We thank the reviewer for the insightful review that raises significant issues on the new method that we propose to use to calculate the susceptibility. The discussion section is expanded, and an appendix section is added to address the issues raised by the reviewer. The Appendix, where most of the revisions based on the reviewer's comments were made, is attached at the end of this document.

The reviewer's concerns are summarized in the two questions:

Is the susceptibility and the procedure to calculate it described in the paper an unbiased estimator of the exponent β?

Can the concept of susceptibility be applied to non-precipitating clouds?

Our short response to the first question is that we agree with the reviewer that the new method for calculating the susceptibility does not provide an unbiased estimate in all cases. We show below that it is dependent on the distribution of the data by following the example of the reviewer and calculating the susceptibility on a random dataset of aerosol concentration N and associated precipitation rate R. We then argue that the aerosol concentrations in the aircraft data from VOCALS closely follow a lognormal distribution, a distribution and that data distributed this way show little bias in derived exponents.

We then argue why the $R=a[(N/N_0)^{-\beta} -1]$ model is inappropriate to study the effect of including non-precipitating clouds in the calculations. We attempt to explain the behavior of the susceptibility in Fig. 3 of the reviewer's comments. Using the initial $R=aN^{-\beta}$ model, we then explore how our method performs when noise is added to the data and how it compares to other methods of calculating β. Finally, we address the second question and discuss the implications of including non-precipitating clouds in calculating susceptibility. We find that no method consistently provides an unbiased estimate of β.

We arrive at the conclusion that including non-precipitating clouds doesn't necessarily give you a more accurate value of β, but more of a richer picture. The susceptibility estimate that incorporates non-precipitating clouds also doesn't run as great a risk of being strongly scale dependent. At the smallest scales f is either 1 or 0. A susceptibility calculated based on just precipitating clouds disregards those cloud segments where $f=0$. However, as we increase the length of the segments over which we average the precipitation rate, those 0's are increasingly incorporated into calculating the segment mean precipitation rate. Therefore, with coarser resolutions an increasing amount of non-precipitating clouds will be incorporated into an analysis that includes only 'precipitating' segments.

Comment 1:
1. *Is the susceptibility and the procedure to calculate it described in the paper an unbiased estimator of the exponent β?*

No, but as the analysis below will show, none of the existing methods are unbiased estimators in all cases.

As the reviewer correctly points out in Section 2.a, even if the R-N relationship indeed followed the form $R=aN^{-\beta}$, the new method that we propose does not generally estimate the underlying relationship without bias. When we recreate a set of points

following the relationship $R=aN^{\beta}$, like the reviewer, we find that the bias is highly dependent on the distribution of N, i.e. whether N is distributed uniformly, normally, or log-normally.

As can be seen in Fig. 1, when N is distributed uniformly or even normally, the number of points in the low N/high R region is small. As a result, when we take the arithmetic mean of the top tercile (red circle) the point does not lie on top of the distribution. Note that the estimated slope is higher, because of the shift of the point that corresponds to low-N and high-R.
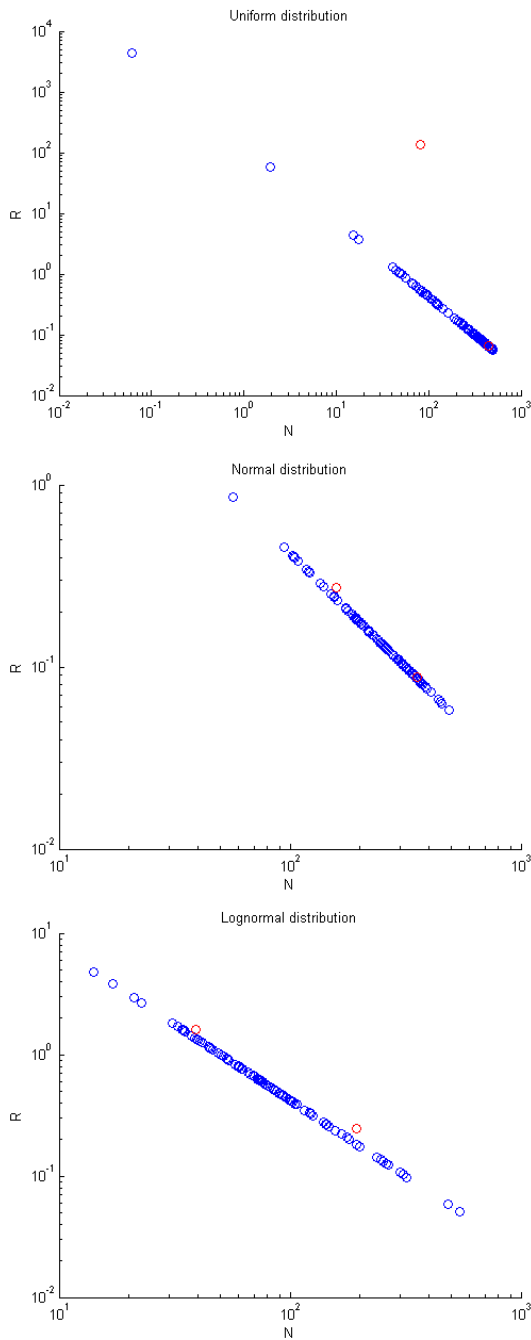
Figure 1: Distribution of R and N, where N is distributed uniformly, normally, and log-normally. Red circles represent the mean of R and N in the upper and lower tercile of N.

If we set the parameters such that $\beta=1.25$ and $a=50^{1.25}$, calculate the susceptibility using the method that we propose, and repeat this 100 times, much like the test that the reviewer conducted, we get a distribution of susceptibility estimates for the three types of N distributions, as shown in Fig. 2. As expected from Fig. 1, the new method significantly overestimates the β value when N is distributed uniformly or normally. On the other hand, when N is distributed lognormally, the new method estimates β with little or no bias. While we have only shown the case where $\beta=1.25$, a similar behavior is observed for β values up to 3. The only difference is that the width of the distribution of susceptibility estimates increases with increasing β.
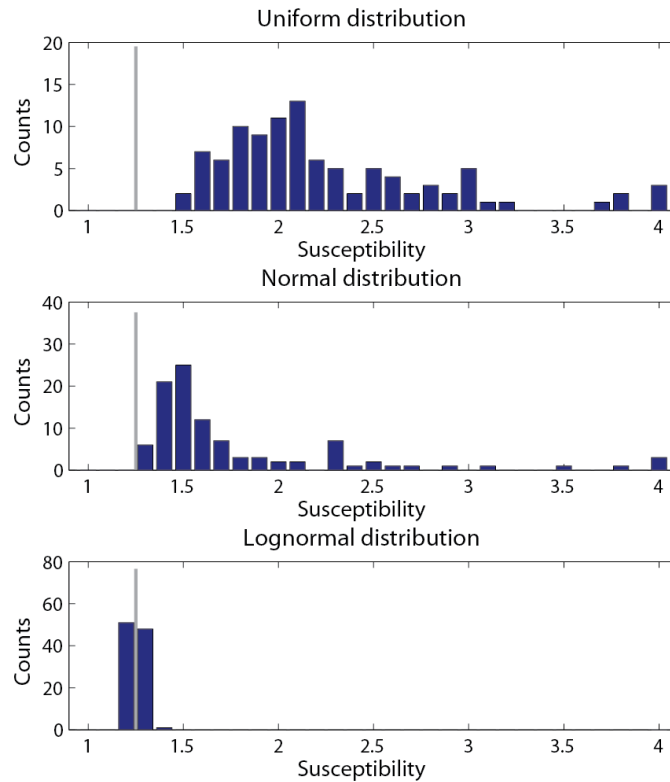

