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 #1

Received and published: 14 September 2017

General comments:

Accurate knowledge of the relationship between clouds and precipitation is a key aspect for climate and earth-system models. Hence, many research studies previously addressed this issue by exploring cloud-precipitation feedbacks using, e.g. numerical and synoptic approaches, on various scales. The manuscript acp-2017-612 revisits this topic and aims at improving the present understanding of daytime cloud-precipitation co-variability by looking at collocated cloud and precipitation observations over a very large spatial domain, i.e. tropical oceans and land, with improved spatial and temporal matching compared to previous studies. The presented method is well suited for analysing this particular coupling and the authors provide an exhaustive analysis based on available data advancing the knowledge on this topic. Many interesting question are addressed and the authors draw plausible conclusions such as the stronger positive correlation between cumulonimbus clouds and heavy precipitation over oceans as opposed to over land. The manuscript is generally well written and the results are presented in a good manner. The manuscript fits well within the scope of Atmospheric Chemistry and Physics and I recommend publishing it after the authors have corrected several minor issues. In general, the manuscript contains heaps of detail but it is unfortunately kept very descriptive, and generally lacks more conclusions and practical implications of the presented findings. In addition, it would be highly beneficial if the authors stress the importance of their results and explicitly state what new scientific insights have been discovered by means of their analysis and which previous knowledge could be confirmed / or rejected. The conclusion section should be suitable for this. A critical examination of shortcomings of the method and data sets is quite brief and it is not entirely clear to what extent the individual sensitivity of the chosen data products for clouds and precipitation may bias the results. In specific, would the authors come to similar conclusions with different data products?

We thank the reviewer for the overall positive assessment. At the outset of the study we aimed to answer specific questions (P3 L3-6) and we believe we generated sufficient results to shed light on these issues. However, we may have not always provided direct answers, something we're trying to improve upon in the revised version. We think that the value of this study rests in employing a novel methodology where cloud histograms and matched precip histograms are jointly analyzed to re-examine in an extensive semi-global-domain results previously reported in local/regional studies. We quote many specific results in the concluding section, where every cloud type's correlation with precipitation is summarized. We also consider the conclusion that shallow continental clouds are better anti-correlated with heavy precipitation than correlated with light precipitation very specific and useful from the perspective of current precip and cloud measurement capabilities and our ability to predict precip from cloud information. Following the suggestion of the reviewer, the higher-level summary has been refined as well ("While some … this conclusion"). Regarding the biases of the specific datasets, explanations are provided below in responses to specific comments.

(P17 L8) "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. We argue that the insensitivity of cloud-precipitation relationships to location (supplementary Fig. 4) and precipitation dataset (initial tests with recent GPM-IMERG data that may be presented in a future study yielded similar results) strengthen the validity of this conclusion."

Specific comments:

P2 L26: Please provide the motivation for your study already here and explain in more detail what is missing in previous research studies. At this point it is not clear to the reader why this topic needs to be touched 'once again'.

Motivation was provided in the paragraph from P2 L26 to P3 L7. See also response to related question below.

P2 L32: Move this part up accordingly.

To address the above comment, the paragraph has been re-written:

(P2 L28) "We thus strive for generality of results by covering the entire tropics and for overcoming the ambiguity of CR-based studies by taking advantage of the ability to break down individual grid-box cloud fractions with the aid of joint cloud histograms. Hence, our paper revisits and explores anew the mesoscale cloud-precipitation relationship via the synoptic approach by employing a Moderate Resolution Imaging Spectroradiometer (MODIS) gridded cloud dataset (King et al., 2003; Platnick et al. 2003) and the TRMM Multi-satellite Precipitation Analysis (TMPA) dataset (Huffman et al., 2007, 2010). While the MODIS Level-3 data are provided at 1°×1° resolution, the 2D joint histogram of cloud optical thickness (τ) and cloud top pressure (p_c) contains pixel-level cloud information which can be combined with the sub-grid variability of precipitation at the 1°×1° scale, available by virtue of the finer 0.25°×0.25° spatial resolution of TMPA. While still coarser than the TRMM PR dataset, the combined MODIS and

TMPA dataset covers the entire tropics every single day, allowing better generalization of the daytime relationship between clouds and precipitation. We seek to answer questions such as: …"

P3 L1: What is meant by 'ambiguity' exactly?

This is in reference to lines 13-14 of p. 2 where it is noted that the CR internal variability can be quite large.

P3: Please elaborate in detail why the MODIS and TRMM TMPA were chosen. It would be highly beneficial to discuss and argue why these two data sets are more suitable than other similar data sets for your study. What about other global precipitation products providing 3-hourly rain estimates such as CMORPH (Joyce et al., 2004), PERSIANN (Sorooshian et al., 2000) or others? The sensitivity to precipitation and different cloud types may be very different amongst these products which could potentially strongly affect your findings.

The main reason for choosing MODIS and TMPA was our more intimate knowledge of these datasets. Moreover, our working assumption was that in such well-characterized Level-3 datasets, severe biases (if they do exist) are limited or well-known. On the precipitation side, we applied our methodology to GPM-IMERG and CMORPH for three recent years, and found that the correlation pyramid plots are quite similar to the one shown in this study. Only minor differences in the coefficient numbers were seen, and different peak locations for light precipitations, particularly with GPM-IMERG (as expected, since TMPA exhibits a weakness in light rainfalls as pointed out in the manuscript). However, at least for the P3, P4, P5 precipitation ranges, we didn't find any fundamental deviations from our current TMPA precipitation dataset. A sentence is added to the concluding section related to this issue. With regards to clouds, MODIS is considered the state-of-the-art provider of Level-3 1° gridded cloud histograms. Had we used ISCCP, we would have resorted to a coarser resolution of 280km and smaller overlapping period with TMPA.

(P17 L11) "We argue that the insensitivity of cloud-precipitation relationships to location (supplementary Fig. 4) and precipitation dataset (initial tests with recent GPM-IMERG data that may be presented in a future study yielded similar results) strengthen the validity of this conclusion."

P4 L24: Please be consistent in naming and differentiating between the grid at 1° resolution and a lower resolution grid at 0.25° resolution throughout the paper (grid cell, sub-grid cell, sub-grid, sub-cell, etc.)

Thanks for pointing this out. We now use consistently "1°×1°" and "sub-grid."

P5 L3-11: The spatial and temporal collocation is crucial for this type of study as precipitation and rain patterns may vary quickly. Please provide a better explanation of the temporal matching of both data sets and provide a reference for the time *conversion of the MODIS data, if possible. The TRMM TMPA 3B42 3-hourly product provides the satellite observation time for each grid cell. Was this information used for the matching? Note that actual observation times for each grid pixel may vary +-90 minutes within a 3-hourly data file. For example, if the TMPA 3B42 12UTC data file is chosen for collocation with MODIS Aqua data the maximum time difference between the MODIS and TMPA data could be more than 1.5 hours. It is generally not quite clear to the reader how the non-trivial collocation of the data sets is performed across all longitudes.*

We are basically matching MODIS UTC time to the closest UTC time in the TMPA dataset. As a first step, we need to calculate an approximate UTC time for MODIS since it is not provided in the L3 data. We accomplish this using the grid-mean solar zenith angle θ_0 and the following equation:

$$h = \cos^{-1}\left(\frac{\cos\theta_0 - \sin\varphi\sin\delta}{\cos\varphi\cos\delta}\right)$$

where h is the hour angle, φ is latitude, and δ is solar inclination angle. [Reference: Liou, K. N.: An introduction to atmospheric radiation, 2. ed., International Geophysics Series. Vol 84, Acad. Press, San Diego, 2002]. The solar inclination angle for a particular latitude is a function of the Julian date. Once the hour angle (h) is obtained it is converted to UTC. With the MODIS UTC at hand, we search for the TMPA 3-hour interval that contains it. For example, a TMPA data point designated as 12pm UTC contains rainfall observations from 10:30am UTC to 1:30 pm UTC, and is selected for that 1°×1° grid cell if the MODIS UTC also falls within that time interval. Hence, there is indeed possibility for a maximum time difference greater than 1.5 hours. However, we don't think that this fact can affect our results significantly because: 1) The distribution of satellite observation times within a 1°×1° grid cell and a 3-hour interval is probably closer to a Gaussian distribution rather than Uniform distribution because, when two or more satellite rainfall observations are available, the one closest to the middle of 3-hour period is selected by the algorithm, and 2) Time mismatch biases are random and largely cancel out due to huge sample size. The text in p. 5 has been changed accordingly to further elaborate on these issues, and a statement about maximum time difference has been removed.

(P5 L11) "The UTC of each grid cell can be estimated from the mean solar zenith angle (SZA) available as a MODIS Level-3 variable, and the latitude and time information for each grid cell."

(P5 L15) "Since the TMPA data is available at 3 hour-intervals, TMPA data centered, say, at 12 pm UTC, will be matched with MODIS data having UTC between 10:30 am and 1:30pm."

P5 L19-21: What could be the explanation for this?

For comparing Aqua matched precipitation against all available precipitation, differences in missing is 5.3% (Aqua: 5.34%, All: 0.04%), and in "No Rain" is -4.19% (Aqua: 85.28%, All: 89.47%). Hence, 1.11% of All data (5.3%-4.19%) is distributed to various precipitation histogram bins, 0.07% per bin on average, thus it is normal that All available data (green bar) is slightly higher than Aqua-matched data (red bar).

For the weak-to-moderate precipitation rate, the difference is slightly larger despite the intrinsic difference explained above. We interpret this as light rain being more frequent outside the time window around noon. We have rephrased the relevant passage ("This appears in Fig. 1 … overpasses") in order to reduce any possible confusion.

(P5 L24) "This appears in Fig. 1 as Terra-matched precipitation having smaller frequencies than the original and the Aqua-matched precipitation, although it is somewhat improper to directly compare Terra- or Aqua-matched data with fully sampled data because the higher ratio of available (non-missing) data in the fully sampled data propagates as higher relative frequency in the various precipitation bins. It is also notable that, for weak-to-moderate precipitation rate (less than 1mm/hr), even Aqua-matched precipitation is (slightly) lower in percentage terms than fully-sampled TMPA precipitation, which can be interpreted as weak-to-moderate precipitation being more frequent outside the time windows of Terra and Aqua overpasses."

P6 L7-9: Why did you choose to consider the MODIS Terra and Aqua as a single ensemble, even though initial results from Fig. 1 point at notable differences in precipitation during the different overpass times? Would you argue that this has no effect on the found cloud-precipitation relationships? Also, it is not clear to the reader how the exact matching of MODIS and TRMM TMPA data is performed. See remark above.

The Terra- and Aqua-matched precipitation does indeed look slightly different in terms of the sample distribution of Fig. 1. At the same time, the cloud type distribution should also not be exactly the same between Terra and Aqua considering the diurnal cycle of cloudiness in some locations (especially continents). However, in this study, we are looking for a general relationship between clouds and precipitation that signifies physical processes (that may be different between ocean and land), and we therefore assume that this relationship is NOT affected by the frequency of specific types of cloud or precipitation. This is the main reason discrimination between Terra and Aqua data was not considered a priority in this study. We actually verified behind the scenes that separate correlations for Aqua and Terra are not too different. A second reason is that we are already breaking down the results by land/ocean, and an additional decomposition would make the presentation cumbersome and hard to follow.

P7 L32-33: Please explain why this is the case for P5 and not for P4 and name the common characteristics with MCS explicitly.

In the manuscript, we interpreted MCS as the combined system of strong stratiform precipitation paired with Cb clouds. In the P4 and P5 PC-TAU composites of Figs. 5 and 6, the P4 composite looks more closely related to Cs than Cb, while the P5 composite shows as many above average Cb clouds as Cs clouds. This enhanced Cb CF is the reason why we relate P5 to MCS and not P4. The sentence has been rephrased to clarify this.

(P8 L12) "The P5 composite patterns of cloud and precipitation shown in Fig. 5 are in accordance with such MCS characteristics, i.e. strong convective clouds and a broad spectrum of precipitation."

P9 L11: It would be worthwhile to explain the effect of autocorrelation between neighboring grid cells in more detail and how this is accounted for.

We have added a sentence ("Consideration for … underestimated") explaining that when not accounting for the autocorrelation the degrees of freedom are overestimated and thus the significance level underestimated. In other words, because of fewer independent measurements it is harder to surpass the threshold of significance. We refer to the citations for explanation since we don't think a digression is appropriate in this case.

(P9 L23) "Consideration for the effect of neighboring grid cells is important because neighboring grid cells are usually *not* independent (e.g., a cloud system can occupy multiple grid cells); without this consideration, the degree of freedom will be overestimated, and thus the significance level underestimated."

P9: How certain are the authors that the calculated correlation coefficients between the cloud types and precipitation data can be interpreted as a 'general relationship' and not just representing the sensitivity of the TRMM TMPA algorithm to different cloud types?

As stated in response to a previous comment, we repeated our analysis with GPM-IMERG data and CMORPH, and didn't find any fundamental deviations from the results shown here, which supports the "general relationship" interpretation.

P13 L2: It would have been quite interesting to see how the correlation coefficients for Fig. 11 change if precipitation frequencies are not progressively added for each bin. Could you provide such a Figure or give a reason why this may not be useful?

Response:

This information is actually plotted, albeit at coarse binning: it exists in the P1>0, … P5 >0 (rightmost of each row) panels of the pyramid plots of Figs. 8 and 9. The more detailed graph that the reviewer suggests can now be found in the supplementary Fig. 6. It is the same format as Fig. 11, but the x-axis starting at the 1st-3rd bins, running sum of three consecutive histogram bins. The plot looks consistent to Fig. 11, and with the other results of our study.

P16 L27-28: This sentence sounds a bit strange, suggesting that you might not have chosen the optimal data sets for your study in the first place. It would be better to discuss in more detail how your results could be validated against results derived from other or future data products.

This study provides a methodology of how to quantify the cloud-precipitation co-variability, and optimal datasets can vary depending on the purpose of a study (e.g., high resolution data would be need for a regional/seasonal study). Optimal data may not even exist now, but may become available in the future, e.g., cloud and precipitation observations of higher sensitivity from the same observational platform which would eliminate a lot of the uncertainty of inexact spatiotemporal matching. This sentence also alludes to the possibility that a higher spatial and temporal resolution precipitation dataset such as GPM-IMERG could be used for this type of analysis once the period of availability has extended substantially. The last sentence of the text has been rephrased to convey what we had in mind more clearly.

(P17 L17) "In addition, more effort should be extended to apply the framework in this study to various case studies with more appropriate datasets (e.g., using higher resolution precipitation dataset for regional/seasonal studies, or longer period dataset for climate studies) in order to increase further our degree of confidence about the cloud-rainfall relationships."

Technical corrections:

P1 L22-23: Please rephrase to make points clearer

We have rephrased as follows:

(P1 L22) "Weak correlations between weaker rainfall and clouds indicate poor predictability for precipitation when cloud types are known, and this is even more true over land than over ocean."

P1 L30: 'models' instead of 'model'; or better rephrase the first part of the sentence

We think the sentence reads fine after changing "model" and "AGCM" to plural.

P2 L18: Please rephrase very long sentence

We're not sure which sentence the reviewer is referring to here. None in the paragraph starting in L15 seem overly long. Nevertheless, we broke the sentence starting in L17 into three sentences. The relevant text now looks like this:

(P2 L16) "An example of this is the "cloud and precipitation feature database" of Liu et al. (2008). The database was derived from observations by the precipitation radar (PR), the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI), the Visible and Infrared Scanner (VIRS), and the Lightning Imaging System (LIS) aboard the TRMM satellite. The authors performed several case studies with this dataset that contrasted continental and oceanic precipitating cloud systems, and found that …"

P2 L20: 'larger extent' - please state whether horizontal or vertical extent is meant

Horizontal extent. We have now clarified.

P2 L26: Please specify references to the datasets used

We added the references. We also provide references to the datasets later in section 2.

