Reply to interactive comment on

"Contrasting the Co-variability of Daytime Cloud and Precipitation over Tropical Land and Ocean"

by Daeho Jin et al.

Anonymous Referee #2

Received and published: 13 November 2017

General comments:

In "Contrasting the Co-variability of Daytime Cloud and Precipitation over Tropical Land and Ocean", Jin et al. analyze satellite datasets of cloud and precipitation retrievals with the aim of determining the relationship between cloud and rain. The topic is highly relevant for understanding the behavior of the atmosphere, both in physical reality and in parameterized cloud and precipitation in models. To my knowledge, the authors are the first to use this particular technique of regime-based cloud classification to analyze the relationship between cloud and precipitation, and I therefore recommend the analysis be published. The authors have chosen a fairly non-straightforward analysis method, and I hope the comments below will help them clarify a few points for the reader.

We thank the reviewer for the overall positive assessment of our work and the recognition that we have introduced an original analysis approach.

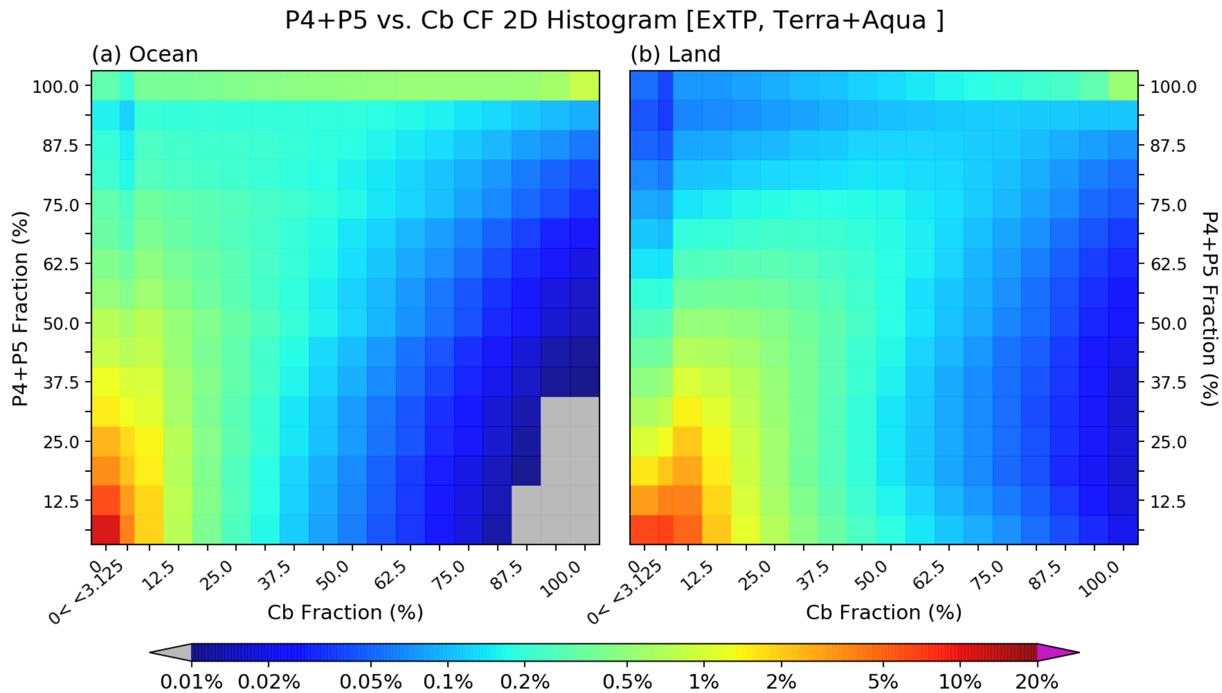
Major comments:

