Anonymous Referee #2

The authors appreciate the reviewers very much for reviewing our manuscript and providing constructive comments. As suggested, we carefully revised the manuscript thoroughly according to the valuable advices, as well as proof-read the manuscript to minimize typographical, grammatical, and bibliographical errors. Our replies to the comments and our actions taken to revise the paper (in blue) are given below (the original comments are copied here). The figures added in the reply is represented by 'Figure', which is distinguished from 'Fig.' in the manuscript.

Major comments:

The main purpose of this paper is to clearly show the ability and/or inability of their model to simulate the observed spatial and temporal variation in the concentration of the chemical species in the atmosphere. Based on those findings the authors and also the readers of the paper can understand what kind applications are suitable for this model and what aspect of the model should be further improved in order to apply it to a particular issue. From this point of view, the self-evaluation about the ability of the model by authors were often insufficient and unclear. The good points and also the shortcomings of the model should be described more specifically in the text. I pointed out some of those points as specific comments in the following, but I strongly recommend the authors to reexamine the descriptions particularly in the model evaluation parts.

Reply: We greatly appreciate the reviewer for insight comments on the manuscript. To respond to the reviewer's major concerns, we made thoroughly revisions and corrections according to all the insight comments of the reviewers. Besides, more crucial information and analysis will be added in the revised manuscript, as follows:

(1) A new evaluation with the WDCGG datasets only for 2014 is updated, and the evaluation is more quantitative.

(2) We provide more information to discuss the model's performance on the underestimation of CO over ocean, including a comparison of the profile concentration of OH with other models.

(3) More studies and descriptions focused on the simulation of ozone. The bias of inter-models and model-observation are discussed. In particular, the poor performance on the seasonal cycle in NH land is interpreted. We further showed the seasonal cycle of

ozone compared against sites separated by the terrain.

(4) The model's simulating ability on aerosol formation is discussed in detail. Especially, the SOA formation mechanism and the multiphase processes in the model are described.

(5) Overall, more analysis on the model's performance are shown. The discussion includes the uncertainty of emissions, the limitation of global model resolution, the superiority of the model, the shortcomings of chemical scheme, the impact of meteorology and deposition. On the basis of these comparison and analysis, some suggests are put forward in the model's further improvement.

Specific comments:

- L24: What are the aerosol effects here?

Reply: The aerosol effects refer to climate effect (direct, semi-direct and indirect effect) and health effect (mainly of respiratory diseases, cardiovascular risk and lung cancer).

- L38: Only R-value can not ensure the accuracy of the simulation. How about MB or NMB?

Reply: The NMB for the nested simulation are within ±0.5. It is supplemented in the abstract.

-L58: Why didn't you cite the latest AR5 report here?

Reply: Good suggestion, we have updated the citation to the latest report.

- L79: Typo? e.g.

Reply: 'For example' is used to introduce the work by Badia et al. (2017), Mann et al. (2010) and Tsigaridis et al. (2014) instead of 'e.g.' in the manuscript.

- L86: EA, this should be defined at its first appearance in the text.

Reply: Thanks, the definition has been added at its first appearance in the revised manuscript.

- L108-109: What do you mean here? Could you use more words to explain "localization of the process parameterization"?

Reply: In the dust module, the deflation mechanism and dust loading parameterization are based on a detailed analysis of the meteorological conditions, landform, and climatology from daily weather records at about 300 local stations in north China. For the heterogeneous chemistry scheme, the parameterization of uptake coefficients considered the meteorological condition of relative humidity in China. It has been added in the manuscript.

- L135-139: Are there citable references for CoLM, , and IAP-OBGCM?

Reply: the references for CoLM (Dai et al., 2003), and IAP-OBGCM (Li et al., 2012) has been supplemented in the revised manuscript.

-L158: What is the main difference between these two models (GNAQPMS and IAP-AACM)?

Reply: Generally, IAP-AACM is similar to GNAQPMS. IAP-AACM has the same model framework with GNAQPMS but has some improvements. The model was renamed when it joined the CAS-ESM.

- L204: What does "synchronous time step" mean ?

Reply: It is the time step of model's integration calculation. In order to keep the stability of calculation in the model, the integration time step will be cut into shorter sub-integration time step in different modules (e.g., advection and gas chemistry processes). So the synchronous time step means the model's integration calculation.

- L206: What is the reason for choosing the year 2014 as the focal year?

Reply: Choosing the year 2014 is a compromise between little change in emissions and more observation data published in China. The year 2014 is near now that there are more comprehensive observation data of both trace gases and particles to obtain over China. Chinese national environmental monitoring network (CNEMC) started to publish data since 2013. The site records of aerosols for 2014 are also available in China. Besides, China has implemented strict air pollution control measures since 2013. Emission sources haven't change much from 2010 to 2014, except SO_2. Zheng et al. (2018).

- L224-225: Emission data used in the study are not up-to-date, the base year of each database is a bit old. Therefore, adjusting the emission data to input them to the model is suitable for the purpose of this study. However, you only mentioned about the adjustment of SO2 emission in China in the text. Did you adjust other species emission?

Reply: No, we only adjust the emission of SO_2 in China for its dramatic variation in the past years. During 2010~2014, the fluctuation of emissions are not severe globally except China, due to a strict controlling policy known to all. The study by Zheng et al. (2018) shows that relative change rates of China's anthropogenic emissions during 2010–2017 are estimated as follows: -62% for SO2, -17% for NOx, -27% for CO, -27% for BC and -35% for OC. But the emissions decreased by 59% for SO2, 21% for NOx, 23% for CO, 28% for BC and 32% for OC during 2013-2017. The dramatic reduction of emissions is mostly happened after 2013

(shown in Figure 1) for China's Clean Air Action implemented during 2013-2017. Compared to 2010, emissions of trace gas in 2014 decreased slightly except SO_2. So we only adjust the emission of SO_2.

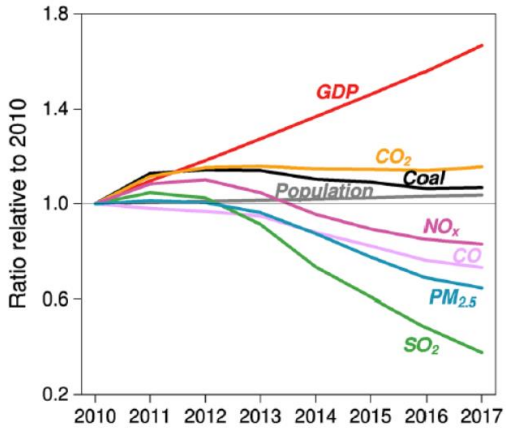

Figure 1 Emission trends and underlying social and economic factors from 2010 to 2017 by Zheng et al. Zheng et al. (2018)

- Figure2: Why did you compare with NCEP R1, not with NCEP-FNL? What is the purpose of it?

Reply: Because the meteorological field of WRF is nudged to FNL datasets. We used NCEP R1 to compare with the simulation considering data independence.

- L244: This statement is not correct. The difference in RH2 between WRF and Reanalysis is much larger in general as shown in Fig2 over land area.

