Reviewer3

[C3-1] In this study, the authors quantify the impact of a double-moment cloud microphysics scheme on aerosols in a high-resolution global model, comparing to its single-moment counterpart in the same model. The description of results and comparison between the two and with published results in the literature are quite comprehensive, but the explanations of differences are often handwaving. Particularly interesting result is the much higher ERFaci in NDW6 than in NSW6. In my opinion, more in-depth analyses in terms of how the two schemes treat aerosol indirect effects on clouds and precipitation in the model are needed. Also, what's the implication of such results for the use of double-moment microphysics in high-resolution climate models?

[A3-1] We appreciate your great contributions to improve our manuscript. Your comments and suggestions are very helpful and motivate us to investigate the results more deeply, especially the analysis of aerosol-cloud interaction in our models in Section 4.2 in the revised manuscript. We believe our revision is acceptable for you.
The comment of "what's the implication of such results for the use of double-moment microphysics in high-resolution climate models?" is also important point. The high-resolution is certainly important to simulate aerosols and clouds more realistically, as shown in this study and our previous study (Goto et al., 2020). The advantage of using double-moment cloud microphysics scheme is that the cloud microphysics representation of double-moment microphysics (NDW6) is more elaborate than that of single-moment microphysics (NSW6). In addition, this study showed that the NSW6-simulated cloud droplet radii for water clouds (CDR) was overestimated due to the inability to predict cloud droplet number concentrations in NSW6. These points were discussed in section 4.2 and mentioned in section 5 in the revised manuscript.
Through the revision, we modified figures and tables as follows:

- Figures 1, 2, 5, E1 and E2: We changed the color of the zonal averages (NDW6 blue, NSW6 orange, as in the other figures).
- Figures 2 and 5: We replotted the model results in white for grids with missing satellite data. We replotted the zonal averages of the model results by eliminating the grids with missing satellite data.
- Figure 4: We changed the caption named "references" to "AeroCom".
- Figure 7: We changed the subtitle named "AOD" to "AOT".
- Figure 9: We provided IRFari of all and each aerosol for shortwave & longwave at the TOA & the surface under all & clear sky conditions. We changed the caption named "references" to "Kinne19" and "Thorsen21".
- Figure 10: We removed the results of IRFari because they were shown in Figure 9. Instead, we newly added the results of ERFari and sum of ERFari and ERFaci. We also modified the ERFaci for shortwave to net ERFaci (for both shortwave and longwave).
- Figure 11: We modified ∂AOT to ∂CCN. We added new parameters such as ∂CDNC, ∂CDR, ∂CA, ∂CF, and net ERFaci to further explore ACI.
- Figure 12: We replaced Table 3 in the original manuscript to Figure 12 in the revised manuscript by adding relevant parameters such as ∂CCN, ∂CDNC, ∂CDR, ∂CA, ∂CF, and net ERFaci.
- Figure 13: To explain possible overestimations of the NSW6-simulated Twomey effect, we newly plotted global budgets of the annual averages of the NDW6- and NSW6-simulated Qc (mixing ratio of cloud droplets) and CDNC (cloud droplet number

concentrations).

- Table 1 in the original manuscript: We removed it and added a paragraph to explain the HRM and LRM as references in section 2.5 in the revised manuscript.
- Table 1 in the revised manuscript: We simply moved Table 2 in the original manuscript to Table 1 in the revised manuscript.
- Table 2 in the revised manuscript: We showed global and annual mean values of ERFari, ERFaci, and the sum of ERFari and ERFaci for shortwave, longwave, and net radiation under both all-sky and clear-sky conditions.
- Table A1: We newly added the statistical metric to compare results in this study with the references by Goto et al. (2020).
- Table A2: We simply moved Table A1 in the original manuscript to Table A2 in the revised manuscript, with two exceptions. One, we changed "References" to "References from model results". Second, we changed the SO_2 production value from 67.5 to 67.7.

Please note that some English was corrected in the revised manuscript.

[C3-2] It's hard to find specific information regarding the setup of experiments (NDW6, NSW6, HRM, LRM), except for the resolutions and microphysics schemes in Table 1. Which specific simulation years? What are the aerosol emissions used for the simulation years, mean or year-specific? Such information is important to determine whether the comparison of simulations results with observations and other models in the literature is valid. Please include the details in Table 1.

