

Response to the editor:

We thank the Editor for a very prompt help on nominating referees, and we also appreciate the editor's thoughtful review and helpful suggestions to our manuscript. These efforts help a lot to improve the quality of our manuscript. Following are the responses to the reviewer's comments, and related revises have been incorporated into the updated manuscript.

Anonymous Referee #1

General comment: The manuscript submitted by Dong et al. reports a basic statistical analysis of 6 simulations from HTAP2 global modelling exercise, aimed at assessing the simulated impact of long-range transport of pollutants from Europe and Russia on China's haze events. The scope of the work is well defined, I think there is some gap that may be filled in terms of link with the existing literature, and there is generally no attempt by the authors in explaining the reasons for inter-model differences. The manuscript is basically a description, sometimes lengthy, of the materials presented in the figures and the tables. Considering the relevance of the topic, I think the manuscript could be published on ACP, after considering some suggestions given below, and after careful English editing.

Response: We appreciate the referee for the overall positive comment and providing the helpful detailed suggestions. We have rewrote unnecessary long sentences and read through the whole manuscript for English editing.

Specific comments:

1. In the introduction the authors very briefly review the literature regarding existing studies on haze in China. It is mentioned that long-range transport contribution to haze episodes is poorly documented (indeed they do not insert any reference). However, the literature on long-range transport to China is not null, and part of the phenomenology and underlying mechanisms might be in common with period of haze episodes. From a very quick literature search I identified, as potential references:

- Lee et al., Heavy metals and Pb isotopic composition of aerosols in urban and suburban areas of Hong Kong and Guangzhou, South China. Evidence of the long-range transport of air contaminants, *Atmospheric Environment*, Volume 41, Issue 2, January 2007, Pages 432-447
- Kong et al., Receptor modeling of PM_{2.5}, PM₁₀ and TSP in different seasons and long-range transport analysis at a coastal site of Tianjin, China, *Science of The Total Environment*, Volume 408, Issue 20, 15 September 2010, Pages 4681-4694
- Akimoto, *Global Air Quality and Pollution*, Science 05 Dec 2003: Vol. 302, Issue 5651, pp. 1716-1719 (and references therein)

I suggest to deepen the review of the literature on long-range transport from Europe to East Asia and put it into the fourth paragraph of the introduction. The same material might be subsequently used in the interpretation of some of the results illustrate afterwards (e.g. in section 3.1 and 3.2).

Response: The references listed above are closely related to our study thus they have been properly cited in the revised manuscript. The objective of this study is to evaluate the contribution of long-range transport to PM_{2.5}, with special focus on the haze episode. Previous studies about long-range transport mainly focused on the exported air pollutants from China to other areas or the transport of O₃. Although some studies (e.g., the Akimoto 2003 publication, and the HTAP Phase I report) pointed out that mitigating global air pollution requires international participations of multiple countries or continents, the contribution of long-range transport to PM_{2.5} in China remains poorly documented. The references suggested by the referee are very helpful. We also added a more detailed description of the research status about long-range transport of air pollutants to China.

2. page 4, lines 1-5: I think these very general statements, without any specific reference, on physical processes should be avoided in the manuscript. Please add proper reference and try to be more specific on the region and the situation you are referring to.

Response: Thanks for the reminder, we have added proper references (Eckhardt et al., 2003; Sthol et al., 2002) and related descriptions.

3. section 2.2: all the data versions and source of data are missing. Please add the exact product names of the data used, the web source used, and the version of the algorithms. This is necessary for the reproducibility of the work.

Response: Thanks for the comment, we have added all the products names, versions, and web sources in the revised manuscript. These details are also summarized in the following table (API, EANET, and EBAS has no version updates information, the data is downloaded from the web source).

Table. Version details and web sources of the data used for model evaluation

Data used	Web source
AERONET (Level2.0, version2)	http://aeronet.gsfc.nasa.gov
API	http://datacenter.mep.gov.cn
EANET	http://www.eanet.asia/
EBAS	http://ebas.nilu.no
MODIS (MOD08, MYD08)	https://modis.gsfc.nasa.gov

4. Figures 2 and 3 and related comments: there are some apparent inconsistency between the results presented in these figures. For example, PM_{2.5} is overestimated by GEOSCHEMADJOINT and underestimated by CHASER, but then AOD at AERONET sites has the opposite bias for these models. Why is that? Perhaps it could be useful to include a comparison only for some specific station for which all the datasets are available, or at least within the same model grid. From Figure 1 it seems to be possible for some stations.

Response: This is a very interesting point and we thank the referee for mention it. Figure 2 shows PM_{2.5} was overestimated by GEOSCHEMADJOINT by 7.5 µg/m³ (63%) in EAS (EANET and API stations), by 8.6 µg/m³ (66%) in EUR (EBAS stations). Figure 3 suggests that GEOSCHEMADJOINT underestimates AERONET-AOD by -0.08 (-23%) in EAS and overestimates AOD by 0.004(4%) in EUR. As suggested by the referee, we selected EANET-Oki (36.28°N, 133.18°E) as the PM_{2.5} site and AERONET-Osaka (34.65°N, 135.59°E) as the AOD site in EAS region, and selected EBAS-Revin (49.90°N, 4.63°E) as the PM_{2.5} site and AERONET-Brussels (50.78°N, 4.35°E) as the AOD site in EUR region. These are the closest nearby sites in each of the domain. Simulation bias of GEOSCHEMADJOINT at these sites are shown in the following figure.

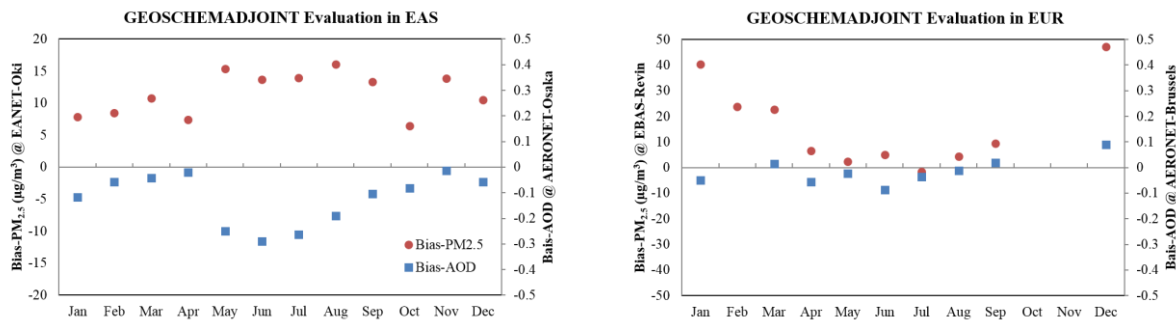


Figure. Simulation bias of GEOSCHEMADJOINT for PM_{2.5} (solid red circles) and AOD (solid blue squares)

As shown in the figure, GEOSCHEMADJOINT overestimated PM_{2.5} but underestimate AOD throughout the full year 2010 in EAS region. We examined this issue and found out there are two reasons: first, there are relatively less PM_{2.5} observation sites (2 in EAS, 5 in EUR) compared to large number of AOD

observation sites (15 in EAS, 73 in EUR). The EANET-Oki station was located on a small island ~50 miles from west coast of Japan thus represents the background concentration, while the AERONET-Osaka site is located in the downtown area of Osaka City. Evaluation in EUR region has the similar condition, the EBAS-Revin site is in a national park, while the AERONET-Brussels site is close to downtown. Although some AERONET sites are also located in remote areas, it generally has a more comprehensive representation of different surroundings including both rural and urban, but the PM_{2.5} data used in this study are most located in rural area. Second, GEOSCHEMADJOINT are reported as tend to overestimate the surface layer PM_{2.5} concentration in Asia (Figure 2 in Gu and Liao, 2016; Figure 2 in Xu et al., 2015) and underestimate the column density AOD (Figure 4 in Choi et al., 2009) in East Asia, although the explicit reason for this inconsistency hasn't been well documented. So generally the performance and evaluation results of this HTAP Phase 2 modeling effort is consistent with those previous studies.

References:

Yi-Xuan GU & Hong LIAO (2016) Response of fine particulate matter to reductions in anthropogenic emissions in Beijing during the 2014 Asia–Pacific Economic Cooperation summit, *Atmospheric and Oceanic Science Letters*, 9:6, 411-419, DOI: 10.1080/16742834.2016.1230465

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, <https://doi.org/10.5194/acp-15-13133-2015>, 2015.

5. Figure 4 and related comments: the modelled AOD over China and elsewhere in the domain differ among models by more than a factor of two. As for previous results on point measurements, there is no attempt to explain the differences. For example, considering the same anthropogenic emissions, the difference over China CHASER and SPRINTARS is quite remarkable.

Response: We also notice the large difference between model performances with the same emission inputs. Explicitly clarify the causes of the difference would require deep detailed investigation of the model schemes, algorithms, and parameterization, which is not within the scope of this study. But the other HTAP Phase2 related studies (Im et al., 2018; Palacios-Peña et al., 2018; Astitha et al., 2018) do present a few investigations into the multi-model comparison between the models used in this study, and the different model performance are attributed to meteorology (in particular wind speed and PBL height), different aerosol mechanisms, treatment of wind-blown dust emission, and biomass burning emission injection heights. Previous multi-modeling efforts such as the AEROCOM also pointed out that these aspects can lead to modeled AOD and surface PM concentration differ by a factor of 2 and 10 respectively, although the some AEROCOM participating models are different from HTAP. We agree with the referee that briefly explain the difference is necessary as our discussion is based on multi-model simulations, so we have added a short discussion in the revised manuscript. The above-mentioned references are also added into the revised manuscript.

References:

Im, U., Christensen, J. H., Geels, C., Hansen, K. M., Brandt, J., Solazzo, E., Alyuz, U., Balzarini, A., Baro, R., Bellasio, R., Bianconi, R., Bieser, J., Colette, A., Curci, G., Farrow, A., Flemming, J., Fraser, A., Jimenez-Guerrero, P., Kitwiroon, N., Liu, P., Nopmongkol, U., Palacios-Peña, L., Pirovano, G., Pozzoli, L., Prank, M., Rose, R., Sokhi, R., Tuccella, P., Unal, A., Vivanco, M. G., Yarwood, G., Hogrefe, C., and Galmarini, S.: Influence of anthropogenic emissions and boundary conditions on multi-model simulations of major air pollutants over Europe and North America in the framework of AQMEII3, *Atmos. Chem. Phys.*, 18, 8929-8952, <https://doi.org/10.5194/acp-18-8929-2018>, 2018.

Palacios-Peña, L., Jiménez-Guerrero, P., Baró, R., Balzarini, A., Bianconi, R., Curci, G., Landi, T. C., Pirovano, G., Prank, M., Riccio, A., Tuccella, P., and Galmarini, S.: Aerosol optical properties over Europe: an evaluation of the AQMEII Phase 3 simulations against satellite observations, *Atmos. Chem. Phys. Discuss.*, <https://doi.org/10.5194/acp-2017-1119>, in review, 2018.

Astitha, M., Kioutsoukis, I., Fisseha, G. A., Bianconi, R., Bieser, J., Christensen, J. H., Cooper, O., Galmarini, S., Hogrefe, C., Im, U., Johnson, B., Liu, P., Nopmongcol, U., Petropavlovskikh, I., Solazzo, E., Tarasick, D. W., and Yarwood, G.: Seasonal ozone vertical profiles over North America using the AQMEII group of air quality models: model inter-comparison and stratospheric intrusions, *Atmos. Chem. Phys. Discuss.*, <https://doi.org/10.5194/acp-2018-98>, in review, 2018.

6. Figures 5-6 and related comments. The figures are interesting because they nicely illustrate the model diversity. For example, the seasonal cycle of contributions from some models is opposite to that of others (e.g. CAM-Chem peaks in summer, CHASER in winter, and GEOS5 in spring). It would be useful to have some inspection of these difference. I suspect that differences in the meteorological fields used in these models are responsible for the large variability.

Response: CAM-chem showed the largest PM response in summer under EUR emission reduction scenario, and SPRINTARS showed the largest PM response in summer under RBU emission reduction scenario, while the other models all showed larger PM responses in winter or spring. We agree with the referee that meteorology difference might be one of the reasons for simulation diversity. We examined the surface air temperature used by the participating models. Domain averages of monthly temperature over EUR and RBU are shown in the following figure.

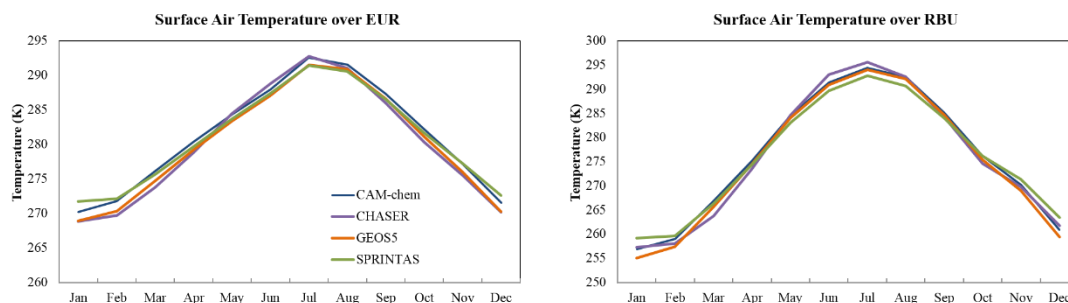


Figure. Monthly surface air temperature from CAM-chem, CHASER, and GEOS5

In EUR region, CAM-chem and SPRINTARS simulated surface air temperatures are systematically higher than other models by $\sim 2.5\text{K}$ in winter. A higher temperature in the emission source region may facilitate the PM precursors' chemistry and subsequently allow less precursors enter long-range transport. In RBU region, SPRINTARS simulated temperature is $\sim 2\text{K}$ lower than other models in summer, which may lead to more precursors transport into EAS and subsequently induce larger PM response. But on the other hand, temperature is apparently not the only influencing factor as CAM-chem showed highest temperature in summer over EUR region yet largest PM response in summer too. Wind speed and PBL height may play a more important role as indicated by Im et al. (2018), but unfortunately only one of the participating model provided wind and PBL data. Explicitly identify the impact of meteorology on modeled PM response would require a set of more detailed experiments, and this is beyond the scope of HTAP program. CAM-chem applied the modified Zhang-McFarlane approach (Zhang and McFarlane, 1995) with shallow convection follows Hack et al. (2006). GEOS5 applied the modified scheme by Moorthi and Suarez (1992), which is a relaxed Arakawa-Schubert algorithm. These schemes are functionalized and parameterized substantially different and will subsequently lead to differences of aerosol vertical distribution, lifetime, transport, and total suspended aerosol concentration in the atmosphere (Stjern et al., 2016). Aerosol parameterization also lead to different $\text{PM}_{2.5}$ formula. CAM-chem simulates secondary organic aerosol (SOA) with the 2-product approach using laboratory-determined yields from photooxidation of monoterpenes, isoprene and aromatics, while GEOS5 has no SOA. The differences of aerosol scheme, heterogeneous chemistry, treatment of OC, OA, and SOA lead to additional inter-model variability. In addition, grid resolutions diversity is also responsible as Molod et al (2015) demonstrated that different grid resolutions will result in different scavenging aerosol even with

the same model. In fact, not only the PM_{2.5} responses but also the baseline PM_{2.5} concentrations show prominent different seasonality among the models in both the HTAP Phase1 (Dentener et al., 2010) and Phase2 program, and this is also why multi-model mean is used to estimate the source-receptor relationship. We have added the abovementioned discussion and references in the revised manuscript.

7. Figures 9-10: some panels look patchy, for example EMEP, SPRINTARS and all in Figure 10. Why is that?

Response: Figures 9-10 are designed to demonstrate the full impact of long-range transport during the haze episode, so the NCDC surface observation data is used to identify where and when haze exist. For those with finer grid resolution such as EMEP (0.5×0.5°) and SPRINTARS (1.1×1.1°), there are some model grids having no NCDC observation site, and these grids are filled with missing value, and this makes the figures look patchy. Although haze (visibility) can also be estimated with aerosol extinction coefficient, using the direct measurements from NCDC is apparently a more solid method to identify haze, and only SPRINTARS provides the aerosol extinction coefficient data. We have added a short sentence in the figure caption to explain the patchy panels to avoid misleading.

8. I recommend English editing of the manuscript. The use of language is imaginative and makes understanding difficult. A few random examples:

- p. 3, l. 40-41: "These datasets are essential to estimate surface PM response compare the aerosol transport in different atmosphere layers". What is "response compare"? "atmosphere" → "atmospheric"

Response: We agree that English editing is necessary, it's also pointed out by another referee. This sentence is removed because it is not necessary.

- p. 5, l. 6-7: "the models all tend to underestimate the high peaks in spring (Mar.-Apr.) and low bottoms in summer". Not clear what "low bottoms" means.

Response: The sentence is changed to "Temporal variation of O₃ is also simulated well in EAS, although the models all tend to underestimate the high values in spring (Mar.-Apr.) and low concentrations in summer (Jul.-Sep.)"

- note 2 on caption of Figure 2: "PM_{2.5} observations in EUR and EAS region have no standard because there are no sites with valid measurements fall into the same model ensemble mean grid". Very difficult to understand: why a standard deviation cannot be calculated even if stations are not in the same model cell?

Response: The standard deviation is calculated between the observations from different sites in the same model grid, we have mentioned in the caption of Figure 2 that "vertical error bars depict the standard deviation across the sites in the same ensemble grid."

Anonymous Referee #2

General comment: This paper presented a work analyzing the contribution from Europe to China's atmospheric particle concentrations and haze events, with intensive chemistry transport modeling. The authors made great efforts on incorporating multiple transport models to understand the difference between models and to reduce the uncertainty of simulation. They also evaluated the impacts of emission inventory on the simulation, as the accuracy of emission inventory for anthropogenic pollutants is always a big concern on the air quality research community. Before it can be accepted as a final atmospheric chemistry physics paper, however, the following issues need to be further discussed or stressed.

Response: We thank the referee for the encouraging comment and providing insightful suggestions.

1. The significance of the paper needs to be reconsidered and relevant statement should be revised. In current format, the authors stated that there were limited studies conducted on regional transport to China and it might be important as China is controlling its emissions. The results, however, show that the impacts of Europe was very few, and the studying period was before 2010, during which the emissions in East Asia were expected to still increase. It seems that the current work did not fully answer the question they raised. The most serious haze events after 2010 were not included in this study?

Response: We have added some references documenting the long-range transport into China (Lee et al., 2007; Kong et al., 2010; Akimoto 2003; Fu et al., 2012) and a short introduction of their findings. In our study, Table 2 summarized the annual average long-range transport contribution from EUR and RBU regions to EAS region in year 2000 and 2008-2010. Table 3 summarized the long-range transport contributions during the haze episode. We raised the research question in "Introduction" section that "the background concentrations of PM and the contributions from outside China import of air pollutants to the haze problem, is poorly documented." So the question is answered by Tables 2-3 and the related discussions. The severe haze event in 2010 is included in this study but not specifically highlighted. Some places in China has more than 300 days with haze identified with NCDC observation. We analyzed the annual total haze events and reported the contribution of long-range transport to these events, as shown with Figures 9-10. An overview of the haze events for full year 2010 is provided in supplementary material Figure S1.