Reply: The statement is correct. The figure is wrong and it has been replaced.

- Table2 and L246-248: If you want to mention only the correlation coefficient of annual mean values, you should remove Table 2. If you want to retain Table2, you should explain the table more precisely here. Table 2 is hard to read and insufficient caption.

Reply: Thank you for your good suggestion. More descriptions (see below) about Table 2 is given in the revised manuscript, and captions and units are added in Table 2.

The simulation of the meteorological factors are close to the site records in different season, with mean bias (MB) of -0.3 ~ 0 ℃, -0.8 ~ -0.5 m/s and -4~ -2.3% for T_2, W_{10} and RH_2 respectively. The model underestimates T_2 in all the seasons. Thus the summer showed the largest negative bias with Root Mean Square Error (RMSE) of 2 ℃, for temperature reaches the peak in summer. As for W_{10}, it's also underestimated the most in summer, with MB of -0.8m/s and RMSE of 1.9 m/s. As for RH2, the underestimation is more obvious in summer

(MB= -3.2%) and autumn (MB= -3.2%), mainly stem from the insufficient precipitation. Overall, the simulation in summer is more underestimated than other seasons. The agreement in T_2 and RH_2 with observations is better than that of W_{10}, with annual correlation coefficients (R) of 0.98, 0.84 and 0.53, respectively. Generally, the meteorology calculated by WRF can rational reproduce the characteristics of observations.

- L270: Why did you take average of 2006-2015 only for WDCGG?

Reply: The WDCGG datasets provides a large number of trace gases observations globally. The datasets can help to evaluate model performance of CO and ozone in different regions, but some sites are without invalid records in 2014. To get more data to evaluate the model over the world, we expanded the time range to ten years (2006-2015) and take the average of 2006-2015 as the statement of the air in the initial evaluation.

We have re-selected the observation data for 2014 to comparison with model results. Overall, the results have not changed much in terms of the evaluation of model's simulation capability. The simulation bias is reduced in some regions while it is increased in some other regions. The simulation of NO_2 performed better with the NMB of Asia and Europe closer to zero. The underestimation of CO in Antarctica disappeared due to the change of the observed value. There are some changes in the trend of the seasonal variation of O_3 in Northern Hemisphere. All the figures (as shown in Figure 2~ Figure 4) and tables related to these changes are updated in the manuscript, and the corresponding analysis is updated in the manuscript, too.

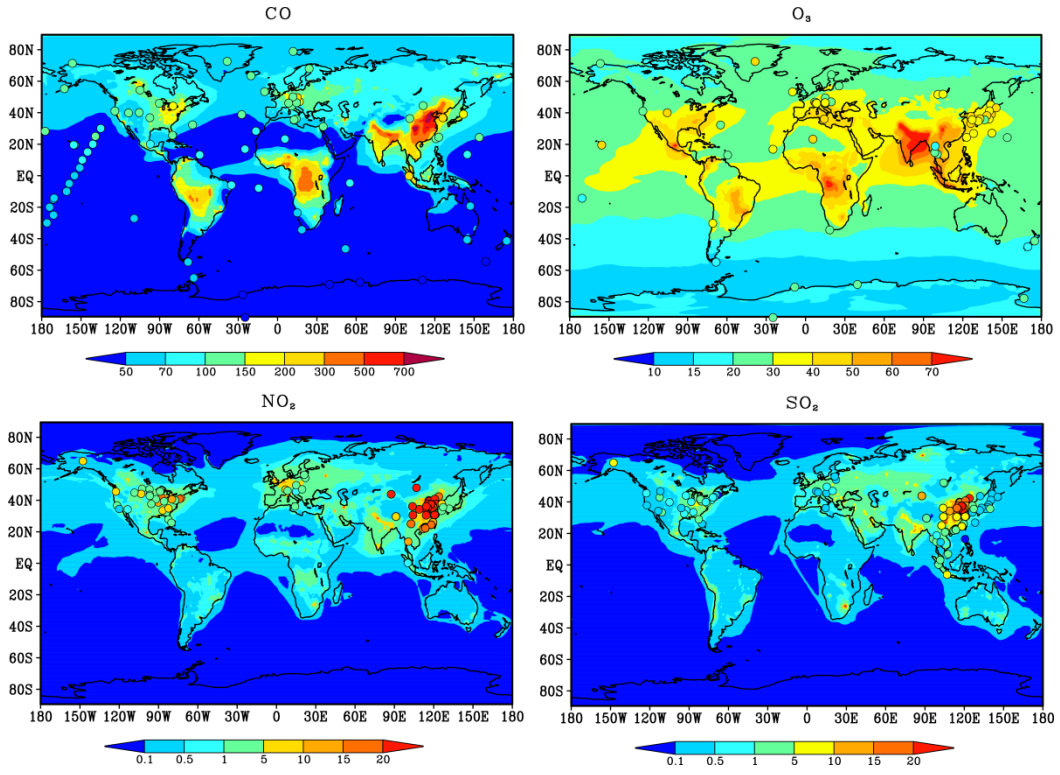

Figure 2 Annual mean concentration (ppb) of the surface layer in IAP-AACM. The circles represent site observations. The first row is CO and O_3, the bottom row is NO_2 and SO_2.

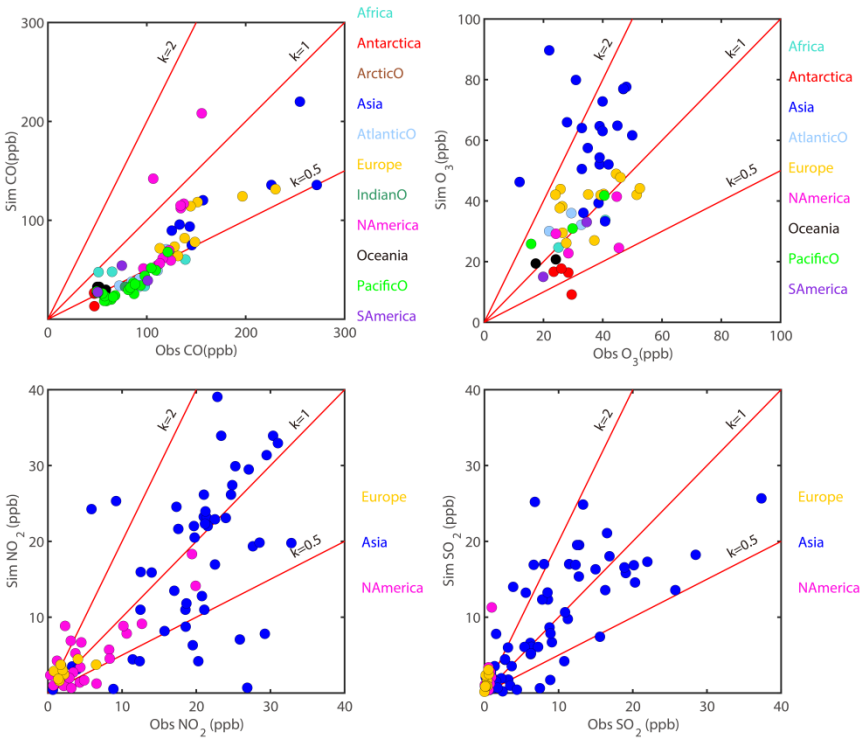

Figure 3 Scatter plots of annual mean concentrations (ppb) in Africa, Antarctica, Arctic Ocean (ArcticO), Asia, Atlantic Ocean (AtlanticO), Europe, Indian Ocean (IndianO), North America (NAmerica), South America (SAmerica), Oceania and Pacific Ocean (PacificO). The abscissa shows the observation and the ordinate shows the simulation. The color of the points represents different regions. (a) ~ (d) show CO, O_3, NO_2 and SO_2 respectively.