[A3-2] Thank you for your comment on the setup of experiments. As you pointed, some of explanation in the original manuscript was missed. So, we added section 2.3 named "Experimental conditions" including the information of the setup and emission inventories to the revised manuscript (Lines 201-216) as follows:

"All experiments with both NDW6 and NSW6 are carried out for 6-years after the 1-month spin-up calculation. The simulation results are climatological runs, because the model does not nudge meteorological fields such as wind and temperatures but nudges the sea surface temperature (SST) and sea ice by the results of the NICAM from Kodama et al. (2015). The initial conditions for the model spin-up are obtained from the end of the 1-year aerosol simulations coupled to NSW6 without nudging the meteorological fields under the present era.

The emission fluxes used in this study are the Hemispheric Transport of Air Pollution (HTAP)-v2.2 (Janssen-Maenhout et al., 2015) for BC, organic carbon (OC) and SO_2 from anthropogenic sources in 2010 and the Global Fire Emission Database (GFED) version 4 (van der Werf et al., 2017) for BC, OC and SO_2 from biomass burning in climatological average from 2005 to 2014. The ratio of OC to OM is set at 1.6 for anthropogenic activities and 2.6 for biomass burning (Tsigaridis et al., 2014). Secondary organic aerosols (SOAs) are assumed to form particles, which are calculated by multiplying the emission fluxes of isoprene and terpenes provided by the Global Emissions Initiative (GEIA) (Guenther et al., 1990) using constant factors. SO_2 is emitted from volcanic eruptions (Diehl et al., 2012) and is also formed from DMS, which is interactively emitted in the aerosol module (Bates et al., 1987). Sulfate is formed from SO_2 oxidation with a 3-dimensional distribution of monthly oxidants (ozone, H_2O_2 and OH) provided by a chemical transport model (CHASER) coupled to MIROC (Sudo et al., 2002). Emission

fluxes for dust (Takemura et al., 2009) and sea salt (Monahan et al., 1986) are interactively calculated in the model using mainly the wind speed at a height of 10 m."
In addition, we added a section 2.5 in the revised manuscript to explain the information of the models, i.e., HRM and LRM in Goto et al (2020) as references. Because we added this section to the revised manuscript, we removed Table 1 in the original manuscript. The section 2.5 in the revised manuscript (Lines 281-295) is as follows:
"Our previous model results provided in Goto et al. (2020) using NICAM.16 at a global 14-km high resolution (hereafter referred to as the HRM) and a global 56-km low resolution (hereafter referred to as the LRM) are used as references to compare the NICAM results. As mentioned in section 2.1, the number of vertical layers is set at 38, and the timestep is 1 minute in both the HRM and LRM. The integration periods in both the HRM and LRM are 3 years as climatological runs. The emission inventories, i.e., 2010 for anthropogenic sources, climatological average in 2005-2014 for biomass burning, and natural sources in the present era, and the nudged SST and sea ice in this study are identical to those in both the HRM and LRM, but the initial conditions in this study are different from those in both the HRM and LRM, which use the model results at the end of December after a 1.5-month spin-up. The initial conditions for the model spin-up are prepared by the reanalysis datasets of the National Centers for Environmental Prediction (NCEP) Final (FNL) (Kalnay et al., 1996) in November 2011. In the cloud microphysics and autoconversion modules, NDW6 coupled to Seifert and Beheng (2006) and NSW6 coupled to Khairoutdinov and Kogan (2000) are used in this study, whereas NSW6 coupled to Berry (1967) is used in both the HRM and LRM. The improvement in the aerosol module described in section 2.2 is also different from that in the HRM and LRM. The results of the HRM and LRM are useful for evaluating the current model results because the observations are limited in some parameters, such as aerosol global budgets and radiative forcings."

[C3-3] At many places, IRFari and ERFaci are referred to as "shortwave" aerosol forcing. If the longwave component is not considered at all, I don't think they are comparable to the cited values in the literature, which mostly include both shortwave and longwave components and are referred to as net radiative forcing. Please confirm and clarify.

[A3-3] Thank you for your suggestion. Surely, we didn't show the results of IRF and ERF for longwave as well as net. We put these estimates in the revised Figure 10 and Table 2. We modified Figure 10 and added net ERFaci values to the revised manuscript.