2. Lines 9-10, Page 2: this results is quite old, there are recently more studies on health impacts of China's air pollution.

Response: We have added several most recent studies that reported the premature deaths attributable to PM_{2.5} pollution in China from 2013 to 2017, these references include: Huang et al., 2018; Cao et al., 2017; Li et al., 2018.

3. Lines 13-16, Page 6: Please define how the MNB was calculated. Is there any criterion indicating the acceptable range of MNB?

Response: The MNB is calculated as:

$$MNB = \frac{1}{S} \sum_{i=1}^S \sum_{j=1}^T \frac{sim_{ij} - obs_{ij}}{obs_{ij}}$$

Where S is the number of observation stations, T is the total number of month, obs_{ij} is the observed value at the station i and month j , and sim_{ij} is the corresponding simulation value at the closet grid point to the station. There is no well documented threshold for an acceptable MNB, but the AERCOM research program have been frequently cited by the research community, so we used their MNB values to demonstrate our participating models' performance.

4. There are limited PM_{2.5} observations used in model evaluation for China. I understand that the official data were not available until 2013. However, could the data published in previous studies be available and could the evaluation be improved?

Response: This is a very insightful comment and we agree that there many some publications reporting the measured PM_{2.5} concentrations in China (Zhang and Cao, 2016; Lowson and Conway, 2016), but the diversities in instrument, measuring method, and sampling period make it difficult to develop a consistent observation database from the literatures. In addition to the potential uncertainties within each individual measurement literature, these measurements are usually presented with charts or figures so we would have to roughly read the values from the figures, which may introduce more uncertainty and is not proper for the HTAP program as it requires applying official downloadable data by all participating groups so all experiments and analysis could be reproduced. In contrast, the EANET data used in this study provides measurement collected with the same type of instrument and method. Considering the limited number of EANET sites, we also included AERONET and MODIS AOD which are all public accessible for model evaluation with better spatial and temporal coverage. In addition, since examining long-range transport of surface PM_{2.5} into China is the main objective of this study, evaluating models performances with literature review collected data would require intensive efforts and make the manuscript lengthy.

5. Figures 5 and 6 illustrated the surface aerosol response under EURALL and RBUALL. Can you explain why the seasonal variations were different between models? In Fig 6, for example, larger response was found in summer for CAM-chem and SPRINTARS, while smaller was found in summer in EMEP and GEOSCHEMADJOINT. By the way, caption of Figure 6 should be revised (Figure XXX?).

Response: We apologize for the typo in the figure caption, it is corrected in the revised manuscript. We agree with the referee that prominent different seasonality was found between CAM-chem and other models. Despite applying the same emission inputs, several aspects of the participating models lead to the different seasonality of PM_{2.5} response. One of the other two referees also pointed out this issue, and we briefly probe into these aspects. These aspects including the meteorology, aerosol mechanisms, and convection mechanisms. We first examined the meteorology differences by comparing the model simulated air temperature the following figure shows domain average monthly mean surface air temperature from CAM-chem, CHASER, GEOS5, and SPRINTARS. In EUR region, CAM-chem and SPRINTARS simulated surface air temperatures are systematically higher than other models by ~2.5K in winter. A higher temperature in the emission source region may facilitate the PM precursors' chemistry and subsequently allow less precursors enter long-range transport. In RBU region, SPRINTARS simulated temperature is ~2K lower than other models in summer, which may lead to more precursors transport into EAS and subsequently induce larger PM response. But on the other hand, temperature is apparently not the only influencing factor as CAM-chem showed highest temperature in summer over EUR region yet largest PM response in summer too. Wind speed and PBL height may play a more important role as indicated by Im et al. (2018) but unfortunately only one of the participating model provided wind and PBL data.

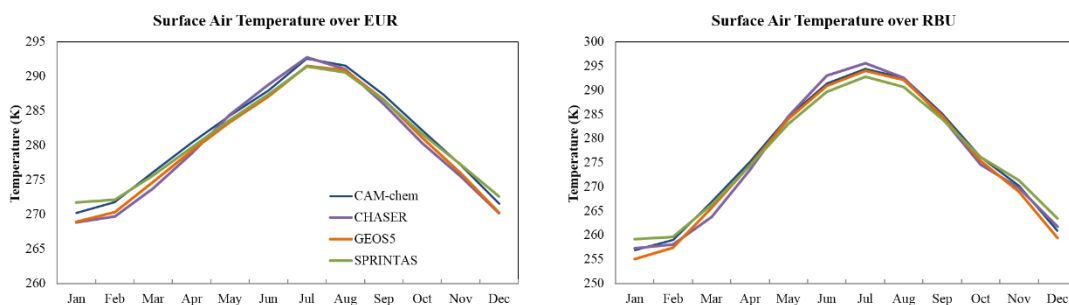


Figure. Monthly surface air temperature from CAM-chem, CHASER, and GEOS5

We then examined the convection schemes among models. CAM-chem applied the modified Zhang-McFarlane approach (Zhang and McFarlane, 1995) with shallow convection follows Hack et al. (2006). GEOS5 applied the modified scheme by Moorthi and Suarez (1992), which is a relaxed Arakawa-Schubert algorithm. These schemes are functionalized and parameterized substantially different and will subsequently lead to differences of aerosol vertical distribution, lifetime, transport, and total suspended aerosol concentration in the atmosphere (Stjern et al., 2016). Aerosol parameterization also lead to different PM_{2.5} formula. CAM-chem simulates secondary organic aerosol (SOA) with the 2-product approach using laboratory-determined yields from photooxidation of monoterpenes, isoprene and aromatics, while GEOS5 has no SOA. The differences of aerosol scheme, heterogeneous chemistry, treatment of OC, OA, and SOA lead to additional inter-model variability. In addition, grid resolutions diversity is also responsible as Molod et al (2015) demonstrated that different grid resolutions will result in different scavenging aerosol even with the same model. In short summary, the abovementioned aspects may all contribute to the different seasonality of PM_{2.5} response, and more a set of more specifically designed model experiments is necessary to explicitly identify their influences, yet this is beyond the current scope of HTAP program. We have added a short discussion of the seasonality in the revised manuscript to point out this issue with the clues mentioned here.

References:

Hack, J. J., Caron, J. M., Yeager, S. G., Oleson, K. W., Holland, M. M., Truesdale, J. E., and Rasch, P. J.: Simulation of the Global Hydrological Cycle in the CCSM Community Atmosphere Model Version 3 (CAM3): Mean Features, *J. Climate*, 19, 2199–2221, doi:10.1175/JCLI3755.1, 2006.

Molod, A., Takacs, L., Suarez, M., and Bacmeister, J.: Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2, *Geosci. Model Dev.*, 8, 1339–1356, doi:10.5194/gmd-8-1339-2015, 2015.

Moorthi, S. and Suarez, M. J.: Relaxed Arakawa-Schubert. A Parameterization of Moist Convection for General Circulation Models, *Mon. Weather Rev.*, 120, 978–1002, doi:10.1175/1520-0493(1992)120<0978:RASAPO>2.0.CO;2, 1992.

Zhang, G. J. and McFarlane, N. A.: Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian climate center general circulation model, *Atmos.-Ocean*, 33, 407–446, 1995.

6. Lines 23-27, Page 10. I am not persuaded by the authors by the linearity assumption. They estimated the full impact by scaling the PM responses under the 20% emissions perturbation conditions by a factor of 5. If this is the case, why not directly estimate the PM response by removing 100% emissions for given region? Nonlinearity of PM_{2.5} concentrations to precursor emissions was strong, and the uncertainty of the assumption should be carefully analyzed and quantified.

Response: We agree with the referee that PM_{2.5} concentrations to precursor emissions are strong, and this is the merit of applying atmospheric models to simulate the “real” aerosol response instead of simply estimating it with a certain emission change ratio. The 20% emission perturbation is the first priority of model experiment designed by the HTAP Phase2 program because it is a relatively reasonable and applicable control rate for air quality management. The impact of long-range transport however, indicates the overall contribution of the total emission in the source regions, so 100% emission reduction would be a stronger but unrealistic experiment. While the 100% emission perturbation simulation is not available, the “full impact” calculated from 20% emission perturbation is the best estimates we can derive. This method was applied by several investigations for estimating inter-continental transport of O₃ (Fiore et al., 2009; Lin et al., 2010; West et al., 2009; Zhang et al., 2009), which has an even more significant nonlinear response to the precursors. But as the referee mentioned, this method may introduce uncertainty due to the nonlinear response, and we also noticed this issue while analyzing the modeling data. We have applied the Response Surface Method (RSM) with Hemispheric-CMAQ model to quantify the source-

receptor relationship with more detailed simulation design than HTAP. We are analyzing the data now and expect to report the findings later.

Anonymous Referee #3

General comments: The manuscript submitted by Dong et al. assesses how changes in aerosol emissions in Europe and Russia influences haze events in China, using simulations from the HTAP2 project. Analyses include a thorough model evaluation towards various surface and satellite-based observation data, presentations of the seasonality of the long-range impacts from the two regions on China, evaluation of how the long-range impacts are distributed between within- and above-PBL layers, comparison of results to findings for earlier years, as well as an analysis of estimated horizontal visibility and how this variable is affected by the two source regions. The authors have performed many and rigorous analyses, and the results are likely to be of broad interest to the community. There are, however, some issues that need to be resolved before the paper should be accepted for publishing.

General comments: -The language of the manuscript could greatly benefit from a thorough read-through by a person fluent in English. - The manuscript is at times unnecessarily lengthy. I have suggested several sentences that could be removed, but going through the manuscript and removing sentences and statements that contain irrelevant information or information that has already been given, will help the reader.

- In the Introduction, it would be good to see a bit more background on haze in China – for instance, write out in more detail what the references around lines 15-20 find. Do that Wang studies referred to on line 19 look at sources in China only, or is there an element of long-range influences here that could be relevant for this study?

Response: We intended to provide a brief introduction of the research topics related with haze in China, so detailed findings of these references are not described. The findings of these publications are not used in our study. We listed them in the “Introduction” section as part of the background information. None of the literatures intended to quantify the contribution of long-range transport to haze in China, and this is why we conduct our study.

- The “Results and Discussion” section is at times too much description of figures and numbers, and too little discussion of results. I believe a lot of the numbers could be put in a table so that more time can be spent on the main highlights and how they agree/differ from other findings. There are several interesting results and features here that deserve to be accentuated.

Response: We agree with the referee that language of the original submission shall be greatly improved, and thank the referee for providing other detailed suggestions and comments. We have carefully go through the whole manuscript to improve the writing, by removing some of the descriptions of the numbers and adding more in-depth discussions, such as the possible reason for model diversity.

Specific comments:

1. P1 L37: add “from EUR” after “long-range transport”?

Response: It has been added in the revised manuscript.

2. P1 L38: change “aerosol response” to “the aerosol response in EAS”

Response: It has been added in the revised manuscript.

3. P1 L44: to compare how much 1-3 days change in haze frequency is to the percentages given above, please consider changing 1-3 days to percent change

Response: The percent change has been added in the revised manuscript.

4. P2 L12: It is a bit difficult to catch the meaning of the sentence starting with “Although” – a rewording would be good!

Response: The long sentence has been revised as “Some pilot studies have tried to explore the understanding of haze in China.”

5. P2 L28: Not necessary to introduce the AQMEII and MICS-Asia projects, as data from these are not used in the present study? Instead, line 26 could instead start with “One of these is the Task Force on Hemispheric ...”

Response: The description of AQMEII and MICS-Asia is removed and the sentence has been revised.

6. P2 L34: These last two sentences are not strictly necessary.

Response: These sentences are removed.

7. P3 L27: The first part of this sentence “To quantify ... sensitivity simulations” is superfluous – one could instead start directly at “Emission perturbations are conducted with all..”

Response: That sentence is changed according to the comment.

8. P3 L31: Fix reference Guido R. can der Werf?

Response: The reference is removed, it shall be: “(1) BASE scenario with all baseline emissions”

9. P3 L41: the sentence starting with “These datasets are essential” can also be removed.

Response: The sentence is removed.

10. P4 L2: Here you could stop after “descend into the PBL.” And then start a new sentence motivating the remaining text by stating the relevance of the PBL-analysis to haze (for instance, that pollutants within the PBL give more haze, and therefore it is necessary to understand the contributions of within- and above- PBL).

Response: These two sentences have been revised according to the comment, and necessary references are added according to the suggestion of another referee.

11. P4 L37: I may have missed something, but P3 L19 says that all models have resolution of 0.1x0.1 – where does the 2.8x2.8 come from? Please clarify.

Response: Page4-Line19 states the resolution of the emission inventory, to avoid misleading, that sentence is changed to “The emissions are compiled from several regional inventories for the year 2010 with monthly temporal resolution and 0.1°x0.1° grid resolution”. Page4-Line37 states that the grid resolution of ensemble mean is 2.8°x2.8°, that sentence is changed to “ ... model ensemble mean, calculated as the average of all participating models at 2.8°x2.8° grid resolution.”

12. P5 L1: Please define MB

Response: “MB” refers to “mean bias”, it has been added in the updated version.

13. P5 L5: Please consider replacing all uses of “temporal” in this section with “seasonal”, as the “temporal” gives an impression of temporal (year-to-year) development.

Response: It has been replaced in the updated version.

14. P5 L6: you write that models tend to underestimate the high peaks in spring, but Fig. 2d seems to me to show that models overestimate in spring (or at least all models are higher in March, and the observations are in the midst of the models in April)?

Response: The referee is correct, we intended to emphasize that “model overestimate high values in spring”, it has been fixed in the updated version.

15. P5 L15: Remove “shows significantly ... than the others”, which is given from the previous sentence.

Response: The sentence has been removed.

16. P5 L16: Do you have any data on the occurrence or tendency for wildfires near this specific stations? If not, this comment should perhaps be removed.

Response: The sentence has been removed.

17. P9 Section 3.3 heading: I am a bit skeptical to the use of the word “Trend” in this heading and in the section text, as a trend can hardly be quantified based on a comparison between the years 2000 and 2010 (data for years 2008 and 2009 helps, but the data are scarce). Consider changing “trend” to “change” or something similar.

Response: It has been changed in Section3.3.

18. P9 L39: Please add a reference after “the past decade.”

Response: The reference “Li et al., 2018” is added because it compares emission changes in China.

19. P11 L25: How would the results look if you use CAM-chem only for all the years?

Response: The numbers (in Table 2 for RBU scenario in 2008 and 2009) are indeed calculated with CAM-chem only.

20. P12 L16: ECE → CEC?

Response: Yes, it has been changed to “CEC”

21. P13 L10: “The participating models ... to 5.5%” can be removed as it has just been said above.

Response: It has been removed.

22. P13 L14: It says Frequency_Full_Impact15 twice :)

Response: The second one is removed.

23. P13 L34: Please give this in % change as well.

Response: The percent change is added.

24. P14 L37: Please add references after “recent years”.

Response: The following references are added:

Chen, Y., Schleicher, N., Fricker, M., Cen, K., Liu, X. L., Kaminski, U., Yu, Y., Wu, X. F., and Norra, S.: Long-term variation of black carbon and PM_{2.5} in Beijing, China with respect to meteorological conditions and governmental measures, *Environ Pollut*, 212, 269-278, 10.1016/j.envpol.2016.01.008, 2016.

Feng, J. L., Zhong, M., Xu, B. H., Du, Y., Wu, M. H., Wang, H. L., and Chen, C. H.: Concentrations, seasonal and diurnal variations of black carbon in PM_{2.5} in Shanghai, China, *Atmos Res*, 147, 1-9, 10.1016/j.atmosres.2014.04.018, 2014.

Lu, Z., Streets, D. G., Zhang, Q., Wang, S., Carmichael, G. R., Cheng, Y. F., Wei, C., Chin, M., Diehl, T., and Tan, Q.: Sulfur dioxide emissions in China and sulfur trends in East Asia since 2000, *Atmos Chem Phys*, 10, 6311-6331, 10.5194/acp-10-6311-2010, 2010.

Zhu, J. L., Liao, H., and Li, J. P.: Increases in aerosol concentrations over eastern China due to the decadal-scale weakening of the East Asian summer monsoon, *Geophys Res Lett*, 39, L0980910.1029/2012gl051428, 2012.

Long-range Transport Impacts on Surface Aerosol Concentrations and the Contributions to Haze Events in China: an HTAP2 Multi-Model Study

5 Xinyi Dong¹, Joshua S. Fu¹, Qingzhao Zhu¹, Jian Sun¹, Jiani Tan¹, Terry Keating², Takashi Sekiya³, Kengo Sudo³, Louisa Emmons⁴, Simone Tilmes⁴, Jan Eiof Jonson⁵, Michael Schulz⁵, Huisheng Bian⁶, Mian Chin⁷, Yanko Davila⁸, Daven Henze⁸, Toshihiko Takemura⁹, Anna Maria Katarina Benedictow⁵, Kan Huang^{1,10}

¹Department of Civil and Environmental Engineering, The University of Tennessee, Knoxville, Tennessee, USA

10 ²Environmental Protection Agency, Applied Science and Education Division, National Center for Environmental Research, Office of Research and Development, Headquarters, Federal Triangle, Washington, DC 20460, USA

³Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Japan

15 ⁴Atmospheric Chemistry Observations and Modeling Laboratory, National Center for Atmospheric Research, Boulder, Colorado, USA

⁵Norwegian Meteorological Institute, Oslo, Norway

⁶Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore, MD, USA

⁷Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, MD, USA

⁸Department of Mechanical Engineering, University of Colorado, Boulder, CO, USA

20 ⁹Research Institute for Applied Mechanics, Kyushu University, Fukuoka, Japan

¹⁰Center for Atmospheric Chemistry Study, Department of Environmental Science and Engineering, Fudan University, Shanghai 200433, China

Correspondence to: Joshua S. Fu (jsfu@utk.edu)

25

Abstract

Haze has been severely affecting the densely populated areas in China ~~during-recently-recent-years~~. While many of the ~~pilot-studies-efforts~~ have been devoted to investigate the ~~contributions-impact-of-from~~ local anthropogenic emission, limited attention has been paid to the ~~influence-contribution~~ from long-range transport. In this study, we ~~use-apply~~ simulations from 6 participating models supplied through the Task Force on Hemispheric Transport of Air Pollution Phase 2 (HTAP2) exercise to investigate the long-range transport impact of Europe ~~(Tie-et-al.)(EUR)-(Tie-et-al.)~~ and Russia/Belarus/Ukraine (RBU) on the surface air quality in East Asia (EAS), with special focus on their contributions during the haze episodes ~~over-in~~ China. The impact of 20% anthropogenic emission perturbation from the source region is extrapolated by a factor of 5 to estimate the full impact. We find that the full impacts from EUR and RBU are $0.99\mu\text{g}/\text{m}^3$ (3.1%) and $1.32\mu\text{g}/\text{m}^3$ (4.1%) respectively during haze episodes, while the annual averaged full impacts are only $0.35\mu\text{g}/\text{m}^3$ (1.7%) and $0.53\mu\text{g}/\text{m}^3$ (2.6%) respectively. By estimating the aerosol response within and above the planetary boundary layer (PBL), we find that long-range transport ~~from EUR~~ within the PBL contributes to 22-38% of the total column density of aerosol response ~~in EAS~~. Comparison with the HTAP Phase 1 (HTAP1) assessment reveals that from 2000 to 2010, the long-range transport from Europe to East Asia has decreased significantly by a factor of 2-10 for surface aerosol mass concentration due to the simultaneous emission reduction in source region and emission increase in the receptor region. ~~By investigating the visibility response, we~~ We also find ~~that~~ the long-range transport from the Europe and RBU region increases the number of haze events in China by 0.15% and 0.11% respectively, and the North China Plain and southeast China ~~receives-gets~~ 1-3 extra haze days (~~<3%~~). This study is the first investigation into the contribution of long-range transport to haze in China with multi-model experiments.

