

We thank you for providing helpful and constructive comments and suggestions. We have revised our manuscript accordingly. We hope that these revisions satisfactorily address all the points you have raised. Our point-by-point responses are provided below, and revisions are indicated in blue in the revised manuscript.

General comments

This study presents an interesting analysis of modeled wet deposition in Southeast Asian and proposes an approach to improve wet deposition estimates using two precipitation datasets. The manuscript is well written and clearly structured, the figures and tables are of high quality, and the analysis methods and underlying model simulations are sound. My comments mainly focus on the interpretation of the results which I feel could be expanded and improved upon prior to publication.

Reply:

We appreciate this positive comment. To address your concerns related to the interpretation of the results, we have revised the manuscript according to your comments.

Specific comments

Page 4, lines 5 – 12: While the details of the model configurations have been published before, I think it would be useful to add a table summarizing some key aspects (model resolution, chemical mechanism, aerosol scheme, wet and dry deposition scheme) of the models used in this study. This would help in the interpretation of model performance, model-to-model variability, ensemble construction, and needs for future model development.

Reply:

Thank you for this helpful comment. We have included a new Table 1 to present descriptions (horizontal and vertical advection/diffusion, gas and aerosol chemistry, dry and wet deposition schemes, and boundary condition) of the seven models used in this study. We also added the relevant references in this new Table 1. As the introduction to this new

table, we have added the following sentences in Section 2.1 on P4, L16.

“In this study, seven models (M1, M2, M4, M5, M6, M11, and M12) that using the same meteorological fields simulated by the Weather Research and Forecasting (WRF) model version 3.4.1 (Skamarock et al., 2008) over the unified modeling domain were selected. The unified modeling domain covered the whole of Asia with a horizontal grid resolution of 45 km and 40 vertical layers from the surface up to 10 hPa. Descriptions of the seven models are listed in Table 1”

Page 4, line 17: My interpretation of Solazzo et al. (2012) is that ensemble approaches other than averaging over all available models with equal weights performed better than the ensemble mean. Were such other ensemble approaches explored in MICS-Asia Phase 3, especially given that 5 of the 7 models used in this study were CMAQ and may have had greater similarity than the other two models?

Reply:

Thank you for this insightful comment. In our overview paper (Itahashi et al., 2020), we tried to apply a simple ensemble average and weighted ensemble approach. To clarify this point, we have added the following sentences in Section 2.2 on P5, L10.

“Solazzo et al. (2012) proposed a method to produce a better ensemble. In the deposition analysis of MICS-Asia Phase III, a simple ensemble and a weighted ensemble were performed using the correlation coefficient (R) between the modeled and observed wet deposition (Itahashi et al., 2020). It was found that R was always improved by the weighted ensemble; however, biases can be worse in a weighted ensemble for some cases.”

Page 5, Lines 35 – 37: please clarify how the aggregation of weekly or ten day observations to monthly values was handled when a sampling period spanned across two months, and whether the same approach was used to calculate monthly model values in these instances.

Reply:

We appreciate this question. To explain the methodology used for monthly averaging, we have added the following sentences on P5, L27.

“The monthly accumulated wet deposition at each site were used for the model evaluation. For weekly or 10-day observational data, the central observation day was regarded to represent the corresponding month, and then the monthly accumulated wet deposition was calculated.”

Additionally, to clarify this information, the sampling intervals at each observation site have been added in Table 2.

For models, the monthly average was simply calculated based on the date. We have added this point on P5, L29.

“Meanwhile, the model results were simply calculated from the calendar date.”

Page 7, line 96: can you please state what criteria were used to define an “acceptable level”?

Reply:

Because there are no criteria to define this level, we have removed this point in the revised manuscript.

Page 8, lines 48 – 50: I think it would be really important to expand this discussion and provide a motivation why the precipitation-adjustment approach was viewed as the most appropriate and effective avenue for improving model performance. I agree that biases in precipitation are critical to consider when evaluating and trying to improve modeled wet deposition, but I would also like to see a discussion (maybe a summary of Itahashi et al., 2020) of other potential drivers of model biases, not just in wet deposition but also the actual concentrations of the compounds analyzed in this study. For example, what is known from other MICS-Asia work about whether some of the wet deposition performance issues might be caused by errors in the emissions of oxidized and reduced nitrogen and uncertainties in the representation of atmospheric chemistry in the models? The model-to-model variability indicated by the whiskers in Figures 1–7 is substantial and is not driven by precipitation since all models used the same WRF fields, yet no mention of this variability is made in sections 3.1.1 – 3.1.7 and its implications for improving the modeled wet deposition fields are not discussed in this section. An acknowledgment that precipitation errors are likely not the sole drivers of error in wet deposition should also be added to the abstract (page 2, lines, 39-40) and conclusions (page 13, lines

04-05).