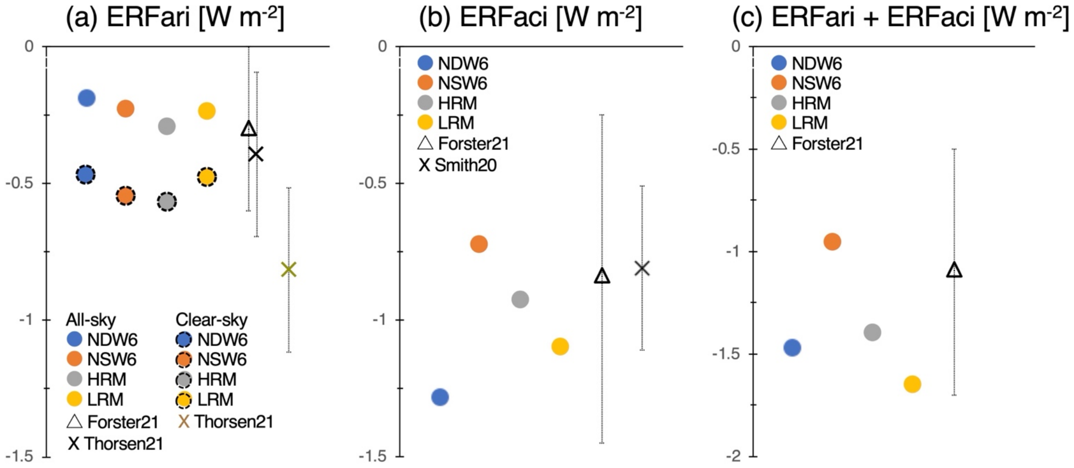

Figure 10: Global and annual mean values of (a) effective radiative forcing for anthropogenic aerosol-radiation interaction (ERFari) for shortwave and net (sum of shortwave and longwave) radiation, (b) ERFaci for anthropogenic aerosol-cloud interaction, and (c) the net ERF (sum of ERFari and ERFaci). All units are in W m^{-2}. In ERFari, the reference of Forster21 is estimated in the net radiation by IPCC-AR6 or Forster et al. (2021), whereas the reference of Thortsen21 is estimated in the shortwave radiation by Thorsen et al. (2021). The reference for Smith20 is Smith et al. (2020). The values are also listed in Table 2.

Table 2: Global and annual mean values of ERFari for anthropogenic aerosol, ERFaci for anthropogenic aerosol-, and the net ERF (sum of ERFari and ERFaci) for shortwave, longwave, and net (sum of shortwave and longwave) radiation under the all-sky abnd clear-sky conditions. All units are in W m^{-2}.

	ERFari under the all-sky conditions			
	NDW6	NSW6	HRM	LRM
Shortwave	-0.22	-0.26	-0.33	-0.26
Longwave	0.03	0.03	0.04	0.02
Net	-0.19	-0.23	-0.29	-0.24
	ERFari under the clear-sky conditions			
Shortwave	-0.52	-0.60	-0.63	-0.51
Longwave	0.05	0.05	0.05	0.03
Net	-0.47	-0.55	-0.57	-0.48
	ERFaci			
Shortwave	-1.34	-0.63	-0.81	-1.17
Longwave	0.06	-0.10	-0.12	0.07
Net	-1.28	-0.73	-0.93	-1.10
	ERFari+ERFaci			
Shortwave	-1.56	-0.89	-1.15	-1.43
Longwave	0.09	-0.07	-0.08	0.09
Net	-1.47	-0.96	-1.23	-1.34

[C3-4] Both NDW6 and NSW6 appear to simulate a too weak SWCRF and LWCRF (shown in Table 2), as compared to observations and historical mean of CMIP models. Is this due to interannual variability or model bias? How does this affect the evaluation of aerosols and their associated forcings in NDW6? Please include a discussion on this issue.