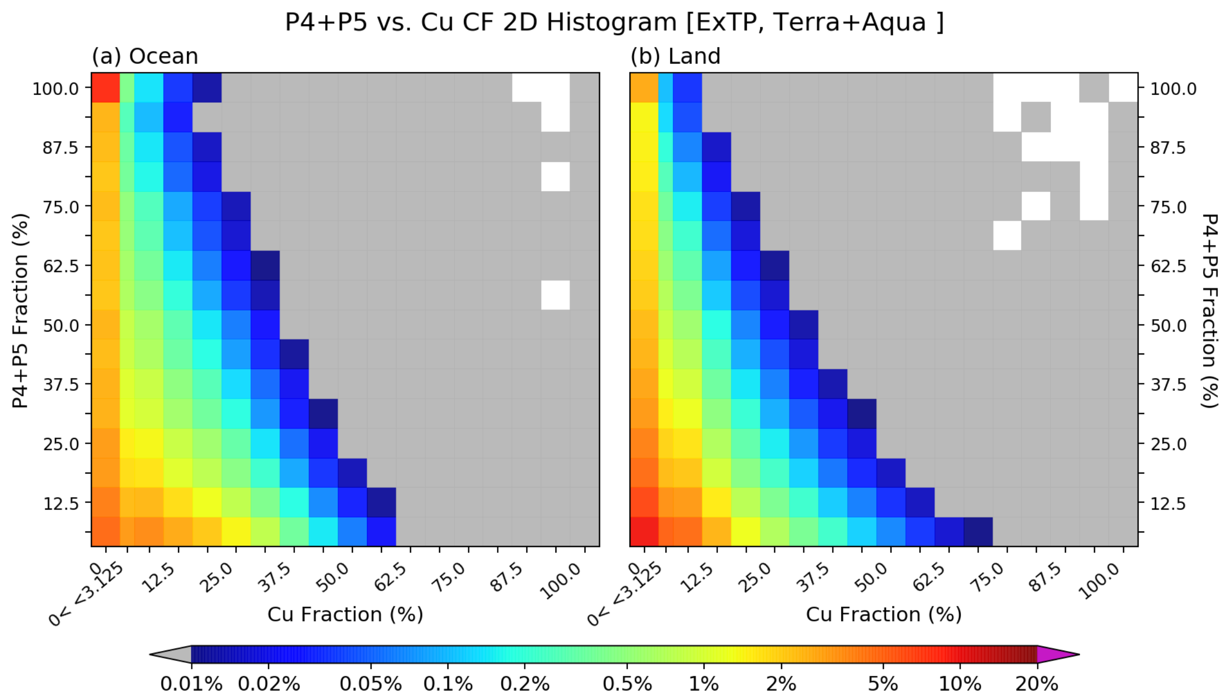
The main potential weaknesses of the analysis are the following:

1. As the authors themselves point out, using cloud optical thickness and cloud-top pressure to define cloud regimes is an essentially ad-hoc classification based on arbitrary choices. In the conclusions, they then describe the regime classification as "widely accepted". It is true that these regimes are widely used, subject to the known caveats that the authors correctly state (e.g., that the regime names are not to be taken to correspond literally to actual cloud types); however, this acceptance is based on the regimes' usefulness having been demonstrated for each particular application, for example by showing that susceptibilities to aerosol are very different across regimes (e.g., Gryspeerdt et al, ACP 2014). In my opinion, this paper provides some interesting indications that the regime classification does indeed differentiate between cloud type of very different behavior regarding precipitation, but this is the case for only three (out of nine) regimes, so I think it will take some extra work (perhaps the unpublished paper referred to in the Conclusions) before the field will "widely accept" the use of these regimes for precipitation studies.