Figure 2: Histogram of susceptibility estimates using TLD for three different distributions of data: uniform (top), normal (middle), and lognormal (bottom). The histograms are based on 100 estimates calculated from 100 different samples of the same underlying distribution. The gray line at 1.25 shows the value of β in the underlying relationship $R=aN^{\beta}$.

From this short exercise we've found that our method is indeed not always an accurate predictor of the underlying β dependence, except for when the data is distributed log-normally. We also have found that the bias is a result of the inability of the arithmetic mean to be coincident with the geometric mean in the low-N, high-R range. The high bias

of the new method that the reviewer finds in Fig. 1 of the review is a result of the *N* distribution of the data.

We find that the distribution of the *N* data in our dataset is closer to a log-normal distribution than to either a normal or a uniform distribution, as shown in Fig. 3. The red line indicates the theoretical log-normal distribution, based on the arithmetic mean and standard deviation of the data in each of the four cloud thickness bins. The lognormal distribution of the aircraft *N* data indicates that we should expect relatively little bias in the *β* that we estimate. To address this issue, we have inserted a section on data distribution in the Appendix, attached below.
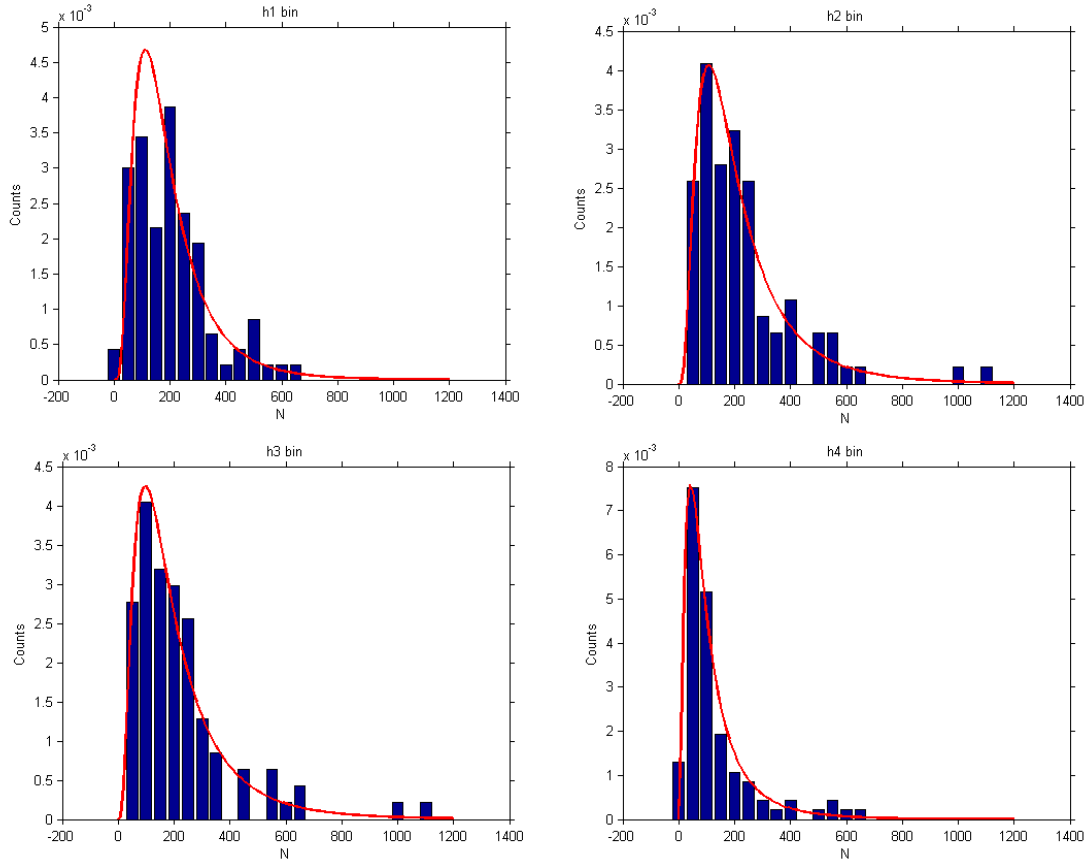

Figure 3: Histogram of the distribution of *N* in the data in the four *h* bins. The red line is the theoretical lognormal distribution, based on the arithmetic mean and standard deviation of the *N* data in each bin.

The problem, in my opinion, is the averaging performed on the data, which does not preserve non-linear relationships, i.e., the procedure cannot provide an unbiased estimate of β and the result should maybe not be called precipitation susceptibility. If anything, the authors introduce a new metric which might or might not be a useful one.

Since the *β* values estimated by the new method overestimate the *β* dependence when the data is not distributed lognormally, there is valid concern that perhaps this new

metric should not be called precipitation susceptibility. However, the newly proposed method is an estimator of the precipitation susceptibility, albeit a potentially biased one in some circumstances. The concern with our method would be that we are incorporating clouds with $R=0$ in calculating a metric that takes the logarithm of R; the logarithm of zero is negative infinity. However, as long as there are precipitating clouds in the high N range, the susceptibility introduced in this study captures the effect of aerosols in changing the mean precipitation rate of clouds of a given thickness in the framework of the precipitation susceptibility.

The true precipitation susceptibility, defined as $S = -\left(\frac{\partial \ln R}{\partial \ln N}\right)_{LWP}$ is a continuous function of N, and as such there is no one method that can ever determine this function from discrete observational data without some risk of that estimator being biased. The previous estimators, e.g., Sorooshian et al. (2009), do not provide enough information to ascertain what estimator of S was used and whether that too may be biased. If indeed Sorooshian et al. (2009) used a linear regression of log R on log N for fixed LWP bins (where N in this case is the CCN proxy), then we show below that such a regression is not necessarily an unbiased estimator of S if R and N have uncertainty (measurement and system noise) and minimum thresholds associated with them.

