We acknowledge the referees for their insightful comments. Please find our responses to referees' comments in blue.

Report #1 from Referee #2

The revised manuscript well addressed most of my concerns. I have a few minor comments.

1. Line 62, Page 2. The authors may add some pertinent studies for high-density mapping of $PM_{2.5}$ over East Asia (e.g., Chen et al., 2019).

Added.

2. Line 84, Page 3. The statement of "little study" may be revised. In fact, there have been many studies in East Asia, as reviewed in Li et al., 2017.

We deleted the sentence. We added reference "Li et al., 2017" and two publications cited by "Li et al., 2017", and revised lines 84-86 to: "Airborne measurements of aerosol vertical profiles (without species information) in East Asia are limited (Zhang et al., 2006; Liu et al., 2009; Zhang et al., 2009; Sun et al., 2013; Li et al., 2017), and speciated vertical profiles are rarer."

References:

Zhang, Q., Zhao, C., Tie, X., Wei, Q., Huang, M., Li, G., Ying, Z., and Li, C.: Characterizations of aerosols over the Beijing region: A case study of aircraft measurements, Atmos. Environ., 40, 4513-4527, https://doi.org/10.1016/j.atmosenv.2006.04.032, 2006.

Zhang, Q., Ma, X., Tie, X., Huang, M., and Zhao, C.: Vertical distributions of aerosols under different weather conditions: Analysis of in-situ aircraft measurements in Beijing, China, Atmos. Environ., 43, 5526-5535, https://doi.org/10.1016/j.atmosenv.2009.05.037, 2009.

3. Line 89, Page 3. The authors may revise the statement. First, this study mainly focused on the $PM_{2.5}$ rather than coarse PM. Second, multiple studies have investigated the contribution of coarse aerosol to AOD, which also can be represented by fine mode fraction (Eck et al., 2010).

We clarified lines 89-91 to: “Finally, although there are studies on the optical depth of coarse mode desert dust (Eck et al., 2010; Ridley et al., 2016), there has been to our knowledge no study of how coarse anthropogenic PM may contribute to the AOD measurements.”

Reference:

Ridley, D. A., Heald, C. L., Kok, J. F., and Zhao, C.: An observationally constrained estimate of global dust aerosol optical depth, Atmos. Chem. Phys., 16, 15097-15117, 10.5194/acp-16-15097-2016, 2016.

4. Section 3.1. The authors define 0-2 km as the average PBL. Note that the average daily maximum PBL depth is around 2 km. PBL during the morning and evening is much lower. Does Figure 1 only present the noontime results or the whole day average?

We have added ‘aircraft’ in line 219 and added a line 220: “The KORUS-AQ aircraft sampled during the daytime, mainly between 9 am and 3 pm local time.”

Line 335-337 were revised to: “This is because the model bias is mainly driven by nighttime conditions (Figure 5), while aircraft samples in the daytime during KORUS-AQ.”

5. Table 1. The authors may briefly explain why choose these sites, since there are many other PM stations over East Asia.

We have changed ‘North China’ to ‘East China’ everywhere in Section 2. The number of sites in East China is changed to 598 in Table 1. We in addition added in Section 5 in lines 363-366: “The Figure gives normalized mean biases (*NMB*s) relative to ground-based measurements from AERONET and from the $PM_{2.5}$ surface networks (shown as circles) over the North China region (115.5-122° E, 34.5-40.5° N) and South Korea. The North China region is defined to overlap with the domain of the geostationary satellite AOD, and to ensure consistent seasonal variations within its narrow latitude.”

References:

Eck, T.F., Holben, B.N., Sinyuk, A., Pinker, R.T., Goloub, P., Chen, H., Chatenet, B., Li, Z., Singh, R.P., Tripathi, S.N. and Reid, J.S., 2010. Climatological aspects of the optical properties of fine/coarse mode aerosol mixtures. Journal of Geophysical Research: Atmospheres, 115(D19).

Li, Z. J. Guo, A. Ding, H. Liao, J. Liu, Y. Sun, T. Wang, H. Xue, H. Zhang, and B. Zhu, 2017: Aerosols and boundary-layer interactions and impact on air quality, Natl. Sci. Rev., 4, 810-833, doi:10.1093/nsr/nwx117.

Chen, J., Yin, J., Zang, L., Zhang, T., and Zhao, M.: Stacking machine learning model for estimating hourly $PM_{2.5}$ in China based on Himawari-8 aerosol optical depth data, Sci. Total Environ., 697, 134021, https://doi.org/10.1016/j.scitotenv.2019.134021, 2019.

Report #3 from Referee #1

Line 73: Better to say artificial intelligence algorithms and involve some literature using deep learning models. Also, a typical study focusing on hourly PM estimates using Himawari-8 (https://doi.org/10.5194/acp-21-7863-2021) should be considered.

Changed to “artificial intelligence algorithms”. Wei et al. 2021a is cited in line 62 and line 75.

Figure 6: I would like to see the comparison in spatial patterns between satellite and model $PM_{2.5}$ observations in a specific region. Maybe you can use the public CHAP dataset for a simple comparison in Eastern China, available at: https://doi.org/10.5281/zenodo.4660858.

We believe that it would be misleading and detrimental to do this because of our point in the paper that models such as GEOS-Chem need to fix their overestimate of nitrate and the underestimate of coarse PM before attempting to produce a $PM_{2.5}$ product (lines 470-473 in the manuscript) – we don’t actually show AOD-inferred $PM_{2.5}$ anywhere. We are now working to improve nitrate and coarse PM in GEOS-Chem, and after we do, we intend to generate a $PM_{2.5}$ product using GEOS-Chem combined with GOCI. Showing a flawed product at this stage would be counterproductive.

Additional revisions

1. We updated Jun Meng’s affiliation.

2. We changed the color bars in Figure 4, Figure 6, and Figure S7, and the red colors in Figure 7.