Anonymous Referee #2

General Summary:

This study compares the performances of 12 regional models from MICS-Asia III in simulating particulate matter during 2010. Performance is evaluated in terms of PM_{10}, $PM_{2.5}$, sulphate, nitrate, and ammonium. The mean bias of multimodel mean range from -30% to 12% for different species. The analysis focuses on four processes that are likely contributing to the inter-model differences namely natural emissions, gas to particle conversion of sulphur and oxidized nitrogen, the role of dust emissions, and deposition parameterization. I think the subject of this paper is suitable for publication in ACP but I have several concerns about the organization and analysis in the paper. Therefore, I recommend major revisions before the paper can be published in ACP.

Firstly, the supplementary information is overwhelming and contains a lot of the material that can be moved to the main paper. Specifically, I suggest moving the evaluation part to the main paper and several key discussions (e.g., supplementary sect. S2.5, table S1, and fig. S8) related to the process analysis.

Second, I have some concerns using CO and BC as surrogates for anthropogenic emissions (see specific comment on lines 258-259).

Third, I have some concerns regarding evaluation with MODIS AOD (see specific comments on the evaluation part).

In addition, there are many grammatical errors that I believe can be removed upon reading by a native English speaker.

Response: We would like to thank the reviewer for the valuable comments. Following are the responses to the comments.

Here are my specific and minor comments:

2Q1: Comment: Abstract: Line 32-34: Why would natural emissions affect BC concentrations and CO loading?

Response: Thank you for pointing out the question. All models used the same anthropogenic and natural emissions. Thus, the impact of different natural emissions on model performance is not an application in this study. However, mismatch during the temporal and vertical treatment of emission files by different modelling groups has caused differences in the model inputs (Itahashi et al., 2019). In addition, different model set-up such as the heights of first vertical layer and spatial resolutions of model grids also affect the inter-model comparison on direct model outputs such as surface PM concentrations. Therefore, we decided to use indicators (such as SOR) instead of direct model outputs to facilitate comparison on model mechanisms. For this reason, we removed our discussion on the impacts of emissions and IC/BC on model performance (section 3.2 in the old manuscript) and focused on the comparison of the three processes: gas-aerosol portioning, dust mechanism and wet and dry deposition efficiency.

Reference: Itahashi, S., Ge, B., Sato, K., Fu, J. S., Wang, X., Yamaji, K., Nagashima, T., Li, J., Kajino, M., Liao, H., Zhang, M., Wang, Z., Li, M., Kurokawa, J., Carmichael, G. R., and Wang, Z.: MICS-Asia III: Overview of model inter-comparison and evaluation of acid deposition over Asia, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-624, in review, 2019.

2Q2: Comment: Minor comments: Line 58: Change “pollutions” to “pollution”.

Response: We have changed it in the manuscript.

2Q3: Comment: Line 61: Change “impact evaluation on” to “impact evaluation of”

Response: We have revised the manuscript.

2Q4: Comment: Line 86: Change “view on” to “view of”

Response: We have made the change in the manuscript.

2Q5: Comment: Lines 88-89: Spell out different acronyms.

Response: We added the full names of the models in the manuscript.

2Q6: Comment: Table 1: Do you mean AERO5 instead of AEO5 in 6th row.

Response: Table 1 has been removed from the manuscript.

2Q7: Comment: Line 154-155: This is not correct. GOCART in WRF-Chem accounts for ammonium mass in calculation of $PM_{2.5}$ and PM_{10} by multiplying sulfate concentrations by 1.375 to account for missing ammonium. GOCART also includes fine dust and sea-salt in $PM_{2.5}$.

Response: Thank you for pointing out this problem. We removed the description on the model mechanism as they have already been introduced in detail in our companion paper (Chen et al., 2019). The description of the GOCART model is updated in that paper.

Reference: Chen, L., Gao, Y., Zhang, M., Fu, J. S., Zhu, J., Liao, H., Li, J., Huang, K., Ge, B., Wang, X., Lam, Y. F., Lin, C.-Y., Itahashi, S., Nagashima, T., Kajino, M., Yamaji, K., Wang, Z., and Kurokawa, J.: MICS-Asia III: multi-model comparison and evaluation of aerosol over East Asia, Atmos. Chem. Phys., 19, 11911–11937, https://doi.org/10.5194/acp-19-11911-2019, 2019.

2Q8: Comment: Line 161: Change “transport” to “travel”.

Response: We have made the change in the manuscript.

2Q9: Comment: Line 163: Change “taken” to “take”

Response: We have made the change.

2Q10: Comment: Line 176: How about turbulent diffusion?

Response: We have revised the description on the mechanism of dry deposition as follows:

Line 137: “Dry deposition is mainly driven by turbulent and molecular diffusion processes.”

2Q11: Comment: Line 174-180: Please discuss wet deposition parameterization as well.

Response: We added the following sentences to introduce the wet deposition parameterization:

Lines 124-136: "Wet deposition removes gases and aerosols from the atmosphere by rain droplets, involving both in-cloud scavenging (rainout) and below-cloud scavenging (washout). The gases in the atmosphere are dissolved in the raindrop and then removed from the atmosphere. For the non-reactive gases, the removal rate depends on the solubility of gases and is a function of Henry's Law. Particles participate in the cloud condensation nuclei in the presence of supersaturation water vapor and then grow into cloud droplets. In this study, only M2, M4, M6, M11 and M12 have submitted the main components of S and N depositions. All these models use the same wet deposition scheme based on Henry's law. The efficiency of wet deposition is assessed by the so-called "washout ratio", calculated as the ratio of particle concentrations in the deposition to particle concentrations in surface air as shown in Eq. 1.

$$\lambda_{wet} = \frac{C_{depo}}{C_{surface\_air}} \times 100\% \qquad (1)$$

where λ_{wet} is the washout ratio for wet deposition, C_{depo} is the concentration of particles in deposition and $C_{surface\_air}$ is the concentration of particles at near surface atmosphere."

2Q12: Comment: Line 184: Change "summaries" to "summarizes"

Response: We have made the change in the manuscript.

2Q13: Comment: Figure 1: Except for panel (a), I am unable to see the markers for observations in the maps of $PM_{2.5}$, sulfate, nitrate, and ammonium. I suggest showing standard deviations in absolute values as well because the current plots give an impression that large variability (more than 80%) exists over the regions of lowest concentrations (e.g., Tibetan Plateau) but this is simply a result of division by a very small number.

Response: The model performance on aerosols has been updated and discussed in more detail in our companion paper of Chen et al., 2019. All the discussion and figures/tables related to model evaluation have been removed from the manuscript.

The absolute values for standard deviations are shown in supplementary Fig. S2.

2Q14: Comment: Line 195: Can you explain in more detail why dust emissions are so different among different models? Is it because of wind speed or soil moisture or source functions?

Response: We added an explanation to the model differences on dust as follows:

Lines 292-294 "The differences among the models mainly come from the different parameterizations such as source functions, dust-lifting mechanisms and size distributions of particles (Chen et al., 2019)."

2Q15: Comment: Line 215: Remove "slightly".

Response: We have removed it from the manuscript.

2Q16: Comment: Figure S6: Why isn't there a complete seasonal cycle of MODIS AOD at many locations?

Response: The model performance on AOD has been evaluated in our companion paper (Chen et al., 2019) with all available MODIS data. We removed the model evaluation on AOD from this manuscript.

2Q17: Comment: Figure S7: Do white spaces in the maps correspond to AOD below 0.1 or do they also correspond to missing data? Have the model and MODIS AOD been collocated before comparison? Did you use Level 2 MODIS AOD in this comparison?

Response: The model and MODIS AOD haven't collocated before comparison. And Level 2 MODIS AOD is used for comparison. We updated the model performance on AOD in our

companion paper (Chen et al., 2019) and removed the related text and figures/tables from this manuscript.

2Q18: Comment: Lines 219-225: Again, could you please say more about the inter model differences here? Are differences in AOD controlled by differences in aerosol chemical composition or by differences in aerosol size distributions or assumed aerosol optical properties?

Response: AOD is affected by both the chemical composition and particle size distribution. In addition, different model treatments and set-up (as we mentioned in our answer to Q1) could also affect the model results of AOD. We decided to use indicators instead of direct model output like AOD for inter-model comparison. Thus, the discussion related to AOD has been removed from the manuscript.

2Q19: Comment: Line 234: At line 190, MB in PM_{10} for this study is reported as -25 ug/m3 but here it is reported as -11.2 ug/m3. Which one is correct?

Response: Thank you for raising up this question. This study (MICS-Asia III) used three observational datasets: EANET, API, and Ref to evaluate the PM_{10}, and HTAP II used only EANET and API datasets. To make a fair comparison, we recalculated the MICS-Asia III multi-model performance at EANET and API sites. The MB in PM_{10} is changed from -11.2 ug/m^3 to -18.6 ug/m^3.

2Q20: Comment: Line 240: Change "so-call" to "so-called".

Response: We have made the change in the manuscript.

2Q21: Comment: Lines 258-259: It is difficult to understand why it is so hard to get access to the emissions used by different groups. Using model simulated CO and BC as surrogates for emissions is not a good approach because they are strongly influenced not only by emissions but also by transport, deposition, and chemistry (in case of CO). Use of column CO is especially

concerning because CO distribution in the free troposphere is primarily controlled by transport (inflow from domain boundaries in case of regional models) and not by emissions. Similarly, regions downwind of strong emission sources will exhibit disproportionately larger concentrations of BC and CO than the emissions. Therefore, I recommend using actual emission fluxes in your analysis in Section 3.2.

Response: We agree with the reviewer than CO and BC are not appropriate to indicate distributions of emission. Model treatments have caused differences in the model inputs (see the answer for Q1). Thus, we decided to compare indicators instead of direct model outputs in this manuscript. For this reason, section 3.2 is removed from the manuscript.

2Q22: Comment: Lines 325-326: If the models were able to capture SO_2 concentrations, would they have overestimated SO_4 considering that they are simulating SO_4 reasonably well even with underestimated SO_2 concentrations.

Response: As shown in the following figure, the SOR values of models are generally higher than the observation at sites E42, E43 and E44. This high SOR values reveal that models have converted more SO_2 to SO_4^{2-} than the actual condition. In this case, it is possible that SO_4^{2-} concentrations are overestimated while SO_2 concentrations are well fitted to observations.

2Q23: Comment: Line 364: Sea-salt emissions are controlled via a namelist option in WRF-Chem. Was there a specific reason for turning off sea-salt emissions in M7 and M8?

Response: We have confirmed with the model groups that sea-salt emissions are not included in the simulations by M7 and M8.

2Q24: Comment: Line 390: µg cm-3 or µg m-3?

Response: We have changed to µg m^{-3} and kept it consistent in the manuscript.

Anonymous Referee #3

Received and published: 21 November 2019

General Comments: The manuscript examines the spatial and temporal variability in aerosol concentration and composition simulated by 12 regional models. The model predictions are evaluated with measurements from the different monitoring networks. It also quantifies the ensemble mean bias through several processes, namely model inputs, gas to particle conversion, the impact of sea-salt and dust emissions parameterization, as well as deposition processes. The topic and overall approach fit with ACP. However, I feel that there are some aspects need to be discussed in more detail and the paper needs some improvement before being published in ACP.

As Reviewer #2 already mentioned, the supplementary material contains a very large number of figures that could easily be added in the main text. In the current version, the readers have to go back and forth from the main text to the supplementary material.

It would be nice to have a more in-depth investigation of possible causes that could explain the substantial differences between simulations. As an example, there are large differences between simulations performed using the same model (see Figs.2, S9, S12 and S13). The M4 and M5 models seem to share similar setup and use the same input data, except the organic chemistry treatment. Then, how could the authors explain the different modelled spatial distribution of black carbon, SOR or C(NO_2), as well as the different modelled concentrations of NO_2 and NO_3?

The logical connection between the manuscript and the SM needs to be improved. For example, the discussion associated with Fig. S8 appears after the one associated with Fig. S12.

Response: We would like to thank the reviewer for the valuable comments. Following are the responses to the comments.

Specific comments

3Q1: Lines 104-105: it repeats information already provided in the introduction

Response: We have revised the manuscript and deleted the redundant information.

3Q2: Line 115: use “IC/BC” instead “BC”

Response: We have changed it.

3Q3: Lines 164-165: Why the online calculation of natural emissions of dust and sea salt particles fully embedded within the WRF-Chem is not included?

Response: We have confirmed with the modelling groups that the two WRF-Chem models did not include dust and sea-salt emissions in the simulations.

3Q4: Line 185: move this information into the main text

Response: We updated the model performance on aerosols in our companion paper (Chen et al., 2019). We removed the section of model evaluation from this manuscript and focused more on the inter-model comparison.

We also moved the figures and tables from the supplementary material to the main text to make it easier for the readers.

Reference: Chen, L., Gao, Y., Zhang, M., Fu, J. S., Zhu, J., Liao, H., Li, J., Huang, K., Ge, B., Wang, X., Lam, Y. F., Lin, C.-Y., Itahashi, S., Nagashima, T., Kajino, M., Yamaji, K., Wang, Z., and Kurokawa, J.: MICS-Asia III: multi-model comparison and evaluation of aerosol over East Asia, Atmos. Chem. Phys., 19, 11911–11937, https://doi.org/10.5194/acp-19-11911-2019, 2019.

3Q5: Lines 185-186: briefly describe the main findings of Chen et al., 2019

Response: We added the description on the main findings of Chen et al., 2019 as follows:

Lines 176-188: “Evaluation of model performance on aerosols can be found in our companion paper (Chen et al., 2019). The following are the main findings: PM_{10} concentrations were generally underestimated over the simulation domain. $PM_{2.5}$ concentrations were also underestimated over Eastern EA but were well simulated in Central EA. Models failed to reproduce the high peaks of SO_4^{2-} concentration in Central EA, probably due to missing SO_4^{2-} formation mechanisms (such as

heterogeneous SO_4^{2-} chemistry), which has been reported as an important formation pathway of SO_4^{2-} in China. NO_3^- concentrations were overpredicted by most models over the simulation domain and were associated with the underestimation of SO_4^{2-}. M7 and M8 models produced significantly lower NO_3^- concentrations than observations and other models, due to model bias in simulating the NH_3 concentrations and missing the N_2O_5 heterogeneous reaction that severs as an import formation pathway of NO_3^-. The spatial distributions of AOD was generally well simulated except underestimation around the Himalaya mountains, Taklamakan Desert and Gobi Desert."

3Q6: Lines 191-193: should be an NMB of -300%?

Response: We removed section 3.2 from the manuscript (reasons are listed in our response to Q9). This sentence has been removed from the manuscript.

3Q7: Lines 196-197: Fig. S3 instead of Fig. S2?

Response: Here we want to refer to the location of HBT in figure S2 (now figure 1 in the revised manuscript).

3Q8: Lines 205-208: For northern and south-eastern EA these findings are not applicable. Could the authors explain why this behaviour is seen only for eastern EA, but not for the other analysed regions?

