Response to Reviewer #2's Comments:

Bida Jian et al. (Author)

We are very grateful for the Reviewer #2' detailed comments and suggestions, which help us improve this paper significantly. Based on the comments and suggestions from the editor and two Reviewers, we reorganize the datasets and methods, results and conclusion sections, and add some interpretations in each section in order to make the manuscript clearer. In addition, some superfluous information in each section is deleted.

Please see our point-by-point reply to comments. In addition, all revisions were highlighted in revised manuscript by using track changes.

General responses:

1. **Insufficient consideration of meteorological factors.** Only two meteorological factors are considered in this study: omega900 and RH850. However, subtropical stratocumulus (including cloud optical depth/thickness, LWP, and cloud droplet effective radius) are impacted by several other important meteorological factors, including sea surface temperature, estimated inversion strength, horizontal surface temperature advection, and wind speed (e.g. Fuchs et al. 2018, Scott et al. 2020, and references therein). Therefore, the consideration of only omega900 and RH850 in this study is inadequate. The omission of the other meteorological factors noted above may in fact greatly affect the results of Section 3.2, "The impacts of different aerosol types and meteorological factors on cloud albedo changes", due to confounding effects. The authors state that "If the correlation between the cloud albedo and a [predictor] candidate is significant at a 90% confidence level, the variable was considered as a predictor factor." But which candidates were considered? Inversion strength and advection have been shown to be the dominant meteorological controls on interannual changes in cloud optical depth in eastern ocean stratocumulus regions (Scott. et al. 2020). Therefore, I find it hard to believe that these two cloud-controlling factors are not significantly correlated with cloud albedo. Chen et al. (2014) investigated the effects of aerosols on marine warm clouds using observations. They found that the response of LWP to aerosol loading strongly depends on lower tropospheric stability and free-tropospheric moisture. This is additional evidence that

the omission of several meteorological factors, especially the inversion strength, is a crucial oversight in the present study. Finally, the choices of 900 hPa and 850 hPa as vertical levels for omega and RH are not justified. These levels are not external to the boundary layer. Omega700 (or 500) and RH700 (or 500) should be used instead, as is standard, since they represent free tropospheric vertical velocity and humidity. The authors should use these vertical levels instead, unless they can provide a compelling justification for their unusual choice of vertical levels.

Response: We are very grateful for the detailed comments and suggestions from reviewer. Indeed, the consideration of meteorological factors in this investigation is insufficient. In the revised manuscript, the relative humidity at 700hPa (RH700)， vertical velocity at 900hPa and 700hPa (omega900 and omega700), estimated inversion strength (EIS) and horizontal temperature advection at the surface (SSTadv) are added in the multilinear regression model. Vertical winds below the clouds can affect the exchange between the clouds and the air layer below the clouds, allowing more aerosols to enter the clouds, which serve as CCN, and causing more cloud droplets (Yang et al., 2019). In the study, aerosol mass concentrations at the 910hPa level are employed. Here, the omega900 factor is retained to assess the effect of vertical velocity under the cloud on the cloud albedo. We also add detailed discussions in the revised manuscript. Please see the section 2 and section 3.

Yang, Y., Zhao, C. F., Dong, X. B., Fan, G. C., Zhou, Y. Q., Wang, Y., Zhao, L. J., Lv, F., and Yan, F.: Toward understanding the process-level impacts of aerosols on microphysical properties of shallow cumulus cloud using aircraft observations, Atmos. Res., 221, 27-33, https://doi.org/10.1016/j.atmosres.2019.01.027, 2019.

2. **Lack of analysis of satellite simulator output.** Modern analyses of cloud fraction in GCMs should incorporate at least some analysis of satellite simulator output, such as ISCCP simulator cloud fraction. However, the current study compares the raw GCM cloud fraction with satellite cloud fraction, which is a somewhat outdated approach. Some of the differences between GCMs and the observations found in the paper may be due to different definitions of cloud fraction. The authors do note that "this study… employed the total cloud fractions as

there are no available MODIS simulator outputs for CMIP6." However, ISCCP simulator output is available for several CMIP5 and CMIP6 models. MODIS cloud fraction is more comparable to ISCCP simulator cloud fraction than it is to the raw GCM cloud fraction.

