

Interactive comment on “Twelve-month, 12 km resolution North American WRF-Chem v3.4 air quality simulation: performance evaluation” by C. W. Tessum et al.

C. W. Tessum et al.

tess0050@umn.edu

Received and published: 24 February 2015

Comment: Overall I found the manuscript to be generally well written and organized.

Response: We thank the reviewer for this comment.

Comment: As is typically the case when attempting to evaluate model simulations that span large spatial domains and time periods, the difficulty becomes in summarizing the results in a meaningful way that does not overwhelm the reader with statistics and numbers. Here, the authors present annual mean performance metrics for the entire domain, along with regional/seasonal statistics. I generally don't find the annual statistics to be helpful in any way, other than perhaps initially to make sure there isn't some

C3409

huge gross error in the model results. Otherwise, bulk annual/domain-wide statistics are typically difficult to interpret due to often compensating seasonal biases (e.g. particulate nitrate is often underestimated in the summer and overestimated in the winter). To their credit, the authors do acknowledge this issue with the bulk statistics. I'm wondering if the manuscript would benefit from dropping the annual domain-wide statistics and just focus on presenting the seasonal and regional statistics. I will leave this decision to the authors, but just note that I think most readers would find a much value in the annual/domain-wide stats and would immediately focus on the seasonal/regional stats.

Response: We agree with the review that the seasonal and regional statistics are important, but we agree with Reviewer 2 that the annual/domain-wide statistics are also useful: they show an overall summary of model performance. For this reason we would like to keep the annual/domain-wide statistics in the main manuscript.

Comment: It might be nice to move some of the seasonal/regional plots for the specified PM_{2.5} components from the supplemental material to the main text.

Response: This is a good idea. We have moved the seasonal/regional scatterplots for PM_{2.5} subspecies to the main text.

Comment: Finally, the authors need to support some of their statements with references, specifically regarding difference in sampling protocols and/or analysis techniques between the different networks.

Response: In response to this comment, we have added a citation of the variability in OC analysis methods. We also fixed two inconsistencies in our processing of the data, the result of which that some of this text was no longer relevant and so was removed.

Comment: Abstract: Perhaps mention the modeling year earlier in the abstract.

Response: We added “year 2005” to the first sentence of the abstract.

Comment: Provide some examples of “contemporary models”.

C3410

Response: We thank the reviewer for this suggestion. We provide examples of contemporary models and their performance in the main text. (Adding examples in the abstract (and defining the lengthy acronyms that make up their names) would add to the word count and distract from the main messages of the abstract.) In the abstract, we changed “contemporary models” to “contemporary modeling efforts” to better reflect the comparisons that we do in the manuscript.

Comment: Again, bulk annual average statistics are not all the useful. Maybe replace these with more meaningful seasonal/regional metrics.

Response: We respect this viewpoint. As mentioned above, we feel that both types of statistics are useful.

Comment: It's a little strange to look at 24h average ozone, given the large biases that typically can occur with ozone overnight. It might be better to present a different, more meaningful metric for ozone here (e.g. daily 8hr average maximum).

Response: In response to this comment, we clarified the abstract text to state that average daytime and daily peak concentrations are more relevant for health effects and regulatory analysis, and the model performance is better for those metrics.

Comment: Page 8435, lines 13-15: It might be a little disingenuous to refer to 12-km as “finescale”. Understanding that scale is relative thing (15 years ago, 12-km was “fine-scale”), 12-km is probably better referred to as regional-scale at this point in time, considering that more and more modeling is taking place at 4-km and below.

Response: In response to this comment, we changed the text “fine-scale (12 km or better)” to “12 km or finer scale”.

Comment: Page 8436, line 21: 28 layers seems like it's on the low-end of layer structures these days. Were the computer limitations the deciding factor in going with 28 layers instead of something closer to say 40 or even 50? Do the authors feel that increasing the number of vertical layers (and in particular using the smaller first layer) would significantly impact the results?

C3411

Response: We thank the reviewer for this comment. In response, we added the text “Previous studies (e.g., Appel et al., 2012; Yahya et al., 2014) have used 34 vertical layers; our choice of 28 vertical layers represents a tradeoff between vertical grid resolution and computational expense”. We have not investigated the question of how the results would be impacted by increasing or decreasing the number of vertical layers; that issue is important and worthy of further consideration but for the present article is outside the scope of our investigation.

Comment: Page 8437, line 13: What exactly constitutes “miscellaneous PM2.5”?

Response: In response to this comment, we changed “miscellaneous PM2.5” to “unclassified PM2.5”.

Comment: Page 8438, lines 7-9: The 2008 NEI has been available for quite some time now (and 2011 NEI is now available too). It seems like 2005 is a fairly old year to simulate at this point. When the authors say that the 2005 NEI was most recent available it makes it seem like this work started a long time ago. Has it just taken that long from start to finish for this modeling exercise?

Response: We thank the reviewer for this question. This manuscript is part of a larger modeling exercise, which has taken a number of years to complete. The other part of this study was recently published here: <http://www.pnas.org/content/111/52/18490.abstract>

Comment: Page 8439, Line 23: A 50-60 meter first layer height seems quite large, especially since nighttime boundary layers can often reach 50m or below. What impact do the authors feel there is from having such a deep first layer?

Response: Testing the impact of the number of layers on model performance is outside of the scope of this study. We note in the Discussion that the investigation of model parameters on performance is an important area for future research.

Comment: Page 8445, Lines 9-10: Exactly what differences are there between the network measurement techniques and why would they result in such larger differences?

C3412

IMPROVE sites are rural, so perhaps background SO₂/SO₄ is greatly overestimated.

Response: In response to this comment, we revisited the measurement data documentation and found that the IMPROVE network reports elemental sulfur concentrations instead of SO₄ concentrations. Adjusting our calculations to account for this decreases the differences between measurement networks for SO₄.

Comment: Page 8445: First, the authors state a MFB = -110

Response: We thank the reviewer for calling this to our attention. -110

Comment: The nitrate biases reported are really large. Do the authors have any explanation as to why nitrate is underpredicted by so much (especially in the west where nitrate makes up a greater percentage of the total PM_{2.5} than in the east)?

Response: Particulate nitrate formation is strongly temperature dependent, and as we discuss in the article, many model performance evaluations only cover the summer months. We state in the article that nitrate predictive performance is better in the summer than in the winter. In Table A2 we compare our results to another full-year, contiguous U.S. modeling simulation. Predictive performance for nitrate in that study is similar to our results.

Comment: Page 8446: The OC underestimation at CSN sites is really large too. How is it that the differences don't appear to be rural vs. urban, since the urban CSN sites have an OC MFB = -113

Response: In response to this comment we reviewed our calculations and found and fixed a configuration error which was partially responsible for the difference between networks. As noted in the text, figure A12 shows that the difference between networks is similar when considering only urban vs. only rural locations. We have added a reference that discusses the variability in measured values of OC when using different analysis methods, which can be up a factor of 5.

Comment: Page 8446, Line 26: Change "lower" to "worse".

C3413

Response: We thank the reviewer for this comment, but we think that he or she may have misinterpreted Table A2. We state that for most pollutants and networks, NME is lower in our study than in Yahya et al. The numbers in Table A2 support this statement. Since lower error is generally considered to be better than higher error, we don't feel that it would be appropriate to change "lower" to "worse". To clarify this, we changed the text to "lower (i.e., better)".

Comment: Table A2: Are these annual values being reported?

Response: We edited the title of Table A2 to clarify that the values are for annual average performance.

Interactive comment on Geosci. Model Dev. Discuss., 7, 8433, 2014.

C3414