We thank the reviewer for the constructive comments and suggestions. Individual responses to each of the reviewer's comments can be found below. Since some of the comments of the reviewer are addressed in the newly added Appendix, it is attached at the bottom of this document.

(i) The work presents highly valuable susceptibility numbers, which currently are scarce in the literature, especially when derived from aircraft measurements. At least one other stratocumulus dataset has been used to quantify precipitation susceptibility values (Lu et al., 2009). It would be helpful to compare the values in the VOCALS region to those in the Lu study, which also used airborne measurements. When doing this comparison, it should also be noted that different data analysis methodologies may have been applied in the two studies. Another recent study examining precipitation susceptibility that should be mentioned is that of Bangert et al. (2011) (i.e. see their Figure 11).

Lu et al. (2009), Marine stratocumulus aerosol-cloud relationships in the MASE-II experiment: Precipitation susceptibility in eastern Pacific marine stratocumulus, J. Geophys. Res., 114, D24203, doi:10.1029/2009JD012774.
Bangert et al. (2011), Regional scale effects of the aerosol cloud interaction simulated with an online coupled comprehensive chemistry model. Atmos. Chem. Phys., 11(9): 4411-4423.

At the end of the second paragraph under Discussion and conclusions we have inserted: The precipitation susceptibility estimates in this study are generally higher than previous airborne studies of marine stratocumulus clouds, such as those of Lu et al. (2009). In their study, they found susceptibility estimates of 0.46 from MASE I and 0.63 from MASE II, based on their separately measured sensitivities of precipitation to cloud droplet number concentration and of cloud droplet number concentration to aerosol concentration.

Bangert et al. (2011) provides precipitation susceptibility estimates that include precipitation from both warm and mixed-phase clouds. Therefore, the susceptibility estimates that they report include effects from both warm and mixed-phase microphysics. Since the discussion in the paper is focused on warm cloud precipitation, we have not included their results in the discussion.

(ii) Do the authors believe that the "SI" metric (rather than "SR") is more similar to the precipitation susceptibility quantified in previous studies that were cited by the authors (e.g. Sorooshian et al., 2009; Jiang et al., 2010), albeit probably with different minimum rain rate thresholds? If so, it may be worth mentioning this and considering this at least in the comparison with data from previous studies such as the Lu et al. study in the comment above.

Yes, S_I is similar to the precipitation susceptibility studied previously. The analysis carried out for the Appendix (see Fig. A6 and A4 below) appears to show that S_I and the linear regression in log-space have similar values and respond similarly to changes in threshold.

We have inserted the following sentence in the Appendix: S_I, which is more akin to the susceptibilities reported in previous studies, quantifies the effect of aerosols on how intense a cloud precipitates.

In this regard, it is interesting that the authors noted (Pg 33397 Line 3-20) that SI exhibited a maximum at an intermediate LWP value in Figure 7 for the 5 km analysis. Figure 8 also shows a susceptibility maximum at an intermediate LWP value. Was this behavior in SI evident in the 10 km and 20 km analyses?

Our explanation in this section was unclear. Although not statistically robust, we find that S_I actually gradually *increases* with increasing LWP across the three bins. This is a feature that is evident in the 10 km and 20 km analyses. We have attached a figure that shows S_I with increasing cloud LWP. Since the increase in S_I is larger than the decrease in S_f between the second and third LWP bin, a maximum in the S_R appears. To clarify this point, we have added the following sentence after "… third LWP bin (not shown).": S_I also increases from the third to fourth LWP bin, but the decrease in S_f is larger, which leads to the decrease in S_R.

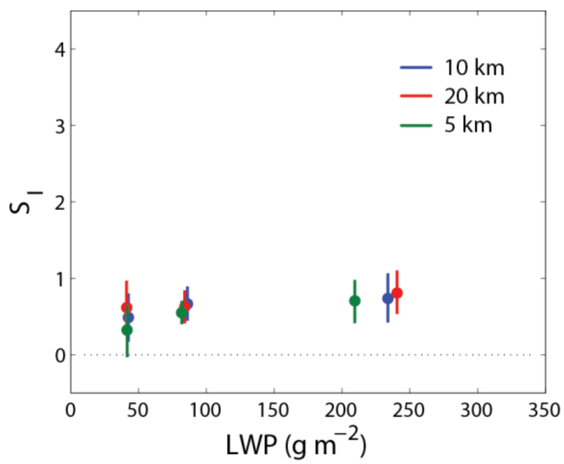

Figure: S_I as a function of LWP. Note: S_I from the first bin is not shown.