45

1. Introduction

Frequent low visibility due to heavy haze has been one of the most important environmental concerns in China ~~recently during recent years~~. Long-term monitoring data suggests that visibility degradation has been identified during the past 30 years over North China Plain, Pearl River Delta, and Yangtze River Delta (Fu et al., 2014; Wang et al., 2014a), where more than 40% of the national population is hosted. As the most apparent symptom of air pollution, visibility degradations induced by haze not only interrupt highway and airline operations, but also indicate critical deterioration of public health. The China Ministry of Environmental Protection (MEP) reported that air quality in 265 of the 338 major cities failed to attain the national air quality standard in 2015 (Wang, 2017), and ~~World Bank studies~~ also suggests that 350,000-400,000 annually premature deaths are attributable to air pollution exposure (WorldBank, 2007; Cao et al., 2017; Li et al., 2018) in China ~~during the past decade~~.

China haze is usually associated with high concentrations and rapid hygroscopic growth of fine particulate matters ⁰¹(Im et al., 2018). ~~Although only recently has public attention been centered on PM_{2.5} with national level regular measurement data released since 2013, many~~ Some pilot studies have ~~already been conducted for years~~ tried to explore the understanding of haze in China. The topics attracted most of the ~~focused on the~~ research topics including ~~efforts are focused on:~~ ambient air quality conditions under haze condition (Huang et al., 2012; Wang et al., 2015), spatial distribution and long-term trend of haze in China (Fu et al., 2014), meteorology conditions that favor the formation of haze (Wang et al., 2014a), chemical components and size distributions of aerosols (Guo et al., 2014; Ho et al., 2016; Shen et al., 2017; Yin et al., 2012; Zhang et al., 2012), source apportionment of fine particles during haze episodes (Hua et al., 2015; Wang et al., 2014b; Wang et al., 2014c), and also public health impact of haze (Gao et al., 2017; Tie et al., 2009; Xu et al., 2013).

Although these ~~pilot studies shed light on~~ helped improve the fundamental understanding of ~~fundamental characteristics of haze in China~~, very limited attention has been paid to reveal the role of long-range transport ~~estimate the contribution from outside the country through long range transport. (Huang et al., 2012).~~ Research community has realized the hemispheric transport could also exacerbate local regional and even inter-continental air quality problems since early of the 20th century (Akimoto, 2003), and several international collaborated programs have been initiated to investigate the long-range transport of air pollutants since then (Carmichael et al., 2008; Rao et al., 2011). One of these is ~~The~~ Task Force on Hemispheric Transport of Air Pollution (TF HTAP) ~~is~~, designated to advance the understanding of inter-continental transport of air pollutants in the Northern Hemisphere (~~Dentener et al., 2010~~ Streets et al., 2010); ~~Air Quality Modeling Evaluation International Initiative (AQMEII) coordinated by the Joint Research Centre/IES, Environment Canada, and United States Environmental Protection Agency (U.S. EPA) aims at promoting research on transatlantic transport of air pollutants (Rao et al., 2011), the Model Inter-Comparison Study Asia (MICS Asia) organized by multiple research institutions from East Asia and Southeast Asia countries devoted their efforts to develop the local anthropogenic emissions and estimate the source apportionment of acid depositions and air pollutants transport within Asian (Carmichael et al., 2008). These cooperative efforts pave the way for the research community to probe into the air pollution problem with an international perspective, and enable the exchange of information between countries/regions for keeping in focus of policy needs as well.~~

The abovementioned prior efforts however, have limited assessment of long-range transport impact on haze, ~~which is likely due to the fact that haze is a relatively recent topic and mostly causing problems~~

~~in China and India.~~ In order to achieve a better air quality condition and reduce the frequency of haze events, China is ~~spending investing~~ billions to reduce the local anthropogenic emissions (Li and Zhu, 2014; Liu et al., 2015) ~~and trying to stop scarifying the environment benefit for economic growth.~~ However, the background concentrations of PM and the contributions from ~~long-range transport outside China import of air pollutants to the haze problem,~~ is poorly documented. ~~A few studies have demonstrated the existence of long-range transport into China with campaign measurements (Lee 2007; Kong 2010) and attempted to quantify O₃ response in East Asia due to intercontinental transport (Fu et al., 2012), but the contribution of external emissions to China's PM_{2.5} pollution remains unknown.~~ Understanding of the long-range transport impact is essential to estimate the background concentrations of air pollutants and estimate the efficiency and effectiveness of local emission control, it is also an important scientific support for policy makers to better organize the international collaborations.

In this study, we evaluate the long-range transport impact on haze in China by estimating the PM concentration response and visibility change based on multi-model data provided ~~by through~~ the second phase of HTAP (HTAP2). We focused on transport from two source regions designed by the HTAP2 framework: Europe ~~(Tie et al.)(Tie et al.)(EUR)~~ and Russia/Belarus/Ukraine (RBU) since they are the most important upper wind areas with respect to East Asia (EAS) as the receptor region. Modeling framework and baseline evaluation is described in section 2. Results and discussions are summarized in section 3, including the demonstration of long-range transport seasonality, comparison of PM transport above and within the planetary boundary layer (PBL), the assessment of full impact and relative importance of long-range transport, and also the contributions during haze episodes in China. Conclusions are summarized in section 4.

2. Method

2.1 Models, emissions, and simulations configuration

The HTAP2 participating models all utilize the same anthropogenic emission inventories for SO₂, NO_x, CO, non-methane VOC (NMVOC), NH₃, PM₁₀, PM_{2.5}, black carbon (BC) and organic carbon (OC), ~~which~~ The emissions are compiled from several regional inventories for the year 2010 with monthly temporal resolution and 0.1°×0.1° grid resolution, with more details reported in Janssens-Maenhout et al. ~~(Janssens-Maenhout et al., 2015).~~ Emissions of year 2008 and 2009 are also prepared in the same format as that of 2010 through the HTAP2 effort, yet model simulations for these two years are of lower priorities. So in this study we ~~will~~ mainly focus on the 2010 model experiments, ~~but and briefly also~~ probe into the inter-annual variability by utilizing the 2008 and 2009 ~~model experiments, which is only available from a few participating models data.~~ Emissions from biomass burning and natural sources are not prescribed by the HTAP2 framework, but most of the participating models used the recommended Global Fire Emission Database version 3 (GFED3) and Model of Emissions of Gases and Aerosols from Nature (MEGAN) for biomass burning and biogenic emissions respectively. ~~To quantify the contribution from each source region with sensitivity simulations,~~ Emission perturbation is conducted with all anthropogenic emissions cut off by 20% over the source region. To examine the relative importance of long-range transport as compared to local emission change, emission perturbation is also performed for the receptor region only. ~~T~~ So this study utilizes the simulations from four scenarios: ~~(Guido R. van der Werf) (Guido R. van der Werf) (1)~~ BASE scenario with all baseline emissions; (2) EURALL scenario with all anthropogenic emissions from EUR reduced by 20%, (3) RBUALL scenario with all anthropogenic emissions from RBU reduced by 20%, and (4) EASALL scenario with all anthropogenic emissions from EAS reduced by 20%. Domain configurations

of these regions are shown in Fig.1. Note that all model experiments are conducted at global scale but the analysis of this study will focus on EUR, RBU, and EAS only.

This study takes input from 6 global models with their grid resolution, meteorology, and references listed in Table 1. These models are selected because ~~of the model level PM mass concentrations data availability, they have the simulations from the BASE, EURALL, RBUALL, and EASALL scenarios of the model level PM mass concentrations. These datasets are essential to for estimating surface PM response compare due to long range transport the aerosol transport inat different atmospheric layers.~~ Long-range transport of air pollutants may occur near the planetary boundary layer (PBL) or occur in the upper free troposphere and then descend into the PBL, ~~while air pollutants are transported from Europe to Asia (Eckhardt et al., 2003;Stohl et al., 2002), thus it is necessary to understand the contributions through PBL and free troposphere respectively.~~ Since ~~near surface aerosol plays a more important role in haze event than that in the upper air, it is necessary to understand the contributions from within- and above-PBL, vertical distribution of suspended particles, especially as relative to cloud height, plays an important role in determining radiative forcing disturbance, investigating contributions of long range transport in PBL and free troposphere will also help to examine the long range transport impact on regional climate, which has been reported in Stjern et al. (Stjern et al., 2016). These participating models have grid resolution around $1 \times 1^\circ$ and are generally sufficient to demonstrate the broad impacts from one continent to another (Dentener et al., 2010). Although these simulations are relatively coarse as compared to regional modeling, they are able to probe into the transport of upwind pollutants in both low and high altitudes (Fiore et al., 2009).~~

2.2 Model evaluation

~~Before A~~analyzing the source-receptor (S/R) relationship, ~~we applied measurements from multiple observation networks almost purely relies on model simulations, thus the performances of the models, especially over the source and receptor regions, determine the reliability of long range transport assessment. To understand the accuracy and uncertainty of the simulations, measurements from multiple observation networks are employed in this study~~ to evaluate the models performances at EUR, RBU, and EAS regions respectively. Surface observations are collected from four programs: EBAS from the Norwegian Institute for Air Research (~~NILU, <http://ebas.nilu.no>~~), Air Pollution Index (API) from the China Ministry of Environmental Protection (~~<http://datacenter.mep.gov.cn/>~~), Acid Deposition Monitoring Network in East Asia (EANET, 2007), and the AERONET (~~<http://aeronet.gsfc.nasa.gov>~~) from NASA. EBAS (Torseth et al., 2012) sites are all located in Europe so the data is used for model evaluation in EUR. API includes PM₁₀ concentrations from 86 cities over China (Dong et al., 2016), and EANET has observations of PM_{2.5}, PM₁₀, O₃, CO, SO₂, NH₃, NO₂, SO₄²⁻, NO₃⁻, and NH₄⁺ at more than 30 sites over East Asia countries (Dong and Fu, 2015a, b), so these two datasets are used for model evaluation in EAS. AERONET (~~level2.0, version2~~) has AOD measurements at more than 1,400 sites with a global coverage (Dubovik et al., 2000), ~~so the AERONET data is categorized into EUR, RBU, and EAS region first and then applied for the model evaluation at the corresponding region.~~ As some of the sites may not have valid measurements during the simulation period, only those with valid data are used and their locations are shown in Fig.1. Satellite retrieved AOD is collected from the ~~daily~~ MODIS product (~~MOD08, MYD08, <https://modis.gsfc.nasa.gov/>~~) with $0.25^\circ \times 0.25^\circ$ grid resolution to investigate the spatial distributions and column densities of aerosol simulated by the participating models.

Monthly mean surface concentrations from participating models are sampled at their own model grid cells containing the observational sites, and the corresponding measurements are also averaged on monthly scale to facilitate the evaluation. ~~As no~~ No valid data is found for surface measurements of air pollutants in the RBU region. ~~and the~~ The monthly ~~trend variations~~ trend variations of surface O₃, PM_{2.5}, and PM₁₀ are shown ~~only~~ only for EUR and EAS in Fig.2. ~~The e~~ Evaluation statistics including mean bias ~~(Simpson et al.)~~ (MB) ~~(Simpson et al.)~~ and coefficient of determination (R²) are indicated in Fig.2 for the model ensemble mean ~~only~~, calculated as the average of all participating models at ~~the coarsest grid resolution (2.8°×2.8°)~~ grid resolution. ~~Although the mass concentrations in~~ Measurements of aerosol sub-species including sulfate (SO₄²⁻), nitrate (NO₃⁻), ammonium (NH₄⁺), organic aerosols (OA) and gas-phase species such as CO, NH₃, NO₂, and SO₂ are also available at some of the EBAS and EANET stations. ~~But,~~ But the data coverage is very sparse in terms of both number of sites and sampling periods, so the evaluations of these species are not discussed here but presented in the supplementary material (Table S1). In general, all participating models successfully reproduce the seasonal cycle of O₃ in EUR and EAS. The model ensemble mean shows ~~mean bias (Simpson et al.)~~ MB of only 4.4 μg/m³ as compared to the EBAS observation in EUR. Relatively large biases (8-15 μg/m³) are indicated in warmer months ~~(from Jun. to Sep.)~~, but ~~But~~ meanwhile the standard deviation of measurement (indicated by vertical error bars in Fig.2) is even larger (10-15 μg/m³), indicating that the measured O₃ concentrations vary significantly among the EBAS sites in the same model ensemble grid ~~due to the coarse resolution~~. ~~Temporal~~ Seasonal variation of O₃ is also simulated well in EAS, ~~although the models all tend to underestimate the high peaks values in spring (Mar. Apr.) and low bottoms concentrations in summer (Jul. Sep.) with moderate overestimation throughout the year.~~

Simulations of surface PM_{2.5} concentrations are consistent among the participating models except that GEOSCHEMADJOINT suggests larger ~~temporal seasonal~~ variation than the other models. In EUR, the model ensemble mean shows MB as -4.6 μg/m³ against EBAS measurements. ~~The seasonal cycle of PM_{2.5} is less prominent as that of O₃ as indicated by the observations, but the ensemble mean and~~ generally captures the monthly changes with R² ~~as of~~ 0.7. Underestimation of surface PM_{2.5} concentration in EUR might be due to the fact that some of the measurements are affected by the local ~~sources environments~~ sources. ~~As demonstrated in Fig.2(b), PM_{2.5} are available from five 5 EBAS stations. By examining their locations, and we find that~~ one of the stations is close to highway (49.90°N, 4.63°E) ~~and shows significantly higher PM_{2.5} measurements than the others, which shall be attributed to the influences from not only the traffic emissions but also the wild fires from the forest nearby.~~ These local impacts can hardly be captured by global models due to their coarse grid resolutions. In EAS region, model ensemble mean shows a small MB as -1.6 μg/m³ but poor correlation with measurement as ~~the~~ R² is ~~only~~ 0.2. The monthly dynamics of PM_{2.5} is more prominent in EAS as that in EUR and the models tend to miss the high peaks in spring (Apr.-May). As the anthropogenic emission in Asia is developed with top-town method, the predefined ~~temporal seasonal~~ profile ~~applied during the modeling~~ have been demonstrated to affect the model's capability of reproducing the ~~seasonal temporal~~ changes of PM_{2.5} (Dong and Fu, 2015a). Simulation of PM₁₀ concentration shows good agreement between the model ensemble mean and the measurements in EUR, with MB of ~~only~~ -0.7 μg/m³. The models systematically underestimate surface PM₁₀ by -30.7 μg/m³ in EAS but successfully reproduce the seasonal cycle. This is likely due to the fact that majority of the API and EANET stations are ~~located in the urban area and thus get frequently affected by the local sources. Previous studies (Dong et al., 2015a) also suggested that the anthropogenic emission of primary PM₁₀ might be underestimated in China and subsequently lead to negative MB.~~

As no surface measurement of air pollutants is available in the RBU region, we evaluate the model simulated AOD against AERONET measurement and MODIS satellite product on monthly scale in all the three regions as shown in Fig.3. Most of the models fall into the two-fold range at both AERONET stations and MODIS grid cells. ~~The participating models tend to overestimate AOD in the EUR region as compared to the AERONET observation, as with 0.1 MB and 0.3 R² for the model ensemble mean shows MB of 0.1 and R² of 0.3.~~ In the RBU region, the model ensemble mean shows MB of only 0.05 yet the R² is only 0.2, indicating that there is a large discrepancy between model simulation and AERONET in terms of the ~~temporal-seasonal~~ changes of AOD. The model ensemble mean shows best performance in EAS among all the three regions with MB of 0.1 and R² of 0.6, suggesting that models have good agreement with AERONET observation for both the level and the seasonal cycle of AOD. The simulated AOD are generally consistent between models, except that CHASER is always 1-2 times higher than the others. The validations against MODIS product suggest slightly better ~~performance of the models~~ model performance, as the model ensemble mean shows R² values as 0.5, 0.4, and 0.6 in EUR, RBU, and EAS respectively. In contrast to the overall overestimation indicated by ~~evaluation against~~ AERONET, ~~the evaluation against~~ MODIS suggests models tend to slightly underestimate the AOD in all three regions with MB of -0.02, -0.04, and -0.03 in the EUR, RBU, and EAS regions respectively. This shall be due to the fact that AERONET has limited number of stations – there are 73, 11, and 15 stations in the EUR, RBU, and EAS regions respectively that have valid observations covering the simulation period – while MODIS has more comparable grid cells over the study domain.

The discrepancy between AERONET observations and MODIS product indicates that limited number of surface observations may not be sufficient to judge the overall performance of model since there is a high chance that observation may get affected by the local sources and subsequently biasing the assessment. ~~To achieve a more solid understanding of the model performance, s~~ Spatial distributions of the simulated AOD from all participating models and the MODIS product are compared as shown in Fig.4. The Aerosol Comparisons between Observations and Models (AEROCOM) project has conducted a thoroughly evaluation of 14 global models and suggested the simulated AOD is in a two-fold range of the observations with mean normalized bias (MNB) varied between -44% and 27% (Huneus et al., 2011). As presented in Fig.4, the model ensemble mean in this study shows good agreement with the MODIS production in terms of spatial distribution, and the MNB values are 9.3%, 18.1%, and 44.9% in the EUR, RBU, and EAS regions respectively. These evaluation statistics are consistent with ~~the evaluations by~~ AEROCOM. But we also find some exceptions as CHASER significantly overestimate the AOD in China especially over the central and east coastal areas, indicating that the simulation bias may be generated by the model's treatment of the intensive anthropogenic emission over these areas. ~~The~~ SPRINTARS is also found to significantly overestimate AOD over the Taklamakan Desert area, indicating that the bias shall be attributed to the ~~model's capability of treatment of predicting~~ wind-blown dust.

