The paper presents comparisons of CALIOP AOD retrievals with NAAPS model results after assimilation of MODIS and MISR AOD data, with focus on night-time observations. Global as well as regional datasets for the entire year of 2007 are discussed. Furthermore, comparisons with ground-based lidar observations in the Caribbean are shown. The authors find significant discrepancies between measured and modeled AOD values as well as between daytime and night-time CALIOP retrievals. The paper is based on a careful data analysis and thus is a valuable contribution to the CALIOP validation efforts. However, the interpretation and discussion of the findings could be significantly improved if the authors went more into the background of CALIOP extinction retrievals, namely the aerosol-typing issue and its consequences. Here, it would also be useful to relate the findings better to other validation and comparison studies, e.g., with ground-based observations at different sites, available in the literature.

MAJOR COMMENTS

1) The authors compare CALIOP and NAAPS data based on a differentiation of surfaces (ocean, coastal, land) on a global scale and on the selection of specific regions. This approach is probably driven by procedures for passive remote sensing where the surface plays a dominating role in the retrievals. However, as the authors correctly state, CALIOP extinction and AOD retrievals strongly depend on the proper selection of the lidar ratio and thus on the aerosol typing. A respective discussion is missing almost throughout the paper. Only at the end of the manuscript consequences of the aerosol-typing approach are tackled in the specific comparisons with ground-based measurements in the Caribbean (see Fig. 10). To my feeling, a much better reasoning of the obtained discrepancies could be given by rigorous investigation of differences in dependence on aerosol type used in the CALIOP retrievals in conjunction with a critical discussion of the selected lidar ratios (see also comments below).

2) The authors should put their study in better context to other validation efforts for CALIPSO. In particular, ground-based observations are more or less ignored at the moment. The reference list is dominated by self-citations. Considering the facts that CALIPSO is a joint effort of NASA and CNES and the manuscript is submitted to a European journal, it is somewhat disconcerting to see that only two out of more than 60 citations refer to European research efforts.

SPECIFIC COMMENTS

1) Pages 2753-2755: The description of CALIOP data evaluation is difficult to understand and should be improved. In particular, the meaning of the QC flags and of the Extinction_Coefficient_Uncertainty needs better explanation. Some questions in this context:
Page 2753: What is r^*?
Page 2754, lines 9 and 10: What means 99 without unit?
Page 2754, line 11: …*of all 2010 nighttime*… nighttime what?

2) Page 2755, discussion on attenuation:
…limitation for nadir-aligned profiling… Why specifically for nadir?
The approach of omitting profiles when no aerosol is detected close to the surface because of signal attenuation (cut of optically very thick layers) leads to a bias in the dataset, doesn't it? The consequences, e.g., for strong aerosol sources (deserts, fires) should be discussed.

3) Page 2759, line 3: *...where each bin value is treated as a single normalized point…*What does it mean?

4) Page 2759, line 12: *...AOD threshold…* threshold for what?

5) Page 2759, lines 22-23: *...the west African coast, including the Caribbean, Brazil…* is a geographically strange description.

6) Page 2759, line 26 ff.: *...Saharan desert dusts, where a-priori assignment of the lidar ratio for extinction retrieval is relatively stable…* This statement is contradictory to the ongoing discussion on the correctness of the lidar-ratio assignment for desert dust (see, e.g., Schuster et al., Atmos. Chem. Phys. Discuss., 12, 11641–11697, 2012). The a-priori lidar ratio of dust used in the CALIPSO retrievals is considerably smaller than values measured in pure Saharan dust from ground (see, e.g., Tesche et al., Tellus B, 61, 144-164 2009; Wandinger et al., Geophys. Res. Lett., 37, L10801, 2011; Tesche et al., Tellus B, 63, 649-676, 2011). Thus an underestimation of the global dust AOD from CALIOP is very likely. This could be the reason for the discrepancies obtained along the dust belt (Sahara to Gobi) in Fig. 4.

7) Page 2759, line 29: *...biased high…* against what?

8) Section 3.1 in general: Most of the discussion of the discrepancies between CALIOP and NAAPS AOD in this section is driven by surface aspects. As mentioned above, CALIOP AOD determination strongly depends on aerosol typing and the respective a-priori selection of the lidar ratio. A discussion of this aspect would be very helpful here.

9) Section 3.2: How and why are these specific regions selected? All of them seem to be somehow influenced by marine conditions. Why are pure continental sites missing? Again, in view of the interpretation of data with respect to aerosol types, it would be very helpful to select regions with relatively "pure" aerosol types as used in the CALIOP retrievals and perform the respective comparisons. Also in Sec. 3.2, the discussion should not only focus on the surface aspects.

10) Section 3.3: Only here, when comparing with ground-based instruments, discussion on the aerosol-type-dependent lidar-ratio assumption is started. However, a clear reasoning for the obtained discrepancies is missing. A discussion on how "polluted dust" is identified from CALIOP level 1 data, what it means for the lidar-ratio selection and thus the extinction calculation, and why this approach fails when dust is mixed into the marine PBL, as it is obviously the case in the Caribbean, would pave the way for some interesting conclusions here.

OTHER CORRECTIONS

Page 2753, line 27: *...avalue…*
Page 2757, line24: *...at over land…*
Page 2761, line 2: *...is found the tropical…*
Page 2765, line 29: *...xceeding…*
Page 2766, sentence starting at line 12 is not complete