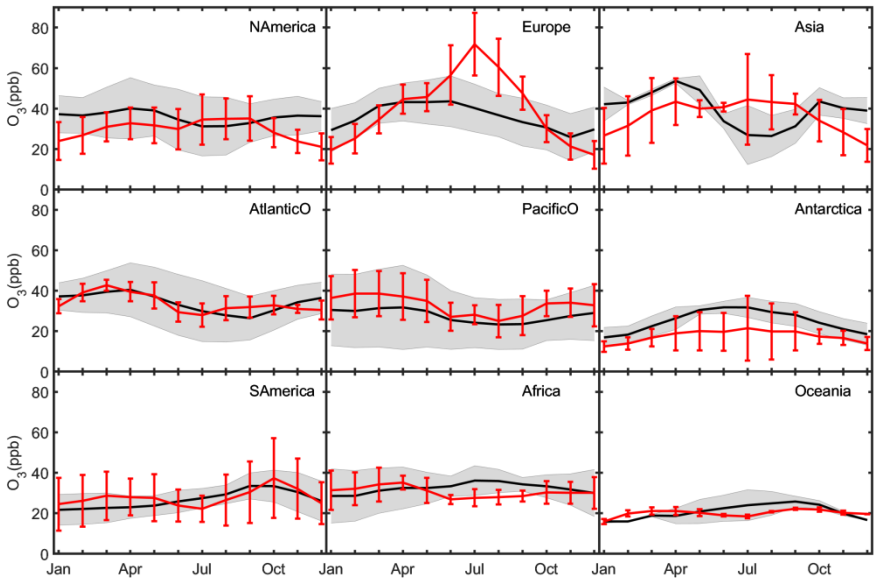

Figure 4 Mean seasonal variation of O_3 (ppb) over NAmerica, Europe, Asia, AtlanticO, PacificO, Antarctica, SAmerica, Africa and Oceania sites. Black lines and red lines represent the average of observations and simulations respectively. Gray shaded areas and red vertical bars show 1 standard deviation over the sites for observations and for model results respectively.

- Table3: It is better to include the information of region of each observation datasets.

Reply: That's a good suggestion and we have added the region of each observation datasets.

- L298: The value (23.3) differs from that in Table4.

Reply: Yes, it's a writing mistake. We have corrected it to 22.8 in the sentence.

- Table5: Table 5 is not completely filled with the necessary information for OM (other sources and total sink are missing).

Reply: Other sources and total sink are added to Table 5.

- Figure4: Fig4c and 4d should be switched to be in accordance with the order of panels in Fig3.

Reply: The subplot in Fig. 4 has been switched to the same order as Fig. 3.

- L363: Typo?: Northern Hemisphere

Reply: Yes, we have revised it.

L368-369: Why did IAP-AACM show the lowest concentration of CO over ocean among the models considered here?

Reply: It potentially reflects a difference in emissions. The natural sources of CO over ocean are included in the HTAP models whereas they are not considered in IAP-AACM. Besides, it may reflect differences in chemical transformation between models. The tropospheric (200hpa to the surface) mean OH concentration of IAP-AACM is 13.0×10^5 molec cm^{-3}. It is a little higher than the mean OH concentration study ($11.1\pm1.6\times10^5$ molec cm^{-3}) from 16 ACCMIP models for 2000 by Naik et al. (2013). It potentially leads to strong atmospheric oxidation. As shown in Figure 5, there is a slightly higher peak concentration of 30-35 molec cm^{-3} in IAP-AACM, compared with the other models (under 30 molec cm^{-3}) (Huijnen et al., 2010; Badia et al., 2017). Due to the sink reaction of CO ($CO + OH \rightarrow CO_2 + H$), the CO loss will be faster in IAP-AACM.

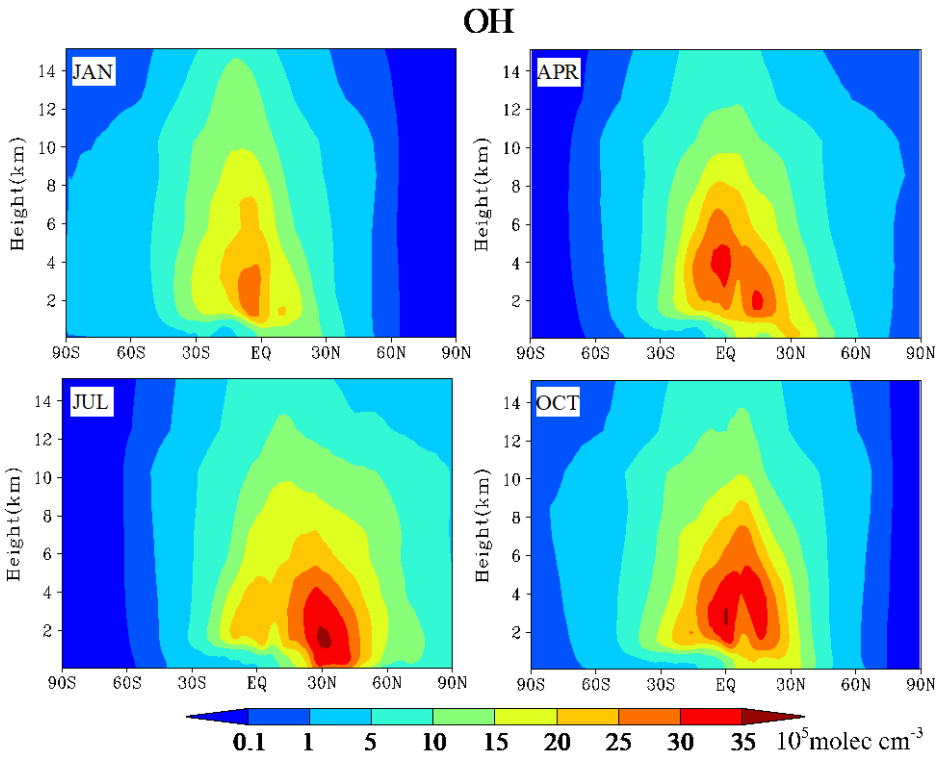

Figure 5 Zonal monthly mean concentration of OH for January, April, July and October by the IAP-AACM. The unit is 10^5 molecule cm^{-3}.

- L378-380: The seasonal variation of surface O3 should be different in different environment even in the same region. So, I recommend the authors to compare separately for different environment (e.g. maritime area vs mountainous area). Otherwise, I can not regard the Fig6 as an evidence that the model can well simulate the seasonal variation of surface O3.