This also raises the question as to whether direct comparisons should be made for the LWP-dependent (or H-dependent) behavior of susceptibility for different warm cloud regimes that also may have different ranges in LWP or thickness. I suggest that if the authors have more thoughts about this, it would be worthwhile to add some discussion in the last paragraph of the manuscript where they start to get into these issues. This could be helpful to inform future studies examining precipitation susceptibility.

The LWP-dependent behavior of susceptibility may be different depending on the cloud regime (cumulus clouds vs. stratocumulus clouds), especially given the different types of entrainment the two cloud regimes undergo, the condensate loadings and timescales available for precipitation formation in each case. Therefore, we have expounded the second to last sentence of the paragraph such that it now reads: While the results here support the theoretical arguments of Wood et al. (2009) … and that other factors, such as thermodynamic environment and cloud type, may play a role.

Also, can the authors clarify what is meant by sampling artifacts (Line 18, pg 33397)?

Since the wording was confusing, we have replaced "to be an artifact of the sampling" with: to not be due to a change in microphysical processes.

(iii) Page 33397, line 7: Do the authors intend to say "from the first to the second LWP bin"? This is what Figure 7 indicates to this reviewer.

The susceptibility estimates from the first LWP bin are not shown because of the lack of precipitating data points. Therefore, the first point seen in Fig. 7 corresponds to the second LWP bin. We have clarified this part by adding the following sentence such that it now reads: "… third LWP bin (not shown). Note that S_R from the first bin is not shown for lack of precipitating clouds. To investigate …"

(iv) Page 33399, line 22-23: Can another issue be that different LWP-dependent (or H-dependent) behavior of precipitation susceptibility exists for different cloud types evolving in different meteorological/thermodynamic conditions?

We cannot disregard the possibility that the susceptibility responds with LWP differently in different cloud types. Since entrainment of dry air can play a large role in determining cloud droplet sizes, the thermodynamic conditions in which the clouds form are a big factor. We have expanded the sentence such that it now reads: We attribute the difference to whether or not non-precipitating clouds are included in calculating the susceptibility, though another explanation for the difference may be that the precipitation susceptibility behaves differently with cloud LWP in different thermodynamic environments and cloud regimes.

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 β = 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.