[A3-4] Thank you for your comment on CRF. We checked the results of CRF in all six years and found that the standard deviations were smaller than the difference between NDW6 and CERES. The standard deviations were calculated to be 0.2 Wm^{-2} (NDW6, Annual), 0.6 Wm^{-2} (NDW6, January), and 0.2 Wm^{-2} (NDW6, July). Therefore, the difference between NICAM and CERES is not caused by the internal variability.
In the beginning of Section 4.2, we roughly mentioned the results of SWCRF and the bias, and discussed the spatial and zonal distributions in Appendix E. As mentioned in Appendix A, the SWCRF results are generally consistent to the LWP results. As mentioned there and in section 3.1 related to LWP, the global averages in NDW6 are not closer to the observation compared to those in NSW6, but this can be caused by compensation errors in space (In the original manuscript, we mentioned the error in space and time, but after checking again, we found that it was caused by the error only in space. We modified this in the revised manuscript (Lines 328-329, 340, 515, 621, and 659)). The NICAM results of both SWCRF and LWP are underestimated compared to the observation in most regions, but the NSW6 has some positive bias of SWCRF and LWP in some regions such as the western Pacific Ocean. As a result, the zonal and global averages of SWCRF and LWP in NDW6 are underestimated relative to NSW6. Therefore, we cannot conclude that the bias of SWCRF and LWP in NDW6 is larger than that in NSW6.
The large underestimation of SWCRF in the western Pacific Ocean is related to the underestimation of the clouds. The underestimation of the low-level clouds in NICAM were realized in NICAM studies (e.g., Kodama et al., 2021). Since the NICAM-simulated precipitation is close to the observation, the NICAM-simulated cloud-to-precipitation conversion is overestimated, as shown in Figure 5. This causes the overestimation of the wet deposition for aerosols as shown in Figure 7; therefore, the underestimation of the NICAM-simulated SWCRF can cause the underestimation of the simulated aerosols in such tropic regions. This is a possible bias to be caused by the difference in the SWCRF between the model and observation.
Over land where the observed LWP information is very limited, the observed SWCRF may be useful to evaluate the model. In high aerosol loading areas such as China, Europe, and the United State, the simulated SWCRF tends to be underestimated compared to the observation. This indicates the underestimation of the simulated LWP and/or the overestimation of the simulated cloud droplet effective radius (CDR). When the simulated LWP is underestimated, the simulated aerosols are underestimated, according to the above discussion. When the simulated CDR is overestimated, the simulated CCN is underestimated. This is consistent to the underestimation of the simulated aerosol. Therefore, if negative biases in the simulated SWCRF are eliminated, the simulated aerosols will increase. In the revised manuscript (Lines 516-522), we modified this as follows:
"The NDW6-estimated SWCRF values are concluded to be better than the NSW6 results, but the underestimation of the simulated SWCRF can have an impact on the aerosol simulations. The underestimation of the simulated SWCRF indicates the underestimation of the simulated LWP and/or the overestimation of the simulated CDR. When the

simulated LWP is underestimated, the simulated aerosols are also underestimated because the simulated precipitation is generally comparable to the observations in this study, as shown in Figure 1. When the simulated CDR is overestimated, the simulated CCN must be underestimated. This is consistent with the underestimation of the simulated aerosol. Therefore, if the negative biases in the simulated SWCRF are eliminated, the simulated aerosols will increase."

As for longwave CRF, the impacts of the longwave CRF (LWCRF) on aerosols are small in this study, because the LWCRF is mainly determined from the high-level clouds and convective clouds and our model ignored the direct interaction between aerosols and ice crystal (as an ice nuclei). Therefore, if negative biases in the simulated LWCRF in this study are eliminated, the simulated high-level clouds increase but the simulated aerosols may not be changed. Because the direct interaction between aerosols and ice clouds is not considered in this model, the impacts of the bias of the simulated LWCRF are unclear. In the revised manuscript (Lines 523-525), we modified this as follows:

"The underestimation of the simulated LWCRF is caused by the underestimation of the simulated high-level clouds, but the impacts of this negative biases in the simulated LWCRF on the aerosol simulations are unclear due to ignorance of the interaction between aerosols and ice crystals (as ice nuclei) in this model."

Minor comments and technical corrections:

[C3-5] L118-119: If CDNC is fully prognostic in the NDW6 double-moment scheme, I assume it's always updated with source and sink tendencies. Why is there an additional constraint by CCN that depends on supersaturation as well? please clarify.

[A3-5] We apologize for the lack of explanation about this part "In addition, a CDNC value is assumed to be updated to a CCN value only when the CCN value exceeds the CDNC value in a grid box". The CCN value only when the CCN value exceeds the CDNC value in a grid box is an aerosol activation process and a source term. The sink tendencies are accretion, autoconversion, and evaporation for water clouds. As you mentioned, the CDNC is determine by source and sink tendencies.

The additional constraint by CCN that depends on supersaturation is needed to nucleate water clouds (not to nucleate ice clouds). One reason of this is an aerosol activation process parameterized by Abdul-Razzak and Ghan (2000) is applicable only for water clouds but calculate in any clouds even below supersaturation in the aerosol physic model.

