Response to Comment 2 by Anonymous Reviewer 2

Review of MICS-Asia III: Overview of model inter-comparison and evaluation of acid deposition over Asia by Itahashi et al.

We thank you for providing helpful and constructive comments and suggestions. We have revised our manuscript accordingly. For convenience, we have divided your general comments using numbers. We hope that these revisions have satisfactorily addressed all the points you have raised. Our point-by-point responses are provided below, and revisions are indicated in blue in the revised manuscript.

1. The paper deals with an intermodel comparison of wet deposition patterns for East Asia given present day conditions. The models used in this study are chemistry transport models and meteorological data is provided by external model fields (for most models form WRF). In general this study is worth publication, however the gain of new scientific knowledge from this effort is rather small. Consequently, would have proposed that this manuscript is much better suited for GMD instead of ACP as a typical evaluation paper. The paper is decently structured and despite some minor required language improvements (proof-reading) could be easily published. The model and observation data description sections are adequate; however if it might be possible to add the observation data file in the supplement more model applicants would benefit from this study.

Reply:

Thank you for your journal recommendation; however, we believe that the ACP is an appropriate selection. The "Aims and Scope" of GMD and ACP are as follows:

GMD

- **geoscientific model descriptions, from statistical models to box models to GCMs;**
- **development and technical papers, describing developments such as new parameterizations or technical aspects of running models such as the reproducibility of results;**
- **new methods for assessment of models, including work developing new metrics for**

assessing model performance and novel ways of comparing model results with observational data;

- **papers describing new standard experiments for assessing model performance or novel ways of comparing model results with observational data;**
- **model experiment descriptions, including experimental details and project protocols;**
- **full evaluations of previously published models**

ACP

ACP covers the altitude range from the land and ocean surface up to the tropopause, including the troposphere, stratosphere, and mesosphere. The main subject areas comprise atmospheric modeling, field measurements, remote sensing, and laboratory studies of gases, aerosols, clouds and precipitation, isotopes, radiation, dynamics, biosphere interactions, and hydrosphere interactions.

The scope of ACP is focused on studies with general implications for atmospheric science, rather than investigations that are primarily of local or technical interest, and authors should thus pay particular attention to whether their paper would fit better in other journals published by Copernicus Publications such as GMD.

We agree that our manuscript may fall within the scope of GMD; however, we would like to submit our manuscript for the special issue of ACP entitled "Regional assessment of air pollution and climate change over East and Southeast Asia: results from MICS-Asia Phase III". Our manuscript describes the Model Inter-comparison Study for Asia (MICS-Asia) Phase III, and the model results are fully evaluated using data from the Acid Deposition Monitoring Network in East Asia (EANET) covering the whole of Asia. The manuscript contains not only model evaluation, but also the following three main discussion points: 1. acid deposition mapping over Asia; 2. an ensemble approach; and 3. a precipitation-adjusted approach. The latter two discussion points are aimed at improving modeling performance, whereas the first point is strongly related to the importance of the EANET observation network, as well as the balance between emissions (input) and depositions (output). Our manuscript concerns the impacts of Asian air

quality not only locally but also globally, so we believe that it has general implications for atmospheric science rather than investigations that are primarily of local or technical interest and that it will be of interest to readers of ACP. To increase the suitability of our manuscript for publication in ACP, we have sought to further improve it based on the reviewers' comments.

The EANET observational dataset used in this study is freely available from a publicly accessible website. Therefore, we believe that it is not necessary to include data as supplemental material. However, we have added a "Data Availability" section to the end of the revised manuscript in which we provide the URL where the EANET dataset can be accessed.

2. Despite the model domain is very similar (if not identical) in the simulations and the models use unified emissions, substantial differences in the deposition maps occur. Unfortunately, the authors do not show a comparison of the simulated concentrations or tropospheric vertical columns of the species the study focuses on. Therefore, it is difficult to judge, whether the differences occur from slightly different assumptions in the wet deposition schemes, the simulated precipitation amount or the simulated concentrations of the trace species and their precursors. As the applied chemistry and aerosol schemes differ to a certain degree, this could already be a major cause for the differences in the wet deposition patterns.

