We acknowledge the referees for their insightful comments. We have made efforts to improve the manuscript accordingly. Please find our responses to referees' comments in blue. Newly added references are listed at the end of this document.

RC1 by Referee #3

This paper attempts to understand the relationship between AOD and $PM_{2.5}$. However, after reading through, I feel that the paper is more of a GEOS-Chem validation and uncertainty analysis work, rather than offering physical explanation of the AOD-$PM_{2.5}$ relationship.

The title has been revised to: "Relating geostationary satellite measurements of aerosol optical depth (AOD) over East Asia to fine particulate matter ($PM_{2.5}$): insights from the KORUS-AQ aircraft campaign and GEOS-Chem model simulations".

We have rephrased lines 34-38 in the abstract: "Geostationary satellite measurements of aerosol optical depth (AOD) over East Asia from the GOCI and AHI instruments can augment surface monitoring of fine particulate matter ($PM_{2.5}$) air quality, but this requires better understanding of the AOD-$PM_{2.5}$ relationship. Here we use the GEOS-Chem chemical transport model to analyze the critical variables determining the AOD-$PM_{2.5}$ relationship over East Asia by simulation of observations from satellite, aircraft, and ground-based datasets."

We added lines 67-68: "This enables us to identify critical variables and uncertainties for inferring $PM_{2.5}$ from satellite AOD data."

We rephrased lines 433-439 in the conclusions section: "Geostationary satellite observations of aerosol optical depth (AOD) over East Asia may usefully complement $PM_{2.5}$ air quality networks if the local relationship between AOD and $PM_{2.5}$ can be inferred from a physical and/or statistical model. Here we analyzed the ability of the GEOS-Chem chemical transport model to provide this relationship by using a new fused GOCI/AHI geostationary satellite product together with AERONET ground-based AOD measurements, aerosol vertical profiles over South Korea from the KORUS-AQ aircraft campaign (May-June 2016), and surface network observations. This allowed us to identify the critical features and limitations of the model for successful representing the AOD-$PM_{2.5}$ relationship."

Specifically, could the authors clarify, perhaps with additional analysis, how different factors, such as PBL height, RH, organic matter fraction, etc, contribute to the uncertainty in AOD-$PM_{2.5}$ relationship? How does the role of each factor vary with region (e.g., Korea vs. China)? The only clear conclusion is that AOD and $PM_{2.5}$ have reversed seasonality because of seasonally varying PBL height, but this is already well known.

We quantified in the abstract (lines 43-47): "We updated SNA and organic aerosol size distributions in GEOS-Chem to represent aerosol optical properties over East Asia by using in-situ measurements of particle size distributions from KORUS-AQ. We find that SNA and organic aerosols over East Asia have larger size (number median radius of 0.11 μm with geometric standard deviation of 1.4) and 20% larger mass extinction efficiency as compared to aerosols over North America (default setting in GEOS-Chem)."

We quantified in lines 286-287: "The model underestimates extinction coefficients by 20% below 1 km altitude, leading to a 10% underestimate of aircraft inferred AOD, although there is no such underestimate in aerosol mass."

We quantified in lines 323-324 to: "Therefore, about half of the GEOS-Chem underestimate of total AOD can be attributed to missing coarse PM, with the other half comes from negative RH bias."

We added a Figure 7 and lines 399-402: "The correlations of these three pairs are similar over South Korea and North China, except that GEOS-Chem overestimates springtime $PM_{2.5}$ in South Korea but not over North China, possibly due to a model overestimate of the long-range transport of $PM_{2.5}$ from China to South Korea in spring."