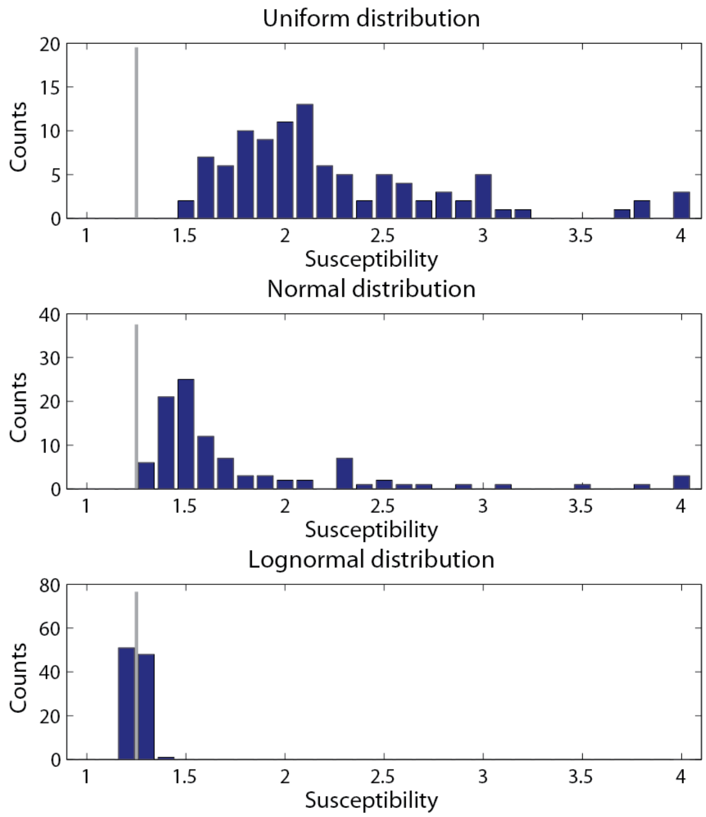

Figure A1: Histogram of susceptibility estimates using TLD for three different distributions of data: uniform (top), normal (middle), and lognormal (bottom). The histograms are based on 100 estimates calculated from 100 different samples of the same underlying distribution. The gray line at 1.25 shows the value of β in the underlying relationship $R = a\,N^{\beta}$.

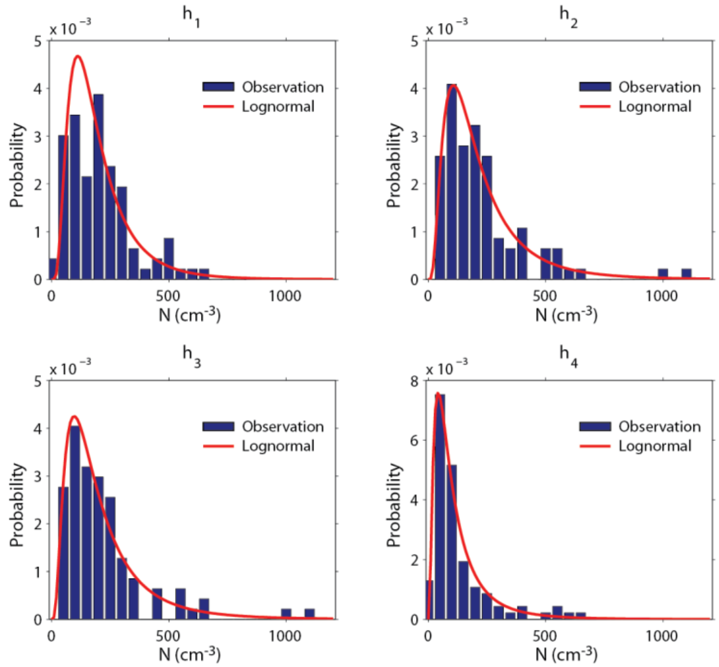

Figure A2: Histogram of 10 km-averaged PCASP aerosol concentrations N that are used in the susceptibility analysis. Each panel shows the distribution of N in the four cloud thickness bins. The red line shows the probability density function of a theoretical lognormal distribution, based on the arithmetic 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.

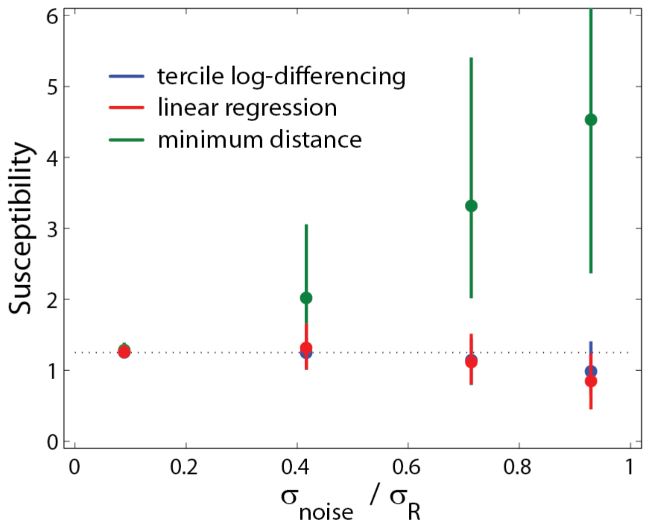

Figure A3: Susceptibility estimates from TLD (blue), linear regression (red), and minimum distance (green) with increasing noise level. Dots represent mean susceptibilities based on 100 estimates, and the lines show middle 95 % interval. The abscissa shows the ratio between the standard deviation of 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.

