We are grateful to the evaluations from the reviewers, which have allowed us to clarify and improve the manuscript. Below we addressed the reviewer comments, with the reviewer comments in italic and our response in bold.

J.Quaas (Referee #1)

Received and published: 7 November 2017

Bai et al. present a comprehensive calculation of statistics of precipitation vs. aerosol index/cloud droplet concentration from A-Train satellite retrievals. The document differences in the linear regression metrics when applying different (microwave vs. visible-near infrared) LWP retrievals; different precipitation retrievals (albeit all from CloudSat), different aerosol metrics (aerosol index from different retrievals vs. cloud droplet number concentration estimates), and different thresholds.

Although this work does not provide breakthrough science itself, documenting these differences in a consistent way is useful to the debate. The study is in general performed diligently and is pertinent to ACP

I have two main modifications I recommend, and several specific comments.
I have not highlighted semantic or orthographic mistakes since I assume the Copernicus copy editing will take care of this.

Main remarks:
1. The authors should expand their discussion of the state of the art, especially they need to discuss the various other aerosol-precipitation interactions beyond the "lifetime effect". It is in particular necessary that the authors discuss the role of aerosol scavenging when interpreting the metrics they investigate.
Thanks for your comments! We now added more discussion in Section 4 to acknowledge aerosol-precipitaiton interactions beyond the "lifetime effects", and it reads: "It should be noted that precipitation susceptibility in our study is based on Eq. (7) and is derived by linear regression between precipitation fields and CDNC/AI in log-log space. The negative/positive correlation between precipitation frequency/intensity and aerosols may not be readily explained as aerosol effects on precipitation. For example, a negative correlation between precipitation frequency and aerosols may come from the wet scavenging effects of aerosols (more precipitation leads to less aerosols) but not aerosol suppression of precipitation. However, in our study, we not only calculate precipitation susceptibility with respect to AI (S_{X_AI}), but also with respect to CDNC (S_{X_CDNC}) and the later one is expected to be less affected by the wet scavenging effects. The broad consistency between these two estimates shown in our results (Fig. 13), especially for the estimate of S_{POP}, lends the support to the limited

influence of wet scavenging in our estimate. Further support for this comes from the fact that precipitation susceptibility estimates based on the 1 degree L3 MODIS aerosol products are similar to those based on the 10 km L2 MODIS aerosol products (Fig. 4), as we would expect the wet scavenging effects are more important at smaller scales if the wet scavenging effects are a dominating factor. Nevertheless, the effects of wet scavenging can still be important in satellite studies of aerosol-cloud-precipitation interactions, and should be better quantified in future, perhaps in combination with model simulations.".

2. I am a bit astonished on how rather poorly the figures are done. The authors should take care revising these so that the content is more readily understandable.
Thanks a lot for your suggestions! We have revised most figures in the paper according to the comments provided by two referees. This includes adding 95% confidence intervals and zero lines to most figures, and narrowing the range of y-axis in most figures.

Specific remarks
1. P1 l15: Here and at a plenty of instances in the text, the relationship between aerosol and precipitation derived from the observations is overly readily interpreted in a cause-effect manner. If only this science was so easy, then a plenty of issues wouldn't exist. I urge the authors to thoroughly revise their text and imply causality only where they can prove it, or at least where they can corroborate cause-effect relationships. Why not interpret a negative aerosol index – POP relationship, for example, as showing the wet scavenging precipitation effect on aerosol?
The sentence is now reformulated to "We find that S_{POP} strongly depends on atmospheric stability, with larger values under more stable environments". We have added more discussion in Section 4 for addressing this issue. See our reply to your main remaks#1.

2. P1 l26: I suggest the authors adapt to the IPCC AR5 language and define the radiative forcing due to aerosol-cloud interactions ("cloud albedo effect") and cloud adjustments (all subsequent modifications). It is necessary that the authors put the "cloud lifetime effect" hypothesis into context of the manifold other hypotheses.
Thanks for this suggestion! We now adapted the IPCC AR5 language in the manuscript. The text in the first paragraph of the introduction now reads "Aerosol-cloud interactions play an important role in the climate system and affect the global energy budget and hydrological cycle. The effective radiative forcing from aerosol-cloud interactions (ERFaci), which includes the instantaneous effect on cloud albedo from changes in cloud condensation nuclei (CCN) or ice nuclei and all subsequent changes to cloud lifetime and thermodynamics, remains one of the largest

uncertainties in our estimates of anthropogenic radiative forcing (Boucher et al., 2013)." We also removed all other references to "cloud lifetime effect" in the manuscript and replaced it by "cloud water response to aerosol perturbations".