P2 L33: 'examines' instead of 'examine'

Done.

P3 L3: 'Is there a closer' instead of 'Is there more close'

Fixed.

P3 L11: Please provide references for the MODIS instrument and specify the exact name of the dataset.

In the sentence that immediately follows we provided such a reference "The MODIS cloud dataset (King et al., 2003) provides Level-3 …" We also added Platnick et al. (2003) and now call out the MODIS cloud dataset specifically.

(P3 L16) "The MODIS cloud dataset (MOD08_D3 and MYD08_D3; King et al., 2003; Platnick et al., 2003) provides Level-3 cloud products at daily time scales with 1°×1° horizontal resolution."

P3 L22: Please rephrase and/or explain why a lower number of bins was chosen.

The reason of reducing dimension is provided in subsection 2.3 when we explain the correlation method. Without coarsening we don't have correspondence with the ISCCP cloud types. The added sentence reads as follows:

(P3 L26) "In this study, the joint histogram bins are coarsened from 42 bins to 9 cloud types because of practical considerations (see subsection 2.3) as well as our desire to draw an analogy with the ISCCP cloud types (Chen et al., 2000; Rossow and Schiffer, 1999)."

P3 L26: leave out 'best'

We just follow the original expression in Huffman et al. (2007): "…, with the goal that the final product will have a calibration traceable back to the single "best" satellite estimate." Our intention was to convey the philosophy of TMPA.

P4 L1: Specify the overpass time of MODIS Terra and Aqua in local time / equator crossing time.

The equator crossing time, 10:30am for Terra and 1:30pm for Aqua (LST) has already been provided in the sentence.

P4 L14: Please rephrase part with 'algorithmic variations'

Changed to:

(P4 L20) "by these algorithm differences"

P4 L19: Please rephrase the first part of the sentence

The first part of the sentence has been re-written as follows:

(P4 L25) "Because the 3B42 dataset has higher spatial resolution than the MODIS Level-3 cloud dataset, we resample it to the…"

P4 L21-22: Use consistent format of grid resolution, i.e. either 1° or 1x1°

We now use 1°×1° throughout.

P4 L24: 'of a histogram', instead of 'of histogram'

Fixed.

P4 L24: 'without missing values' instead of 'when no missing values exist'

Thank you for suggestion, but it is actually rephrased as:

(P4 L31) "when there are no missing values"

P4 L32: 'shows the distribution' instead of 'shows distribution'

Fixed.

P5 L3: Please rephrase the sentence, for example: '... the TMPA and MODIS observations also need to be matched in time.'

Thank you, it does indeed sound better as suggested.

(P5 L8) "…, the TMPA and MODIS observations also need to be matched in time."

P5 L15: Please rephrase 'in explaining'

The sentence has been re-written as follows:

(P5 L21) "Other differences in occurrence frequencies between original and matched data are probably due to the diurnal cycle of precipitation."

P5 L16: Please rephrase 'For example' at the beginning of the sentence

"For example" has been removed.

P5 L17: What is meant exactly by: 'relatively suppressed'?

Changed to "relatively weak".

P5 L28: Please rephrase sentence starting with "This is simply ... " as this would have been probably possible and could, in fact, provide additional insight, but was not pursued for practical reasons.

We rephrased as follows:

(P6 L3) "Analysis and visualization of such a large number of coefficients are impractical, hence we pursue an analysis where both the cloud and precipitation histograms are coarsened.

P6 L4-7: Please rephrase last part of sentence starting at "with no confusion resulting"

Rephrased as follows:

(P6 L16) "For simplicity, the same symbols are henceforth also used to represent the frequency of occurrence within these groups, since their meaning is always clear by the context."

P6 L13: Please explain what it is meant by the co-variability of anomalies.

This is a convoluted way to simply say "correlations" (calculated from deviations from the mean), so now we just say "correlations".

P6 L27: Replace 'of no-rain case' with 'of the no-rain case'

Changed as suggested.

P7 L3: Please indicate in which section the issue of less rain over land is analysed.

Done.

(P7 L16) "The issue of less rain over land is also covered in the next composite plots (Figs. 5 and 6)."

P7 L4: Please make it clearer to reader what you mean by the 'composite mean cloud and precipitation histogram'

The composite means are cloud and precipitation histograms that were conditionally averaged. The condition in this particular case was that at least one occurrence of P4 or P5 precipitation existed in the 1°×1° grid cell, as described in the line that follows.

P7 L19: 'in the P4 group' instead of 'due to the P4 group'

Rephrased as suggested.

P8 L13 and L16: is the increase really linear

"linearly" was replaced by "monotonically"

P8 L29: Please rephrase 'a factor affecting the Fig. 7 results"

Rephrased to:

(P9 L9) "…, may be affecting the land results of Fig. 7".

P10 L6: Please rephrase 'the peak negative value is a weaker value of'

Rewritten as follows:

(P10 L27) "…, the peak negative value weakens to -0.23 and …"

P10 L24: Please rephrase the second part of the sentence to make clear what you mean exactly.

Rephrased as follows for clarity:

(P11 L10) "…, but similar peak correlations over land occur even for Cs and Cb."

P12 L8: Please rephrase 'but suffice it to say here'

We don't think that this needs to be rephrased.

P13 L11: Please rephrase the beginning of the sentence (not start with 'But')

"But" is replaced by "However,"

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.

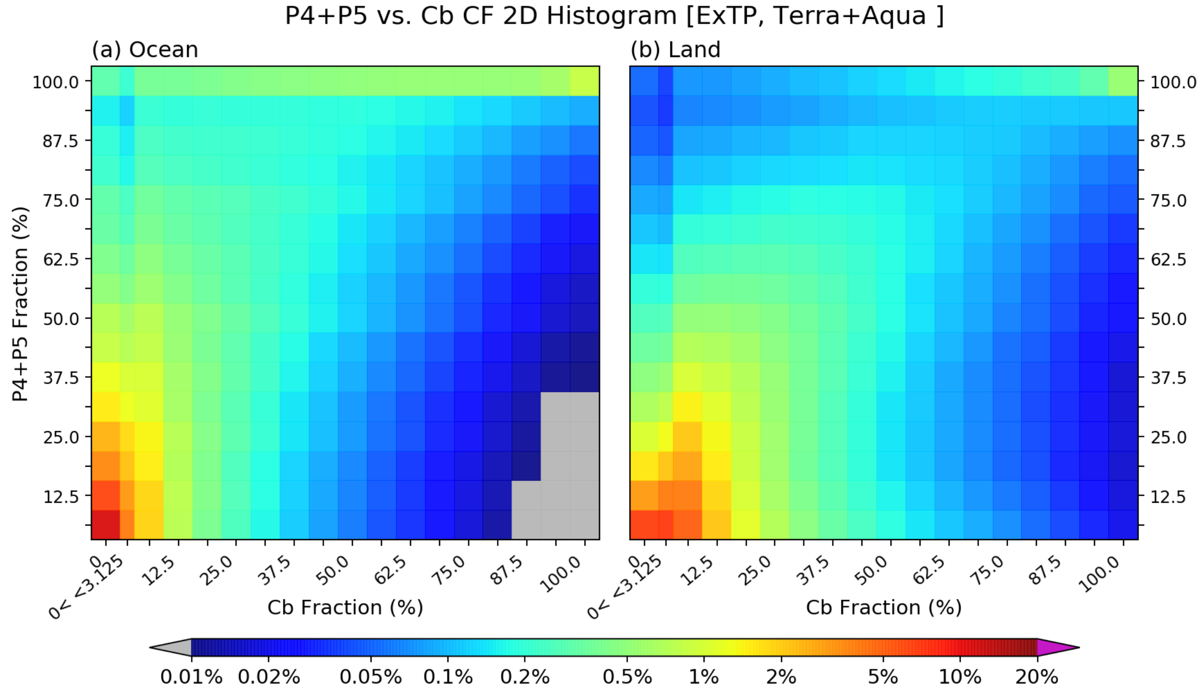
Major comments:

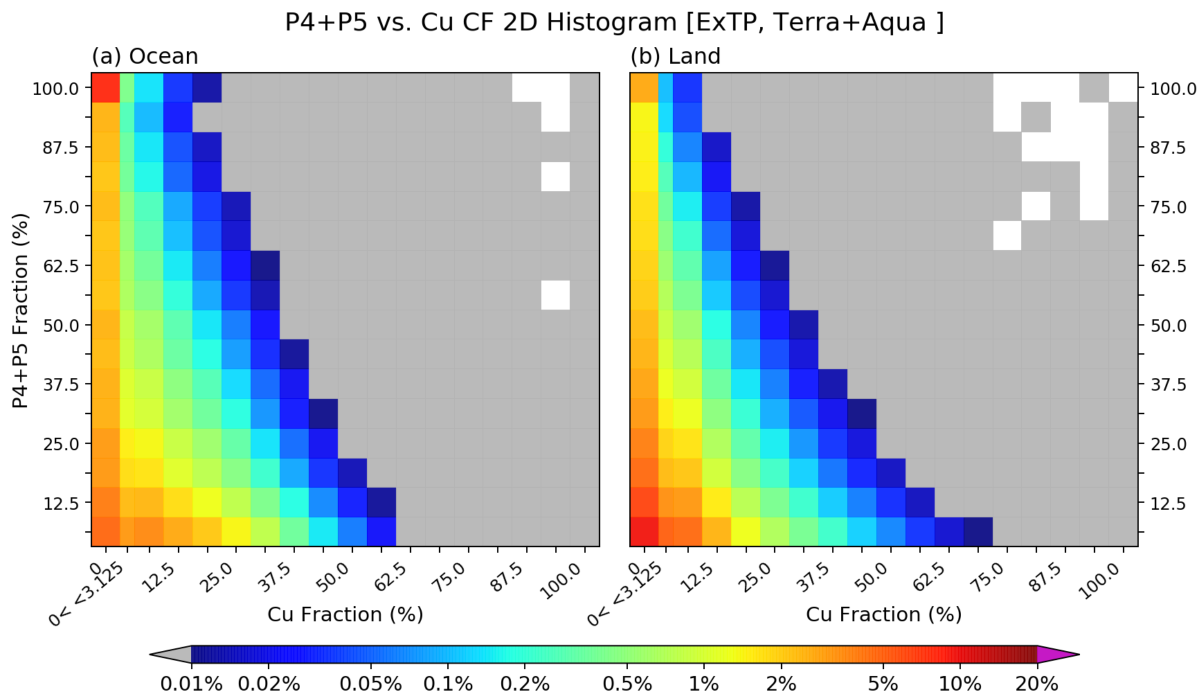
The main potential weaknesses of the analysis are the following:

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.

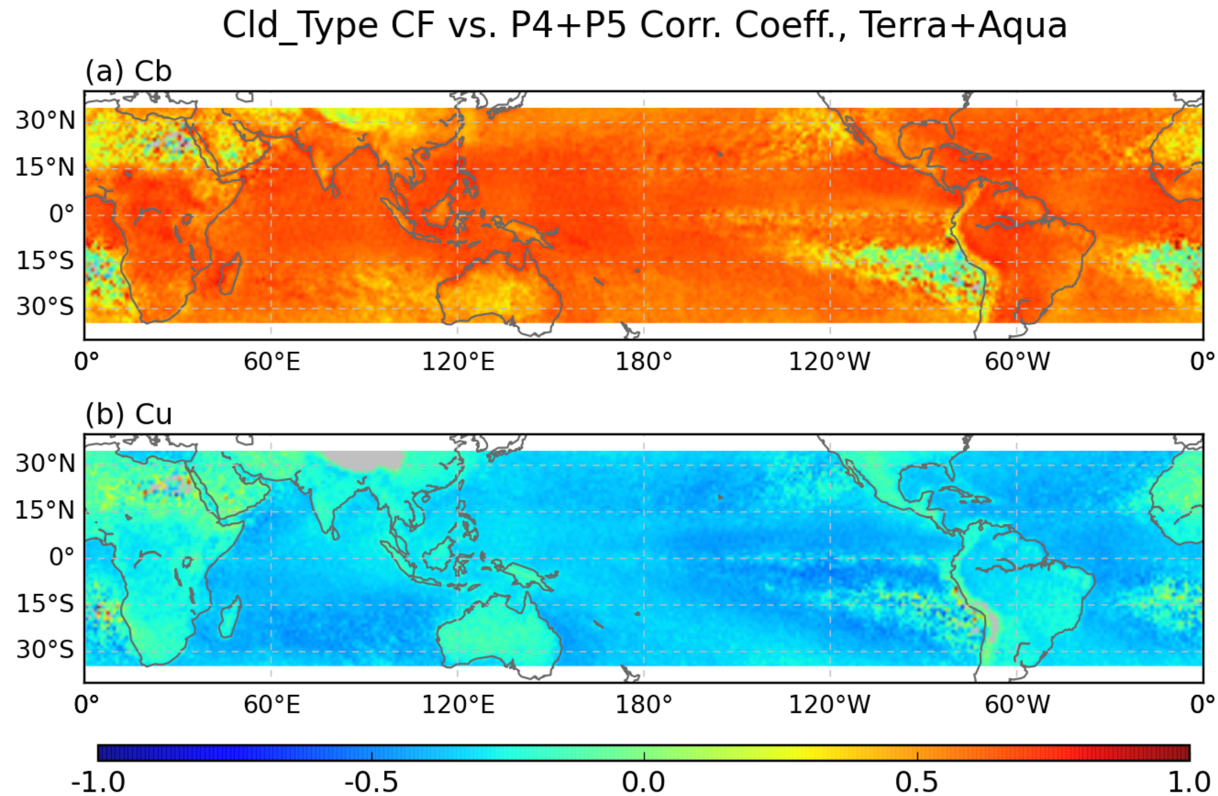
We take the opportunity here to clarify that our work uses the concept of "cloud type" rather than that of "cloud regime" (aka "weather state") as in Oreopoulos et al. (2014, 2016), the Gryspeerdt work quoted above, and previously in Rossow et al. (2005), among others. This is an important distinction. 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.

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

Daeho Jin[1,2], Lazaros Oreopoulos[2], Dongmin Lee[3,2], Nayeong Cho[1,2], and Jackson Tan[1,2]

[1]University Space Research Association, Columbia, MD, USA
[2]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[3]Morgan State University, Baltimore MD, USA

Correspondence to: Daeho Jin (Daeho.Jin@nasa.gov)

Abstract. The co-variability of cloud and precipitation in the extended tropics (35°N−35°S) is investigated using contemporaneous datasets for a 13-year period. The goal is to quantify potential relationships between cloud type amounts and precipitation events of particular strength. Particular attention is paid to whether the relationships exhibit different characteristics over tropical land and ocean. A primary analysis metric is the correlation coefficient between fractions of individual cloud types and frequencies within precipitation histogram bins that have been matched in time and space. The cloud type fractions are derived from Moderate Resolution Imaging Spectroradiometer (MODIS) joint histograms of cloud top pressure and cloud optical thickness in one-degree grid cells, and the precipitation frequencies come from the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) dataset aggregated to the same grid.

It is found that the strongest coupling (positive correlation) between clouds and precipitation occurs over ocean for cumulonimbus clouds and the heaviest rainfall. While the same cloud type and rainfall bin are also best correlated over land compared to other combinations, the correlation magnitude is weaker than over ocean. The difference is attributed to the greater size of convective systems over ocean. It is also found that both over ocean and land the anti-correlation of strong precipitation with "weak" (i.e., thin and/or low) cloud types is of greater absolute strength than positive correlations between weak cloud types and weak precipitation. Cloud type co-occurrence relationships explain some of the cloud-precipitation anti-correlations. Weak correlations between weaker rainfall and clouds indicate poor predictability for precipitation when cloud types are known, and this is even more true over land than over ocean.