Reply:

We are grateful for this insightful comment on our approach. As pointed out, we omitted the discussion of the ambient concentration itself, so we carefully revised our manuscript to mention this point. In terms of the model evaluation for ambient concentrations, our companion paper (Chen et al., 2019) overviewed the modeling performance of MICS-Asia Phase III. Moreover, the variation among the seven models has now been added in each discussion section. The additional discussions of the inter-model variability are as follows:

For Myanmar, in Section 3.1.1 on P6, L21:

“At the Yangon (No. 30) site, the model variation (shown by whiskers in Fig. 2) was small for the wet depositions of S, N, and A; this indicates that the overestimation during the dry season and underestimation during the wet season was common among all models.”

For Thailand, in Section 3.1.2 on P7, L16:

“Large inter-model variability in the modeled wet deposition was found in some months at Khanchanaburi (No. 34). This could be related to the difference in the ambient concentration and the difference in the mechanisms of the wet deposition scheme because all models used the same meteorological field. It should be noted that all models always showed a large wet deposition in February, March, and November, despite the observed zero wet deposition amount in these months (due to the lack of precipitation during the dry season). This suggests that the discrepancy in the simulated precipitation amount could be the cause of the inaccurate simulation of wet deposition”

For Cambodia, in Section 3.1.3 on P8, L1:

“All models commonly underestimated the wet deposition during the wet season.”

For Vietnam, in Section 3.1.4 on P8, L12:

“There were large inter-model differences when the precipitation was high. This result suggests that heavy rain events may lead to large inter-model variability in the simulated wet deposition, and the mechanisms should be further investigated. As concluded in the overview paper of Itahashi et al. (2020), this is one of the lessons learned in MICS-Asia

Phase III, and this will be addressed as part of the next MICS-Asia”

For the Philippines, in Section 3.1.5 on P8, L26:

“Because of this precipitation overestimation, the ENS also tended to overestimate the wet depositions of S, N, and A. Compared with other countries, the inter-model differences were larger for the sites in the Philippines. Further seeking of model wet deposition schemes focused on this region will be needed.”

For Malaysia, in Section 3.1.6 on P9, L15:

“This tendency was common as indicated by the model-to-model variability. At these two sites, observations showed a small wet deposition of N, and the balance between cations and anions should be carefully examined.”

and

“The inter-model variability was small; hence, this overestimation could be connected to the overestimation of precipitation.”

For Indonesia, in Section 3.1.7 on P10, L2:

“As was found in the Philippines, the inter-model variation was large, except for Maros (No. 54) , and further study focusing on this region will also be required.”

In the conclusion section on P16, L12:

“Additionally, greater inter-model variation was noted in the Philippines and Indonesia, especially during months with heavy precipitation. To investigate the differences on model wet deposition scheme, such heavy rainy events with finer spatio-temporal resolution should be pursued in the future MICS-Asia Phase IV.”

Regarding the modeling performance for ambient concentration, we added the following discussion in Section 4.1 on P10, L10:

“The errors in the simulated values of wet deposition are associated with ambient concentration and/or precipitation. Our previous overview paper (Itahashi et al., 2020) presented two approaches for improving the modeling of wet deposition, namely, the

ensemble approach and the precipitation-adjusted approach. The former approach was used in this study. In terms of the modeling performance for the ambient concentrations of aerosols of SO_4^{2-} , NO_3^- , and NH_4^+ , our companion paper reported better performance over Southeast Asia compared with North and East Asia (Chen et al., 2019). As noted in Section 3, the model generally overestimated precipitation as well as wet deposition during the dry season. Additionally, the model sometimes simulated non-zero precipitation, and consequently non-zero wet deposition, despite the absence of wet deposition due to the absence of precipitation. Based on these findings in MICS-Asia III, the difficulty stemmed from the inaccuracy of the modeled precipitation, which is fundamentally important for simulating the wet deposition. The precipitation-adjustment method is expected to improve model performance.”