Response: We very thank reviewer for providing detailed comments and suggestions. Indeed, the ISCCP simulator output is available for several CMIP5 and CMIP6 models. But the ISCCP output can only provide the monthly mean cloud fraction (variable name: cltisccp) but not corresponding monthly mean shortwave flux. This makes it impossible to calculate the corresponding cloud albedo based on the ISCCP simulator outputs. We tried to compare the AMIP6 cloud fraction with ISCCP simulator cloud fraction, and compared it with satellite observations (See Figure R1). From the correlation coefficients, the results of ISCCP simulator outputs did not outperform the AMIP outputs. There is a good agreement between the ISCCP simulator outputs and AMIP outputs (R3>0.8). In addition, we compared the ISCCP simulator cloud fraction with ISCCP observed cloud fraction (See Figure R2). However, the performance of ISCCP simulator is poor to reproduce the ISCCP observed cloud fraction. Considering that the AMIP outputs have complete corresponding shortwave flux data, we still use the GCM cloud fraction in this study.

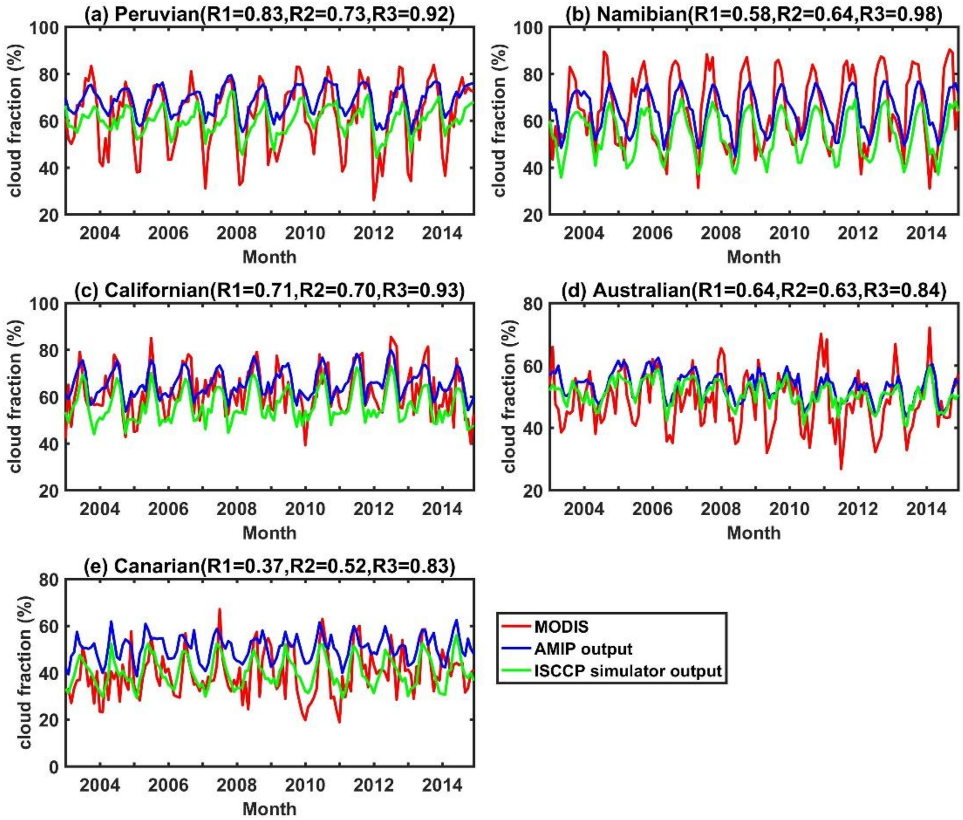

Figure R1: Monthly mean time series of cloud fraction from AMIP6 MEM outputs and ISCCP simulator outputs MEM during 2003-2014 compared with satellite observations (MODIS), over the (a) Peruvian, (b) Namibian (c) Californian, (d) Australian and (e) Canarian regions. The R1 and R2 indicate the temporal correlations between satellite observations and AMIP6 outputs and ISCCP simulator outputs, respectively. The R3 indicates the temporal correlations between AMIP6 outputs and ISCCP simulator outputs. Here, the ISCCP MEM is calculated based on 11 model outputs (see the Table R1 below).

