

Final Response on Referee Comments (RC 1)

"Nature and extent of shallow marine convection in subtropical regions: detection with airborne and space borne lidar-systems over the tropical North Atlantic Ocean"

by Gutleben, M., Groß, S., Wirth, M., Ewald, F. and Schäfler, A.

We want to thank the Referee for carefully reading the manuscript and for the helpful suggestions and comments. The comments and questions will be answered by direct response **(bold)**.

General comments and recommendation

This manuscript compares cloud top height and cloud size distributions in a typical trade-wind cumulus region from measurements by a lidar aboard the HALO aircraft and aboard the CALIPSO satellite. I was very excited to read the manuscript, because it exploits new lidar measurements to look at a type of clouds that is important for understanding climate. I was hoping to read either a discussion on the technical capabilities and limitations of either lidar system at accurately detecting these clouds, or learn something new about the distribution of these clouds when viewed from above - with some of the best instruments around. But the manuscript largely disappointed me.

We realized that we failed to present the main objectives and results of our study. Spaceborne lidar systems were frequently used to derive distributions of cloud macro-physical properties of shallow marine trade wind convection. However, up to now, the applicability of spaceborne lidar measurements regarding the detection of shallow marine cumulus convection by means of direct comparisons to airborne lidar measurements, were not tested. Spaceborne lidar measurements have a resolution of a few hundred meters, compared to our lidar measurements. This may cause differences in observed cloud distributions. We thoroughly reworked the manuscript to better present main objectives and key findings of this study.

What does the manuscript want to achieve? Given that this is submitted to a journal that focuses on a discussion of atmospheric measurement techniques I had expected to read much more details about the actual measurement and analysis methods. But details on why the authors have chosen specific thresholds on backscatter ratios, or how the molecular and total backscatter are calculated, are missing. The authors also do not discuss how the different lidars and their footprints explain measured differences in CTH distributions, even if small. The fact that lidars measure clouds only in the along-track direction, effectively missing information on the 2D dimensions of clouds, and thus possibly leading to a bias in cloud length or gap distribution, is never even mentioned.

We thoroughly reworked the text to provide more information on the analysis methods. Hereby, we focused on direct comparisons between airborne and spaceborne lidar measurements to highlight limitations of spaceborne lidar measurements for the detection of shallow marine cumulus clouds. We do not think that the 2-dimensional nature of our measurements significantly biases our distributions, due to the large amount of collected data and measurements in all geographic directions.

I also did not gain any new physical insight from this manuscript, which largely glances over a rich body of literature that has looked at the distribution of cloud layers from these clouds (from ground-based measurements e.g., most notably the Barbados Cloud Observatory, but also from space-borne sensors e.g., MISR, ASTER), and which does not consider another rich body of literature that has described the evolution of cloud types and cloud depth along trade-wind trajectories. The discussion and summary that touch upon influences of the aerosol on the change in cloud top heights as one moves towards the equator feels misplaced.

In the course of the revision of the manuscript, we also focused on the discussion of our findings and put them in relationship with former findings.

A lack of instrumental details and technical insights on the one hand, and on the other hand, a lack of interpretation and putting the measurements into perspective of existing work, has led to my decision to reject the manuscript. Below I have outlined a few specific points that motivated and might further clarify my decision.

We reworked the manuscript to answer the main points raised by this referee. In particular we better described the instrumentation and method we used, focused on direct WALES and CALIOP comparisons and put our results in perspective to existing work.

Specific Comments

The abstract describes that this paper will test the utilization of CALIOP at observed trade-wind cumuli, yet there is no sentence that specifically states how well CALIOP performs, and which remaining biases in CALIOP data - if any - remain. The abstract ends with two sentences that discuss cloud gaps, and which have no specific meaning when the authors do not explain what their findings on cloud gaps implies for cloud cover, for the heterogeneity of the cloud field, or for something else. These sentences refer to findings that are discussed with as much as two sentences in the main body of the paper. Is that worthy of ending the abstract?

We agree with this referee, that the abstract fails to highlight the main objective and findings of this study. We reworked the abstract following the suggestions of this referee.

The introduction: P2L24: what do the authors understand by systematic evaluation? And in L26-35: How is step 1 of this systematic evaluation different from step 2? Furthermore, what is meant with "small-scale" character (L27)? This could be cloud droplet number concentrations, dimensions of clouds, cloud cover, irregularity of cloud edges, etcetera.

We agree that the terms 'systematic evaluation' and 'small-scale character' can be misleading. We thus avoided using these terms in the revised manuscript.

Section 2.3: P5: How are the total and molecular backscatter defined? I could not repeat the author's analysis from this description alone. How variable is the molecular component, e.g., is it absolute necessary to account for this? If so, what is the order of magnitude with which this varies? How do the authors motivate their BSR thresholds? L31 reads: "BSR values for clouds were found to be much higher". What independent source did the authors use to note that at such values they are dealing with clouds?