For this model with a precipitation threshold N_0 neither the standard approach for susceptibility nor the new method provide an unbiased estimate of $\beta = 1.25$. Both methods result in a susceptibility estimate which is much larger.

The proposed model follows the form

$$R = a\left[\left(\frac{N}{N_0}\right)^{-\beta} - 1\right] \qquad \text{for } N < N_0, \text{ and}$$
$$R = 0 \qquad\qquad\qquad\qquad \text{for } N \geq N_0.$$

We find that the model above is inappropriate to test the performance of the method in calculating the underlying β value. From the dataset alone, we cannot deduce the exact nature of a, and we do not know whether such a threshold value N_0 exists. For example, Fig. 3 of the satellite study of L'Ecuyer et al. (2009) suggests that in observed warm maritime clouds there is no such threshold behavior, where precipitation is 0% below some threshold value and 100% above another threshold value.

Furthermore, if we plot values of R and N that follow this relationship in a log-log plot, as in Fig. 4, we can see that the slope of the relationship in log-space steepens at higher N values. The slope starts off relatively close to the $-\beta$ value, but steepens considerably when $a(N/N_0)^{-\beta}$ becomes comparable to a. Any kind of linear fitting method in log space will calculate a slope that is larger than β.

L'Ecuyer, T. S., Berg, W., Haynes, J., Lebsock, M., & Takemura, T. (2009). Global observations of aerosol impacts on precipitation occurrence in warm maritime clouds. *Journal of Geophysical Research*, 114(D9), 1-15. doi:10.1029/2008JD011273
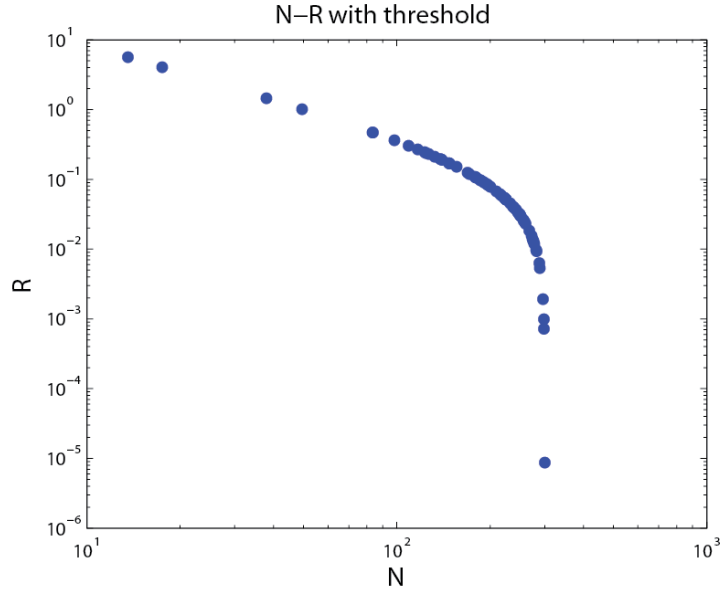
Figure 4: Plot of N vs. R in log axes for the relationship that includes a threshold (i.e. $R=a[(N/N_0)^{-\beta} -1]$).

This Fig. 3 shows in fact a certain similarity to Fig. 3 of the paper, and I would ask the authors to provide further evidence that their results are not an artifact of the analysis method.

The result from Fig. 3 of the reviewer's comments, where β is calculated in the four bins, is a result of the change in slope of the underlying relationship, as seen in Fig. 4. Whether we add noise to this relationship or not, the underlying slope will still steepen with increasing N, and hence result in a behavior similar to Fig. 3 of the reviewer's comments.

We have tried to recreate the 'dataset' with a threshold behavior ($R=a[(N/N_0)^{-\beta} -1]$) and added noise (normal distribution) with a mean of 0 and standard deviation of 0.01. $N_0 = 450$, a is calculated such that R ($N=50$) $= 1$. We created four bins with the bin edges of [50 150 250 350 450], as in the reviewer's comments. We then created a set of 100 points with N in each of those bins following a random uniform distribution. The corresponding R was calculated ($R=a[(N/N_0)^{-\beta} -1]$), and then noise was added to the R value. Then we calculated the susceptibility, as described in our paper. This was repeated 100 times to obtain 100 susceptibility estimates in each of the four bins.

The total susceptibility (blue in Fig. 5) looks like Fig. 3 of the reviewer's comments and indeed follows a similar behavior as the susceptibility in Fig. 3 of the paper. If, however, we look at the behavior of S_I (red in Fig. 5), the β calculated when non-precipitating points are excluded, we see that S_I decreases with increasing N, a feature absent in Fig. 3 of the paper. Therefore, we do not believe that the behavior of the susceptibility in Fig. 3 of the reviewer's comments is a direct comparison of the susceptibility in Fig. 3 of the paper.
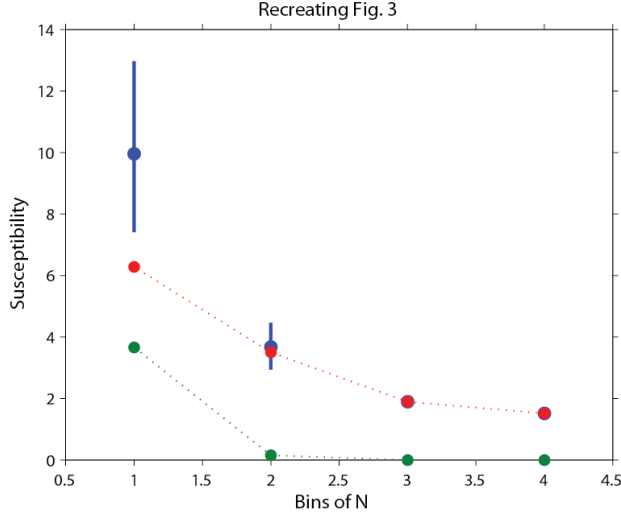
Figure 5: Susceptibility estimates from four N bins, where noise is added to R, given the underlying relationship $R=a[(N/N_0)^{-\beta}$ -1. The ranges of N for each of the bins from left to right are 350 to 450 for bin 1, 250 to 350 for bin 2, 150 to 250 for bin 3, and 50 to 150 for bin 4. Total susceptibility, S_R, is in blue, susceptibility of drizzle fraction S_f is in red, and susceptibility of drizzle intensity S_I is in green.

It remains important to investigate how linear regression in log space and our new estimator are able to estimate an underlying β value (if it exists) if noise also exists in the data. We propose using the initial model $R=aN^\beta$, and adding noise to this distribution so that some non-precipitating points ($R=0$) are introduced. In creating this model, we can also examine how N changes with the changes in precipitation threshold. When creating this 'noisy dataset,' we use a lognormal distribution of N because (a) we have determined there is little if any bias when calculating β using the method in a lognormal distribution, and (b) because observed N is distributed approximately lognormally (Fig. 3 above). Any biases in the new method, would therefore be caused by the addition of noise in the data.