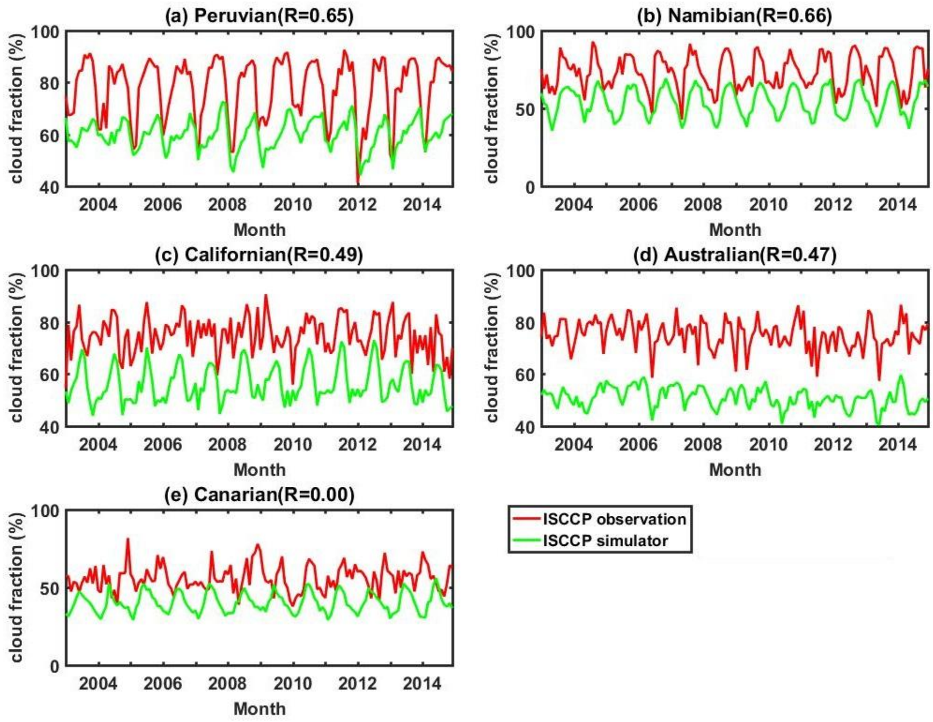

Figure R2: Monthly mean time series of cloud fraction from ISCCP simulator outputs MEM (green lines) during 2003-2014 compared with satellite observations (ISCCP, red lines), over the (a) Peruvian, (b) Namibian (c) Californian, (d) Australian and (e) Canarian regions. The R indicate the temporal correlations between satellite observations and ISCCP simulator outputs.

Table R1: The list of CMIP6 models with ISCCP simulator outputs.

	Model name	Origin
1	BCC-CSM2-MR	Beijing Climate Center, China
2	CESM2	National Center for Atmospheric Research, Climate and Global Dynamics Laboratory, USA
3	CanESM5	Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada, Canada
4	E2SM-1-0	LLNL, ANL, BNL, LANL, LBNL, ORNL, PNNL and SNL, USA
5	GFDL-CM4	National Oceanic and Atmospheric Administration, Geophysical Fluid Dynamics Laboratory, USA
6	GISS-E2-1-G	NASA/Goddard Institute for Space Studies, USA
7	IPSL-CM6A-LR	Institut Pierre Simon Laplace, France
8	MIROC6	AORI, NIES and JAMSTEC, Japan

9	MRI-ESM2-0	Meteorological Research Institute, Japan
10	NorESM2-LM	Norwegian Climate Centre, Norway
11	TaiESM1	Research Center for Environmental Changes, Academia Sinica, Taiwan

3. **Lack of analysis of low-level cloud fraction.** The authors should verify that their key observational results are valid for low-level cloud fraction. The regions chosen are dominated by low clouds, but high cloud variability may impact some of the results.

   Response: We agree with reviewer. High clouds are a source of uncertainty in this study. However, based on the available data, we cannot exclude its effect and analyze the low cloud albedo independently. As a reference, we used the daytime low-level cloud fraction from CERES SSF1deg product to verify the linear relationship between cloud fraction and planetary albedo (see Figure R3). However, in most regions, low-level cloud fraction didn't reproduce the same linear relationship as total cloud fraction. Charlson et al., (2007) used lidar reflectivity as a proxy for albedo (lidar albedo). They found lidar profiles with high clouds will distort the relationship between integrated attenuated lidar backscatter and low clouds. Fortunately, Bender et al., (2011) quantified the monthly mean regional-scale albedo of marine sratiform clouds based on MODIS and CALIPSO satellite observations, and found that the CALIPSO estimated cloud albedo (excluding the number of profiles with high cloud) is considerable consistent with that of MODIS. Here, the MODIS provided cloud fraction is total cloud fraction. For MODIS, the overlying high clouds didn't contaminate the linear relation. This means that the high clouds have little effect on the calculation of cloud albedo. Considering the consistency of both datasets, we still consider the estimated cloud albedo to be sufficiently representative of the cloud albedo in these stratocumulus regions.