1 Introduction

Attempts to estimate precipitation from cloud observations have a long history dating back to the era of first passive thermal infrared observations of clouds (e.g., Richards and Arkin 1981). Enlisting numerical models to help with the interpretation of observations has not been as helpful as hoped since these models generally do not produce coherent relationships between clouds and precipitation (e.g., Stephens et al. 2010; Gianotti et al. 2012; Jiang et al. 2015), with even cloud-resolving models explicitly representing precipitation processes facing challenges in that respect (e.g., Kooperman et al. 2016; Matsui et al. 2016). In the case of atmospheric global circulation models (AGCMs), it is nearly impossible to resolve individual

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: the relationship

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: types

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: relationship exhibits

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: major

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: coefficients

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: over ocean

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: over land

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: that

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Couplings

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: are also distinct in ocean vs. land, with

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: predictability

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: type is

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: being quite poor in general, particularly

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: model (AGCM

precipitating processes due to the sub-grid nature of the problem and the excessive computational burden. Hence, for AGCM evaluation, and also for observation-based water budget studies, a synoptic approach for identifying the relationships between cloud and precipitation has been deemed an inevitable compromise.

One example of employing a synoptic approach is the use of the concept of a "cloud regime" (CR) also known as "weather state" (WS; Jakob and Tselioudis 2003; Rossow et al. 2005; Oreopoulos and Rossow 2011; Tselioudis et al. 2013; Oreopoulos et al. 2014, 2016) to study precipitation characteristics. CRs represent the dominant mixtures of cloud types, and can be used as a framework to categorize cloud data in a grid (e.g., Level-3 satellite products). Using the International Satellite Cloud Climatology Project (ISCCP) WSs defined in the extended tropics (35°S−35°N), Lee et al. (2013) provided a comprehensive picture of precipitation characteristics for each WS, with an additional focus on the relationship between the most convective regime (WS1) and precipitation. Rossow et al. (2013) also conducted similar analysis but for precipitation extremes using ISCCP WSs for the deep tropical zone of 15°S−15°N. While such CR-based approaches provide valuable information about the cloud-precipitation relationship at large scales, the precipitation composites by CR encompass large spreads which obscure details of the relationship. Since CRs contain mixtures of clouds types by design, and therefore contain considerable cloud variability, ambiguities in the cloud-precipitation relationships are hard to resolve.

Cloud-precipitation relationships can, however, be examined at a more detailed level with coincident precipitation profile and cloud measurements. An example of this is the "cloud and precipitation feature database" of Liu et al. (2008). The database was derived from observations by the precipitation radar (PR), the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI), the Visible and Infrared Scanner (VIRS), and the Lightning Imaging System (LIS) aboard the TRMM satellite. The authors performed several case studies with this dataset that contrasted continental and oceanic precipitating cloud systems, and found that oceanic storms were generally horizontally larger at 2 km altitudes, but continental storms tended to be vertically more coherent, with a higher top and more severe rainfall. Houze et al. (2015) also reported similar results using solely vertical rainfall profiles from the TRMM PR. While these studies provided a more detailed look at the cloud-precipitation relationship thanks to the high resolution of the TRMM PR (4-5km footprint at nadir), the penalty was narrow horizontal coverage (swath widths of 215 km before orbit boost and 247 km after orbit boost). Our study aims to go beyond widely known cloud-precipitation associations (such as geometrically deep and optically thick clouds producing stronger rainfall), and to examine instead more carefully the details of the connections between clouds and precipitation for situations that also include non-heavy precipitation. We thus strive for generality of results by covering the entire tropics and for overcoming the ambiguity of CR-based studies by taking advantage of the ability to break down individual grid-box cloud fractions with the aid of joint cloud histograms. Hence, our paper revisits and explores anew the mesoscale cloud-precipitation relationship via the synoptic approach by employing a Moderate Resolution Imaging Spectroradiometer (MODIS) gridded cloud dataset (King et al., 2003; Platnick et al. 2003) and the TRMM Multi-satellite Precipitation Analysis (TMPA) dataset (Huffman et al., 2007, 2010). While the MODIS Level-3 data are provided at 1°×1° resolution, the 2D joint histogram of cloud optical thickness (τ) and cloud top pressure (p_c) contains pixel-level cloud information which can be combined with the sub-grid variability of precipitation at the 1°×1° scale, available by virtue of the

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted:) built

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: TRMM

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Tropical Rainfall Measuring Mission (TRMM), and used to perform

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: land

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: ocean

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: . The authors showed

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: have

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: extent

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: tend

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: provide

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: is

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Our paper revisits and explores once again the potential of the synoptic approach. We employ Moderate Resolution Imaging Spectroradiometer (MODIS) gridded cloud dataset and the TRMM Multi-satellite Precipitation Analysis (TMPA) dataset to explore the mesoscale cloud-precipitation relationship. While the MODIS Level-3 data are provided at 1°×1° resolution, the 2D joint histogram of cloud optical thickness (τ) and cloud top pressure (p_c) contains pixel-level cloud information which can be combined with the sub-grid variability of precipitation at the 1° scale, available by virtue of the finer 0.25° spatial resolution of TMPA. While still coarser than TRMM PR dataset, the combined MODIS and TMPA dataset covers the entire tropics every single day, allowing better generalization of the daytime relationship between clouds and precipitation.

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: also

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: the 1° cloud fraction by cloud type.

finer 0.25°×0.25° spatial resolution of TMPA. While still coarser than the TRMM PR dataset, the combined MODIS and TMPA dataset covers the entire tropics every single day, allowing better generalization of the daytime relationship between clouds and precipitation. We seek to answer questions such as: What are the general expectations and limitations in predicting precipitation given a cloud type in the extended tropics? Is there a closer relationship between certain precipitation rates and cloud types? Do answers to the above questions differ substantially between oceans and continents?

The next section introduces the concept of "precipitation histogram" and how it can be matched and correlated to sub-grid cloud type fractions at the grid level. A comprehensive examination and interpretation of cloud and precipitation co-variability over tropical land and ocean follows in Section 3. In addition to summarizing the results, the concluding Section 4 calls attention to the new insights that emerge from this study and challenges that remain to be addressed about the nature of cloud-precipitation coupling.

2 Data and Methodology

2.1 Cloud and Precipitation Data

Our passive cloud retrievals come from the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument aboard the Terra and Aqua satellites. The MODIS cloud dataset (MOD08_D3 and MYD08_D3; King et al., 2003; Platnick et al., 2003) provides Level-3 cloud products at daily time scales with 1°×1° horizontal resolution. Among various cloud products, we focus on the 2D joint histogram of cloud optical thickness (τ) and cloud top pressure (p_c). The histogram is composed of cloud fraction (CF) values along 7 classes of p_c and 6 classes of τ (for a total 42 histogram bins), and contains pixel-level cloud variability information at the 1° scale. The most recent version of the MODIS atmospheric datasets, known as "Collection 6" (Platnick et al., 2017), provides a separate histogram for "partially cloudy" (PCL) pixels, flagged as such by the so-called "clear-sky restoral" algorithm (Pincus et al., 2012; Zhang and Platnick, 2011). The PCL pixels represent usually cloud edge pixels for which the cloud property retrievals are deemed more uncertain (Cho et al., 2015). We opted to include PCL pixels in our analysis by adding the PCL histogram to the nominal histogram because, by doing so, the MODIS cloud climatology becomes more consistent (see Oreopoulos et al. 2014) to that by ISCCP (Rossow and Schiffer, 1991, 1999), which has a long track record in cloud research and can potentially be used in a study similar to this one. In this study, the joint histogram bins are coarsened from 42 bins to 9 cloud types because of practical considerations (see subsection 2.3) as well as our desire to draw an analogy with the ISCCP cloud types (Chen et al., 2000; Rossow and Schiffer, 1999).

The precipitation dataset used in our study is the 3B42 research product (version 7) of Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) (Huffman et al., 2007, 2010; Huffman and Bolvin, 2015). The TMPA pursues the "best" satellite precipitation estimates using TRMM Microwave Imager (TMI) and Precipitation Radar (PR) data as calibrators in merging measurements from several microwave and infrared sensors, and monthly gauge data (over land) from the Global Precipitation Climatology Centre (GPCC; Huffman et al. 2007). The horizontal resolution of TMPA is 0.25°×0.25° covering 50°S to 50°N. TMPA is available from January 1998 with 3-hourly resolution, but we use

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: more close

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: subgrid

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: links that emerge from this study

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: As we will see later,

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: , and

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted:), led us to coarsen the histogram from 42 to 9 bins.

only the period from December 2002 to November 2015 which overlaps temporally with Aqua and Terra MODIS data. Since we pursue the co-variability of cloud and precipitation, and one of the essential pieces of cloud information is the optical thickness which is only available during daytime, our study relies on measurements only around the Terra and Aqua overpasses of 10:30 am and 1:30 pm local solar time (LST), respectively. We restrict our study to the extended tropical region (35°N − 35°S) to avoid ambiguities in the interpretation of the MODIS joint histograms which include progressively more temporal variability towards higher latitudes as data from successive spatially overlapping orbits fall within the same 1°×1° grid cell. Still, we should note that when various aspects of the analysis were tested on the full TMPA spatial coverage (50°N − 50°S), the results were not substantially different. Lastly, since it is well-established that precipitation properties over land and ocean are quite different (e.g., Williams and Stanfill 2002; Zipser et al. 2006; Matsui et al. 2016), we maintain via the MODIS land-water mask (Carroll et al., 2009) distinct land and ocean results throughout our analysis. At the 1°×1° resolution, a grid cell is marked as ocean when the water mask area is greater than 90%, while it is marked as land when the water mask area is smaller than 10%. For our extended tropics domain this definition assigns 71.1% of the grid cells to the ocean and 24.1% to the land category.

The quality of the TMPA product differs between land and ocean, mainly due to two factors: (1) Gauge adjustment which reduces systematic biases in land precipitation, and (2) Satellite retrieval algorithm differences which result in lower random errors over ocean (Liu, 2016; Sapiano and Arkin, 2009; Tian and Peters-Lidard, 2010). We assert that our findings about ocean-land differences are not much affected by these algorithm differences because, first, random errors should be suppressed due to large sample size, and second, our analysis is largely based on deviations from the mean state. Nevertheless, it is understood that TMPA overall performs less reliably in certain situations such as continental warm rains (Kidder and Vonder Haar, 1995; Kummerow et al., 2015).

2.2 Matching Precipitation Data to Cloud Grid

Because the 3B42 dataset has higher spatial resolution than the MODIS Level-3 cloud dataset, we resample it to the 1°×1° resolution of the MODIS dataset. Previous studies averaged precipitation rates to a single value representing grid mean (e.g., Lee et al. 2013; Rossow et al. 2013). In this study, a marginal histogram of 3B42 0.25°×0.25° grid precipitation rates is created for each 1°×1° grid cell. The idea of such 1°×1° precipitation histograms was drawn from our other main data set, the MODIS joint 2D histogram of p_c–τ, which preserves a certain degree of sub-grid cloud information (although not of the actual spatial distribution of the sub-grid variability). So, in a sense, sub-grid information about precipitation rate can also be preserved in the form of a histogram by assigning the 16 values (when there are no missing values) of precipitation rate at 0.25°×0.25° resolution to pre-defined bins to create a marginal histogram at 1°×1° grid cell. The histogram is normalized by dividing each bin count by the total count in the histogram bins, i.e. 16 in the default case of no missing value. 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. One of the 16 precipitation histogram bins corresponds to "no-rain" and the remaining 15 bins to rain rates greater than zero. Histogram bin boundaries are selected with fifteen logarithmically-spaced intervals to

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: algorithmic variations
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: conceivable
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: has
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: reliable data
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: The
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: is of
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: that of
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: data, and
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: therefore
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: the 3B42 data
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: exist) values
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted:
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: if
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: . If the number of bins in the histogram is chosen to be also 16
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: the sum of all histogram bins at 1° grid cell is equal to 1,
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: thus converted to

ensure a more even distribution of counts (see Fig. 1). Figure 1 shows the distribution of precipitation rate of the original TMPA data in our extended tropics domain according to this histogram binning approach. We see that the amount of missing data is negligible, and that the "no-rain" bin has an 89.5% share of all data points. The rain rate around 1 mm/hr has a maximum share near 1.1%, and extreme values are below 0.4% at both low and high rain rates.

In addition to the trivial matching of grid cells, the TMPA and MODIS observations also need to be matched in time. Since MODIS Level-3 cloud data come from the aggregation of retrieved satellite observation along the Terra or Aqua paths, and since these satellites are in a sun-synchronous orbit, each grid cell of a daily MODIS map has a limited range of nominal LST, but has a varying Coordinated Universal Time (UTC), the time-keeping system of TMPA. The UTC of each grid cell can be estimated from the mean solar zenith angle (SZA) available as a MODIS Level-3 variable, and the latitude and time information for each grid cell. Because of minimal overlap of satellite orbits in the tropics, the mean SZA value is a result of mostly (small) spatial variations within the 1°×1° grid cell. After identifying the UTC corresponding to the grid cell of cloud data, the proper TMPA data points can be extracted. Since the TMPA data is available at 3 hour-intervals, TMPA data centered, say, at 12 pm UTC, will be matched with MODIS data having UTC between 10:30 am and 1:30pm.

The histograms of TMPA tropical rainfall rate that matches Terra and Aqua paths spatially and temporally are also shown in Fig. 1. One notable change from the original TMPA data to Terra- or Aqua-matched data is that the portion of missing data now surges to over 5% of total data points. Most of these missing data are traced back to unavailable Level-3 MODIS data, for reasons such as absence of clouds or gaps between consecutive Terra-Aqua orbits at low latitudes. Other differences in occurrence frequencies between original and matched data are probably due to the diurnal cycle of precipitation. At the Terra overpass time of around 10:30 am (LST), precipitation is relatively weak over both land and ocean (e.g., Yang and Smith 2006; Kikuchi and Wang 2008). This appears in Fig. 1 as Terra-matched precipitation having smaller frequencies than the original and the Aqua-matched precipitation, although it is somewhat improper to directly compare Terra- or Aqua-matched data with fully sampled data because the higher ratio of available (non-missing) data in the fully sampled data propagates as higher relative frequency in the various precipitation bins. It is also notable that, for weak-to-moderate precipitation rate (less than 1mm/hr), even Aqua-matched precipitation is (slightly) lower in percentage terms than fully-sampled TMPA precipitation, which can be interpreted as weak-to-moderate precipitation being more frequent outside the time windows of Terra and Aqua overpasses.

2.3 Analysis Method and Simplification of Cloud and Precipitation Histograms

The simplest and most straightforward method to measure the co-variability of two variables is to calculate their cross-correlation coefficients, namely Pearson's r. In this study, the cloud fraction values in each bin of the p_c-τ joint histogram and the relative frequencies in the precipitation histogram form large arrays (O(1,000,000)) in the spatio-temporal domain, from which we can calculate correlation coefficients as time and location varies. The original resolution of the p_c-τ and precipitation histograms yields 672 (= 42 CF bins × 16 precipitation bins) correlation coefficients. Analysis and visualization

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: domain of the
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: time of the
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: needs
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted:
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: .
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: 3B42
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: the maximum time difference
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: cloud
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: precipitation observations is
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: .5 hours in this study
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: In explaining other
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: , the effect of
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: should be considered. For example, at
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: suppressed
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: .
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: -wise
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: of
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: the
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: coefficient between cloud and precipitation bin values
DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: This is simply too

of such a large number of coefficients are impractical, hence we pursue an analysis where both the cloud and precipitation histograms are coarsened.

Reducing the 42 bins of the cloud histogram allows us to make a more intuitive physical connection with the 9 standard ISCCP cloud types of Rossow and Schiffer (1999). 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. Figure 2 shows the p_c and τ range for each cloud type. Low and mid-level cloud types are composed of 4 CF bins (= 2 p_c classes × 2 τ classes) while high cloud types are composed of 6 CF bins (= 3 p_c classes × 2 τ classes). Hence, the CF value of each cloud type comes from the summation of either 4 or 6 CF bin values of the original 2D joint histogram.