For a consistent comparison of the WALES and CALIOP instruments data sets need to be converted into a uniform unit. In principle it would not be necessary to convert the signals into the backscatter ratio (which is the ratio of the total backscatter to the molecular backscatter coefficient). We could also use backscatter coefficients in this study. However, we preferred to use the backscatter ratio as it is much more comprehensive and has convenient values compared to the backscatter coefficient. To calculate the backscatter ratio, the particle/total and the molecular backscatter coefficient have to be known. The particle/total backscatter coefficient depends on amount and type of observed aerosol in the atmosphere. The profile of the molecular backscatter coefficient can be calculated analog Rayleigh scattering theory from known temperature and pressure profiles. The molecular backscatter coefficient varies with height as it depends on temperature and pressure profiles, but it should not show significant variation in a height range or a limited altitude range within the same meteorological regime. Taking this into account it would not be necessary to use the backscatter ratio; however, due to its more comprehensive nature we still prefer to use it in this study. For this work we used a threshold of 90. This is an empirical value determined when examining our data. An aerosol free and cloud free atmosphere has a backscatter ratio of 1 and thus aerosols and clouds can be distinguished quite easily from aerosol/cloud free regions. The difficult part in this study is to distinguish cloud signatures from aerosol signatures; especially as the clouds we are interested in, form right on top the aerosol loaded marine boundary layer. We found BSR values between ~5 and ~10, when investigating the backscatter ratio of the aerosol loaded boundary layer. Hygroscopic growth of the particles can lead to even higher values. We found significantly larger values (>80) when clouds were detected. To clearly distinguish clouds from swollen aerosols we defined a threshold of 90. We did a sensitivity study and used

a wide range of threshold values. However, we found no significant changes in the retrieved distributions. Thus, we concluded that the result is less sensitive to the chosen threshold except when choosing thresholds similar to the backscatter ratio found for aerosols (20-30). We reworked the manuscript to make this clearer.

Section 2.4: P6: How relevant is taking into account the spherical nature of the Earth when most clouds tend to be smaller than 1 km? Or, how many long clouds did you detect that made you choose to take the sphericity into account, and what error would you have made if not doing so?

As our main focus is on clouds of small horizontal extent, neglecting the spherical nature of the Earth would have just little consequences. However, we see no point why we should change that as we already considered it in our calculation.

Section 3.1: P6 and the satellite image in Figure 3. L25 read: "small irregularly scattered clouds dominate the area over Barbados and the Atlantic Ocean". When I look at the satellite image, I actually observe a large number of very large clouds clustered together in specific areas, and I see very few small cumuli. One should either zoom in to see the clouds you are talking about, or change the interpretation of this image.

As requested by this reviewer we changed the satellite image to better account for the measurement situation.

Section 3.1: P8L14: the stratiform cloud in Figure 4 does appear to have some gaps in between – is this really 125 km and counted as one cloud?

The phrase 'stratiform-like cloud structure' in the manuscript is misleading. Thus we removed this term and refer to an elevated cloud regime at about 2.5 km altitude.

Section 3.2: Can you put your findings into perspective of recent studies looking at similar statistics but using ground-based lidars at the Barbados Cloud Observatory? What are the downsides of flying at high altitudes and looking down, compared to being on the ground and looking up? What are you missing?

We extended the manuscript to put our work in perspective to the findings of existing studies also including ground-based instruments. Regarding the questions of the different advantages of airborne measurements to ground-based measurements, the potential to directly derive cloud length and cloud gap length is certainly on the side of the airborne measurements as ground-based measurements have to make assumptions on wind direction and speed when estimating cloud and cloud gap lengths. Another advantage is the possibility to

perform satellite underflights to directly compare the results of the different platforms - which is the main objective of this paper. As the lidar can rarely penetrate water clouds, the derived properties of ground-based and spaceborne lidar measurements are furthermore not directly comparable. However, of course the number of flights and measurement periods of airborne campaigns are limited, especially when compared to long-term measurements at the Barbados Cloud Observatory.

Section 3.3 and Figure 6: There are many studies that have looked at the change in cloud properties along trade-wind trajectories and which have done very detailed work on trying to decipher what controls cloud break-up and cloud deepening when moving closer to the Equator. Think of all the studies that discuss stratocumulus to cumulus transitions, and the GPCP transition cases.

We modified the manuscript with the intention to put our work in perspective to existing studies. However, our main focus was on the verification that CALIOP measurements (which were used in a number of these studies) are capable to study the shallow marine trade-wind clouds.

Section 3.4 and Figure 7: This figure clearly shows that (unlike suggested in preceding text) that when cloud top heights are detected near 2 - 2.5 km that they are as often associated with individual small but deep clouds as they are associated with extended stratiform-like layers.

There was an error in the labeling of this plot which leads to the misinterpretation. The plot is replaced in the revised manuscript.

Discussion and Summary: Any differences you find between different periods are of course related to changing synoptic conditions and changing seasons, hence, those should no longer be just assumptions. Several other studies using both ground-based and satellite data have demonstrated a seasonality in trade-wind cumuli in the region, as well as explained what meteorological factors (and not the aerosol) are responsible.

We totally agree that the changing synoptic conditions are the main drivers in the seasonality in cloud distribution. We removed these misleading sentences.