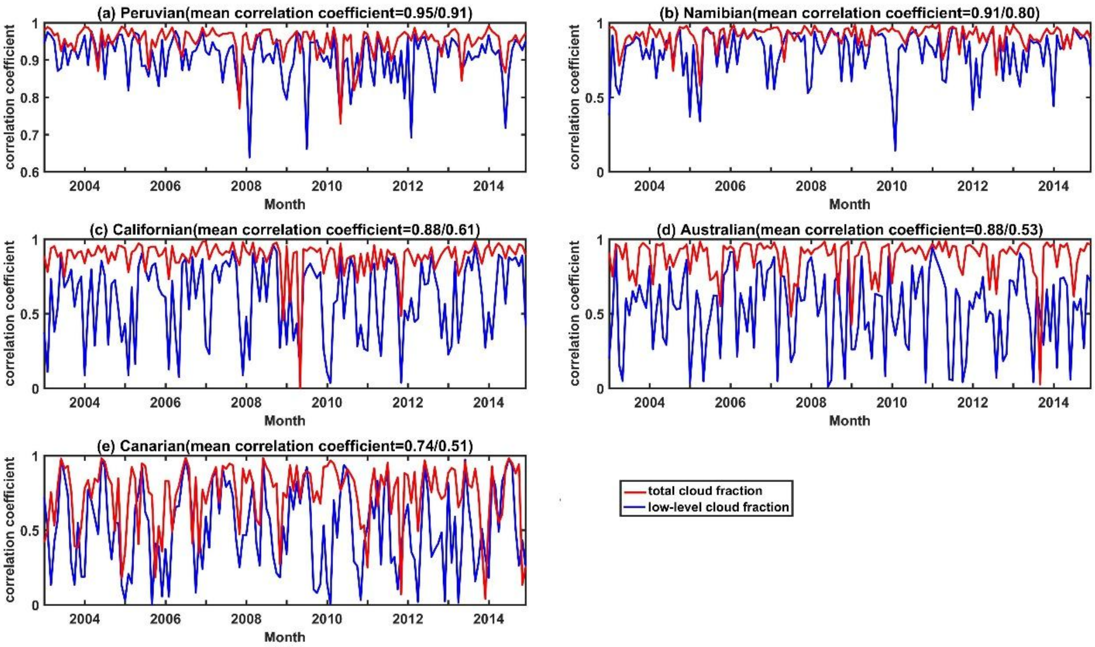

Figure R3: Monthly mean time series of correlation coefficients between total cloud fraction (red lines) and low-level cloud fraction (blue lines) and planetary albedo from CERES observations during 2003-2014, over the (a) Peruvian, (b) Namibian (c) Californian, (d) Australian and (e) Canarian regions. The monthly mean correlation coefficients are given in the title. The left side of the slash ('/') is the monthly mean correlation coefficients of the total cloud fraction, and the left side of the slash ('/') is that of low-level cloud fraction.

Bender, F. A. M., Charlson, R. J., Ekman, A. M. L., and Leahy, L. V.: Quantification of Monthly Mean Regional-Scale Albedo of Marine Stratiform Clouds in Satellite Observations and GCMs, J. Appl. Meteorol. Clim., 50, 2139-2148, https://doi.org/10.1175/jamc-d-11-049.1, 2011.

Charlson, R.J., Ackerman, A.S., Bender, F.A. M., Anderson, T.L. and Liu, Z.: On the climate forcing consequences of the albedo continuum between cloudy and clear air, Tellus B, 59, 715-727, https://doi.org/10.1111/j.1600-0889.2007.00297.x, 2007.

4. **Lack of verification of results with additional observational data.** MODIS is the state-of-the-art passive satellite cloud dataset, but, given that multi-linear regression can be sensitive to the input data, the authors should examine additional satellite data (such as ISCCP cloud fraction and the Multisensor Advanced Climatology of Liquid Water Path [Elsaesser et

al. (2017)]) to corroborate their results and establish robustness. Additional reanalyses should be considered as well for the meteorological data. ERA5 is considered to be the most state-of-the-art reanalysis.