Similarly, the 16 histogram bins of precipitation are reduced to 6 groups. The "no-rain" bin is unchanged, and the other 15 bins of measurable rainfall are resampled to 5 precipitation groups (each called as a "P-group" hereafter) by summing three consecutive precipitation bins, as shown at the bottom of Fig. 1. Each P-group is labelled from P1 to P5, with P1 representing the lightest precipitation, and P5 representing the heaviest precipitation. For simplicity, the same symbols are henceforth also used to represent the frequency of occurrence within these groups, since their meaning is always clear by the context.

Our histogram coarsening reduces the number of correlation coefficients to 54 (= 9 cloud type CF values × 6 P-group frequencies). Since the Terra and Aqua data (and matched precipitation data) are considered as a single ensemble, our results represent the local cloud-precipitation co-variability for the 6-hour daytime period spanning 1.5 hour before the Terra overpass to 1.5 hour after the Aqua overpass.

3 Land-Ocean Difference of cloud-precipitation relationships

3.1 Basic Statistics and Composite Means of Cloud and Precipitation Data

Before examining correlations between cloud and precipitation data, it is illuminating to examine the basic statistical information and mean states of both histograms from which correlations are extracted. First, we examine the P-groups that co-exist with certain cloud type fractions at the grid-level. Figures 3 and 4 show the conditional probability of P-group occurrence under the condition that a particular cloud type exists over ocean (Fig. 3) and land (Fig. 4). For example, for all oceanic 1°×1° grid cells with Cumulonimbus (*Cb*) clouds occurring, about 52% of the grid cells report P5 precipitation at one or more 0.25° sub-grid cell(s) (Fig. 3a, upper-right bin). The threshold CF that determines cloud occurrence is set to 6.25%, i.e. the same threshold fraction (1/16) that defines precipitation occurrence. We note that P-groups are not mutually exclusive because several P-groups can occur simultaneously in a 1°×1° grid cell.

Over ocean, the cloud type co-occurring the most with precipitation rates of medium to heavy intensity is, not surprisingly, *Cb*. The P-group most likely to occur alongside *Cb* clouds is P4 with a probability of 0.77 (Fig. 3b). The probability of P5 group occurrence is lower at 0.52, but also comes with an overall P5 population smaller than that of P4 (Fig. 1 and Table 1).

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: a

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: to analyze, visualize, and make sense of. Hence, a coarsening of

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: before correlations are calculated is in order per the procedure described below

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: First,

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: bins are reduced from 42 bins

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: defined in

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted:

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Henceforth, for

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: for

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: with no confusion resulting, as

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: is invariably clear. As a result,

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: decreases

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: We also note that

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: so

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: and

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: **relationship**

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: co-variability of anomalies is

When precipitation of any intensity is considered (Fig. 3f), besides *Cb* having the highest probability of precipitation, 0.90, oceanic Nimbostratus (*Ns*) also emerges with a high probability of 0.75. The no-rain occurrences are, not surprisingly, better associated with thin and/or low clouds (so-called "weak" clouds), topped by the 0.82 probability for Cumulus (*Cu*) clouds. It is notable that no-rain probabilities are clearly distinguishable from those of the weak P1 or P2 rain groups not only by the probability of these P-groups occurring (we note that the population of the no-rain case is much larger), but also by how the probability varies with cloud type within the precipitation group (e.g., compare *Cu* and *Ns* in Figs. 3e and 3g as an extreme contrast). Comparing Figs. 3 and 4, we see that land clouds generally have a smaller chance of precipitation co-existing with clouds at the 1° scale. Even the P4 precipitation probability of *Cb* clouds is only 0.54 (Fig. 4b), far lower than its oceanic counterpart of 0.77. For the case of rainfall with any intensity (Fig. 4f), the precipitation probability of Ns is only 0.35 compared to 0.75 over ocean. The precipitation probability of mid-level Altostratus (*As*) also decreases from 0.53 to 0.31, so mid-level clouds seem particularly less active precipitation producers over land. In addition, the lightest rain group P1 over land is not associated with any particular cloud type (Fig. 4e vs. Fig. 3e) while the no-rain case exhibits strong probability dependence on cloud type. The issue of less rain over land is also covered in the next composite plots (Figs. 5 and 6).

Figures 5 (ocean) and 6 (land) show composite mean cloud and precipitation histograms, for occurrences of the strongest precipitation groups P5 and P4 (i.e., at least one of the sub-grids within the 1°×1° grid cell has a precipitation rate belonging to the P5 or P4 class). When P5 occurs over ocean (Fig. 5), both cloud and rainy fractions exceed those of the P4 cases. On the cloud side, *Cb* exhibits the largest increases in CF when moving from the P4 to the P5 composite. For the P5 composite, the largest CFs (red color) are located in the bins with p_c below 310 hPa and the τ bins extending from 9.4 to 60, while in the P4 composite, CF peaks in the bin bounded by 310 and 180 hPa, and with τ between 3.6 to 23. Conversely, thin ($t < 3.6$) cloud CFs as well as stratocumulus (*Sc*) CF are smaller in the P5 composite than the P4 composite. However, it cannot be determined from this analysis alone whether the increased amounts of thin and *Sc* clouds in the P4 composite are directly linked with the occurrence of P4 precipitation, or are a consequence of increased chance of co-existence with other clouds producing P4 precipitation. The CFs of mid-level clouds increase only slightly from P5 to P4 composites in terms of absolute values, but these increases are quite large in a relative sense because absolute CF values for these clouds are very small in the MODIS climatology.

Consistent with the CF changes, the total rainy fraction, defined as the sum of the 15 precipitation histogram bin frequencies excluding the "No-rain" bin in 1°×1° grid cells, also increases in the P5 composite (0.794 vs. 0.627). The mean precipitation histogram in the P5 composite (Fig. 5 top right) exhibits a peak within the P5 group, but the fraction of total precipitation in the P4 group is larger. This does not come as a surprise because, first, the absolute population of P4 is higher than that of P5 (Fig. 1) and second, most P5 precipitation events co-occur with P4 precipitation events at 1°×1° resolution (Table 1). The P4 fractional contribution in the P5 composite is also larger than the P4 contribution in the P4 composite (Fig. 5 bottom right), while the light to moderate P-group (P1-P3) fractions are slightly larger in the P4 composite compared to the P5 composite. This indicates that stronger precipitation events also have a weaker tail towards lower rainfall rates. In terms of total rainy fraction, considering that approximately 38% of the P4 composite population overlaps with the P5 composite (Table 1), we

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: will be revisited later.

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: subgrids

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: of

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: , respectively

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: are larger compared to

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: which is

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: fraction

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: bins

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: shows that the

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: occurs

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: due to

see that the spatial extents of oceanic rain systems producing P4 but not P5 are often much smaller than systems producing P5.

We also examined the geographical distributions of P4 and P5 occurrence frequency (supplementary Fig. 1), and found that the distribution maps look very similar, with the regions significantly skewed towards one of P4 or P5 being very few. This result suggests that the P4 and P5 composites in Fig. 5 are related and likely respectively capture the developing and mature stage of mesoscale convective systems (MCSs). The review of Houze (2004) and Chapter 9 of Houze (2014) describe MCS as the combined system of a large region of stratiform precipitation paired with individual or clustered Cb clouds, yielding thus a variety of cloud and precipitation structures (Houze et al., 1990). The P5 composite patterns of cloud and precipitation shown in Fig. 5 are in accordance with such MCS characteristics, i.e. strong convective clouds and a broad spectrum of precipitation.

Figure 6 shows the same P4 and P5 composite means as Fig. 5, but over land. Comparing the top and bottom panels of Fig. 6, we see that the general characteristics of the differences between P4 and P5 land composites are similar to their oceanic counterparts. For example, the total CF and rainy fraction increase from the P4 to the P5 composites, accompanied by larger CFs of *Cb* clouds, and P4 group fractional contribution in the P5 composite. However, there are also notable differences, such as total CF difference between P4 and P5 composites being smaller over land than over ocean. Over land, smaller CFs can produce P5-magnitude precipitation while larger CFs are needed for P4-magnitude precipitation compared to ocean. The total rainy fractions of P4 and P5 composites are also smaller over land. For example, when P5 occurs, 79% of oceanic sub-grids at 1° scales are precipitating, while the same is true for only 59% of continental sub-grids. For P4 composites, the values are 63% over ocean vs. 47% over land. These are strong indications that continental systems producing heavy rainfall are in general smaller in size than their oceanic brethren (Liu et al. 2008; Houze 2014, Chapter 9; Houze et al. 2015).

The distributions of total rainy fraction as well as grid mean cloud properties by P-group are further examined in Fig. 7, which shows boxplots of total rainy fraction, CF, log10(τ), and p_c distributions. Over ocean, both total rainy fraction and CF generally increase monotonically with precipitation. However, the picture is somewhat different for land. From P2 to P5, both the median and mean values of land CF are quite similar (Fig. 7b). As a result, in the P2 case, the land CF median is nearly 10% greater than the ocean CF median, while it becomes 5% smaller than its ocean CF counterpart in the case of P5. At the same time, the total rainy fraction over land appears to be monotonically increasing in the same way as over ocean, albeit with a notably smaller absolute value and slope of growth. Hence, it appears that over land, similar amounts of CF (e.g., 70% to 80%) in a 1°×1° grid cell are involved in a broad range of precipitation rates, while the fraction of raining clouds in the grid cell is much smaller compared to ocean. Collectively, these results indicate reduced predictability of precipitation from knowledge of CF over land, at least with the precipitation dataset at hand.

Houze (2014, Chapter 9) and Houze et al. (2015) noted that shallow and isolated clouds producing "warm rain" are mostly oceanic phenomena, while the size of MCSs is generally larger over ocean than land. These two different precipitating sources can explain the big contrast over ocean of total rainy fraction between P2 (median 35%) and P5 (median 85%); over land the difference is less than 30%. In addition, Figs. 7c and 7d show that light-to-medium precipitation groups (P1-P3)

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: horizontal sizes

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted:), respectively.

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: cells

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: cells

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: linearly

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: is

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: linearly

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: .

over ocean are associated with optically thinner and shallower clouds, another evidence of the prevalence of marine "warm" rain processes. Note also the larger variability (taller inter-quartile box) of oceanic p_c for light precipitation compared to heavy precipitation, indicating that the former is harder to relate to particular cloud types. For continental light precipitation, associated clouds are optically thicker and of higher altitude than oceanic counterparts, but TMPA's potential weakness in identifying such precipitation, as described in section 2.1, may be affecting the land results of Fig. 7.

In summary, the 1°×1° spatio-temporally matched cloud and precipitation data suggest that prevailing features such as contrasting horizontal size of oceanic and land MCSs can be clearly detected by this study's methods. In the next subsection, the covariability of cloud and precipitation is examined in detail using explicit correlation analysis.

3.2 Correlations between Cloud and Precipitation Fractions

As stated previously, to measure the co-variability of cloud and precipitation, we calculate cross-correlation coefficients between the CFs of the 9 cloud types and the normalized frequencies (equivalent to fraction of precipitating area) within the 5 P-groups. Figures 8 and 9 show correlations of cloud types for each P-group as well as all combinations of consecutive cumulative P-group frequencies over the oceanic and land regions of our extended tropical domain. We note that, when the fraction sum of specific P-group(s) is zero, the data point is excluded from the calculation of correlations. Hence, for example, correlation coefficients with P5 over ocean (Fig. 8a) are calculated with approximately 4.5% of the total data available. Even over land, the sample size for this case (Fig. 9a) still exceeds one million, placing the 99% significance level to less than 0.005 in correlation coefficient absolute value. The statistical significance level was calculated here using a bootstrapping method which randomly shuffles the array, but in a way that considers the effect of autocorrelations between neighboring grid cells (i.e., shuffling by "chunks"; Kunsch 1989; Léger et al. 1992; Liu and Singh 1992). Consideration for the effect of neighboring grid cells is important because neighboring grid cells are usually *not* independent (e.g., a cloud system can occupy multiple grid cells); without this consideration, the degree of freedom will be overestimated, and thus the significance level underestimated. With the significance level quoted above, all correlations in Figures 8 and 9 are statistically significant.

Examining first oceanic cloud-precipitation coupling, Fig. 8 reveals that strong correlations, both negative and positive, occur in the panels on the left, whilst correlations weaken as one moves to the right. The leftmost column panels consist of P-group(s) that include P5, the group of heaviest precipitation, while as one moves to the right, heavier precipitation is progressively excluded. The overall picture then is that of strong correlations corresponding mostly to heavy precipitation and of light precipitation correlating poorly with all cloud types. The leftmost column panels of Fig. 8 indicates that positive coefficients occur for high cloud types of moderate to strong optical thickness, namely Cirrostratus (*Cs*; probably includes many anvils) and *Cb* (deep convection core), while negative values occur for low cloud types that are also optically thin. In the 5 panels of the leftmost column, *Cb* clouds always have strong positive correlations with precipitation, a result that comes as no surprise. For the correlation of *Cs* clouds to become positive and then increase, lighter precipitation has to be added to P5. For example, when only P5 values are used (Fig. 8a), the correlation coefficient of *Cs* clouds is negative (-0.16),

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: it

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: can

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: a factor

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Fig. 7

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: with our binning conventions

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: consecutively

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: nevertheless

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: 1

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Looking

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: at

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: correspond

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: while

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: correlates

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: any

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: type

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: should not come

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: a

but changes to 0.22 after P4 is added to P5 (Fig. 8b). This suggests that lighter rain in the vicinity of the heaviest rain is more closely related to *Cs* (or anvil) clouds. A similar trend of stronger correlations when lighter precipitation is added also ensues for low and thin *Cu* clouds, although with a negative sign in this case. In Fig. 8a, a strong negative correlation is seen with (high and thin) Cirrus (*Ci*) clouds, and as lighter precipitation is added, the peak of negative correlations moves towards lower *Cu* clouds.

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).

Notable patterns in correlation coefficients are also detected in the second left column panels which show correlations with P4 precipitation included, but without P5. Similar to the leftmost column panels, *Cb*, *Cs*, and *Cu* clouds show the stronger correlations with positive or negative signs. One difference from the P5 cases is that, in Figs. 8e, 8h, and 8l, the positive correlations of *Cs* clouds are stronger than those of the thicker *Cb* clouds. The correlation coefficient values of *Cs* clouds in these panels are quite similar to the values for the same clouds in the leftmost column (which includes P5 precipitation). This result suggests that it is actually *Cs* clouds that are related the most to the variability of P4 and lighter precipitation.

Moving now to the land regions of our extended tropical domain, we use the same "correlation pyramid" to note that the relationship between high and optically thick *Cb* clouds and P5 heavy precipitation is of positive strength similar to that over oceans (Fig. 9). However, other details are quite different between land and ocean. First, the negative correlations of *Cu* clouds in the leftmost column panels are weaker. In Fig. 8, the peak negative correlation values reached -0.40 and occurred in panels (d) and (g) which include the moderate to weak precipitation of the P3 and P2 groups. In Fig. 9 on the other hand, the peak negative value weakens to -0.23 and occurs in panel (b), which represents the sum of only P4 and P5 precipitation; the negative correlations weaken as lighter precipitation is added. This result suggests better chances of *Cu* clouds and P5 precipitation co-existing in $1^\circ \times 1^\circ$ grid cells over land compared to ocean. This observation may be related to our earlier finding inferred from Figs. 5-7 that the size of precipitating systems is much smaller over land than ocean.