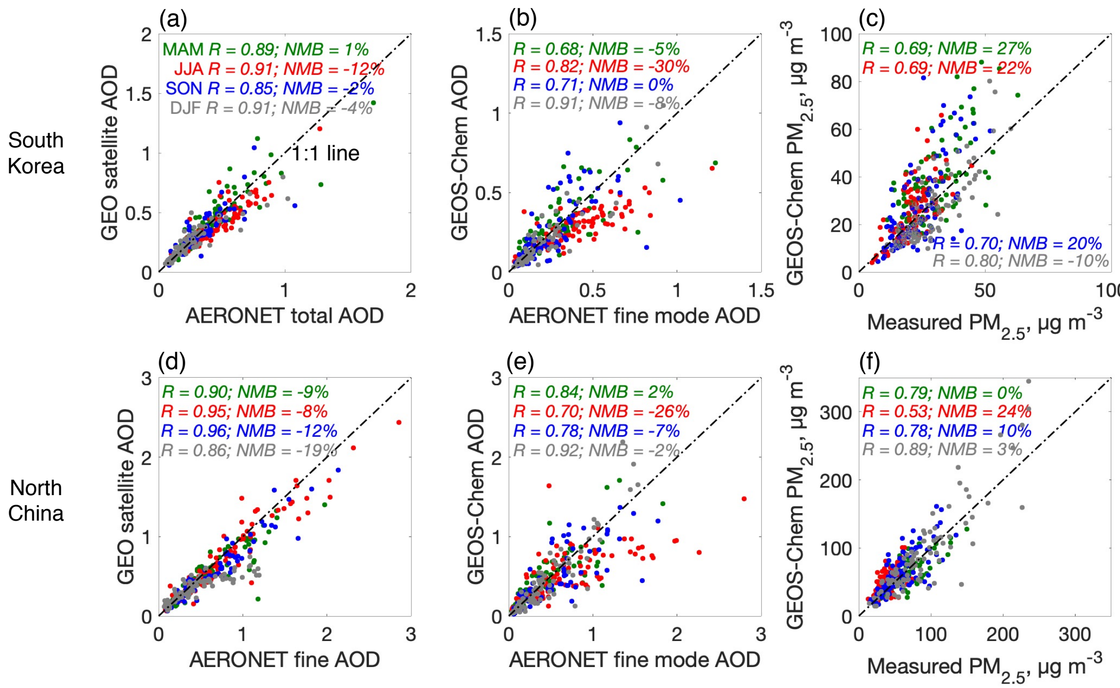

Figure 7. Scatter plots of regional mean daily (a and d) GEO satellite AOD vs. AERONET total AOD, (b and e) GEOS-Chem AOD vs. AERONET fine-model AOD, and (c and f) GEOS-Chem PM2.5 vs. measured PM2.5 over South Korea (a-c) and North China (d-f). Different colors represent different seasons. Values inset are correlation coefficients (R) and normalized mean biases (NMB) between surface measurements and GEO satellite or GEOS-Chem values.

We quantified in lines 412-414: "The GEOS-Chem AOD is ~ 20% biased low in summer and this is largely due to a low RH bias (Figure S8), as seen previously in the KORUS-AQ comparisons but amplified by the high RH in summer that drives hygroscopic growth (Latimer and Martin, 2019)."

We added analysis in lines 420-423: "GEOS-FP daytime PBL height also shows a stronger seasonality over North China than over South Korea (Figure S8), generally consistent with the CALIPSO daytime PBL height (Su et al., 2018). Previous studies have shown opposite seasonality between MODIS AOD and surface $PM_{2.5}$ over North China and attributed this to the seasonality in PBL height and RH (Qu et al., 2016; Xu et al., 2019)."

We quantified in lines 443-446 in the conclusion section: "We find that GEOS-Chem aerosol optical properties based on measurements over the North America (default model setting) underestimate KORUS-AQ aerosol mass extinction efficiency by around 20%. In addition, a low bias in GEOS-FP RH below 1 km leads to a 10% underestimate of AOD inferred from the aircraft profile."

We added lines 451-452 in the conclusions section: "The remaining 15% underestimate of AERONET fine-mode AOD by GEOS-Chem can be attributed to the RH low bias."

We quantified in lines 460-462 in the conclusions section: "GEOS-Chem simulates successfully the seasonality of measured $PM_{2.5}$ but is ~ 20% biased low in summer for AOD, due again to RH low bias like that during KORUS-AQ, amplified by the high RH in summer that drives hygroscopic growth (Latimer and Martin, 2019)."

RC2 by Referee #2

The manuscript investigated the physical relationships between AOD and $PM_{2.5}$ over East Asia by using the model simulation and comprehensive observation. The results indicate that the aerosols over this region are largely contributed by the sulfate-nitrate-ammonium and organic aerosols within the PBL. Meanwhile, the dust in the free troposphere also has an important contribution to column AOD. The seasonality of AOD and $PM_{2.5}$ has been specifically discussed. In general, this paper is well-written with a good logical connection. Thus, I recommend the manuscript for publication in Atmospheric Chemistry and Physics, after addressing the following comments.

Specific Comments:

1. The current introduction section may be insufficient to demonstrate the significance of this paper. The authors need to clearly explain the limitation of previous studies and the advantage of this study.

We revised the whole introduction:

"$PM_{2.5}$ (particulate matter with aerodynamic diameter less than 2.5 μm) in surface air is a severe public health concern in East Asia, but surface monitoring networks are too sparse to thoroughly assess population exposure. Satellite observations of aerosol optical depth (AOD) can provide a valuable complement (Van Donkelaar et al., 2015). Geostationary satellite sensors, including the Geostationary Ocean Color Imager (GOCI) launched by the Korea Aerospace Research Institute (KARI) in 2011 (Choi et al., 2016, 2018, 2019) and the Advanced Himawari Imager (AHI) launched by the Japanese Meteorological Agency (JMA) in 2014 (Lim et al., 2018, 2021), offer the potential for high-density mapping of $PM_{2.5}$ over East Asia. However, more confidence is needed in relating AOD to $PM_{2.5}$. Here we evaluate the capability of the GEOS-Chem chemical transport model (CTM) to simulate AOD-$PM_{2.5}$ relationships over East Asia, exploiting in-situ aircraft measurements of vertical aerosol profiles and optical properties from the joint NASA-NIER Korea - United States Air Quality (KORUS-AQ) field study in May-June 2016 (Crawford et al., 2021; Peterson et al., 2019; Jordan et al., 2020) together with GOCI/AHI geostationary satellite data and surface measurement networks. This enables us to identify critical variables and uncertainties for inferring $PM_{2.5}$ from satellite AOD data.

