Dear Reviewer 3,

this is an update to our earlier reply to your full review during the quick access phase. Thank you again for your review. We now included a new analysis as reaction to your and other reviewers requests: A synthetic cloud test field from a cloud resolving model and a simulation of measurements with the 3D radiative transfer code demonstrate how O2A derived distances could be “calibrated” for certain cloud types as long as the type of cloud geometry expected can be provided by cloud modelling. The results largely corroborate our earlier conclusions.

Please find below our replies to open points of your review.

Best regards,
Tobias Zinner (and co-authors)

Reply to reviewer 3:

Reviewer comments are highlighted in gray.

General comments

Your open point ...

The paper in its present form describes a technique that currently does not work, and its results are artificially forced to agree with stereo dataset for "validation". This is certainly is not worth publishing. In my opinion the only way to save the paper is to follow the authors' own suggestion (last paragraph of Conclusions Sec.) and incorporate stereo measurements into A-band algorithm as an occasional "calibration" source. (Note that stereo measurements may not be used for validation then.) Only after this is successfully done (which is a "major change") paper can be accepted for publication.

We already do suggest the combination of both data sources for our purpose and discuss the limitations. That means, the method works for our setup of instrumentation to close the gap regarding cloud distance measurements for the campaign in question. Now we added a demonstration how stereo data could be replaced by a statistically generated set of Monte Carlo simulations for modeled cloud geometry with given typical computational capabilities. In addition, we think that the manuscript lays out the way for the community to minimize this approach’s remaining uncertainties using future increased computational capabilities.

Our main goal is a determination of cloud surface orientation and cloud points’ vertical height for our cloud side view for a specific campaign data set (ACRIDICON-CHUVA). For a plausibility check we compare our data to stereo-points. We find an offset (3.8 km) which is mainly caused by 3D effects. Apart from this deviation, errors are small and within the expected and described ranges. We now demonstrate that the offset found lies in the range of offsets caused solely by the typical deviation of 3D cloud surface orientations.

On purpose, we did not use the word “validation” in the manuscript for this comparison of stereo and oxygen-A derived distance, because we do not consider the stereo distances to be free of uncertainty and even systematic effects. Candidates there are imperfect orientation information of the camera, window distortion effects or the bias of the stereo method towards the selection of specific high contrast points (often related to shadows and strong distance gradients).

Open points:

p.11, Fig.5 and ll.3-17:
Important: the absolute accuracies (in meters, not %) of distance measurement must be presented instead of relative ones. Measured distance is only an intermediate step towards derivation of cloud shape and geographic location, which are independent of the sensor position. Thus, only the absolute values of uncertainties in derived distance have physical meaning. The same, say, 600 m accuracy of cloud location determination (the only thing that matters) can be 1% if the distance between cloud and sensor is 60 km or 10% if this distance is 6 km, thus, speculation that 1% is better than 10% does not make any sense.

After discussion, we still think the use of relative values shows differences more nicely and is more intuitive to understand. Using absolute values would basically only show an increasing stddev with increasing distance. In comparison literature usually both are provided.

p.13, Fig.6(c,d):
It is difficult to see comparison between A-band and stereo distances/heights since the former are shown by color, while the latter - by numbers. I suggest to include (in addition to this figure) several single-scan plots showing 2D cross-section of cloud surface derived from A-band measurements with stereo points plotted over. This could also show how/if stereo points can be used for "calibration" of A-band retrievals on single scan or single cloud scene bases.

I still stick to my first reply … You mean a 2D cross section in x-z plane? x being the horizontal direction along line-of-sight, across-track. That would mean to show one x-z cross-section through the oxygen-A derived surface as a line for each stereo value (sometimes two), because the surface changes quickly in y (flight) direction. We are not sure, if another figure improves this comparison much. A third chance to directly compare them is given in Fig 7 where the turquoise points show the values again (now mentioned in the text).

p.14, Fig.7:
Make both plots square (since x and y have the same physical meaning and scale) and set them side by side. Provide means and standard deviations of (y-x).

Done now. Numbers provided.

p.15, ll.18-19; p.16, ll.17-18:
"In general, such a "calibration" of the method could also be reached using a limited number of synthetic cloud model based simulations." Here the authors try to downplay the complexity of 3D RT simulations contradicting their own words in the first paragraph of Sec. 2.2.2. I do not believe in "universal" 3D-RT LUT based on a few of cloud scenes.

I tried to clarify and explain already for the discussion version. Now we even added the new analysis of a single cloud scene to demonstrate the mentioned possibility.