Secondary but still noteworthy differences between land and ocean are identified in the correlations between *Cs* clouds and precipitation that includes the P4 and P3 groups (second and third column panels from left). Previously in Fig. 8, the maximum correlation values in the second-from-left column were the ones correlated with *Cs* clouds, up to 0.36. In the third column associated with P3 precipitation, correlations with *Cs* clouds weaken to 0.16. In contrast, Fig. 9 shows that the strongest correlations of the second column are those for *Cb* clouds, not *Cs*. In the third column, the correlations with *Cs* clouds do not weaken as much, with a 0.21 correlation value being reached in P-groups that include P3. This pattern indicates that continental high clouds are better correlated with lighter precipitation. It is also notable that correlation coefficients with *Cs* clouds in the first column of Fig. 9 (including P5 over land) reach just 0.25, while those in Fig. 8

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: included

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: not

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: is a weaker value of

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: is seen

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: drawn

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: , 6 and

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: raining

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: are

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: weakened

(ocean) are as high as 0.39. A possible explanation of the above correlation results is that thick anvils of continental MCS (Cetrone and Houze, 2009; Yuan et al., 2011) are more frequently classified as *Cb* rather than *Cs* (as defined in this study). For light precipitation, the absolute values of correlation coefficients are smaller than those for heavy precipitation commonly found over land and ocean, reflecting the fact that the mechanisms and origins of light precipitation exhibit a greater variety. Nevertheless, a meaningful difference between land and ocean can be seen in Figs. 8 and 9. When comparing Figs. 8n and 8o with 9n and 9o, peak correlations around 0.1 occur for Stratus (*St*) over ocean, but similar peak correlations over land occur even for *Cs* and *Cb*. This result suggests that over land even light precipitation is more frequently related to strong convective activity while oceanic light precipitation has a greater chance of being produced by "warm rain" mechanisms, as noted at the end of subsection 3.1.

In summary, continental *Cs* and *Cb* clouds co-exist with a broader range of precipitation, but are also more weakly correlated with them, compared to their oceanic brethren. This result is consistent with the previously noted climatological features of grid mean cloud properties shown in Fig. 7. For example, the median p_c for the P2 group over land in Fig. 7d was already below 440hPa, while for oceanic clouds the median p_c reached such values when precipitation was strong enough to belong to the P4 group. The optical thickness was also generally larger for land clouds (Fig. 7c). It is possible that TMPA is missing some "warm" rain events over land due to microwave retrieval inadequacies as stated in subsections 2.1 and 3.1. For heavy precipitation, Level-2 TRMM observations led Liu et al. (2008) to conclude that tropical land storms are more vertically developed, i.e. optically thicker clouds with higher tops, but also spatially more confined than oceanic storms (see also Houze et al. 2015; Matsui et al. 2016). Hence, precipitation over land occupies a smaller area, resulting in weaker correlations at scales of one-degree. Differences in correlations between Figs. 8 and 9 therefore reflect land/ocean differences in the nature of tropical storms or MCSs.

There are other intriguing aspects of cloud-precipitation co-variability in land and ocean, and these are examined more closely in the next subsection: (1) the origin of negative correlations and (2) correlation sensitivity to precipitation strength.

3.3 Further Investigation for Correlation Features

3.3.1 Negative correlations between precipitation and thin clouds

In Figs. 8 and 9, we saw thin clouds having negative correlations with heavy precipitation. These negative correlations can be interpreted as thin clouds being rarer when heavy precipitation occurs, an interpretation that is consistent with empirical observation and expectations. However, since it is also seen that heavy precipitation is strongly related to thick and high-level clouds (e.g., *Cb*), the negative correlation of optically thin clouds with heavy precipitation can also be interpreted as a contemporaneous negative co-occurrence relationship between optically thin low and optically thick high clouds. Please note that for a cloud type to be always (i.e., regardless of precipitation strength) anti-correlated with precipitation, its occurrence must be anti-correlated with that of other cloud types that are positively correlated with precipitation of a certain range. In order to examine these issues, we calculate internal correlations among cloud types based on the spatiotemporal variability of

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: the possibility

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: by the definition of cloud types

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: .

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: such values remain with

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: over land

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: is

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: reaches

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: is

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: is

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: horizontally smaller

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: is spread over

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: the level

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: areas

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: are

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: likely good indicators of

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: found that

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: have

their CFs. From all possible combinations, we elect to show results where one of the cloud types is either *Cs* or *Cb* when P4 or P5 precipitation group occurs. These results for both land and ocean are shown in Fig. 10. For example, Figs. 10a (ocean) and 10d (land) show correlation coefficients between the CF of *Cs* and the CFs of all other cloud types for grid cells reporting P4 precipitation. We note that the samples used for Fig. 10 correlations are the same as those used for cloud-precipitation correlations shown in Figs. 8 and 9, given the same precipitation conditions, namely P4 or P5 values greater than zero (i.e., Figs 10a, 10b, and 8c).

Figures 10a, 10b, and 10c show correlation coefficients based on oceanic *Cs* and *Cb* CFs. The *Cs* clouds are strongly anti-correlated with *Cu* and *Sc* clouds, while *Cb* clouds are furthermore also strongly anti-correlated with *Ci* clouds. In the cases of P5 precipitation presence (Fig. 10c), the anti-correlation between *Cb* and *Ci* CFs becomes even stronger. Actually, in this case, *Cb* clouds are anti-correlated with all other cloud types; i.e., when *Cb* CF increases at 1°×1° grid cell, CFs of other clouds decreases, and vice versa. These cloud type correlation patterns remind us of Figs. 8a, 8b, and 8c. For example, a comparison between Figs. 10c and 8a shows that the anti-correlation ordering by strength is the same, with *Ci* clouds coming first, *Cu* second, and *Sc* clouds third. This finding suggests that in tropical oceans P5 precipitation is mainly related to *Cb* clouds, and its anti-correlation with thin clouds is another expression of the anti-correlation between *Cb* and thin clouds. The exact nature of the anti-correlation are unknown because a passive sensor such as MODIS has limited skill in distinguishing between cases where the mid- and low-level clouds are absent and cases where they are obscured by high clouds.

When focus shifts to the weaker P4 precipitation class, both *Cs* and *Cb* clouds anti-correlate with low *Cu* and *Sc* clouds, and the anti-correlation is only slightly weaker for *Sc* than *Cu* (Figs. 10a and 10b). Previously however, Fig. 8c indicated that the anti-correlation between P4-class and *Sc* cloud is much weaker than that between P4 and *Cu* cloud (-0.15 vs. -0.28). This discrepancy is also seen in all panels of Fig. 8 representing correlations with moderate to heavy precipitation classes (third column from left), but is not seen over land (Fig. 9). While this issue will be discussed further in the next subsection which deals with correlation sensitivity, suffice it to say here that cloud-precipitation anti-correlations cannot be exclusively attributed to cloud type co-occurrence anti-correlations.

When comparing oceanic and continental correlation patterns in Fig. 10 (top row vs. bottom row), we see the correlation patterns being quite similar, but with weaker correlation magnitudes over land. For example, *Cs* clouds in Fig. 10d remain strongly anti-correlated with *Cu* and *Sc* clouds, and *Cb* clouds in Fig. 10f are still anti-correlated with all other cloud types. Yet, differences between ocean and land clouds also emerge. First, particularly in the presence of non-zero P4 precipitation (Figs. 10d and e), there are stronger anti-correlations between *Cb* or *Cs* clouds and mid-level clouds over land. Previously in Fig. 6, we noted that mid-level clouds have greater CFs over land compared to ocean (even though their absolute value is much smaller than high clouds). The increased CFs of mid-level clouds over land may be related to a closer relationship with high-thick clouds, thus affecting the correlation strength.

Another difference between ocean and land is the correlation between *Cb* and *Cs* in the presence of P5-class precipitation. Comparing Figs. 10c and 10f, the notable anti-correlation value of -0.27 over ocean weakens to -0.16 over land. This result indicates that *Cb* and *Cs* clouds are less mutually exclusive over land. Since the overcast condition (100% CF) in a 1°×1°

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: samples…s those used for cl ... [1]

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: when…f P5 precipitation is ... [2]

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: is shifted…hifts to the weak ... [3]

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Comparing…hen comparing ... [4]

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Cs…*b* and Cb ... [5]

grid cell is more frequent over ocean (Fig. 7b), indicating that oceanic MCS can grow to sizes larger than 1°, there is a greater chance of competition between *Cb* and *Cs* over ocean to fill the grid cell.

Lastly, we return to our previous point that the anti-correlation of CFs among cloud types does not explain all features of the anti-correlation between cloud and precipitation shown in Figs. 8 and 9. For example, comparing Figs. 8a and 9a, anti-correlation between P5 and *Cu* cloud weakens from -0.25 (ocean) to -0.20 (land). However, Figs. 10c and 10f indicate that the anti-correlations between *Cb* and *Cu* clouds are almost the same for ocean and land (-0.37 vs. -0.36). This further supports the hypothesis that the frequencies of the P5 precipitation group and *Cb* CFs are more weakly coupled over land.

3.3.2 Correlation sensitivity to heavy precipitation

Correlations between cloud and precipitation shown previously in Figs. 8 and 9 indicated that the heaviest precipitation group has a solid relationship (correlation or anti-correlation) with cloud types, while weaker precipitation groups do not. This fundamental finding is examined more closely with more detailed CF-precipitation correlations. Figure 11 shows correlation coefficients over both ocean and land between the CF of various cloud types and the frequency of cumulative precipitation within original precipitation bins, from the 7th bin onward (i.e., 0.251 mm/hr and above). Hence, at the start of the x-axis the precipitation frequency corresponds only to the 7th bin, and as one moves along the axis precipitation frequencies for subsequent bins are progressively added until the end of the axis where the precipitation frequency represents the sum of all values from the 7th to 15th bin, namely the sum of frequencies of the P3, P4 and P5 groups. When compared to Figs. 8 or 9, Fig.11 shows essentially in more detail the evolution of correlation coefficients for the third row of the "pyramid", i.e. correlation changes as one moves from Fig. 8f (9f) to 8e (9e) and then to 8d (9d) over ocean (land).

Figure 11a shows the correlation change of high clouds (*Ci*, *Cs*, and *Cb*). Over ocean (solid line), the correlation of *Cb* cloud increases monotonically as heavy precipitation is added, while that of *Cs* cloud peaks when the 13th bin (2.51−3.98 mm/hr) is added; further additions of heavier precipitation results in correlation coefficients trending downward. Similar patterns are also seen for the land clouds in this category. However, one prominent difference between ocean and land is that the land clouds in this group tend to be more strongly correlated with weaker precipitation. For example, continental *Cb* clouds correlate better than oceanic *Cb* to precipitation up to the 13th bin. However, the correlation curve for oceanic *Cb* clouds exhibits a steeper slope after adding the 11th bin, and ends up surpassing continental *Cb* clouds with the heaviest precipitation. In the case of *Cs* cloud, the continental correlation curve peaks with the addition of the 11th bin (1.0−1.58 mm/hr), while the oceanic peaks upon addition of the 13th bin. This result indicates that P4 precipitation over land tends to be more related with *Cb* than *Cs* clouds, contrasting what happens over ocean. In the case of *Ci* clouds, the anti-correlation is stronger at weak precipitation over land, consistent with the above argument, but the difference between land and ocean is not very pronounced given the small absolute values of coefficients compared to *Cs* and *Cb* clouds.

For the mid-height cloud group shown in Fig.11b, a notable difference between ocean and land is seen for the *As* and *Ns* clouds. Oceanic *Ns* clouds have broad positive correlations around 0.1 for all precipitation bins. Oceanic *As* also have positive correlations with moderate-to-heavy precipitation bins, although they decrease to zero as heaviest precipitation is

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Cs

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Cb

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: /

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: over

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: But

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: has been added

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: when

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: is added

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: which contrasts

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: the value of

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: the

added. On the other hand, continental *As* and *Ns* clouds show only negative correlations with all precipitation strengths. *As* and *Ns* occurrences are smaller over ocean (3.8%, 1.7%) than over land (5.3%, 2.5%) in Figs. 5 and 6, P4>0), but shallower convection over ocean seems sufficiently strong to produce moderate-to-heavy precipitation from *As* and *Ns* clouds.

In the case of the low cloud group shown in Fig.11c, first, the thickest *St* cloud's correlation evolution pattern looks similar to that of *As* cloud above, although the presence of *St* cloud over ocean is even smaller than *As* (*St* CF=1.3% vs. *As* CF=3.8% when P4>0 in Fig. 5). The correlation pyramid of Fig. 8 has shown that the positive correlation of *St* cloud is stronger when it is related to weak precipitation classes (P1 or P2) which are not included here (but are included in Supplementary Fig. 5). Secondly, also notable is the contrasting correlation evolutions of oceanic *Cu* and *Sc* clouds, previously mentioned to have different magnitudes of anti-correlation. Oceanic *Sc* clouds have slightly positive correlations with the 7th and 7th-to-8th precipitation bins which then become negative as heavier precipitation is added. Similar to the *St* cloud, the positive correlation of *Sc* cloud is expected to strengthen with even lighter precipitation (Fig. 8 and Supplementary Fig. 5). For the oceanic shallow convection, our results of low and mid-level cloud correlations consistently indicate that shallower and thinner clouds (e.g., *Sc*) relate better to lighter precipitation, while higher and thicker clouds (e.g., *Ns*) relate better to heavier precipitation. In the case of *Cu*, the correlation coefficient is roughly the same between ocean and land for the 7th precipitation bin, but the correlation curves diverge as heavier precipitation is added. By the time the frequencies of all precipitation bins from 7th to 15th have been added, oceanic *Cu* clouds have twice as strong anti-correlation compared to their continental counterparts. As discussed previously in the context of Fig. 10, correlations among cloud fraction co-occurrence, i.e. [*Cu* vs. *Cs*] or [*Cu* vs. *Cb*], are not as different between ocean and land as those shown here between clouds and precipitation. The weaker anti-correlation of continental *Cu* cloud with rainfall reflects then, at least partly, the less robust relationship between heavy precipitation and continental high clouds.

3.4 Limiting factors and uncertainties

3.4.1 Uncertainty of cloud type classification

In this study, MODIS-observed clouds are classified into 9 cloud types adopted from previous ISCCP conventions (Chen et al., 2000; Rossow and Schiffer, 1999) for the sake of convenience. This classification is, strictly speaking, based on arbitrary τ and p_c thresholds, and clouds assigned to each pair of bin boundaries will only loosely represent cloud types originally defined from morphological features seen by surface observers. Previously we noted that continental MCSs often include thick anvils (Cetrone and Houze, 2009; Yuan et al., 2011), but we can not confirm that these anvils are classified as *Cs* or *Cb* without knowledge of the cloud vertical extinction profile. Moreover, a passive sensor like MODIS has intrinsic limitations in identifying certain cloud types. Recent studies examining the nature of MODIS Cloud Regimes with active sensor observations from CloudSat and the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) show that similar MODIS joint histograms can have a variety of cloud vertical structures (Oreopoulos et al. 2017). In addition, Wang et al. (2016) showed that defining cloud types from CloudSat-CALIPSO observations where cloud vertical extent is

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: ([3.8%, 1.7%] ocean vs. [

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: %] land

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: see

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: 2

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: 2

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: correspond

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: correspond

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: identify whether those

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: 2017, submitted to *J. Geophys.*

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Moved down [1]: Res.

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: *Atmos.*).

better known can yield large disagreements with cloud type definitions from the MODIS p_c-t joint histogram. Such ambiguous definitions of cloud types from passive measurements may be the reason for substantial correlations between *As* and certain ranges of precipitation even though *As* is usually thought of as a non-precipitating cloud type. In summary, the 9 cloud types in this study may not strictly correspond to their traditional, ground-based classification, so their relationship with precipitation should not be taken literally or juxtaposed with empirical knowledge. They are simply a convenient framework to organize findings about cloud-precipitation co-variability at 1° scales.

Furthermore, the passive MODIS observations suffer from low skill in detecting multi-layer clouds. Specifically, MODIS generally only detects the cloud top of the highest cloud, so high clouds such as cirrus or stratiform anvil will mask the presence of shallow clouds. This may be a contributing factor to the negative correlations by *Cu* and *Cs* in Fig. 10. Unfortunately, this is a shortcoming of passive cloud observations that we have to live with in exchange for wider coverage.