A number of past studies have used satellite AOD data to infer surface $PM_{2.5}$ using physical and statistical models. The standard geophysical approach has been to use a CTM, such as GEOS-Chem, to compute the $PM_{2.5}$/AOD ratio (Liu et al., 2004; van Donkelaar et al., 2006; van Donkelaar et al., 2015; Xu et al., 2015; Geng et al., 2017), with recent applications correcting for CTM biases using available $PM_{2.5}$ surface network data (Brauer et al., 2016; Van Donkelaar et al., 2016; van Donkelaar et al., 2019; Hammer et al., 2020). An alternative approach is to use machine-learning algorithms to relate satellite AOD to $PM_{2.5}$ by training on the surface network data (Hu et al., 2017; Chen et al., 2018; Xiao et al., 2018; Wei et al., 2021; Pendergrass et al., 2021), and sometimes including CTM values as predictors (Di et al., 2019; Xue et al., 2019). Yet another approach is to assimilate the satellite-measured AODs in a CTM and correct in this manner the $PM_{2.5}$ simulation, although this requires attribution of model AOD errors to specific model parameters (Kumar et al., 2019; Saide et al., 2014; Sekiyama et al., 2010; Cheng et al., 2019). In all of these approaches, a better physical understanding of the AOD-$PM_{2.5}$ relationship as simulated by CTMs can greatly enhance the capability to infer PM¬2.5 from AOD data.

AOD measures aerosol extinction (scattering and absorption) integrated over the atmospheric column, so that its relationship to 24-hr average surface $PM_{2.5}$ (the standard air quality metric) depends on the aerosol vertical distribution and optical properties, ambient relative humidity (RH), diurnal variation of $PM_{2.5}$, and contribution from coarse particulate matter to AOD. Little study of these factors has been conducted for East Asia. Airborne measurements of aerosol vertical profiles in East Asia are very limited (Liu et al., 2009; Sun et al., 2013). AOD is highly sensitive to RH (Brock et al., 2016; Latimer and Martin et al., 2019; Saide et al., 2020), but the impact from RH uncertainty on AOD simulation lacks evaluation. In addition, because the AOD is a daytime measurement that needs to be related to 24-h average $PM_{2.5}$, the diurnal variation of $PM_{2.5}$ needs to be understood (Guo et al., 2017; Lennartson et al., 2018). Finally, there has been to our knowledge no study of how coarse anthropogenic PM may contribute to the AOD measurements. Coarse anthropogenic PM (distinct from desert dust) is known to be high over East Asia (Chen et al., 2015; Dai et al., 2018)."

We revised lines 191-195: "Therefore, we re-computed the diagnostic AOD using updated log-normal size distributions for SNA and organic aerosol with number median radius $R_{N,med}$ = 0.11 μm and geometric standard deviation σ = 1.4 based on KORUS-AQ observations, instead as compared tof $R_{N,med}$ = 0.058 μm and σ = 1.6 in the standard model version 12.7.1, which is derived from IMPROVE network measurements of aerosol mass scattering efficiency over North America (Latimer and Martin, 2019)."

2. The analyses of this study are closely associated with the model simulation of GEOS-Chem, while the title only mentioned the observations. There are some disconnections between the title and the main text.

The title has been revised to: "Relating geostationary satellite measurements of aerosol optical depth (AOD) over East Asia to fine particulate matter ($PM_{2.5}$): insights from the KORUS-AQ aircraft campaign and GEOS-Chem model simulations".

3. Line 203, Page 8. The PBL varies significantly during the different periods. It is risky to define the 0-2 km as the PBL. The authors should give more justifications for this definition.

We rewrote line 221 to: "…, which we define as the average planetary boundary layer (PBL) during KORUS-AQ, …"

We have line 232: "KORUS-AQ aerosol component profiles for different meteorological regimes is presented in Park et al. (2021)."

4. The seasonality of AOD and $PM_{2.5}$ and its association with PBLH have been discussed previously (e.g., Guo et al., 2017; Su et al., 2018). I suggest the authors acknowledge these works.

References:

Su, T., Li, Z. and Kahn, R., 2018. Relationships between the planetary boundary layer height and surface pollutants derived from lidar observations over China: regional pattern and influencing factors. Atmospheric Chemistry and Physics, 18(21), pp.15921-15935.

Guo, J., Xia, F., Zhang, Y., Liu, H., Li, J., Lou, M., He, J., Yan, Y., Wang, F., Min, M. and Zhai, P., 2017. Impact of diurnal variability and meteorological factors on the $PM_{2.5}$-AOD relationship: Implications for $PM_{2.5}$ remote sensing. Environmental Pollution, 221, pp.94-104.

Reply: We cited Su et al. (2018) in two places:

Lines 420-422: "GEOS-FP daytime PBL height also shows a stronger seasonality over North China than over South Korea (Figure S8), generally consistent with the CALIPSO daytime PBL height (Su et al., 2018)."