Therefore, we modified this in the revised manuscript (Lines 122-126) as follows:

"In addition, a source term of CDNC is assumed to be updated to a CCN value only when the CCN value exceeds the CDNC value in a grid box. The CDNC is updated with source (aerosol activation) and sink (autoconversion, accretion, and evaporation for water clouds) in NDW6 (Seiki and Nakajima, 2014). The balance of source and sink tendencies determines the CDNC in NDW6."

[C3-6] L123: Please clarify on "which" is updated in this study. NDW6 or NSW6?

[A3-6] Thanks you for your comment. This is NSW6. To avoid unclear expression, we removed the part 'which is updated in this study (cf.., Seiki and Roh, 2020)" from the revised manuscript.

[C3-7] L132: Is there no shallow convection parameterization at 14-km grid spacing? Please justify.

[A3-7] We also do not use shallow convection schemes, like other studies using NICAM with 14-km grid spacing (e.g., Satoh et al., 2010; Kodama et al., 2021). We modified this in the revised manuscript (Lines 143-144) as follows:
"As in previous studies using the NICAM (e.g., Satoh et al., 2010; Kodama et al., 2021), no parameterization schemes for deep and shallow convection are used in this study."
We realized possible biases caused by not including a shallow convection parameterization. For example, the simulated low-level clouds and shortwave radiation such as OSR and SWCRF in NICAM tend to be underestimated compared to the observation (Kodama et al., 2021). This can also influence SST and is very important for atmosphere-ocean coupling models to predict SST. For example, Masunaga et al. (2023) introduced a flux adjustment to the atmosphere-ocean coupled NICAM without shallow convection parameterization to avoid SST drift. However, we don't think the exclusion of the shallow convection parameterization is so important for atmospheric circulation models with fixed SST. Therefore, we don't use any shallow convection parameterization like previous studies using the NICAM as an atmospheric circulation model (e.g., Satoh et al., 2010; Kodama et al., 2021).

Reference:
Masunaga, R., Miyakawa, T., Kawasaki, T., Yashiro, H.: Flux adjustment on seasonal-scale sea surface temperature drift in NICOCO, Journal of the Meteorological Society of Japan. Ser. II, 101(3), 175-189, doi:10.2151/jmsj.2023-010, 2023

[C3-8] L157-159: This might be relevant to the comment for L118-119. This treatment needs more clarification and justification. If the aerosol scheme and the NDW6 are fully coupled, I don't see why this constraint is justified for certain conditions only. CCN number, which is only meaningful with supersaturation specified, can be larger than CDNC before the activation tendency is updated to CDNC. Otherwise, CCN diagnosed from interstitial aerosols should be mostly smaller than CDNC in clouds. I wonder whether partial-grid clouds matter here.

[A3-8]
We apologize for confusing you due to the lack of explanation about this. Like our answer in [A3-5], we would like to mention the source tendency of the CDNC. Therefore, we modified this in the revised manuscript (Lines 154-156) as follows:
"First, when the CCN number concentration is higher than the CDNC calculated online in the aerosol module, the value of water supersaturation is positive, and the atmospheric pressure is above 300 hPa, the CCN number concentration becomes an input of source tendency for CDNC."

[C3-9] L174: Change anthropogenic "materials" to "sources" or "activities".

[A3-9] Thanks. Changed.

[C3-10] L207: Does that mean there is no BC, OC and SO2 emission in the extra experiment except for fire and volcanos? I wonder how this experiment is different from the preindustrial-condition experiment, which should also have emissions in the anthropogenic sectors (e.g., residential, agricultural waste burning, etc.)

