Answer to Anonymous referee #1

We would like to thank anonymous referee #1 for the careful reading of the manuscript and the useful comments and recommendations. We accepted all suggestions. Below you will find our replies and short descriptions of the changes we've made in the text. Referee comments are in red and start with **"R:"** and our replies are in black and start with **"A:".** Original manuscript text is shown in **blue**, with new text highlighted in <mark>yellow</mark>.

We would like to warn referee #1, however that we are still working on our data analysis to accommodate all suggestions from both referees and hence the final numbers might still change. Referee #2 urged us to change our definition of cirrus clouds according to (Campbell et al., 2015). More important, however, is the request by Referee #1 to do a full multiple-scattering correction, instead of the approximate correction we have now. The new manuscript will be uploaded as soon as we finish all these changes.

R1: The authors should be less aggressive when speculating on their statistics and reading into differences that are not statistically significant.

A: We agree with the recommendation. We have reviewed our statistical analysis; particularly we have calculated the statistical significance before making strong statements.

R2: Much of the paper is spent discussing season and diurnal differences in the cirrus statistics. However, little attenuation is paid to whether these differences are actually statistically significant. Some figures/tables give the standard deviations, but little discussion of them is given in the text leaving the reader to determine significance themselves. In Table 2, it appears that non of the statistics differ significantly: i.e. one cannot say that the cirrus differ at all from season to season. Similarly in Figure 2, the box plot reveals that there is no significant seasonal cycle in frequency of occurrence either. In other figures where histograms are given, a statistical test should be applied to ensure difference among distributions are statistical significant before they are discussed. It is only appropriate/worthwhile to discuss differences that are statically significant.

A: We agree that it only makes sense to discuss differences that are statistically significant, and we have indeed tried to do that. It seems, however, that the captions and text discussion about figure 2 and table 2 were not well explained, leading the reviewer to misinterpret their content.

Table 2 shows the values of the mean and of the standard deviation of the sample (for individual 5-min observations). When comparing mean values, however, we should use the standard deviation of the mean, which is roughly the standard deviation of the sample divided by square-root(N). As shown in Figure S.1 (supplement), there are about 1500 good profiles per month, or ~6000 per season. However, as multiple layers are considered as different clouds (if spaced by more than 400m), the number of cloud layers detected in any season is about ~10000, hence the standard deviation of the mean values are about 1/100th of the values shown in brackets in table 2. That is why the differences of seasonal mean values are statistically significant for most parameters shown.

We agree, however, that we did not give enough information, and we apologize particularly for not having included the values of N for each column in Table 2. To resolve that issue, we moved the information from Fig S.1 into table 2, and we also included the number of cloud layers detected. We have also included an * (asterisk) to indicate when the difference of mean values between the dry and wet seasons is statistically significant with a 95% confidence level. The caption of Table 2 now reads:

Table 2. Mean cirrus cloud properties and standard deviation <mark>of the sample</mark> (in parenthesis) for all cirrus clouds, for cirrus clouds above and below 14 km and cirrus clouds with base below and top above 14 km <mark>are shown</mark>. <mark>Values are given</mark> to <mark>the total time of observation,</mark> <mark>as well as the wet,</mark> transition and dry seasons. <mark>Statistics are based on 5-min observations. The number of clouds layers detected was used to calculate the standard deviation of the mean, and then determine if the seasonal differences (wet-dry) are statistically significant to the 95% confidence level (indicated as *).</mark>

One important point to note is that table 2 gives statistical properties (mean and std) based on 5-min profiles, and hence the frequency of occurrence is a single number without an associated standard deviation. The box-plot in Figure 2, on the other hand, was calculated from daily averages as was discussed in the text and in the figure caption. For instance, the caption reads: *"..Red dashes (black x) in the boxplots are the median (mean) of the daily frequency in each month…"*. Hence, the box-plot indicates the day-to-day variability in

each month. For example, during March 2012, we never observed any single day with less than 52% or more than 95% frequency of occurrence. In fact, 50% of the daily averages are within the interval 79-89%. As noted before, to compare the mean values (in this case of the daily occurrences) one should use the standard deviation of the mean, not of the sample. And when comparing the whole distribution, one should also note that even if two distributions have the same mean values, it doesn't mean they are the same. In this sense, the boxes in Figure 2 show that Jul and Aug 2011 have the same distribution and mean values, while Jan and Jun 2012 have close medians but different distribution.