Reply: That's a helpful suggestion. The seasonal cycle of ozone shows different characteristics in different topographic conditions due to different control factors. We separate the observational sites as maritime area, inland and mountain due to the altitudes (shown in Figure 6). At the mountain sites the model tends to underestimate O_3 concentrations more. As high-altitude sites more frequently sample free tropospheric air (Fiore et al., 2009), it is more likely to be influenced by foreign emissions. The weaker intercontinental pollutant transport within troposphere in the model may be responsible for the underestimation. Moreover, steep topographic gradients at local may not be represented using coarse resolution models. Steep topographic gradients are averaged out within one model grid cell. It is difficult to capture the spatio-temporal variation of ozone caused by topographically driven flows or sharp gradients in mixing depths. For inland, the model tends to overestimate O_3 concentrations in summer time. Uncertainties in volatile organic compounds (VOCs)-NOx-O_3 chemistry may contribute. The natural source of isoprene from

vegetation is important in the O_3 formation due to its high proportion of VOCs emission in summer (as estimated to be 40.9 Tg/yr in China by Fu et al., 2012)

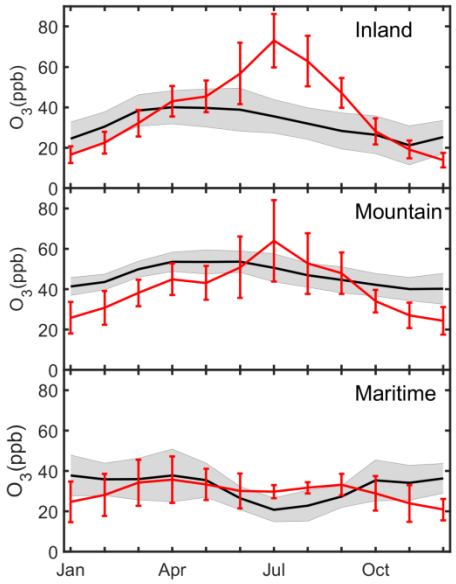

Figure 6 Mean seasonal variation of O_3 (ppb) over inland, mountain and maritime area in Northern Hemisphere compared with site records. Black lines and red lines represent the average of observations and simulations respectively. Gray shaded areas and red vertical bars show 1 standard deviation over the sites for observations and for model results respectively.

- L385: Underestimation in Antarctica is not small. Such an underestimation could be seen in the other CTMs. Can you use more words about this issue here?

Reply: In IAP-AACM, ozone concentration is about 10~15 ppb lower than site observations in Antarctica. An Evaluation of ACCMIP ozone simulations shows most models underestimate ozone concentration at high latitudes in the southern hemisphere at 800 hPa (Miyazaki & Bowman, 2017). As displayed in their seasonal comparison results, the underestimation of ozone is more significant (10~20 ppb) during April-August, which is in accordance with our study. It may be caused by a lack of halogen chemistry in our model. Remarkable ozone depletion events which is driven by halogen chemistry (mostly notably as bromine) is observed in the polar boundary layer (Simpson et al., 2007). Furthermore, Falk & Sinnhuber (2018) used EMAC v2.52 to interpret the significant underestimation of situ observed ozone in Antarctica, indicating that there are missing sources of bromine release from ice and snow in the model. The over prediction of dry deposition velocity to sea ice also plays a role in the underestimation of ozone. The dry deposition velocity to ice is under 0.02 cm s^{-1} across 15 HTAP models (Ganzeveld et al., 2009). In IAP-AACM, it's obviously higher (0.035~0.048 cm s^{-1}) than those models, as shown in Figure 7.

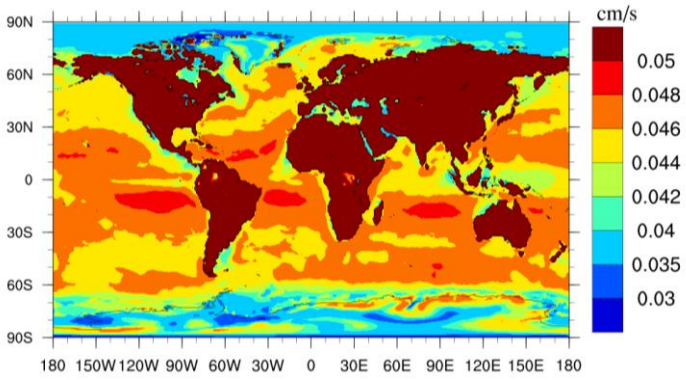

Figure 7 Annual mean dry deposition velocity of ozone in IAP-AACM. The unit is cm s^{-1}.

- L385-387: In the NH land area, it seems that the model completely failed to represent the seasonal cycle of surface O3, but the author regarded it as just a positive bias during July-September. More words for this issue are necessary. For example, what do you think the apparent underestimation in cold season in NAmerica and Europe?

Reply: Yes, the model showed poor performance on the seasonal cycle of surface ozone in the NH land area, with overestimation in Europe and EA in summer while underestimation in winter in NH land, as shown in Figure 4 (the comparison drawn with WDCGG observation only for 2014).

According to Figure 6, the underestimation of ozone in cold seasons mainly occurs at sites of mountain and marine area, where is relatively clear and tend to be impacted by foreign pollutions. The underestimation in winter may relate to the weaker intercontinental pollutant transport within troposphere in the model. Winds are generally stronger in winter than in summer, causing intercontinental transport to be more rapid during winter months (HTAP, 2010). The observation at those sites will be influenced by source emissions more frequently.

The surface O_3 are also underestimated in spring over NH land. In IAP-AACM, the stratospheric-tropospheric exchange is not considered. It will lead to a large negative bias in simulation. To date it has become apparent that the measured annual cycle of ozone shows a distinct maximum during spring. The stratosphere-to-troposphere ozone transport event occurs widely across mid-latitudes in the NH (Monks et al., 2000; Akritidis et al., 2018). Since the magnitude and frequency of the transport through tropopause is still not clear. There are large uncertainties in simulating the flux. Some researches (Munzert et al, 1985; Austin and Follows, 1991) showed that the maximum in the stratosphere to troposphere flux occurs in late winter/spring. It may partly be responsible for the underestimation of O_3 in winter, too.

The surface O_3 concentrations over East Asia (sites mainly located in Japan) are overestimated in summer and early autumn. The same pattern is also found in the multi-model

inter-comparison of 21 HTAP models (Fiore et al., 2009). The simulations in island countries of EA are sensitive to the timing and extent of the Asian summer monsoon (Han et al., 2008). The positive model bias in this season may stem from inadequate representation of southwesterly inflow of clean marine air.

- L388-390: In Badia et al (2017), they suspected the excessive emission height of NOx which will cause low NOx at surface and consequently might lead to weak NO titration. Do the same things happen in your model?

Reply: Yes, there is the same situation in our model. In Badia et al. (2017), all the land-based anthropogenic emissions are emitted in the first 500m of the model. In IAP-AACM, the energy emissions and industry emissions are emitted in the first five layers considering the stack height, which the top injection height is over 500m.