3.4.2 Uncertainty of TMPA and its temporal matching to MODIS

As noted in subsection 2.1, TMPA quality varies by location. Over land, the strong surface emissivity forces microwave retrievals of precipitation to rely on the ice scattering signature, which may not be present for warm (or shallow) rain. While there are gauge adjustments over land, they depend on the quality and density of the gauges used and operate at monthly time scales—thus may not be able to correct the precipitation rates for individual rain events. Over ocean, gauge adjustment is unavailable, leading to potential systematic errors in the precipitation estimates. Furthermore, the retrieval of remotely sensed precipitation relies on algorithms that estimate surface precipitation rates from passive microwave brightness temperature, a task that remains challenging. In addition, due to the intermittency of passive microwave sensors on low-Earth orbit satellites, gaps in the microwave field are filled in by infrared-based precipitation estimates, which have poor accuracy as infrared brightness temperature in isolation is only indirectly related to precipitation (it is as if we were trying to correlate precipitation here with one-dimensional p_c histograms). Hence, TMPA estimates possess considerable uncertainties.

Furthermore, precipitating systems can develop quickly, especially over land. For example, a tropical squall line can develop in a few hours, so it is possible that MODIS and TMPA observe different stage of a system given that a 1.5 hour difference is possible in spite of our temporally matching. This situation can result in decreased correlation coefficients between high-thick cloud and heavy precipitation over land. We are somewhat less concerned about this sampling issue because the lead/lag time between MODIS and TMPA is expected to be random, and therefore hopefully not a source of systematic bias. In the future, this concern can be ameliorated by using a higher temporal resolution precipitation dataset such as the Integrated Multi-satellitE Retrievals for GPM (IMERG; Huffman et al., 2015) instead of TMPA.

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: not

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: covariability

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: multilayer

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: contend

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: a challenge

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted:

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: despite

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: will

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: in place

4 Summary and Conclusion

The total amount, intensity, and frequency of precipitation should be organically related to the properties of the clouds from which it originates. However, due to different radiative signal strengths of hydrometeors at particular parts of the electromagnetic spectrum, precipitation and cloud observations are significantly decoupled, necessitating joint analysis of products developed for different purposes and from imperfectly matched observations. 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.

In order to advance the problem of understanding cloud-precipitation co-variability, we use contemporaneous multi-year datasets, widely-accepted concepts about how to classify clouds into various types from passive observations, and a combination of compositing and correlation analysis. We try to preserve some sub-grid variability information at one-degree scales by employing precipitation histograms built from the TMPA dataset, as well as MODIS joint histograms of cloud top pressure and cloud optical thickness, both of which are matched spatiotemporally.

We find, not surprisingly, that correlations between deep convective clouds and heavy rainfall are strong and stand out clearly, dwarfing all other correlation combinations for both land and ocean. Land-ocean differences are also remarkable. For example, oceanic deep convection systems (e.g., mesoscale convective systems) are more likely to attain overcast conditions and to have larger fractions of rainy sub-grids within 1°×1° grid cells, both indicative of larger horizontal size than their continental counterparts, consistent with previous studies. Over land on the other hand, *Cb* and *Cs* clouds are related not only with heavy precipitation, but rather with a broader range of rainfall which translates to weaker correlations.

Thin clouds, particularly *Cu* clouds (as defined here) are anti-correlated with moderate-to-heavy precipitation. The anti-correlation is stronger over ocean, and the magnitude is comparable to the anti-correlation between *Cu* and high-thick clouds (*Cb* or *Cs*). The fact that oceanic deep convection often fills and outgrows the 1°×1° reference grid cell, is ultimately the cause of clearer relationships (less uncertainty) among heavy precipitation, high-thick clouds, and low-thin clouds.

Over ocean, low-to-mid level clouds also exhibit positive correlations with precipitation of certain ranges, which represents shallow convection and warm rain processes. Among those clouds, the relatively higher and thicker *Ns* clouds relate better to mid to heavy precipitation, while lower and thinner *Sc* clouds relate better to light precipitation. In the end, positive correlations indicate that oceanic precipitation comes from a variety of cloud types and rain formation processes (warm rain) while most precipitation over land requires the presence of high clouds. Notably, the shallow continental clouds show better anti-correlations with heavy precipitation rather than positive correlations with light precipitation. It is conceivable that this result can change once detection of low clouds in the presence of high clouds and of warm rain over land improves (Field and Heymsfield, 2015; Mülmenstädt et al., 2015).

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: we

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: aspire

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: what a cloud type is

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: cells

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: .

Collectively, we make a strong case that rainfall predictability is better over oceans than continents when cloud information is available. But even over oceans, there are significant uncertainties in linking certain ranges of precipitation with specific cloud types, at least with our approach. 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. We argue that the insensitivity of cloud-precipitation relationships to location (supplementary Fig. 4) and precipitation dataset (initial tests with recent GPM-IMERG data that may be presented in a future study yielded similar results) strengthen the validity of this conclusion.

We expect that our study has the potential to form the basis of enhanced evaluation of precipitation in GCMs. A regime-based analysis in the deep tropics by Tan et al. (2017) suggests that clouds and precipitation are more decoupled in models than in observations (see also Jing et al., 2017; Suzuki et al., 2015). Confirming that conclusion with the approach introduced in this study is a possible next endeavour. In addition, more effort should be extended to apply the framework in this study to various case studies with more appropriate datasets (e.g., using higher resolution precipitation dataset for regional/seasonal studies, or longer period dataset for climate studies) in order to increase further our degree of confidence about the cloud-rainfall relationships.

Acknowledgements

We acknowledge funding from the following NASA programs: "The Science of Terra and Aqua"; "Modeling Analysis and Prediction (MAP)". JT was supported by an appointment to the NASA Postdoctoral Program at Goddard Space Flight Center, administrated by USRA through a contract with NASA (NNH15CO48B). We thank our NASA colleague George Huffman for helpful discussions. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center.

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: force

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: resort to

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: , sacrificing perhaps

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: (but suffered from reduced coverage). However

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: allow us to

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: can

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: ; submitted to *Clim. Dyn.*)

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: .

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: identify

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: optimal

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: that contain sufficient degree of detail to go along with extensive coverage

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: characterize with more

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: *is*

References

Carroll, M. L., Townshend, J. R., DiMiceli, C. M., Noojipady, P. and Sohlberg, R. A.: A new global raster water mask at 250 m resolution, Int. J. Digit. Earth, 2(4), 291–308, doi:10.1080/17538940902951401, 2009.

Cetrone, J. and Houze, R. A.: Anvil clouds of tropical mesoscale convective systems in monsoon regions, Q. J. R. Meteorol. Soc., 135(639), 305–317, doi:10.1002/qj.389, 2009.

Chen, T., Rossow, W. B. and Zhang, Y.: Radiative Effects of Cloud-Type Variations, J. Clim., 13(1), 264–286, doi:10.1175/1520-0442(2000)013<0264:REOCTV>2.0.CO;2, 2000.

Cho, H.-M., Zhang, Z., Meyer, K., Lebsock, M., Platnick, S., Ackerman, A. S., Di Girolamo, L., C.-Labonnote, L., Cornet, C., Riedi, J. and Holz, R. E.: Frequency and causes of failed MODIS cloud property retrievals for liquid phase clouds over global oceans, J. Geophys. Res. Atmospheres, 120(9), 4132–4154, doi:10.1002/2015JD023161, 2015.

Field, P. R. and Heymsfield, A. J.: Importance of snow to global precipitation, Geophys. Res. Lett., 42(21), 9512–9520, doi:10.1002/2015GL065497, 2015.

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Moved (insertion) [1]

Gianotti, R. L., Zhang, D. and Eltahir, E. A. B.: Assessment of the Regional Climate Model Version 3 over the Maritime Continent Using Different Cumulus Parameterization and Land Surface Schemes, J. Clim., 25(2), 638–656, doi:10.1175/JCLI-D-11-00025.1, 2012.

Houze, R. A.: Mesoscale convective systems, Rev. Geophys., 42(4), RG4003, doi:10.1029/2004RG000150, 2004.

Houze, R. A.: Cloud dynamics, 2. ed., Elsevier/Acad. Press, Amsterdam., 2014.

Houze, R. A., Smull, B. F. and Dodge, P.: Mesoscale Organization of Springtime Rainstorms in Oklahoma, Mon. Weather Rev., 118(3), 613–654, doi:10.1175/1520-0493(1990)118<0613:MOOSRI>2.0.CO;2, 1990.

Houze, R. A., Rasmussen, K. L., Zuluaga, M. D. and Brodzik, S. R.: The variable nature of convection in the tropics and subtropics: A legacy of 16 years of the Tropical Rainfall Measuring Mission satellite, Rev. Geophys., 53(3), 994–1021, doi:10.1002/2015RG000488, 2015.

Huffman, G. J. and Bolvin, D. T.: TRMM and Other Data Precipitation Data Set Documentation, in NASA Global Change Master Directory Doc., p. 44, Mesoscale Atmospheric Processes Laboratory, NASA GSFC. [online] Available from: https://pmm.nasa.gov/sites/default/files/document_files/3B42_3B43_doc_V7.pdf, 2015.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P. and Stocker, E. F.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales, J. Hydrometeorol., 8(1), 38–55, doi:10.1175/JHM560.1, 2007.

Huffman, G. J., Adler, R. F., Bolvin, D. T. and Nelkin, E. J.: The TRMM Multi-Satellite Precipitation Analysis (TMPA), in Satellite Rainfall Applications for Surface Hydrology, edited by M. Gebremichael and F. Hossain, pp. 3–22, Springer Netherlands, Dordrecht. [online] Available from: http://www.springerlink.com/index/10.1007/978-90-481-2915-7_1 (Accessed 30 June 2016), 2010.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R. and Xie, P.: NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG), [online] Available from: http://disc.sci.gsfc.nasa.gov/precipitation/documentation/gpm, 2015.

Jakob, C. and Tselioudis, G.: Objective identification of cloud regimes in the Tropical Western Pacific, Geophys. Res. Lett., 30(21), 2082, doi:10.1029/2003GL018367, 2003.

Jing, X., Suzuki, K., Guo, H., Goto, D., Ogura, T., Koshiro, T. and Mülmenstädt, J.: A Multimodel Study on Warm Precipitation Biases in Global Models Compared to Satellite Observations, J. Geophys. Res. Atmospheres, 122(21), 11806–11824, doi:10.1002/2017JD027310, 2017.

Jiang, X., Waliser, D. E., Xavier, P. K., Petch, J., Klingaman, N. P., Woolnough, S. J., Guan, B., Bellon, G., Crueger, T., DeMott, C., Hannay, C., Lin, H., Hu, W., Kim, D., Lappen, C.-L., Lu, M.-M., Ma, H.-Y., Miyakawa, T., Ridout, J. A., Schubert, S. D., Scinocca, J., Seo, K.-H., Shindo, E., Song, X., Stan, C., Tseng, W.-L., Wang, W., Wu, T., Wu, X., Wyser, K., Zhang, G. J. and Zhu, H.: Vertical structure and physical processes of the Madden-Julian oscillation: Exploring key model physics in climate simulations, J. Geophys. Res. Atmospheres, 120(10), 4718–4748, doi:10.1002/2014JD022375, 2015.

Kidder, S. Q. and Vonder Haar, T. H.: Satellite Meteorology: An Introduction, Academic Press. [online] Available from: https://books.google.com/books?id=ZUlbHlmc3_YC, 1995.

Kikuchi, K. and Wang, B.: Diurnal Precipitation Regimes in the Global Tropics, J. Clim., 21(11), 2680–2696, doi:10.1175/2007JCLI2051.1, 2008.

King, M. D., Menzel, W. P., Kaufman, Y. J., Tanre, D., Bo-Cai Gao, Platnick, S., Ackerman, S. A., Remer, L. A., Pincus, R. and Hubanks, P. A.: Cloud and aerosol properties, precipitable water, and profiles of temperature and water vapor from MODIS, IEEE Trans. Geosci. Remote Sens., 41(2), 442–458, doi:10.1109/TGRS.2002.808226, 2003.

Kooperman, G. J., Pritchard, M. S., Burt, M. A., Branson, M. D. and Randall, D. A.: Robust effects of cloud superparameterization on simulated daily rainfall intensity statistics across multiple versions of the Community Earth System Model, J. Adv. Model. Earth Syst., 8(1), 140–165, doi:10.1002/2015MS000574, 2016.

Kummerow, C. D., Randel, D. L., Kulie, M., Wang, N.-Y., Ferraro, R., Joseph Munchak, S. and Petkovic, V.: The Evolution of the Goddard Profiling Algorithm to a Fully Parametric Scheme, J. Atmospheric Ocean. Technol., 32(12), 2265–2280, doi:10.1175/JTECH-D-15-0039.1, 2015.

Kunsch, H. R.: The Jackknife and the Bootstrap for General Stationary Observations, Ann. Stat., 17(3), 1217–1241, 1989.

Lee, D., Oreopoulos, L., Huffman, G. J., Rossow, W. B. and Kang, I.-S.: The Precipitation Characteristics of ISCCP Tropical Weather States, J. Clim., 26(3), 772–788, doi:10.1175/JCLI-D-11-00718.1, 2013.

Léger, C., Politis, D. N. and Romano, oseph P.: Bootstrap Technology and Applications, Technometrics, 34(4), 378–398, doi:10.1080/00401706.1992.10484950, 1992.

Liu, C., Zipser, E. J., Cecil, D. J., Nesbitt, S. W. and Sherwood, S.: A Cloud and Precipitation Feature Database from Nine Years of TRMM Observations, J. Appl. Meteorol. Climatol., 47(10), 2712–2728, doi:10.1175/2008JAMC1890.1, 2008.

Liu, R. Y. and Singh, K.: Moving Blocks Jackknife and Bootstrap Capture Weak Dependence, in Exploring the Limits of Bootstrap, pp. 225–248, Wiley, New York., 1992.

Liu, Z.: Comparison of Integrated Multisatellite Retrievals for GPM (IMERG) and TRMM Multisatellite Precipitation Analysis (TMPA) Monthly Precipitation Products: Initial Results, J. Hydrometeorol., 17(3), 777–790, doi:10.1175/JHM-D-15-0068.1, 2016.

Mülmenstädt, J., Sourdeval, O., Delanoë, J. and Quaas, J.: Frequency of occurrence of rain from liquid-, mixed-, and ice-phase clouds derived from A-Train satellite retrievals, Geophys. Res. Lett., 42(15), 6502–6509, doi:10.1002/2015GL064604, 2015.

Matsui, T., Chern, J.-D., Tao, W.-K., Lang, S., Satoh, M., Hashino, T. and Kubota, T.: On the Land–Ocean Contrast of Tropical Convection and Microphysics Statistics Derived from TRMM Satellite Signals and Global Storm-Resolving Models, J. Hydrometeorol., 17(5), 1425–1445, doi:10.1175/JHM-D-15-0111.1, 2016.

Oreopoulos, L. and Rossow, W. B.: The cloud radiative effects of International Satellite Cloud Climatology Project weather states, J. Geophys. Res., 116(D12), D12202, doi:10.1029/2010JD015472, 2011.

Oreopoulos, L., Cho, N., Lee, D., Kato, S. and Huffman, G. J.: An examination of the nature of global MODIS cloud regimes, J. Geophys. Res. Atmospheres, 119(13), 8362–8383, doi:10.1002/2013JD021409, 2014.

Oreopoulos, L., Cho, N., Lee, D. and Kato, S.: Radiative effects of global MODIS cloud regimes, J. Geophys. Res. Atmospheres, 121(5), 2299–2317, doi:10.1002/2015JD024502, 2016.

Oreopoulos, L., Cho, N. and Lee, D.: New Insights about Cloud Vertical Structure from CloudSat and CALIPSO, J. Geophys. Res.-Atmospheres, 2017.