Response: Thanks for your suggestions. To verify the sensitive of the results to input data, we employ different datasets to perform the multi-linear regression. Based the ISCCP total cloud fraction (all time) and shortwave flux data (from the ISCCP-H and ISCCP-FH products), we calculated the monthly cloud albedo (see Figures R4-5). However, the ISCCP-H product can't provide daytime total cloud fraction which may introduce errors into the estimated cloud albedo. And the radiative fluxes data provided by ISCCP-FH product is derived from the model output rather than direct observations. In addition, the linear relationship between the cloud fraction and planetary albedo from ISCCP observations is not as stable as that of MODIS and CERES observations (see Figures R4). Considering these uncertainties, the estimated cloud albedo based on ISCCP observations don't consider as inputs in the multiple regression model.

The monthly Multisensor Advanced Climatology of Liquid Water Path (MAC-LWP) is used to test the sensitive of the results to input LWP data. Considering the differences in retrieval methods and values of the MODIS LWP and MACLWP datasets (Greenwald, 2009), we used the threshold of 55 g m^{-2} for MACLWP to better split the samples evenly. The regressed results are given in Figure R6. For comparison, the results from MODIS LWP are also given below (see Figure R7). We can see that the results did not change significantly, indicating that the regressed results are relatively robust. For the reanalyzed dataset, indeed, ERA5 is considered to be the most state-of-the-art reanalysis with higher temporal and spatial resolutions. However, as the aerosol data from MERRA-2, we think it will be better to use corresponding MERRA-2 meteorological data. We also used the ERA5 data to perform the multiple regression model (see Figure R8). Although the results change slightly, the changed results do not affect the main conclusions. Therefore, the MERRA-2 data is used in the revised manuscript.

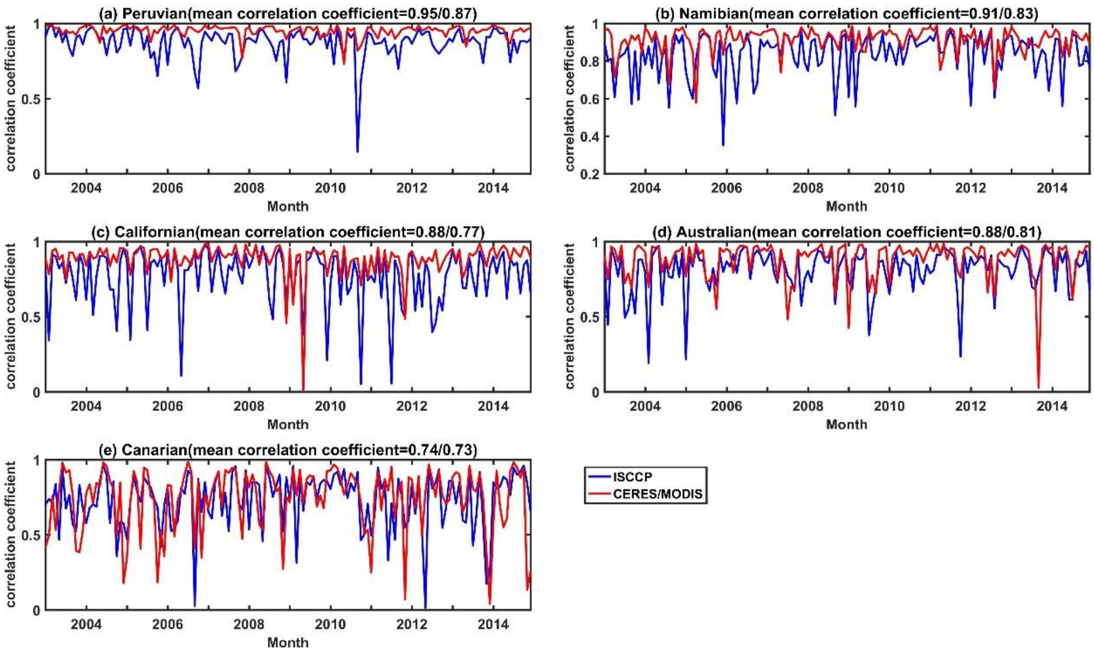

Figure R4: Monthly mean time series of correlation coefficients between cloud fraction and planetary albedo from CERES/MODIS observations (red lines) and ISCCP observations (blue lines) during 2003-2014, over the (a) Peruvian, (b) Namibian (c) Californian, (d) Australian and (e) Canarian regions. The monthly mean correlation coefficients are given in the title. The left side of the slash ('/') is the monthly mean correlation coefficients of CERES/MODIS observation, and the left side of the slash ('/') is that of ISCCP.