3. P2 l4: Also, the relationship would need to be linear (or more generally, of known, universal, monotonic functional form).

Thanks! We agree and now removed this sentence.

4. P4 l17: But L3 is at 1◦, so far from the stated 5 km resolution.

We are sorry for the confusion here. We now further clarified in the revised manuscript about how the collocation among different datasets is done (See the first paragraph in the section 2.1), including 1 degree MODIS L3 dataset. The reason why we used L3 aerosol product is because we would like to examine how aerosol homogeneity might affect the estimate of precipitation susceptibility. This is now added in the first paragraph of Section 2.2.1 and it reads "This MODIS Level 3 dataset has been used in previous studies to examine aerosol-cloud-precipitation interactions (e.g., L'Ecuyer et al., 2009; Wang et al., 2012) and is compared here with results form the MODIS Level 2 aerosol product to examine how aerosol homogeneity might affect precipitation susceptibility estimates.". The comparison between the MODIS L3 and L2 products is documented in details in Section 3.2.

5. P4 l25: again, why the two, and not only the L2 data?

Please see our reply to comment#4.

6. P5 l1: The authors should report exactly how the colocation is done.

Thanks for the suggestion. We now clarified this in Sec 2.1 and the text reads "MODIS cloud product and CPR radar reflectivity observations used in this study are both provided from the Caltrack datasets, which resample observations from many sensors under CALIOP subtrack with the horizontal resolution of 5km (see the website of http://www.icare.univ-lille1.fr/projects/calxtract/products for more information). For other aerosol and cloud products, including MODIS/CALIOP aerosol products and AMSR-E cloud products, they are further collocated into the CALIOP subtracks in the Caltrack dataset. For each CALIOP subtrack, the closest aerosol/ cloud retrieval sample within one-degree grid box (1°×1°) centered at this subtrack is chosen. To reduce the uncertainty in cloud retrievals, only samples where MODIS cloud fraction is equal to 100% are selected. "

7. P5 l5: A discussion of Christensen et al. doi 10.5194/acp-2017-450 would be useful here.

Thanks for bringing this paper to our attention! This is highly relevant to our study and this paper is now added and discussed in the revised manuscript and it reads "Retrievals of aerosol properties from passive sensors and lidar observation are both affected by clouds near the aerosol, and thereby result in overestimation for aerosol property (Chand et al., 2012; Tackett and Di Girolamo, 2009; Christensen et al., 2017). The extent of this overestimation may be different among different sensors, and depends on how far aerosol pixels chosen are from the corresponding cloud pixels (Christensen et al., 2017).".

8. P5 l7: Wood and Hartmann is a good paper, but is it a pertinent reference here?
Thanks! We checked and this is now removed.

9. P5 l12: What is a "pixel" here? A 1-km MODIS cloud retrieval, or rather an aggregated CALTRACK 5 km grid box?
Pixel here is an aggregated CALTRACK 5 km grid box as MODIS cloud product is provided from the Caltrack datasets. Sorry for the confusion. The sentence is reformulated to "To reduce the uncertainty when deriving CDNC, cloud pixels (identified by Caltrack-MODIS cloud product with the horizontal resolution of 5 km) where cloud optical depth is less than 3 and cloud fraction is less than 100% are excluded (Cho et al., 2015; Zhang and Platnick, 2011)."

10. P5 l15: Is this statement tested/implemented? Or is it just taken for granted from the Kubar study?
Thanks! We now examined the percentage of single layer clouds in our study and it is 94%, consistent with Kubar et al. (2009), and we now updated the text and it reads: "Additionally, we limit our analysis to warm clouds by screening cloud pixels with cloud top temperature warmer than 273K. Under these screening criteria, our results show that 94% warm clouds are single layered (93% in Kubar et al., 2009). Therefore, our analysis mainly focuses on single-layer clouds."

11. P6 l29: It is nonsense that LTS is able to clearly distinguish cloud regimes (e.g. Nam and Quaas doi 10.1002/grl.50945). Klein and Hartmann only show that the seasonal cycles of cloud fraction and LTS correlate.
We agreed that it is still a challenging task to find a unique metric to clearly distinguish different cloud regimes. In previous studies, several metrics were applied to define different cloud regimes. For instance, by using LTSS and vertical pressure velocity, Zhang et al., (2016) divided descending regimes into stratocumulus, transitional clouds and trade wind cumulus regimes. Webb et al., (2015) developed an index (ALPI) based on LTSS and precipitation to distinguish cloud regimes.