Pincus, R., Platnick, S., Ackerman, S. A., Hemler, R. S. and Hofmann, R. J. P.: Reconciling Simulated and Observed Views of Clouds: MODIS, ISCCP, and the Limits of Instrument Simulators, J. Clim., 25(13), 4699–4720, doi:10.1175/JCLI-D-11-00267.1, 2012.

Platnick, S., King, M. D., Ackerman, S. A., Menzel, W. P., Baum, B. A., Riedi, J. C. and Frey, R. A.: The MODIS cloud products: algorithms and examples from terra, IEEE Trans. Geosci. Remote Sens., 41(2), 459–473, doi:10.1109/TGRS.2002.808301, 2003.

Platnick, S., Meyer, K. G., King, M. D., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Holz, R. E., Yang, P., Ridgway, W. L. and Riedi, J.: The MODIS Cloud Optical and Microphysical Products: Collection 6 Updates and Examples From Terra and Aqua, IEEE Trans. Geosci. Remote Sens., 55(1), 502–525, doi:10.1109/TGRS.2016.2610522, 2017.

Richards, F. and Arkin, P.: On the Relationship between Satellite-Observed Cloud Cover and Precipitation, Mon. Weather Rev., 109(5), 1081–1093, doi:10.1175/1520-0493(1981)109<1081:OTRBSO>2.0.CO;2, 1981.

Rossow, W. B. and Schiffer, R. A.: ISCCP Cloud Data Products, Bull. Am. Meteorol. Soc., 72(1), 2–20, doi:10.1175/1520-0477(1991)072<0002:ICDP>2.0.CO;2, 1991.

Rossow, W. B. and Schiffer, R. A.: Advances in Understanding Clouds from ISCCP, Bull. Am. Meteorol. Soc., 80(11), 2261–2287, doi:10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2, 1999.

Rossow, W. B., Tselioudis, G., Polak, A. and Jakob, C.: Tropical climate described as a distribution of weather states indicated by distinct mesoscale cloud property mixtures, Geophys. Res. Lett., 32(21), L21812, doi:10.1029/2005GL024584, 2005.

Rossow, W. B., Mekonnen, A., Pearl, C. and Goncalves, W.: Tropical Precipitation Extremes, J. Clim., 26(4), 1457–1466, doi:10.1175/JCLI-D-11-00725.1, 2013.

Sapiano, M. R. P. and Arkin, P. A.: An Intercomparison and Validation of High-Resolution Satellite Precipitation Estimates with 3-Hourly Gauge Data, J. Hydrometeorol., 10(1), 149–166, doi:10.1175/2008JHM1052.1, 2009.

Stephens, G. L., L'Ecuyer, T., Forbes, R., Gettelmen, A., Golaz, J.-C., Bodas-Salcedo, A., Suzuki, K., Gabriel, P. and Haynes, J.: Dreary state of precipitation in global models, J. Geophys. Res. Atmospheres, 115(D24), D24211, doi:10.1029/2010JD014532, 2010.

Suzuki, K., Stephens, G., Bodas-Salcedo, A., Wang, M., Golaz, J.-C., Yokohata, T. and Koshiro, T.: Evaluation of the Warm Rain Formation Process in Global Models with Satellite Observations, J. Atmospheric Sci., 72(10), 3996–4014, doi:10.1175/JAS-D-14-0265.1, 2015.

Tan, J., Oreopoulos, L., Jakob, C. and Jin, D.: Evaluating rainfall errors in global climate models through cloud regimes, Clim. Dyn., doi:10.1007/s00382-017-3806-7, 2017.

Tian, Y. and Peters-Lidard, C. D.: A global map of uncertainties in satellite-based precipitation measurements, Geophys. Res. Lett., 37(24), L24407, doi:10.1029/2010GL046008, 2010.

Tselioudis, G., Rossow, W., Zhang, Y. and Konsta, D.: Global Weather States and Their Properties from Passive and Active Satellite Cloud Retrievals, J. Clim., 26(19), 7734–7746, doi:10.1175/JCLI-D-13-00024.1, 2013.

Wang, T., Fetzer, E. J., Wong, S., Kahn, B. H. and Yue, Q.: Validation of MODIS cloud mask and multilayer flag using CloudSat-CALIPSO cloud profiles and a cross-reference of their cloud classifications, J. Geophys. Res. Atmospheres, 121(19), 11,620-11,635, doi:10.1002/2016JD025239, 2016.

Williams, E. and Stanfill, S.: The physical origin of the land–ocean contrast in lightning activity, Comptes Rendus Phys., 3(10), 1277–1292, doi:10.1016/S1631-0705(02)01407-X, 2002.

Yang, S. and Smith, E. A.: Mechanisms for Diurnal Variability of Global Tropical Rainfall Observed from TRMM, J. Clim., 19(20), 5190–5226, doi:10.1175/JCLI3883.1, 2006.

Yuan, J., Houze, R. A. and Heymsfield, A. J.: Vertical Structures of Anvil Clouds of Tropical Mesoscale Convective Systems Observed by *CloudSat*, J. Atmospheric Sci., 68(8), 1653–1674, doi:10.1175/2011JAS3687.1, 2011.

Zhang, Z. and Platnick, S.: An assessment of differences between cloud effective particle radius retrievals for marine water clouds from three MODIS spectral bands, J. Geophys. Res., 116(D20), doi:10.1029/2011JD016216, 2011.

Zipser, E. J., Liu, C., Cecil, D. J., Nesbitt, S. W. and Yorty, D. P.: Where are the most intense thundersotorms on Earth?, Bull. Am. Meteorol. Soc., 87(8), 1057–1071, doi:10.1175/BAMS-87-8-1057, 2006.

Tables

Table 1: Population percentages of grid cells with specific precipitation characteristics over ocean and land from 13 years of data in our 35°S-35°N extended tropics domain.

	Ocean	Land
P0>0.5	85.21%	89.95%
P4>0	11.13%	9.73%
P5>0	4.52%	4.23%
P4+P5>0	11.41%	10.13%
P4>0 and P5>0	4.27%	3.83%

Figures

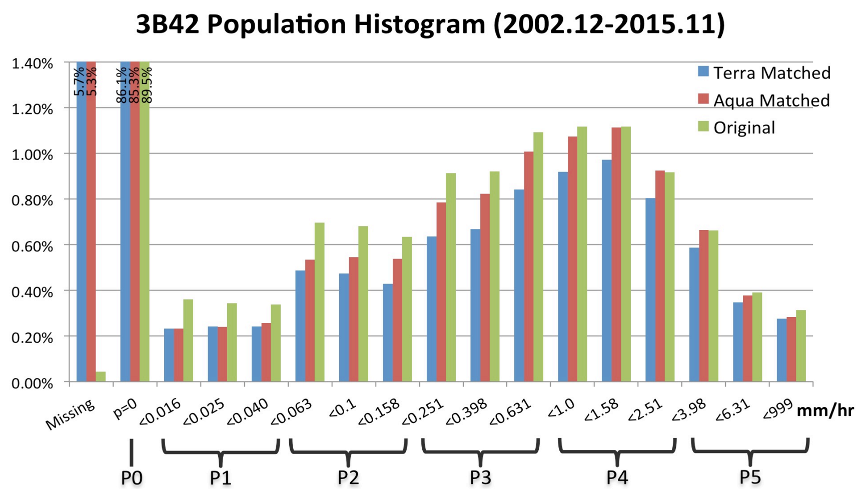


Figure 1: Histograms of TMPA original 0.25°×0.25° 3-hourly 3B42 precipitation data (green), and subsets matched with daytime Terra (blue) and Aqua (red), from December 2002 to November 2015 in the extended tropics domain. The boundaries that define the six simplified precipitation groups are shown at the bottom.

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Definitions of

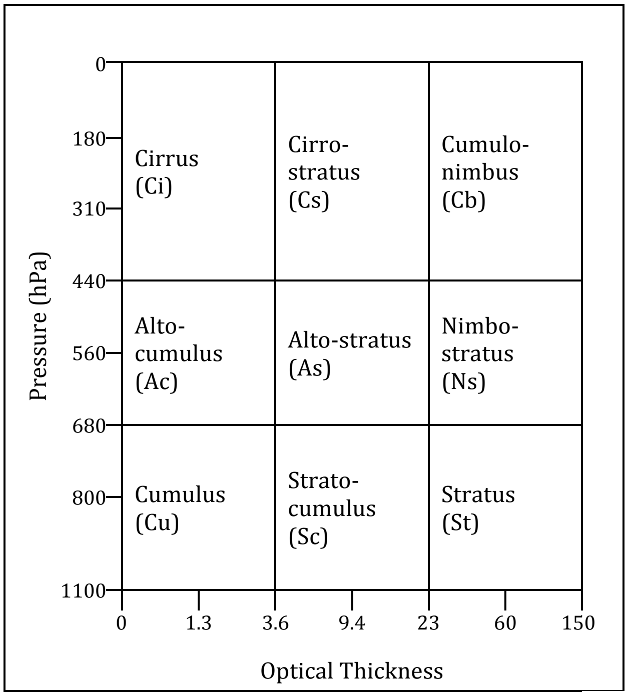


Figure 2: ISCCP cloud types assigned to groups of bins in MODIS joint histogram of τ–p_c.

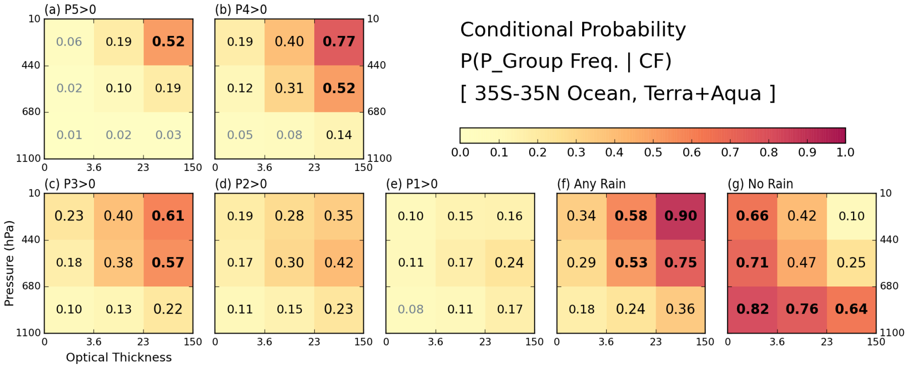


Figure 3: (a) to (e): Conditional probabilities of precipitation within a P-group (from TMPA) given occurrences of a cloud type (from MODIS) over ocean in the extended tropics from December 2002 to November 2015; (f): Conditional probabilities of any rain amount (sum of all P-group frequencies); (g): Conditional probabilities of no rain co-occurring with cloud. The CF threshold for cloud type occurrence is 6.25%.

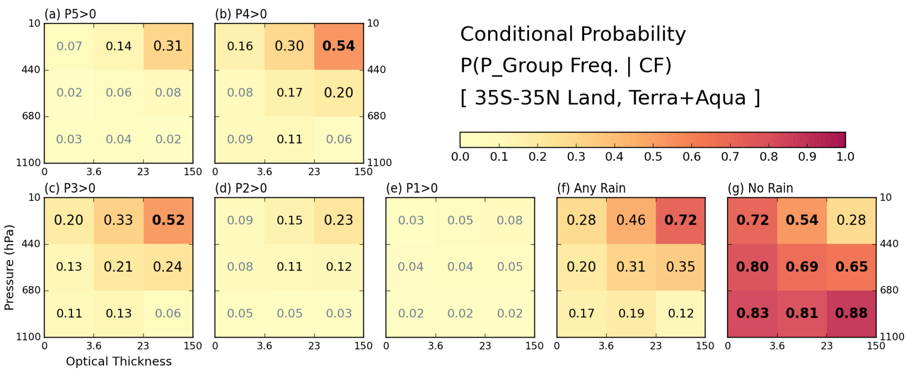


Figure 4: Same as Fig. 3, but over land

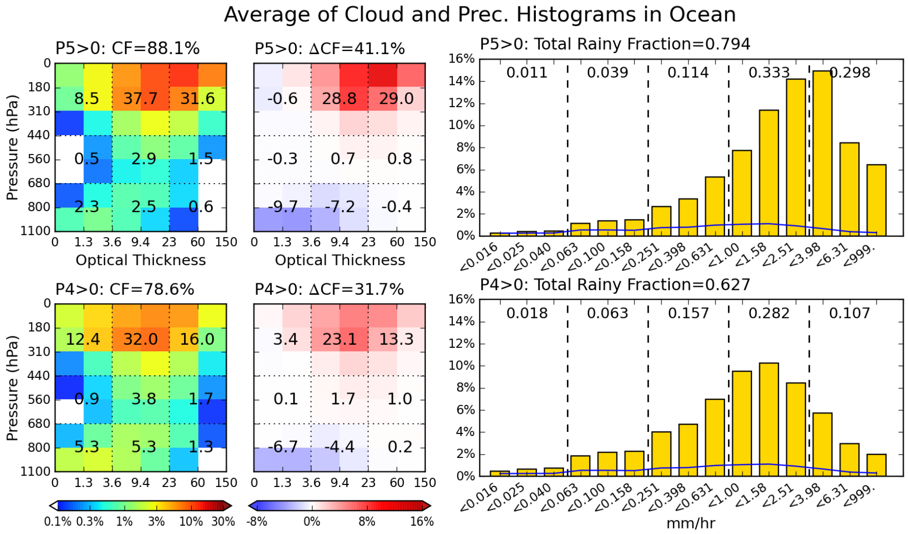


Figure 5: Conditional composite mean of 2D joint histogram of p_c and τ (left column), differences from overall (unconditional) mean (middle column) and precipitation histogram (right column) over the extended tropical oceans for 13 years. Top row is for P5, while bottom row is for P4 precipitation. Blue lines in precipitation histograms indicate the overall mean. Both cloud and precipitation overall means correspond to the entire domain, and not just ocean. Numbers on cloud histograms are the cloud fraction (CF; %) of each cloud type, which is the sum of 4 or 6 histogram bin values assigned to the cloud type. The sum of all values is equal to the total cloud fraction provided above each panel. Numbers on precipitation histograms are the fraction of each P-group, P1 (left) to P5 (right), obtained as the sum of three individual bin values. Total rainy fraction is the sum of all P-groups' fractions (i.e., sum of 15 individual bin values).

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Composite

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: climatology

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: climatology.