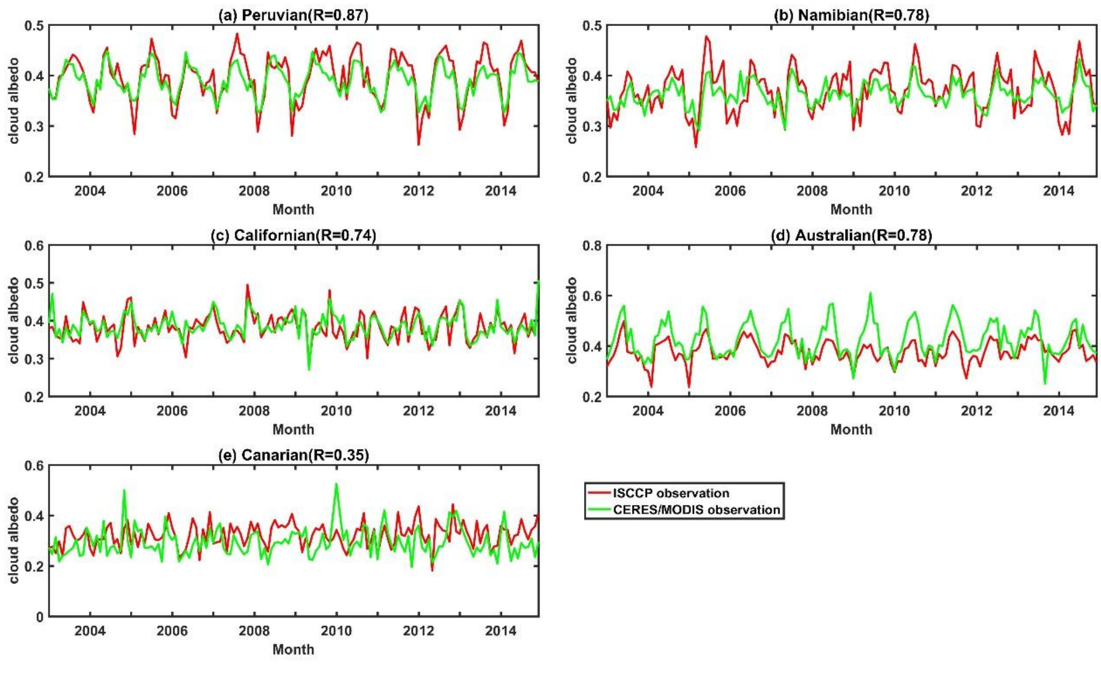

Figure R5: Monthly mean time series of cloud albedo based on ISCCP observation (red lines) during 2003-2014 compared with that of CERES/MODIS observations (green lines), over the

(a) Peruvian, (b) Namibian (c) Californian, (d) Australian and (e) Canarian regions. The R indicate the temporal correlations between satellite observations and ISCCP simulator outputs.