Response: On lines 205-208 (manuscript before revision), we described the model performance on simulating the trends of SO_4^{2-} and NO_3^- concentrations in eastern EA. Model underestimations in several months caused the correlations between model ensemble and observation to be low. This is also the reason for low R values in simulating SO_4^{2-} and NO_3^- in northern EA and southern EA.

This sentence has been removed from the manuscript.

3Q9: Lines 258-260: As Reviewer #1 already mentioned, using CO and BC concentration as proxies for emissions won't give a good indication of the emission's spatial distribution.

Response: After checking with the modelling groups, we notice that all models used the same anthropogenic and natural emissions. However, mismatch during the temporal and vertical treatment of emission files by different modelling groups have caused differences in the model inputs (Itahashi et al., 2019). In addition, different model set-up such as the heights of first vertical layer and spatial resolutions of model grids also affect the inter-model comparison on direct model outputs such as surface PM concentrations. Therefore, we decided to use indicators (such as SOR) instead of direct model outputs to facilitate comparison on model mechanisms. For this reason, we removed our discussion on the impacts of emissions and IC/BC on model performance (section 3.2 in the old manuscript) and focused on the comparison of the three processes: gas-aerosol portioning, dust mechanism and wet and dry deposition efficiency.

Reference: Itahashi, S., Ge, B., Sato, K., Fu, J. S., Wang, X., Yamaji, K., Nagashima, T., Li, J., Kajino, M., Liao, H., Zhang, M., Wang, Z., Li, M., Kurokawa, J., Carmichael, G. R., and Wang, Z.: MICS-Asia III: Overview of model inter-comparison and evaluation of acid deposition over Asia, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-624, in review, 2019.

3Q10: Lines 266-269: the models indeed do not provide the layer height in meters, but this can be easily calculated

Response: Thank you for the suggestion. As the reviewer mentioned in Q9, CO is not an appropriate indicator to reveal the vertical distributions of emissions, we removed the discussion and figure related to CO vertical distribution from the manuscript. However, discussion on the inter-model differences on the vertical layers can be found in our companion study (Li et al., 2019).

Reference: Li, J., Nagashima, T., Kong, L., Ge, B., Yamaji, K., Fu, J. S., Wang, X., Fan, Q., Itahashi, S., Lee, H.-J., Kim, C.-H., Lin, C.-Y., Zhang, M., Tao, Z., Kajino, M., Liao, H., Li, M., Woo, J.-H., Kurokawa, J., Wang, Z., Wu, Q., Akimoto, H., Carmichael, G. R., and Wang, Z.: Model evaluation and intercomparison of surface-level ozone and relevant species in East Asia

in the context of MICS-Asia Phase III – Part 1: Overview, Atmos. Chem. Phys., 19, 12993–13015, https://doi.org/10.5194/acp-19-12993-2019, 2019.

3Q11: Lines 273-276: these differences represent averaged values over the domain? It is difficult to connect these numbers with Fig S11. What do you mean by 2-6 ug m^{-3} for NO_3? Do these numbers represent the range of differences? From Fig, S11 the range seems to be between -4 and 6 ug m^{-3}.

Response: The numbers represent the ranges of values over the continental regions. For NO_3^-, the range is -4 to 6 μg m^{-3}.

This sentence has been removed from the manuscript.

3Q12: Lines 277-279: have you compared these values with observations?

Response: Lines 277-279 explain the differences in PM components around the edge of the simulation domain between the M1 and M2 is caused by using different boundary condition as model input. M1 used downscale results from GEOS-Chem and thus considered the impacts of emission and pollutions from outside of the research domain. M2 used the default values in the CMAQ model as a boundary condition, which is much smaller than those from GEOS-Chem. This result is not comparable to observations.

This sentence has been removed from the manuscript.

3Q13: Lines 297-300: Why? A different aerosol treatment?

Response: Yes, there is an update in the formation pathway of SO_4^{2-} particles in CMAQv502. We added the following sentences in the manuscript for explanation:

Lines 229-234: "The X-CMAQ models (including WRF-CMAQ and RAMS-CMAQ) produce similar *SOR* patterns, except that the CMAQv5.0.2 models (M1 and M2) predict 10% higher

SOR in the HBT region than the CMAQv4.7.1 models (M4, M5, and M6). CMAQv502 updated the production of SO_4^{2-} in the aqueous reaction of the older version (Appel et al., 2013). The explicit treatment of Fe and Mn allows more consistent treatment of aqueous reaction from SO_2 to SO_4^{2-}."

3Q14: Figs. 2 and S13 and subsequent discussion: Could the authors provide an explanation for the really low NO_3 concentration modelled with M8? Have they considered the M8 model when the ensemble mean was compared against observation? Using such outliers, the MMM mean can even be deteriorated compared to individual simulations.

Response: The reasons have been discussed on page 11920 of our companion paper (Chen et al., 2019). One possible reason is the low production of NH_3 by M8 (Fig S4 of Chen et al., 2019). Another reason is not including the N_2O_5 heterogeneous reaction in M8, which is an important pathway for the formation of NO_3^- (Chen et al., 2019).

Yes, M8 is included in the ensemble mean. We agree if the MMM is used to show the model performance, the NO_3^- of M8 should be taken out. However, this paper aims at identifying the differences among models. The MMM is used as a reference to evaluate the differences between models. Thus, we kept including M8 in the MMM.

3Q15: Fig. 2 c) M3 should be M4

Response: We have modified the figure.

3Q16: Lines 323-330: if the S emissions were insufficient, why the M13 and M14 models provide reasonable results?

Response: Our current analysis can only demonstrate that the SOR values of models are generally higher than observations. Since model inputs are not identical in different models due to the reason we mentioned in our answer to Q9, we can't confirm if the underestimation of SO_2

is caused by insufficient S emissions. For this reason, we removed the related discussion from the manuscript.

3Q17: Section 3.4: same as before, why not using the online calculation of natural emissions of dust particles within WRF-Chem?

Response: We have confirmed that the two WRF-Chem models do not include dust emissions during the simulation.

3Q18: Line 374: "perfectly" is a strong word

Response: We have changed the word to "reasonably".

3Q19: Lines 432-433: Table 4 and the associated discussion. Please be consistent when calculating the total wet and dry deposition of S/N and doing subsequent analysis.

Response: We have modified the manuscript and checked the calculation to be consistent and clear.

3Q20: Lines 460-473: In the case of wet deposition of S and N, M11 shows a very different spatial pattern compared with the other models. What is the reason for this? It would be nice if the authors will try to analyse in detail the causes of these differences.

Response: We added the explanation in the manuscript as follows:

Lines 403-405: "Both λ_{wet} and V_d of M11 are much lower than the other models, especially over eastern EA. And this is a possible reason for the biased performance of M11 on wet deposition (Fig. 7)"

Why models perform differently on particulate matter over East Asia? – A multi-model intercomparison study for MICS-Asia III

Jiani Tan[1], Joshua S. Fu[1], Gregory R. Carmichael[2], Syuichi Itahashi[3], Zhining Tao[4], Kan Huang[1,5], Xinyi Dong[1], Kazuyo Yamaji[6], Tatsuya Nagashima[7], Xuemei Wang[8], Yiming Liu[8], Hyo-Jung Lee[9], Chuan-Yao Lin[10], Baozhu Ge[11], Mizuo Kajino[12], Jia Zhu[11], Meigen Zhang[11], Hong Liao[13] and Zifa Wang[11]

[1] Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, TN, 37996, USA
[2] Center for Global and Regional Environmental Research, University of Iowa, Iowa City, IA, 52242, USA
[3] Central Research Institute of Electric Power Industry, Abiko, Chiba, 270-1194, Japan
[4] Universities Space Research Association, Columbia, MD, 21046, USA
[5] Department of Environmental Science and Engineering, Fudan University, Shanghai, 200433, China
[6] Graduate School of Maritime Sciences, Kobe University, Kobe, Hyogo, 658-0022, Japan
[7] National Institute for Environmental Studies, Tsukuba, Ibaraki, 305-8506, Japan
[8] Institute for Environment and Climate Research, Jinan University, Guangzhou, 511443, China
[9] Department of Atmospheric Sciences, Pusan National University, Busan, 609-735, South Korea
[10] Research Center for Environmental Changes Academia Sinica, 11529, Taiwan
[11] Institute of Atmospheric Physics, Chinese Academy of Science, 100029, China
[12] Meteorological Research Institute, Japan Meteorological Agency, 305-0052, Japan
[13] School of Environmental, Science and Engineering, Nanjing University of Information Science & Technology, Nanjing, 210044, China

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

Abstract. This study compares the performances of twelve regional chemical transport models (CTM) from the third phase of Model Inter-Comparison Study for Asia (MICS-Asia III) on simulating the particulate matter (PM) over East Asia (EA) in 2010. The participating models include WRF-CMAQ (v4.7.1 and v5.0.2), WRF-Chem (v3.6.1 and v3.7.1), GEOS-Chem, NHM-Chem, NAQPMS and NU-WRF. This study investigates three model processes as the possible reasons for different model performances on PM: (1) Models perform very differently in the gas-particle conversion of sulphur (S) and oxidized nitrogen (N). The model differences in sulphur oxidation ratio (50%) is of the same magnitude as that in SO_4^{2-} concentrations. The gas-particle conversion is one the main reasons for different model performances on fine mode PM. (2) Models without dust emissions/modules can perform well on PM_{10} at non-dust-affected sites, but largely underestimate (upmost 50%) the PM_{10} concentrations at dust sites. The implementation of dust emissions/modules in models has largely improved the model accuracies at dust sites (reduce model bias to -20%). However, both the magnitudes and distributions of dust pollution are not fully captured. (3) The amounts of modelled depositions vary among models by 75%, 39%, 21%

and 38% for S wet, S dry, N wet and N dry depositions, respectively. Large inter-model differences are found in the washout ratios of wet deposition (at most 170% in India) and dry deposition velocities (general 0.3-2 cm s^{-1} differences over inland regions). This study investigates the reasons for different model performances on PM over EA and offers suggestions for future model development.

1 Introduction

Atmospheric pollution due to particulate matter (PM) has raised world-wide attention for its relationship with environmental and public health issues (Fuzzi et al., 2015;Nel, 2005). Fine particles ($PM_{2.5}$) are associated with cardiovascular and respiratory related cancer and premature deaths (Hoek and Raaschou-Nielsen, 2014;Knol et al., 2009). Outdoor $PM_{2.5}$ pollution is estimated to cause 2.1-5.2 million premature deaths worldwide annually (Lelieveld et al., 2015;Rao et al., 2012;Silva et al., 2013). It accounts for eight percent of global mortality in 2015 and ranks fifth in the global mortality risk (Cohen et al., 2017). East Asia (EA) has been suffering from severe PM pollution due to anthropogenic emissions and natural dust emissions (Akimoto, 2003). China and India are the top two countries suffering from outdoor air pollution, which altogether account for 20% of global mortalities caused by $PM_{2.5}$ exposure in 2010 (Lelieveld et al., 2015). The mixing of dust with anthropogenic pollutants can even enlarge the effects of pollution (Li et al., 2012). However, the impact evaluation of PM pollution is of high uncertainty due to unclearness in the toxicity of PM components (Lippmann, 2014) and difficulty in the measurement and prediction of PM concentrations.

For a better understanding of PM pollution, modelling approach has been adopted to study the spatial distributions of PM with the aid of measurements. Multi-model ensemble approach, which interprets modelling results with combined information from several models, has been proven to increase the reliability of model accuracy (Tebaldi and Knutti, 2007). This method has been widely used for studies in Europe (Bessagnet et al., 2016;Vivanco et al., 2017) and at global scales (Lamarque et al., 2013;Galmarini et al., 2017) on air quality issues. The Model Inter-Comparison Study Asia Phase (MICS-Asia) aims at understanding the air quality issues over EA. The first phase of MICS-Asia (MICS-Asia I) was carried out in the 1990s with eight regional chemical transport models (CTMs). The study focused on air pollution issues related to sulphur (S) (including SO_2, SO_4^{2-} and wet SO_4^{2-} deposition). The second phase of MICS-Asia (MICS-Asia

II) was launched in early 2000s with nine CTMs (Carmichael and Ueda, 2008). The study covered the chemistry and transport of S, nitrogen (N), PM and acid deposition. Multi-model results on SO_4^{2-}, NO_3^- and NH_4^+ (SNA) were evaluated with measurements from fourteen sites of Acid Deposition Monitoring Network in East Asia (EANET) and the Fukue site in Japan. However, a non-exhaustive evaluation on PM_{10} concentrations in China with scarce datasets left an unclear view of models' ability in this area, a region recognized as one of the most heavily polluted in EA. Meanwhile, model results were found with high inconsistencies on simulating both gas and aerosol phases of S and N (Hayami et al., 2008). Further efforts are needed to investigate the reasons for model differences to improve model accuracies.

This study compares the performances of twelve regional models participated in the third phase of MICS-Asia (MICS-Asia III) on simulating PM over EA. The comparison among models aims at identifying the reasons for different model performances. The models involved in this study include Weather Research and Forecasting Model (WRF) coupled with Community Multiscale Air Quality Modeling (CMAQ) (version 4.7.1 and v5.0.2), WRF model coupled with Chemistry (WRF-Chem) (v3.6.1 and v3.7.1), Goddard Earth Observing System coupled with Chemistry (GEOS-Chem), Non-Hydrostatic Model coupled with Chemistry (NHM-Chem), Nested Air Quality Prediction Modeling System (NAQPMS) and NASA-Unified WRF (NU-WRF). The model performance on simulating PM has been reported in a companion paper (Chen et al., 2019). The main findings are described in sect. 3.1. Sections 3.2-3.4 examine the influences of three model processes on model performances: (1) Formation of fine particles (PMF): model differences in the gas-particle conversion. (2) Formation of coarse particles (PMC): model improvements by implementing dust emissions/modules on simulating PM and the remaining problems. (3) Removal processes of particles from the atmosphere: uncertainties lay on the efficiencies of wet and dry depositions. Section 4 concludes the findings of this study and provides suggestion for further study.

2 Methodology

2.1 Framework of MICS-Asia

MICS-Asia is a model intercomparison study with contributions from international modelling groups to simulate the air quality and deposition over EA. MICS-Asia I focused on air quality issues related to S. The multi-model performances on simulating SO_2 and SO_4^{2-} concentrations and

SO_4^{2-} wet deposition were evaluated with observation from eighteen stations (Carmichael et al., 2002). A source-receptor relationship of S deposition was developed based on the sensitivity simulations for seven prescribed receptor regions: Komae, Oki, Fukue, Yangyang, Beijing, Nanjing and Taichung (Carmichael et al., 2002).