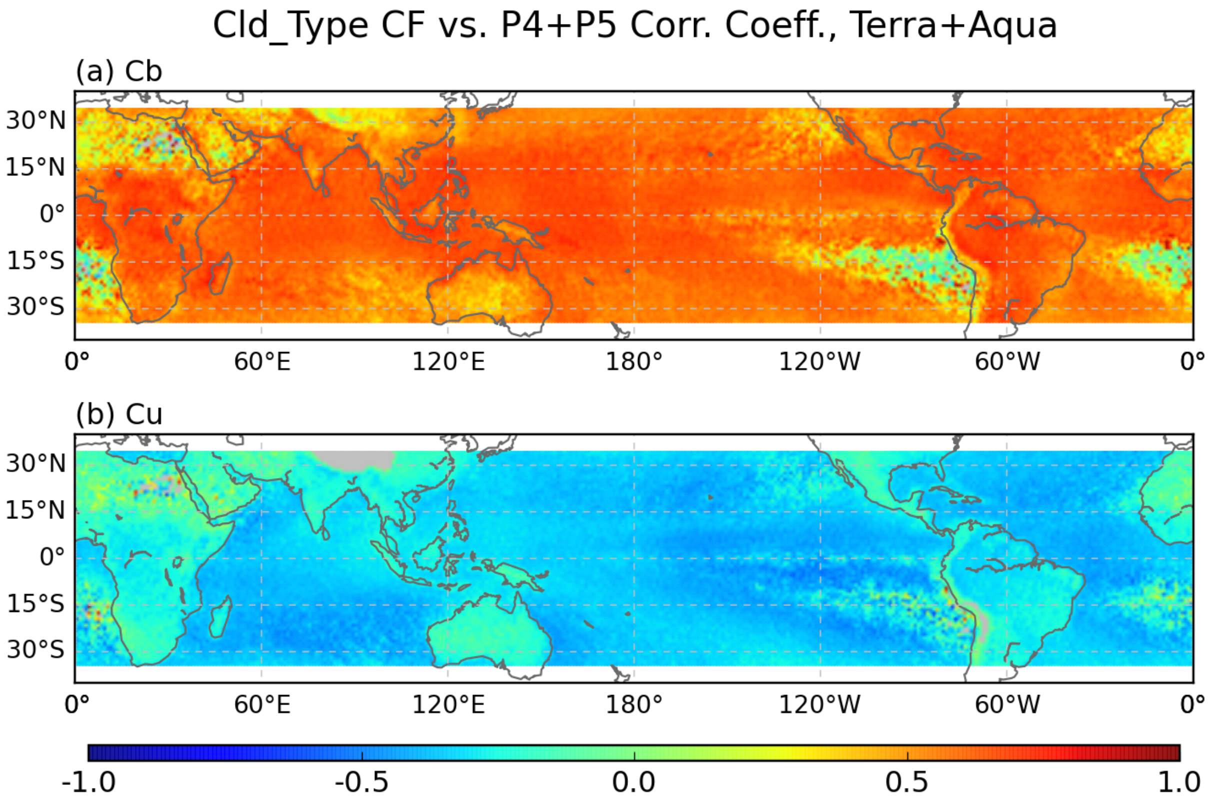
We take the opportunity here to clarify that our work uses the concept of "cloud type" rather than that of "cloud regime" (aka "weather state") as in Oreopoulos et al. (2014, 2016), the Gryspeerdt work quoted above, and previously in Rossow et al. (2005), among others. This is an important distinction. The presence of cloud types in a grid cell is simply described by the fraction of pixels within appropriate CTP and COT boundaries. On the other hand, when discussing cloud regimes, each cloudy grid cell can belong only to one cloud regime (derived from clustering analysis), specifically the one whose centroid histogram minimizes the Euclidean distance from the grid cell's particular joint histogram occurrence. In other words, a regime represents a mixture of cloud types, with usually one of them being dominant. A combined cloud regime-precipitation analysis was conducted by Lee et al. (2013) and was completely different in character than the present study. Had we used cloud regimes we could have conceivably correlated each grid cell's total CF with the values in the 16 different precipitation bins (or 5 precipitation groups), and composited the results by regime. But we chose instead the cloud type approach since there is greater familiarity in associating cloud types with precipitation. While cloud types have initially been defined in terms of cloud appearance as viewed from surface human observers, the ISCCP definition of assigning cloud types from space-based passive observations using detected cloud extinction and vertical location is also considered quite standard, see https://isccp.giss.nasa.gov/cloudtypes.html (an early version appeared in Rossow and Schiffer, BAMS 1991, Fig. 4). We hope that this background clarifies the use of the expression "widely-accepted concepts about how to classify clouds into various types from passive observations" in p. 16, line 15 that the reviewer seems to allude to. What is widely accepted is not our analysis method, but the particular definition of cloud types. (Or perhaps the reviewer's comment was prompted by our statement that "Our study aims to go beyond widely known cloud-precipitation associations…" in p. 2, line 26?) In any case, we appreciate the opportunity to provide clarifications on the distinction between cloud types and cloud regimes, both derived from the same CTP-COT histograms.

