Detailed response to reviewer #2' comments

We would like to thank the reviewer#2 for the insightful comments which helped improve this paper. A list of our response is given below (in Italic).

In this paper, the authors investigate cloud properties as seen by different A-train satellites over the Southeast Asia. They further divide clouds into 5 cloud types as a function of their cloud phase and different overlapping possibilities among the cloud-type layers. In the last part of the manuscript, they study possible links between these 5 types and MJO and ENSO conditions. While the topic of this paper aligns with the scope of the journal, it is difficult to judge the novelty of the analysis since most of it is a kind of climatology rather than new results. However, I acknowledge a tremendous amount of work from the authors. Yet, the paper is too long and descriptive, which makes it difficult to follow. In addition to trimming the manuscript, I have a couple more concerns to address before recommending this paper for publication. A more detailed explanation is provided below.

The paper is indeed a climatological study, focusing on cloud phase characteristics. Since the climatological characteristics of cloud phase have not been addressed, the results are novel. The abstract highlights several of the new key results, and we note that Referee#1 commented on the novelty of our results. The length and descriptive nature of the manuscript is addressed below.

Main concerns

1) This paper is too long and descriptive. It is hard to follow and I often lost track of the goal of the sections. Every section should be reduced in size and I would recommend focusing on specific findings relevant to the topic of the study rather than describing every subplot of the figures as well as the behavior of each cloud types.

Thank you for this comment.

We have worked to reduce the length of most sections of the manuscript in response to this comment. We have also provided additional edits throughout the manuscript to improve readability. In reducing the text, we did aim to let the figures and tables speak for themselves, but key findings from the figures do need to be discussed. We note that Referee#1 stated "The results are discussed adequately, ..." so we tried to strike a balance between the referees' comments that are at odds with one another.

2) If I understand correctly there is no filtering of the data whatsoever to make them consistent with each other. I find this a little concerning. For example, in section 3.1.5 when comparing the spatial heterogeneity index with CloudSat-CALIPSO, all pixels are used including those where CC and MODIS cloud masks disagree. This may result in large biases as explained by the authors later on. It would be best to keep in the main analysis the pixels where CC and MODIS agree.

Thanks for this comment. We did examine this issue (results shown in figure below), and some of these issues are discussed in the paper (as noted by the reviewer), with results on CC and MODIS both being "clear" shown in Table 2. If we forced the analysis to be the same class consistency, then that leaves us vulnerable to carrying MODIS cloud detection and classification errors into our analysis, which we didn't want to do. The point is just to focus on CC *classification as a function of H_σ, without the additional issues that MODIS cloud detection and classification would bring to the interpretation (i.e., discussion as to why the red and black curves look different below is entirely due to MODIS cloud detection and classification limitations—miss some thin or small clouds).*

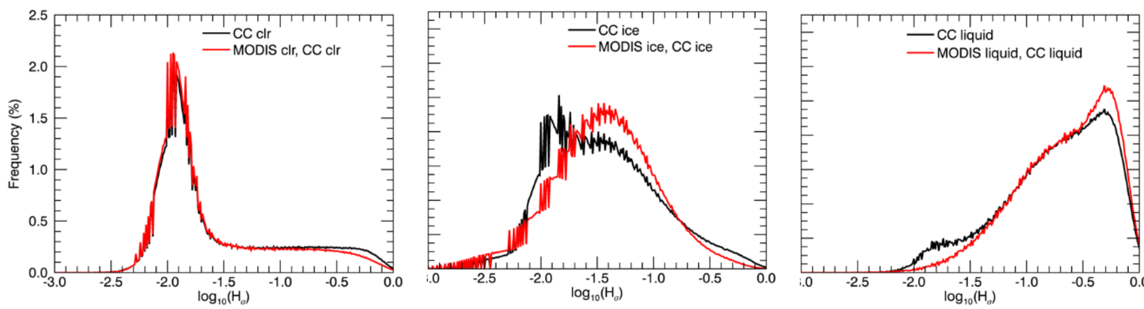

Fig. the H_σ PDF for clear, ice and liquid cloudy skies: black for CC detections, and red for the samples agreed by both CC and MODIS.