Lines 472-474: "Besides the factors discussed in this study, topography might be another important factor influencing surface $PM_{2.5}$ and its vertical mixing (Su et al., 2018), and this also requires future investigation."

We cited Guo et al. (2017) at lines 87-89: "In addition, because the AOD is a daytime measurement that needs to be related to 24-h average $PM_{2.5}$, the diurnal variation of $PM_{2.5}$ needs to be understood (Guo et al., 2017; Lennartson et al., 2018)."

RC3 by Referee #1

The relationships between AOD and PM in East Asia are discussed by using ground-based and aircraft observations, but the whole study focus on direct validation and comparison, lacking in-depth analysis and literature support. In addition, there may be some problems in the use of satellite data. I suggest that the authors add more analysis to enrich the study.

Thank you for the insightful comments. In addition to the added analysis illustrated in the responses below, a bunch of references (listed at the end of this document) have been added in the introduction section.

Abstract: Line 35: Himawari-8/AHI provides AOD products at 500 nm, not 550 nm.

Reply: We have deleted 'at 500 nm' in line 35 in the abstract.

Meanwhile, we detailed in lines 122-126: "Geostationary satellite AOD at 550 nm are retrieved by the Yonsei Aerosol Retrieval (YAER) algorithm for the GOCI (Choi et al., 2016, 2018) and AHI (Lim et al. 2018) instruments, with GOCI covering East China and South Korea and AHI covering the broad East Asia region. AOD from GOCI and AHI have a 6 km × 6 km spatial resolution and 1-hour (GOCI) to 2.5-minute (AHI) temporal resolution for 8 hours per day (09:30 to 16:30 local time)."

Introduction

It is too short and the authors are suggested to summarize previous studies on investigating the relationships between $PM_{2.5}$ and AOD, especially those focusing on Asia.

In addition, studies on PM estimation from satellite AOD products need to be summarized, especially those using geostationary satellites.

Finally, the author should highlight the innovation and difference between the current study and previous related studies, and discuss the importance of understanding the relationships between PM and AOD in these studies.

The revised introduction is pasted as below:

"$PM_{2.5}$ (particulate matter with aerodynamic diameter less than 2.5 μm) in surface air is a severe public health concern in East Asia, but surface monitoring networks are too sparse to thoroughly assess population exposure. Satellite observations of aerosol optical depth (AOD) can provide a valuable complement (Van Donkelaar et al., 2015). Geostationary satellite sensors, including the Geostationary Ocean Color Imager (GOCI) launched by the Korea Aerospace Research Institute (KARI) in 2011 (Choi et al., 2016, 2018, 2019) and the Advanced Himawari Imager (AHI) launched by the Japanese Meteorological Agency (JMA) in 2014 (Lim et al., 2018, 2021), offer the potential for high-density mapping of $PM_{2.5}$ over East Asia. However, more confidence is needed in relating AOD to $PM_{2.5}$. Here we evaluate the capability of the GEOS-Chem chemical transport model (CTM) to simulate AOD-$PM_{2.5}$ relationships over East Asia, exploiting in-situ aircraft measurements of vertical aerosol profiles and optical properties from the joint NASA-NIER Korea - United States Air Quality (KORUS-AQ) field study in May-June 2016 (Crawford et al., 2021; Peterson et al., 2019; Jordan et al., 2020) together with GOCI/AHI geostationary satellite data and surface measurement networks. This enables us to identify critical variables and uncertainties for inferring $PM_{2.5}$ from satellite AOD data.

A number of past studies have used satellite AOD data to infer surface $PM_{2.5}$ using physical and statistical models. The standard geophysical approach has been to use a CTM, such as GEOS-Chem, to compute the $PM_{2.5}$/AOD ratio (Liu et al., 2004; van Donkelaar et al., 2006; van Donkelaar et al., 2015; Xu et al., 2015; Geng et al., 2017), with recent applications correcting for CTM biases using available $PM_{2.5}$ surface network data (Brauer et al., 2016; Van Donkelaar et al., 2016; van Donkelaar et al., 2019; Hammer et al., 2020). An alternative approach is to use machine-learning algorithms to relate satellite AOD to $PM_{2.5}$ by training on the surface network data (Hu et al., 2017; Chen et al., 2018; Xiao et al., 2018; Wei et al., 2021; Pendergrass et al., 2021), and sometimes including CTM values as predictors (Di et al., 2019; Xue et al., 2019). Yet another approach is to assimilate the satellite-measured AODs in a CTM and correct in this manner the $PM_{2.5}$ simulation, although this requires attribution of model AOD errors to specific model parameters (Kumar et al., 2019; Saide et al., 2014; Sekiyama et al., 2010; Cheng et al., 2019). In all of these approaches, a better physical understanding of the AOD-$PM_{2.5}$ relationship as simulated by CTMs can greatly enhance the capability to infer PM¬2.5 from AOD data.