3. Result and Discussion

3.1 Seasonality of long-range transport impacts at surface layer

We start evaluating the long-range transport of PM_{2.5} from the EUR and RBU source regions to the EAS receptor region by estimating the surface PM_{2.5} concentration response on domain average scale under

the emission perturbation scenarios. PM response (ΔPM) is defined as the concentrations difference between the baseline scenario and the perturbation scenarios as:

$$\Delta PM_{EURALL} = PM_{BASE} - PM_{EURALL}$$

$$\Delta PM_{RBUALL} = PM_{BASE} - PM_{RBUALL}$$

5 To also understand the responses of aerosol sub-species, simulations of SO_4^{2-} , NO_3^- , NH_4^+ , OA, and black carbon (BC) are collected from each of the participating models if it is available. Dust and sea salt are not analyzed in this study because emission perturbations are performed for anthropogenic sectors only. So in this study we assume that $\Delta PM_{2.5} = \Delta SO_4^{2-} + \Delta OA + \Delta BC + \Delta NO_3^- + \Delta NH_4^+$. For those models reporting organic carbon (OC) instead of OA, an OC-to-OA conversion factor as 1.8 is applied to estimate
10 OA following the method discussed in Stjern et al. (2016). For those models reporting only some of the sub-species and total $PM_{2.5}$, an extra species “other” is defined as subtracting the available sub-species from $PM_{2.5}$. For example, GEOS5 and SPRINTARS report mass concentrations of SO_4^{2-} , OA, BC, and $PM_{2.5}$, then for these two models we use: Other = $PM_{2.5} - (SO_4^{2-} + OA + BC)$. Note that the CAM-chem model reports sub-species for all scenarios but NO_3^- for BASE scenario only, so no $\Delta Other$ is estimated for this
15 model.

Long-range transport impacts from the EUR region are presented in Fig.5. Large variations of the simulated $PM_{2.5}$ responses are found among the models. The largest estimation of $\Delta PM_{2.5}$ is $0.16 \mu g/m^3$ estimated by GEOS5 in March, and the smallest $\Delta PM_{2.5}$ is $0.01 \mu g/m^3$ estimated by EMEP in July. Regarding the seasonal cycle, majority of the models suggest the long-range transport has higher impact in winter and spring and lower impact in summer, well consistent with the O_3 long-range transport seasonality reported by the HTAP1 assessment (Dentener et al., 2010; Streets et al., 2010). In contrast to other models that show most significant responses in winter or spring, CAM-chem does suggest higher values of $\Delta SO_4^{2-} + \Delta OA + \Delta BC + \Delta NH_4^+$ in July, but the seasonal cycle of $\Delta PM_{2.5}$ is unknown because ΔNO_3^- is not available. The prominent difference in seasonality may attributed to the model diversity in terms of meteorology, aerosol mechanisms, and convection scheme. CAM-chem simulated surface air temperature is $\sim 2K$ higher than other models in EUR region. (Im et al., 2018) suggested wind speed and PBL height may play a more important role in resulting model diversities of aerosol burden, but unfortunately only one of the participating model (SPRINTARS) provides the PBL data. Stjern et al. (2016) suggested that the differences of aerosol schemes and treatment of OC, OA, and SOA lead to additional inter-model variability. Additional specifically designed model experiment is necessary to explicitly identify the causes of inter-model variability. For most of the participating models, ΔSO_4^{2-} and/or ΔOA make larger contributions to $\Delta PM_{2.5}$ and show more prominent monthly changes than other sub-species. CAM-chem and GEOSCHEMADJOINT simulated ΔSO_4^{2-} shows monthly variations with a factor of 5, and GEOS5 suggests the monthly dynamics of ΔOA is with a factor of 8. The model ensemble mean suggests that the largest long-range transport impact of $\Delta PM_{2.5}$ is $0.064 \mu g/m^3$ in March and the smallest impact is $0.035 \mu g/m^3$ in September, and the contributions from ΔBC , ΔSO_4^{2-} , ΔOA , ΔNO_3^- , and ΔNH_4^+ are 3%, 45%, 19%, 17%, and 16% respectively.
20
25
30
35

Long-range transport from the RBU to the EAS region is presented in Fig.6. The highest $\Delta PM_{2.5}$ is estimated by GEOS5 as $0.19 \mu g/m^3$ in March, while—and the lowest ΔPM is indicated by GEOSCHEMADJOINT as $0.018 \mu g/m^3$ in July. Similar to the response under EURALL scenario, long-
40

range transport from the RBU region is also ~~substantially~~ mainly contributed by ΔSO_4^{2-} , but ΔNO_3^- and ΔNH_4^+ share more significant portions in $\Delta\text{PM}_{2.5}$. Most of the models suggest relatively lower values of ΔOA except for GEOS5, which suggests up to $0.1\mu\text{g}/\text{m}^3$ ΔOA in March. The model ensemble mean suggests maxima of $\Delta\text{PM}_{2.5}$ as $0.101\mu\text{g}/\text{m}^3$ in March and the minima as $0.065\mu\text{g}/\text{m}^3$ in August, and the contributions from ΔBC , ΔSO_4^{2-} , ΔOA , ΔNO_3^- , and ΔNH_4^+ are 2%, 43%, 14%, 20%, and 21% respectively. Percentage contributions are generally less than 3%, yet the highest contributions could be up to 3-4% for ΔSO_4^{2-} , ΔNO_3^- , and ΔNH_4^+ as suggested by EMEP. The relatively lower contribution of ΔOA and higher contributions of ΔNO_3^- and ΔNH_4^+ ~~under the RBUALL scenario~~ is probably due to the low temperature in the RBU source region, which may extend the lifetime of gas-phase precursors (SO_2 , NO_x , and NH_3) and enhance the export of secondary inorganic aerosols produced during the journey of long-range transport. ~~In fact, the low~~ Low temperature also favors SOA production from VOC due to the partitioning to the condensed phase. CAM-chem suggests the contribution of ΔSOA in ΔOA is 32% under the RBUALL scenario and 28% under the EURALL scenario, and model ensemble mean also shows that more OA is transported from RBU ($0.01\mu\text{g}/\text{m}^3$) than that from EUR ($0.008\mu\text{g}/\text{m}^3$), although the anthropogenic NMVOC and OC emission from EUR is 10% and 70% higher respectively. But the low temperature seems affect the SO_2 , NO_x , and NH_3 more by influencing the chemical kinetics and slow down the production of PM at the source region, which may allow more uplift motion of the gas-phase precursors, and finally result in more ΔSO_4^{2-} , ΔNO_3^- , and ΔNH_4^+ produced during the long-range transport pathway. More research effort is necessary to explicitly understand the export of precursors and secondary inorganic aerosols traveling from high latitude areas.

3.2 Long-range transport above and within the PBL

The HTAP phase 1 (HTAP1) report (Streets et al., 2010) suggests that long-range transport of air pollutants from Europe to Asia are identified at two major different heights: within and above 3km respectively, and the upper path is believed to be more important due to the existence of the Westerlies, especially when the emission source area is close to the jet stream. ~~While the modeling effort~~ (Eckhardt et al., 2003;Stohl et al., 2002) ~~referenced by the HTAP1 report is mainly investigating the influence of North Atlantic Oscillation (NAO) on air pollutants transport towards the Arctic,~~ The Europe to Asia transport pathways are identified based on spatial distributions of simulated CO column density, and the contributions from upper and lower levels transport remain unknown. The transport pathways above and within 3km are commonly used by previous studies in order to distinguish the long-range transport above and within the free troposphere, but 3km was apparently a rough estimation of the PBL height. Although PBL transport plays a dominant role in air pollutants dispersion at local scale, ~~The intensity of long-range transport exclusively within the PBL is believed to be negligible because it is frequently affected by the land surface, turbulence, and exchange with the free troposphere. The transport from Europe to Asia estimated with model experiment in this study however, may also exist~~ show some significances within the PBL since the emission perturbation is performed on continental scale, and there is a large portion of remote areas with flat topography in the Asia-Stan region laying between Europe and East Asia. ~~As very limited modeling efforts have been devoted to investigate the transport within the PBL, we compare the amount of PM responses within and above the PBL in this study to examine the contributions of long-range transport in different atmosphere layers.~~ Annual average PBL height is about 1.5km (880hPa-850hPa) above surface ground over our study domain on annual average scale, and instead of assuming a constant PBL height, we

use the monthly PBL data from the SPRINTARS model because it is the only one that uploads. To enable the comparison of PM transported within and above the PBL, we use the column density instead of mass concentration, defined as below:

$$\Delta PM_{within} = \sum_{layer=surface\ layer}^{PBL} \Delta PMC_{layer} \times HT_{layer}$$

$$\Delta PM_{above} = \sum_{layer=PBL+1}^{model\ top} \Delta PMC_{layer} \times HT_{layer}$$

where ΔPM_{above} (ΔPM_{within}) is the ΔPM transported above (within) the PBL, ΔPMC is the mass concentrations responses under the perturbation scenarios at each layer, and HT is the model layer thickness. Fig.7 presents the spatial distributions of model simulated ΔPM_{within} and ΔPM_{above} under the EURALL scenario, as well as the longitude-pressure cross sections of ΔPMC estimated by the participating models. It is important to note that PM mentioned in this section refers to the lump sum of SO_4^{2-} , OA, and BC (because these are the sub-species available from all participating models) to enable the inter-model comparison. ~~between the models.~~

Transport from the EUR to the EAS region shows generally consistent spatial distributions ~~by~~ between all participating models. Long-range transport of PM above the PBL is mainly distributed along 40°N and higher latitude, where the impact can reach even further towards the west Pacific. The lower latitude (30°N-40°N) transport of PM is blocked by the Pamirs, Tianshan, and Altay Mountains due to the elevated topography along the western boundary of China. ~~The~~ long-range transport of PM within PBL is mostly blocked shortly after exported from Europe at the eastern side of Black Sea along Iran, Georgia, and Armenia, while the rest of it travels along 45°N and above latitudes towards East Asia. All participating models suggest that PM is firstly carried from EUR towards northeast direction over Siberia, Mongolia and Northeast part of China, and then down to lower latitude areas over North China Plain (NCP). This transport pathway is well consistent with ~~the findings by~~ the HTAP1 assessment (~~Dentener et al., 2010~~ Streets et al., 2010). ΔPM_{above} is found substantially higher than ΔPM_{within} over the EAS receptor region. Large values of ΔPM_{above} suggest that the long-range transport may also play an important role in affecting the shortwave radiative forcing budget since the aerosol may suspend above the cloud. Deposition of PM from upper air down to the surface layer may also subsequently affect to the near surface layer air quality. Most models show gradually decreased ΔPM_{above} and ΔPM_{within} from EUR to EAS, but SPRINTARS shows non-negligible PM changes along the southeast coast of China, which could be due to the production of secondary SO_4^{2-} converted from long-range transport SO_2 , discussed earlier in section 3.1. The largest long-range transport impact is estimated by CHASER and smallest impact is estimated by EMEP, but no significant model diversities are found ~~among the models regarding the intensity of ΔPM_{above} and ΔPM_{within} .~~ The longitude-pressure cross sections of the PM responses present a clear depict of the long-range transport from EUR to EAS at different height. The PM responses along the longitude can reach up to higher than 500hPa over the EUR region (10°E-40°E), indicating a significant uplift motion of the air pollutants over Europe. Majority of the eastward transport PM is blocked at 45°E-50°E due to the elevated topography. In the upper layer above 800hPa however, PM is slightly less affected by the topography and can transport further towards the EAS region, where it deposits to near surface layer subsequently. Both the spatial distributions of ΔPM_{within} and the cross sections of ΔPMC suggested that the inter-continental transport of

aerosol does occur within PBL, although the intensity is less significant as compared to that above PBL. Under the ERUALL scenario, $\Delta\text{PM}_{\text{within}}$ contribution to the total column density of ΔPM is 34% estimated by the model ensemble mean, with the lowest contribution estimated by EMEP as 22% and highest contribution estimated by GEOSCHEMADJOINT as 38%.

5

Long-range transport from RBU follows the similar pathway as that from EUR to EAS, as shown in Fig.8, which is likely because most of the RBU anthropogenic emissions are located at the European part of Russia and Ukraine. PM responses are also relatively more significant in the upper air above PBL, which spread along 45°N and higher latitude and affect the north part of China, North Korea, South Korea, and Japan. Long-range transport from RBU is slightly larger than that from EUR for both above and within the PBL. Spatial distributions of $\Delta\text{PM}_{\text{above}}$ and $\Delta\text{PM}_{\text{within}}$ suggest that RBU exported air pollutants can travel further towards the west Pacific. Cross sections of PM concentrations suggest that RBU emitted PM shows a much lower plume rise height in the source region as compared to that over EUR. PM response under the RBUALL scenario is also found to exist at up to 500hPa in the source region, but majority of plume is within 800hPa.

15

3.3 ~~Trend Change~~ and inter-annual variability of the long-range transport

The global anthropogenic emissions have changed significantly especially over East Asia during the past decade (~~Li et al.,~~) (Li et al., 2017), thus the long-range transport impact and its relative importance may have also changed as well. In this section, we compare the impact estimated ~~in this study~~ for the year 2010 with the assessment reported by HTAP1 for the year 2000 ~~to reveal the trend change of inter-continental transport~~. We also analyze the HTAP2 simulations for the year 2008 and 2009 to probe into the inter-annual variability ~~of the long-range transport~~. To properly interpret the HTAP1 report and the HTAP2 modeling results, it is important to realize that the regions definitions are moderately different between the two ~~modeling~~ experiments. HTAP1 used straight latitude and longitude boundaries to define the domain coverage of each region (Fiore et al., 2009), while HTAP2 applies national boundaries (one exception in the Northern Hemisphere is the Arctic region, defined as being North of 66°N latitude), thus the spatial coverage of “EU” (25°N-65°N; 10°W-50°E) defined by HTAP1 is slightly different from “EUR” defined by HTAP2, although both of them represent the European region. A similar discrepancy exist for the definition of East Asia between the two experiments, as the HTAP1 defined “EA” (15°N-50°N; 95°E-160°E) is smaller than the EAS region with less coverage on the west and north side of China. Consequently, when referring to “long-range transport from Europe to East Asia”, neither the source (Europe) nor the receptor ~~region~~ (East Asia) region share exactly the same meaning between HTAP1 and HTAP2. ~~These different region definitions will determine how to interpret the modeling results as will be discussed later in this section.~~ In addition, emission perturbations in source regions performed in both HTAP1 and HTAP2 experiments are 20% instead of 100%, thus the full contributions from the EUR or RBU to the EAS region remain unknown. Although the PM response is not exactly proportional to emission perturbation, previous studies (Leibensperger et al., 2011; Liu et al., 2008) suggested that it is reasonable to linearly extrapolate it when evaluating the inter-continental source-receptor relationship because the non-linear relationship between precursor emission changes and PM responses is only effective locally. The HTAP1 assessment reported that surface SO_4^{2-} concentrations is reduced by 12%-14% from 20% local emission reduction in

20

25

30

35

40

East Asia, Europe, and North America, corresponding to 60%-70% reduction under 100% local emission reduction if the responses are extrapolated linearly. Yet model experiments show that the real 100% emission perturbation simulations suggest 80-82% surface SO_4^{2-} concentrations reduction due to “oxidant limitation” over these polluted areas. However, this relationship becomes linear during trans-oceanic transport due to the relatively short lifetime of precursors as compared to the travel duration. So in this study, we use the *Full_Impact* to represent the PM responses from 100% emission perturbation at EUR and RBU by scaling the PM responses under the 20% emission perturbation conditions by a factor of 5, which provide a rough but direct estimation of the full contributions of long-range transport. This method has been applied by the HTAP1 related studies to estimate the long-range transport of O_3 (Fiore et al., 2009; West et al., 2009; Zhang et al., 2009). ~~which provide a rough but direct estimation of the full contributions of long-range transport:~~

$$Full_Impact_{EUR} = 5 \times \Delta PM_{EUR}$$

$$Full_Impact_{RBU} = 5 \times \Delta PM_{RBU}$$

and:

$$Full_Impact_{EUR\%} = \frac{Full_Impact_{EUR}}{PM_{BASE}} \times 100\%$$

$$Full_Impact_{RBU\%} = \frac{Full_Impact_{RBU}}{PM_{BASE}} \times 100\%$$

In addition, we also defined the *Relative_Impact* in this study to represent the relative importance of long-range transport in contrast to the local emission, as the ratio of PM responses under 20% emission perturbation in source region (i.e. EUR, RBU) to the PM responses under 20% emission perturbation in the receptor region (i.e. EAS):

$$Relative_Impact_{EUR\%} = \frac{\Delta PM_{EUR}}{\Delta PM_{EAS}} \times 100\%$$

$$Relative_Impact_{RBU\%} = \frac{\Delta PM_{RBU}}{\Delta PM_{EAS}} \times 100\%$$

Full impact and relative impact are calculated with model ensemble mean to represent the averages, and with individual modeling results to estimate the minima and maxima, as summarized in Table 2. The HTAP1 experiment only reported the assessment of SO_4^{2-} , BC and OA, so this section will focus on the analysis and comparison of these species. As mentioned earlier, the EAS region is different from the EA region defined in HTAP1, so we also calculate the full impact and relative impact for the EA region but with HTAP2 modeling data to enable the comparison ~~between the assessments from the two experiments.~~ We first compare the 2000 EU impact on EA with the 2010 EUR impact on EA. The long-range transport shows prominent decreasing trend-change for all investigated species ~~as shown in Table 2.~~ The full impact of Europe long-range transport on surface SO_4^{2-} concentration decreased from $0.15\mu\text{g}/\text{m}^3$ (5.0%) in 2000 to $0.02\mu\text{g}/\text{m}^3$ (0.5%) in 2010, which shall be due to the significant reduction of SO_2 anthropogenic emission in Europe from 9.95Tg in 2000 to 6.18Tg in 2010 (anthropogenic emissions are summarized in Table S2). The full impacts of Europe long-range transport on surface BC and OA ~~are also found to also~~ decreased by a factor of 2-5 for both absolute concentrations and percentage contributions during the 10 years period. Anthropogenic emissions of BC, OC, NMVOC, and primary PM in Europe are decreased by 21%, 4%,

37%, and 2% respectively and their emissions in East Asia are increased by 39%, 21%, 38%, and 32% respectively from 2000 to 2010. The emission increase in East Asia shall be response for the enhanced surface PM concentrations simulated under the baseline scenario, and the emission reductions in EUR are consistent with the decreasing trend-change of the long-range transport contributions estimated by the models.