This point was also added in the abstract to explain why the adjustment method was applied in this study. The additional sentences are as follows on P2, L4:

“Considering the model performance for ambient aerosol concentrations over Southeast Asia, this failure of models is considered to be related to the difficulty in capturing the precipitation in Southeast Asia, especially during the dry and wet seasons. Generally, meteorological field overestimated the precipitation during the dry season, which leads to the overestimation of wet deposition during this season. To overcome this, a precipitation-adjusted approach that scaled the modeled precipitation to the observed value was applied, and it was demonstrated that the model performance was improved.”

Page 9, lines 73 – 80: I like the idea of including both the EANET observations and TRMM satellite data in this analysis. What I find interesting and would like to see discussed in more detail is the fact that the squared correlation coefficient between EANET and TRMM data is only about 0.5, i.e. only about 50% of the precipitation variability seen in the point observations is captured by the satellite product. To the extent that these point observations (both for precipitation and wet deposition) are used to evaluate the model, what does this relatively low correlation say about the spatial scales represented by the three different datasets (observations, model simulations, and TRMM), their commensurability, and inferences drawn from differences between them?

Reply:

We are grateful for this helpful comment on the differences in precipitation. We have added the following point regarding the requirements for future studies on P11, L26.

“It should be noted that even though satellite and ground-based observations showed differences in the precipitation amount, this result indicates that further consideration of the how well precipitation is represented by the spatial resolution (broader observation by satellites and point-specific observations using ground-based monitoring) is important. Accordingly, the effect of the modeling spatial resolution on the simulated precipitation should be considered in future studies.”

Page 9, lines 82-83: I am not sure I follow this argument. If the EANET observations are the ground truth and TRMM has a correlation of only 0.7 with them, why would one consider the TRMM-based AS adjustment AS over AO? I realize that the conceptual benefit is that AS can be applied across the entire domain while AO is limited to specific stations as discussed at the beginning of section 4.3, but this is not the argument I'm reading here.

Reply:

The simulated wet deposition based on the original simulated precipitation shown in Figure 9 (left) has been evaluated in Figures 2–8. Based on the precipitation-adjustment via Eq. (5), the simulated wet deposition is revised. Satellite measurement compared to EANET ground observations shown in Figure 9 (right) obtained a more accurate precipitation amount rather than the simulation. The better estimation of precipitation itself will lead to a better simulation of wet deposition. In this sentence, we did not consider the better performance of AS compared to AO; rather, we only discussed the model improvements relative to the original simulation results.

To avoid misunderstanding, this point has been revised as follows on P11, L25.

“From this result, it is expected that precipitation-adjustment based on satellite measurements also has the potential to improve the original simulation of wet deposition.”

Page 13, lines 77 – 83, discussion of Figure 17. I recommend avoiding the terms “overestimation” and “underestimation” when discussing the spatial patterns of wet deposition rather than precipitation

results. I suggest to instead use terms like “higher values” and “lower values”. Aside from the EANET station analysis already presented in Figures 10 – 16, no observational data is available to judge whether the AS deposition patterns are higher or lower than reality.

Reply:

We agree with this suggestion. We have revised the manuscript to use “higher/lower values” for the discussion of Figure 17.

Page 13, lines 89 – 90: Since no actual country-wide observations are available aside from the EANET point measurements, I recommend avoiding terms like “reproducibility” and “accurate estimation” when discussing Figure 18.

Reply:

We have removed the use of “reproducibility” from the revised manuscript by deleting the redundant sentence. The expression “accurate estimation” was not used in relation to Figure 18, and we would like to retain the use of this expression elsewhere.

Technical corrections

Page 2, Line 40: change “scaling” to “scaled”

Reply:

We have corrected this.

Page 2, Line 42: suggesting changing “spatio-and-temporal” to “spatio-temporal”

Reply:

We have corrected this.

Page 3, Line 97: change “The participating model was requested” to “The participating models were requested”

Reply:

We have corrected this.

Page 6, line 79: suggest changing “calculated” to “modeled”

Reply:

We have corrected this.

Page 13, lines 01 – 02: please double check the wording of this sentence, it reads like a contradiction to me “Generally, the ensemble model can capture the observed wet deposition; however, the models failed to capture the wet deposition, even the ensemble mean ...”

Reply:

We appreciate this careful checking. We have revised this sentence as follows in P15, L22.

“Generally, the ensemble model could capture the observed wet deposition; however, sometimes failed to capture the wet deposition and obtained low correlations and/or large biases and errors.”