AOD measures aerosol extinction (scattering and absorption) integrated over the atmospheric column, so that its relationship to 24-hr average surface $PM_{2.5}$ (the standard air quality metric) depends on the aerosol vertical distribution and optical properties, ambient relative humidity (RH), diurnal variation of $PM_{2.5}$, and contribution from coarse particulate matter to AOD. Little study of these factors has been conducted for East Asia. Airborne measurements of aerosol vertical profiles in East Asia are very limited (Liu et al., 2009; Sun et al., 2013). AOD is highly sensitive to RH (Brock et al., 2016; Latimer and Martin et al., 2019; Saide et al., 2020), but the impact from RH uncertainty on AOD simulation lacks evaluation. In addition, because the AOD is a daytime measurement that needs to be related to 24-h average $PM_{2.5}$, the diurnal variation of $PM_{2.5}$ needs to be understood (Guo et al., 2017; Lennartson et al., 2018). Finally, there has been to our knowledge no study of how coarse anthropogenic PM may contribute to the AOD measurements. Coarse anthropogenic PM (distinct from desert dust) is known to be high over East Asia (Chen et al., 2015; Dai et al., 2018)."

We revised lines 191-195: "Therefore, we re-computed the diagnostic AOD using updated log-normal size distributions for SNA and organic aerosol with number median radius $R_{N,med}$ = 0.11 μm and geometric standard deviation σ = 1.4 based on KORUS-AQ observations, instead as compared tof $R_{N,med}$ = 0.058 μm and σ = 1.6 in the standard model version 12.7.1, which is derived from IMPROVE network measurements of aerosol mass scattering efficiency over North America (Latimer and Martin, 2019)."

Lines 85-86: Ångström Exponents at 500 nm? AE refers to a wavelength range. Reference is needed here.

We detailed in lines 98-101: "We use total and fine-mode AODs at 500 nm wavelength from the AERONET Version 3; Spectral Deconvolution Algorithm (SDA) Version 4.1 Retrieval Level 2.0 database (Giles et al., 2019; O'Neill et al., 2003). The AERONET AODs at 500 nm are converted to 550 nm using total and fine mode Ångström Exponents at 500 nm for consistency with the satellite AOD data."

Lines 107-110: Himawari-8/AHI: Which version do you use? Reference is needed. Again, Himawari-8/AHI provides AOD products at 500 nm. I am not sure about GOCI (should be 550 nm). Are they the same? If not, does the wavelength difference be taken into account in the data fusion?

We detailed in lines 122-126: "Geostationary satellite AOD at 550 nm are retrieved by the Yonsei Aerosol Retrieval (YAER) algorithm for the GOCI (Choi et al., 2016, 2018) and AHI (Lim et al. 2018) instruments, with GOCI covering East China and South Korea and AHI covering the broad East Asia region. AOD from GOCI and AHI have a 6 km × 6 km spatial resolution and 1-hour (GOCI) to 2.5-minute (AHI) temporal resolution for 8 hours per day (09:30 to 16:30 local time)."

Lines 288-290: What are the potential reasons? Is it the aerosol algorithm or the difference caused by sample matching at different wavelengths?

We explained in lines 308-312: "The low biases in the SMA could be due to high-concentration aerosol pixels mis-identified as clouds and/or possible issues with the aerosol type assumption in the aerosol retrieval, while the high biases on the Yellow Sea islands could result from uncertainties in the assumption of ocean surface reflectance, as has been discussed by Choi et al. (2016, 2018) and Lim et al. (2018, 2021)."

Lines 295-296: What are the potential reasons?

We rephrased lines 317-319: “GEOS-Chem reproduces the satellite AOD enhancements along the west coast of South Korea but the values are lower than observed, which we attribute to unaccounted coarse PM and negative RH bias as discussed below.”

We quantified in lines 323-325: “Therefore, about half of the GEOS-Chem underestimate of total AOD can be attributed to missing coarse PM, with the other half comes from negative RH bias.”

Lines 297-312: Is there any relevant published literature to support the author's explanations of reasons for these differences between GEOS-Chem and satellites observations?

We added in lines 85-91 in the introduction: “AOD is highly sensitive to RH (Brock et al., 2016; Latimer and Martin et al., 2019; Saide et al., 2020), but the impact from RH uncertainty on AOD simulation lacks evaluation. In addition, because the AOD is a daytime measurement that needs to be related to 24-h average $PM_{2.5}$, the diurnal variation of $PM_{2.5}$ needs to be understood (Guo et al., 2017; Lennartson et al., 2018). Finally, there has been to our knowledge no study of how coarse anthropogenic PM may contribute to the AOD measurements. Coarse anthropogenic PM (distinct from desert dust) is known to be high over East Asia (Chen et al., 2015; Dai et al., 2018).”

I also suggest adding some scatter plots to validate and compare the satellite-based and modeled AODs, $PM_{2.5}$, and other parameters if possible, so that readers can see their differences more clearly.