2. Many of the conclusions are based on regime-composite Pearson correlation coefficients between cloud area fraction and precipitation intensity percentiles. The Pearson correlation coefficient is fraught with pitfalls. The authors would greatly assist the reader in his or her assessment of the robustness of the conclusions by providing:
(a) a representative scatter plot of the correlated variables in the case of a strong positive correlation and a strong negative correlation and
(b) a geographic map of correlation strengths for the strongly positively and negatively correlated cloud/precip categories to see, e.g., whether the subsidence regions, ITCZ, warm pool, SPCZ, and maritime continent contribute as expected to the global-mean positive and negative correlations.





We thank the reviewer for the thoughtful comment that prompted us to conduct additional analysis. The figures above (now the supplementary Figs. 2 and 3) show 2-dimensional histograms of P4+P5 rainfall fraction and cloud type fraction co-occurrence (2D histograms are more appropriate than scatterplots given the large number of points). The upper panels are for *Cb*, and the lower panels for *Cu*. The samples are conditional to P4+P5>0, same as for the correlation shown in Figs. 8b or 9b. Gray color indicates very small percentage less than 0.01%, and white color indicates 0%. The histogram bin size is 1/16 (=6.25%), and the bin labeled as "50%" indicates bin boundaries from 46.875% to 53.125%. One can see that (for *Cb* CF=0 bin) heavy precipitation can occur at instances even when there is no *Cb* cloud, and that the probability of strong precipitation is much higher for small or zero *Cu* fractions (which manifests as an anti-correlation in our correlation "pyramid" plots.



The above figures (also the supplementary Fig. 4) show correlation coefficients for each grid cell at 1°×1° scale. The regions of abnormally high or low correlation values (e.g., *Sc*-dominant regions, the Sahara, the Himalaya, etc.) usually have small sample size for both these cloud types. Positive correlation between heavy rainfall and *Cb* cloud appears independent of location when there are enough samples. The anti-correlation between heavy rainfall and Cu cloud is more notable in oceanic subsidence regions, and weaker over land or convective regions (e.g., warm pool region, ITCZ, SPCZ).

The relevant text added is as follows:

(P10 L10) "In order to get a sense of the physical reality represented by Pearson's *r,* we examined two-dimensional histograms of cloud type CF and P-group for both strong positive and strong negative correlations (Supplementary Figs. 2 and 3). We note that more samples are available for zero or small amount of cloud type fraction for each case, and the distribution patterns look otherwise reasonable. We also examined the geographical dependence of these correlations and found them generally insensitive to location (Supplementary Fig. 4)."

3. According to the authors, the TMPA precipitation dataset uses cloud-top temperature to fill in precipitation information where radar is not available. Since cloud-top height information is also used in the regime definitions, I would expect some amount of potentially spurious correlation. Discussion of whether this effect has been considered would be appropriate in the text.

This is a valid point. Spurious correlations may indeed arise due to the use of the IR information to identify both cloud type and estimate surface precipitation rates. Since there is a physical relationship between cloud top temperature and precipitation, it is difficult to disentangle the physical effect form the spurious effect.