MICS-Asia II was initiated in 2003. Nine regional models simulated the air qualities for four months (March, July and December of 2001 and March of 2002) to study the chemistry and transport of air pollutants and acid deposition (Carmichael and Ueda, 2008). All modelling groups were enforced to use the same emission: the Transport and chemical Evolution over the Pacific (TRACE-P) emission of 2000, and common IC/BC to facilitate a comparison on the physical and chemical mechanisms of models. The modelling species expanded to S, N, O_3, PM and acid deposition. Model evaluations and major findings can be found in literature (Carmichael et al., 2008;Fu et al., 2008;Han et al., 2008;Hayami et al., 2008).

MICS-Asia III is launched in 2010. The simulation time covers the whole year of 2010. All modelling groups are required to use the prescribed anthropogenic and natural inputs (Li et al., 2017). Three purposes are set for this project– topic I: evaluating the strengths and weaknesses of current multi-scale air quality models in simulating air qualities over EA and providing suggestion to reduce uncertainty for future simulations, topic II: developing a reliable anthropogenic emission inventory for EA, topic III: investigating the interaction of aerosol-weather-climate by using online coupled air quality models. This study focuses on topic I.

2.2 Model configurations

The model set-up can be found in Table 1 of Chen et al. (2019). Fourteen modeling groups (M1-M14) participated, but M3 and M9 are not included in this study due to uncompleted model submission. M14 model has a smaller simulation domain than the others, therefore it is not included in the multi-model mean (MMM) results. The gas and aerosol modules and dust schemes employed by the participating models were introduced in detail in sector 2.1 of Chen et al., 2019. Following are the descriptions on the model set-up for wet and dry deposition.

Wet deposition removes gases and aerosols from the atmosphere by rain droplets, involving both in-cloud scavenging (rainout) and below-cloud scavenging (washout). The gases in the

atmosphere are dissolved in the raindrop and then removed from the atmosphere. For the non-reactive gases, the removal rate depends on the solubility of gases and is a function of the Henry's Law. Particles participate in the cloud condensation nuclei in the presence of supersaturation water vapor and then grow into cloud droplets. In this study, only M2, M4, M6, M11 and M12 have submitted the main components of S and N depositions. All these models use the same wet deposition scheme based on Henry's law. The efficiency of wet deposition is assessed by the so-called "washout ratio", calculated as the ratio of particle concentrations in deposition to particle concentrations in surface air as shown in Eq. 1.

$$\lambda_{wet} = \frac{C_{depo}}{C_{surface\_air}} \times 100\% \quad (1)$$

where λ_{wet} is the washout ratio for wet deposition, C_{depo} is the concentration of particles in deposition and $C_{surface\_air}$ is the concentration of particles at near surface atmosphere.

Dry deposition is mainly driven by turbulent and molecular diffusion processes. All models except M12 use the same dry deposition scheme from Wesely (1989). The dry deposition flux is proportional to the concentration of pollutants at height. The dry deposition velocity is calculated with Eq. 2.

$$V_d = -F_c \,/\, C_a \quad (2)$$

$$V_d = \frac{1.0}{r_{surf} + r_a + r_{bc}} \quad (3)$$

where F_c is the dry deposition flux, V_d is the deposition velocity and C_a is the concentration of species at height. The negative mark indicates the direction of the dry deposition velocity. The V_d is determined by the resistance of air layer (r). The total r is composed of three factors (Eq. 3): the aerodynamic resistance (r_a), boundary layer resistance (r_{bc}) and canopy resistance (r_{surf}).

M12 uses the general approach from Wesely (1989) and updates by Zhang et al. (2003). Zhang et al. (2003) updates the value of non-stomatal resistance (R_{ns}), which is a component of R_{surf} related to the soil uptake and cuticle uptake of dry deposition. Model evaluation shows the updates can improve the model prediction on dry deposition velocities of SO_2 (Zhang et al., 2003).

2.3 Observation data

To make the discussion clear, we define the regions used in the following analysis here: northern EA (Russia and Mongolia), central EA (China), western EA (Japan and Korea) and southern EA (Cambodia, Lao PDR, Myanmar, Thailand, Vietnam, Indonesia, Malaysia and Philippines). Following monitoring datasets are used in the analysis in sects. 3.2-3.4: Air Pollution Indices (APIs) provides monthly average PM_{10} data from eighty-six sites (A1-A86 in Fig 1) (http://datacenter.mep.gov.cn/). This dataset has been widely used to study the PM pollution (Qu et al., 2010;Chen et al., 2008;Deng et al., 2011) as well as model evaluation (Wang et al., 2012;Xing et al., 2015) in China. It is replaced by the Air Quality Index (AQI) after 2013. The APIs data covers the eastern China well with intensively located sites, but sites in western China are very limited. EANET (E1-E54) provides monthly average concentrations of PM_{10}, SNA and S and N depositions from fifty-four sites (http://www.eanet.asia/, last access: 28 May 2018). For PM_{10}, this dataset has very limited number of sites in China. The sites are generally located along the east coast of China and couldn't well cover the areas with high PM_{10} pollution, such as the Hebei-Beijing-Tianjin (HBT) region (Fig. 1). And the data completeness in northern EA is not as satisfied as the other regions. Only three sites located in Rishiri (E15), Ochiishi (E16) and Oki (E21) in Japan have $PM_{2.5}$ observation during our study period. R1-R35 (green) are thirty-five Reference (Ref) sites provided by the Institute of Atmospheric Physics Chinese Academy of Science (IAP_CAS). The sites are intensively located in three regions: HBT region, Pearl River Delta (PRD) and Taiwan.

3 Result and discussion

3.1 Brief results of model performance evaluation

All models have submitted the monthly average concentrations of PM_{10}, $PM_{2.5}$ and SNA at surface layer except PM_{10} from M13 and NO_3^- and NH_4^+ from M10. Evaluation of model performance on aerosols can be found in our companion paper (Chen et al., 2019). Following are the main findings: PM_{10} concentrations were generally underestimated over the simulation domain. $PM_{2.5}$ concentrations were also underestimated over Eastern EA, but were well simulated in Central EA. Models failed to reproduce the high peaks of SO_4^{2-} concentration in Central EA, probably due to

missing SO_4^{2-} formation mechanisms (such as heterogeneous SO_4^{2-} chemistry), which has been reported as an important formation pathway of SO_4^{2-} in China. NO_3^- concentrations were overpredicted by most models over the simulation domain and were associated with the underestimation of SO_4^{2-}. M7 and M8 models produced significantly lower NO_3^- concentrations than observations and other models, due to model bias in simulating the NH_3 concentrations and missing the N_2O_5 heterogeneous reaction that sever as an import formation pathway of NO_3^-. The spatial distributions of AOD were generally well simulated except the underestimation around the Himalaya mountains, Taklamakan Desert and Gobi Desert.

This study compares the model performances with global-scale model study. The Task Force on Hemispheric Transport of Air Pollution (TF HTAP) is an inter-comparison study of global and regional models to assess the impact of hemispheric transport of air pollutants on regional atmosphere. The second phase of HTAP (HTAP-II) involved more than twenty global models to simulate the air quality in 2010 (Galmarini et al., 2017). Most models utilize coarse-resolution grids at about 2°-3°. The HTAP-II and MICS-Asia III share some common points like using the same emission inventory in East Asia (Li et al., 2017) and using the same observation dataset to evaluate PM_{10} (more than 100 EANET and API sites) and $PM_{2.5}$ (two EANET sites) (Dong et al., 2018). The mean bias (MB) of PM_{10} over EA is -30.7 μg m^{-3} and -18.6 μg m^{-3} for HTAP-II and this study, respectively (values for sites used by both studies). And the MB of $PM_{2.5}$ is -1.6 μg m^{-3} and -4.3 μg m^{-3} for HTAP-II and this study, respectively. Both studies find underestimation of PM_{10} concentrations, while $PM_{2.5}$ concentrations are well produced. Models of MICS-Asia III perform slightly better than those of HTAP-II with lower model bias in PM_{10}, probably taking the advantage of finer resolutions of model grids.

The so-called "diagnostic evaluation" approach is adopted to check the model bias oriented by individual process (Dennis et al., 2010). Although all modelling group are required to use the prescribed emission inventory, but mismatch was found during the temporal and vertical treatments of emission files by different modelling group and has caused differences in the model inputs (Itahashi et al., 2019). To avoid the possible impacts on inter-model comparison, we compare the indicators (i.e. sulphur oxidation ratio (SOR)) instead of direct model outputs (i.e. SO_4^{2-} concentrations) to focus on the differences caused by model mechanisms. The following three processes are examined:

(1) Formation of PMF: sect. 3.2 investigates the differences in the gas-particle conversion of S and N among different models.

(2) Formation of PMC: sect. 3.3 assesses the model abilities in reproducing the spatial and temporal distributions of PM in regions affected by dust storm. A comparison is conducted between models with and without dust emissions/modules.

(3) Removal of particles from the atmosphere: sect. 3.4 compares the model performances in simulating the amounts of deposition and the efficiencies of wet and dry depositions.

3.2 Gas-particle conversion

The following two indicators are calculated to illustrate the gas-particle conversions of S and N.

$$SOR = \frac{n\text{-}SO_4^{2-}}{n\text{-}SO_4^{2-} + n\text{-}SO_2} \quad (4)$$

$$C(NO_2) = \frac{n\text{-}NO_3^{-}}{n\text{-}NO_3^{-} + n\text{-}NO_2} \quad (5)$$

where $n\text{-}SO_4^{2-}$, $n\text{-}SO_2$, $n\text{-}NO_3^{-}$ and $n\text{-}NO_2$ are the mole concentrations of SO_4^{2-} particle, SO_2 gas, NO_3^{-} particle and NO_2 gas. The $C(NO_2)$ indicator only has NO_3^{-} and NO_2 in the denominator due to the limitation of observation data. But it still can portrait the conversion of N between gas phase and particle phase.

Figures 2 and 3 show the distributions of *SOR* and $C(NO_2)$ values of models. The *SOR* values are lowest around the HBT region in north-eastern China (10-40%) and highest in south-western China (60-80%) (Fig. 2). The X-CMAQ models (including WRF-CMAQ and RAMS-CMAQ) produce similar *SOR* patterns, except that the CMAQv5.0.2 models (M1 and M2) predict 10% higher *SOR* in the HBT region than the CMAQv4.7.1 models (M4, M5 and M6). CMAQv502 updated the production of SO_4^{2-} in the aqueous reaction of the older version (Appel et al., 2013; Fountoukis and Nenes, 2007). The explicit treatment of Fe and Mn allows more consistent treatment of aqueous reaction from SO_2 to SO_4^{2-}. For the X-Chem models (including WRF-Chem, GEOS-Chem and NHM-Chem), the two WRF-Chem models (M7 and M8) produce similar magnitudes and distributions of *SOR* in all regions, except the south-western China (around Tibet in Fig. 1) and the open oceans, while the NHM-Chem (M12) and GEOS-Chem (M13) models produce slightly higher *SOR* values over the whole simulation domain. The differences between

239 the X-CMAQ and the X-Chem models are significant over the inland regions of northern and
240 eastern China, Japan and southern EA. The X-CMAQ models generally predict 5-20% higher *SOR*
241 values than the X-Chem models. Similarly, the X-CMAQ models generally give 20% higher
242 $C(NO_2)$ values than the WRF-Chem models, especially in eastern EA (Fig. 3). The $C(NO_2)$ of M8
243 is extremely low due to unreasonably low NO_3^- concentrations.

244 Figure 4 shows the gas-particle conversions of S and N by models and observation at the
245 EANET sites. The red bars represent concentrations of gases and the black bars represent
246 concentrations of aerosols. The values with blue color above the bars are observed and modelled
247 *SOR* and $C(NO_2)$ values. Results for individual sites are available in supplementary Fig. S1.
248 According to Fig. 4(a), the total amount of S (SO_2 gas+SO_4^{2-} particle) is about 0.15 μmole(S) m^{-3}.
249 Most models have biases on this value, especially the moderate underestimation by M7, M8 and
250 M13. On the other hand, the *SOR* value (0.25) is well simulated by M1 (0.26), M2 (0.20), M10
251 (0.29) and M13 (0.26). Other models generally under-predict the *SOR* value except M12 (0.33)
252 and M14 (0.57). The WRF-CMAQv5.0.2 models (M1 and M2) produce higher *SOR* than WRF-
253 CMAQv4.7.1 models (M4, M5 and M6), probably attributed to the updates in the formation
254 pathway of SO_4^{2-}.

255 Figure 4(b-e) show the results in different regions. In northern EA, the total amount of S is
256 underestimated by all models except M13 and M14. However, the *SOR* value (0.12) is well
257 reproduced by most models (0.08-0.20) except M12 (0.25) and M10 (0.32). There is only one site
258 available for central EA. Most models (except M12 and M13) have largely underestimated the
259 *SOR* value, while M14 has largely overestimated it. For eastern EA, the total amount of S is well
260 captured by all models except M11, M12 and M14. The *SOR* value (0.55) is generally
261 underestimated by all models except M10 (0.55) and M14 (0.71). For southern EA, the total
262 amount of S is generally overestimated by all models except M13, while the *SOR* value is
263 underestimated by all models except M13 and M14. Overall, the models have both positive and
264 negative biases in simulating the total amounts of S, but generally underestimated the *SOR* values
265 in all regions. Furthermore, the modelled *SOR* values vary largely among models (ranging from
266 0.12 to 0.57), resulting in a large inter-model difference (1sd% = 50%). This variation is of the
267 same magnitude as the variation of SO_4^{2-} concentration (1sd% = 50% in supplementary Fig. S2).

The results suggest that differences in gas-particle conversion among models could account largely for the models' inconsistency in simulating the SO_4^{2-} concentrations.

Figure 4(f-h) compares the gas-particle conversion of N with the *C(NO_2)* indicator. Only one site in China and one site in Japan have both NO_2 and NO_3^- observations. At the Hongwen sites in China, all models except M5 underestimate the sum of NO_2 and NO_3^-, but the modelled *C(NO_2)* values are close to the observation (0.18) except M5 (0.07), M8 (0.00) and M12 (0.40). Similar to the results of S conversion, the newer version of WRF-CMAQ model generally produces higher *C(NO_2)* than the older version, but the differences between the two are smaller. At the Banryu site in Japan, the sum of NO_2 and NO_3^- is well simulated by all models except M8. The *C(NO_2)* (0.19) value is also well simulated by all models except M8 (0.00), M12 (0.53) and M14 (0.77). Overall, the model accuracy on *C(NO_2)* is slightly higher than that on *SOR* according to the comparison with observed values. Models also have higher consistencies on *C(NO_2)* than *SOR*. However, further validation is required due to the limited number of observations for the conversion of N.