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: climatologies

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: for

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: shown in

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: titles

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: which is

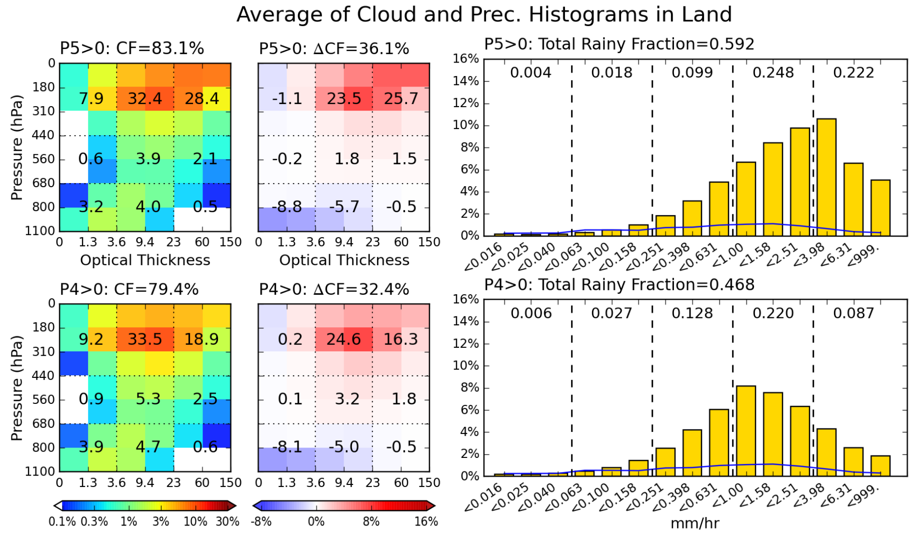
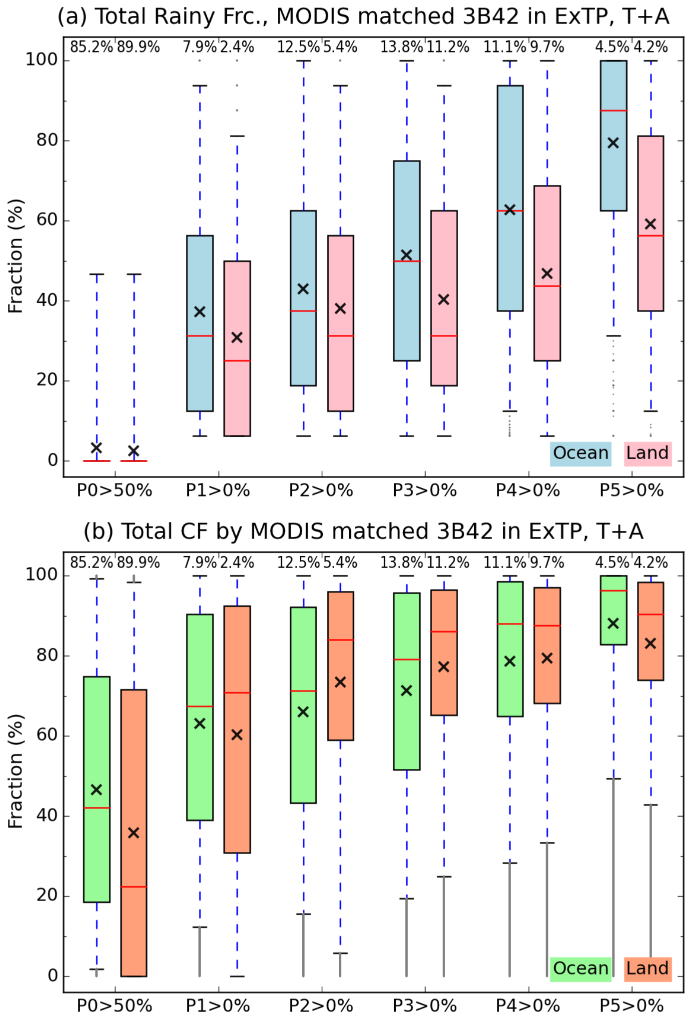


Figure 6: Same as Fig. 5, but over land.


(a) Total Rainy Frc., MODIS matched 3B42 in ExTP, T+A
85.2% 89.9% 7.9% 2.4% 12.5% 5.4% 13.8% 11.2% 11.1% 9.7% 4.5% 4.2%
Fraction (%)
0 20 40 60 80 100
P0>50% P1>0% P2>0% P3>0% P4>0% P5>0%
Ocean
Land
(b) Total CF by MODIS matched 3B42 in ExTP, T+A
85.2% 89.9% 7.9% 2.4% 12.5% 5.4% 13.8% 11.2% 11.1% 9.7% 4.5% 4.2%
Fraction (%)
0 20 40 60 80 100
P0>50% P1>0% P2>0% P3>0% P4>0% P5>0%
Ocean
Land

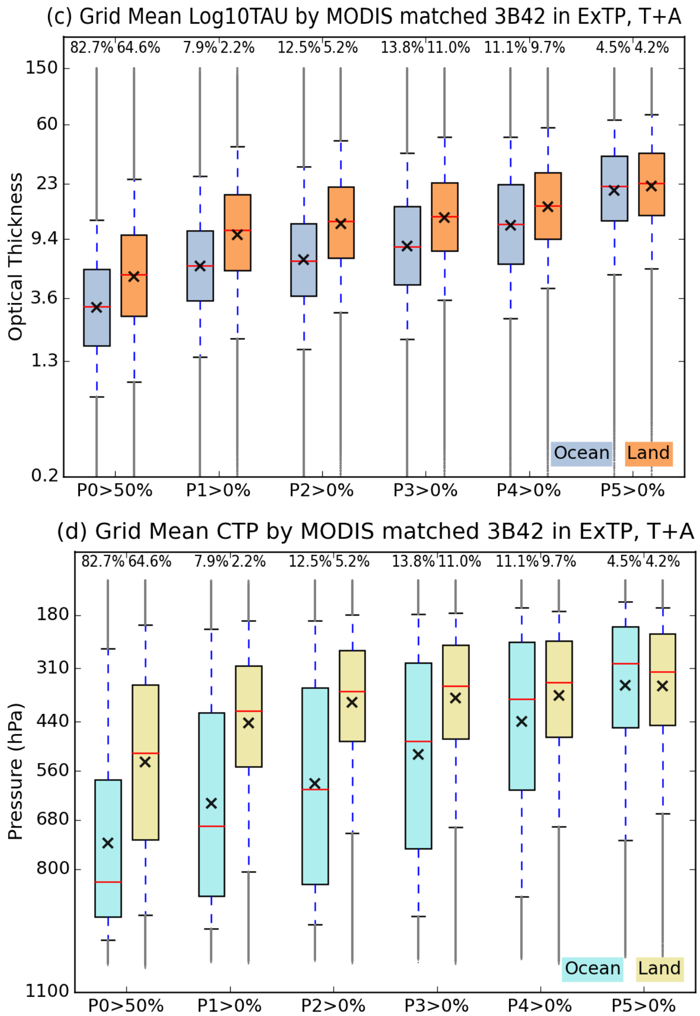


Figure 7: Box-whisker plot of (a) the total rainy fraction, (b) the total cloud fraction, (c) the grid mean *log10(τ)*, and (d) the grid mean p_c conditioned by precipitation groups, separately for ocean and land. The median values are shown as red horizontal lines, and the mean values are shown as black crosses. The vertical width of the boxes indicates the interquartile range (25th-75th percentile), and the whiskers extend from 5% to 95% values. Percentage numbers above the boxes indicate the occurrence ratio of each P-group relative to the total ocean or land grid cells.

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted:

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: on top of

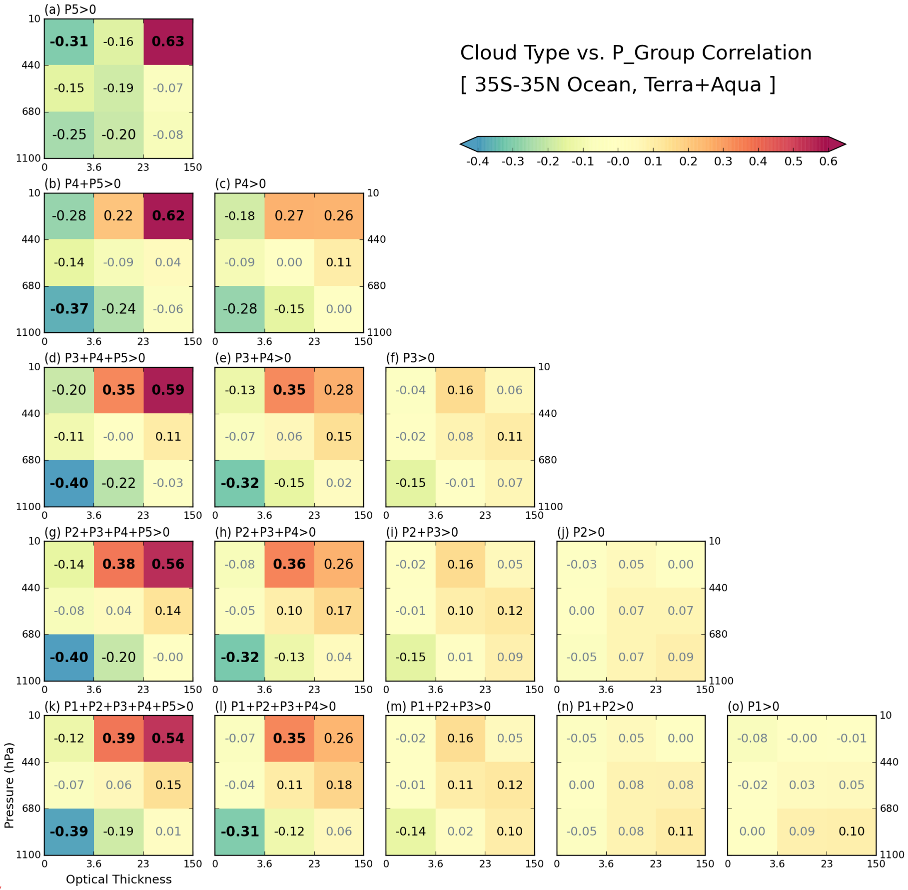


Figure 8: Cross-correlation coefficients in the extended tropical oceans for 13 years calculated between CFs of cloud types and precipitation group (individual or cumulative P-groups) values. The sum of all five precipitation groups shown in panel (k) corresponds to the total rainy fraction.

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: … [6]

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: indicates

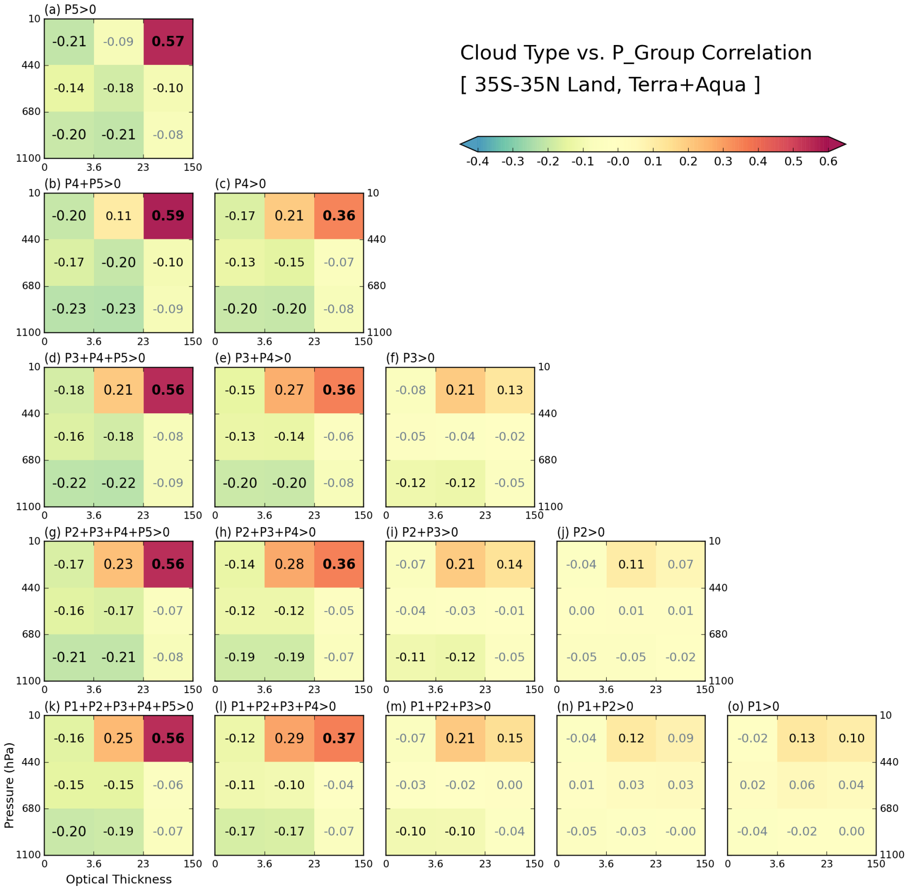


Figure 9: Same as Fig. 8, but over land.

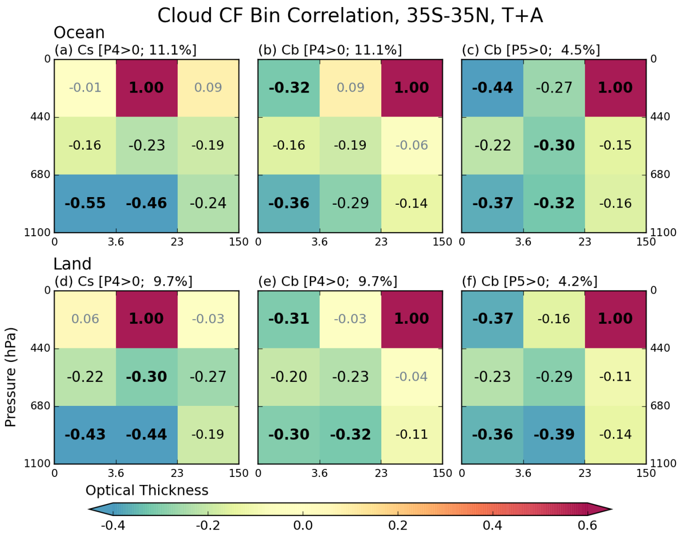


Figure 10. Conditional cross-correlation coefficients between cloud joint histogram bin CF values calculated for 13 years, based on (a) Cs CF over Ocean when P4>0, (b) Cb CF over Ocean when P4>0, (c) Cb CF over Ocean when P5>0, (d) Cs CF over Land when P4>0, (e) Cb CF over Land when P4>0, and (f) Cb CF over Land when P5>0. The percentage numbers above each panel are sample size ratios relative to the total number of ocean or land grid cells.

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Cross

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: are calculated

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Percentage of subtitle indicates

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: ratio over

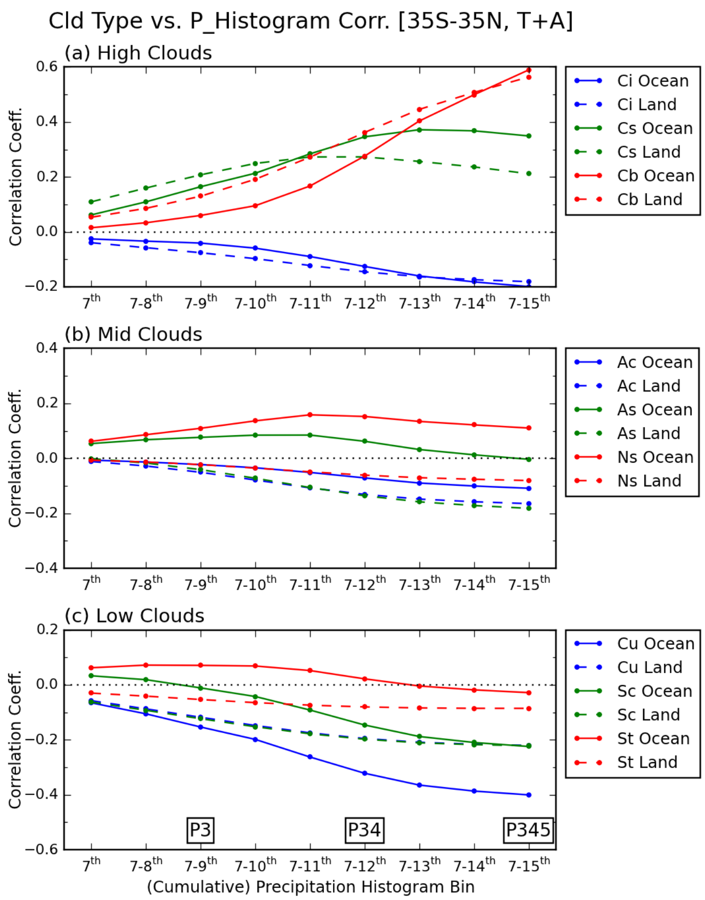


Figure 11: Correlation coefficients between cloud type CF and precipitation histogram values, for (a) high clouds (Ci, Cs, and Cb), (b) Mid-level clouds (Ac, As, and Cu), and (c) low clouds (Cu, Sc, and St). Precipitation histogram values are added cumulatively from the 7th bin onward, so the sum from the 7th to the 9th bin corresponds to P3, and so on. Oceanic cloud results are shown in solid and continental cloud results are shown in dashed lines.

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted:

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: accumulated

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: is the same

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: Ocean clouds

DAEHO JIN (GSF…, 12/21/2017 3:53 PM
Deleted: land clouds