However, we expect the spurious effect to be sufficiently tempered for two reasons. First, TMPA and MODIS use different sources (and wavelengths) of IR information. TMPA uses the microwave precipitation rates to calibrate IR brightness temperatures and establish spatially varying relationships to IR-based precipitation rates. On the other hand, MODIS uses longer wavelengths of IR data ("CO2 slicing" for high and mid-level clouds producing most of the precipitation) for estimating cloud top altitude than those used for precipitation, which are window brightness temperatures. Moreover, cloud optical thickness, the other dimension used to define cloud type does not come from the IR, but from shorter, near-visible solar wavelengths. The fact that TMPA precipitation is based primarily on the (radar-calibrated) microwave observations, which in turn calibrate the IR brightness temperatures, helps a lot in making the datasets substantially independent.

Minor comments:

• *Section 2.2, "If the number of bins in the histogram is chosen to be also 16, each bin value falls between 0 and 1 in multiples of 1/16, the sum of all histogram bins at 1° grid cell is equal to 1, and sub-grid precipitation rates are thus converted to areal fractions of specific ranges of precipitation rates": I don't quite follow the why 16 is a magic number in the link between the number of bins and area fraction; since we end up with 6, not 16, bins in Sec. 2.3, are those not area fractions anymore?*

The reviewer is correct, and 16 is indeed not a magic number. We correct the sentence accordingly.

(P5 L1) "Hence, each bin value falls between 0 and 1 in multiples of 1/16, and sub-grid precipitation rates are interpreted as *areal fractions* of specific ranges of precipitation rates."

• *Fig. 5: define what is meant by "climatology".*

We changed the caption of Fig. 5 to avoid the ambiguous term "climatology".

"Conditional composite mean of 2D joint histogram of pc and τ (left column), differences from overall (unconditional) mean (middle column)…"

• *p. 11, first paragraph: I find the claimed link between P4, P5, and MCS tenuous; for example, if P5 indicates MCS (where we expect clouds at all levels), why are both Cb and Cs anticorrelated with low- and mid-level clouds?*

According to our definition of *Cb*, it has high cloud top altitude with large optical thickness, which means that the cloud vertical structure is quite deep and probably extends to low altitude. Hence, at the 1°×1° grid level, more fraction of *Cb* means less fraction of all other cloud types. This is less true for *Cs* cloud, so it is more likely that *Cs* and low cloud co-exist at the same location than for *Cb*. However, the low cloud under the high cloud can remain undetected by MODIS due to the inherent limitations of passive sensors. Moreover, we should re-emphasize that the separation between *Cs* and *Cb* is quite inexact given the way these cloud types were defined by ISCCP.

• *Anticorrelation in the Cu case: I am surprised that Cu is so anti-correlated with rain; I always thought (perhaps my thinking is guided by the regime name, which the authors caution against) that this would be the regime that clouds with high in-cloud water content but low area fraction (hence low grid-scale optical thickness).*

One advantage of using MODIS 2D joint histogram of cloud is that the data preserves the sensor's pixel level information at the grid scale, so no averaging (or interpolating) pixel information to grid level takes place. As the reviewer realizes while heeding our advice for caution, the *Cu* cloud in our study is different from the *Cu* cloud that comes to mind in other situations (like the cumulus congestus he/she seems to visualize). The *Cu* cloud from MODIS joint histogram (according to ISCCP definitions) has low optical thickness even at small spatial scales (MODIS pixel resolution is approximately 1km). We added a sentence about this point.

(P6 L7) "While these cloud types were given the same names as the standard cloud types seen by human observers from the ground and have some affinity with them, they are only loosely connected with the widely recognized traditional cloud types."