Reply:

We agree that the simulated concentrations are needed to explain the differences in wet deposition analyzed in this study. Because our result is a part of the MICS-Asia project, we prefer not to explicitly show the simulated concentration fields within this manuscript to avoid redundancy with respect to our companion papers that were published in a special issue of MICS-Asia Phase III. Instead, to address your comment, we have added references to our companion papers (Chen et al., 2019; Tan et al., 2019). Additionally, we have conducted an evaluation of wet deposition to be consistent in the use of the same observation sites and prepared one additional figure (Figure S1 in the revised supporting material) and three additional tables (Tables S1, S2, and S3 in the revised supporting material). These points have now been addressed as independent

paragraphs in Sections 3.2.1, 3.2.2, and 3.2.3.

In Section 3.2.1:

"The model performances for atmospheric concentrations were presented in our companion paper (Figs. 3 and 5 and Table 2 of Chen et al., 2019). For consistency with that companion paper, we also performed the model evaluation at the same sites used for the analysis of atmospheric concentrations. The results are shown in Table S1 and the correspondence between the NMB of atmospheric concentration and wet deposition is shown in Figure S1 (a). The modeling performances were generally similar for the wet deposition of S using all data (Table 1) and using limited data (Table S1). Models M1, M2, M4, M5, M6, and M13 underestimated atmospheric concentrations of SO_4^{2-} over Asia, with an NMB of around –30 to –20% (Table S1) and accordingly these models also underestimated the wet deposition of SO_4^{2-}. Only models M12 and M14 overestimated atmospheric concentrations of SO_4^{2-} and model M14 was also distinguished by the overestimation of the atmospheric concentration of SO_4^{2-} over coastal regions, such as over Korea and Japan. Model M11 was the only model to overestimate the wet deposition of SO_4^{2-} and produced the largest underestimation of the atmospheric concentration of SO_4^{2-}, with an NMB of –34.5%. These overestimations (underestimations) of the atmospheric concentration of SO_4^{2-} are closely related to the underestimation (overestimation) of wet deposition of $SO4^{2-}$ found in models M11, M12, and M14. The close relationship between atmospheric concentration and wet deposition was also observed in a model inter-comparison study in Japan (Itahashi et al., 2018c). The atmospheric concentration of SO_4^{2-} was underestimated, especially in winter (Fig. 5 of Chen et al., 2019). Another companion paper (Tan et al., 2019) investigated the sulfur oxidation ratio, which represents the conversion rate from SO_2 to SO_4^{2-}. The observation-based ratio was 0.25. Models M1 and M13 both predicted a comparable sulfur oxidation ratio of 0.26; however, models M2, M4, M5, and M6 underestimated the ratio, giving values of around 0.16–0.20, suggesting the insufficient oxidation from the precursor of SO_2 (Fig. 2 of Tan et al., 2019). The sulfur oxidation ratio was strongly underestimated by model M11, which gave a value of 0.12. This underestimation can be corrected by refining the treatment of catalysis using O_2, introducing the aqueous-phase production of SO_4^{2-} using NO_2, or newly established gas-phase oxidation by the stabilized Criegee intermediate (Itahashi et al., 2018b, 2018c, 2019). Moreover, another study pointed out that heterogeneous chemistry is a possible explanation for the missing

production of SO_4^{2-} in models (Zheng et al., 2015; Shao et al., 2019). The modeled sulfur oxidation ratios of model M12 and M14 were 0.33 and 0.57, respectively; that is, the ratios were overestimated. This overestimation is one reason for the overestimation of the atmospheric concentration of SO_4^{2-} by models M12 and M14. In summary, for models M11, M12, and M14, the model performance for the wet deposition of S is characterized by a close relationship with either the overestimation or underestimation of atmospheric concentration, and models M1, M2, M4, M5, M6, and M13 underestimated both atmospheric concentration and wet deposition of S. "

In Section 3.2.2:

"The model performances for the atmospheric concentration of NO_3^- in our companion paper also showed large differences between models (Figs. 3 and 5 and Table 2 of Chen et al., 2019). The model evaluation for the analysis of atmospheric concentrations for N was conducted at the same sites as those for S (Table S1), as shown in Table S2. The correspondence between the NMB of atmospheric concentration and the NMB of wet deposition is shown in Figure S1 (b). Models M2, M4, M5, and M6 showed underestimation, whereas models M1, M11, M12, M13, and M14 showed overestimation. Models M1, M2, and M5 showed better performance in terms of NMB (NMBs of between –10% and 10%) (Table S2). If both H_2SO_4 and HNO_3 are present, H_2SO_4 preferentially reacts with NH_3, and therefore NH_4NO_3 is produced only if excess NH_3 is present. The underestimation of the atmospheric concentration of SO_4^{2-} can lead to the overestimation of the atmospheric concentration of NO_3^-. This can explain the performance of models M1, M11 and M13 but not that of models M12 and M14 because they overestimated the atmospheric concentrations of both SO_4^{2-} and NO_3^-. Another companion paper revealed that models M12 and M14 also used a higher nitrogen oxidation ratio (i.e., the ratio of oxidation from NO_2 to NO_3^-) than that of other models and observation, in addition to using a higher sulfur oxidation ratio (Fig. 2 of Tan et al., 2019). The higher oxidation capacity in models M12 and M14 is connected to the overestimation of the atmospheric concentration of both SO_4^{2-} and NO_3^-. On the other hand, models M2, M4, M5, and M6 underestimated the atmospheric concentration of both SO_4^{2-} and NO_3^-. These four models of M2, M4, M5, and M6 also had lowernitrogen oxidation ratios of between 0.08 and 0.14 than the observed value of 0.18 (Fig. 2 of Tan et al., 2019). In summary, for the wet

deposition of N, all models except M5 and M11 underestimated this parameter; however, the relationship between the wet deposition of N and the atmospheric concentration of was not obvious different from this relationship for S. Because the correlation coefficient for the model performance of the wet deposition of N is lower than that for S, future studies should focus on N and related species in greater detail. Our future companion paper will attempt a detailed analysis of N using an intensive observation network over China."

In Section 3.2.3:

"Our companion paper reported model performances for the atmospheric concentration of NH_4^+ (Figs. 3 and 5 and Table 2 of Chen et al., 2019). The model evaluation for the analysis of atmospheric concentrations for A was conducted at the same sites as those for S and N (Tables S1 and S2), as shown in Table S3, and the correspondence between the NMB of atmospheric concentration and that of wet deposition is shown in Figure S1 (c). Generally, the behavior of NH_4^+ is associated with the atmospheric concentrations of SO_4^{2-} and NO_3^- as counterions. The studied models generally underestimated the atmospheric concentration of S and overestimated the atmospheric concentration of N; consequently, all models except M4 overestimated the atmospheric concentration of A. The reason for the different behavior of model M4 is that this model underestimated atmospheric concentrations of NO_3^-. In general, the models overestimated the atmospheric concentration of A and underestimated the wet deposition of A (Fig. S1 (c)); this indicates a close relationship between atmospheric concentration and wet deposition processes. "

Finally, we also added the following to the concluding section:

"Comparisons of atmospheric concentrations revealed that model performances are either characterized by a close relationship between the overestimation (underestimation) of the wet deposition of S and the underestimation (overestimation) of the atmospheric concentration of S or characterized by the underestimation of both the atmospheric concentration and wet deposition of S species. The general features for underestimation of wet deposition of A and overestimation of atmospheric concentration of A were clarified.

The relationships between atmospheric concentration and wet deposition of N are complicated and further research focusing on nitrogen species especially targeting the nitrogen cycle is required."

3. The evaluation of the precipitation is unfortunately only superficial. As only monthly mean precipitation is compared with the simulation results, the corresponding frequency distribution, i.e. the number, duration and intensity of the events cannot be determined. However, this is crucial for wet deposition, as a few short but intense rain events result in less deposition compared to longer precipitation events of average intensity. This will also substantially impact the precipitation adjustment (see comment below!).

Reply:

We agree that the evaluation of monthly mean precipitation is not appropriate in terms of the frequency distribution, number, duration, and intensity of rain events. However, the sampling interval of wet deposition is mostly daily over North and East Asia and weekly over Southeast Asia (Table 2). Additionally, the sampling periods of these measurements were not consistent across site. To allow for a consistent analysis period for the whole dataset, we believe that taking the monthly-mean precipitation is an appropriate approach and this analysis could provide a broad overview in Asia.

In response to your comment about the impact on precipitation adjustment, please see our reply to comment 8.