[A3-10] Thank you for your comment on the experimental setup in preindustrial condition in this study. Yes, as you said, our condition is slightly different from the preindustrial condition in the other studies, because the other studies use the 1850 (or 1750) aerosol emissions. As you mentioned, the residential sector in the preindustrial time has a large contribution to the total amount (Hoesly et al., 2018). This point is important but was not discussed in the original manuscript. We wanted to recalculate the simulations using the 1850 aerosol emissions, but unfortunately unable to do them due to the limitation of our computer resources. So, we would like to speculate the uncertainty of the ERF caused by the different assumption of the preindustrial era. As a reference, the results in Hoesly et al. (2018) are shown in Table. We added the following discussions to the revised manuscript (Lines 217-227):
"In the preindustrial experiments, the anthropogenic emission fluxes of BC, OC and SO_2 are assumed to be zero in this study. Hoesly et al. (2018) estimated global averages of the differences in the emission amounts of anthropogenic sources between 1850 and 2010 to be 2.1% (sulfate), 12.0% (BC), and 22.7% (OC). The residential sector has the largest contribution to the total anthropogenic emissions in the preindustrial era. Takemura (2020) calculated the IRFari due to anthropogenic sulfate under the conditions of 0% and 30% of the present emissions and found that the difference in the IRFari was within 0.03 Wm^{-2}. Therefore, differences in the assumptions for the preindustrial era between this study and other studies, such as IPCC-AR6 (Szopa et al., 2021), will result in a difference in the IRFari due to anthropogenic sources of at most 0.05 Wm^{-2}. Takemura (2020) also calculated ERFari and ERFaci due to anthropogenic sulfate under the conditions of 0% and 30% of the present emissions and found that the difference in ERFari plus ERFaci was within 0.2 Wm^{-2}. These are possible uncertainties in the estimated radiative forcings due to anthropogenic sources in this study, but these magnitudes are smaller than the difference between NDW6 and NSW6 in this study, as shown in section 4."
We also added the following sentence to the summary (Lines 651-657) in the revised manuscript:
"As mentioned in section 2.3, the assumption of the preindustrial conditions for aerosols can cause possible differences in the aerosol radiative forcing due to the anthropogenic sources between this study and other studies, such as IPCC-AR6 (Szopa et al., 2021). This study assumes that the anthropogenic emission fluxes of BC, OC and SO_2 are zero in the preindustrial conditions, whereas other studies often use them in 1750 or 1850 provided by Hoesly et al. (2018). Using the results of MIROC by Takemura (2020), the possible difference in the aerosol radiative forcing due to the anthropogenic source will be at most 0.05 Wm^{-2} (IRFari) and 0.2 Wm^{-2} (ERFari plus ERFaci)."

Table Emission amounts in 1850 and 2010 years by Hoesly et al. (2018)

Anthropogenic Source	1850	2010	2010-1850 [%]
SO_2 [ktSO_2]	2,481	115,487	2.1
BC [ktC]	934	7,755	12.0
OC [ktC]	4,262	18,755	22.7

Reference:

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750-2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), Geosci. Model Dev., 11, 369-408, doi:10.5194/gmd-11-369-2018, 2018.

[C3-11] L247-249: the use of "products" and "results" for model simulations and satellite, respectively, should be the other way around, i.e., referred to as simulation results and satellite products.

[A3-11] Thanks. Changed.

[C3-12] L261: What does the "horizontal biases" mean? Spatial or regional biases?

[A3-12] Thanks for your comment. Yes, this "horizontal" means "spatial" or "regional". We removed "horizontal biases" from the revised manuscript. Related to this comment, we found that the compensation errors were found only in the horizontal distribution (not found in seasonal distribution). Therefore, we modified the analysis of the differences in the LWP between NDW6 and NSW6 to the revised manuscript (Lines 318-329) as follows:
"In the tropics where the LWP is larger than the other areas, the NDW6-simulated LWP is lower and not closer to the MAC results than the NSW6-simulated LWP. Notably, the MAC results contain regional biases of up to 25%, especially in the tropics (Elsaesser et al., 2017), but even with the largest errors, the NDW6- and NSW6-simulated LWPs in the tropics are still underestimated compared to the MAC results. In the horizontal distribution over the eastern Pacific Ocean and Southern Atlantic Ocean at lower latitudes (30°S-0), the NDW6-simulated LWP is lower than the NSW6 results but comparable to the MAC results. However, over the western Pacific Ocean and Indian Ocean at the lower latitudes, both NDW6- and NSW6-simulated LWPs are lower than the MAC results. Therefore, the overestimation of the NSW6-simulated LWP in the eastern Pacific Ocean and Southern Atlantic Ocean effectively balanced the underestimation of the zonal averages of the simulated LWP and unexpectedly led to zonal LWP values closer to the MAC results. This situation also occurs in the northern hemisphere at lower latitudes (30°N-0). Therefore, in the lower latitudes (30°S-30°N), the zonal averages of the NSW6-simulated LWP look closer to the MAC results, but this is attributed to the compensation errors in the regional distribution. As a result, the global and annual mean values of the NSW6-simulated LWP appear closer to the MAC results."

[C3-13] Figure 7: "AOD" is used in the figure labels, but "AOT" is used in the figure caption, tables (e.g., Table 3), and the main text. Please make them all consistent.

[A3-13] We use 'AOT' in this manuscript, so modified Figure 7.