We thank Reviewer #2 for his/her analysis and comments on the paper. The responses to major and minor comments are given below. We marked the reviewer's and the author's comments by “RC:” and “AC:”, respectively.

**Major comments**

RC: These findings are quite valuable to understand how to interpret both data sets and also valuable to construct longer time records than those obtained by lidar on a single satellite. There is a lack of clarifications in the current form of manuscript.

AC: We thank the Reviewer for pointing out the importance of the work for merging the different spaceborne datasets into one long-term record. As for the clarifications, we have added the definition of the Scattering Ratio, the necessary formalism to convert the scattering ratio from 532 to 355nm and the definition of the different variables (Sect. 3). We have also updated the figures and the corresponding text, and we have addressed all the comments of all the reviewers.

RC: Theoretical justification of using the simple SR conversion factor method between 355nm and 532nm in Equation (1) is not sufficient.

AC: We agree with this statement. We have added a section with all necessary definitions and conversion formulae. This section also appears to be helpful in the discussion of the potential sources of discrepancy between CALIPSO and ALADIN. The collocated dataset has been reprocessed and the conversion has been re-calculated and analyzed.

RC: When model outputs are used, there is no need to rely on the conversion factor and the SR for 355nm and 532nm/1064nm can be estimated independently.

AC: This is true, but we do not used this conversion factor for the model+simulator part. We have re-written the simulation section and we added a flowchart to clarify the steps of this simulation experiment.

RC: The choice in Equation (1) seems to be essential to the theoretical derived value (0.81) for cloud detection agreement between CALIOP and ALADIN. That is, the treatment
of model output as well as cloud detection algorithm affect the estimation of the value of 0.81.

AC: Please, see the answer to the previous question. The theoretically estimate of the best achievable normalized cloud detection agreement (= value of 0.81, refined in this version) does not use Eq. 1. As we show in Fig. 4 of the new version of the manuscript, the value is mostly determined by difference in observation geometry and orbital parameters leading to non-ideal collocation.

RC: There are no descriptions about the output parameters for EAMv1 model used in this article.

AC: The outputs of the EAMv1 model are the usual standard inputs for COSP/lidar (e.g. Chepfer et al. 2008; Tang et al. 2019). But, we added several modifications to a standard model+COSP/lidar simulation for this study. Those are presented in the flowchart (Fig. 3) and described in Section 4: (a) subscale horizontal cloud variability; (b) instrumental noises for ALADIN and CALIOP; (c) diurnal variation of cloud fraction.

RC: The actual signals in the CALIOP and ALADIN contain the aerosols as well as clouds and molecules. Aerosol signals at 355nm might be larger than those at 532nm and it is naturally expected that the discrimination between clouds and aerosols is more challenging at ALADIN compared with CALIOP.

AC: First, we did not try to build the cloud detection scheme based on ALADIN-defined SR (see Eq. 2 in new version). As for the CALIOP-like defined SRs (new Fig. 5), the SRs from CALIOP are equal or larger than those estimated from ALADIN, so the cloud-aerosol discrimination problem mentioned in the question is not revealed.

RC: It is not clear how to incorporate the wavelength dependence of aerosols into the equation (1). It is not clear whether aerosols are contained in the EAMv1 model or not. There is no description about how multiple scattering effects for CALIOP and ALADIN are treated in the simulations in section 2.

AC: Again, the simulation experiment does not use Eq. 1. We apologize for a lack of clarity in the previous version of the manuscript regarding the simulations and we hope the new Sect. 4 is helpful. However, the question about multiple scattering is relevant and it is included into the present version of the manuscript in its new theoretical part (Sect. 2) as well as in the discussion of possible reasons for the discrepancy of low-level clouds.

RC: It seems to be possible to apply practically the same cloud detection algorithms used in the ALADIN L2A as well as CALIOP GOCCP products in the theoretical analyses in section 2. If one will do so, it would give a different cloud detection agreement of 0.77. The above-mentioned information is important to interpret the results in section 3 and conclusions.

AC: Since we did not convert the SRs for the simulation study (but only for the actual observations), we actually apply the same detection algorithms to the ALADIN an CALIPSO theoretical analyses. We agree that it was not well described in the previous of the manuscript, we hope the new Sect. 4 and the flowchart help.

RC: There are also lack of clarifications in the treatment of CALIOP clouds for the comparisons. It seems there is no sub-grid scale treatment for 87km-ALADIN L2A products so that 0 or 1 cloud fraction for each 87km-grid.

AC: First, the sub-grid treatment of ALADIN is a part of a Prototype v_3.10 algorithm from ESA, which is not available for the end user. The current end-user ALADIN dataset
contains the backscatter and extinction profiles at 355\,nm that are standard for an HSRLidar (but not for non-HSRL like CALIOP). There’s no 0 or 1 in this ALADIN dataset nor does it define the cloud fraction itself. Therefore, we performed a conversion from ALADIN’s backscatter and extinction at 355 to SR’\_532 and apply the uniformly defined cloud detection threshold on this SR’\_532 profile (see Section 2 in the updated version of the manuscript). Second, we used high-resolution CALIOP data on 333m grid, averaged its AMB(z) and ATB(z) profiles at the same vertical and horizontal resolution as ALADIN and calculated SR\_532(z). These procedures ensure that the two averaged profiles (SR’\_532 derived from ALADIN and SR\_532 derived from CALIOP) are comparable.