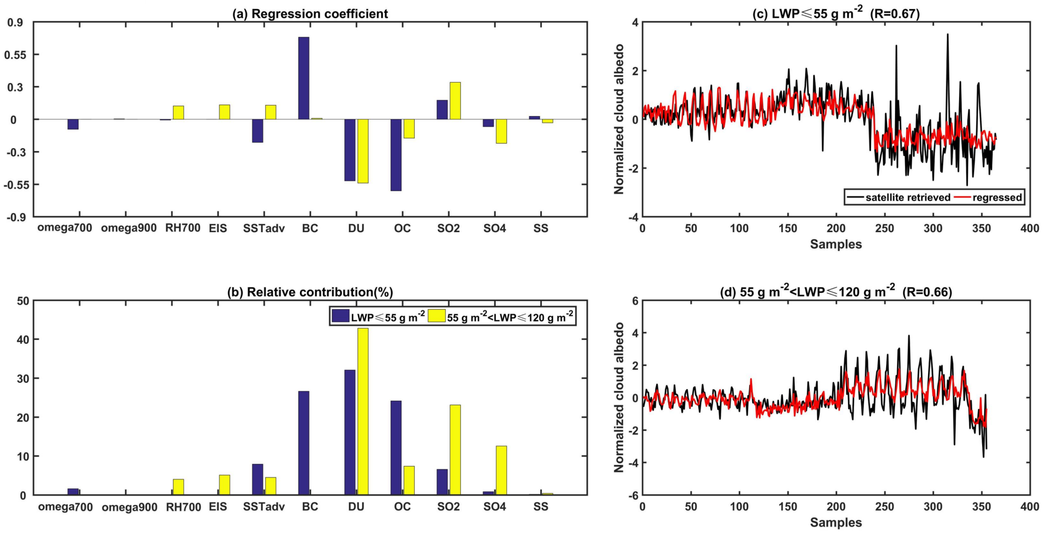

Figure R6: The (a) regression coefficients and corresponding (b) relative contribution of each predictor variables relating to cloud albedo from the multilinear regression models under two MACLWP conditions: LWP ⩽ 55 g m^{-2} (blue) and 55 g m^{-2} < LWP ⩽ 120 g m^{-2} (yellow). Note that for ease of comparison, eight variables are given in the figure, variables without values are not predictive variables of the sample group. And the satellite- and model-driven normalized cloud albedo trained in two sample groups: (c) LWP ⩽ 55 g m^{-2} and (d) 55 g m^{-2} < LWP ⩽ 120 g m^{-2}. The correlations (R value) between satellite- and model-driven normalized cloud albedo are given in parentheses.

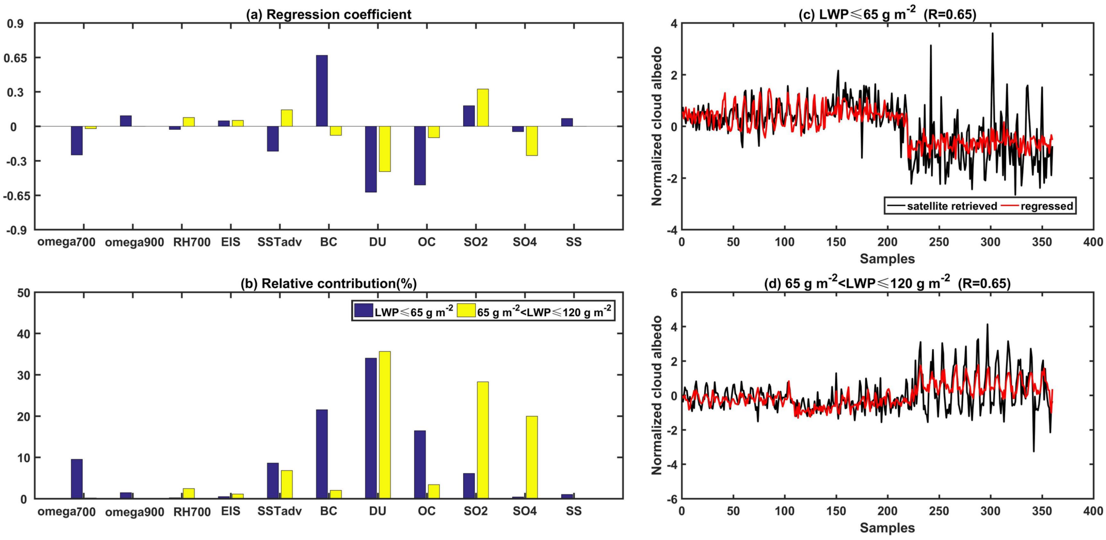

Figure R7 (Fig. 6 in the revised manuscript): Similar to Figure R6, but for MODIS LWP (LWP threshold: 65 g m^{-2}).