To study the impact of noise on the different methods estimating β, we take 100 random samples of N, taken from a lognormal distribution with an arithmetic mean of 150 and standard deviation of 75 and calculate R as before ($R=aN^\beta$, $a=50^{1.25}$, $\beta=1.25$). The standard deviation of R before adding the noise is typically 0.22. To the R value we then add noise taken from a normal distribution with a mean of zero and standard deviation σ_{noise}. If R is negative after adding the noise, then R is set to zero, since R represents a precipitation rate. We then calculate the susceptibility using three methods, and repeat the process 100 times to obtain a distribution of susceptibilities for each method and for four different noise levels (σ_{noise}= 0.02, 0.1, 0.2, and 0.3). The three methods are 1) the method used in the paper, which we call TLD or tercile log-differencing, 2) standard linear regression in log-space, and 3) a linear regression fit in log-space based on minimizing the perpendicular distance between the fit and the data, as discussed by Reed (1992). Each data point is equally weighted in all cases.

The sensitivity to noise (Fig. A3) shows that all three methods accurately estimate the underlying β value for low noise, but as σ_{noise} increases, the minimum distance method increasingly overestimates the β value, while both TLD and the standard linear

regression method capture the true β value with minimal bias. The standard linear regression most likely outperforms the minimum distance method, because noise is only added to R, and one of the main assumptions of the standard linear regression is that errors in N are zero or negligible. The minimum distance method assumes that errors exist in both R and N. We have not carried out a test where we add noise to N.

Reed, B. C.: Linear least-squares fits with errors in both coordinates. II: Comments on parameter variances. *American Journal of Physics*, 60(1), 59, doi:10.1119/1.17044, 1992.
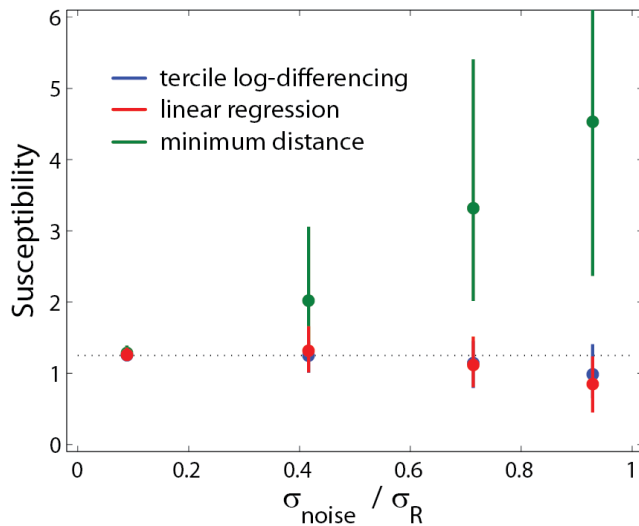
Figure A3: Susceptibility estimates from TLD (blue), linear regression (red), and minimum distance (green) with increasing noise level. Dots represent mean susceptibilities based on 100 estimates, and the lines show middle 95 % interval. The abscissa shows the ratio between the standard deviation of the distribution from which the noise is taken, σ_{noise}, and the standard deviation of R, σ_R, after the noise has been added. The dotted line represents the underlying β value.

For one last experiment, we apply a minimum threshold for the 'measurable' precipitation. This test recreates a situation similar to real data, where a minimum threshold for precipitation usually exists.

We test how accurately the three estimators are able to capture β when we apply a minimum threshold to R. We use the same underlying lognormal distribution as in the previous test to obtain 100 random samples of N, and the same relationship between R and N. We maintain the noise level at $\sigma_{noise} = 0.3$ and vary the minimum threshold of R such that values of R less than the threshold are set to zero. We choose $\sigma_{noise} = 0.3$, because this gives a σ_{noise}-to-σ_R ratio that is similar to those found in the VOCALS observations. The mean value of R following the addition of noise, but before the threshold is applied, is typically 0.37. The four threshold R values we use are 0.01, 0.2, 0.4, and 0.6.

Susceptibility estimates from all three methods (Fig. A4) are sensitive to the threshold value. The linear regression method increasingly underestimates the underlying

β value as the threshold increases. This result is consistent with that of Jiang et al. (2010) and Duong et al. (2011), who both found that increasing the minimum threshold for precipitation decreased the susceptibility estimate. Though the minimum distance method overestimates β when the threshold is near zero, it also follows the same trend of decreasing susceptibility estimates with increasing threshold value. The TLD, on the other hand, overestimates β with increasing threshold value. The susceptibility estimate positively correlates with the fraction of non-precipitating points in each set (Fig. A5). In general, higher susceptibility values are found with increasing fraction of non-precipitating points. From this analysis alone, however, we cannot determine what non-precipitating fraction would always give an unbiased estimate of β.

If we split the susceptibility S_R of the TLD method into S_f and S_I, as done in the body of the manuscript, we find that S_f increases with increasing fraction of non-precipitating points and provides the vast majority of the trend in S_R (Fig. A6). On the other hand, S_I decreases with increasing fraction of non-precipitating points, much like the standard linear regression in Fig. A4. No method consistently gives an unbiased estimate of the underlying β in cases where a significant number of data values are determined to be non-precipitating.
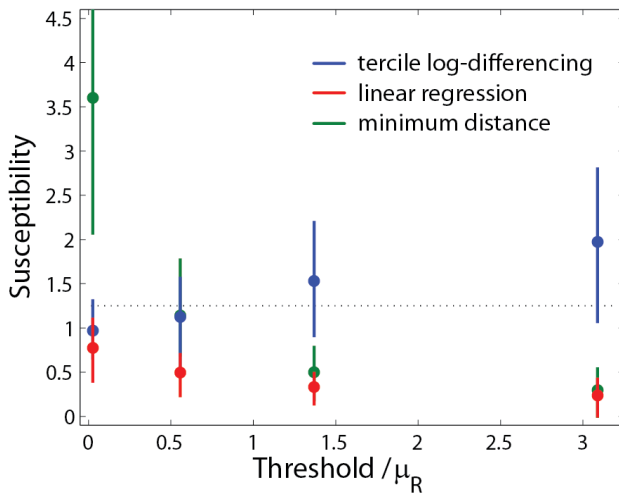

