Response to reviewer 1
The authors are very grateful for the helpful comments. Our responses are provided below in blue.

Major comment
In figures 3a and 4, inter-model relationships between present-day LCC and dLCC/dSST (or dLCC/dSST and dCRE/dSST) should not be simplified too much. The negative correlations are clearly found when MIROC5 (in the upper-right corners of these panels) is set aside. However, physical reasons why MIROC5 can be omitted from discussion here are not clearly explained. The authors should extend their notes here to physics-based discussion. Why is MIROC5 so unique among the 14 models? Is there any unique physics scheme implemented in that model? To obtain any physical explanations, the authors can contact the model developers and/or developers of CALIPSO simulator and discuss with them.
We agree that we did not provide a detailed explanation for the behavior of outliers in this section. Before addressing the reviewer concern, we want to emphasize that the outlier in Fig. 3a is MIROC5 whereas the outlier in Fig. 4a is IPSL5A. Although we think it is beyond the scope of our paper - focused in detail mainly on GISS model evaluation - to contact any modeling center to inquire about a particular behavior evident in our analysis, we do attempt to provide a physically-based explanation using the existing literature for each of these models to address the reviewer's concern.

First, regarding Fig. 3a MIROC5:
The slight correlation that exists between ΔLCC/ΔSST and LCC in Fig. 3a is tied to whether or not models simulate i) Sc clouds in the correct regions and 2) their transition to shallow cumulus clouds. First, to produce Sc clouds, a model needs a turbulent scheme that accounts for moist processes in the PBL in which clouds maintain the turbulent mixing that sustains them. From Watanabe et al. (2010), it appears that MIROC5 does not have a moisture-aware turbulence scheme. Therefore, the model does not produce large, persistent decks of low clouds in Sc regions (off the western coasts of continents) and its ΔLCC/ΔSST is small. Second, a model has to develop shallow convection over a warmer sea surface to vent the PBL and transition from Sc to shallow cumulus clouds. However, MIROC5 lacks any such transition owing to "insufficient vertical mixing of the humid air in the PBL and the dry air in the free troposphere (FT)" as pointed out by Ogura et al. (2017) and Tatebe et al. (2018). This insufficient mixing results in too much low-cloud cover in the trade-wind regions and small spatial variability over the tropical oceans (see Fig. R1), consistent with a relatively large LCC value in Fig. 3a, as well as a small amplitude of ΔLCC/ΔSST because SST changes poorly propagate through the PBL to the FT (consistent with what found by Sherwood et al., 2014).
"The explanation of the MIROC5 behavior is twofold. First, similar to the other unconstrained models, MIROC5 evidently lacks the model physics to produce Sc-type clouds, i.e., simulating moist processes in the PBL in which stratiform clouds maintain the turbulent mixing that sustains them. As a result, its ΔLCC/ΔSST is small. Second, MIROC5 suffers from an insufficient vertical mixing of the humid air in the PBL and the dry air in the free troposphere (Ogura et al.,2017; Tatebe et al., 2018), which generates too large of an LCC over trade-wind regions compared to CALIPSO-GOCCP (not shown) and the largest mean LCC among the models. In addition, this poor vertical mixing may also explain the small amplitude of ΔLCC/ΔSST because SST changes

poorly propagate through the PBL to the free troposphere (consistent with what found by Sherwood et al., 2014)".