3.3 Implementation of dust emissions/modules in models

The PMC concentrations at surface layer are calculated by subtracting $PM_{2.5}$ from PM_{10}. Figure 5 shows the spatial distribution of annual average PMC of models. Most models show very low (< 2μg m^{-3}) concentrations of PMC around the Takalmakan Desert and the Gobi Desert in northern China except M10, M11 and M14. These three models use dust emissions/modules in simulations (Chen et al., 2019). M12 also includes dust emissions, but its PM_{10} concentrations over northern China are much lower than the three models. The predicted PMC concentrations of the three models differ largely. The domain-average concentrations of PMC are 21, 7 and 12 μg m^{-3} for M10, M11 and M14, respectively. The distributions of PMC also differ largely over north-west China, where the impacts of dust are most significant. The differences among the models mainly comes from the different parameterizations such as source functions, dust-lifting mechanisms and size distributions of particles (Chen et al., 2019). Different PMC concentrations are also found over oceans, mainly attributed to the sea-salt emissions in this study. The sea-salt emissions are parameterized in the models with various formula (Chen et al., 2019). In this study, the WRF-Chem models (M7 and M8) do not account for sea-salt emissions, thus their PMC concentrations

over the oceans and seas are not defined. The two WRF-CMAQ models use the in-line sea-salt emission module of Gong (2003) and updated by Kelly et al. (2010). They predict consistent distributions of PMC over oceans. M10 and M11 use the same module as the CMAQ models (Gong, 2003), but produce higher PMC on oceans. M12 adopts the method of breaking wave over seashore by Clarke et al. (2006) and produces the highest PMC over oceans among all models.

The implementation of dust emission is expected to improve the model performances, but how significant could the improvement be? And can models predict the PM concentrations reasonably at regions affected by dust with current dust emissions/modules? To answer these questions, all sites are grouped to dust and non-dust sites according to their locations. The sites located in regions that have been reported to receive severe impacts and rapid deposition of dust are marked as dust sites (Shao and Dong, 2006) (grey-color shaded areas in Fig. 1). Figure 6(a-b) and Table 1 compare the model performances at the dust and non-dust sites. For the non-dust sites (Fig. 6(b)), most models have well captured the magnitudes of PM_{10} at the "API non-coastal, non-dust" sites (MB = -8% and NMB = -8%). The sites marked as "API coastal" sites, which are located close to the coastal regions, are all slightly underestimated by about 25 μg m^{-3} (30%). Similarly, the PRD and Taiwan sites, which are also located near the coastal regions, are all underestimated by about 20 μg m^{-3} (37%). Bias in sea-salt emissions is the possible reason. Sea-salt emission is reported to contribute to 20-40% of SNA and PM_{10} over coastal regions (Liu et al., 2015). Including the sea-salt emission in model simulation can improve the model accuracy with 8-20% increase in PM_{10}, SNA, Na^+ and Cl^- (Kelly et al., 2010;Im, 2013). The influence of sea-salt emission is not the focus of this study, but further study is strongly recommended.

For the dust sites (Fig. 6(a)), most models have generally underestimated the PM_{10} concentrations by 10-40 μg m^{-3} (15-50%). And the three models with dust module perform better than the others at the dust sites, especially site A2, A30, A68, A69, R5 and R18. However, they miss the high PM_{10} concentrations at sites like R1-R3 and R11, and overestimate the PM_{10} concentrations at sites such as A60 and A80. This indicates that the dust emissions/modules involved in this study can't fully capture the magnitudes and distributions of dust pollution over EA. In addition, the modelled PMC differ a lot with different dust emissions/modules (Fig. 5). M10 model produces very high PMC over the whole eastern China, while M11 model only predicts high PMC around the HBT region. Overall, the model performance on PM over dust regions can

be improved largely by including dust emissions/modules. However, the concentrations and distributions are not yet well captured and large inconsistencies are found among different dust emissions/modules.

Figure 6(c-d) compares the modelled monthly trends of PM_{10} with observations at the dust and non-dust sites and Fig. 6(e) shows the correlations (R) values between models and observation. For the non-dust sites (Fig. 6(d)), the trends are well caught by most models. The R values are close to 0.70 for all models except M7 (0.62), M8 (0.58) and M14 (0.63). The WRF-Chem models (M7 and M8) simulate too low PM_{10} concentrations in winter. M14 model overestimates the PM_{10} concentrations during March to May. Most models have much lower R values at the dust sites than the non-dust sites (Fig. 6(e)), due to underestimation of the PM_{10} concentrations during winter. For instance, R values of M10 drop from 0.7 at the non-dust sites to 0.11 at the dust sites. Spring (March, April and May) has the largest model biases at the dust sites, which is coincident with the dust storm season in Asia (Arimoto et al., 2006). M10 and M14 models perform well in most months at both the dust and non-dust sites, taking the advantage of their dust emissions/modules. But their R values at the dust sites are very low. Future study is strongly suggested on a better understanding of the seasonal variations of dust pollution.

3.4 Wet and dry depositions

Figure 7 and Table 2 show the model performance on wet deposition. For wet SO_4^{2-} deposition, despite that the two sites with highest deposition (E2 and E3) in China are underestimated, the other sites are generally well simulated by MMM with a low MB of -8%. The individual model bias varies from -22% to 41%. The CMAQ models (M2, M4 and M6) all underestimate the wet SO_4^{2-} deposition. There are large differences between CMAQv4.7.1 and CMAQv5.0.2 in JP, where the CMAQv4.7.1 models (M4 and M6) slightly overestimate the wet SO_4^{2-} deposition at E19 and E23, while the CMAQv5.0.2 model (M2) slightly underestimates the value at these sites. The M11 model produces considerably higher wet deposition of SO_4^{2-} and NO_3^- than the other models in East EA. The possible reasons are discussed later. The MMM underestimates the NO_3^- wet deposition by 29%, due to large under-prediction in southern EA. The southern EA has several sites with very high deposition, such as E29 site in MY and E35 and E36 sites in PH, but all models fail to catch those high peaks. The individual model bias varies from -59% to 30% among models. M2, M4, M6 and M12 perform similarly with high underestimation ranging from 39% to 59%.

The M11 is the only model that succeed to capture the high wet NO_3^- deposition at E2 and E3 in CH, but it overestimates most sites in CH, JP and KR. In case of wet NH_4^+ deposition, the MMM generally underestimates the amount at all sites with a bias of -40%, especially at E2-E4 in CH, E45 in TH and E35 and E36 in PH. The individual model bias varies from -10% to -37%. The M2, M4 and M6 models perform similarly, while M11 and M12 models predict higher depositions at all sites. Overall, large inter-model disagreements are found in eastern EA for wet deposition of SO_4^{2-} and NO_3^- and in southern EA for the wet NH_4^+ deposition. The observation of dry deposition is composed by observed concentration of air pollutants and simulated deposition velocity. Since the EANET network only provides the former one, complete evaluation of the dry deposition is not available in this study (complete dry deposition with velocity is available after 2013).

Table 3 lists the domain-total annual-accumulated amounts of S and N depositions by models. The total wet S deposition (D_{Swet}) includes wet depositions of SO_2, H_2SO_4 and SO_4^{2-}. The total dry S deposition (D_{Sdry}) includes dry deposition of SO_2, H_2SO_4 and SO_4^{2-}. The total wet N deposition (D_{Nwet}) includes wet depositions of NO_3^-, NH_4^+, HNO_3, NH_3. The total dry N depositions (D_{Nwet}) includes dry deposition of NO, NO_2, NO_3^-, NH_4^+, HNO_3 and NH_3. D_{Swet} values range from 10.5 to 31.3 Tg(S) yr^{-1} among models (1sd%=75%). The estimation by M11 model is two folds higher than the other four models. The inter-model difference is significant even among the same type of models with different versions. The CMAQv4.7.1 models (M4 and M6) produce 12.5 Tg(S) yr^{-1} (M4) and 13.8 Tg(S) yr^{-1} (M6) of D_{Swet}, while the prediction by CMAQv5.0.2 model (M2) is 25% lower. Despite the large discrepancies in the total amount, all five models agree that over 95% of D_{Swet} is wet SO_4^{2-} deposition. The total amounts of D_{Sdry} range from 4.3 to 10.6 Tg(S) yr^{-1} among models (1sd% =39%). M11 predicts higher D_{Sdry} than other models and the CMAQv5.0.2 model (M2) predicts 45% lower D_{Sdry} than the two CMAQv4.7.1 models (M4 and M6). Similar to D_{Swet}, all models have high agreements on the proportions of the components. D_{Nwet} range from 12.2 to 20.0 Tg(N) yr^{-1} among models (1sd%=21%). The CMAQ models (M2, M4 and M6) simulate close results (12-15 Tg(N) yr^{-1}), while M11 (20.0 Tg(N) yr^{-1}) and M12 (16.5 Tg(N) yr^{-1}) simulate slightly higher amounts. As for the proportion of components, M2, M4, M6 and M12 models predict high proportions of wet NO_3^- and wet NH_4^+ depositions (particle phase), while M11 model produces higher percentages of wet HNO_3 and wet NH_3 depositions (gas phase). D_{Ndry} range from 3.9 to 14.1 Tg(N) yr^{-1} (1sd%=38%). M12 gives a considerably lower amount than the other models. Models are quite consistent on the proportions of components.

Figure 8(a-e) show λ_{wet} of S deposition (λs_{wet}) by models. The CMAQ models (M2, M4 and M6) have similar patterns in λs_{wet} over the inland regions, while M12 model predicts 30-90% lower ratios in India. M11 model generally predicts about 20-70% lower λs_{wet} than the other four models except India, where the difference could reach upmost 170%. For λ_{wet} of N deposition (λ_{Nwet}) (Fig. 8(f-j)), the CMAQv4.7.1 models (M4 and M6) and M12 perform similarly, but the CMAQv5.0.2 model (M2) predicts 30% lower λ_{Nwet} in India, Japan and Korea. M11 generally predicts lower ratios in India (60% lower), Indonesia and Philippines (120% lower) than the CMAQ models. Figure 9 shows the spatial distributions of V_d. For V_d of S deposition (V_{Sd}) (Fig. 9(a-e)), the CMAQ models (M2, M4 and M6) simulate very similar spatial distributions. M11 and M12 models predict 0.5 cm s^{-1} lower V_{Sd} than the CMAQ models over the whole inland regions, especially in east China and India peninsular. For V_d of N deposition (V_{Nd}) (Fig. 9(f-j)), the CMAQ models (M2, M4 and M6) predict very similar distributions. M11 and M12 predict about 0.3 cm s^{-1} and 1-2 cm s^{-1} lower V_{Nd} than the CMAQ models over the inland regions. Both λ_{wet} and V_d of M11 are much lower than the other models, especially over eastern EA. And this is a possible reason for the biased performance of M11 on wet deposition (Fig. 7). Overall, large inter-model differences are found in predicting both the amounts of depositions and the efficiencies of depositions.

4 Conclusion

The topic I of the MICS-Asia III aims at (i) evaluating the strengths and weaknesses of current multiscale air quality models in simulating concentration and deposition fields over East Asia and (ii) providing suggestions for future model developments. This study compares the performances of twelve regional models for the prediction of PM concentrations over EA. The participating models includes WRF-CMAQ (v4.7.1 and v5.0.2), WRF-Chem (v3.6.1 and v3.7.1), GEOS-Chem, NHM-Chem, NAQPMS and NU-WRF. Three processes/mechanisms are investigated to identify the causes of inter-model differences:

(1) For the formations of PMF, *SOR* and $C(NO_2)$ values are used to demonstrate the inter-model differences in gas-particle conversions. The *SOR* values are generally underestimated by most models at the EANET sites. A generally trend is found that the WRF-CMAQv5.0.2 models produce the highest *SOR* values among all models, followed by the WRF-CMAQv4.7.1 models (10% lower in HBT region), the WRF-Chem models and other models (5-20% lower over

inland regions). The inter-model variation on *SOR* (1sd% =50%) is of the same magnitude as that on SO_4^{2-} concentration. Similar results are found in $C(NO_2)$, but models have higher agreements on $C(NO_2)$ than *SOR*. The different treatments of gas-particle conversions account largely for the different model performances on PMF.

(2) For the formations of PMC, the models without dust emissions/modules generate very low (<2μg m^{-3}) PMC concentrations. They can well capture the PM_{10} concentrations at non-dust-affected sites but underestimate the PM_{10} concentrations at sites affected by dust storms by upmost 50%. This underestimation is largely improved by implementing dust emissions/modules (bias reduced to around -20%). However, both the magnitudes and distributions of dust pollution are not fully captured. In addition, models employing different dust emissions/modules show large disagreements on the distributions of PMC.

(3) For the removal of PM from the atmosphere, the amounts of atmospheric deposition vary largely among models (1sd%) by 75%, 39%, 21% and 38% for D_{Swet}, D_{Sdry}, D_{Nwet} and D_{Ndry}, respectively. The λ_{wet} and V_d indicators are used to exclude the influences brought by model inputs. For λ_{wet}, models agree more on the D_{Swet} than D_{Nwet}. The largest model inconsistencies are found in India (upmost 170%), Indonesia and Philippines (upmost 120%). For V_d, models differ more on D_{Ndry} than D_{Sdry}, which is opposite to λ_{wet}. The inter-model differences are widely found over the inland regions for D_{Sdry} (about 0.5 cm s^{-1}) and D_{Ndry} (0.3-2 cm s^{-1}).

The main contributions of this study are: (1) comparing the conversions of S and N between gas and particle phases among different models as well as with observations. The comparison with observation makes it possible to both quantify the inter-model differences and tell which module might be more reasonable; (2) Several new updates on dust modules have been published in recent literature, but there is limited study on the inter-comparison. This study provides an opportunity to bring together the new updates on dust modules/emission and review their performance in EA; (3) providing a comprehensive view on the total budget of S and N aerosols, by including the analysis on the removal processes. It turns out that this process brings significant uncertainties to inter-model differences. It should be noted that other factors such as vertical diffusion can also contribute to model differences. Meanwhile, this study focuses on comparing the model abilities in simulating PM in 2010. The chemical regimes may have changed drastically due the rapid changes of emissions and implementation of control policies in Asia. Studies on more recent years and heavily polluted episodes are under preparation.

Author Contributions. JT and JSf designed the study. JT processed and analysed the data. JSF, GRC, SI and ZT contributed to the results and discussions. JSF, ZT, KH, SI, KY, TN, YM, XW, YL, HJL, JEK, CYL, BG, MK, JZ, MZ, LH and ZW provided modelling data. All co-authors provided comments to the manuscript.

Data Availability. The observation data are introduced with details in supplementary sect. S2.1 with web links of public available datasets. The model data are available upon request.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank all participating modeling groups of MICS-Asia III. We acknowledge the support by the Advanced Computing Facility (ACF), Joint Institute for Computational Sciences of the University of Tennessee. We also thank the support by the Oak Ridge National Laboratory for computational resources, which is supported by the Office of Science of the U.S. Department of Energy (contract DE-AC05-00OR22725).

Reference

Akimoto, H.: Global air quality and pollution, Science, 302, 1716-1719, DOI: 10.1126/science.1092666, 2003.

Appel, K. W., Pouliot, G. A., Simon, H., Sarwar, G., Pye, H. O. T., Napelenok, S. L., Akhtar, F., and Roselle, S. J.: Evaluation of dust and trace metal estimates from the Community Multiscale Air Quality (CMAQ) model version 5.0, Geoscientific Model Development, 6, 883-899, 10.5194/gmd-6-883-2013, 2013.