However, we agree that the text/captions/figures could be modified to make this point more clear. Hence, in Figure 2, we added the standard deviation of the mean as an error bar centered at the mean daily frequency (cross) and we created a secondary panel to show the precipitation. The caption of Fig. 2 now reads:

Figure 2. The upper panel shows the monthly frequency of occurrence of cirrus clouds from July 2011 to June 2012 (blue line) for all data in each month. The boxplot is for the daily frequencies and hence show the day-to-day variability. Red dash, black cross (x) and error-bar indicate the median, the mean and the standard deviation of the mean, for the daily frequencies in each month, respectively. The edges of the boxes are the 25th and 75th percentiles, and the whiskers extend to the most extreme daily values. Accumulated rainfall is shown in the lower panel based on data from TRMM 3B42 version 7.

The text around lines 207-210 was modified accordingly:

The boxplot in Figure 2a show the variability of the daily frequency of occurrence for each month. There is a high day-to-day variation (i.e. dispersion of the daily frequencies), which is maximum in May-Aug and lowest in Nov-Apr. The mean monthly cirrus cloud frequency follows the same seasonal pattern as the accumulated precipitation (Figure 2b), with values during the wet months higher by a statistical significant amount than those during dry months (notice the small standard deviation of the mean despite the high variability).

R3: I would also caution the authors against extrapolating too much from their relatively limited data. An example of this is using the lidar ratio to infer the ice crystal habit. The lidar ratio alone cannot be used to identify the ice crystal habit since it also depends on the particle orientation relative to the laser beam. In addition, theoretical studies of ice crystal phase functions vary wildly so there is no real consensus on what the lidar ratio even is for different ice crystals.

A: We thank the reviewer for pointing out this issue. After his/her comments, we revised the literature once again and found, as he/she mentioned, no consensus about what the LR should be for different crystal habits. We have hence removed any reference to actual crystal shapes. However, we kept the argument that the bimodal distribution of the LR for some seasons is an indicative of mixture of different shapes, although we can't tell which habit from our limited dataset.

R4: The authors note that this ground-based site is unique compared to others reported in previous work, yet rely heavily on previous work to explain their results. The paper would be greatly enhanced by making a more quantitative effort the explain their data. For example, instead of speculating on the sources of moisture for the cirrus in different seasons, a more convincing approach would be to run back trajectories to show the reader where the air came from.

A: We thank the reviewer for pointing this out. We note, however, that the first intent of our paper was do document the diurnal and seasonal cycles of geometrical and optical properties of cirrus clouds in the central Amazon. This is a region known for having an important link with climate, but also known for the lack of continuous long-term observations. This is the reason why most papers rely on satellite-based observations, which are mostly polar-orbit and hence, have low time resolution. We said that "our site is unique" because it is allowing us, finally, to perform long-term observations with high temporal resolution. The obvious drawback is that we don't have the spatial coverage of the satellites.

Although our initial intention wasn't the identification of the sources of the cirrus clouds, we agree that giving a more quantitative explanation would strengthen our paper. Following the referee's suggestion, we did back-trajectory analysis using Hysplit forced by GDAS winds (1deg resolution), starting every 6h from 14.5 km over the site during the dry season period. Each of the 480 back-trajectories were integrated for 7 days. Figure below shows the result of this analysis. In the top panel, we show the individual trajectories just for the 0:00 of each day and there are so many lines that clutter the plot. The lower panel shows the trajectory density, i.e., the number of trajectories in a point divided by the total number of trajectories (a number [0-1]). In this case we used a log-scale because the density will obviously be much higher closer to the trajectory