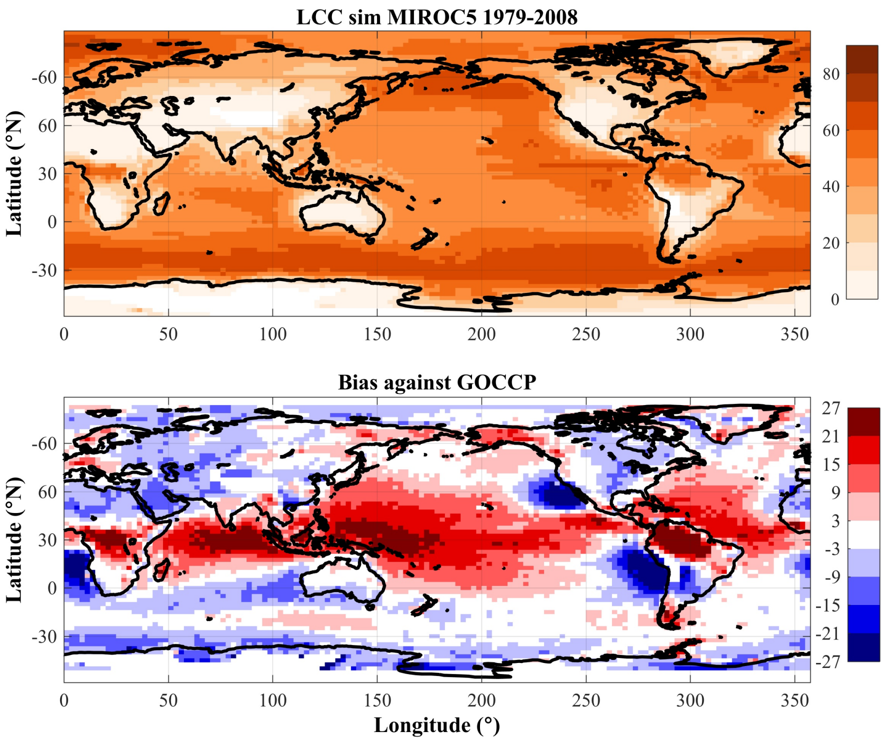

Figure R1: Maps of (a) low cloud cover (%) as simulated by MIROC5 through the lidar simulator and (b) the bias against CALIPSO-GOCCP. Note the lack of clouds over the Sc regions (attributable to the dry turbulence scheme) and the small spatial variability between the trade-wind regions and the rest of the tropics (attributable to insufficient vertical mixing between the PBL and the FT).

Second, Fig. 4a, IPSL5A:

For the IPSL5A model, it is a different problem. Although the physics for Sc clouds is again lacking, the tuning strategy of IPSL5A was focused on getting close to the observed SW CRE at TOA (i.e., Hourdin et al., 2006: *"particular care was given to the tuning of the cloud radiative forcing, and in particular to its latitudinal variations"*). Because their cloud amount is so small, they had to make the clouds very bright to meet the observational target during their tuning process. This problem, also known as "the too few too bright" problem, explains the strong sensitivity of the ΔCRE_{SW} of IPSL5A model to ΔLCC. To verify this, we computed 2D-histogram of SW CRE as a function of the respective LCC for all the models and for the observations in Figure R2. As expected, the IPSL5A model is the only model to show such a particular behavior, that is to say very low cloud amount having very large SW CRE. The deficiency is now better explained in the manuscript in section 4.3: *"This so-called "too few too bright" problem may explain the particular behavior of IPSL5A model in Fig. 4a. In this model, the SW CRE for a given LCC value is far too large compared to the observations and any other model (Fig. S6). This may be why the sensitivity of ΔCRE_{SW} to ΔLCC ($\Delta CRE_{SW}/\Delta LCC$, Tab. 2) is too large and far-off the correlation line in Fig. 4a."*