- L390-391: The AACM apparently showed larger concentration of surface O3 in the tropical regions (central Africa, South America, and Southeast Asia) than the other models. However, the concentration of O3 precursor species (CO and NOx) in these regions are not so different among the models. Can you give discussion about the issue here?

Reply: Yes, the concentrations of CO and NOx in the tropical regions are not so different among the models. There are several uncertainties in the model performance. The same module schemes applied in different models may display different result (Tsigaridis et al. 2014; Hardecre et al., 2015). Furthermore, the meteorological conditions also play a important role in the simulation. The chemical reactions and dynamical processes (transportation and diffusion) of the matters are sensitive to meteorological field (e.g., wind, precipitation, temperature). In addition, the biomass burning emissions used in IAP-AACM is different from the other models. For multi-model activities of HTAP, groups use GFED3 data as the biomass burning emissions (Galmarini et al., 2017). In IAP-AACM, we use GFED4. A comparison of different versions of GFED emissions (Werf et al., 2017) shows the impact of a minor reduction in burned area and decreasing fuel consumption.

- L391-394: These two sentences are not consistent to each other. In general, the region of high O3 concentration can be different in different season. If you look at the "annual mean" concentration, the highest O3 usually occur in the source region in summer, but that in the downwind region in winter. However, if you see the different index such as MD8H O3, you can see completely different seasonal cycle. I strongly recommend the author to carefully revise these sentences.

Reply: We totally agree with your comment. For ozone, the seasonal cycle of high value area

between annual mean value and MD8H O_3 always show different patterns. As for the surface distribution of annual mean concentration shown in Fig. 5, the four models shows the common features of the the NOx titration effect that high concentrations mainly occur downwind of highly polluted areas. But compared with the other models, IAP-AACM exhibits a higher concentration in the source regions and a lower concentration downwind. As the NO_2 emission is emitted at a higher altitude in the model, it is transported to a further distance over the ocean. Thus the concentration of O_3 is lower in the source area (e.g., East China) due to a weaker NOx titration effect, and it's higher in the downwind area. We have revised those sentences.

- L402-403: An overall evaluation of O3 dry deposition in global CTMs can be seen in Hardecre et al. (2015). I recommend to check it out. Hardacre et al. (2015) An evaluation of ozone dry deposition in global scale chemistry climate model, Atmos. Chem. Phys., 15, 6419–6436, doi:10.5194/acp-15-6419-2015.

Reply: Thanks for your suggestion. According to Hardecre et al. (2015), the dry deposition velocity to sea varies little (around 0.05 cm s^{-1}) in different CTMs models using the deposition scheme by Wesely (1989). Besides, the study of Ganzeveld et al. (2009) shows that surface ozone differed by up to 60% if the deposition velocity of ozone varies from 0.01 to 0.05 cm s^{-1}. In IAP-AACM, the deposition velocity over the oceans varies from 0.042 to 0.05 cm s^{-1}, as shown in Figure 7. The variation in absolute terms between IAP-AACM and the other models is smaller than 0.008 cm s^{-1}. Hence the difference of surface ozone caused by dry deposition should be less than 12%. We revised sentences of line 402-403 in the manuscript.

- L416-418: The concentration of NOx over oceanic areas are larger in AACM than in other models, which might stem from larger emission or longer life time of NOx in AACM than the other models. I recommend to discuss this issue further here.

Reply: We totally agree with reviewer's suggestion. Compared with the other models shown in Fig. 5, the surface NO_2 over ocean is larger in IAP-AACM. This may reflect larger emission or less sinks of NO_2 in IAP-AACM. From our research, the most likely cause is the discrepancy in chemical conversion. As displayed in Fig. 7, nitrate is underestimated on the sea of East Asia, which implicates the oxidation of NOx to nitrate is insufficient in the model. Consequently, the higher NO_2 over ocean also leads to higher concentration of surface ozone over equatorial oceans, too. In addition, the higher injection height of emission sources leads to farther transportation distance. This is added in the revised manuscript.

- L436-438: This is misleading statement. The model results are not generally with in a factor of two, but they apparently tend to overestimate the observation in all the three regions. The NMB value for sulfate in Europe, 0.11, is incorrect which is 1.1 in Table 6.

Reply: The '0.11' is a typo, we have corrected it and renewed a new description about the simulation of sulfate as follows. In general, the surface distribution of SNA in IAP-AACM is close to the site records as shown in Fig. 7. When it comes to the site bias, Sulfate is overestimated more or less as shown in Fig. 8. Specifically, in Asia, the simulations at most sites here are within a factor of two of observations, with NMB of 0.36. However, In NAmerica and Europe, it's significantly overestimated, with NMB of 1.94 and 1.1 respectively.

- L438-439: How can you conclude like this (2ugm-3 higher)? What is the ground of this statement?

Reply: We calculated the sites average value in the model and compared it against the observation but we didn't mention it, now we have added this description in the revised paper.

- L442-443: What aspect of the observation do you think your model can reproduce? You should be more specific.

Reply: As shown in Fig. 7, IAP-AACM reproduces the nitrate distribution in western America well but overrates it in eastern America. Hence the west-east gradient over America is overestimated. The model doesn't fully capture the north-south gradient over Europe due to an overestimation at most of the sites. As for Asia, there is an underestimation in Southeast Asia and Japan. We have provided a detailed description of the distribution in the revised manuscript.

- L446-449: About the simulation of ammonium, I can see obvious underestimation in NAmerica and overestimation in Asia and Europe.

Reply: The simulation of ammonium is more diverse since there are more uncertainties in the emission of NH3 (precursor of ammonium) from croplands (Xu et al., 2019). There is slight negative bias in America and positive bias in Asia, with NMB less than ±1 (-0.46 and 0.85 respectively). In Europe, there is significant positive bias with NMB of 1.49. This has been added to the revised manuscript.

- L455-457: The concentration of OC were obviously underestimated by the model.

Reply: Yes, the meteorological conditions and emission inventories in the model are inconsistent with the observation year (2006) of carbonaceous in China. This may be partially

responsible for the bias of OC. According to recent study, there is a slightly increasing (less than 0.1Tg) of both BC and OC emissions from 2006 to 2010 in China (Lu et al., 2011; Fu et al., 2012). As the analysis of the CAWNET observation over China (Zhang et al., 2015), there is no significant changes happened in the proportion of chemical component of PM_{10} from 2006 to 2013, which means the source of carbonaceous are also changed slightly. However, as shown in Fig. 8, the simulation of BC at most sites are close to observations while the simulation of OC is significantly underestimated. The study by Fu et al. (2012) showed a significant underestimation of OC emissions over China. Furthermore, Zhao et al. (2016) found that the pathway of intermediate volatile organic compounds (IVOC) to SOA is very important for the formation of SOA. Their experiments in three dimension model with new SOA scheme called 2B-VBS suggest that OA aging and intermediate-volatility emissions increased OA concentrations in Eastern China by 42%. IVOCs constitute over 40% of OA concentrations, and over half of SOA concentrations. Yang et al. (2018) also showed the significant increase of SOA concentration in an observation-based box model which included the IVOCs reactions. The reaction for IVOC is not included in our SOA module. The SOA module in IAP-AACM is Two-Product scheme. The simulations using Two-Product model substantially underestimated SOA and OA compared with observations. According to recent modeling studies with Two-Product scheme in China, it is estimated to an underestimation of OA by 40-78% (Lin et al., 2016; Han et al., 2016). This can explain the closely simulating of BC but greatly underestimating of OC and throw light on an improvement of SOA formation in the model.