3) The authors use the version R04 of the 2B-CLDCLASS-LIDAR product. This version is not free of uncertainties in particular when it comes to detecting shallow cumulus clouds. It's been shown that this version overestimates the amount of shallow cumulus clouds (https://www.earth-syst-sci-data.net/11/1745/2019/essd-11-1745-2019.html). Similarly, nothing is said about any kind of uncertainty in the cloud phase retrieval of this product. For example, has this product been evaluated against other cloud phase dataset (ground-based, satellite or in situ?). I know that the cloud phase confidence considerably decreases when the lidar is totally attenuated and the decision tree only relies on the radar signal. I would suggest the authors to mention these at least.

Thanks for pointing us to the new version 2B-CLDCLASS-LIDAR data (R05), which was not available at the time of our original analysis.

1) We have updated all our results using Release 05 (R05) of both 2B-CLDCLASS-LIDAR and 2C-ICE. Compared to the R04 version, R05 shows more ice-only (0.8%↑), mixed-only (0.5%↑) and ice-above-mixed (1.9%↑) clouds, but less liquid-only (0.9%↓) and ice-above-liquid (0.8%↓) clouds in the Southeast Asia region (compare Table 3 in the revised and discussion paper). Also, the new 2B-CLDCLASS-LIDAR product displays a better interannual variations of cloud phase associated with ENSO (see subfigure in Fig. 14d between 07/08 and 07/09). Fortunately, the small changes didn't impact any of our conclusions.

2) We agree with the reviewer's concern on the uncertainties of the 2B-CLDCLASS-LIDAR cloud phase retrieval. Characterizing uncertainties in classification does require a truth to compare against, which doesn't exist for cloud phase. When such truths are lacking, the standard approach in validating cloud classification results is to validate the thresholds used in the classification algorithm (Rossow et al., 1989). The thresholds used in the classification algorithm for 2B-CLDCLASS-LIDAR cloud phase is discussed in Section 2.1 and references therein. Still, this doesn't achieve quantitative uncertainty characterization on cloud phase that a *comparison to "truth" can give. To date, there are not any evaluation of the CC cloud phase against any other cloud datasets. This is why we performed a comparison of CC and MODIS cloud phase as shown in Table 2 and described in Sect. 2.3 in the paper. Overall, most of CC ice-only, ice-above-liquid, ice-above-mixed and mixed-only clouds are reported to be ice by MODIS, and most of CC liquid-only clouds are also detected to be liquid by MODIS. This comparison allows us to better interpret our results in later sections (e.g., Sect. 3.1.6 and Sect. 3.3).*

3) We agree that when lidar signal is totally attenuated, confidence level of cloud phase is lowered down. We mention this information in Page 4, Lines 34-35.
'When the lidar signal is totally attenuated, the cloud phase is determined only by Ze and temperature, which lowers down the confidence level.'

4) Finally, the authors consistently mention that their results could be used for model evaluation but fail to explain how. I understand it is tempting to sell any observational result as a possible constraint for model, but if the authors want to do so, they need to explain how and why, which is not done here.

In the paper (Page20, L12-24), we added the following paragraph to make the model-observation comparison clearer.

"Finally, we note that our results may be used to evaluate a model's verisimilitude in capturing cloud properties, particularly phase and spectral characteristics. For example, we show summaries of spectral radiance at the TOA segregated by cloud phase and overlap conditions that can serve as a basis for comparing to those computed from model outputs—a similar approach given by previous research (e.g. Hashino et al., 2013; Masunaga et al., 2007;Yao et al.,2020). Since these models also use the plane-parallel assumption in computing the spectral radiation leaving the TOA, careful comparisons between model and observations can use Hσ as a measure of departure from the plane-parallel assumption in a manner similar to Loveridge and Davies, (2019), where they used Hσ within their analysis in examining GCM clouds in different sectors of southern hemisphere cyclones. The use of Hσ also extends its application to gauge biases in other satellite products used in model evaluation (e.g. Gettelman et al., 2015; Song et al., 2018), such as cloud optical depth and effective radius, since biases in these products have been noted to covary with H_σ(Fu et al., 2019; Zhang et al., 2016)."

Minor comments
Throughout the manuscript, the authors use plural with the term cloud phase, I feel like most of the time, it would be better to use singular.