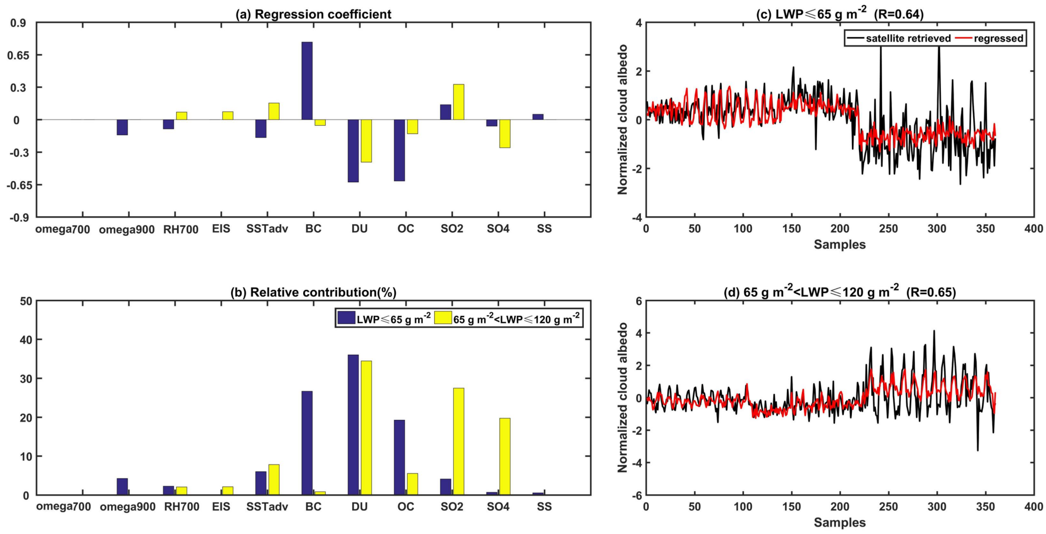

Figure R8：Similar to Figure R7, but for ERA5.

Greenwald, T. J.: A 2 year comparison of AMSR - E and MODIS cloud liquid water path observations, Geophys. Res. Lett., 36, L20805, https://doi.org/10.1029/2009GL040394, 2009.

5. How is the threshold of 60 g m^{-2} for LWP chosen?

Response: Thanks for your comments. To maintain a sufficient sample for both groups, the threshold of 60 g m^{-2} for LWP is chose in the previous manuscript. In order to better split the samples evenly, the threshold of LWP was modified to 65 g m^{-2} in the revised manuscript. Hence the sample size for both datasets is 360. In fact, the statistical results from the thresholds of 60 g m^{-2} and 65 g m^{-2} for LWP didn't exhibit obvious differences. Please see Figure R7 (LWP threshold: 65 g m^{-2}) and R9 (LWP threshold: 60 g m^{-2}).

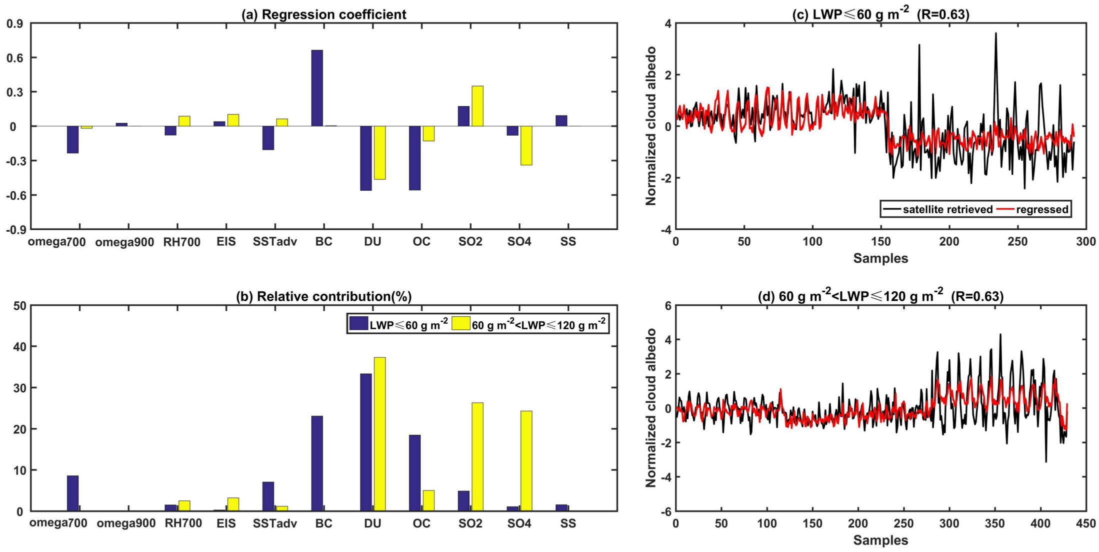

Figure R9: Similar to Figure R7, but the threshold of 60 g m^{-2} for LWP is used.

6. **Excessive detail in discussion of results.** I was quite overwhelmed with the amount of detail discussed in the results section of the paper concerning results for individual models and individual regions and for specific performance metrics. Even after reading the paper a few times, I cannot answer the basic question, "Has the simulation of CMIP6 stratocumulus cloud albedo changed in any major way compared to CMIP5, or is it fundamentally unchanged relative CMIP5?" The paper would be improved by identifying the key differences and similarities between CMIP5 and CMIP6, rather than discussing a detailed and hard-to-remember list of very specific results.

Response: We very thank reviewer for providing detailed comments and suggestions. Based on the comments, we reorganize the Result section (Section 3) and the superfluous information is deleted in the revised manuscript. Please see Section 3.1.