Review to acp-2021-946: "Limitations in representation of physical processes prevent successful simulation of $PM_{2.5}$ during KORUS-AQ" By Travis et al.

This paper presented a detailed model study of the aerosol composition in Korean during the NASA KORUS-AQ aircraft campaign. The study tried to improve the model simulated PM2.5 and use it to quantify the contribution from long range transport and local emission to the observed PM composition. The topic is relevant to ACP and study is well designed and presented. I suggest ACP publish it with some clarification and improvement.

Science related:

(1) Nitrate, sulfate, and ammonium often show a dynamic and complicated balance. Maybe the analysis can focus more on total sulfate+nitrate+ammonium, where overall model shows better agreement with the observations. Your conclusion on missing local source during the so called 'Transport/Haze' period is probably still hold.

(2) How is the model vs observation comparison for SO2? Only SOR is showed in Figure 12. How is the direct comparison of SO2, esp. in different test runs? SO2 is the primary pollutant and precursor of sulfate, its performance should be relevant.

(3) Heterogenous reaction of SO2 to sulfate has been proposed and studied for long time. Early work should be acknowledged.
Summarized by: Ravishankara, Science, Heterogeneous and Multiphase Chemistry in the Troposphere, 1997.
Initial work: Chameides & Davis, The free radical chemistry of cloud droplets and its impact upon the composition of rain, JGR, https://doi.org/10.1029/JC087iC07p04863, 1982,
Application in the CTM: P. Kasibhatla,W. L. Chameides,J. St. John, A three-dimensional global model investigation of seasonal variations in the atmospheric burden of anthropogenic sulfate aerosols, https://doi.org/10.1029/96JD03084, 1997

(4) Based on Figure 11, the 'Increased nighttime mixing' is not necessary improved the model performance systematically. On 05-23 and 05-24 nighttime improvement can been seen for NO2 and ozone. But for other days, the blue line compared worse to observations: phase shift for NO2 on 05-27, 05-28, 05-29 nighttime, and for Ozone on 05-26, 05-27, 05-28, 05-29 'Increased nighttime mixing' degraded the performance.

(5) Please list all model experiments you have done in a table. Quite some model experiments results are discussed, e.g. 'Model', 'Increased nighttime mixing', 'HetSO2', 'No nighttime production', 'Old wet scavenging', '5x dry deposition', etc. What is the difference between those experiments? There is only one thing different between those runs vs the control run, or they are accumulated? 5x dry deposition is global? Or just for urban area or certain particular landuse? Test runs are for both coarse and fine resolution?

(6) Section 5.3 discussion can be shortened and focused on the new findings from this analysis.

As you have shown, even just for the inorganic aerosol components, there are many parameters you can adjust to move some aspect of the model simulation closer to observations. Overall, the model sulfate+nitrate+ammonium show reasonable agreement with observations. If your purpose is to significantly improve the nitrate and sulfate simulation, more evaluations are required, including other region and time.

Model related:

(0) GEOS-Chem has been used in many studies, including multiple aircraft campaigns. It typically shows good performance. It seems the main GEOS-Chem development team is not part of this paper. Have you discussed your findings with the GEOS-Chem development group? 5 times increase of dry deposition and adding a new chemical pathway of SO2 to sulfate seem quite significant changes for a mature model like GEOS-Chem.

(1) Model limitation. This study used a global CTM GEOS-Chem. The model resolution, 0.25x0.3125 is high for a global model, while it might not be the best choice for urban pollution study. Based on Figure 1, most of the ground stations probably are within one grid point of the model. What is the landuse and topography of that model point? What is the reported setting (landuse and elevation) of the stations showed in Figure 1? Do they generally agree? Since GOES-Chem should be able to capture the chemical and physical processes on the regional scale, maybe the model-observation comparison should focus on sites that are representative over the regional conditions?

(2) Model set up: this is a nested simulation, right? From 2x2.5 to 0.25x0.3125? Nesting is a good idea but based on experience with nested model like WRF/WRF-Chem the spatial resolution of ~ 1:3 or 1:4 will provide a smoother and more consistent transition between outer and inner domains. 1:8 is probably too much, for both dynamic and chemical processes.

Other: Figure 12, ‘from aircraft below 1km for the same domain as Fig. 3.’ Figure 3 didn’t show domain info about the aircraft.