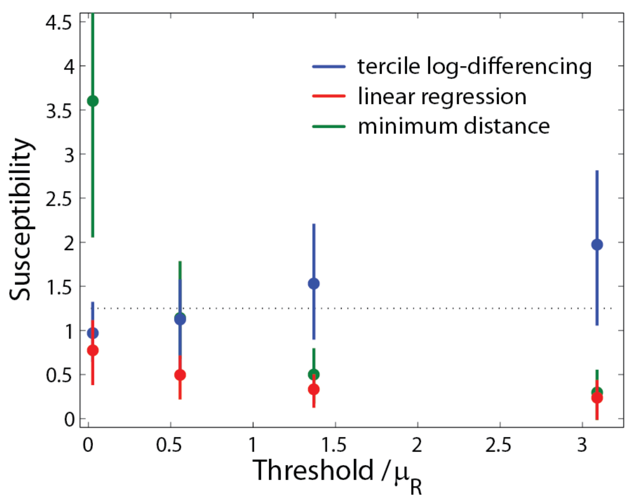

Figure A4: Susceptibility estimates from TLD (blue), linear regression (red), and minimum distance (green) with increasing threshold level. Dots represent mean susceptibility from 100 estimates and the lines show middle 95 % interval. The abscissa shows the ratio between the threshold value and the mean, μ_R, of R after applying the threshold. The dotted line represents the underlying β value.

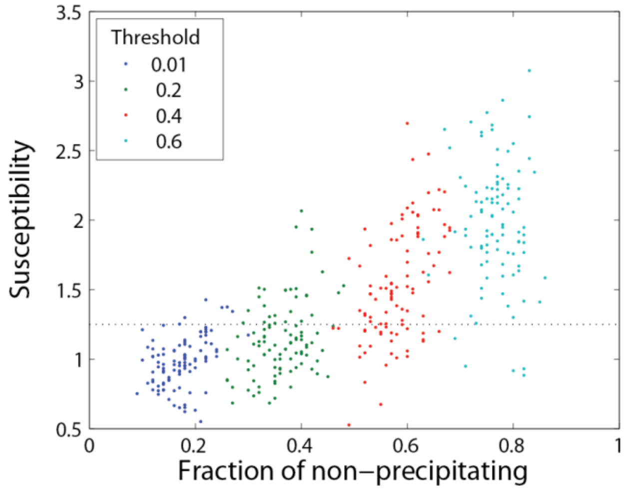

Figure A5: Susceptibility estimates from TLD as a function of the number of non-precipitating points that went into calculating the susceptibility. Colors indicate different levels of the thresholds: 0.01 (blue), 0.2 (green), 0.4 (red), and 0.6 (cyan). The dotted line represents the underlying β value.

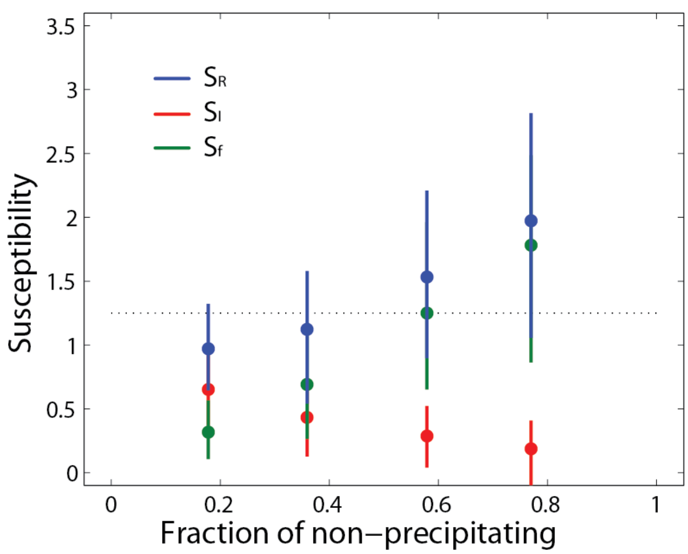

Figure A6: The precipitation susceptibility S_R (blue), susceptibility of drizzle intensity S_I (red), and susceptibility of drizzle fraction S_f (green) with increasing threshold value, calculated using the TLD method. Instead of the ratio between the threshold and mean precipitation rate, the average 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.