- L491-492: The highest AOD in DJF in east China is not clearly seen both in satellite and model AOD.

Reply: Yes, it's an incorrect expression here and we have deleted it. In fact, the highest AOD may not be in DJF, it often appears in MAM. On one hand, China is frequently affected by dust in spring. On the other hand, AOD is an optical characteristic of aerosols for the whole vertical layer. It is not equivalent to the surface aerosol mass concentration.

- Figure10: It's better to show scatter plots too, at least as a supplement figure.

Reply: To be more quantitative, we provided scatter plots of the species in Figure 8. As shown in Figure 8, model results for SO_2, PM_{10} and $PM_{2.5}$ are mostly within the factor of two with NMB within ±0.52, while NO_2 concentration are a bit underestimated (NMB= -0.63).

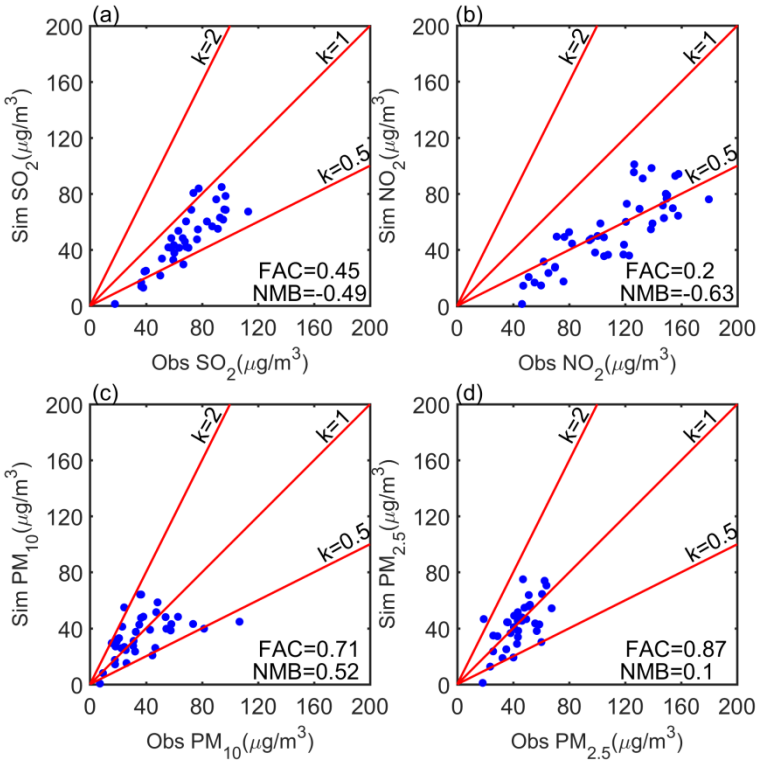

Figure 8 Scatter plots of annual mean concentrations (μg m^{-3}) in nested domain. (a)~(f) is SO_2, NO_2, PM_{10} and $PM_{2.5}$ respectively. The abscissa shows the observation and the ordinate shows the simulation.

- Figure11: The area and the map projection of the figures for all models should be united.

Reply: The model results have been adapted to the same area and projection (shown in Figure 9). Also, a more detailed analysis has been displayed in the manuscript.

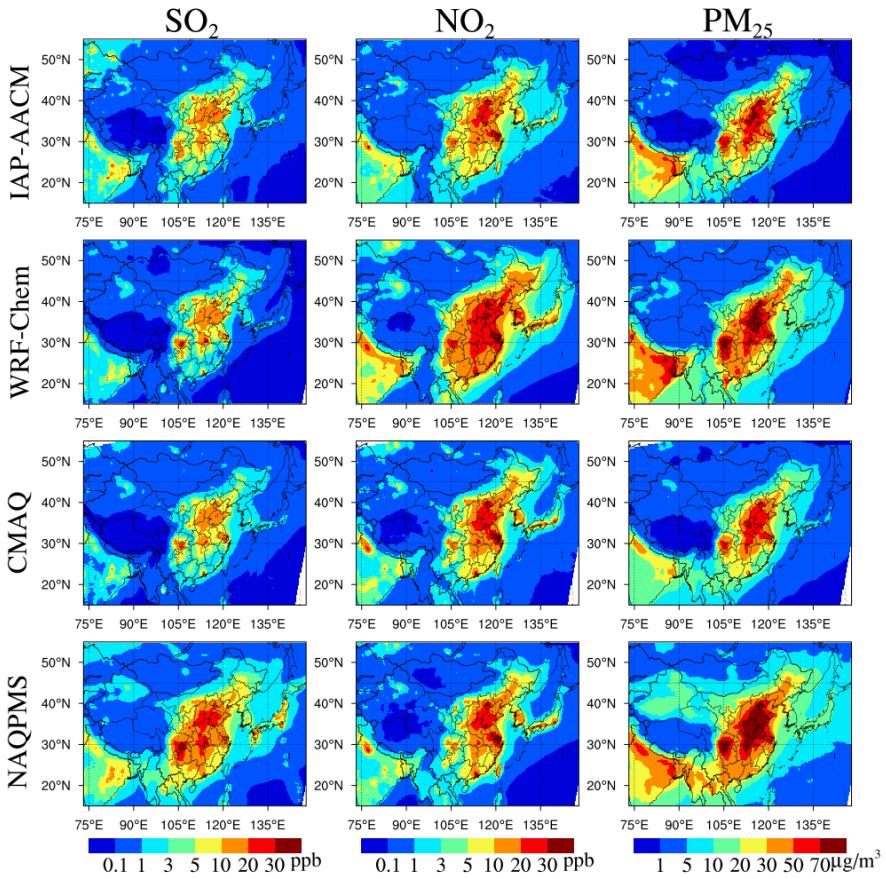

Figure 9 Annual surface distributions from nested IAP-AACM compared with regional models from MICS-Asia. Each row from top to bottom represents IAP-AACM, WRF-Chem, CMAQ and NAQPMS respectively. The left column is SO_2, the middle column is NO_2 and the right column is $PM_{2.5}$. The unit for gases is ppb and for particles is μg m^{-3}.

- L517-519: I'm sorry I can not understand what you want to mean here.

Reply: As shown in Fig. 11, the nested simulation of SO_2 in IAP-AACM has consistent spatial distribution of pollutants in the far northwest of the domain, which cannot be reproduced in the regional models. The $PM_{2.5}$ from NQAPMS is higher than IAP-AACM in Northwest of China because it includes dust aerosol in NQAPMS.

L545-547: What do you want to mean here? Your model overestimated the NO2 in summer in NC and YRD regions. If you don't use the NO2* observation, the model's overestimation should become worse.