We then investigate the inter-annual variability of the long-range transport by examining the EUR to EAS and the RBU to EAS impact from 2008 to 2010. The model estimated $Full\ Impact_{EUR\%}$ shows annual moderate changes by 15%-30% for all species from year to year, with no significant trend change is found. The $Full\ Impact_{RBU\%}$ shows relatively larger inter-annual changes. As the anthropogenic emissions from the RBU region has steadily decreased by ~9% from 2008 to 2010, the large dynamics of $Full\ Impact_{RBU\%}$ is more likely due to the fact that only one model (CAM-chem) is available to estimate the RBU impact in 2008 and 2009 and thus the assessment may be biased. While the estimation for 2010 is calculated with multi-model ensemble mean, the estimations for the other two years are determined by one model CAM-chem only and need to be further validated.

We finally analyze the relative importance of long-range transport. The HTAP1 reported that the overall contribution to SO_4^{2-} and OA from EU to EA is 2.9% in 2000, and in this study the estimated relative impact in 2010 is 2.2%, indicating that long-range transport is playing a less important role as compared to the local anthropogenic emission in terms of affecting the surface air quality in East Asia. In contrast, 20% anthropogenic emission reductions in the EAS region lead to surface concentration of $SO_4^{2-}+OA$ decreased by 16.8% in 2000 and 14.1% in 2010, suggesting that the non-linear relationship between precursor and PM becomes more significant when the anthropogenic emissions increase. It also indicates that to achieve a better air quality with lower PM concentrations, more efforts shall be devoted to reduce the emissions in 2010 because the top 20% emission reduction would lead to less PM response as compared to that in 2000.

3.4 Long-range transport impact during the haze episode

As the annual average full impact for aerosol sub-species are presented in last section, in this section we evaluate the full impact during the haze episodes for $PM_{2.5}$. We first use the National Climate Data Center (NCDC) observations to identify the locations and periods of haze in China, and then analyze the long-range transport impacts during these identified haze episodes. Haze is can be quantitatively defined identified with a visibility less than 10km and relative humidity less than 90% (Fu et al., 2014). As most of the haze (locations of NCDC sites and full map of haze shown in Fig.S1) are located over central and eastern part of China (CEC), in this section we focus the analysis of long-range transport impacts on the CEC subdomain (20°N-55°N; 100°E-135°E). The full impacts during the haze episodes (HAZE) are estimated and compared with the annual averaged full impacts throughout the year of 2010 (AAVG), as shown in Table 3.

CAM-chem and GEOS5 has no daily surface data available so data from the rest 4 participating models are analyzed in this section. The models suggest that the $PM_{2.5}$ baseline concentrations during haze episodes are substantially higher than the annual averages as shown in Table 3, with the largest difference between HAZE and AAVG is estimated by CHAER as $27.27\mu g/m^3$ and the smallest difference estimated by GEOSCHEMADJOINT as $2.56\mu g/m^3$. The full impacts of long-range transport from the source regions are also higher during the haze episodes by a factor of 2-3 than the annual averages. As estimated by the model ensemble mean, on annual average scale the $Full\ Impact_{EUR}$ is $0.35\mu g/m^3$, contributing to 1.7% of

the surface $PM_{2.5}$ in the EAS region. During the haze episode however, $Full_Impact_{EUR}$ is $0.99\mu g/m^3$, contributing to 3.1% of the surface $PM_{2.5}$ in the ECE CEC region. The impact from the RBU region is also found more significant during haze episodes, as the $Full_Impact_{RBU}$ increased from $0.53\mu g/m^3$ (2.6%) to $1.32\mu g/m^3$ (4.1%). Higher values of $Full_Impact_{EUR}$ and $Full_Impact_{RBU}$ suggest that more fine particles are transported from the EUR and RBU source regions when China is suffering from haze.

As shown in Fig.9. The spatial distributions of the long-range transport full impacts during the haze episodes demonstrate a very similar pattern among the participating models. The $Full_Impact_{EUR\%}$ is most significant over the northeast corner of China, and gradually decreases towards the southeast direction. The intensity of $Full_Impact_{EUR\%}$ estimated by models however, show large difference as the maxima estimated by SPRINTARS is 10.5% and the minima estimated by EMEP is 0.4%. ~~The spatial distributions of $Full_Impact_{RBU\%}$ shown by the models are similar to that of $Full_Impact_{RBU\%}$, but the intensity of long-range transport from the RBU region is generally larger.~~ The numbers presented in Table 3 have demonstrated the general full impacts during all haze episodes, but we are still unaware of how those individual haze episodes are affected by the long-range transport. So we also summarize the histograms of daily full impacts during the haze episodes. The frequency of the histogram is calculated as:

$$Frequency_{Full_Impact=i\%} = \frac{\#HazeEvent_{i\%}}{\sum_{i=1}^{MaxFI=15} \#HazeEvent_{i\%}} \times 100\%$$

and it satisfies:

$$\sum_{i=1}^{MaxFI=15} Frequency_{Full_Impact=i\%} = 100\%$$

We define $MaxFI = 15$ to represent the upper boundary as $Full_Impact \geq 15\%$. This value (i.e. 15%) contribution is selected in order to compare the full impact from long-range transport against the $PM_{2.5}$ response under 20% local emission control in the EAS region. As shown in Table 2, surface concentration of $SO_4^{2-}+OA$ is reduced by ~15% under the EASALL scenario. So if $Full_Impact_{EUR} \geq 15\%$, it indicates that the long-range transport from EUR may have an equivalent or even more significant contribution to the surface $PM_{2.5}$ as that produced from 20% of the local anthropogenic emission. We define $\#HazeEvent_{i\%}$ as the number of haze events that satisfies: $(i - 1)\% < Full_Impact \leq i\%$ and is calculated as:

$$HazeEvent_{i\%} = \sum_{d=1}^{365} H_{d,r,c}$$

$H_{d,r,c}$ is the haze event at day d , row r , and column c , defined as:

$$H_{d,r,c} = \begin{cases} 1, & \text{if } RH_{d,r,c} < 90\% \text{ and } visibility_{d,r,c} < 10km, \text{ and } i\% < Full_Impact_{d,r,c} \leq (i + 1)\% \\ 0, & \text{otherwise} \end{cases}$$

So with $Frequency_{Full_Impact=i\%}$, we can estimate the percentage of the haze episodes during which the long-range transport contributes to $i\%$ of the surface $PM_{2.5}$. The values of $Frequency_{Full_Impact=15\%}$ are indicated in the histogram plots as shown in Fig.9. The SPRINTARS estimated $Frequency_{Full_Impact=15\%}$ is 5.5%, suggesting that during almost 5.5% of the haze episodes in China, long-range transport from Europe contributed to at least the equivalent amount of surface $PM_{2.5}$ concentration as that generated from 20% of local anthropogenic emission, while the other models' estimations range from 0.01% to 1.9%. ~~The~~

participating models suggest that $Frequency_{Full_Impact=15\%}$ ranges from 0.1% to 5.5% with the model ensemble mean estimates as 1.8%. The influence from the RBU region shows slightly higher value of $Frequency_{Full_Impact=15\%}$ as ~~the model ensemble mean estimates as~~ 2.2%. Although significant variations are found among the model estimations, all participating models suggest that non-negligible values of $Frequency_{Full_Impact=15\%}$ ~~and $Frequency_{Full_Impact=15\%}$~~ , indicating the important contributions of long-range transport to haze episodes in China.

~~Although~~ The high surface $PM_{2.5}$ is believed to be the most direct reason for causing haze condition. ~~But~~, visibility cannot be represented by $PM_{2.5}$ mass concentration only since it is also determined by the optical properties, number concentrations, and size distributions of the aerosols. Thus the analysis of PM concentration response depicts only partially of the impact of long-range transport during haze episodes. Calculating model predicted visibility requires the detailed aerosol information mentioned above which is not available from any of the participating models. So we use the Koschmieder equation (Han et al., 2013) to estimate the model simulated visibility from aerosol extinction coefficient (β) as:

$$Visibility = \frac{3.912}{\beta}$$

Modeled visibility is calculated for SPRINTARS only since the other participating models has no surface layer extinction coefficient available. The long-range transport impact on visibility change and number of haze days change are shown in Fig.10. It shall be noticed that SPRINTARS estimated long-range transport impact of surface $PM_{2.5}$ is the highest among the participating models, thus the analysis of visibility change shown in Fig.10 ~~shall may~~ represent the upper boundary of model estimations. The spatial distribution of visibility changes agree well with ~~the that distribution~~ of surface $PM_{2.5}$ responses. Visibility is reduced by up to 10km along the northeast boundary of China, which is likely due to the fact that these areas receive the most significant amount of the long-range transport aerosols from the EUR and RBU regions. The number of haze days changes however, are mostly ~~distributed prominent~~ in the NCP and along the east coast of China. The long-range transport results in 1-3 days (~~<3%~~) of extra haze over these areas throughout the year. The total number of haze events ($\sum_{i=1}^{MaxFI=15} \#HazeEvent_{i\%}$) estimated by the SPRINTARS model is 18566, 18538, and 18546 under the BASE, EURALL, and RBUALL scenarios, suggesting that that transport from the EUR and RBU region contribute to an extra of 0.15% and 0.11% haze events respectively.

4. Summary and conclusions

To estimate the long-range transport contributions to the surface aerosol concentrations in East Asia, this ~~analysis study~~ uses 6 global models participating in the HTAP2 experiment. Simulations for the year 2010 from baseline scenario and 20% anthropogenic emission perturbation scenarios are explored to estimate the long-range transport from the Europe and Russia/Belarussia/Ukraine source regions respectively. We find that on annual average scale, long-range transport from Europe contributes 0.04-0.06 $\mu\text{g}/\text{m}^3$ (0.2-0.8%) to the surface $PM_{2.5}$ concentration in East Asia as indicated by the 20% emission perturbation experiment, with majority of the transported aerosols are SO_4^{2-} and OA at 43% and 19% respectively. Long-range transport from Russia/Belarussia/Ukraine shows slightly higher impact with contributions of 0.07-0.10 $\mu\text{g}/\text{m}^3$ (0.3-0.9%) to the surface $PM_{2.5}$ in East Asia, within which the NO_3^- and NH_4^+ responses share bigger slices as 20% and 21% respectively, larger than that of OA as 14%. As

5 compared to the impact from Europe to East Asia, more secondary inorganic aerosols are transported from the Russia/Belarus/Ukraine region despite the fact that the 2010 anthropogenic emission from RBU is 40-50% lower than that from EUR for SO₂, NO_x, and NH₃. Our analysis suggests that the lower temperature in RBU may result in extended lifetime of the gas-phase precursors, which are gradually converted to secondary inorganic aerosols during the transport pathway to East Asia, yet further modeling experiment is necessary to explicitly explore the temperature impact on long-range transport.

10 By investigating the PM responses in different atmosphere layers, we find that long-range transport exist both within and above the PBL, although the upper level transport takes a larger portion as 66% of the total PM column density response in East Asia. Spatial distributions of the PM responses suggest that the long-range transport from Europe and Russia/Belarus/Ukraine are both predominantly blocked at western side of China due to the elevated topography of Pamirs, Tianshan, and Altay Mountains, where the rest of the exported pollutants are carried by the Westerlies along 45°N and higher latitude towards China, North Korea, South Korea, Japan, and the west Pacific.

15 Comparison between the HTAP1 assessment and the estimation from this study reveals the 10 years decreasing ~~trend-change~~ of long-range transport from Europe to East Asia. When extrapolating the impact of 20% anthropogenic emission perturbation by a factor of 5 to estimate the full impact, contributions to surface concentrations are decreased from 5.0%, 1.0%, and 0.4% in 2000 to 0.5%, 0.2%, and 0.2% in 2010 for SO₄²⁻, BC, and OA respectively. This comparison may contain uncertainty because of the different model ensemble compositions between HTAP1 and this study, but the ~~trend-change~~ of the long-range transport impacts from 2000 to 2010 found in this study was consistent with the implications from the emissions changes. The simultaneously emission reduction in Europe and emission enhancement in East Asia shall be responsible for the decreasing ~~trendchange~~. The surface concentrations of SO₄²⁻, BC, and OA in East Asia are also increased by 14%, 50%, and 140% from 2000 to 2010, well consistent with many of the local measurements reported in recent years (Chen et al., 2016;Feng et al., 2014;Lu et al., 2010;Zhu et al., 2012). It is important to emphasize that based on the model ensemble mean estimations, despite the fact that baseline of 2010 anthropogenic emission is substantially higher (20-40%) than that in 2000, a same percentage reduction of the local anthropogenic emission will lead to less benefit in terms of reducing the ambient PM concentrations in the 2010 scenario, indicating the increasingly more difficulties for air quality management in East Asia.

30 The long-range transport impact during haze episodes in China are estimated by using the NCDC surface observations to identify the haze events, on top of which the HTAP2 experiments are analyzed to quantify the changes of surface PM_{2.5}, visibility, and number of haze days. Despite the significant discrepancy between the models, all participants demonstrate that the full impacts during haze episodes are more significant than that on annual average scale. Estimations with the model ensemble mean suggest that the full impacts from EUR and RBU are 0.99μg/m³ (3.1%) and 1.32μg/m³ (4.1%) respectively during haze episodes, significantly higher than the annual averages. The model ensemble also suggest that during 5.5-5.7% of the haze episodes, long-range transport can contribute to surface PM_{2.5} as much as that generated from 20% of local anthropogenic emission. Based on analysis with the SPRINTARS model output, visibility is reduced by up to 10km with the largest impact found along northeast China, and the impact gradually decreases towards southeast and causes less than 500m visibility reduction. The enhancement of number of haze days however, is found mainly located at the North China Plain and southeast coast area of China, where most of the places receive extra 1-3 haze days due to the influence of long-range transport. We find that throughout the full year of 2010, number of haze event in our studying domain is increased by

0.15% and 0.11% due to the long-range transport from the Europe and Russia/Belarus/Ukraine region respectively.

5. Acknowledgements:

5 This work was partly supported by Natural Science Foundation of China (41429501). We would like to thank the UN-ECE CLRTAP (EMEP), AMAP, and NILU for supporting the EBAS database with air pollutants measurements. We thank Dr. Keiichi Sato and Dr. Ayako Aoyagi from Asia Center for Air Pollution Research for providing the EANET data. We would also like to acknowledge NOAA NCDC to provide the public accessible meteorology observations. We thank the Oak Ridge Leadership Computing Facility (OLCF) at Oak Ridge National Lab (ORNL) for providing computer sources.

10

References

- Akimoto, H.: Global air quality and pollution, *Science*, 302, 1716-1719, 2003.
- 5 Cao, Q. L., Liang, Y., and Niu, X. T.: China's Air Quality and Respiratory Disease Mortality Based on the Spatial Panel Model, *Int J Env Res Pub He*, 14, [Artn-1081](#)
10.3390/ijerph14091081, 2017.
- Carmichael, G. R., Sakurai, T., Streets, D., Hozumi, Y., Ueda, H., Park, S. U., Fung, C., Han, Z., Kajino, M., Engardt, M., Bennet, C., Hayami, H., Sartelet, K., Holloway, T., Wang, Z., Kannari, A., Fu, J., Matsuda, K., Thongbooncho, N., and Amann, M.: MICS-Asia II: The model intercomparison study for Asia Phase II
10 methodology and overview of findings, *Atmos Environ*, 42, 3468-3490, 2008.
- Chen, Y., Schleicher, N., Fricker, M., Cen, K., Liu, X. L., Kaminski, U., Yu, Y., Wu, X. F., and Norra, S.: Long-term variation of black carbon and PM_{2.5} in Beijing, China with respect to meteorological conditions and governmental measures, *Environ Pollut*, 212, 269-278, [10.1016/j.envpol.2016.01.008](#), 2016.
- 15 Dong, X. Y., and Fu, J. S.: Understanding interannual variations of biomass burning from Peninsular Southeast Asia, part I: Model evaluation and analysis of systematic bias, *Atmos Environ*, 116, 293-307, 2015a.
- Dong, X. Y., and Fu, J. S.: Understanding interannual variations of biomass burning from Peninsular Southeast Asia, part II: Variability and different influences in lower and higher atmosphere levels, *Atmos Environ*, 115, 9-18, 2015b.
- 20 Dong, X. Y., Fu, J. S., Huang, K., Tong, D., and Zhuang, G. S.: Model development of dust emission and heterogeneous chemistry within the Community Multiscale Air Quality modeling system and its application over East Asia, *Atmos Chem Phys*, 16, 8157-8180, 2016.
- Dubovik, O., Smirnov, A., Holben, B. N., King, M. D., Kaufman, Y. J., Eck, T. F., and Slutsker, I.: Accuracy assessments of aerosol optical properties retrieved from Aerosol Robotic Network (AERONET) Sun and sky radiance measurements, *J Geophys Res-Atmos*, 105, 9791-9806, 2000.
- 25 Eckhardt, S., Stohl, A., Beirle, S., Spichtinger, N., James, P., Forster, C., Junker, C., Wagner, T., Platt, U., and Jennings, S. G.: The North Atlantic Oscillation controls air pollution transport to the Arctic, *Atmos Chem Phys*, 3, 1769-1778, 2003.
- Feng, J. L., Zhong, M., Xu, B. H., Du, Y., Wu, M. H., Wang, H. L., and Chen, C. H.: Concentrations, seasonal and diurnal variations of black carbon in PM_{2.5} in Shanghai, China, *Atmos Res*, 147, 1-9, [10.1016/j.atmosres.2014.04.018](#), 2014.
- 30 Fiore, A. M., Dentener, F. J., Wild, O., Cuvelier, C., Schultz, M. G., Hess, P., Textor, C., Schulz, M., Doherty, R. M., Horowitz, L. W., MacKenzie, I. A., Sanderson, M. G., Shindell, D. T., Stevenson, D. S., Szopa, S., Van Dingenen, R., Zeng, G., Atherton, C., Bergmann, D., Bey, I., Carmichael, G., Collins, W. J., Duncan, B. N., Faluvegi, G., Folberth, G., Gauss, M., Gong, S., Hauglustaine, D., Holloway, T., Isaksen, I. S. A., Jacob, D. J., Jonson, J. E., Kaminski, J. W., Keating, T. J., Lupu, A., Marmer, E., Montanaro, V., Park, R. J., Pitari, G., Pringle, K. J., Pyle, J. A., Schroeder, S., Vivanco, M. G., Wind, P., Wojcik, G., Wu, S., and Zuber, A.: Multimodel estimates of intercontinental source-receptor relationships for ozone pollution, *J Geophys Res-Atmos*, 114, [Artn-D04301](#)
40 [10.1029/2008jd010816](#), 2009.
- Fu, G. Q., Xu, W. Y., Yang, R. F., Li, J. B., and Zhao, C. S.: The distribution and trends of fog and haze in the North China Plain over the past 30 years, *Atmos Chem Phys*, 14, 11949-11958, 2014.
- Fu, J. S., Dong, X. Y., Gao, Y., Wong, D. C., and Lam, Y. F.: Sensitivity and linearity analysis of ozone in East Asia: The effects of domestic emission and intercontinental transport, *J Air Waste Manage*, 62, [10.1080/10962247.2012.699014](#), 2012.
- 45

