Atmos. Chem. Phys. Discuss., 14, C6672–C6676, 2014
www.atmos-chem-phys-discuss.net/14/C6672/2014/

Atmospheric Chemistry and Physics
Discussions

Interactive comment on "Long-term particulate matter modeling for health effects studies in California – Part 1: Model performance on temporal and spatial variations" *by* J. Hu et al.

Anonymous Referee #2

Received and published: 8 September 2014

The manuscript by Hu et al., is what appears to be the first in a series of papers to use an air quality model, in this case the "UCD/CIT" model, to develop PM exposure indices for health studies. This one deals with model evaluation. The model domain covers California and some surroundings. They find that while PM2.5 mass is reasonably accurately simulated, nitrate, organic carbon and sulfate are on the low side, while dust is high. They then conduct sensitivity analyses to help explain the issues.

General Comment: The use of air quality models in health studies is a growing trend, though one should tread cautiously. As this paper shows, there can be large biases involved (which is likely a bigger concern than the errors), particularly when those

biases are not thoroughly investigated. In this case, organic carbon, sulfate and nitrate are biased low. As noted below, much of the issue is laid to the emissions, but the modeling approach may have concerns here as well.

Specific Comments: Their sensitivity analysis is not well motivated and done in a rather cursory fashion. They suggest that the reason for the low nitrate is that the RH is low from WRF. That may be the case, in which case one should figure out why RH is low in WRF. Instead, they raise the RH by 30%. The nitrate then increases. It is suggested that they find out why the base meteorological model is providing biased results, and have it corrected in a fashion that will capture linkages important to the meteorology. For one, if you increase RH, shouldn't you get more rain? Arbitrarily increasing RH as an output does not provide this natural link. Instead, they cap the RH at 95%. More rain will lower, not increase nitrate. Thus, their approach is a bit one-sided. How well did WRF simulate rain? So far as issues with the meteorological model, they also increased the friction velocity by 50% to decrease an overestimate of wind speed, but leading to a negative bias in wind speed. Whilst they can cite two of their own papers, this would strike me as to atypical practice or everyone would follow suit. Is this generally accepted? Further, given it leads to a negative bias, is this not too much? Also, might the RH problem be linked to this? Should they use a different meteorological model?

They suggest that the poor sulfate results are due to the uncertainties in the sulfur dioxide emissions. My understanding is that air quality models in the US are best at sulfate because the sulfur oxide emissions are well characterized as they arise from fuels, and fuel usage and composition are well known, or the emissions are measured. Here, they say the emissions are uncertain. They should provide a better justification for this conclusion. Are such emissions very uncertain and why? What other demonstration of this is there, particularly given the findings from other studies? Also, a low sulfate might increase the nitrate, but the nitrate is also low.

Given it is a long term simulation, a question arises as to how well does their model simulate deposition over that same period. I have seen evaluations of other air quality models for deposition. A publication of such an evaluation for the model they have chosen would be of interest. They should provide an update here if possible. This should be done, particularly, considering the modification to the friction velocity and the low RH bias, which probably means a low bias in wet deposition, which should lead to higher pollutant concentrations.

They suggest that many of the problems are due to emissions. This should be better demonstrated. One can get very similar problems if some other issue is driving the problems. What if some other process is not being captured correctly? For example, might it be that the model is dispersing material too rapidly/not rapidly enough? I would have liked to see more support using other approaches, e.g., from recent tunnel studies or satellite data. Their paragraph, starting "Figure 5b and c..." says things work in some places and times, but not other places and times. They conclude the emissions are uncertain, which is a potential cause. The model may also lead to those same biases due to the parameterizations of physical processes, and, indeed, one might find the latter a more likely happenstance. Given the complex meteorological situation in the area, it could be quite challenging for a model to correctly capture dispersion. They hint at this possibility, but do not give it the importance it deserves, and how it might be addressed, or how it might impact the use of model results.

The top boundary of the model is of concern. It is only 5000 m. For such long simulations, and in such complex topography, aren't there periods where various processes might occur that lead to exchange above 5000 m, e.g., convective storms, atmospheric waves over the mountains in the domain? Might stratospheric intrusion play a role?1,2 They should conduct a sensitivity study to assess how the lower model boundary condition impacts their results. It would also seem that the low model height might negate any modeled role of air craft or lightning emissions.

They say "For the first time, a $\sim$ decadal..." I do believe that the US EPA has conducted air quality modeling over the whole US, which includes California, for a fairly long time

period and may have used that for health analyses. I am not sure to what degree they have used their results for health studies, but that should be checked. Even if they have not, the "for health effects studies" is not that relevant in this case since this really is a model evaluation, and I am not sure if the "for health effects studies" changed the model evaluation analysis appreciably. How does this effort differ from any other evaluation of a long term model application? Also, the whole domain they use is not 4 km, just a subset, so they should alter their lead-in sentence, e.g., add "with populated regions modeled at a resolution as fine as 4 km".

Not sure if "These results will be improved in future studies." belongs in an Abstract, or even in the paper. It begs the question, why aren't they improved here? Is this paper premature? Also, it does not appear in the paper.

In the end, too much of the less than desired agreement with observations is put to the emissions and the meteorological model as opposed to potential issues with the model itself. Further, they need to provide some idea of what should be done. In terms of model evaluation, they should look at the European AQMEII effort.3 (They might also look at other model-based air quality model-health studies in Europe and elsewhere.4,5) The US EPA has looked at model evaluation under emission uncertainty.6 I would suggest they look at those efforts a bit more. They should consider raising their model domain height given the complexity of the terrain and the potential for longer range transport, convective storms and stratospheric intrusion to be of an issue. They should try to figure out why their meteorological model provides inadequate inputs in its base formulation, e.g., the need to slow winds down and increase RH.

1. McDonald-Buller EC, Allen DT, Brown N, Jacob DJ, Jaffe D, Kolb CE, Lefohn AS, Oltmans S, Parrish DD, Yarwood G, Zhang L. Establishing policy relevant background (prb) ozone concentrations in the united states. Environ Sci Technol 2011;45:9484-9497. 2. Lefohn AS, Wernli H, Shadwick D, Oltmans SJ, Shapiro M. Quantifying the importance of stratospheric-tropospheric transport on surface ozone concentrations at high- and low-elevation monitoring sites in the united states. Atmos Environ

2012;62:646-656. 3. Hogrefe C, Roselle S, Mathur R, Rao ST, Galmarini S. Space-time analysis of the air quality model evaluation international initiative (aqmeii) phase 1 air quality simulations. J Air Waste Manage Assoc 2014;64:388-405. 4. Beevers SD, Kitwiroon N, Williams ML, Kelly FJ, Anderson HR, Carslaw DC. Air pollution dispersion models for human exposure predictions in london. Journal of Exposure Science and Environmental Epidemiology 2013;23:647-653. 5. Kelly F, Anderson HR, Armstrong B, Atkinson R, Barratt B, Beevers S, Derwent D, Green D, Mudway I, Wilkinson P, Committee HEIHR. The impact of the congestion charging scheme on air quality in london. Part 1. Emissions modeling and analysis of air pollution measurements. Research report (Health Effects Institute) 2011:5-71. 6. Napelenok SL, Foley KM, Kang DW, Mathur R, Pierce T, Rao ST. Dynamic evaluation of regional air quality model's response to emission reductions in the presence of uncertain emission inventories. Atmos Environ 2011;45:4091-4098.

Interactive comment on Atmos. Chem. Phys. Discuss., 14, 20997, 2014.