Lines 396-402: “Figure 7 shows daily correlations of the regional average series between AERONET total AOD and GEO satellite AOD, between AERONET fine mode AOD and GEOS-Chem AOD, as well as between measured $PM_{2.5}$ and GEOS-Chem $PM_{2.5}$. Correlations in Figure 7 are all statistically significant with correlation coefficients (R) ranging from around 0.7 to more than 0.9 and normalized mean biases (NMB) within ± 30%. The correlations of these three pairs are similar over South Korea and North China, except that GEOS-Chem overestimates springtime $PM_{2.5}$ in South Korea but not over North China, possibly due to a model overestimate of the long-range transport of $PM_{2.5}$ from China to South Korea in spring.”

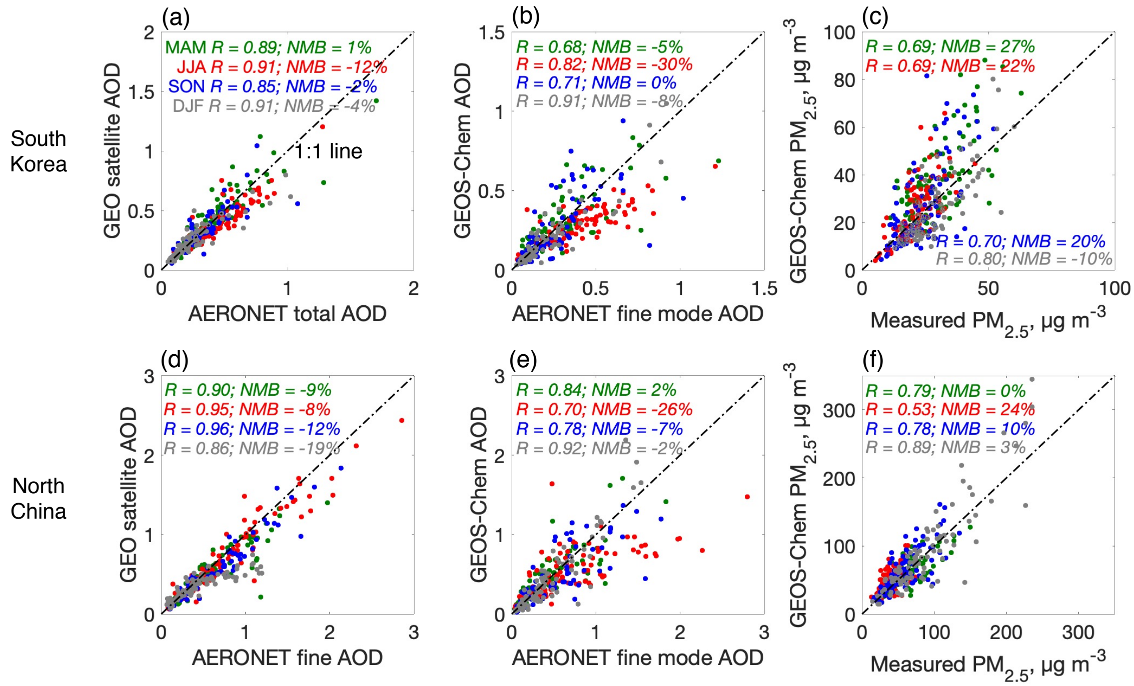

Figure 7. Scatter plots of regional mean daily (a and d) GEO satellite AOD vs. AERONET total AOD, (b and e) GEOS-Chem AOD vs. AERONET fine-model AOD, and (c and f) GEOS-Chem $PM_{2.5}$ vs. measured $PM_{2.5}$ over South Korea (a-c) and North China (d-f). Different colors represent different seasons. Values inset are correlation coefficients (*R*) and normalized mean biases (*NMB*) between surface measurements and GEO satellite or GEOS-Chem values.

Figure 6: I suggest adding some satellite $PM_{2.5}$ estimated results to see the difference with model simulations since there are many available PM products, especially in China.

We added lines 470-472 in the conclusion section: "We have used results from our study in a recent machine-learning reconstruction of daily 2011-present $PM_{2.5}$ over East Asia from GOCI AOD data by identifying critical variables for the machine-learning algorithm and providing blended gap-filling data for cloudy scenes (Pendergrass et al., 2021)."

Last, the authors should consider the impact of other factors, especially BLH, meteorological conditions, and topography, on surface PM, and to see how much impact can they have on the differences between satellite and model results.

We added in the abstract (lines 43-47): "We updated SNA and organic aerosol size distributions in GEOS-Chem to represent aerosol optical properties over East Asia by using in-situ measurements of particle size distributions from KORUS-AQ. We find that SNA and organic aerosols over East Asia have larger size (number median radius of 0.11 µm with geometric standard deviation of 1.4) and

20% larger mass extinction efficiency as compared to aerosols over North America (default setting in GEOS-Chem)."

We quantified in lines 286-287: "The model underestimates extinction coefficients by 20% below 1 km altitude, leading to a 10% underestimate of aircraft inferred AOD, although there is no such underestimate in aerosol mass."