start point. The result is quite interesting as it reveals that many trajectories actually don't follow the average wind pattern (fig. 3 in the manuscript, top panel). On the other hand, many trajectories come from Colombia and Venezuela, exactly where precipitation from deep convection is found (also shown in fig. 3, top), and some even reach towards the ITCZ, far to the east. This comparison could be improved if we select only the trajectories starting at times when we detected a cirrus clouds (yet do be done).
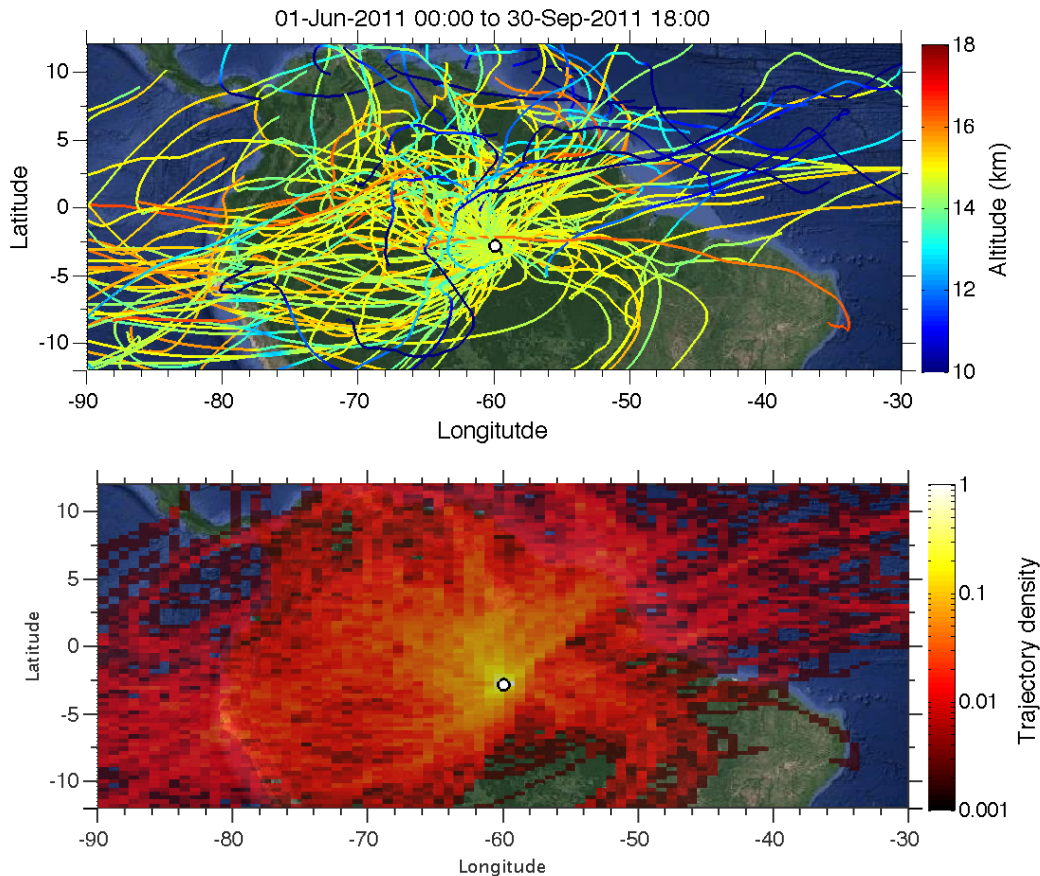

Figure C.1 – Hysplit 7day backward trajectories starting 14.5km above the site every 6h for the four months of the dry season. It should be compared to the top panel of figure 3 in the manuscript.

Although this trajectory analysis (suggested by referee #1) indeed gives further evidence that our cirrus clouds originate from deep convection, it is not a quantitative evidence (as referee #1 wanted). Alternative ways of having a quantitatively analysis would be to run the back trajectories for each cloud layer detected and use GOES images to locate deep convective cells, and then calculate the distance between each trajectories and the surrounding precipitation (as a function of backward time). This is a huge effort and, we believe, deserves its own paper. Another possibility would be to do that, but just for one case study in each season (e.g. as Fourtin et al., 2007JGR). We are not sure, however, how representative and quantitative that could be.

If the reviewer has other suggestions, besides the back-trajectories already shown above, we would be glad to try. Or if the reviewer / editor think the plots above are already enough, we would be happy to include this discussion in the manuscript.

R5: Is there reason the authors don't use the nitrogen signal to retrieve extinction? Not doing so doesn't completely discount the data presented, but it does devalue it somewhat since this paper is just another in a long-line of elastic lidar cirrus studies.

A: Raman inelastic scattering has a cross-section of about $1/1000^{th}$ of the elastic one. For our system, that weak signal is only discernible from the background during nighttime, thus not being appropriate for our study, which wanted to investigate the diurnal cycle. Moreover, the Raman method involves calculating the derivative of the signal, which gives rise to large uncertainties in cases of such low signal-to-noise ratios as typically found at cirrus altitudes. That noise could be reduced by vertical smoothing or by time averaging,

but our clouds are not homogenous in time in space, and the averaging could compromise the results.

We should also mention that our experience with simulated signals have shown us that we can obtain a very precise and accurate LR with Chen's method. We did this analysis of simulated cloudy lidar profiles to evaluate the accuracy of the transmittance method but did not include in the manuscript, or in the supplement (but see discussion following your next question). We believe that these simulations and the fact that Raman cannot be done during daytime justify our method of choice.