As suggested, we correct the 'cloud phases' to 'cloud phase' in most places.

P2 L35: improve GCM performance => improve climate simulations

Corrected.

P4 L8: in the lower troposphere => below 8 km

We change the statement as 'in the lower troposphere (i.e. below 8.2 km)'

P4 L8: in the upper troposphere => above 8 km

We change the statement as 'in the upper troposphere (i.e. above 8.2 km)'

P4 L21: Please specify the version. From your table, I believe you use the version R04.

Data version is added and now all results are updated using R05 data.

P4 second paragraph: You don't describe how the algorithm works when the lidar signal is completely attenuated. The cloud phase is then based only on Ze and T thresholds, which substantially decreases its confidence level. Since the region under study is dominated by convective clouds, this situation may occur very often.

Yes, it is true that Ze and T threshold is used to deduce cloud phase in the radar-only region, which could lower down the confidence level.

This information is mentioned in Page 4, Lines 34-35 in the revised paper:

'When the lidar signal is totally attenuated, the cloud phase is determined only by Ze and temperature, which lowers down the confidence level.'

P8 L13: I don't understand the unit of LTS, it's supposed to be in K (or ËŽC).

The unit of LTS depends on how to define it. We define static stability same as Frierson and Davis, (2011) and Li et al. (2014), i.e., $\frac{\partial\theta}{\partial z}$, *which has the unit of K/km, while in some other studies such as Klein and Hartmann, (1993), they used the definition of* $\Delta\theta = \theta(p = 700mb) - \theta(p = \text{sea level pressure})$, *whose unit is K.*

To avoid the confusion, in Page 8, Lines 1-3, we have revised the text as:

'The lower-troposphere static stability (LTSS=$(\theta_{z=3\ km} - \theta_{z=0})/3\ km$*) and the upper-troposphere static stability (UTSS =* $(\theta_{z=tropopause} - \theta_{z=tropopause-3\ km})/3\ km$*) are shown in Figs 2d1-d4, where the* θ *is potential temperature in unit of K.'*

P8 L18-23: The CLDCLASS-LIDAR product provides a cloud fraction (between 0 and 1) per layer so how do you get cloud and sample numbers?

Whenever the lidar cloud fraction within a radar volume is reported to be greater than zero, we count that radar sample as cloudy. The frequency reported in Figure 3 is the frequency of these samples.

To clarify this information, we now add the text in Page 4, Lines 17-18:

'This product reports the lidar cloud fraction that records how many lidar profiles are contained in a radar resolution'

and in Page 4, Lines 39-40：

'Four years of 2B-CLDCLASS-LIDAR data, version P1_R05 (2007-2010) with lidar cloud fraction greater than zero are used.'

P8 L25: Large => large
Corrected.

Fig. 3: why do you use this particular cross section rather than a zonal mean.

We have updated the results using the zonal mean.

P8 L41: "Also: : : of cirrus". Why do you mention MISR here out of the blue? Also this sentence is confusing.

We have deleted these statements to get rid of the confusion.

P9 L1: Can you elaborate on this statement? DO you mean for that region?

The original statement: " Low-level clouds cloud have a high chance to be covered by the upper ubiquitous ice clouds (Yuan and Oreopoulos, 2013), which is further quantified in next section."

To make this clear, it has been revised in Page 8, Lines 36-38,

'As shown in Yuan and Oreopoulos, (2013), low-level clouds have a high chance to be overlapped by upper clouds in the warm pool region. In the next section, we will examine cloud overlap with a focus on cloud phase.'

P9 L19: Confusing sentence.

The original statement: "However, liquid-only clouds have very small frequencies (< 10%) between 10°S-10°N where widely distribute ice clouds, which indicates that liquid clouds occurring here are likely being covered by ice clouds, hence, they are grouped as ice-above-liquid cloud class."

Now in Page 9, Lines 19-21 , it is rephrased to make in clearer:

'Elsewhere, liquid-only clouds have very small frequencies (< 10%). The annual mean frequency of liquid-only cloud is ~16.0%.'

P11 L13: This should appear in the data section along with the other uncertainties related to the datasets.