We quantified in lines 323-324 to: "Therefore, about half of the GEOS-Chem underestimate of total AOD can be attributed to missing coarse PM, with the other half comes from negative RH bias."

We quantified in lines 412-414: "The GEOS-Chem AOD is ~ 20% biased low in summer and this is largely due to a low RH bias (Figure S8), as seen previously in the KORUS-AQ comparisons but amplified by the high RH in summer that drives hygroscopic growth (Latimer and Martin, 2019)."

We added analysis in lines 420-423: "GEOS-FP daytime PBL height also shows a stronger seasonality over North China than over South Korea (Figure S8), generally consistent with the CALIPSO daytime PBL height (Su et al., 2018). Previous studies have shown opposite seasonality between MODIS AOD and surface $PM_{2.5}$ over North China and attributed this to the seasonality in PBL height and RH (Qu et al., 2016; Xu et al., 2019)."

We quantified in lines 443-446 in the conclusion section: "We find that GEOS-Chem aerosol optical properties based on measurements over the North America (default model setting) underestimate KORUS-AQ aerosol mass extinction efficiency by around 20%. In addition, a low bias in GEOS-FP RH below 1 km leads to a 10% underestimate of AOD inferred from the aircraft profile."

We added lines 451-452 in the conclusions section: "The remaining 15% underestimate of AERONET fine-mode AOD by GEOS-Chem can be attributed to the RH low bias."

We quantified in lines 460-462 in the conclusions section: "GEOS-Chem simulates successfully the seasonality of measured $PM_{2.5}$ but is ~ 20% biased low in summer for AOD, due again to RH low bias like that during KORUS-AQ, amplified by the high RH in summer that drives hygroscopic growth (Latimer and Martin, 2019)."

We added lines 472-474 in the conclusion section: "Besides the factors discussed in this study, topography might be another important factor influencing surface $PM_{2.5}$ and its vertical mixing (Su et al., 2018), and this also requires future investigation."

References (newly added):

Brauer, M., Freedman, G., Frostad, J., van Donkelaar, A., Martin, R. V., Dentener, F., Dingenen, R. v., Estep, K., Amini, H., Apte, J. S., Balakrishnan, K., Barregard, L., Broday, D., Feigin, V., Ghosh, S., Hopke, P. K., Knibbs, L. D., Kokubo, Y., Liu, Y., Ma, S., Morawska, L., Sangrador, J. L. T., Shaddick, G., Anderson, H. R., Vos, T., Forouzanfar, M. H., Burnett, R. T., and Cohen, A.: Ambient Air Pollution Exposure Estimation for the Global Burden of Disease 2013, Environ. Sci. Technol., 50, 79-88, 10.1021/acs.est.5b03709, 2016.

Brock, C. A., Wagner, N. L., Anderson, B. E., Beyersdorf, A., Campuzano-Jost, P., Day, D. A., Diskin, G. S., Gordon, T. D., Jimenez, J. L., Lack, D. A., Liao, J., Markovic, M. Z., Middlebrook, A. M., Perring, A. E., Richardson, M. S., Schwarz, J. P., Welti, A., Ziemba, L. D., and Murphy, D. M.: Aerosol optical properties in the southeastern United States in summer – Part 2: Sensitivity of aerosol optical depth to relative humidity and aerosol parameters, Atmos. Chem. Phys., 16, 5009-5019, 10.5194/acp-16-5009-2016, 2016.

Chen, G., Li, S., Knibbs, L. D., Hamm, N. A. S., Cao, W., Li, T., Guo, J., Ren, H., Abramson, M. J., and Guo, Y.: A machine learning method to estimate $PM_{2.5}$ concentrations across China with remote sensing, meteorological and land use information, Sci. Total Environ., 636, 52-60, https://doi.org/10.1016/j.scitotenv.2018.04.251, 2018.

Chen, W., Tang, H., and Zhao, H.: Diurnal, weekly and monthly spatial variations of air pollutants and air quality of Beijing, Atmos. Environ., 119, 21-34, https://doi.org/10.1016/j.atmosenv.2015.08.040, 2015.

Choi, M., Kim, J., Lee, J., Kim, M., Park, Y. J., Holben, B., Eck, T. F., Li, Z., and Song, C. H.: GOCI Yonsei aerosol retrieval version 2 products: an improved algorithm and error analysis with uncertainty estimation from 5-year validation over East Asia, Atmos. Meas. Tech., 11, 385-408, 10.5194/amt-11-385-2018, 2018.

Guo, J., Xia, F., Zhang, Y., Liu, H., Li, J., Lou, M., He, J., Yan, Y., Wang, F., Min, M., and Zhai, P.: Impact of diurnal variability and meteorological factors on the $PM_{2.5}$ - AOD relationship: Implications for $PM_{2.5}$ remote sensing, Environ. Pollut., 221, 94-104, https://doi.org/10.1016/j.envpol.2016.11.043, 2017.

Hu, X., Belle, J. H., Meng, X., Wildani, A., Waller, L. A., Strickland, M. J., and Liu, Y.: Estimating $PM_{2.5}$ Concentrations in the Conterminous United States Using the Random Forest Approach, Environ. Sci. Technol., 51, 6936-6944, 10.1021/acs.est.7b01210, 2017.