Reply: Yes, the overestimation of NO_2 reflects the shortcoming of multiphase processes in IAP-AACM. The overestimation of NO_2 and underestimation of nitrate in daytime of summer and autumn is related to the over decomposition of nitric acid at high temperature condition in the thermodynamic equilibrium module. Moreover, heterogeneous chemical reactions in the

model should partly be responsible for the NO_2 overestimation in summer. Reactive heterogeneous uptake of gases may be crucial for the formation of secondary aerosols when the other oxidants (e.g. ozone, OH) are in low concentrations level (Jacob, 2000; Martin et al., 2003). The heterogeneous chemical module coupled in IAP-AACM has been tested in North China in winter (Li et al, 2018). The uptake of SO_2 by wet aerosols significantly enhanced sulfate formation under highly polluted conditions, contributing 50%-80% of total concentration of sulfate. The mechanism also reduced the overestimation of nitrate which is also appeared in other models. However, when it comes to the problem here, we checked the simulations excluded heterogeneous chemical processes and found a better performance of NO_2 in summer (shown in Figure 10). It implicates that a more comprehensive mechanism should be considered in model development.

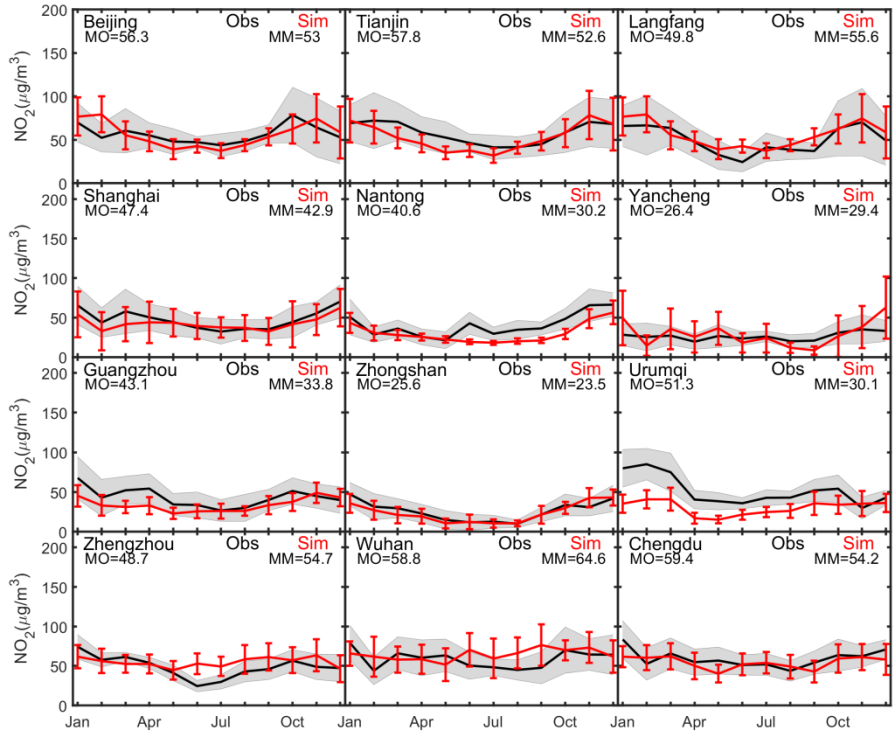

Figure 10 Seasonal cycle of NO_2 (μg m^{-3}) simulated without heterogeneous chemical process over China. The black line and red line represent monthly mean concentration of city-averaged observation and simulation respectively. Gray shaded areas and red vertical bars show 1 standard deviation over the sites for observations and for model results, respectively. MO and MM stand for annual mean concentration of observation and simulation respectively.

- L547-550: I can not understand what aspect of seasonal difference in NO2 column observation were reproduced by your model. You should describe more specifically on it.

Reply: As shown in Fig. S3, the model captured seasonal variations of NO_2 column concentrations in the vertical troposphere well. In China, the NO_2 VTC is higher during September-October-November and December-January-February while lower in June-July-August, likely caused by seasonal human activities such as fuel heating.

- L595: Typo? respects ! aspects

Reply: Yes, we have revised it.

- Conclusions should be revised according to the modifications made to respond the reviewers comments.

Reply: Yes, new conclusions will be updated in the manuscript.

References

Akritidis, D., Katragkou, E., Zanis, P.， et al. A deep stratosphere-to-troposphere ozone transport event over Europe simulated in CAMS global and regional forecast systems: analysis and evaluation. Atmospheric Chemistry and Physics, 18, 15515–15534, doi: 10.5194/acp-18-15515-2018, 2018.

Austin, J. F., Follows, M. J. The ozone record at Payerne: an assessment of the cross-tropopause flux. Atmospheric Envirnoment 25A, 1873-1880, 1991.

Dai, Y. J., Zeng, X. B., Dickinson, R. E. , Baker, I. ,Bonan, G. B. & Bosilovich, M. G.: The common land model. Bulletin of the American Meteorological Society, 84(8), 1013-1023, 2015. doi:10.1175/BAMS-84-8-1013, 2015.

Falk, S., & Sinnhuber, B. M. Polar boundary layer bromine explosion and ozone depletion events in the chemistry-climate model EMAC v2.52: implementation and evaluation of airsnow algorithm. Geoscientific Model Development, 11(3), 1-15, https://doi.org/10.5194/gmd-11-1115-2018, 2018.

Fiore，A. M., Dentener，F. J., Wild, O., et al. Multimodel estimates of intercontinental source-receptor relationships for ozone pollution. JOURNAL OF GEOPHYSICAL RESEARCH, 114, D04301, doi:10.1029/2008JD010816, 2009, 2009.

Fu， T.-M., Cao, J. J., Zhang, X. Y., Lee, S. C., & Henze, D. K.: Carbonaceous aerosols in china: top-down constraints on primary sources and estimation of secondary contribution. Atmospheric Chemistry and Physics, 12(5), 2725-2746, doi:10.5194/acp-12-2725-2012, 2012.

Galmarini, S., Koffi, B., Solazzo, E., Keating, T., Hogrefe, C., & Schulz, M., et al. Technical note: coordination and harmonization of the multi-scale, multi-model activities HTAP2,

AQMEII3, and MICS-Asia3: simulations, emission inventories, boundary conditions, and model output formats. Atmospheric Chemistry and Physics Discuss, 17(2), 1543-1555, 2017.

Han, Z., et al. MICS-Asia II: Model intercomparison and evaluation of ozone and relevant species, Atmos. Environ., 42, 3491 – 3509, doi:10.1016/j.atmosenv.2007.07.031, 2008.

Han, Z., Xie, Z., Wang, G., Zhang, R., & Tao, J. Modeling organic aerosols over east china using a volatility basis-set approach with aging mechanism in a regional air quality model. Atmospheric Environment,124, 186-198, 2016.

Hardacre, C., Wild, O., & Emberson, L.. An evaluation of ozone dry deposition in global scale chemistry climate models. Atmospheric Chemistry and Physics, 15(11), 6419-6436, doi:10.5194/acp-15-6419-2015, 2015.