Figure A4: Susceptibility estimates from TLD (blue), linear regression (red), and minimum distance (green) with increasing threshold level. Dots represent mean susceptibility from 100 estimates and the lines show middle 95 % interval. The abscissa shows the ratio between the threshold value and the mean, μ_R, of R after applying the threshold. The dotted line represents the underlying β value.
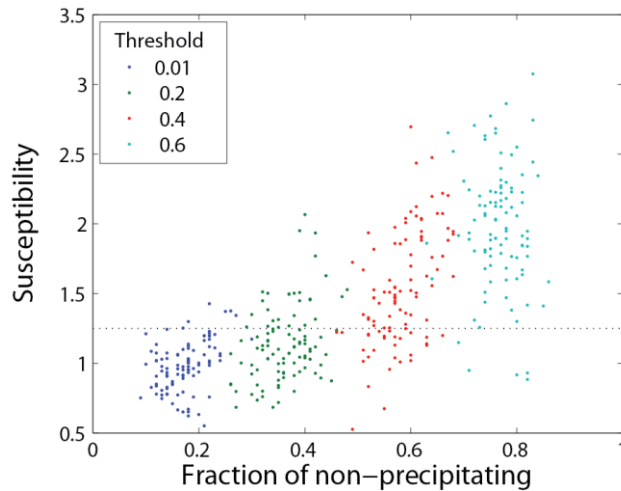
Figure A5: Susceptibility estimates from TLD as a function of the number of non-precipitating points that went into calculating the susceptibility. Colors indicate different levels of the thresholds: 0.01 (blue), 0.2 (green), 0.4 (red), and 0.6 (cyan). The dotted line represents the underlying β value.
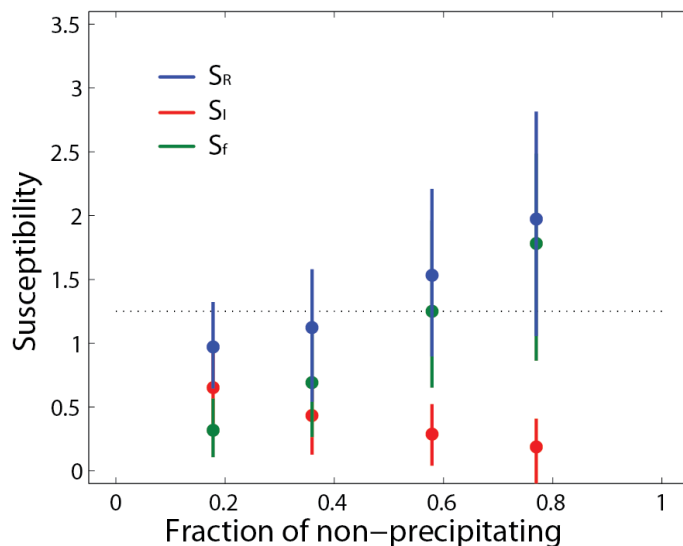

Figure A6: The precipitation susceptibility S_R (blue), susceptibility of drizzle intensity S_I (red), and susceptibility of drizzle fraction S_f (green) with increasing threshold value, calculated using the TLD method. Instead of the ratio between the threshold and mean precipitation rate, the average fraction of non-precipitating points at each threshold is used for the abscissa. The dotted line represents the underlying β value.

What we find in the analysis above is that no method will always give an unbiased estimator of β. The method in the paper, TLD, does appear to overestimate β with increasing numbers of non-precipitating clouds. However, when S_R is split into S_f and S_I, we do see that S_f plays an important role in determining S_R, giving us some insight into how non-precipitating clouds are affecting the overall susceptibility value. From a standard linear regression estimate, which does not incorporate non-precipitating clouds,

one cannot deduce solely from the susceptibility estimate whether one may be underestimating β, as is the case in Fig. A4.

As I have shown, the susceptibility S might lead to spurious results with unrealistic large values when R is close to zero or when non-precipitating clouds are included. Therefore I wonder whether the results for the bin h_1 can be included in the statistics. If this data is excluded, then the dependency of S on cloud depth becomes much less convincing, as the susceptibility is close to 1 for the remaining bins.

The reviewer has indeed demonstrated that the new method may lead to spurious results with large values of S given certain distributions of N. Given the analysis above, however, we find that the no method consistently captures the underlying dependence of R on N. From Fig. A5, we can also see that having a sample of 100% precipitating clouds doesn't assure an unbiased estimate of β. Furthermore, if we look at global observations of maritime warm clouds, as in Fig. 3 of the satellite-based study of L'Ecuyer et al. (2009), we can see that the probability of precipitation does not reach 100% even in the thickest of the observed clouds. This suggests that we cannot just disregard susceptibility estimates from cloud thickness bins where a significant fraction of the clouds are non-precipitating.

I would like to emphasize again that S does not need to be an unbiased estimator of β in Eqs. (2) or (4) to be useful, but it would be important to show what we can actually learn from S and, maybe even more important, it should be emphasized that the actual functional dependency of R on N might be very different from any simple dependency suggested by S.

The reviewer raises the underlying question about the appropriateness of the susceptibility metric as a model to quantify the effect of aerosol or cloud drop concentration on the precipitation. There is no theoretical reason to assume that the precipitation depends on aerosol concentration, following the form $R = aN^\beta$. Given our current lack of understanding of the relationship, however, we find that our current formulation for estimating the susceptibility provides a useful tool to compare the dependence of aerosols to precipitation across different studies.

We have included the following paragraphs in the discussion section to address this issue:

The method by which we calculate the susceptibility (Eq. 4) is different from previous studies that only incorporate precipitating clouds. One concern with our method would be that we are incorporating clouds with $R = 0$ in calculating a metric that takes the logarithm of R; the logarithm of zero is negative infinity. However, as long as there are precipitating clouds in the high N range, the susceptibility introduced in this study captures the effect of aerosols in changing the mean precipitation rate of clouds of a given thickness in the framework of the precipitation susceptibility. We provide a detailed analysis of the method that we use and its limitations in Appendix A. We find that susceptibility estimates from Eq. 4 are sensitive to the number of non-precipitating clouds that go into calculating the susceptibility. We also find that compared to other

methods of calculating the susceptibility calculated using Eq. 4 best captures the uncertainty in the susceptibility estimates.

For much of this study, we have worked under the simple framework for understanding the effect of aerosols on precipitation that follows the model of Eq. 2. The actual dependence of R on N may not be as simple. Regardless of the actual functional dependence of R on N, the utility of the S_R, we find, is in how it distinguishes the effect that aerosol concentrations have on the intensity of precipitation and the effect that they have on the fraction of precipitation.

Appendix

       We use the tercile log-differencing (TLD) method to calculate the precipitation susceptibility in this study so that non-precipitating clouds can be included in an analysis that tries to quantify the effect of aerosols on precipitation suppression. Since none of the methods that calculate the susceptibility using regression in log-space incorporate non-precipitating clouds, they neglect the cases where increased aerosols completely suppress precipitation. In this section, we take a critical look at the TLD method and explore how data distribution, noise, and thresholds can affect the susceptibilities obtained by TLD.