Lennartson, E. M., Wang, J., Gu, J., Castro Garcia, L., Ge, C., Gao, M., Choi, M., Saide, P. E., Carmichael, G. R., Kim, J., and Janz, S. J.: Diurnal variation of aerosol optical depth and $PM_{2.5}$ in South Korea: a synthesis from AERONET, satellite (GOCI), KORUS-AQ observation, and the WRF-Chem model, Atmos. Chem. Phys., 18, 15125-15144, 10.5194/acp-18-15125-2018, 2018.

Lim, H., Go, S., Kim, J., Choi, M., Lee, S., Song, C. K., and Kasai, Y.: Integration of GOCI and AHI Yonsei aerosol optical depth products during the 2016 KORUS-AQ and 2018 EMeRGe campaigns, Atmos. Meas. Tech., 14, 4575-4592, 10.5194/amt-14-4575-2021, 2021.

Liu, P., Zhao, C., Liu, P., Deng, Z., Huang, M., Ma, X., and Tie, X.: Aircraft study of aerosol vertical distributions over Beijing and their optical properties, Tellus B Chem. Phys. Meteorol., 61, 756-767, 10.1111/j.1600-0889.2009.00440.x, 2009.

O'Neill, N. T., Eck, T. F., Smirnov, A., Holben, B. N., and Thulasiraman, S.: Spectral discrimination of coarse and fine mode optical depth, J. Geophys. Res. Atmos., 108, https://doi.org/10.1029/2002JD002975, 2003.

Pendergrass, D. C., Jacob, D. J., Zhai, S., Kim, J., Koo, J. H., Lee, S., Bae, M., Kim, S.: Continuous mapping of fine particulate matter ($PM_{2.5}$) air quality in East Asia at daily 6x6 km^2 resolution by application of a random forest algorithm to 2011-2019 GOCI geostationary satellite data, submitted, 2021.

Qu, W., Wang, J., Zhang, X., Sheng, L., and Wang, W.: Opposite seasonality of the aerosol optical depth and the surface particulate matter concentration over the north China Plain, Atmos. Environ., 127, 90-99, https://doi.org/10.1016/j.atmosenv.2015.11.061, 2016.

Su, T., Li, Z., and Kahn, R.: Relationships between the planetary boundary layer height and surface pollutants derived from lidar observations over China: regional pattern and influencing factors, Atmos. Chem. Phys., 18, 15921-15935, 10.5194/acp-18-15921-2018, 2018.

Sun, X., Yin, Y., Sun, Y., Sun, Y., Liu, W., and Han, Y.: Seasonal and vertical variations in aerosol distribution over Shijiazhuang, China, Atmos. Environ., 81, 245-252, https://doi.org/10.1016/j.atmosenv.2013.08.009, 2013.

van Donkelaar, A., Martin, R. V., Li, C., and Burnett, R. T.: Regional Estimates of Chemical Composition of Fine Particulate Matter Using a Combined Geoscience-Statistical Method with Information from Satellites, Models, and Monitors, Environ. Sci. Technol., 53, 2595-2611, 10.1021/acs.est.8b06392, 2019.

van Donkelaar, A., Martin, R. V., and Park, R. J.: Estimating ground-level $PM_{2.5}$ using aerosol optical depth determined from satellite remote sensing, J. Geophys. Res. Atmos., 111, 10.1029/2005JD006996, 2006.

Xiao, Q., Chang, H. H., Geng, G., and Liu, Y.: An Ensemble Machine-Learning Model To Predict Historical $PM_{2.5}$ Concentrations in China from Satellite Data, Environ. Sci. Technol., 52, 13260-13269, 10.1021/acs.est.8b02917, 2018.

Xu, J., Han, F., Li, M., Zhang, Z., Xiaohui, D., and Wei, P.: On the opposite seasonality of MODIS AOD and surface $PM_{2.5}$ over the Northern China plain, Atmos. Environ., 215, 116909, https://doi.org/10.1016/j.atmosenv.2019.116909, 2019.

Xu, J. W., Martin, R. V., van Donkelaar, A., Kim, J., Choi, M., Zhang, Q., Geng, G., Liu, Y., Ma, Z., Huang, L., Wang, Y., Chen, H., Che, H., Lin, P., and Lin, N.: Estimating ground-level $PM_{2.5}$ in eastern China using aerosol optical depth determined from the GOCI satellite instrument, Atmos. Chem. Phys., 15, 13133-13144, 10.5194/acp-15-13133-2015, 2015.

Xue, T., Zheng, Y., Tong, D., Zheng, B., Li, X., Zhu, T., and Zhang, Q.: Spatiotemporal continuous estimates of $PM_{2.5}$ concentrations in China, 2000–2016: A machine learning method with inputs from satellites, chemical transport model, and ground observations, Environ. Int., 123, 345-357, https://doi.org/10.1016/j.envint.2018.11.075, 2019.