Arimoto, R., Kim, Y. J., Kim, Y. P., Quinn, P. K., Bates, T. S., Anderson, T. L., Gong, S., Uno, I., Chin, M., Huebert, B. J., Clarke, A. D., Shinozuka, Y., Weber, R. J., Anderson, J. R., Guazzotti, S. A., Sullivan, R. C., Sodeman, D. A., Prather, K. A., and Sokolik, I. N.: Characterization of Asian Dust during ACE-Asia, Global Planet Change, 52, 23-56, https://doi.org/10.1016/j.gloplacha.2006.02.013, 2006.

Bessagnet, B., Pirovano, G., Mircea, M., Cuvelier, C., Aulinger, A., Calori, G., Ciarelli, G., Manders, A., Stern, R., Tsyro, S., Vivanco, M. G., Thunis, P., Pay, M. T., Colette, A., Couvidat, F., Meleux, F., Rouil, L., Ung, A., Aksoyoglu, S., Baldasano, J. M., Bieser, J., Briganti, G., Cappelletti, A., D'Isidoro, M., Finardi, S., Kranenburg, R., Silibello, C., Carnevale, C., Aas, W., Dupont, J. C., Fagerli, H., Gonzalez, L., Menut, L., Prevot, A. S. H., Roberts, P., and White, L.: Presentation of the EURODELTA III intercomparison exercise - evaluation of the chemistry transport models' performance on criteria pollutants and joint analysis with meteorology, Atmospheric Chemistry and Physics, 16, 12667-12701, https://doi.org/10.5194/acp-16-12667-2016, 2016.

Carmichael, G. R., Calori, G., Hayami, H., Uno, I., Cho, S. Y., Engardt, M., Kim, S. B., Ichikawa, Y., Ikeda, Y., Woo, J. H., Ueda, H., and Amann, M.: The MICS-Asia study: model intercomparison of long-range transport and sulfur deposition in East Asia, Atmospheric Environment, 36, 175-199, https://doi.org/10.1016/S1352-2310(01)00448-4, 2002.

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., Thongboonchoo, N., and Amann, M.: MICS-Asia II: The model intercomparison study for Asia Phase II methodology and overview of findings, Atmospheric Environment, 42, 3468-3490, http://dx.doi.org/10.1016/j.atmosenv.2007.04.007, 2008.

Carmichael, G. R., and Ueda, H.: MICS-Asia II: The model intercomparison study for Asia phase II, Atmospheric Environment, 42, 3465-3467, http://dx.doi.org/10.1016/j.atmosenv.2007.10.003, 2008.

Chen, L., Gao, Y., Zhang, M., Fu, J. S., Zhu, J., Liao, H., Li, J., Huang, K., Ge, B., Wang, X., Lam, Y. F., Lin, C.-Y., Itahashi, S., Nagashima, T., Kajino, M., Yamaji, K., Wang, Z., and Kurokawa, J.: MICS-Asia III: multi-model comparison and evaluation of aerosol over East Asia, Atmos. Chem. Phys., 19, 11911–11937, https://doi.org/10.5194/acp-19-11911-2019, 2019.

Chen, Z. H., Cheng, S. Y., Li, J. B., Guo, X. R., Wang, W. H., and Chen, D. S.: Relationship between atmospheric pollution processes and synoptic pressure patterns in northern China, Atmospheric Environment, 42, 6078-6087, 10.1016/j.atmosenv.2008.03.043, 2008.

Clarke, A. D., Owens, S. R., and Zhou, J. C.: An ultrafine sea-salt flux from breaking waves: Implications for cloud condensation nuclei in the remote marine atmosphere, J Geophys Res-Atmos, 111, https://doi.org/10.1029/2005JD006565,2006.

Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C. A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C. J. L., and Forouzanfar, M. H.: Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015, Lancet, 389, 1907-1918, 10.1016/s0140-6736(17)30505-6, 2017.

Deng, J. J., Wang, T. J., Jiang, Z. Q., Xie, M., Zhang, R. J., Huang, X. X., and Zhu, J. L.: Characterization of visibility and its affecting factors over Nanjing, China, Atmospheric Research, 101, 681-691, 10.1016/j.atmosres.2011.04.016, 2011.

Dennis, R., Fox, T., Fuentes, M., Gilliland, A., Hanna, S., Hogrefe, C., Irwin, J., Rao, S. T., Scheffe, R., Schere, K., Steyn, D., and Venkatram, A.: A framework for evaluating regional-scale numerical photochemical modeling systems, Environ Fluid Mech, 10, 471-489, doi: 10.1007/s10652-009-9163-2, 2010.

Dong, X. Y., Fu, J. S., Zhu, Q. Z., Sun, J., Tan, J. N., Keating, T., Sekiya, T., Sudo, K., Emmons, L., Tilmes, S., Jonson, J. E., Schulz, M., Bian, H. S., Chin, M., Davila, Y., Henze, D., Takemura, T., Benedictow, A. M. K., and Huang, K.: Long-range transport impacts on surface aerosol concentrations and the contributions to haze events in China: an HTAP2 multi-model study, Atmospheric Chemistry and Physics, 18, 15581-15600, https://doi.org/10.5194/acp-18-15581-2018, 2018.

Fountoukis, C., and Nenes, A.: ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K+-Ca2+-Mg2+-Nh(4)(+)-Na+-SO42--NO3--Cl--H2O aerosols, Atmospheric Chemistry and Physics, 7, 4639-4659, https://doi.org/10.5194/acp-7-4639-2007, 2007.

Fu, J. S., Jang, C. J., Streets, D. G., Li, Z., Kwok, R., Park, R., and Han, Z.: MICS-Asia II: Modeling gaseous pollutants and evaluating an advanced modeling system over East Asia, Atmospheric Environment, 42, 3571-3583, http://dx.doi.org/10.1016/j.atmosenv.2007.07.058, 2008.

Fuzzi, S., Baltensperger, U., Carslaw, K., Decesari, S., van Der Gon, H. D., Facchini, M. C., Fowler, D., Koren, I., Langford, B., Lohmann, U., Nemitz, E., Pandis, S., Riipinen, I., Rudich, Y., Schaap, M., Slowik, J. G., Spracklen, D. V., Vignati, E., Wild, M., Williams, M., and Gilardoni, S.: Particulate matter, air quality and climate: lessons learned and future needs, Atmospheric Chemistry and Physics, 15, 8217-8299, https://doi.org/10.5194/acp-15-8217-2015, 2015.

Galmarini, S., Koffi, B., Solazzo, E., Keating, T., Hogrefe, C., Schulz, M., Benedictow, A., Griesfeller, J. J., Janssens-Maenhout, G., Carmichael, G., Fu, J., and Dentener, F.: Technical note: Coordination and harmonization of the multi-scale, multi-model activities HTAP2, AQMEII3, and MICS-Asia3: simulations, emission inventories, boundary conditions, and model output formats, Atmos. Chem. Phys., 17, 1543-1555, 10.5194/acp-17-1543-2017, 2017.

Gong, S. L.: A parameterization of sea-salt aerosol source function for sub- and super-micron particles, Global Biogeochem Cy, 17, https://doi.org/10.1029/2003GB002079, 2003.

Han, Z., Sakurai, T., Ueda, H., Carmichael, G. R., Streets, D., Hayami, H., Wang, Z., Holloway, T., Engardt, M., Hozumi, Y., Park, S. U., Kajino, M., Sartelet, K., Fung, C., Bennet, C., Thongboonchoo, N., Tang, Y., Chang, A., Matsuda, K., and Amann, M.: MICS-Asia II: Model intercomparison and evaluation of ozone and relevant species, Atmospheric Environment, 42, 3491-3509, http://dx.doi.org/10.1016/j.atmosenv.2007.07.031, 2008.

Hayami, H., Sakurai, T., Han, Z., Ueda, H., Carmichael, G. R., Streets, D., Holloway, T., Wang, Z., Thongboonchoo, N., Engardt, M., Bennet, C., Fung, C., Chang, A., Park, S. U., Kajino, M., Sartelet, K., Matsuda, K., and Amann, M.: MICS-Asia II: Model intercomparison and evaluation of particulate sulfate, nitrate and ammonium, Atmospheric Environment, 42, 3510-3527, http://dx.doi.org/10.1016/j.atmosenv.2007.08.057, 2008.

Hoek, G., and Raaschou-Nielsen, O.: Impact of fine particles in ambient air on lung cancer, Chinese Journal of Cancer, 33, 197-203, 10.5732/cjc.014.10039, 2014.

Im, U.: Impact of sea-salt emissions on the model performance and aerosol chemical composition and deposition in the East Mediterranean coastal regions, Atmospheric Environment, 75, 329-340, https://doi.org/10.1016/j.atmosenv.2013.04.034, 2013.

Itahashi, S., Ge, B., Sato, K., Fu, J. S., Wang, X., Yamaji, K., Nagashima, T., Li, J., Kajino, M., Liao, H., Zhang, M., Wang, Z., Li, M., Kurokawa, J., Carmichael, G. R., and Wang, Z.: MICS-Asia III: Overview of model inter-comparison and evaluation of acid deposition over Asia, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-624, accepted, 2019

Kelly, J. T., Bhave, P. V., Nolte, C. G., Shankar, U., and Foley, K. M.: Simulating emission and chemical evolution of coarse sea-salt particles in the Community Multiscale Air Quality (CMAQ) model, Geoscientific Model Development, 3, 257-273, https://doi.org/10.5194/gmd-3-257-2010, 2010.

Knol, A. B., de Hartog, J. J., Boogaard, H., Slottje, P., van der Sluijs, J. P., Lebret, E., Cassee, F. R., Wardekker, A., Ayres, J. G., Borm, P. J., Brunekreef, B., Donaldson, K., Forastiere, F., Holgate, S. T., Kreyling, W. G., Nemery, B., Pekkanen, J., Stone, V., Wichmann, H. E., and Hoek, G.: Expert elicitation on ultrafine particles: likelihood of health effects and causal pathways, Particle and Fibre Toxicology, 6, 10.1186/1743-8977-6-19, 2009.

Lamarque, J. F., Shindell, D. T., Josse, B., Young, P. J., Cionni, I., Eyring, V., Bergmann, D., Cameron-Smith, P., Collins, W. J., Doherty, R., Dalsoren, S., Faluvegi, G., Folberth, G., Ghan, S. J., Horowitz, L. W., Lee, Y. H., MacKenzie, I. A., Nagashima, T., Naik, V., Plummer, D., Righi, M., Rumbold, S. T., Schulz, M., Skeie, R. B., Stevenson, D. S., Strode, S., Sudo, K., Szopa, S., Voulgarakis, A., and Zeng, G.: The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): overview and description of models, simulations and climate diagnostics, Geoscientific Model Development, 6, 179-206, 10.5194/gmd-6-179-2013, 2013.

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., and Pozzer, A.: The contribution of outdoor air pollution sources to premature mortality on a global scale, Nature, 525, 367-+, 10.1038/nature15371, 2015.

Li, J., Wang, Z. F., Zhuang, G., Luo, G., Sun, Y., and Wang, Q.: Mixing of Asian mineral dust with anthropogenic pollutants over East Asia: a model case study of a super-duststorm in March 2010, Atmospheric Chemistry and Physics, 12, 7591-7607, DOI: 10.5194/acp-12-7591-2012, 2012.

Li, M., Zhang, Q., Kurokawa, J. I., Woo, J. H., He, K., Lu, Z., Ohara, T., Song, Y., Streets, D. G., Carmichael, G. R., Cheng, Y., Hong, C., Huo, H., Jiang, X., Kang, S., 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.

Lippmann, M.: Toxicological and epidemiological studies of cardiovascular effects of ambient air fine particulate matter (PM2.5) and its chemical components: Coherence and public health implications, Crit Rev Toxicol, 44, 299-347, DOI: 10.3109/10408444.2013.861796, 2014.

Liu, Y. M., Zhang, S. T., Fan, Q., Wu, D., Chan, P. W., Wang, X. M., Fan, S. J., Feng, Y. R., and Hong, Y. Y.: Accessing the Impact of Sea-Salt Emissions on Aerosol Chemical Formation and Deposition over Pearl River Delta, China, Aerosol Air Qual Res, 15, 2232-2245, doi: 10.4209/aaqr.2015.02.0127, 2015.

Nel, A.: Air pollution-related illness: Effects of particles, Science, 308, 804-806, DOI: 10.1126/science.1108752, 2005.

Qu, W. J., Arimoto, R., Zhang, X. Y., Zhao, C. H., Wang, Y. Q., Sheng, L. F., and Fu, G.: Spatial distribution and interannual variation of surface PM10 concentrations over eighty-six Chinese cities, Atmospheric Chemistry and Physics, 10, 5641-5662, 10.5194/acp-10-5641-2010, 2010.

Rao, S., Chirkov, V., Dentener, F., Van Dingenen, R., Pachauri, S., Purohit, P., Amann, M., Heyes, C., Kinney, P., Kolp, P., Klimont, Z., Riahi, K., and Schoepp, W.: Environmental Modeling and Methods for Estimation of the Global Health Impacts of Air Pollution, Environmental Modeling & Assessment, 17, 613-622, 10.1007/s10666-012-9317-3, 2012.

Shao, Y., Dong, C. H. A review on East Asian dust storm climate, modeling and moni-toring: Global and Planetary Change, 2006, 52(1-4): 1-22.

Silva, R. A., West, J. J., Zhang, Y. Q., Anenberg, S. C., Lamarque, J. F., Shindell, D. T., Collins, W. J., Dalsoren, S., Faluvegi, G., Folberth, G., Horowitz, L. W., Nagashima, T., Naik, V., Rumbold, S., Skeie,

R., Sudo, K., Takemura, T., Bergmann, D., Cameron-Smith, P., Cionni, I., Doherty, R. M., Eyring, V., Josse, B., MacKenzie, I. A., Plummer, D., Righi, M., Stevenson, D. S., Strode, S., Szopa, S., and Zeng, G.: Global premature mortality due to anthropogenic outdoor air pollution and the contribution of past climate change, Environ Res Lett, 8, 10.1088/1748-9326/8/3/034005, 2013.

Tebaldi, C., and Knutti, R.: The use of the multi-model ensemble in probabilistic climate projections, Philos T R Soc A, 365, 2053-2075, https://doi.org/10.1098/rsta.2007.2076, 2007.

Vivanco, M. G., Bessagnet, B., Cuvelier, C., Theobald, M. R., Tsyro, S., Pirovano, G., Aulinger, A., Bieser, J., Calori, G., Ciarelli, G., Manders, A., Mircea, M., Aksoyoglu, S., Briganti, G., Cappelletti, A., Colette, A., Couvidat, F., D'Insidoro, M., Kranenburg, R., Meleux, F., Menut, L., Pay, M. T., Rouil, L., Silibello, C., Thunis, P., and Ung, A.: Joint analysis of deposition fluxes and atmospheric concentrations of inorganic nitrogen and sulphur compounds predicted by six chemistry transport models in the frame of the EURODELTAIII project, Atmospheric Environment, 151, 152-175, https://doi.org/10.1016/j.atmosenv.2016.11.042, 2017.