- Gao, J. H., Woodward, A., Vardoulakis, S., Kovats, S., Wilkinson, P., Li, L. P., Xu, L., Li, J., Yang, J., Li, J., Cao, L., Liu, X. B., Wu, H. X., and Liu, Q. Y.: Haze, public health and mitigation measures in China: A review of the current evidence for further policy response, *Sci Total Environ*, 578, 148-157, 10.1016/j.scitotenv.2016.10.231, 2017.
- 5 Guido R. van der Werf, J. T. R., Louis Giglio, Thijs T. van Leeuwen, Yang Chen, Brendan M. Rogers, Mingquan Mu, Margreet J. E. van Marle¹, Douglas C. Morton, G. James Collatz, Robert J. Yokelson, and Prasad S. Kasibhatla: Global fire emissions estimates during 1997–2015, *Earth Syst. Sci. Data Discuss.*, <https://doi.org/10.5194/essd-2016-62>, 2017.
- 10 Guo, S., Hu, M., Zamora, M. L., Peng, J. F., Shang, D. J., Zheng, J., Du, Z. F., Wu, Z., Shao, M., Zeng, L. M., Molina, M. J., and Zhang, R. Y.: Elucidating severe urban haze formation in China, *P Natl Acad Sci USA*, 111, 17373-17378, 10.1073/pnas.1419604111, 2014.
- Han, X., Zhang, M. G., Tao, J. H., Wang, L. L., Gao, J., Wang, S. L., and Chai, F. H.: Modeling aerosol impacts on atmospheric visibility in Beijing with RAMS-CMAQ, *Atmos Environ*, 72, 177-191, 10.1016/j.atmosenv.2013.02.030, 2013.
- 15 Henze, D. K., Hakami, A., and Seinfeld, J. H.: Development of the adjoint of GEOS-Chem, *Atmos Chem Phys*, 7, 2413-2433, DOI 10.5194/acp-7-2413-2007, 2007.
- Ho, K. F., Ho, S. S. H., Huang, R. J., Chuang, H. C., Cao, J. J., Han, Y. M., Lui, K. H., Ning, Z., Chuang, K. J., Cheng, T. J., Lee, S. C., Hu, D., Wang, B., and Zhang, R. J.: Chemical composition and bioreactivity of PM_{2.5} during 2013 haze events in China, *Atmos Environ*, 126, 162-170, 2016.
- 20 Hua, Y., Cheng, Z., Wang, S. X., Jiang, J. K., Chen, D. R., Cai, S. Y., Fu, X., Fu, Q. Y., Chen, C. H., Xu, B. Y., and Yu, J. Q.: Characteristics and source apportionment of PM_{2.5} during a fall heavy haze episode in the Yangtze River Delta of China, *Atmos Environ*, 123, 380-391, 10.1016/j.atmosenv.2015.03.046, 2015.
- Huang, K., Zhuang, G., Lin, Y., Fu, J. S., Wang, Q., Liu, T., Zhang, R., Jiang, Y., Deng, C., Fu, Q., Hsu, N. C., and Cao, B.: Typical types and formation mechanisms of haze in an Eastern Asia megacity, Shanghai, *Atmos Chem Phys*, 12, 105-124, 2012.
- 25 Im, U., Christensen, J. H., Geels, C., Hansen, K. M., Brandt, J., Solazzo, E., Alyuz, U., Balzarini, A., Baro, R., Bellasio, R., Bianconi, R., Bieser, J., Colette, A., Curci, G., Farrow, A., Flemming, J., Fraser, A., Jimenez-Guerrero, P., Kitwiroon, N., Liu, P., Nopmongkol, U., Palacios-Pena, L., Pirovano, G., Pozzoli, L., Prank, M., Rose, R., Sokhi, R., Tuccella, P., Unal, A., Vivanco, M. G., Yarwood, G., Hogrefe, C., and Galmarini, S.: Influence of anthropogenic emissions and boundary conditions on multi-model simulations of major air pollutants over Europe and North America in the framework of AQMEII3, *Atmos Chem Phys*, 18, 8929-8952, 10.5194/acp-18-8929-2018, 2018.
- 30 Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., Keating, T., Zhang, Q., Kurokawa, J., Wankmuller, R., van der Gon, H. D., Kuenen, J. J. P., Klimont, Z., Frost, G., Darras, S., Koffi, B., and Li, M.: HTAP_v2.2: a mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution, *Atmos Chem Phys*, 15, 11411-11432, 2015.
- 35 Leibensperger, E. M., Mickley, L. J., Jacob, D. J., and Barrett, S. R. H.: Intercontinental influence of NO_x and CO emissions on particulate matter air quality, *Atmos Environ*, 45, 3318-3324, 10.1016/j.atmosenv.2011.02.023, 2011.
- 40 Li, M., Zhang, Q., Kurokawa, J., Woo, J. H., He, K. B., Lu, Z. F., Ohara, T., Song, Y., Streets, D. G., Carmichael, G. R., Cheng, Y. F., Hong, C. P., Huo, H., Jiang, X. J., Kang, S. C., Liu, F., Su, H., and Zheng, B.: MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP, *Atmos Chem Phys*, 17, 935-963, 10.5194/acp-17-935-2017, 2017.
- 45 Li, W. J., Cao, Y., Li, R. K., Ma, X. M., Chen, J. Y., Wu, Z. L., and Xu, Q.: The spatial variation in the effects of air pollution on cardiovascular mortality in Beijing, China, *J Expo Sci Env Epid*, 28, 297-304, 10.1038/jes.2016.21, 2018.
- Li, Y., and Zhu, L.: Cost of energy saving and CO₂ emissions reduction in China's iron and steel sector, *Appl Energy*, 130, 603-616, 2014.

- Liu, J., Mauzerall, D. L., and Horowitz, L. W.: Source-receptor relationships between East Asian sulfur dioxide emissions and Northern Hemisphere sulfate concentrations, *Atmos Chem Phys*, 8, 3721-3733, DOI 10.5194/acp-8-3721-2008, 2008.
- 5 Liu, L. W., Chen, C. X., Zhao, Y. F., and Zhao, E. D.: China's carbon-emissions trading: Overview, challenges and future, *Renew Sust Energy Rev*, 49, 254-266, 10.1016/j.rser.2015.04.076, 2015.
- Lu, Z., Streets, D. G., Zhang, Q., Wang, S., Carmichael, G. R., Cheng, Y. F., Wei, C., Chin, M., Diehl, T., and Tan, Q.: Sulfur dioxide emissions in China and sulfur trends in East Asia since 2000, *Atmos Chem Phys*, 10, 6311-6331, 10.5194/acp-10-6311-2010, 2010.
- 10 Rao, S. T., Galmarini, S., and Puckett, K.: Air Quality Model Evaluation International Initiative (AQMEII) Advancing the State of the Science in Regional Photochemical Modeling and Its Applications, *B Am Meteorol Soc*, 92, 23-30, 2011.
- Shen, R. R., Schafer, K., Shao, L. Y., Schnelle-Kreis, J., Wang, Y. S., Li, F. X., Liu, Z. R., Emeis, S., and Schmid, H. P.: Chemical characteristics of PM_{2.5} during haze episodes in spring 2013 in Beijing, *Urban Clim*, 22, 51-63, 10.1016/j.uclim.2016.01.003, 2017.
- 15 Simpson, D., Benedictow, A., Berge, H., Bergstrom, R., Emberson, L. D., Fagerli, H., Flechard, C. R., Hayman, G. D., Gauss, M., Jonson, J. E., Jenkin, M. E., Nyiri, A., Richter, C., Semeena, V. S., Tsyro, S., Tuovinen, J. P., Valdebenito, A., and Wind, P.: The EMEP MSC-W chemical transport model - technical description, *Atmos Chem Phys*, 12, 7825-7865, 10.5194/acp-12-7825-2012, 2012.
- Stohl, A., Eckhardt, S., Forster, C., James, P., and Spichtinger, N.: On the pathways and timescales of intercontinental air pollution transport, *J Geophys Res-Atmos*, 107, 2002.
- 20 Sudo, K., Takahashi, M., Kurokawa, J., and Akimoto, H.: CHASER: A global chemical model of the troposphere - 1. Model description, *J Geophys Res-Atmos*, 107, 2002.
- [Streets, D. G., van Aardenne, J., Battye, B., Garivait, S., Grano, D., Guenther, A., Klimont, Z., Lamarque, J. F., Lu, Z., Maenhout, G., Ohara, T., Parrish, D. J., Smith, S. J., and Vallack, H.: Emission Inventories and Projections, in: Hemispheric Transport of Air Pollution \(2010\) Part A: Ozone and Particulate Matter, Chapter 3. Air Pollution Studies No. 17, edited by: Dentener, F., Keating, T., and Akimoto, H., 77-133, UNECE Information Service, Geneva, Switzerland, 2010.](#)
- 25
- Takemura, T., Nozawa, T., Emori, S., Nakajima, T. Y., and Nakajima, T.: Simulation of climate response to aerosol direct and indirect effects with aerosol transport-radiation model, *J Geophys Res-Atmos*, 110, ~~Art~~ D02202
- 30 10.1029/2004jd005029, 2005.
- Tie, X. X., Wu, D., and Brasseur, G.: Lung cancer mortality and exposure to atmospheric aerosol particles in Guangzhou, China, *Atmos Environ*, 43, 2375-2377, 10.1016/j.atmosenv.2009.01.036, 2009.
- 35 Tilmes, S., Lamarque, J. F., Emmons, L. K., Kinnison, D. E., Marsh, D., Garcia, R. R., Smith, A. K., Neely, R. R., Conley, A., Vitt, F., Martin, M. V., Tanimoto, H., Simpson, I., Blake, D. R., and Blake, N.: Representation of the Community Earth System Model (CESM1) CAM4-chem within the Chemistry-Climate Model Initiative (CCMI), *Geosci Model Dev*, 9, 1853-1890, 10.5194/gmd-9-1853-2016, 2016.
- Torseth, K., Aas, W., Breivik, K., Fjaeraa, A. M., Fiebig, M., Hjellbrekke, A. G., Myhre, C. L., Solberg, S., and Yttri, K. E.: Introduction to the European Monitoring and Evaluation Programme (EMEP) and observed atmospheric composition change during 1972-2009, *Atmos Chem Phys*, 12, 5447-5481, 2012.
- 40 Wang, J. D., Wang, S. X., Jiang, J. K., Ding, A. J., Zheng, M., Zhao, B., Wong, D. C., Zhou, W., Zheng, G. J., Wang, L., Pleim, J. E., and Hao, J. M.: Impact of aerosol-meteorology interactions on fine particle pollution during China's severe haze episode in January 2013, *Environ Res Lett*, 9, 2014a.
- 45 Wang, L. T., Wei, Z., Yang, J., Zhang, Y., Zhang, F. F., Su, J., Meng, C. C., and Zhang, Q.: The 2013 severe haze over southern Hebei, China: model evaluation, source apportionment, and policy implications, *Atmos Chem Phys*, 14, 3151-3173, 10.5194/acp-14-3151-2014, 2014b.

- Wang, Q. Z., Zhuang, G. S., Huang, K., Liu, T. N., Deng, C. R., Xu, J., Lin, Y. F., Guo, Z. G., Chen, Y., Fu, Q. Y., Fu, J. S. S., and Chen, J. K.: Probing the severe haze pollution in three typical regions of China: Characteristics, sources and regional impacts, *Atmos Environ*, 120, 76-88, 2015.
- 5 Wang, Y. J., Li, L., Chen, C. H., Huang, C., Huang, H. Y., Feng, J. L., Wang, S. X., Wang, H. L., Zhang, G., Zhou, M., Cheng, P., Wu, M. H., Sheng, G. Y., Fu, J. M., Hu, Y., Russell, A. G., and Wumaer, A.: Source apportionment of fine particulate matter during autumn haze episodes in Shanghai, China, *J Geophys Res-Atmos*, 119, 1903-1914, 10.1002/2013JD019630, 2014c.
- 10 West, J. J., Naik, V., Horowitz, L. W., and Fiore, A. M.: Effect of regional precursor emission controls on long-range ozone transport - Part 1: Short-term changes in ozone air quality, *Atmos Chem Phys*, 9, 6077-6093, DOI 10.5194/acp-9-6077-2009, 2009.
- WorldBank, T.: Cost of Pollution in China, Rural Development, Natural Resources and Environment Management Unit, Washington, D.C., 2007.
- Xu, P., Chen, Y. F., and Ye, X. J.: Haze, air pollution, and health in China, *Lancet*, 382, 2067-2067, Doi 10.1016/S0140-6736(13)62693-8, 2013.
- 15 Yin, L. Q., Niu, Z. C., Chen, X. Q., Chen, J. S., Xu, L. L., and Zhang, F. W.: Chemical compositions of PM_{2.5} aerosol during haze periods in the mountainous city of Yong'an, China, *J Environ Sci-China*, 24, 1225-1233, 2012.
- Zhang, L., Jacob, D. J., Kopacz, M., Henze, D. K., Singh, K., and Jaffe, D. A.: Intercontinental source attribution of ozone pollution at western US sites using an adjoint method, *Geophys Res Lett*, 36, Artn L11810
- 20 10.1029/2009gl037950, 2009.
- Zhang, X. Y., Wang, Y. Q., Niu, T., Zhang, X. C., Gong, S. L., Zhang, Y. M., and Sun, J. Y.: Atmospheric aerosol compositions in China: spatial/temporal variability, chemical signature, regional haze distribution and comparisons with global aerosols, *Atmos Chem Phys*, 12, 779-799, 2012.
- 25 Zhu, J. L., Liao, H., and Li, J. P.: Increases in aerosol concentrations over eastern China due to the decadal-scale weakening of the East Asian summer monsoon, *Geophys Res Lett*, 39, Artn-L09809
- 10.1029/2012gl051428, 2012.

Figure Captions

5 Figure 1. The HTAP2 source and receptor regions for EUR (green), RBU (red), and EAS (grey). Sites marked with the same symbols are from the same observation network: red circles represent API, blue squares represent AERONET, orange diamonds represent EANET, and yellow triangles represent EBAS.

10 Figure 2. Monthly mean surface concentrations of O_3 (left column), $PM_{2.5}$ (center column), and PM_{10} (right column) for the year 2010 in the EUR (upper row) and EAS¹ (lower row) regions from observations and model simulations³. Observations (bold black lines with vertical error bars) represent the averages of all sites falling within the same ensemble grid (bold red lines), and the vertical error bars² depict the standard deviation across the sites in the same ensemble grid. Models are sampled at the nearest grid to each station, multiple stations within the same model grid are averaged to represent the paring observation.

15 Figure 3. Monthly average AOD comparison between the models and AERONET (upper row) and between the models and the MODIS (bottom row) in EUR (left column), RBU (center column), and EAS (right column). Models are represented by markers with different colors and styles. Evaluation statistics (MB and R^2) are indicated for model ensemble mean in the upper left corner of the scatter plot. The solid black line is the 1:1 line whereas the black dash contours represent the 1:2 and 2:1 lines.

Figure 4. Spatial distributions of AOD from MODIS and model simulations. Evaluation statistics of each model are indicated at the lower left corner of the plot.

20 Figure 5. Monthly averages of surface aerosol response in the EAS receptor region under the EURALL scenario. Solid bars with different colors represent the responses of different aerosol.

Figure 6. Same as Figure 5 but under the RBUALL scenario.

Figure 7. Annual averages of PM column density responses (calculated as $\Delta PM = \Delta BC + \Delta SO_4^{2-} + \Delta OA$) under the EURALL scenario within (left column) and above (middle column) PBL, and the corresponding longitude-pressure cross sections of PM concentrations (averaged over 10°N-70°N) estimated by participating models.

25 Figure 8. Same as Figure 7 but under the RUBALL scenario

Figure 9. Spatial distributions and histograms of the long-range transport full impacts during the haze episodes. Model grids with no NCDC observation sites located in are assigned to fill values.

30 Figure 10. Reduction of visibility (left column) and enhancement of number of haze days (right column) under the EURALL (upper row) and RBUALL (lower row) scenarios.

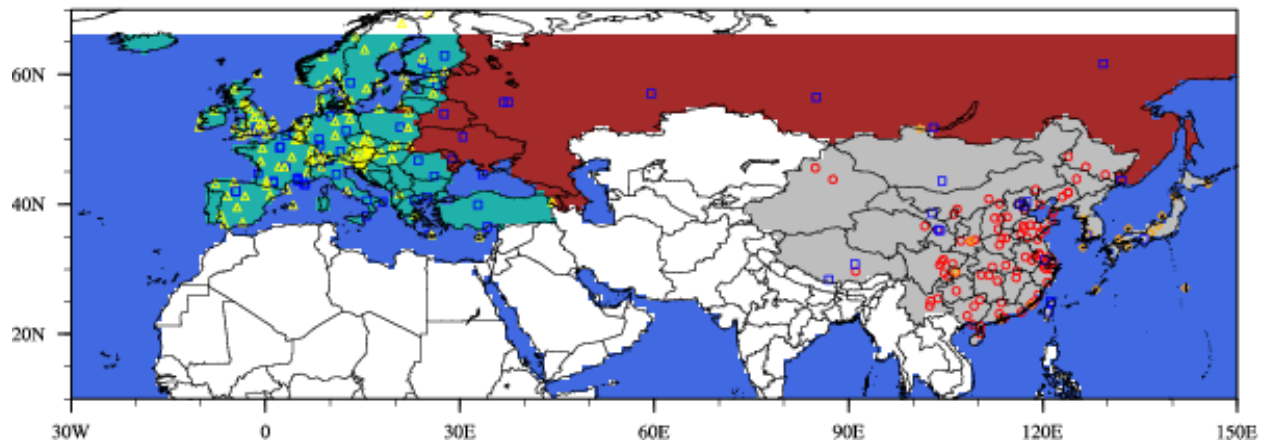


Figure 1. The HTAP2 source and receptor regions for EUR (green), RBU (red), and EAS (grey). Sites marked with the same symbols are from the same observation network: red circles represent API, blue squares represent AERONET, orange diamonds represent EANET, and yellow triangles represent EBAS.