As suggested, we mention these uncertainties due to instrument limitations in Sect. 2.1 (Page 4, Lines 35-37), where we describe the 2B-CLDCLASS-LIDAR data.

'Also, in cases of thick ice clouds attenuating lidar signals over shallow liquid clouds that are missed by the radar, only ice clouds are reported in the profiles. Biases due to instrument limitations are kept in mind in our analysis.'

P11 L22: CLIPASO => CALIPSO
Corrected.

P11 L22: attenuated by clouds with optical thickness greater than 3.
It is revised as suggested.

P12 L14: spatial => spatially
Corrected.

P12 L26: "due to the small optical thickness of the.."
It is revised as suggested.

P12 L27: It's unclear to me why the authors constantly refer to MISR for no reasons since MISR observations are used in this study.
Thanks for pointing out this. In the new version, we remove the contents related to MISR to avoid the confusion.

P12 L36: This sentence needs re-wording.
Original statements: "While these clouds are locally homogenous, hence favoring the plane-parallel assumption in radiation computation (Ham et al., 2015)."

Now in Page 12, Lines 18-19, it is revised as:

'These clouds are locally homogeneous and hence favor the plane-parallel assumption in radiation computation.'

P13 L1-2: I would strongly recommend excluding pixels in which CC and MODIS disagree in the main figures rather than only mentioning it as "not shown".

Comment addressed earlier.

P13 L2-4: There is no main verb in this sentence, please re-word.

Also, are you referring to shallow cumulus clouds? In this case it would be rather easy to validate your hypothesis by focusing on a shallow Cu dominated region, such as the Barbados. However, depending on the MODIS product used, Pincus et al 2012 reported that a substantial amount of these clouds (partially-filled pixels) are excluded of the cloud product. Another important thing to note is that, R04 over-estimate shallow Cu cloud fraction (see main concern comments).

1) Original statement:

"Indeed, many small liquid clouds with size ranging in a few tens to hundreds of meters (e.g., Koren et al., 2008) that are difficult to be measured by MODIS as reported in Zhao and Di Girolamo (2006)."

Now in Page 12, Lines 27-29

'Indeed, many small liquid clouds with size ranging in a few tens to hundreds of meters can go undetected by MODIS (Zhao and Di Girolamo 2006).'

2) Here, we are referring to small and shallow cumulus clouds undetected by MODIS--with their sizes smaller than 1 km. Pixels containing these small clouds could be reported to be clear by both CC and MODIS due to their relatively large spatial resolutions. To validate our hypothesis, we revisit the MODIS and Advanced Space Thermal Emission and Reflection Radiometer (ASTER) data (15 m resolution) used in Zhao and Di Girolamo (2006) over the tropical western Atlantic (Rain in Cumulus over the Ocean field campaign, near Barbados – check Fig. 1 in Rauber et al., (2007)). By excluding the MODIS clear sky pixels that contain ASTER reported clouds, i.e. a focus on MODIS-ASTER clear sky pixels, the long tail of the H_σ *is strongly reduced (see the Cyan line in the following figure). This is consistent with our hypothesis that small liquid clouds contribute to the long tail of* H_σ *PDF.*

To make this clear, now in Page 12, Lines 29-35, we have included the statements:

'We revisit the MODIS and Advanced Space Thermal Emission and Reflection Radiometer (ASTER) data (15-m resolution) used in Zhao and Di Girolamo (2006) over the tropical western Atlantic. The long tail of Hσ PDF is significantly reduced, i.e. frequency change from 0.3% to 0.1% at Hσ ~ 0.1, when the ASTER data is applied to exclude the MODIS clear sky pixels that contain ASTER reported clouds. This further affirms that the undetected clouds in MODIS and CC clear sky pixels contribute to large Hσ values, which at least impact 20% clear-sky samples when Hσ >1 (Fig. 7a)'

3) In terms of the concern on R04 data, we have updated our results with the R05 version data.