Wang, K., Zhang, Y., Nenes, A., and Fountoukis, C.: Implementation of dust emission and chemistry into the Community Multiscale Air Quality modeling system and initial application to an Asian dust storm episode, Atmospheric Chemistry and Physics, 12, 10209-10237, https://doi.org/10.5194/acp-12-10209-2012, 2012.

Wesely, M. L.: Parameterization of Surface Resistances to Gaseous Dry Deposition in Regional-Scale Numerical-Models, Atmospheric Environment, 23, 1293-1304, https://doi.org/10.1016/0004-6981(89)90153-4, 1989.

Xing, J., Mathur, R., Pleim, J., Hogrefe, C., Gan, C. M., Wong, D. C., Wei, C., Gilliam, R., and Pouliot, G.: Observations and modeling of air quality trends over 1990-2010 across the Northern Hemisphere: China, the United States and Europe, Atmospheric Chemistry and Physics, 15, 2723-2747, 10.5194/acp-15-2723-2015, 2015.

Zhang, L., Brook, J. R., and Vet, R.: A revised parameterization for gaseous dry deposition in air-quality models, Atmospheric Chemistry and Physics, 3, 2067-2082, https://doi.org/10.5194/acp-3-2067-2003, 2003.

Figures and tables

Figure 1

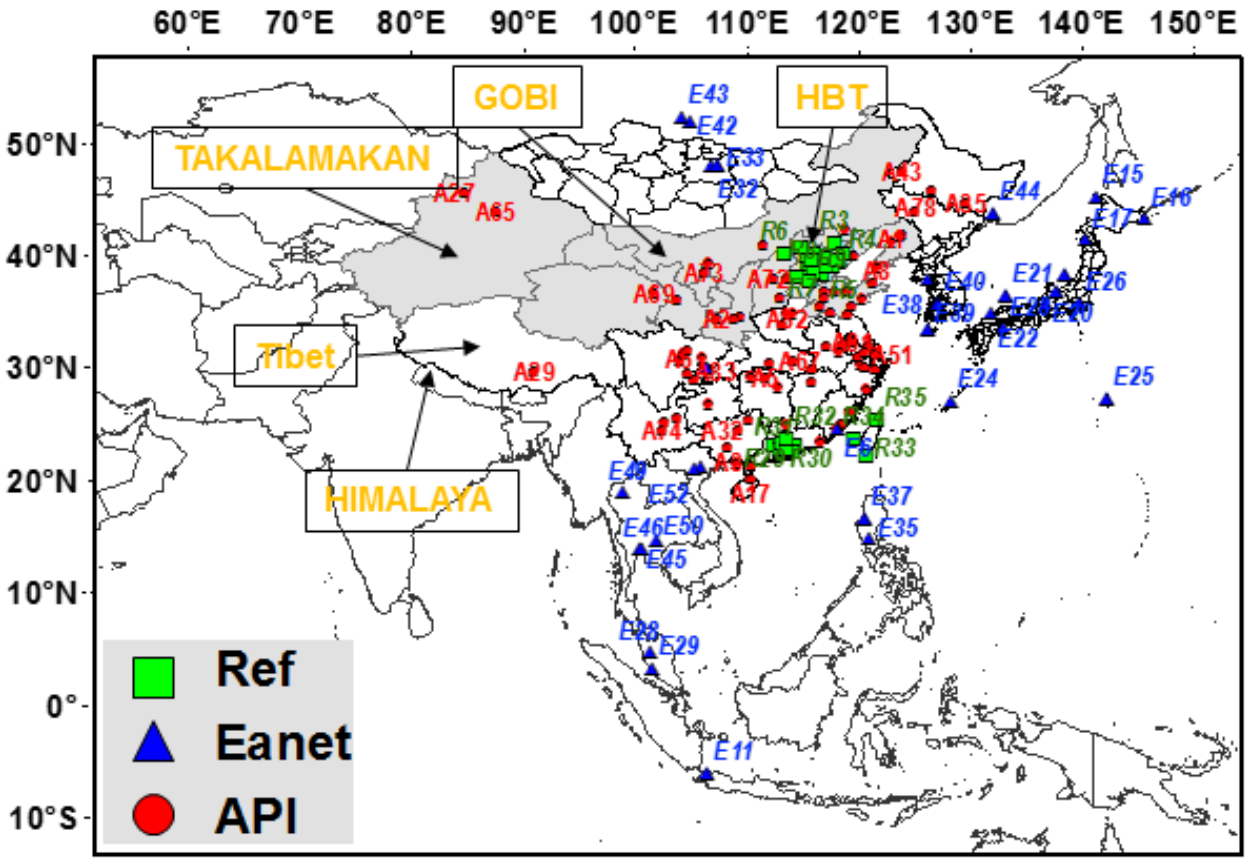


Figure 1 The geographical locations of observation networks of API (red color, A1-A86), EANET (blue color, E1-E54, only sites with available observation during simulation time are shown) and Ref (green color, R1-R35) sites. Grey shaded regions have been reported to be affected by dust storms.

Figure 2

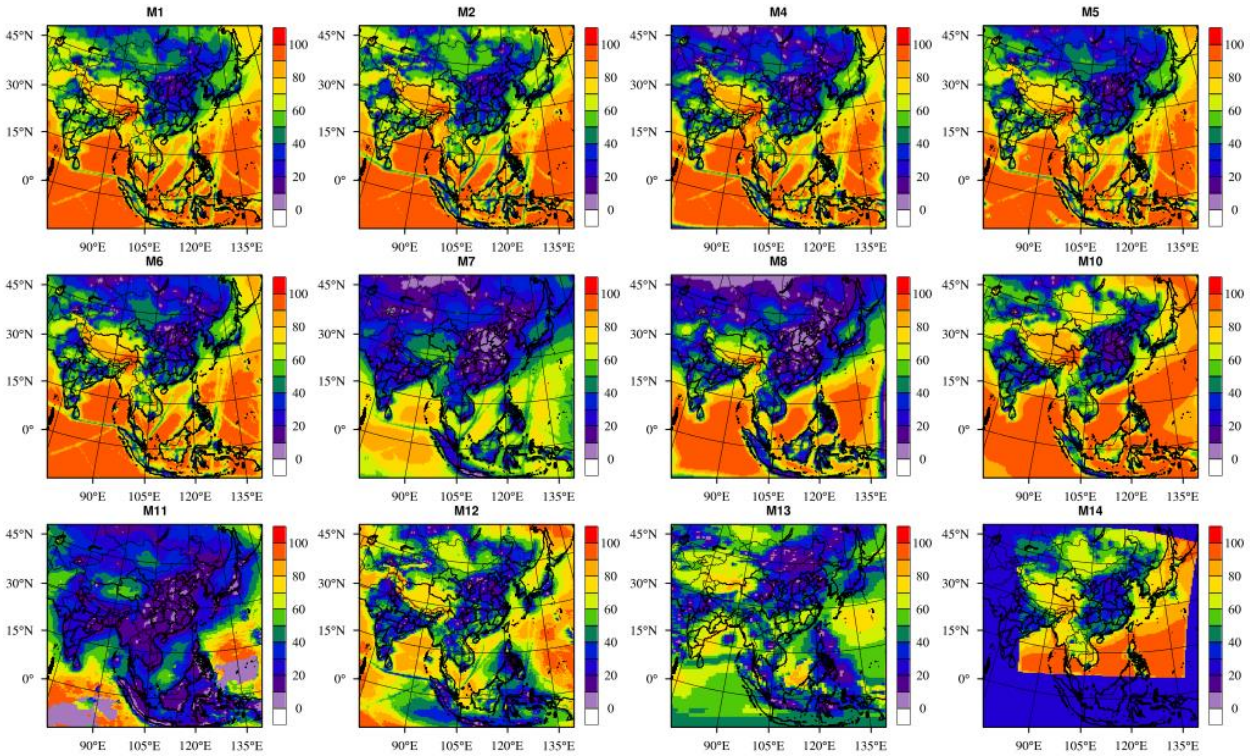


Figure 2 *SOR* values at surface layer for models (unit: %). *SOR* is calculated by $SO_4^{2-}/(SO_2+SO_4^{2-})\times100\%$. The SO_2 and SO_4^{2-} concentrations are transferred from ppb and μg m^{-3} to mole(S) m^{-3} before calculating *SOR*. Values are calculated by annual average data.

Figure 3

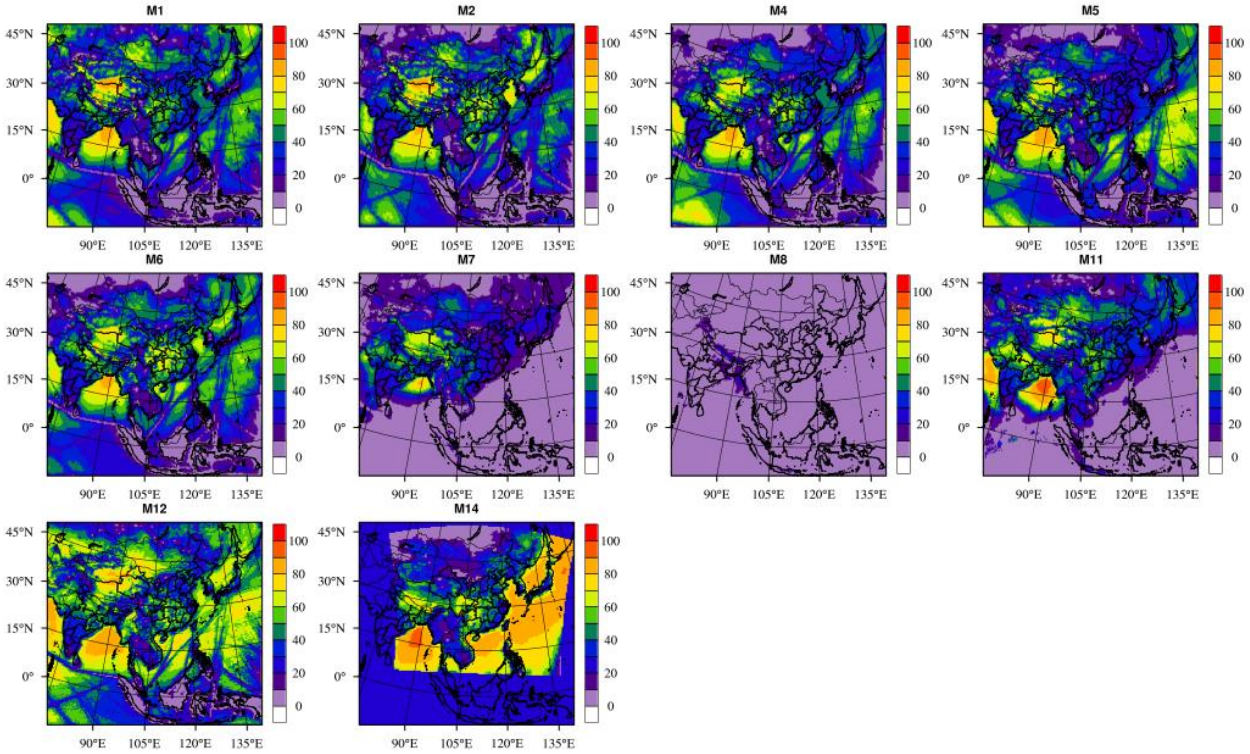


Figure 3 Same as Fig.2 but for *C(NO_2)* (unit: %). *C(NO_2)* is calculated by $NO_3^-/(NO_2+NO_3^-)\times100\%$. The *C($NO_2$)* of M8 is extremely low due to unreasonable low NO_3^- concentration, which is considered as outliner in this study. Values are calculated by annual average data.

Figure 4

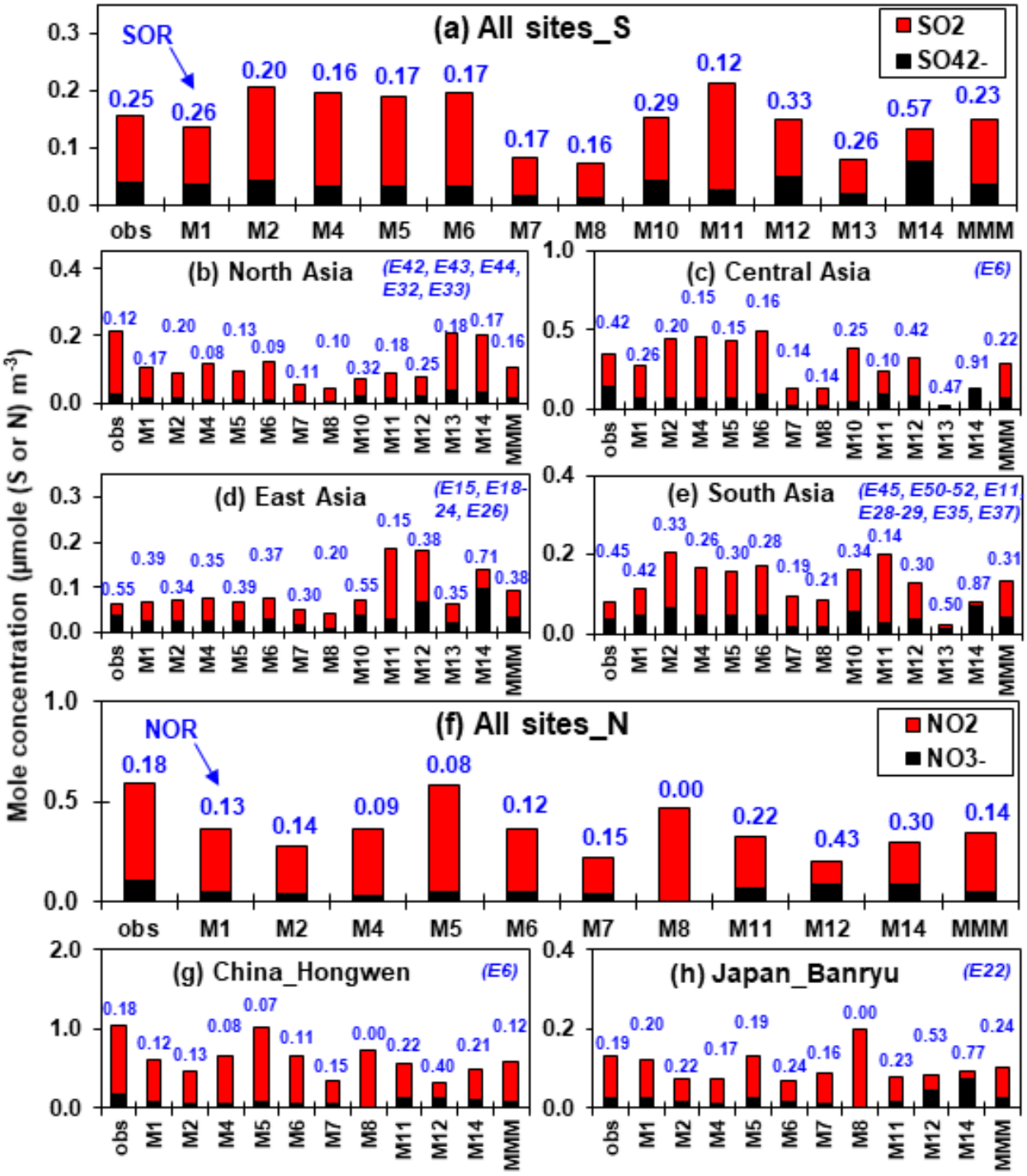


Figure 4 Gas-particle conversions of S and N of observation and models at EANET sites. The unit is μmole (S or N) m^{-3}. The red bars and black bars represent the concentrations of gases and aerosols. The blue-color values above the bars are observed/modelled *SOR* and $C(NO_2)$. Values are calculated with annual average concentrations. The concentrations of gases and aerosols are all transferred to μmole (S or N) m^{-3} before calculation. The blue-color numbers on top-right (e.g. E22) are site numbers. The locations of the sites are illustrated in Fig. 1. Results for individual sites are shown in supplementary Fig. S1.