Finally, we also noticed that this argument was given only at the results section (lines 398-400). We have thus added the following paragraph in section 2.4, at line 181:

We use Chen et al. (2002) instead of the Raman method (Ansman et al., 2002) because our instrument can only detect the nitrogen Raman channel during nighttime. Moreover, the Raman results are very noisy even during nighttime and by analyzing simulated lidar profiles (supplement material) we found that a precise and accurate cirrus LR can be obtained with Chen's method.

R6: In addition, the transmission method is really only accurate for mid-range optical depths. Too thin and there isn't enough transmission signal to get a reliable optical depth. Too thick and there isn't enough molecular signal above the cloud. I encourage the authors to go beyond just checking the SNR above/below the cloud when doing the optical depth retrieval and to fully derive the uncertainty in the optical depth values they report. Figures 5 and 6 show optical depths down to 0.001, which I expect to be extremely uncertain when using the transmission method to retrieval optical depth.

A: We agree with the reviewer's point of view and, in fact, we have already calculated the uncertainty in all optical depth values that we have obtained. However, the plots and tables shown in the manuscript are always for averages over a huge amount of profiles, and hence we choose to report the standard deviation of the mean instead of the errors in individual retrievals.

We also agree that the COD uncertainty is very large if obtained for a single profile with COD = 0.001. In fact, we have done an extensive simulation study to validate the methods we use, which was not included in the manuscript. For COD = 10^{-3}, the relative error in a single retrieval is 120% for S/N = 50 and 1150% for S/N = 3, both large but not enough to change the cirrus category (e.g. from sub-visual to thin). Moreover, averaging over N profiles reduces this uncertainty by a factor of square-root of N. In our study, we analyzed about 37k 5-min profiles, where 21k had S/N > 3 at 12km and in 14k of these we found a cirrus cloud. Thus, the error in the mean cloud optical depth reported in Table 1, or in the histograms in figures 5 and 6, is indeed much lower, typically below 20% even for S/N = 3.

We were planning to have a separate manuscript on AMT about the accuracy and precision of the transmission method for the retrieval of COD and LR from elastic lidars. However, as both referees have questioned about this, we believe that some of that needs to be included in the supplement material. We will consult the editor to see if he/she agrees with this approach.

R7: The treatment and discussion of multiple scattering could be improved. Although, not explicitly stated, I'm guessing the authors use Eq. (10) from Chen et al. (2002) where eta depends on the optical depth of the cloud layer. I'd would encourage against using this equation. Chen et al. provide no physical justification for this equation and the values for larger optical depths quickly approach the wide angle scattering limit of eta=0.5 which is unrealistic for the geometry of a ground based lidar. In addition, for optical depth greater than about 1.2, eta<0.5 which is unphysical. The authors should also keep in mind that the shorter wavelength of 355nm (compared to 532nm as is used in Chen et al. 2002 and many other studies) means much stronger forward scattering and therefore larger amounts of multiple scattering. Typical extinction biases could range from 5-30% and sometimes even larger (see Thorsen and Fu, JTECH 2015 Fig. 13). I would suggest the authors make clear to the reader that their optical depth may contain significant biases due to multiple scattering unless some type of explicit treatment of multiple scattering is performed.

A: We thank the referee for explaining the limitations of the correction proposed by Chen. To appropriately account for the multiple-scattering, we reviewed the work of Platt (1981) and Wandinger (1998) and finally decided to apply a full treatment following the model of Hogan (2008). Our preliminary results are indicating a change of the LR from ~19 sr to ~25 sr. That means that Chen's correction was indeed not valid for our case! Looking at table 1, where we compared our results with some available in the literature, it seems that

Goldfarb et al. (2001), Pace et al. (2003), Cadet et al. (2003), Hoareau et al. (2013), and Pandit et al. (2015) also did the same mistake as we did.

However, reprocessing all the dataset with this more complex algorithm for multiple-scattering is taking much longer than anticipated. Hence, as we mentioned in the beginning, an updated version of the manuscript will be uploaded only after finishing these calculations.

References:

Hogan, R. J., 2008: Fast lidar and radar multiple-scattering models: Part 1: Quasi-small-angle scattering using the photon variance-covariance method. J. Atmos. Sci., 65, 3621-3635.

Platt, C. M. R., 1981: Remote sounding of high clouds. III: Monte Carlo calculations of multiple scattered lidar returns. J. Atmos. Sci., 38, 156–167

Wandinger, U.: Multiple-scattering influence on extinction- and backscatter-coefficient measurements with Raman and high-spectral-resolution lidars, Appl. Optics, 37, 417–427, 1998.