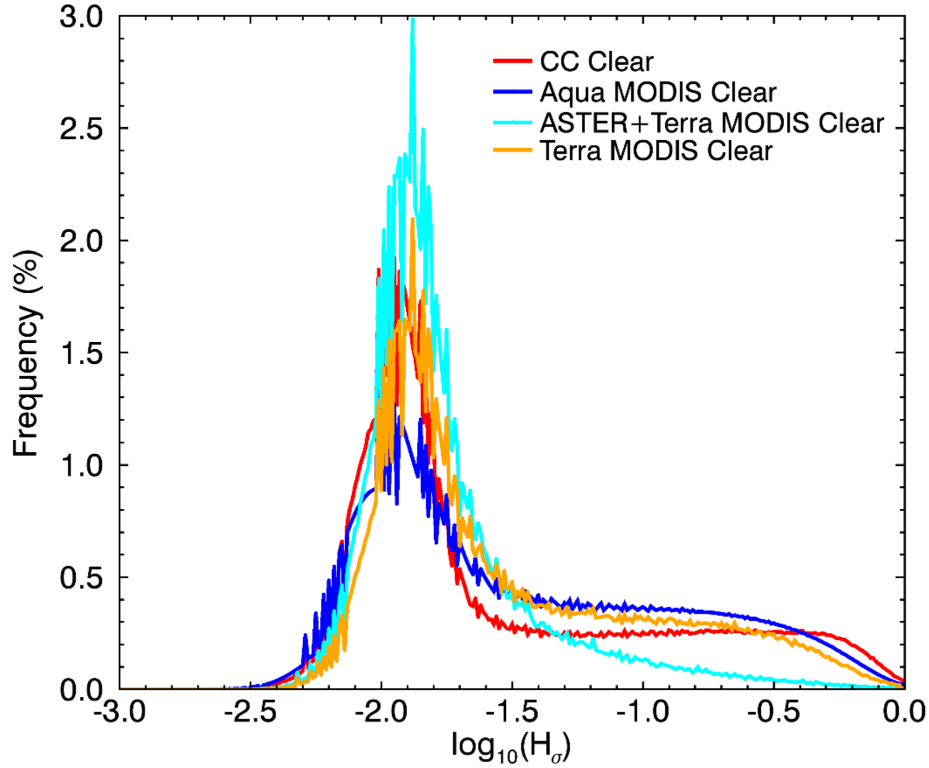

Fig. H_σ *PDF for clear skies obtained from CC and Aqua MODIS over Southeast Asia and from Terra MODIS and ASTER-Terra MODIS over the tropical western Atlantic region.*

P13 L33: Proves seems a bit strong.
We replace 'proves' as 'indicates'

P14 L35: reflecting
It is revised as 'reflective'.

P17: Why are you showing MJO phases? It's been documented in many many studies already. What does this bring to your study?

We agree that MJO is well documented in different aspects, including the related cloud type, radiative, dynamic and thermal dynamic characteristics. However, the cloud phase and the corresponding heterogeneity are less studied.

To make it clear, Page 15, Lines 40-43, we added the following statements:
'This section discusses the features of cloud phase associated with the intraseasonal 30-90 day MJO. Previous studies have provided full overviews of the radiative (in terms of OLR), dynamic and thermal dynamic characteristics of the MJO (Knutson et al., 1986; Riley et al., 2011; Wheeler and Hendon, 2004; Zhang, 2005). The purpose of this study is to focus on how the cloud phase characteristics discussed in previous sections vary with MJO phases.'

P17 L38: "suppressed"

It is corrected.

P18 L3: I don't understand the meaning of this sentence and I don't see how this could be used to validate climate models.

Original statement: 'Overall, the eastward-propagating H_σ patterns-a behavior similar to OLR pattern (e.g., Wheeler and Hendon, 2004), indicate that the H_σ could be useful for MJO studies such as serving as an observed-based parameter that are sensitive to cloud phase, to track MJO position and validate MJO simulations in climate models.'

Now in Page 17, Lines 4-6, it is revised as:
'The eastward-propagating H_σ patterns vary with MJO, indicating that H_σ could be useful for MJO studies, such as serving as an observed-based parameter to track the MJO position.'

P18 L8-20: It basically shows there are more convective clouds.

We remove the reflectance and brightness temperature from Fig. 13 and rewrite the whole paragraph to emphasize the heterogeneity variations (see Page 17, Lines 17-23).

P19 L13-14: Another sentence without meaning, please re-word.