Figure 5

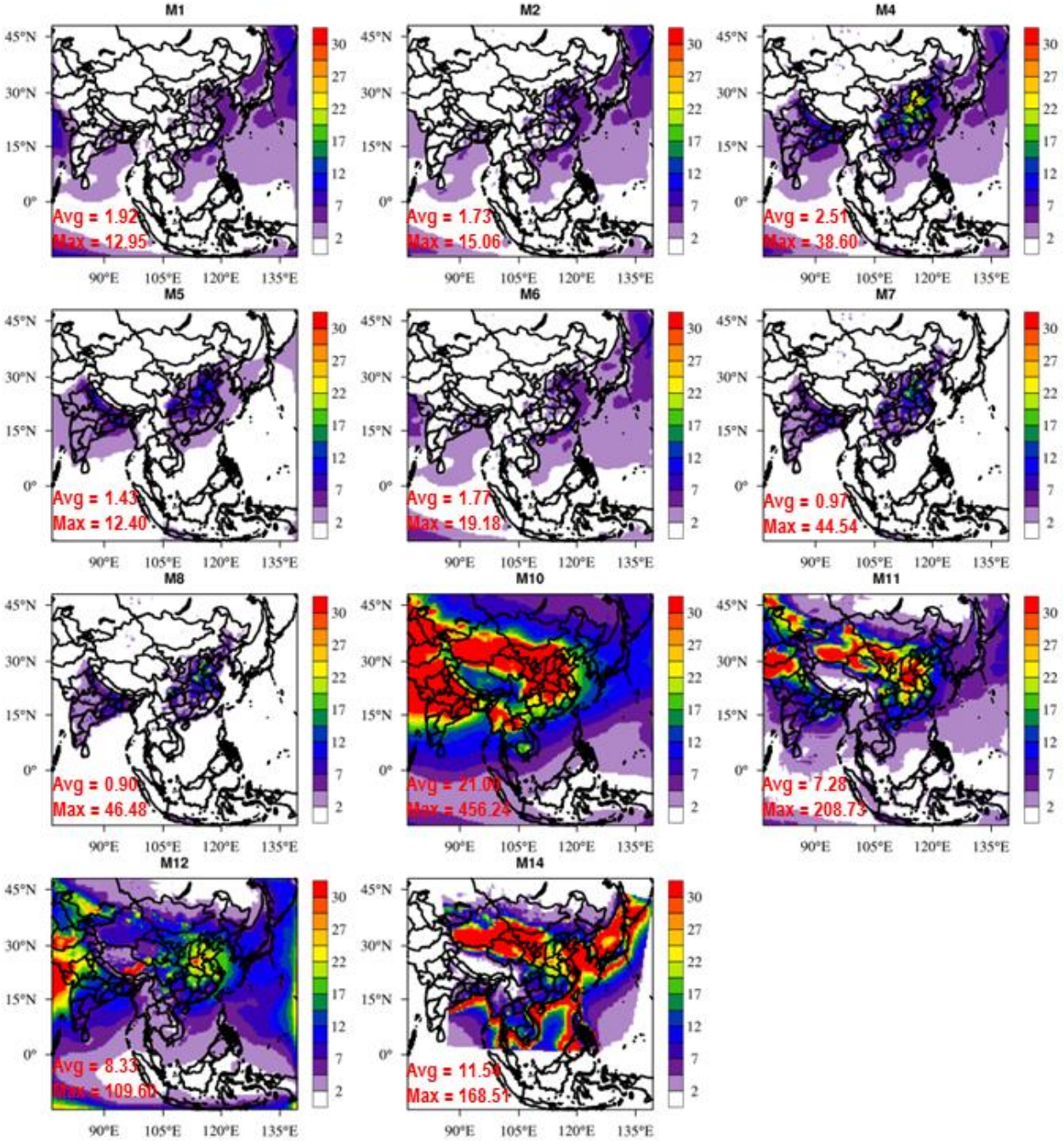


Figure 5 Annual average PMC concentrations at surface layer of individual models (μg m^{-3}). The value is calculated by subtracting $PM_{2.5}$ from PM_{10}. The values in left-bottom are domain average (Avg) and maximum (Max) values.

Figure 6

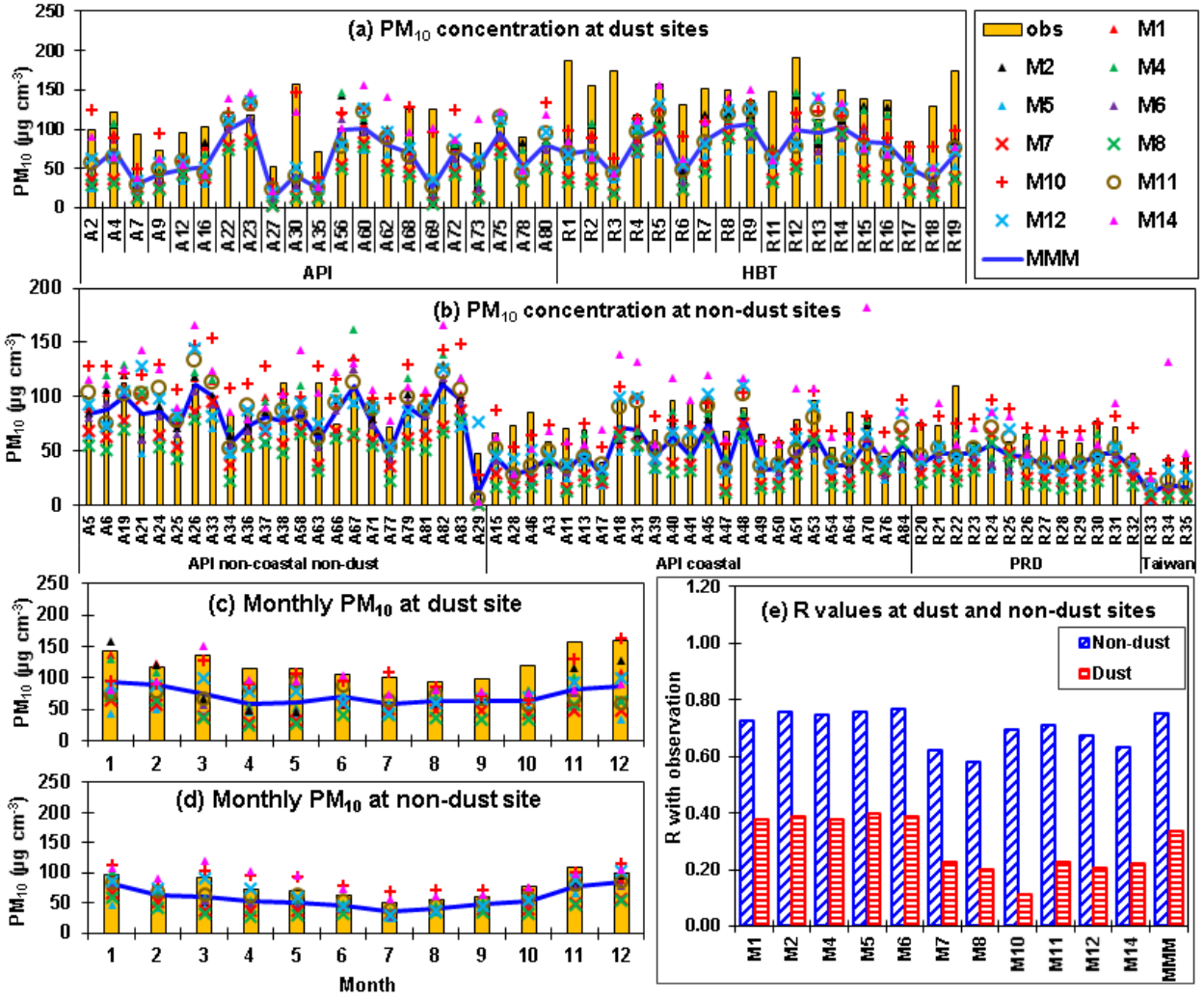


Figure 6 Multi-model performances on (a-b) annual average PM_{10} concentrations at the dust sites and non-dust sites and (c-d) monthly average PM_{10} concentrations at the dust sites and non-dust sites. X axis for (a-b) indicates site numbers. The locations of the sites are illustrated in figure 1. The yellow bars are observations, the blue lines are the MMM and different markers represent individual model results. (e) R values of models with observations at the dust and non-dust sites.

Figure 7

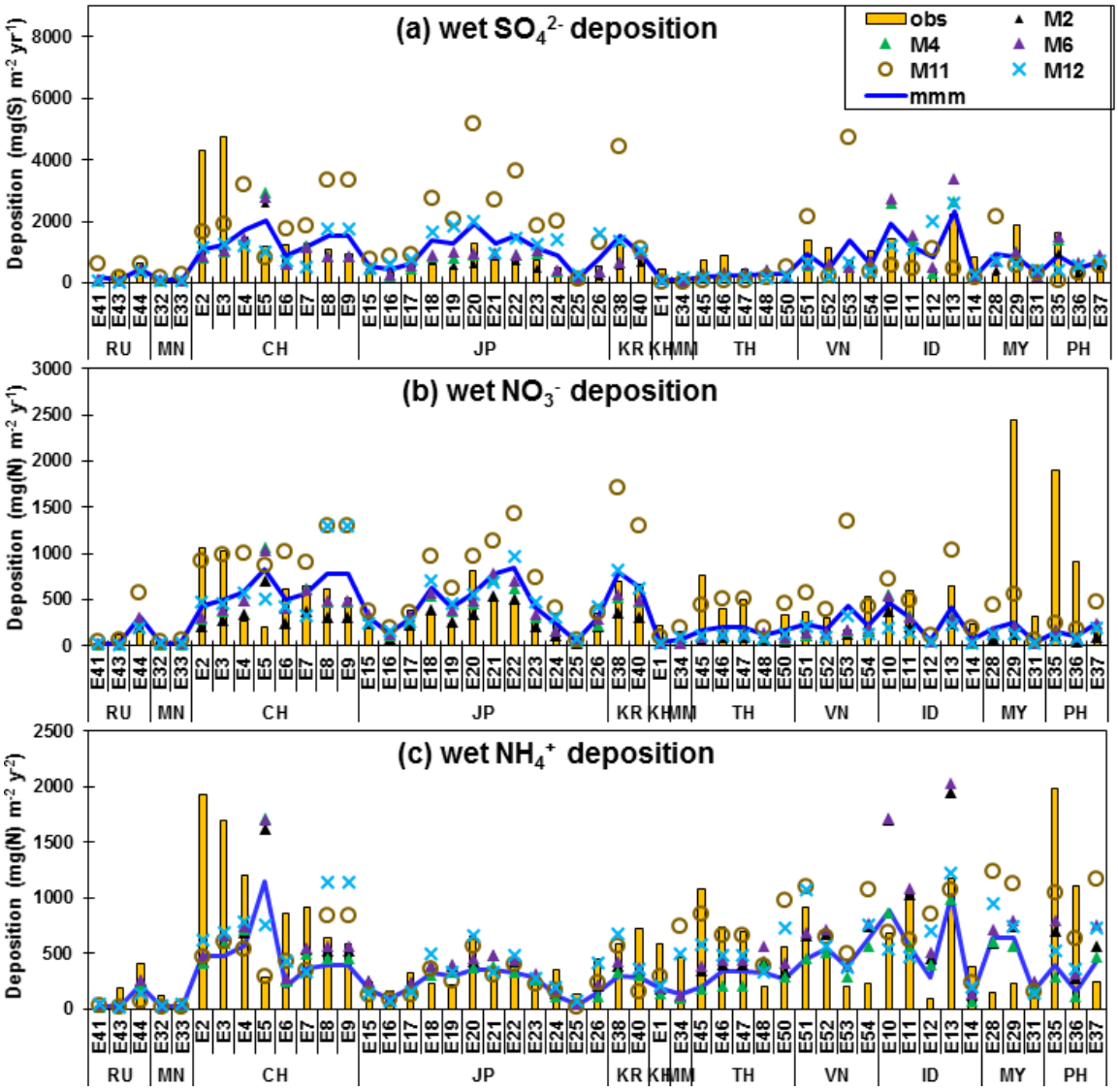
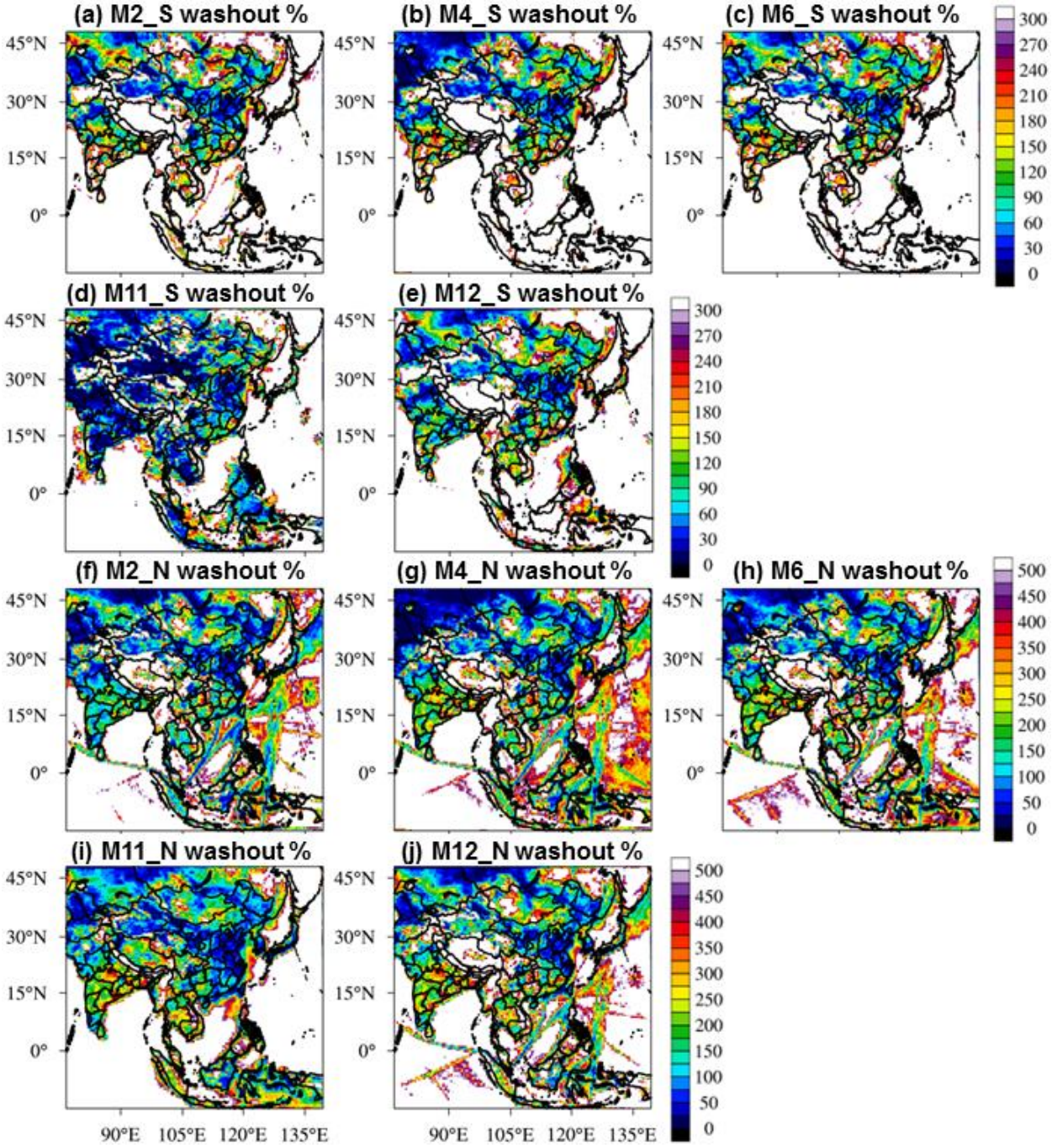