       To test how accurately the TLD method estimates a given underlying dependence of precipitation on aerosol concentration, we create multiple synthetic sample datasets with the relationship $R = a\,N^{\beta}$ and use each of these to estimate β. The synthetic model may not exactly capture the true physical dependence of precipitation on aerosol concentration. However, it is desirable that an analysis method can accurately estimate the value of β. In addition to using TLD, we estimate β using a standard least-squares regression in log-space and a linear regression fit in log-space based on minimizing the perpendicular distance between the fit and the data, as discussed by Reed (1992). Each data point is equally weighted in all cases.

Distribution of the data

       Many variables in the atmosphere are distributed normally; many are not. Depending on the spatial and temporal extent of the dataset, N, our controlling (independent) variable, can in principle be distributed in a number of ways. We find that the nature of this distribution has an important impact upon how effective TLD is in estimating β.

       To study this we create a sample dataset of 100 random N values, where N is distributed uniformly, normally, or lognormally. A sample size of 100 is chosen, because the sample size in each of the cloud thickness bins in the VOCALS data is approximately 100. Corresponding R values are calculated using the relationship $R = a\,N^{\beta}$, where $\beta = 1.25$; this lies between the mean susceptibility that we estimate for the VOCALS data and the susceptibilities estimated in previous studies. We set $a = 50^{1.25}$, but the results for this particular analysis are insensitive to the choice of a. To generate a distribution of susceptibility estimates from the three methods (TLD, linear regression, and minimum distance), we resample the set 100 times from the same underlying distribution and calculate the susceptibility in each case, giving us 100 susceptibility estimates. When N is distributed either uniformly or normally, the susceptibilities from TLD overestimate β, as can be seen in Fig. A1. In these cases, the concentration of points in log-scale is skewed towards higher N, resulting in the overestimate of β. When N is distributed lognormally, the concentration of points is not skewed in log-space, and the value of β is more accurately captured by the susceptibility. In this case, we note that the susceptibilities from the linear regression and minimum distance methods accurately capture β as one would expect for a dependence of R on N that is simply a power law unburdened with noise that is introduced by both measurement uncertainties and additional controlling variables.

Importantly, the 10 km-averaged PCASP aerosol concentration N (the primary independent aerosol variable used in this study), for each of the four cloud thickness bins (h_1, h_2, h_3, and h_4), is distributed approximately lognormally (Fig. A2). Neither a uniform nor a normal distribution describes the data well. This gives us confidence that the susceptibility from the TLD method is not likely to be a strongly biased estimator of β for our observed data.
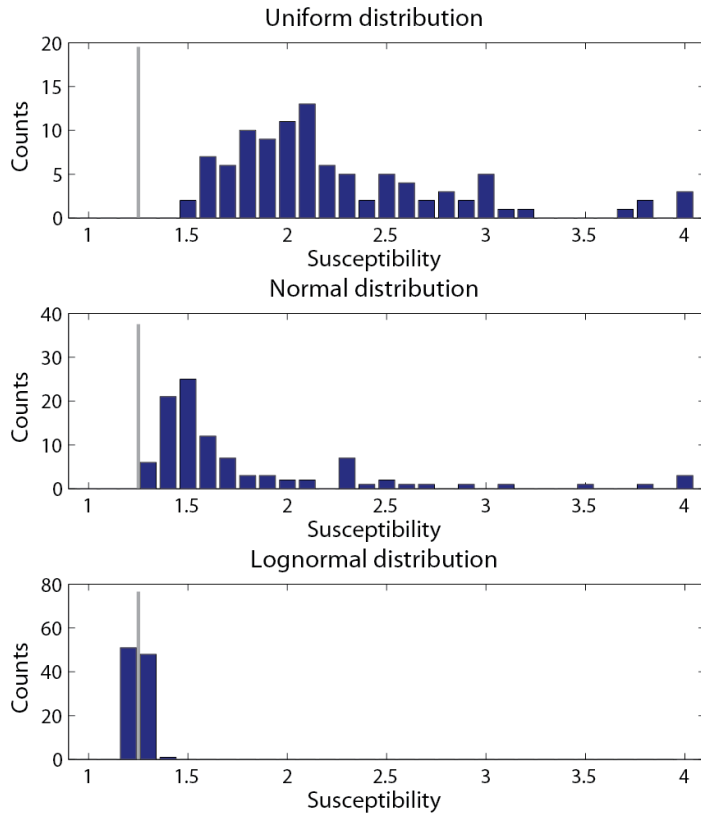

Figure A1: Histogram of susceptibility estimates using TLD for three different distributions of data: uniform (top), normal (middle), and lognormal (bottom). The histograms are based on 100 estimates calculated from 100 different samples of the same underlying distribution. The gray line at 1.25 shows the value of β in the underlying relationship $R = a\, N^{\beta}$.
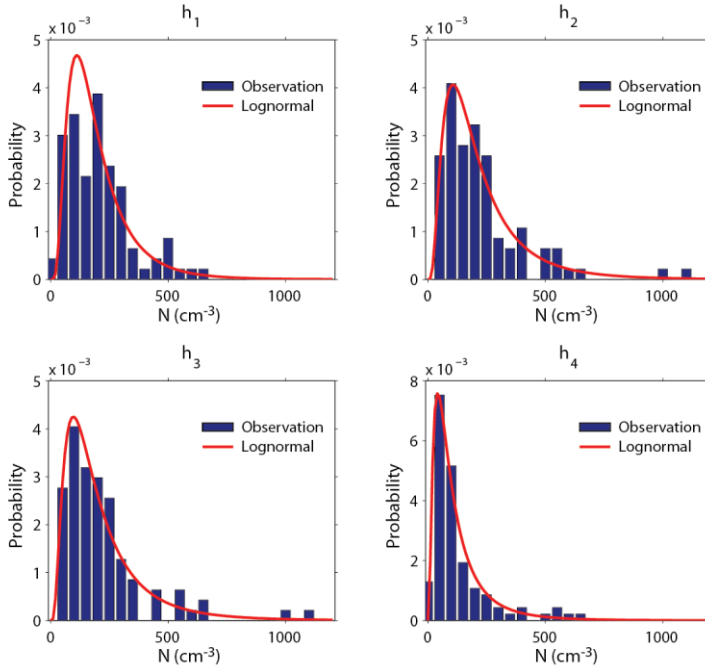
Figure A2: Histogram of 10 km-averaged PCASP aerosol concentrations N that are used in the susceptibility analysis. Each panel shows the distribution of N in the four cloud thickness bins. The red line shows the probability density function of a theoretical lognormal distribution, based on the arithmetic mean and standard deviation of N in each of the bins.

