Review of

Total Column Optical Depths Retrieved from CALIPSO Lidar Ocean Surface Backscatter

Robert A. Ryan, Mark A. Vaughan, Sharon D. Rodier, Jason L. Tackett, John A. Reagan, Richard A. Ferrare, Johnathan W. Hair, Brian J. Getzewich

General comments:

This work basically documents an algorithm that is now being used to produce aerosol column optical depth over oceans from CALIPSO surface returns and MERRA-2 10m wind speed. The first part that describes the method is really more suited as an ATBD (Algorithm Theoretical Basis Document) than a research paper. In general, it is a well written and thought-out paper. A strength is that they present comparisons with MODIS, HSRL, SODA and CALISPO layer optical depths. Also important is their treatment of the CALIOP surface return, realizing that it can be saturated and using a fitting method to the instrument impulse response function to retrieve a more accurate measurement of the surface return magnitude. The filtering of suspect data based on surface signal magnitude, depolarization and wind speed is well done.

However, I think the paper is too long. At around 50 pages it becomes a very tedious read. I would suggest breaking the paper into two parts. The first would present the ODCOD algorithm and the uncertainty analysis and the second part would present results and comparisons. This would allow a better and more in-depth description of the surface return fitting to the IRF (which I think at present is confusing) and the results paper to include more examples and comparisons. A comparison that should be added are some island-based or costal AERONET comparisons of column optical depth.

A general question is why don't you use AMSR wind speeds exclusively and drop MERRA-2 all together? Why use AMSR as a correction to MERRA?

It would add value to include a figure showing the number of observations on a map. So, a figure like figure 11 but only showing number of observations. There must be a low number of observations in very cloudy areas such as 50-60S and the North Atlantic.

Specific comments:

Abstract: In the opening paragraph it should be stated that this paper presents a documentation of an algorithm that now produces column optical depth in the latest version of the CALIPSO data products and state what that version is.

Line 55 -60: Here you should include other pertinent work such as He et al., 2016 and others not cited.

Line 75: Isn't it the ratio of the measured and modeled return, not the difference?

General: How do you ensure that you are getting the bins associated with the surface return? In situations where you have very low clouds or fog and enough attenuation in the troposphere above, these low cloud layers can masquerade as a surface return. Is the surface return found in L2 processing and you accept that as truth?

Also what about multiple scattering which can be large for low clouds? This will increase the surface return magnitude and result in a lower than actual column optical depth. I do not see a correction for multiple scattering. Here you are filtering out clouds, but they are included in the ODCOD product. In general, multiple scattering is low for low optical depth layers like aerosol, but may be important for some optically thick aerosol layers such as Saharan dust over the ocean.

Line 110: I think you need a sentence like this after line 110: "The combined Htan and Gausian fit to the lab measured IRF points is here after called the instrument response model (IRM).

Lines 165 - 170: This discussion is not clear. The downlinked samples I assume are the lidar profile bins. You say: "The sample time delay is the temporal offset between the actual ocean surface signal onset and the midpoint of the CALIOP onboard or downlinked sample range bin." How can this be calculated? The actual surface return will occur in one or more bins. You do not know where in that 30 m bin the surface return occurred. Also what is an onboard average IRM? I thought the IRM is the fit to the lab measured instrument response model? This is not clear either.

I think what is not clear is the concept of the DIRM – a downlinked IRM? Where does it come from? I don't think you've explained the DIRM so that the reader can understand it.

Your terminology is also confusing. You use "downlinked samples", "onboard samples" and "test samples" What is the difference? Please explain what they are more clearly.

Line 200: How is the surface return "base" determined? Is that part of L2 processing? Why do you limit it to 4 bins? Why not 3 or 5?

Line 203: You write: "If a downlinked sample's time delay places it before the onset of the IRM (i.e. a time delay less than zero), that sample may still be part of the surface return as long as one of the averaged onboard samples is part of the surface." Your wording here is confusing. You reference "downlinked sample" and "average onboard samples". What is the difference between the two? And how do you determine that one of the averaged onboard samples is part of the surface?

Line 245: What about highly attenuated surface returns over a rough sea surface? Would not the below surface scattering be significant in that case? What about clear water less than 60 m deep? In that case you would have a second surface return.

Line 248: Why do you say this when you require the 10 m windspeed in order for your algorithm to work? Please add "and the MERRA-2 windspeed" to the sentence.

Line 285: Change "discontinuities will make the uncertainty estimate less certain" to discontinuities will make the uncertainty estimate less accurate.

Line 487: How did you determine these thresholds: SIAB > 0.0413 sr-1 in the daytime and > 0.0353 sr-1 in the nighttime? What happens if you increase them some?

495: The word “retrivals” is misspelled.

Figure 7: The discussion of this figure is somewhat confusing. You say the upper panel shows the distribution of valid ODCOD retrievals, but yet some of the points are above the limits (red line) you choose as a cutoff. Likewise, in the bottom panel, there are points below the cutoff (red line) and greater than 3 m/s that re not included. Why?

Line 505: What is “MAD”? Please explicitly say what it is.

Figures 7 and 8. You are calling these the distribution of valid ODCOD retrievals. Why then are there points below 3 m/s when you say that you are only using data between 3 and 15 m/s?

Figures 10 and 11. Please indicate the resolution of the grid used to make these figures. For instance is it a 1x1 degree grid?

Line 645-650: It is true that more clouds will inadvertently be included in the daytime measurements but this would tend to make the daytime OD greater than nighttime. So why is night OD larger than day as shown in figure 12?

Lines 669-685. I agree with your statement that averaging the profiles before attempting the retrieval helps in the IRM fitting process, but it is a little confusing when you say you average the profiles that did not have a surface detection. What are you averaging when there is no surface detection? Is the surface signal 0 in the averaging?
Also in this discussion, it is confusing that you are talking about profiles where the surface was not detected since I thought that you were only analyzing cloud-free profiles and those profiles should always detect the surface as the column optical depth for those profiles should always be less than ~1-2 and a surface return would always be present.