Figure 7 Modelled annual-accumulated wet deposition of SO_4^{2-}, NO_3^- and NH_4^+ compared with observation from EANET network. The units are mg(S or N) m^{-2} yr^{-1}. Abbreviation for regions: RU-Russia, MN-Mongolia, CH-China, JP-Japan, KR-Korea, KH-Cambodia, MN-Myanmar, TH-Thailand, VN-Vietnam, ID-Indonesia, MY-Malaysia, PH-Philippine.

682 **Figure 8**



683
684 Figure 8 Washout ratios (λ_{wet}) of (a-e) S deposition and (f-j) N deposition of models. Values are calculated with
685 annual accumulated depositions. The unit is %.
686

Figure 9

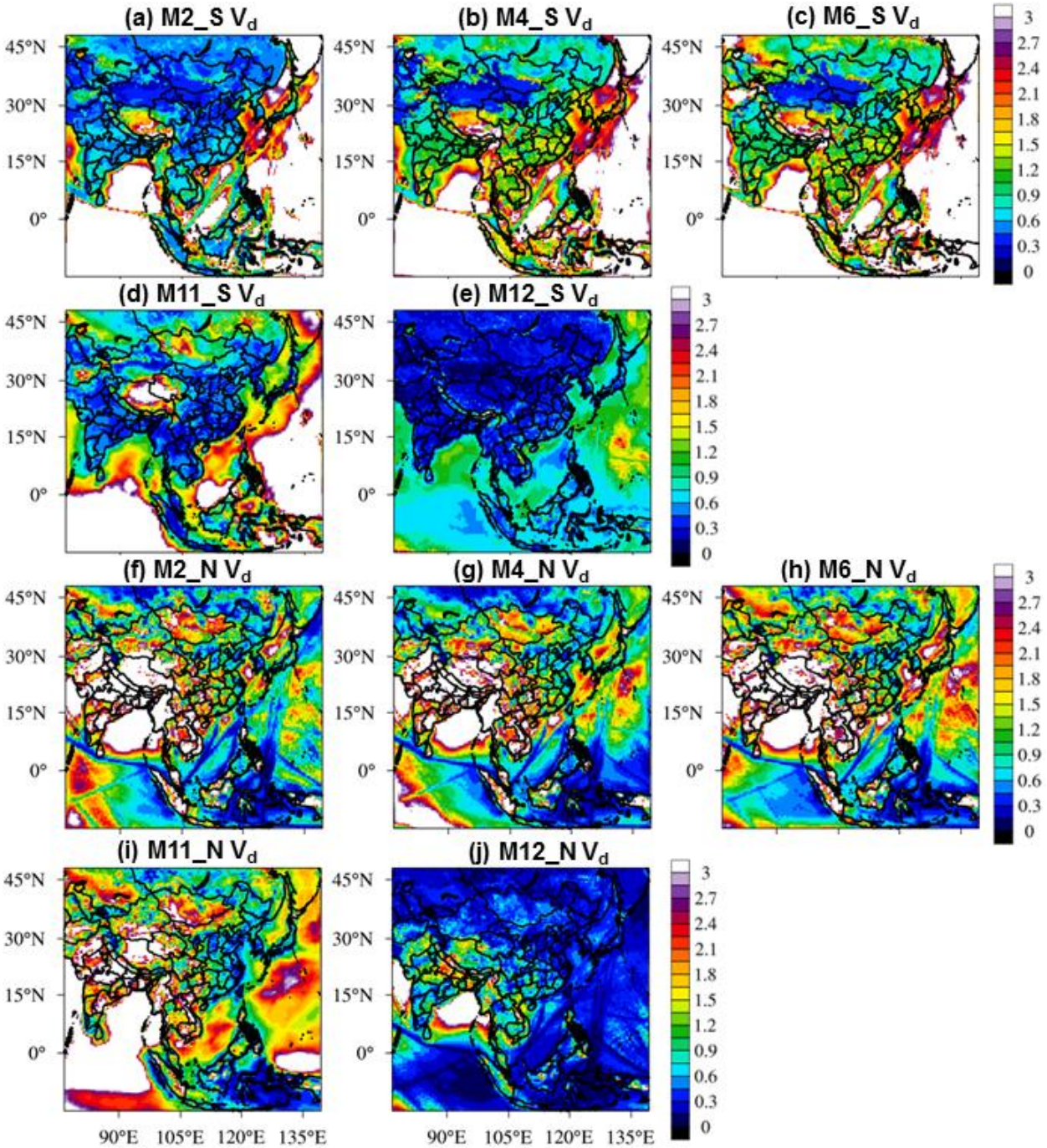


Figure 9 Dry deposition velocities (V_d) of (a-e) S deposition and (f-j) N deposition of models. Values are calculated with annual accumulated depositions. The unit is cm s^{-1}.

692 **Table 1**

693 Table 1 Multi-model performance on annual average concentrations of PM_{10} at the dust and non-
694 dust sites (unit: μg m^{-3})

Dust site	M1	M2	M4	M5	M6	M7	M8	M10	M11	M12	M14	MMM
Mean Obs						120.7						
Mean MMM	77.2	82.2	81.6	51.7	65.6	47.5	44.3	102.5	73.5	77.3	92.1	69.2
S	0.4	0.4	0.4	0.3	0.3	0.2	0.2	0.1	0.2	0.2	0.3	0.3
MB	-43.5	-38.5	-39.2	-69.0	-55.1	-73.2	-76.4	-18.2	-47.2	-43.4	-28.6	-51.5
R	0.4	0.4	0.4	0.4	0.4	0.2	0.2	0.1	0.2	0.2	0.2	0.3
F	66.7	69.2	69.2	38.5	56.4	35.9	33.3	84.6	59.0	66.7	66.7	66.7
NMB (%)	-36.1	-31.9	-32.4	-57.2	-45.7	-60.6	-63.3	-15.1	-39.1	-36.0	-23.7	-42.6
NME (%)	38.3	35.4	36.4	57.2	46.2	60.6	63.3	32.8	42.3	40.5	36.1	42.7
MFB (%)	-49.4	-44.6	-44.6	-83.4	-64.1	-92.9	-98.8	-19.3	-51.8	-46.8	-31.7	-56.9
MFE (%)	51.8	48.3	48.7	83.4	64.7	92.9	98.8	36.1	55.3	51.7	44.5	56.9
Number of Sites						39						

695

696 Table 1 Continued

Non-dust site	M1	M2	M4	M5	M6	M7	M8	M10	M11	M12	M14	MMM
Mean Obs						77.2						
Mean MMM	58.2	58.5	66.5	45.2	55.2	44.8	39.0	90.0	64.4	66.3	89.5	57.8
S	1.0	1.1	1.2	0.8	1.0	0.7	0.6	1.0	1.0	0.9	1.1	0.9
MB	-19.0	-18.7	-10.8	-32.1	-22.1	-32.5	-38.3	12.7	-12.9	-10.9	12.2	-19.4
R	0.7	0.8	0.7	0.8	0.8	0.6	0.6	0.7	0.7	0.7	0.6	0.8
F	82.5	81.0	84.1	66.7	82.5	52.4	46.0	85.7	90.5	93.7	84.1	82.5
NMB (%)	-24.6	-24.2	-14.0	-41.5	-28.6	-42.0	-49.5	16.5	-16.6	-14.1	15.8	-25.1
NME (%)	30.7	30.7	27.3	41.5	31.4	43.9	50.7	25.7	26.3	26.1	30.8	28.0
MFB (%)	-36.8	-37.5	-25.1	-59.2	-41.8	-62.0	-75.0	13.1	-24.9	-20.3	8.3	-34.6
MFE (%)	42.0	42.8	35.3	59.2	44.4	64.0	76.1	23.4	33.5	31.3	29.1	37.5
Number of Sites						63						

697

698

699 **Table 2**

700 Table 2 Multi-model performances on wet deposition (unit: mg(S or N) m^{-2} y^{-1})

	Wet SO_4^{2-} Deposition						Wet NO_3^- Deposition					
	M2	M4	M6	M11	M12	MMM	M2	M4	M6	M11	M12	MMM
Mean Obs	931.3	931.3	931.3	931.3	931.3	931.3	460.9	460.9	460.9	460.9	460.9	460.9
Mean MMM	633.7	724.2	775	1313.2	826.2	854.5	187.5	266.7	279.5	597.8	308.3	328
S	0.3	0.3	0.3	0.3	0.2	0.3	0.1	0.1	0.1	0.2	0.1	0.1
MB	-297.7	-207.1	-156.3	381.9	-105.1	-76.9	-273.4	-194.2	-181.4	137	-152.6	-132.9
R	0.5	0.4	0.4	0.2	0.3	0.4	0.1	0.2	0.2	0.2	0.1	0.2
F	61.2	61.2	61.2	24.5	40.8	51	38.8	49	46.9	44.9	38.8	46.9
NMB	-32	-22.2	-16.8	41	-11.3	-8.3	-59.3	-42.1	-39.4	29.7	-33.1	-28.8
NME	49.3	50.2	51.5	117.3	62.8	53.6	66.2	60.9	60.6	78.4	68.8	58.2
MFB	-37.4	-23.4	-15.8	4.6	-11.4	-4.6	-75.8	-49.8	-42.1	25.8	-40.9	-27.6
MFE	57.8	55.9	53.7	93.8	66.7	57.6	84.9	71.2	69.3	61	74.6	62.3
Number of Sites	49	49	49	49	49	49	49	49	49	49	49	49

701

702 Table 2 Continued

	Wet NH_4^+ Deposition					
	M2	M4	M6	M11	M12	MMM
Mean Obs	558.4	558.4	558.4	558.4	558.4	558.4
Mean MMM	459.9	349.4	497.4	505	478	337.6
S	0.3	0.1	0.3	0.3	0.2	0.2
MB	-98.5	-208.9	-61	-53.4	-80.4	-220.7
R	0.3	0.2	0.3	0.4	0.4	0.3
F	40.8	44.9	44.9	51	46.9	38.8
NMB	-17.6	-37.4	-10.9	-9.6	-14.4	-39.5
NME	64.8	65.5	64.9	58.2	57	63.6
MFB	-21.2	-42.4	-14.4	-18	-12.6	-41.9
MFE	70.7	77.9	69.1	65.9	62.9	76.1
Number of sites	49	49	49	49	49	49

703

704

Table 3

Table 3 Domain-total annual-accumulated S and N depositions of models (Tg(S or N) yr^{-1}). Empty values mean no model submissions or the values are 0.

Model	Wet S deposition				Dry S deposition			
	SO_2	H_2SO_4	SO_4^{2-}	Total Wet S	SO_2	H_2SO_4	SO_4^{2-}	Total Dry S
M1	0.06	-	-	-	-	-	-	-
M2	0.04	-	10.4	10.5	3.4	0.01	0.9	4.3
M4	0.06	-	12.5	12.5	6.6	0.01	1.1	7.6
M5	-	-	-	-	-	-	-	-
M6	0.05	-	13.7	13.8	6.3	0.01	1.4	7.7
M7	-	-	-	-	-	-	-	-
M8	-	-	-	-	-	-	-	-
M10	-	-	-	-	-	-	-	-
M11	1.1	0.3	29.9	31.3	6.9	2.2	1.5	10.6
M12	-	-	16.3	16.3	3.7	-	0.4	4.2
M13	6.0	-	-	-	-	-	-	-
M14	0.02	-	6.2	-	5.4	-	3.2	-

Table 3 Continued

Model	Wet N deposition					Dry N deposition						
	NO_3^-	NH_4^+	HNO_3	NH_3	Total Wet N	NO	NO_2	NO_3^-	NH_4^+	HNO_3	NH_3	Total Dry N
M1	-	-	-	-	-	-	-	-	-	4.3	6.9	-
M2	4.0	8.3	-	-	12.2	0.03	0.4	0.6	0.6	2.0	7.5	11.0
M4	5.4	7.4	-	-	12.8	0.03	0.3	0.7	0.5	2.8	4.7	9.0
M5	-	-	-	-	-	-	0.5	-	-	-	-	-
M6	5.6	9.1	-	-	14.6	0.02	0.3	0.8	0.7	2.9	6.5	11.1
M7	-	-	-	-	-	-	-	-	-	-	-	-
M8	-	-	-	-	-	-	-	-	-	-	-	-
M10	-	-	-	-	-	-	-	-	-	-	-	-
M11	1.5	2.8	8.1	7.6	20.0	-	-	1.3	2.4	3.3	7.1	14.1
M12	5.4	11.0	-	-	16.5	0.04	0.4	0.4	0.3	0.5	2.2	3.9
M13	-	-	4.1	-	-	-	-	-	-	4.5	4.6	-
M14	-	-	-	-	-	-	-	-	-	-	-	-