Original statement:

'As cloud phases vary interannually and hence change the spatial heterogeneity, i.e., being smoother in La Niña year than normal and vice versa in El Niño year.'

In Page 18, Lines 10-11, It is revised as:

'the cloud phase varies interannually, as does H_σ , i.e., being smoother in La Niña years compared to El Niño years.'

P19 L36: "preferentially occur in"
It is revised as suggested.

P19 L38: occur => are
It is revised as suggested.

P21 L10: a comma is missing after ENSO L11.
It is corrected.

Overall, I would suggest rephrasing this sentence because I don't think the authors can claim the heterogeneity index captures MJO or ENSO. At best, it varies for the different ENSO/MJO phases, but it's definitely not well correlated. For the second part of the sentence, unless the authors explain how one could use this for model evaluation, I'd recommend to remove.

Original statement: "The observed H_σ values capture the MJO and ENSO features, implying that the Hσ is able to track MJO and ENSO and provides a way to validate their simulations in GCMs"

In Page 20, Line 6-7, the following sentence replaces the original statement:

'The observed Hσ varies with the ENSO index with a correlation coefficient of 0.49 (significant at confidence level 0.99).'

P21 L14-16: Here again, there is no tool to compare this to models, at least to the best of my knowledge, so unless the authors elaborate on this statement, they should remove this statement. I can envision a qualitative comparison of heterogeneity at best.

Addressed in point 4 above.

References

Frierson, D. M. W. and Davis, N. A.: The seasonal cycle of midlatitude static stability over land and ocean in global reanalyses, Geophys. Res. Lett., 38(13), 1–6, doi:10.1029/2011GL047747, 2011.

Fu, D., Di Girolamo, L., Liang, L. and Zhao, G.: Regional Biases in MODIS Marine Liquid Water Cloud Drop Effective Radius Deduced Through Fusion With MISR, J. Geophys. Res. Atmos., 124(23), 13182–13196, doi:10.1029/2019JD031063, 2019.

Gettelman, A., Morrison, H., Santos, S., Bogenschutz, P. and Caldwell, P. M.: Advanced two-moment bulk microphysics for global models. Part II: Global model solutions and aerosol-cloud interactions, J. Clim., 28(3), 1288–1307, doi:10.1175/JCLI-D-14-00103.1, 2015.

Klein, S. A. and Hartmann, D.: The seasonal cycle of low startiform clouds, J. Clim., 6, 1587–1606, doi:10.1175/1520-0442(1993)006<1587, 1993.

Li, Y., Thompson, D. W. J., Stephens, G. L. and Bony, S.: A global survey of the instantaneous linkages between cloud vertical structure and large-scale climate, J. Geophys. Res. Atmos., 119, 3770–3792, doi:10.1002/2013JD020669, 2014.

Rauber, R. M., Stevens, B., Ochs, H. T., Knight, C., Albrecht, B. a., Blyth, a. M., Fairall, C. W. and Jensen, J. B.: Over the ocean: The RICO campaign, Bull. Am. Meteorol. Soc., (December 2007), 1912–1928, doi:10.1175/BAMS-88-12-1912, 2007.

Rossow, W. B., Garder, L. C. and Lacis, A. A.: Global, Seasonal Cloud Variations from Satellite Radiance Measurements. Part I: Sensitivity of Analysis, J. Clim., 2(5), 419–458, doi:10.1175/1520-0442(1989)002<0419:gscvfs>2.0.co;2, 1989.

Song, H., Zhang, Z., Ma, P. L., Ghan, S. J. and Wang, M.: An evaluation of marine boundary layer cloud property simulations in the Community Atmosphere Model using satellite observations: Conventional subgrid parameterization versus CLUBB, J. Clim., 31(6), 2299–2320, doi:10.1175/JCLI-D-17-0277.1, 2018.

Zhang, Z., Werner, F., Cho, H.-M., Wind, G., Platnick, S., Ackerman, A. S., Di Girolamo, L., Marshak, A. and Meyer, K.: A framework based on 2-D Taylor expansion for quantifying the impacts of subpixel reflectance variance and covariance on cloud optical thickness and effective radius retrievals based on the bispectral method, J. Geophys. Res. Atmos., 121, 7007–7025, doi:10.1002/2016JD024837, 2016.