Noise level

In reality, we rarely expect the data to perfectly fit a model relationship. Instead, we expect there to be noise in the data, representing measurement uncertainties and additional unknown controlling variables. To study the impact of noise on the different methods estimating β, we take 100 random samples of N, taken from a lognormal distribution with an arithmetic mean of 150 and standard deviation of 75 and calculate R as before. The standard deviation of R before adding the noise is typically 0.22. To the R value we then add noise taken from a normal distribution with a mean of zero and standard deviation σ_{noise}. If R is negative after adding the noise, then R is set to zero, since R represents a precipitation rate. We then calculate the susceptibility using the three methods as above, and repeat the process 100 times to obtain a distribution of susceptibilities for each method and for four different noise levels (σ_{noise}= 0.02, 0.1, 0.2, and 0.3).

The sensitivity to noise (Fig. A3) shows that all three methods accurately estimate the underlying β value for low noise, but as σ_{noise} increases, the minimum distance method increasingly overestimates the β value, while both TLD and the standard linear regression method capture the underlying β value with minimal bias. The standard linear regression most likely outperforms the minimum distance method, because noise is only added to R, and one of the main assumptions of the standard linear regression is that

errors in N are zero or negligible. The minimum distance method assumes that errors exist in both R and N. We have not carried out a test where we add noise to N.
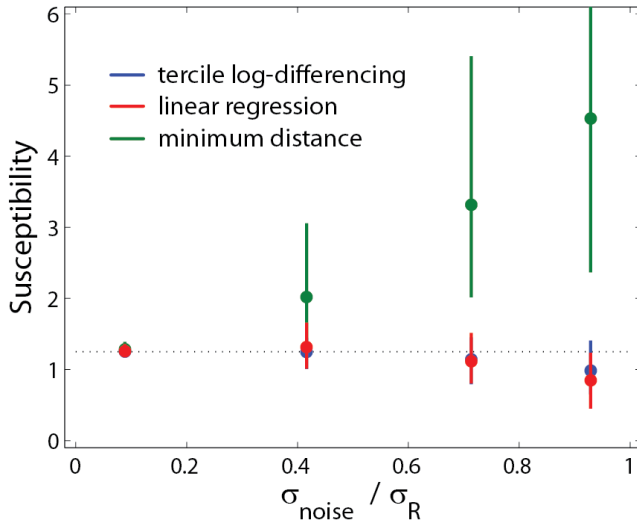

Figure A3: Susceptibility estimates from TLD (blue), linear regression (red), and minimum distance (green) with increasing noise level. Dots represent mean susceptibilities based on 100 estimates, and the lines show middle 95 % interval. The abscissa shows the ratio between the standard deviation of the distribution from which the noise is taken, σ_{noise}, and the standard deviation of R, σ_R, after the noise has been added. The dotted line represents the underlying β value.

Threshold

Previous studies of precipitation susceptibility have imposed different threshold precipitation rates to differentiate precipitating and non-precipitating clouds. Some of the differences are due to instrument sensitivities, others due to the authors' choices. In this study, we choose the -15 dBZ threshold, because precipitation rates above 0.14 mm day^{-1} (the corresponding precipitation rate) begin to have substantial effects (> 4 W m^{-2}) on the energetics of the boundary layer.

We test how accurately the three estimators are able to capture β when we apply a minimum threshold to R. We use the same underlying lognormal distribution as in the previous test to obtain 100 random samples of N, and the same relationship between R and N. We maintain the noise level at $\sigma_{noise} = 0.3$ and vary the minimum threshold of R such that values of R less than the threshold are set to zero. We choose $\sigma_{noise} = 0.3$, because this gives a σ_{noise}-to-σ_R ratio that is similar to those found in the VOCALS observations. The mean value of R following the addition of noise, but before the threshold is applied, is typically 0.37. The four threshold R values we use are 0.01, 0.2, 0.4, and 0.6.

Susceptibility estimates from all three methods (Fig. A4) are sensitive to the threshold value. The linear regression method increasingly underestimates the underlying β value as the threshold increases. This result is consistent with that of Jiang et al. (2010) and Duong et al. (2011), who both found that increasing the minimum threshold for

precipitation decreased the susceptibility estimate. Though the minimum distance method overestimates β when the threshold is near zero, it also follows the same trend of decreasing susceptibility estimates with increasing threshold value. The TLD, on the other hand, overestimates β with increasing threshold value. The susceptibility estimate positively correlates with the fraction of non-precipitating points in each set (Fig. A5). In general, higher susceptibility values are found with increasing fraction of non-precipitating points. From this analysis alone, however, we cannot determine what non-precipitating fraction would always give an unbiased estimate of β.

If we split the susceptibility S_R of the TLD method into S_f and S_I, as done in the body of the manuscript, we find that S_f increases with increasing fraction of non-precipitating points and provides the vast majority of the trend in S_R (Fig. A6). On the other hand, S_I decreases with increasing fraction of non-precipitating points, much like the standard linear regression in Fig. A4. No method consistently gives an unbiased estimate of the underlying β in cases where a significant number of data values are determined to be non-precipitating.
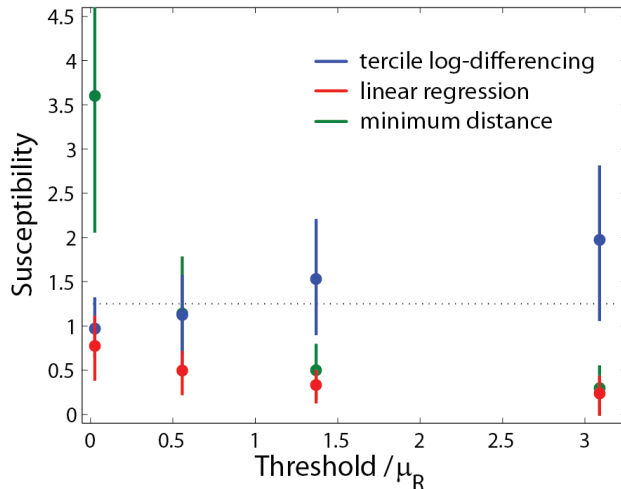

