission and the results here are at most only apply to the East Coast.*
The issue of statistical significance has been addressed in item 2 of the major concerns. We agree that the results presented here are limited to the spatial domain and

the time period studied. Evaluation of the next CMAQ release will utilize simulations over the continental United States for longer time periods. This will allow for analysis of the model performance in the western United States and will also provide information on the sensitivity of the model evaluation results to different meteorological conditions across different seasons. Additionally, there is an evaluation of CMAQv4.7 underway that will be based on an annual simulation for the United States and Europe. See http://aqmeii.jrc.ec.europa.eu/aqmeii2.htm for additional information.

References

Davis, J.M., Bhave, P.V., and Foley, K.M.: Parameterization of N2O5 reaction probabilities on the surface of particles containing ammonium, sulfate, and nitrate, Atmos. Chem. Phys., 8, 5925-5311, 2008.

Interactive comment on Geosci. Model Dev. Discuss., 2, 1245, 2009.

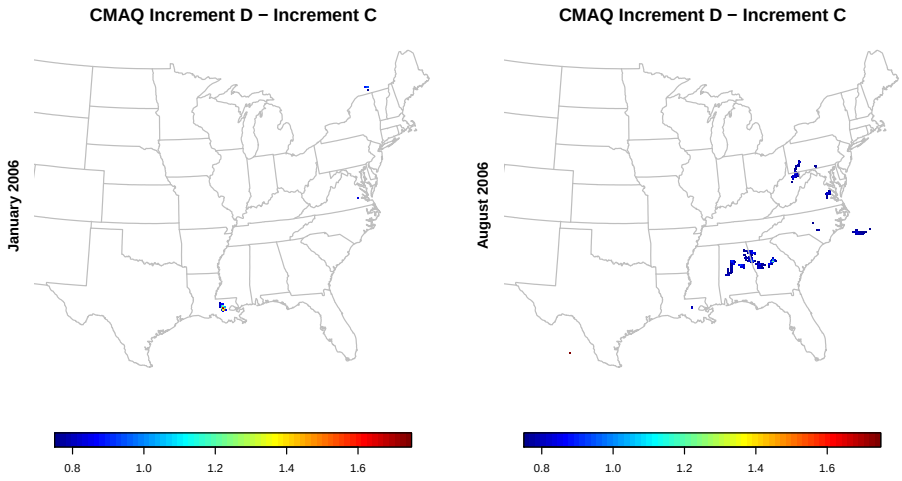

Fig. 1. Largest differences in monthly average SO4 concentrations (ug/mˆ3) for January (left) and August (right) between Increment D (Cloud Increment) and Increment C (Coarse Mode Increment).

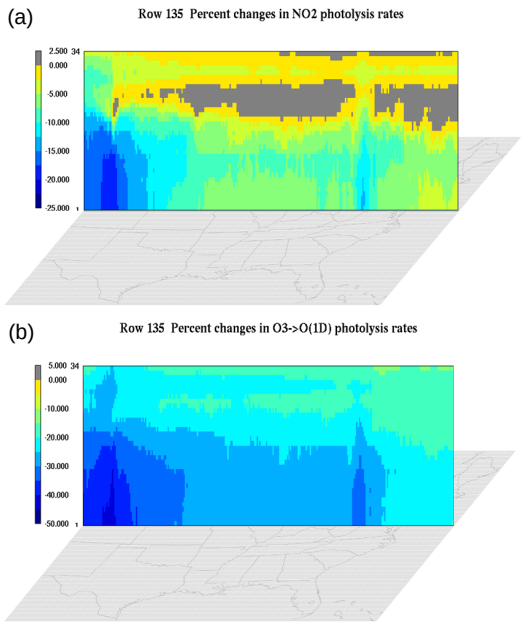

Fig. 2. Vertical cross-sections in the percent change in monthly averaged photolysis rates over daytime hours (14-24Z) for (a) NO2 and (b) O3->O(1D).