5

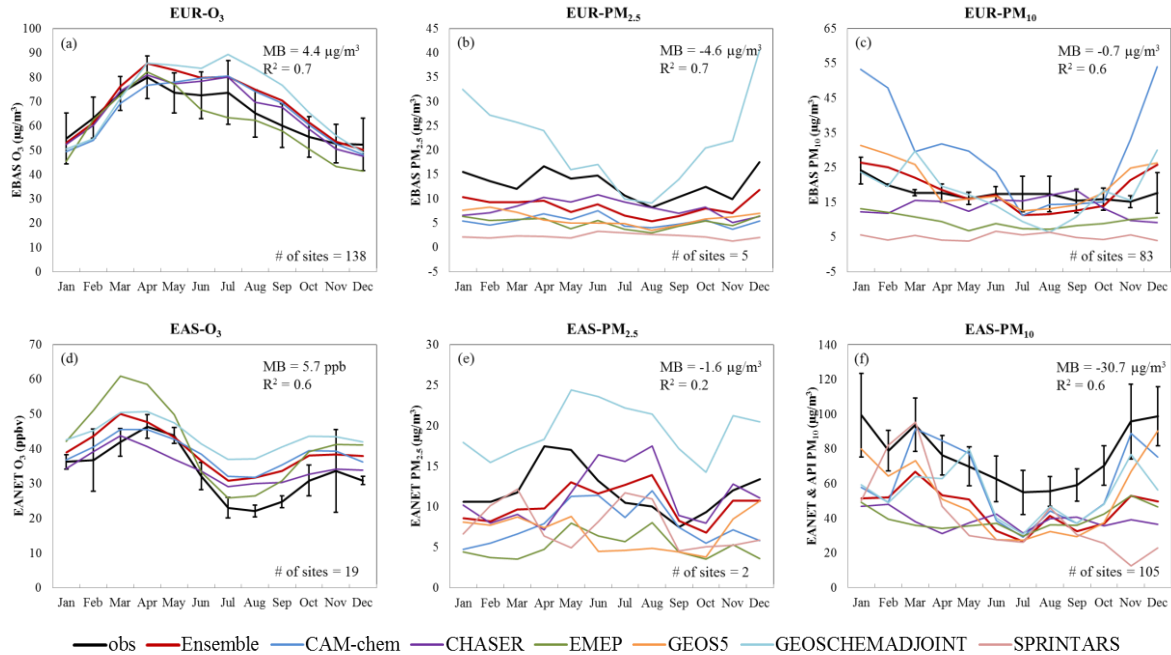
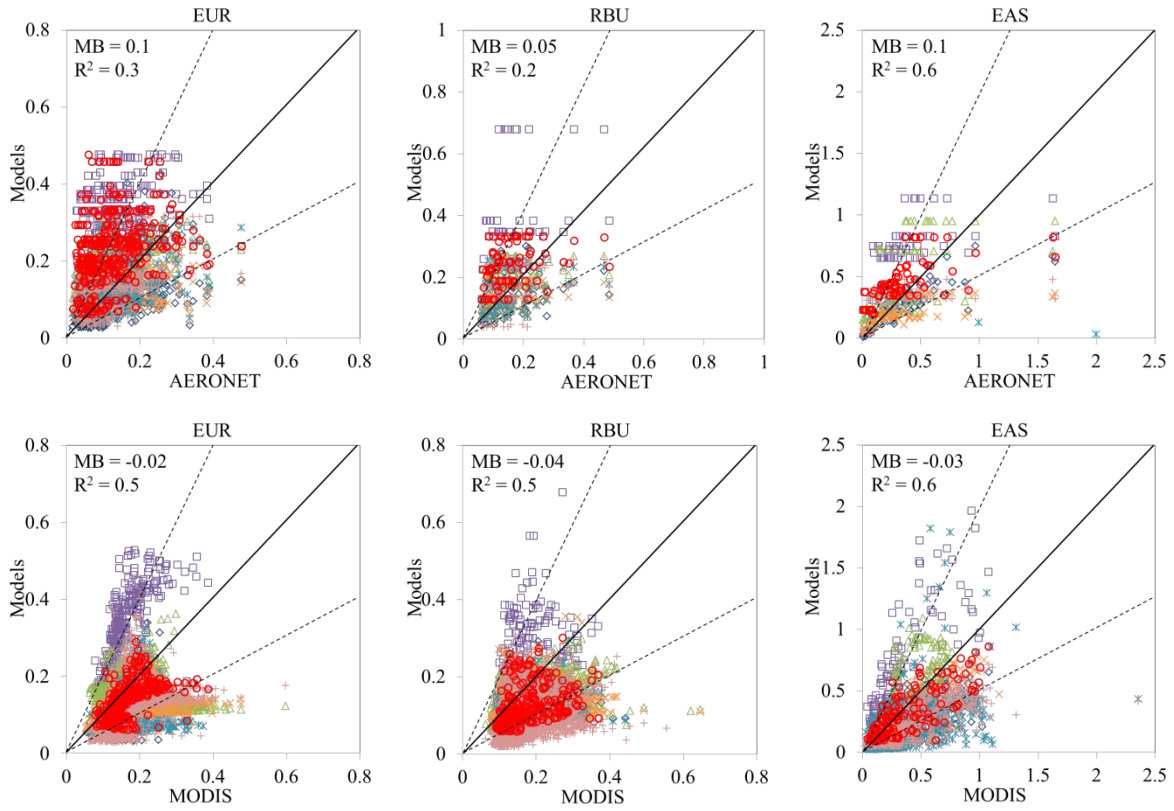


Figure 2. Monthly mean surface concentrations of O_3 (left column), $PM_{2.5}$ (center column), and PM_{10} (right column) for the year 2010 in the EUR (upper row) and EAS¹ (lower row) regions from observations and model simulations³. Observations (bold black lines with vertical error bars) represent the averages of all sites falling within the same ensemble grid (bold red lines), and the vertical error bars² depict the standard deviation across the sites in the same ensemble grid. Models are sampled at the nearest grid to each station, multiple stations within the same model grid are averaged to represent the pairing observation.

¹ PM_{10} from API and EANET are used together to represent the observations in EAS region.

² $PM_{2.5}$ observations in EUR and EAS region have no standard deviation because there are no sites with valid measurements fall into the same model ensemble mean grid.

³Most participating models report the $PM_{2.5}$ mass concentration except that CAM-chem only reports the aerosol sub-species, so we calculate the CAM-chem simulated $PM_{2.5}$ by following the formula described in Silva et al. (2013).



○ Ensemble ◇ CAM-chem □ CHASER △ EMEP × GEOS5 * GEOSCHEMADJOINT + SPRINTARS

Figure 3. Monthly average AOD comparison between the models and AERONET (upper row) and between the models and the MODIS (bottom row) in EUR (left column), RBU (center column), and EAS (right column). Models are represented by markers with different colors and styles. Evaluation statistics (MB and R²) are indicated for model ensemble mean in the upper left corner of the scatter plot. The solid black line is the 1:1 line whereas the black dash contours represent the 1:2 and 2:1 lines.

5

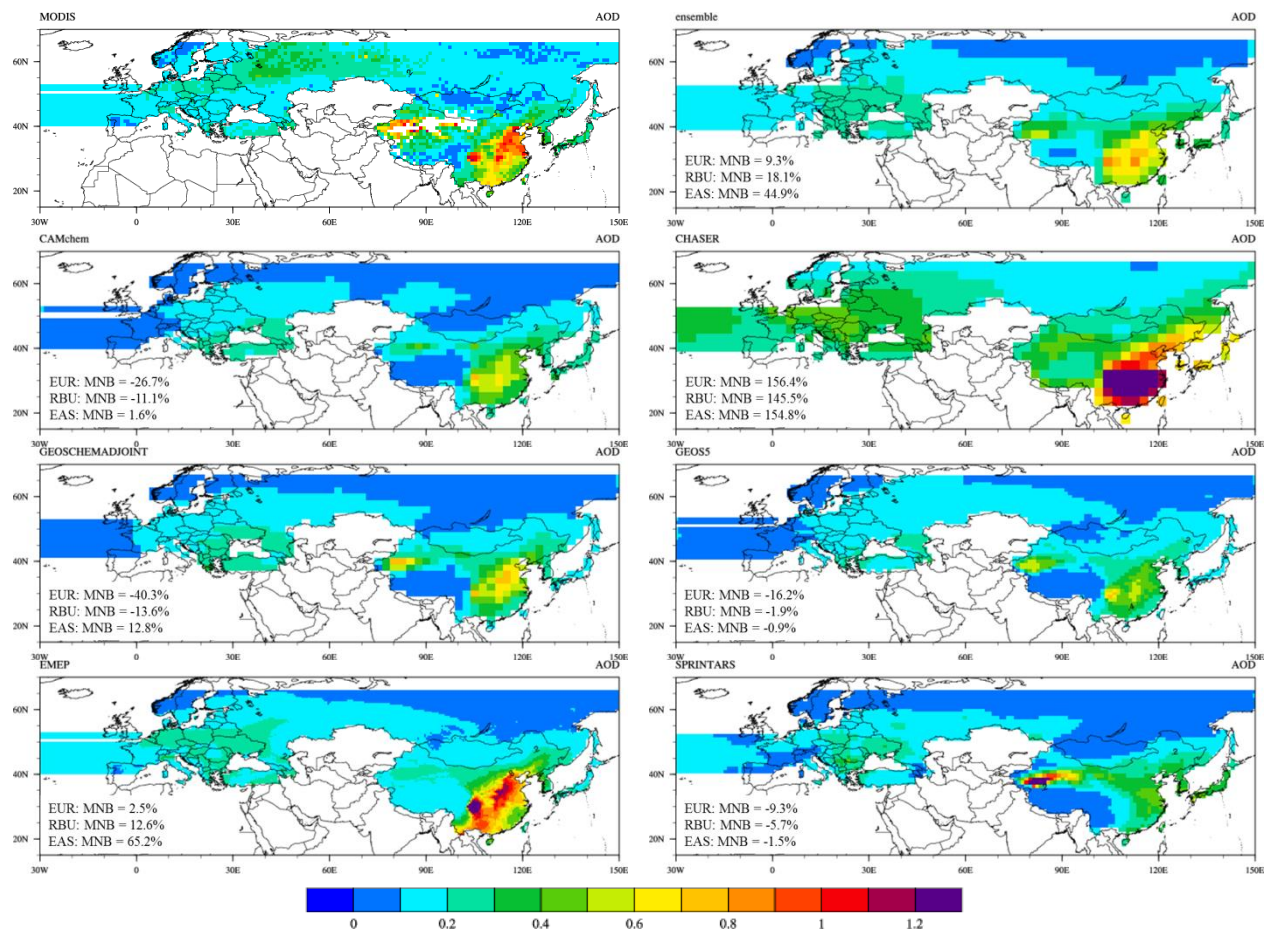


Figure 4. Spatial distributions of AOD from MODIS and model simulations. Evaluation statistics of each model are indicated at the lower left corner of the plot.

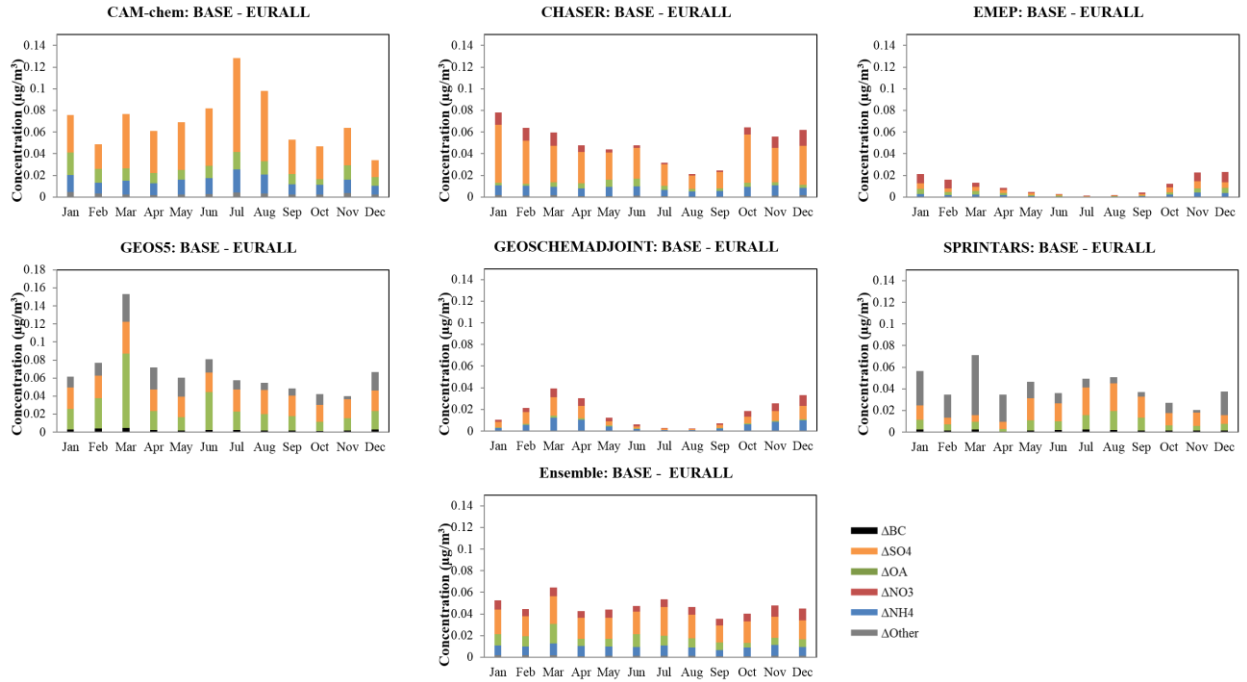


Figure 5. Monthly averages of surface aerosol response in the EAS receptor region under the EURALL scenario. Solid bars with different colors represent the responses of different aerosol.

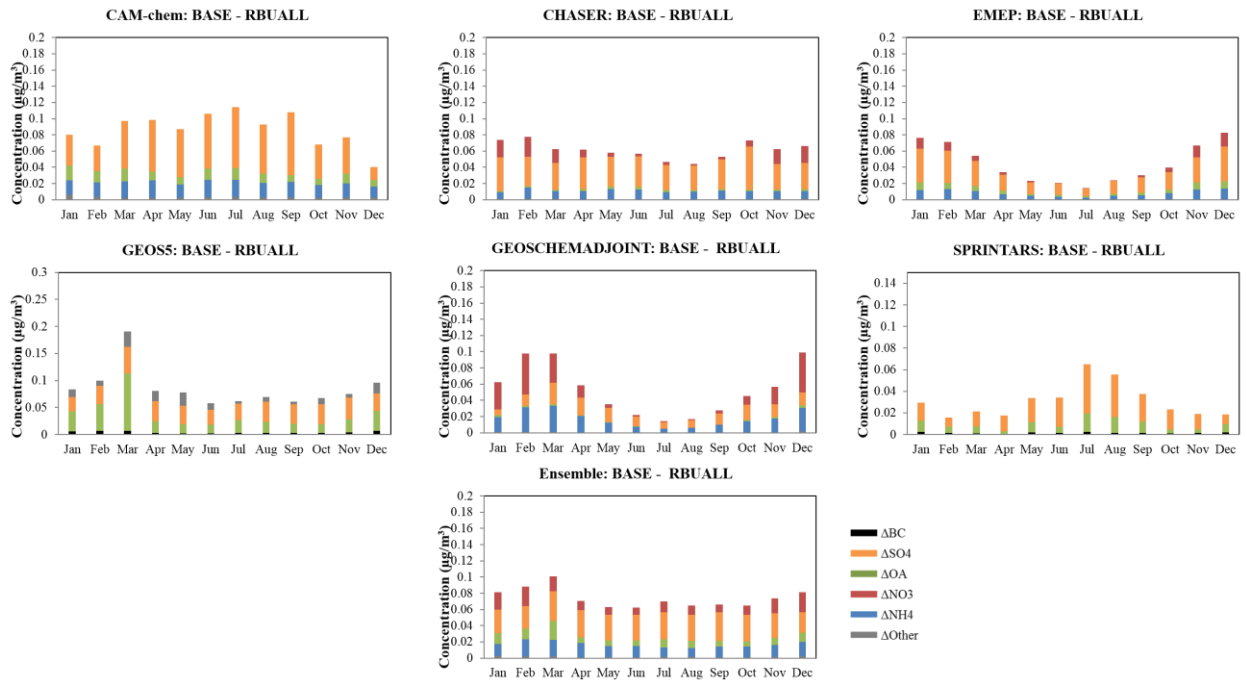


Figure 6. Same as Figure [xxx_5](#) but under the RBUALL scenario.

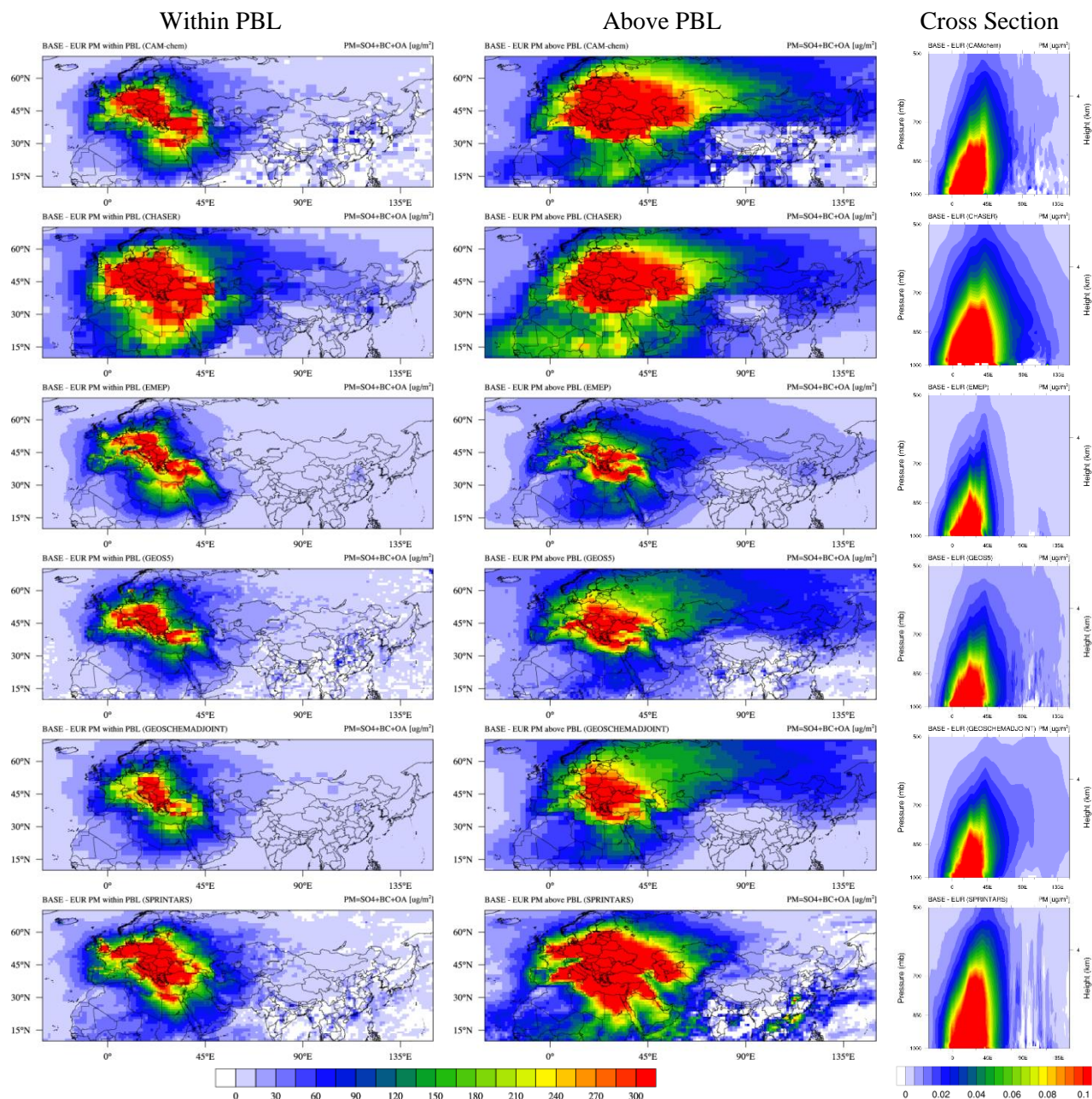


Figure 7. Annual averages of PM column density responses (calculated as $\Delta PM = \Delta BC + \Delta SO_4^{2-} + \Delta OA$) under the EURALL scenario within (left column) and above (middle column) PBL, and the corresponding longitude-pressure cross sections of PM concentrations (averaged over 10°N-70°N) estimated by participating models.