Figure A4: Susceptibility estimates from TLD (blue), linear regression (red), and minimum distance (green) with increasing threshold level. Dots represent mean susceptibility from 100 estimates and the lines show middle 95 % interval. The abscissa shows the ratio between the threshold value and the mean, μ_R, of R after applying the threshold. The dotted line represents the underlying β value.
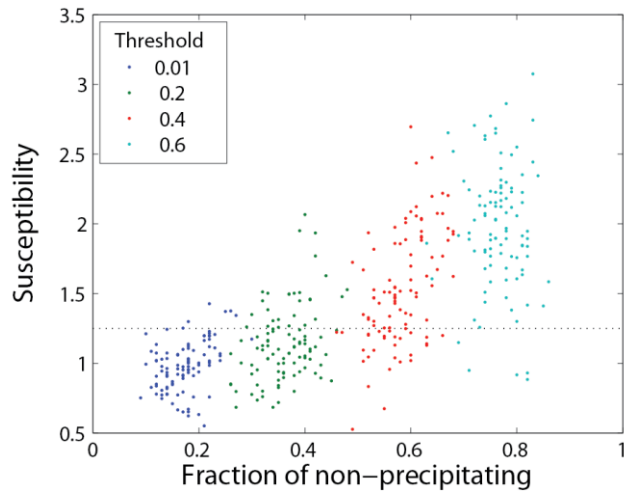
Figure A5: Susceptibility estimates from TLD as a function of the number of non-precipitating points that went into calculating the susceptibility. Colors indicate different levels of the thresholds: 0.01 (blue), 0.2 (green), 0.4 (red), and 0.6 (cyan). The dotted line represents the underlying β value.
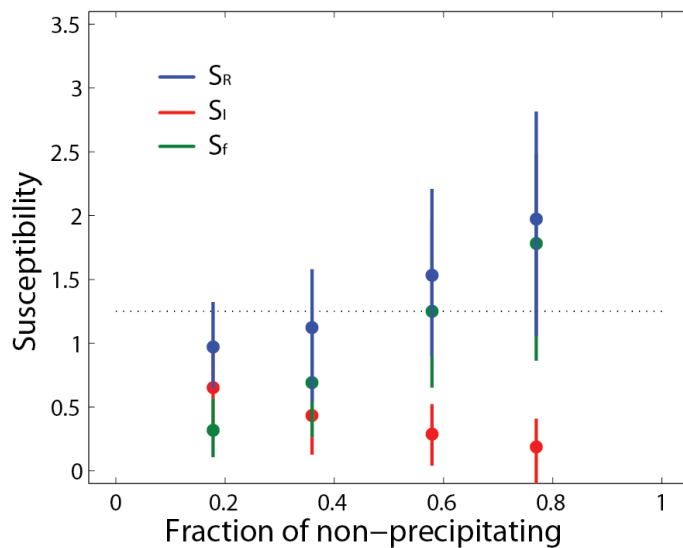

Figure A6: The precipitation susceptibility S_R (blue), susceptibility of drizzle intensity S_I (red), and susceptibility of drizzle fraction S_f (green) with increasing threshold value, calculated using the TLD method. Instead of the ratio between the threshold and mean precipitation rate, the average fraction of non-precipitating points at each threshold is used for the abscissa. The dotted line represents the underlying β value.

Discussion

We now attempt to put the data from VOCALS REx in context of the above analyses. We can see from Fig. A5 and A6 that susceptibility estimates from TLD increase with the fraction of non-precipitating points. Examining the fraction of non-precipitating clouds will give us an indication of the effect of the threshold on our

susceptibility results. The fraction of non-precipitating segments is 0.85, 0.46, 0.14, and 0.04 in the four cloud thickness bins of the 10 km-averaged VOCALS data (h_1 to h_4).

Estimating the 'noise' in the data is more difficult. To obtain some estimate of the noise level in the data, we can take the mean susceptibility values that we obtain in each cloud thickness bin and estimate the noise as the difference between the actual R and the R explained by the susceptibility. This crude estimate of the noise gives us σ_{noise}-to-σ_R ratios of 0.93, 0.91, 0.72, and 0.95 in the four cloud thickness bins (h_1 to h_4). This is not surprising, given that the magnitude of the correlations between N and R are relatively modest in each of the bins: -0.22, -0.28, -0.42, and -0.26 (h_1 to h_4).

We conclude that both threshold and noise play an important role in our dataset. The precipitation variations within each cloud thickness bin are dominated by noise, unexplained by the concentration of aerosol concentrations alone. In such cases, linear regression underestimates the β value. Figure A5 and A6 show that whether TLD method accurately estimates the β value is dependent on the threshold. We also note that Fig. A6 and Fig. 3 are mirror-images of each other, where the difference between the two is that the mean R increases along the abscissa in Fig. 3 and the threshold increases along the abscissa in Fig. A6. S_R increases with the increasing fraction of non-precipitating points. S_f, in both cases, determines the trend of S_R. S_I, on the other hand, display different behaviors in the two figures. S_I in Fig. A6 distinctly increases with decreasing fraction of non-precipitating points; S_I in Fig. 3 does not display such a clear increase. This suggests that the mechanism causing the behavior of the susceptibility in Fig. 3 is not quite identical to that in Fig. A6, though a large part may be due to it.

From the above analysis alone, we cannot disregard the possibility that in Fig. 3, the underlying dependence between aerosols and precipitation is constant and the decreasing trend of the susceptibility is solely because the fraction of non-precipitating clouds is decreasing. When none of the three methods above always give an unbiased estimate of β, the utility of S_R, as calculated using TLD, is found when S_R is taken as the sum of its parts S_f and S_I. It informs us about how both the rate *and* the frequency of precipitation depends upon aerosol concentration. S_I, which is more akin to the susceptibilities reported in previous studies, quantifies the effect of aerosols on how intense a cloud precipitates. S_f, on the other hand, is a metric that quantifies the effect of aerosols on the drizzle fraction, which is identical to the probability of precipitation when we include $f = 0$. L'Ecuyer et al. (2009) found that higher values of aerosol index, which serves as a proxy for columnar concentration of CCN-sized aerosols (Nakajima et al., 2001), tended to decrease the probability of precipitation. They also found that there is no unique liquid water path threshold above which a cloud can be assumed to be precipitating. This interesting finding runs counter to the idea that there is a threshold cloud liquid water path above which all clouds precipitate. S_f in this study attempts to quantify that same effect of aerosol concentrations on the probability of precipitation from the aircraft data from VOCALS. S_R in this study attempts to combine the effect of aerosol concentrations have in determining both the intensity and the probability of precipitation.

L'Ecuyer, T. S., Berg, W., Haynes, J., Lebsock, M., and Takemura, T.: Global observations of aerosol impacts on precipitation occurrence in warm maritime clouds, *J. Geophys. Res.*, 114(D9), 1-15, doi:10.1029/2008JD011273, 2009.

Nakajima, T., Higurashi, A., Kawamoto, K., and Penner, J. E.: A possible correlation between satellite-derived cloud and aerosol microphysical parameters, *Geophys. Res. Lett.*, 28(7), 1171, doi:10.1029/2000GL012186, 2001.

Reed, B. C.: Linear least-squares fits with errors in both coordinates. II: Comments on parameter variances, *American Journal of Physics*, 60(1), 59, doi:10.1119/1.17044, 1992.