Huijnen, V., Williams, J., van Weele, M., van Noije, T., Krol, M., Dentener, F., Segers, A., Houweling, S., Peters, W., de Laat, J., Boersma, F., Bergamaschi, P., van Velthoven, P., Le Sager, P., Eskes, H., Alkemade, F., Scheele, R., Nédélec, P., and Pätz, H.-W.: The global chemistry transport model TM5: description and evaluation of the tropospheric chemistry version 3.0, Geosci. Model Dev., 3, 445–473, doi:10.5194/gmd-3-445-2010, 2010.

Jacob, D. J. . Heterogeneous chemistry and tropospheric ozone. Atmos. Environ. 34, 2131–2159, 2000.

Jiang, X. , & Yoo, E. H. The importance of spatial resolutions of community multiscale air quality (CMAQ) models on health impact assessment. Science of The Total Environment, 627, 1528-1543, doi: 10.1016/j.scitotenv.2018.01.228, 2018.

Lee, Y. H., & Adams, P. J. : Evaluation of aerosol distributions in the GISS-TOMAS global aerosol microphysics model with remote sensing observations. Atmospheric Chemistry & Physics, 10(5), 2129-2144, 2010.

Li, J., Chen, X., Wang, Z., Du, H., Yang, W., & Sun, Y., et al. : Radiative and heterogeneous chemical effects of aerosols on ozone and inorganic aerosols over East Asia, Science of the Total Environment, 622, 1327-1342, https://doi.org/10.1016/j.scitotenv.2017.12.041, 2018.

Li, Y., & Xu, Y..: Uptake and storage of anthropogenic CO_2 in the pacific ocean estimated using two modeling approaches. Advances in Atmospheric Sciences, 29(4), 795-809, 2012. doi: 10.1007/s00376-012-1170-4, 2012.

Jian, L., An, J., Yu, Q., Yong, C., Ying, L., & Tang, Y., et al. Local and distant source contributions to secondary organic aerosol in the Beijing urban area in summer. Atmospheric Environment, 124, 176-185, 2016.

Lu, Z, Zhang, Q, & Streets, D. G.: Sulfur dioxide and primary carbonaceous aerosol emissions in China and India, 1996–2010. Atmospheric Chemistry and Physics, 11(18), 9839-9864, 2011.

Martin, R.V., Jacob, D.J., Yantosca, R.M. . Global and regional decreases in tropospheric oxidants from photochemical effects of aerosols. J. Geophys. Res. 108 (D3):4097. https://doi.org/10.1029/2002JD002622, 2003.

Miyazaki, K. , & Bowman, K. Evaluation of ACCMIP ozone simulations and ozonesonde sampling biases using a satellite-based multi-constituent chemical reanalysis. Atmospheric Chemistry and Physics, 17(13), 8285-8312, 2017.

Monks, P. S. A review of the observations and origins of the spring ozone maximum. Atmospheric Environment, 34(21), 3545-3561, 2000.

Munzert, K., Reiter, R., Kanter, H.-J., Potzl, K. Effect of stratospheric intrusions on the tropospheric ozone. In Proceedings of the Quad. Ozone Symposium, Halkidiki, Reidel, Dordecht, pp. 735-739, 1985.

Naik, V., Voulgarakis, A., Fiore, A. M., Horowitz, L. W., Lamarque, J.-F., Lin, M., Prather, M. J., Young, P. J., Bergmann, D., Cameron-Smith, P. J., Cionni, I., Collins, W. J., Dalsøren, S. B., Doherty, R., Eyring, V., Faluvegi, G., Folberth, G. A., Josse, B., Lee, Y. H., MacKenzie, I. A., Nagashima, T., van Noije, T. P. C., Plummer, D. A., Righi, M., Rumbold, S. T., Skeie, R., Shindell, D. T., Stevenson, D. S., Strode, S., Sudo, K., Szopa, S., and Zeng, G.: Preindustrial to present-day changes in tropospheric hydroxyl radical and methane lifetime from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmos. Chem. Phys., 13, 5277–5298, doi:10.5194/acp-13-5277-2013, 2013.

Simpson, W. R., von Glasow, R., Riedel, K., Anderson, P., Ariya, P., Bottenheim, J., Burrows, J., Carpenter, L. J., Frieß, U., Goodsite, M. E., Heard, D., Hutterli, M., Jacobi, H.-W., Kaleschke, L., Neff, B., Plane, J., Platt, U., Richter, A., Roscoe, H., Sander, R., Shepson, P., Sodeau, J., Steffen, A., Wagner, T., and Wolff, E.: Halogens and their role in polar boundary-layer ozone depletion, Atmos. Chem. Phys., 7, 4375–4418, doi:10.5194/acp-7-4375-2007, 2007.

Task Force on Hemispheric Transport of Air Pollution (HTAP), Hemispheric Transport of Air Pollution, Part A: Ozone and Particulate Matter, edited by F. Dentener, T. Keating, and H. Akimoto, United Nations Publ. ECE/EB.AIR/100, U. N. Econ. Comm. for Eur. Inf. Serv., Geneva, Switzerland. 2010.

Tsigaridis, K., Daskalakis, N., Kanakidou, M., Adams, P. J., & Zhang, X.. The aerocom evaluation and intercomparison of organic aerosol in global models. Atmos. Chem. Phys., 14(19), 6027-6161, doi:10.5194/acp-14-10845-2014, 2014.

Xu, R. T., Tian, H. Q., Pan, S. F., Prior, S. A., Feng, Y. C., & Batchelor, W. D., et al.: Global ammonia emissions from synthetic nitrogen fertilizer applications in agricultural systems: Empirical and process-based estimates and uncertainty. Global Change Biology, 25, 314-325, doi: 10.1111/gcb.14499, 2019.

Wang, L. T., Jang, C., Zhang, Y., Wang, K., Zhang, Q., & Streets, D., et al.: Assessment of air quality benefits from national air pollution control policies in china. Part I: background, emission scenarios and evaluation of meteorological predictions. Atmospheric Environment, 44(28), 3449-3457, 2010.

Wang, L. T., Wei, Z., Yang, J., Zhang, Y., Zhang, F. F., & Su, J., et al. The 2013 severe haze over southern Hebei, China: model evaluation, source apportionment, and policy implications. Atmospheric Chemistry and Physics, 14(11), 28395-28451, doi:10.5194/acp-14-3151-2014,

2014.

Werf, G. R. V. D., Randerson, J. T. , Giglio, L. , Leeuwen, T. T. V. , Chen, Y. , & Rogers, B. M. , et al. Global fire emissions estimates during 1997–2016. Earth System Science Data, 9(2), 697-720. https://doi.org/10.5194/essd-9-697-2017, 2017.

Yang, W. Y., Li, J., Wang, M., Sun Y. L., Wang, Z. F.. A Case Study of Investigating Secondary Organic Aerosol Formation Pathways in Beijing using an Observation-based SOA Box Model, Aerosol and Air Quality Research, 18: 1606–1616, doi: 10.4209/aaqr.2017.10.0415, 2018.

Zhao, B., Wang, S., Donahue, N. M., Jathar, S. H., & Robinson, A. L. .. Quantifying the effect of organic aerosol aging and intermediate-volatility emissions on regional-scale aerosol pollution in china. Scientific Reports, 6, 2016.