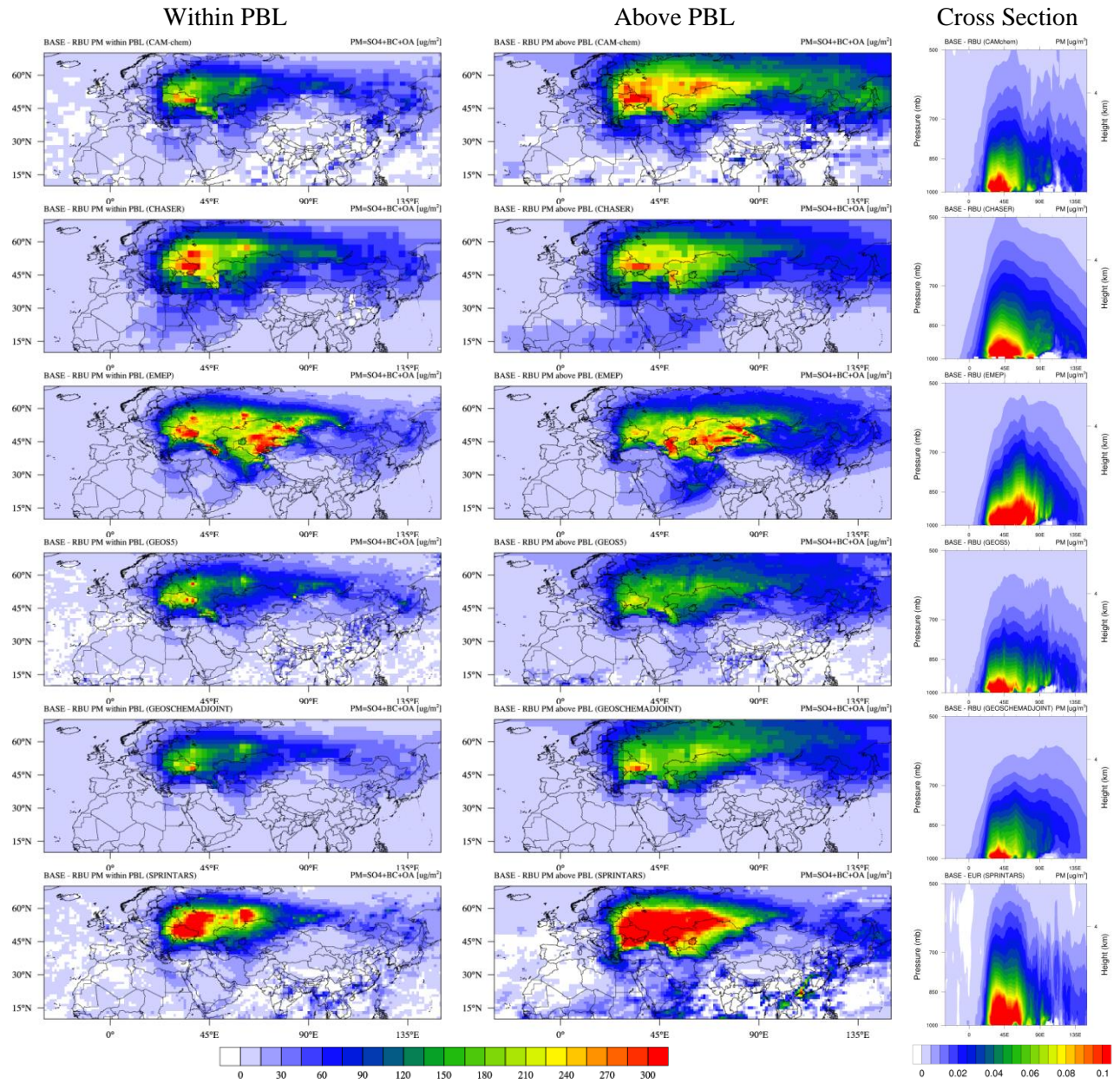
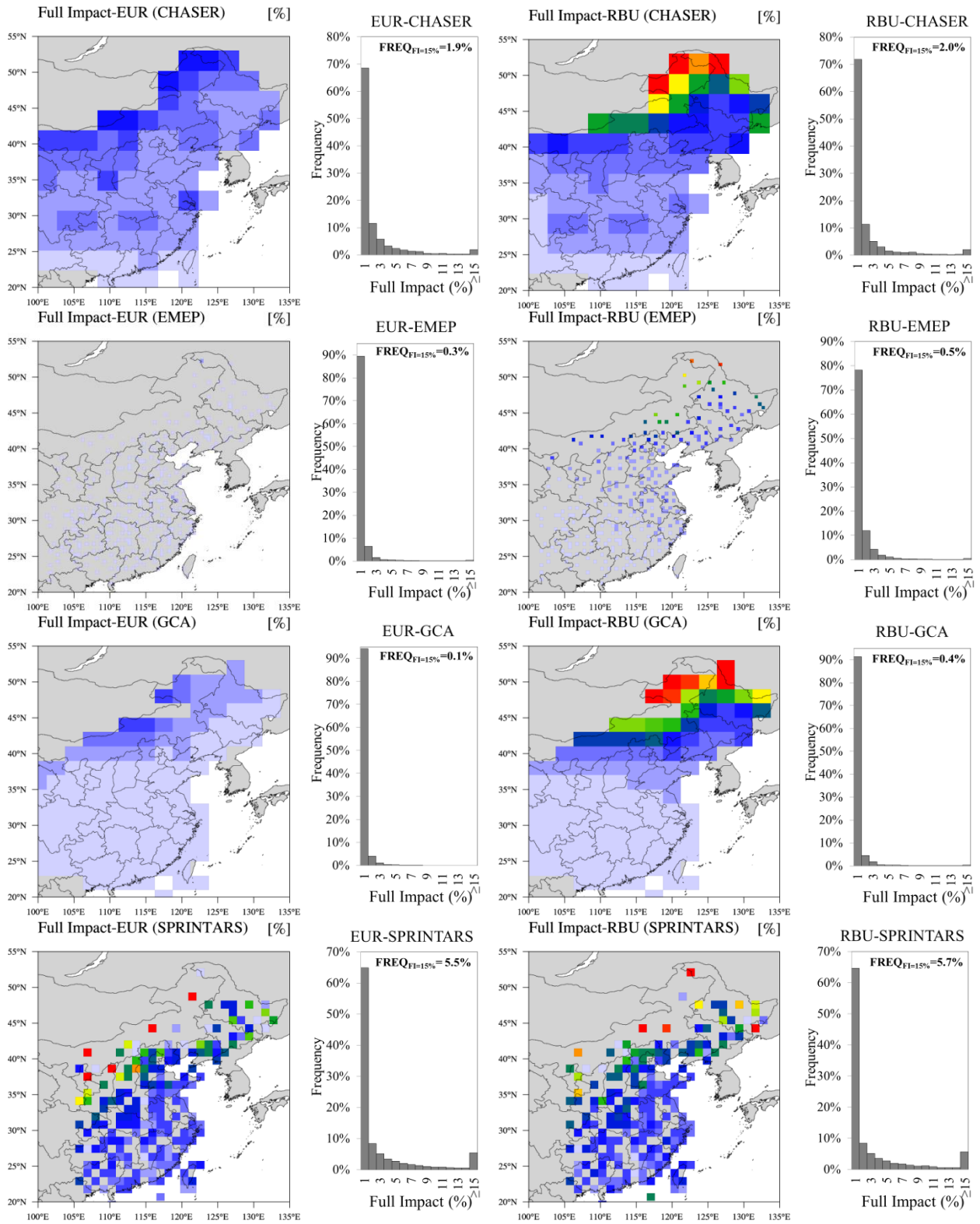


Figure 8. Same as Figure 7 but under the RUBALL scenario



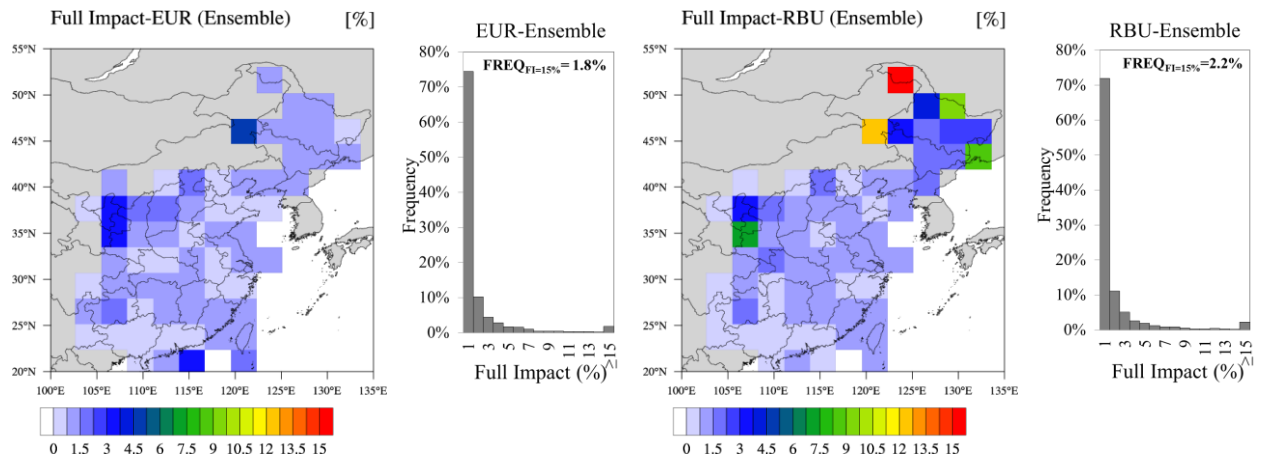


Figure 9. Spatial distributions and histograms of the long-range transport full impacts during the haze episodes. Model grids with no NCDC observation sites located in are assigned to fill values.

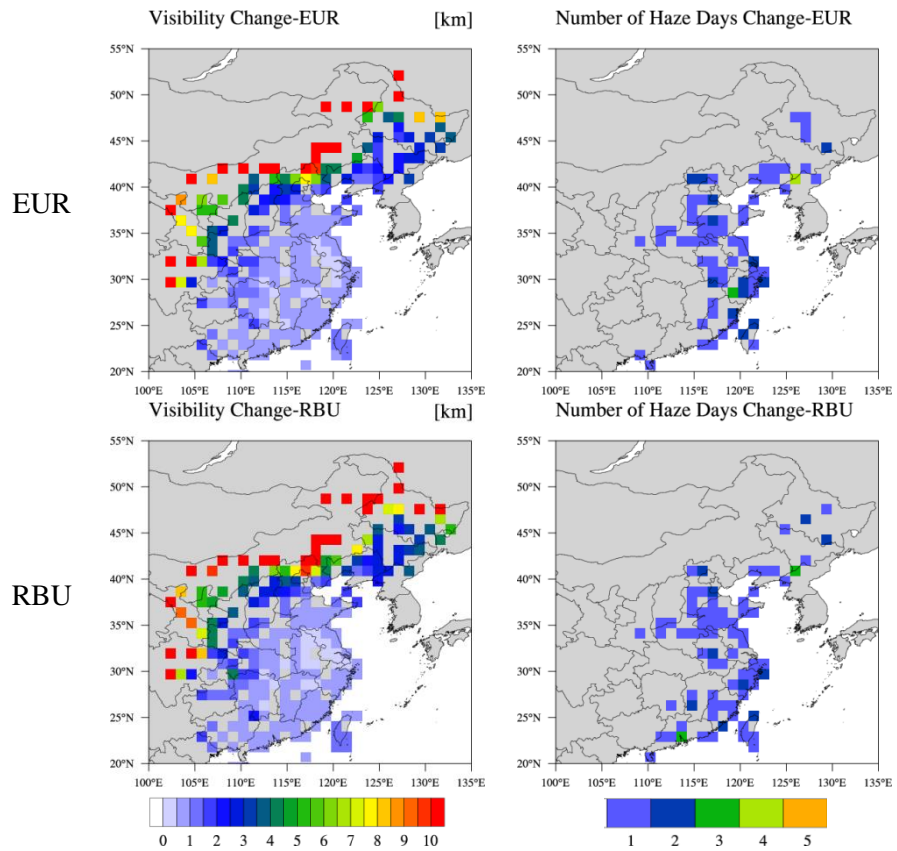


Figure 10. Reduction of visibility (left column) and enhancement of number of haze days (right column) under the EURALL (upper row) and RBUALL (lower row) scenarios.

Table 1. Models used for this study

Model	Resolution (lat/lon/vertical)	Meteorology	Model Reference
CAM-chem	1.9°×2.5°×56	GEOS5 v5.2	(Tilmes et al., 2016)
CHASER	2.8°×2.8°×32	ERA-Interim and HadISST	(Sudo et al., 2002)
EMEP	0.5°×0.5°×20	ECMWF-IFS	(Simpson et al., 2012)
GEOS5	1.0°×0.75°×72	MERRA	Rienecker et al. (2008)
GEOSCHEMADJOINT	2.0°×2.5°×72	MERRA	(Henze et al., 2007)
<u>SPRINTASSPRINTARS</u>	1.1°×1.1°×56	ECMWF Interim	(Takemura et al., 2005)
Model Ensemble Mean	2.8°×2.8°×32	-	-

Table 2. Annual average long-range transport impacts of surface PM concentrations and percentage contributions from the EUR and RBU source regions to the EAS receptor region. Numbers collected from the HTAP1 assessment are presented in *Italic font*, aerosol surface concentrations (Surf. Conc.) under the baseline scenario are presented in **bold font**. Numbers in the parentheses indicate the range of each variable among the participating models.

		Long-range transport Full Impact				
		EA as receptor			EAS as receptor	
		EU→EA	EUR→EA			
		2000 ¹	2010EA ²	2008 ³	2009 ⁴	2010
SO ₄ ²⁻	Surf. Conc. (µg/m³)	<i>2.94 (1.96-4.42)</i>	<i>3.25 (2.07-5.46)</i>	<i>5.9 (5.38-6.51)</i>	<i>5.29</i>	<i>3.80 (1.45-6.67)</i>
	<i>Full_Impact_{EUR%}</i>	<i>5.0 (0.3-9.8)</i>	0.5 (0.1-0.9)	3.5 (2.9-4.1)	4.7	2.7 (0.4-5.6)
	<i>Full_Impact_{RBU%}</i>			5.5	5.2	4.1 (2.6-6.9)
BC	Surf. Conc. (µg/m³)	<i>0.42 (0.28-0.71)</i>	<i>0.56 (0.34-0.74)</i>	<i>1.00 (0.93-1.08)</i>	<i>0.92</i>	<i>0.82 (0.51-1.07)</i>
	<i>Full_Impact_{EUR%}</i>	<i>1.0 (0.5-3.9)</i>	0.2 (0.03-0.3)	1.2 (0.6-1.8)	1.9	1.1 (0.1-2.2)
	<i>Full_Impact_{RBU%}</i>			3.6	1.8	1.1 (0.1-2.5)
OA	Surf. Conc. (µg/m³)	<i>1.46 (0.81-2.52)</i>	<i>3.56 (1.93-6.29)</i>	<i>6.28 (3.51-9.06)</i>	<i>3.37</i>	<i>5.06 (2.1-8.87)</i>
	<i>Full_Impact_{EUR%}</i>	<i>0.4 (0.2-0.9)</i>	0.2 (0.02-0.4)	0.7 (0.3-1.1)	2.1	0.9 (0.1-1.2)
	<i>Full_Impact_{RBU%}</i>			2.5	2.0	1.0 (0.1-3.2)
		Long-range transport Relative Impact				
SO ₄ ²⁻ +OA	<i>Relative_Impact_{EUR%}</i>	2.9	2.2	2.9	2.8	2.7
	<i>Relative_Impact_{RBU%}</i>		3.3 (2.1-5.5)	3.8	3.3	3.7
		Local 20% anthropogenic emission perturbation impact				
SO ₄ ²⁻ +OA	$\frac{\Delta PM_{EAS}}{PM_{BASE}} \times 100\%$	16.8	12.5	14.0	14.1	12

5 ¹Numbers shown for 2000 are collected from the HTAP1 report that representing the long-range transport impact from EU to EA.

²2010EA is calculated with the HTAP2 data by using the HTAP1 domain configuration for EA

³Only two models (CAM-chem and CHASER) data are available for EURALL scenario in 2008, and only one model (CAM-chem) data is available for RBUALL scenario in 2008, so no range is calculated for RBU%.

10 ⁴Only one model (CAM-chem) 2009 data is available so no range is calculated for EUR% and RBU%.

Table 3. Long-range transport full impacts on annual average scale and during the haze episodes. Numbers in the parentheses indicate the percentage contributions.

Models	Base PM _{2.5} [$\mu\text{g}/\text{m}^3$]		EUR Full Impact [$\mu\text{g}/\text{m}^3$ (%)]		RBU Full Impact [$\mu\text{g}/\text{m}^3$ (%)]	
	AAVG	HAZE	AAVG	HAZE	AAVG	HAZE
CHASER	20.46	47.73	0.23 (1.2)	1.00 (2.1)	0.29 (1.4)	0.99 (2.1)
EMEP	17.35	29.34	0.05 (0.3)	0.11 (0.4)	0.23 (1.3)	0.61 (2.1)
GCA ¹	25.47	28.03	0.12 (0.3)	0.29 (1.1)	0.35 (1.4)	0.86 (3.0)
<u>SPRINTASS</u>	17.45	24.80	1.00 (5.7)	2.58 (10.5)	1.26 (7.2)	2.82 (11.4)
<u>PRINTARS</u>						
Ensemble	20.18	32.48	0.35 (1.7)	0.99 (3.1)	0.53 (2.6)	1.32 (4.1)

¹GCA: GEOSCHEMADJOINT