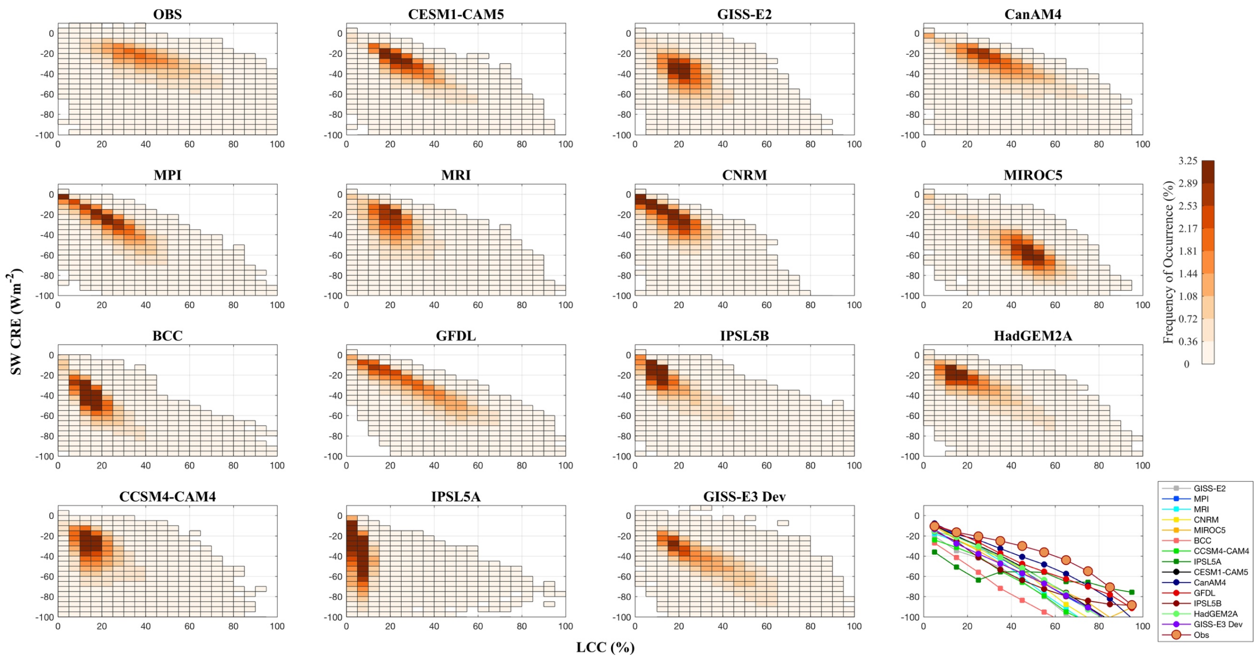

Figure R2: 2D-histograms (frequency of occurrence, %) of shortwave cloud radiative effect (SW CRE, Wm^{-2}, y-axis) as a function of the low-cloud fraction (%, x-axis) for CERES-EBAF and CALIPSO-GOCCP observations (2007-2016) and for the 14 models. The bottom-right panel shows the averaged relationship (see the legend for models' name).

In addition to the overestimation or underestimation of the CRE due to cloud property biases, we also note that the clear-sky portion of cloudy gridboxes may also influence $\Delta CRE/\Delta SST$. For example, artificially increasing the specific humidity of the clear-sky (for radiative transfer only) in GISS-E3 dampens $\Delta CREsw/\Delta SST$ compared to the control version of GISS-E3 because of the increased SW absorption by water vapor. As a result, the SW radiation reflected back to space is reduced, which reduces the SW CRE. Because the SW CRE is smaller, the ratio $\Delta CRE/\Delta LCC$ is reduced, meaning that the same reduction of LCC per K ($\Delta LCC/\Delta SST$) will generate a smaller SW positive feedback ($\Delta CRE_{SW}/\Delta SST$). We added this explanation in the manuscript: *"In addition, the radiative effect of clear-sky portion of the cloudy grid boxes can amplify or dampen the interannual SW cloud feedback ($\Delta CRE_{SW}/\Delta SST$). For example, artificially increasing the specific humidity of the clear-sky in GISS-E3 (for radiative transfer only) reduces the SW CRE at TOA because of the increased SW absorption by water vapor, which ultimately dampens the positive SW cloud feedback with respect to the change in LCC per K ($\Delta CRE_{SW}/\Delta LCC$, not shown)."*

Specific comments

Page 2 line 2: You should combine the first and second paragraphs into a single paragraph

Done.

Page 9 line 14 "the radiative effect of increased CO2 on cloud-top turbulence": Any appropriate citations needed here. Any LES or GCM studies?

This effect is documented by Bretherton 2015, which we added to the manuscript: *"For example, current climate variability does not include the radiative effect of increased CO2 on cloud-top turbulence, which may generate a reduction of stratocumulus cloud amount by increasing* *downwelling LW flux and thus reducing cloud-top radiative cooling (e.g., Bretherton et al., 2015)."*

Page 13 line 6-12: Figure 7 of Su et al. (2013; doi:10.1029/2012JD018575) may be relevant to discussion here

We added the following sentence to the manuscript: *"This is somewhat consistent with the decoupling of the LTS and SST pointed out by Su et al. (2013)."*

Page 13 line 22, 33: "Finally" repeated

We removed it.

Figure 1 caption: "is greater than 10 h Pa/d" should be "equal to 10 h Pa/d over ocean". Please also check supplementary figure caption.

We fixed this in the manuscript and the supplementary material.

References:

Ogura, T., Shiogama, H., Watanabe, M., Yoshimori, M., Yokohata, T., Annan, J. D., Hargreaves, J. C., Ushigami, N., Hirota, K., Someya, Y., Kamae, Y., Tatebe, H., and Kimoto, M.: Effectiveness and limitations of parameter tuning in reducing biases of top-of-atmosphere radiation and clouds in MIROC version 5, Geosci. Model Dev., 10, 4647-4664, https://doi.org/10.5194/gmd-10-4647-2017, 2017.

Sherwood, S. C., S. Bony and J.-L. Dufresne: Spread in model estimates of climate sensitivity traced to atmospheric convective mixing. Nature, 505, 37-42, doi:10.1038/nature12829, 2014.

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., Sudo, K., Sekiguchi, M., Abe, M., Saito, F., Chikira, M., Watanabe, S., Mori, M., Hirota, N., Kawatani, Y., Mochizuki, T., Yoshimura, K., Takata, K., O'ishi, R., Yamazaki, D., Suzuki, T., Kurogi, M., Kataoka, T., Watanabe, M., and Kimoto, M.: Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-155, in review, 2018.