4. Analyzing the wet deposition of sulphur, M11 shows a substantially higher deposition pattern in China. What is the reason for this? This is a typical example for a model inter-comparison study where data is compared, but the causes for the differences are not analysed in detail. Has there been an issue with SO2 emissions or conversion from S(IV) to S(VI)? Is there a bias originating from seasalt sulphate? Is total Sulphur completely overestimated in this model? Or is it completely depleted, as wet deposition is so efficient? These differences require much more analysis for a consistent intercomparison study.

Reply:

This is related to our reply to comment 2. We have added a statement about the sulfur

oxidation ratio by referring to our companion paper (Tan et al., 2019).

5. Especially, when creating ensembles including such outliers, the ensemble mean can even be deteriorated compared to individual simulations. This does not appear to be the case in this study, as the M11 simulation compensates some of the low bias from the majority of the other simulations. A similar behaviour of overestimation is not as obvious for nitrate and ammonium, leading to the impression that this is not necessarily a consequence of the wet deposition scheme.

Reply:

We think that the overestimation of the wet deposition of N is also noticeable in model M11. For the wet deposition of S, model M11 was the only model that overestimated (all of the other models underestimated); however, in the case of the wet deposition of N, there was greater variation in the performance of all models. As a result, the ensemble mean for the wet deposition of N performed better than that of individual models.

6. The weighted ensemble might be a better option to reduce the importance of outliers; however, it simply states that the models which show best agreement with the observations should be used for the ensemble mean. Consequently, it reduces ensemble spread and therefore does not cover the whole range of simulation results properly. Please state explicitly, what you hope to gain from the weighted ensemble mean.

Reply:

In this weighted ensemble mean, we used R as the weighting factor. Therefore, we did not entirely eliminate the model that showed lower performance than observations.

Near the end of Section 4.2, we clearly state the superiority of the weighted ensemble mean as follows:

"In terms of NMB, ENS performed better than WENS; however, WENS could be regarded as a better approach because it takes into account each model performance evaluated by observation using R as the weighting factor and it showed better values than ENS in terms of NME, FAC2, and FAC3."

7. Concerning the total deposition maps, the authors should clearly point out, that underestimated wet deposition can often be compensated by overestimated dry deposition and vice versa, as both processes depend on the atmospheric burden (or near surface concentrations).

Reply:

We appreciate this constructive comment. We have added the following sentence in Section 4.1, accordingly:

"As we have seen (e.g., Fig. S1), the underestimation (overestimation) of wet deposition could be related to the overestimation (underestimation) of atmospheric concentration, and these could be found as dry deposition. The underestimation (overestimation) of wet deposition can be compensated by the overestimation (underestimation) of dry deposition, and may pose the similar total deposition amount. Therefore, this kind of study can give insights into the balance between dry and wet deposition."

8. I (personally) see the option of precipitation adjustment to improve the consistency of the simulation results with observations very critical. This adjustment does not include any kind of frequency distribution of precipitation events, the vertical extent of the precipitation (and hence the accessible fraction of the tracer vertical column for wet deposition). Also it does not include any kind of vertical redistribution by scavenging and subsequent evaporating precipitation and hence tracer release at lower altitude. Of course, I agree that with wrong precipitation amounts it will be impossible to fully match observations, but in my opinion not only the total amount of precipitation, but at least the central moments of the precipitation frequency distribution should be matched. As this correction is applied to the offline data, it could happen that an already strong precipitation event which might have a scavenging rate of 100% (i.e. all sulphate is already removed by the event) is supposed to remove even more sulphate (which is not available, as it is already depleted!). This is not discussed at all, implying that this precipitation adjustment is a useful measure to correct wet deposition for precipitation biases.

Reply:

We agree with your comments on the adjustment approach used in this study. To avoid overstating the usefulness of this approach, we have revised the introduction to this method in Section 4.3 as follows:

“Note that this precipitation-adjusted approach assumes that errors associated with the modeled precipitation are linearly related to errors in wet deposition amounts and the precipitation was adjusted by the total amount of observed precipitation; hence, the modeled convective and sub-grid-scale precipitation was not distinguished. Because current meteorological models have difficulty in capturing the timing of precipitation events, the application of this adjusted approach at a finer temporal resolution will lead to excessive adjustments (e.g., close to zero in the case that observed precipitation is zero or divergent in the case that the modeled precipitation is near zero).”

9. Overall, I think that this study could be published after addressing the points above, but GMD would have been the better journal.

Reply:

We have addressed the appropriateness of our submission for ACP rather than GMD in our reply to comment 1.