• *Anticorrelation in the Cu case (still): It would be interesting to get to the bottom of whether this is a real effect (CAPE/stability) or shadowing artifact, and I think the authors could easily do it by looking at CloudSat profiles (since they are already using MODIS data, not much additional co-location would be needed). If it is an artifact, does that mean all of Fig. 8 could be simplified to just the first row of every 3×3 plot? (By the way, I think the matrix of additive/subtractive Pn>0 subsets in Figs. 8 and 9 is brilliant plotting strategy.) Anyway, my first guess at the source of the anti-correlation was open vs. closed-cell stratocumulus, and it was interesting to learn that that was not the reason.*

The obscuration effect for lower clouds definitely exists as does the tendency of certain cloud type combinations to not co-occur (combined effects are expressed by the results of Fig. 10). Still, the negative relationship between heavy precipitation and *Cu* cloud would likely remain even without these effects, because, as pointed out above, the *Cu* cloud in this study is defined as optically thin cloud with low cloud top which is not expected to precipitate. Given CloudSat's limitations in the detection of boundary layer and puny clouds (e.g., TAU < 3.6), especially under conditions of signal attenuation in the presence of *Cb* hydrometeors, we're not sure whether embarking to such an investigation would pay dividends. The 3×3 plot can probably not be simplified to just the first row because for weaker precipitations the CF of the high cloud types decreases so the likelihood of lower cloud obscuration also decreases. In other words, the degree of obscuration is not independent of the precipitation rate.

• *The other surprise for me in Fig. 8 is that cor(CF, f_prec) never goes above 0.6. All CBs precipitate, so I would expect the Cb CF should correlate much more strongly with f_prec. What am I missing?*

A main reason behind the apparently low correlation value is the non-rigorous definition of *Cb* from the joint histogram. As we showed in Fig. 3, when *Cb* cloud fraction is larger than 6.25% (= 1/16), the probability of any kind of precipitation is 0.9, which seems to be consistent to the reviewer's intuition. However, our correlation comes from area fraction of specific precipitation and cloud type at 1° grid scale. In nature, it is possible that heavy precipitation comes from clouds other than *Cb*, and it is also possible that (at least some part of) *Cb* cloud (as defined here) does not produce heavy precipitation. Furthermore, it is worth noting that, when we tested the same methodology with (temporally and spatially) higher resolution dataset (e.g., GPM-IMERG and CMORPH), we obtained higher (above 0.7) values for the correlation coefficients. So TMPA's spatiotemporal resolution may be dampening the correlations in this case.

• *Sec. 4: it should be clarified that the first paragraph is an aspirational statement about the cloud-physics field as a whole, since this study is an incremental advance*

We have rephrased to provide the clarity that the reviewer is seeking about the reach and scope of our work.

(P16 L7) “Even with such non-ideal data at hand, the community still aspires to answer fundamental questions such as: To what degree can precipitation be predicted given information about clouds? Conversely, with precipitation information at hand, can we provide good guesses about the nature of the clouds responsible? Is precipitation variability associated with cloud variability? Do answers to the above questions differ substantially between ocean and land? This paper seeks to contribute ideas and results that will help us make progress in obtaining concrete answers in the near future, especially if observations also make considerable strides.”

• *p.16 l.16: if "once detection of low clouds in the presence of high clouds and of warm rain over land improves" refers to the use of active rather than passive satellite sensors, the authors may be interested in Field and Heymsfield or Mulmenstadt et al (both 2015, GRL)*

Thank you for the suggestion. We now cite these papers accordingly.

• *p.16 l.20: The authors chose not to use L3 instead of L2 data, presumably for reasons of data management complexity. I don’t think anyone would fault them for this choice, so the defensive tone of this sentence is out of place. Either that, or I misunderstood something about it.*

The intention of the sentences in this paragraph is to contrast this study from others using L2 data. The sentence has been rephrased:

(P17 L7) “Our self-imposed objective to make the study general, multi-year, and applicable to half of the Earth’s surface, led us to Level-3 gridded data as the most appropriate choice. While some of the details seen in previous studies that used Level-2 data will unavoidably be lost, our datasets are good enough to extract major features of cloud-precipitation co-variability and allow us to claim that they are broadly representative of this co-variability in the tropics.”

• *p. 16 l.25: No objection to citing unpublished work, but why not also some published references that show the same thing, e.g., Suzuki et al (2015, J Atmos Sci), Jing et al (2017, JGR)*

The publication Tan et al. has been accepted, and we have updated the citation. Thank you for your other suggestions, which have now been added too.