RC: On the other hands, CALIOP product has finer resolution (333m or 1km). It is not clear how to treat cloud fraction for CALIOP after 67km averaging for the comparisons compared with ALADIN in sections 2 and 3.

AC: We do not use the existing cloud fraction from CALIOP. As mentioned above, we averaged ATB and AMB\,(\,=\,ATBmol\,) over similar resolution as ALADIN and only then do compute SR and apply the cloud detection threshold. We are well aware of the fact that this might lead to an overestimation of cloud fraction in the boundary layer, but we perform this procedure to ensure the comparability of two datasets.

RC: Brief description of Aeolus L2A cloud product is also instructive.

AC: Such a product doesn’t exist (yet), we defined the cloudy or non-cloudy bins by applying the cloud detection threshold to SR\_532(z) values.

RC: The SR for CALIOP was originally estimated to create CALIPOSO GOCCP products where Equation (1) is not needed. It is not convincing why equation (1) is used to simulate SR at 532nm.

AC: In the present version of the manuscript, we do not use Eq. 1 anymore. Instead, we use a more precise recalculation approach presented in Section 3. But, the idea of converting ALADIN’s 355 data to 532nm was to compare apples to apples and apply the same cloud detection threshold to the ‘same’ SR profile at the same spatial resolution.

RC: After reading the manuscript several times, any reasonable explanation was not found why the upper clouds are smaller for ALADIN compared with CALIOP, though CALIOP did not detect most of PSCs where ALADIN detected (in shown in the Figure 4a and f).

AC: Actually, we discussed PSC detection in lines 230-231, 301-303, and 374-376 of the previous version of the manuscript, but in the rest of the manuscript there was a confusing explanation regarding the particulate backscatter and we apologize for this. As we wrote in response to the Reviewer #1’s question, we meant the detection of the particles. Even though the total backscatter is larger at 355nm, the particulate part can be buried in molecular return. If the signal-to-noise ratio is small, then the cross-talk correction (used in High Spectral Resolution lidar) will be noisy and the particulate signal will be retrieved with large uncertainty. We do not know the details of the L2 algorithm computing SR, extinction and backscatter used in ALADIN products, but a common sense tells us that if the signal is noisy then there’s a high chance that the algorithm will reject it. Summarizing, our explanation of smaller ALADIN’s sensitivity to high clouds is linked with a combination of weaker-than-planned SNR and smaller particulate backscatter compared to molecular one.

RC: The authors attributed the lower sensitivity of high clouds for ALADIN to smaller backscatter at 355nm without conducting further analysis.

AC: Please, see the previous answer for the corrected explanation. The text of the
manuscript has been also updated to avoid misunderstanding.

RC: More discussion of the discrepancies in the cloud detections are requested. It is also noted that it is well established that CALIOP has a good capability to detect PSCs so that Figure 4a is strange.

AC: Please, check the new version of Fig. 4 (now Fig. 5) where we show the SRs starting from SR=3. In Fig. 5, one can also see the PSCs detected by CALIOP with SR>5. Note that this threshold is not optimized for PSC that can be optically thin. And, last, but not least, Fig. 8a does contain the PSCs, but their frequency of occurrence is low.

RC: There are several CALIOP based global cloud products, including NASA Langley’s VFM products, GOCCP, DARDAR and KU cloud products and large differences were reported in (Cesana et al., 2016) JGR among GOCCP, NASA standard and KU products, indicating the different cloud detection methods caused the differences. There are several ways to bridge gaps between CALIOP and AEOLUS. Some comments are needed in this regard.

AC: The works mentioned by the reviewer are all using the same source that is L1 collected by CALIPSO. For comparing ALADIN and CALIPSO, the main challenges are because of the difference of nature of their L1 data: (1) ALADIN measures APB and AMB (and not ATB) because it is an HSRL, while CALIPSO measures ATB (and not APB and AMB) because it is a non-HSRL (See Eqs. in Sect. 3), (2) the wavelengths are different (355 nm vs 532nm), (3) the orbits and overpass times are different (see Sect. 2). We tried to state these points more clearly in the new version of the manuscript.

Specific comments

RC: p.6 line 182-184, need clarification for the methods and typical values of noises for Aeolus and CALIOP in the target data sets.

AC: We have updated the methodological part (see new Section 3). As for the noise values, we estimated them from the upper part of the vertical profiles, which are cloud-free and contain only molecular return, which is supposed to be smooth. We added this information to the manuscript (Section 4.1)

RC: p.25 Figure 6, zonal mean cloud frequency for CALIOP and ALADIN would be preferable prior to Figures 6a-d.

AC: Thank you for this suggestion. We added the requested figure and the corresponding text. It is interesting to note that visually the cloud distributions for the compared instruments are much more alike than the SR distributions. But, cloud detection threshold for higher clouds is crossed reached less frequently for ALADIN than for CALIOP.