LTSS may have its limitation for defining different cloud regimes. However, our results show that precipitation susceptibility has clear LTSS-dependence, especially for S_{POP} (Fig 11 and Fig 13). This suggests LTSS provides a feasible way to examine how precipitation susceptibility may depend on cloud regimes. LTSS was also used in many previous studies (e.g.,L'Ecuyer et al., 2009; Terai et al., 2015). Nevertheless, we acknowledged the limitation of LTSS in the revised manuscript in Section 3.6 and it reads: " Our results also suggest that it is important to account for the influence of atmospheric stability owing to the clear dependence of S_{POP} on metrics like LTSS, though it is acknowledged that LTSS alone is an imperfect metric for isolating cloud regimes (e.g., Nam and Quaas, 2013). Different metrics associated with cloud regimes should be examined in future to better understand the effect of cloud regimes on precipitation susceptibility. For instance, LTSS can be combined with vertical pressure velocity to distinguish between different cloud types (Zhang et al., 2016).".

12. P6 l30: The term "unstable" is a misnomer. "unstable" would mean, a negative LTSS.

Given that our study focus on ocean warm clouds mostly with positive LTSS values, we followed the same definition of unstable as L'Ecuyer et al., (2009) and Wang et al., (2012), and they both defined unstable environment by LTSS values less than 13.5K.

13. P7 l10: Why this choice and not deciles?

In our analysis, we keep the LWP bins the same when we compare different satellite products in individual plots in order to facilitate the comparison. So the number of samples for each LWP varies, from 5% to 14%. However, each LWP bin still includes more than ten thousand samples, large enough for producing robust estimate of precipitation susceptibility.

14. P9l5: When using AMSR-E LWP, are the pixels selected overcast at AMSR-E footprint? Or is the AMSR-E interpolated to the CALTRACK grid cells?

The AMSR-E pixels closest to CALTRACK grid cells are selected. We do not require the AMSR-E pixels to be overcast, but clouds from the CALTRACK pixels are overcast with MODIS cloud fraction of 100%. The details of collocation strategy are added to Sec 2.1 in the revised manuscript.

15. P9 l17: Is this may be due to the fact that AMSR-E LWP in fact is cloud fraction times in-cloud LWP, in combination with the fact that CDNC is positively correlated to cloud fraction (Fig. 2)?

Thanks! If this is the case, for a constant AMSR-E LWP shown in Fig. 7f, in-cloud LWP would decrease with increasing CDNC as increasing CDNC means increasing cloud fraction. Smaller in-cloud LWP would then imply lower precipitation intensity, opposite to what is shown in Fig. 7f. Our

results shown in Fig. 8 suggests that this might be related to differences in MODIS and AMSR-E LWP at low MODIS CDNC, but what might cause the discrepancies in two LWP product still needs further investigation in the future.

16. Fig. 1: Since there are only ten bins in LWP, I suggest to label each bin center on the x-axis. Possibly the axis could be chosen irregular then. It would be good to indicate the total amount of data points in the caption. In (a) a zeroline would be helpful.

We have added a zeroline in the Fig. 1. The total number of data points now is included in the caption. Since labels of the x-axis would be dense and overlapped at low LWP if we label each mean value of LWP bin, the x-axis now is divided into smaller intervals. In addition, each mean value of LWP bin is shown in Fig.7.

17. Fig. 4: The authors need to choose a different y-axis that spans only the range of data. As it is now, no details can be distinguished. Again, a zeroline is necessary

The range of y-axis is now narrowed and a zeroline is also added for most figures (Fig.1, Fig.3-Fig.5 and Fig.9-Fig.11)in the revised manuscript.

18. Fig. 5: zeroline would be good

A zeroline is now added in the Fig.5.

19. Fig. 7: a, b, e, f: more x-axis tick marks necessary; e-f: more y-axis tick marks necessary

More x-axis and y-axis tick marks are now added accordingly.

20. Fig. 10 b: zeroline necessary

A zeroline is now added in the Fig.10.

21. Fig. 12: the color code is poorly selected. The colors should be centered around zero (light pink shouldn't indicate positive). Are the LTSS bins chosen so that each contains on average the same amount of pixels (that is the way it should be, else a PDF of LTSS would need to be shown).

We have changed the color code of Fig. 12 and its colors are now centered around zero. The light pink now indicates negative value. In this figure, each LTSS bin now contains on average the same amount of pixels. We also have added this sentence to the caption of Fig. 12.

22. Fig. 14: is it not possible to differentiate likelihoods, e.g. by putting equal weight on each curve entering the shaded area and then varying the color intensity/darkness?

Thanks a lot for your suggestion! We would like to take this suggestion, but as we only have eight curves for each metric shown in Fig. 14, we do not

have large number of curves to show the likelihoods., so we have to keep the figure as it is.