First reply to RC2 comment of Anonymous Referee #2 on the ACPD paper "Satellite observations of aerosols and clouds over South China from 2006 to 2015: analysis of changes and possible interactions" by Nikos Benas et al. (Referee comments are repeated in italics).

General statement:

The authors investigated seasonal and decadal variations of aerosols and clouds over South China using several satellite observation data and GFED biomass burning emissions to understand aerosol cloud interactions and aerosol semi-direct effect. The methodology of combined use of passive and active satellite sensors is useful, but some discussions are not enough clear to understand aerosols cloud interactions.

Reply:
We thank the Referee for reviewing our manuscript. Following are our replies to the general and specific comments:

General comments:

1. The authors used the CALIPSO aerosol classification product to examine atmospheric aerosol composition over South China; however, misclassification of aerosol type in CALIPSO product often occurs. Burton et al. (2013) indicated that 78% of the smoke layers of the CALIPSO product are inferred by the airborne high spectral resolution lidar (HSRL) to be urban (polluted continental) aerosol.

Reply:
The primary purpose of using CALIPSO aerosol data in our analysis was the inclusion of information on the vertical distribution of aerosols, which is crucial in assessing the possibility of aerosol-cloud interactions. The importance of this information is highlighted in Section 3.3.3 specifically for the different semi-direct effect manifestations, which occur with absorbing aerosols present. Regarding the sources of these aerosol types, we infer some possibilities based on comparisons of their changes with corresponding changes in GFED emissions (page 4, lines 39-43 and page 5, lines 1-3). These conclusions, however, are not definitive, and we plan to rephrase these parts accordingly. The comment of the Referee also points to the same ambiguities in these conclusions, since the mentioned misclassification suggests differences in aerosol origins and compositions. It is a crucial addition to the relevant discussion and will be included in the revised manuscript. It should also be noted, however, that the misclassifications of smoke and polluted dust aerosols reported in Burton et al. (2013) most of the times replace an absorbing aerosol type with another absorbing type. While this is crucial when trying to pinpoint the origin of these aerosols, it would not affect significantly the discussion in Section 3.3.3, on how an absorbing aerosol layer would interact with clouds based on their relative position.

2. The accuracy of data products and the uncertainties caused by the different data sampling derived from the different sensors are not discussed in the manuscript. More detailed description about data quality is needed.

Reply:

In order to determine the accuracy of the data products used, validation against reference measurements is required. While this would be beyond the scope of this study, we agree that a relevant discussion is missing, and we plan to refer to relevant studies in the revised manuscript. Uncertainties caused by the different data sampling from the different sensors can also be an important issue. The level of agreement in cloud properties from MODIS and CLARA-A2 (Figs. 4, 5 and 7) suggests that the respective sensors, overpass times, and retrieval algorithms are indeed similar enough that possible discrepancies due to different sampling are small. This is not the case, however, when comparing CALIPSO with passive sensors, as we briefly mention in page 4, lines 8-10. For this reason we tried to minimize possible differences by ensuring that all data used are representative of the same area, months and time periods examined. This was performed by applying the common threshold criteria described in Section 2.3. We intend to emphasize this point in the revised manuscript. We will also include in Section 2 a more detailed description of the quality criteria adopted for the creation of the level 3 data used here, and possible effects of these in our findings.

Specific comments:

P2 line 25: The paper about the CALIPSO level 3 product written by Tackett et al. (2018) was recently published. I would suggest that the authors cite this paper.
Reply:
We thank the Referee for this suggestion. A citation will be included in the revised manuscript.

P2 line 28: Why only three types (dust, smoke, and polluted dust) were used? The CALIPSO aerosol models are consist of six aerosol types (Omar et al. 2009).
Reply:
As explained in our reply to the next comment, we used monthly averaged data products. The corresponding aerosol profile product from CALIPSO is the level 3, version 3 monthly product. It is interesting that, while the CALIPSO aerosol models consist indeed of six types, only three of them are available in level 3 (dust, smoke, and polluted dust). The reason is given in the discussion of the relevant paper (Tackett et al., 2018), where the authors mention that they chose to include a subset of aerosol subtypes for better management of the level 3 file sizes. According to their reply, the selection of these three subtypes was decided based on the algorithm performance in detecting them (Atmos. Meas. Tech. Discuss., doi: 10.5194/amt-2018-97-AC2, 2018). This is an important piece of information regarding this data product, that we plan to include in the revised manuscript. We acknowledge that a similar analysis of all aerosol types would probably provide more robust conclusions. It would require, however, a level of data processing similar to the one used to create the CALIPSO level 3 data set in the first place. This would exceed the scope of the present study. Furthermore, the additional analysis would not alter significantly the conclusions regarding the three aerosol types already analyzed.

P2 line 33: Why monthly GFED data used? I think the daily product is more appropriate for applying the thresholds in section 2.3.
Reply:
All data sets used in this study were monthly averages. We consider this temporal scale appropriate for studying both long-term interannual (many years or decades) as well as seasonal changes.

However, the application of the first threshold (page 3, lines 26-27) requires the number of days used in the calculation of the monthly average. This information is available in all level 3 data sets used. In GFED, the non-zero "daily_fraction" values, available on a daily basis and attributing a daily fraction to the total monthly emissions, were used as a count of the days used. It should be noted here, that monthly values of emissions over the study area were calculated by summing the corresponding pixel values, instead of averaging, which was the case for all other variables. Accordingly, in the case of GFED, this first threshold was applied on a "whole area", rather than a pixel basis. We plan to include these details in the revised manuscript.

P3 line 26: It is not clear to me how several products with different pixel sizes are treated. In addition, how about data sampling? The data sampling of each sensor is different. The swath width of the MODIS sensor is 2330 km, while the CALIPSO lidar only measure the nadir direction from the satellite orbit.
Reply:
Our analysis and corresponding results were based on spatially and temporally averaged values of cloud and aerosol properties, rather than pixel-level and instantaneous comparisons. In this context, we tried to ensure a fair treatment among data sets with different pixel sizes and samplings by applying common threshold criteria to the monthly and spatially averaged values, as described in Section 2.3. We also tried to minimize differences due to temporal sampling by only using data from afternoon satellites: MODIS-Aqua, AVHRRs on NOAA-18 and -19 (for CLARA-A2) and the corresponding daytime product from CALIPSO. It is true, however, that the viewing geometry of CALIPSO and the passive sensors is substantially different and can cause large discrepancies. While the underlying assumption here is that monthly averages smooth out these differences, we also acknowledge this possibility in page 4, lines 8-10. We consider it appropriate, however, to emphasize it further in Section 2.3 of the revised manuscript.

P4 line 9: It is not clear to me why MODIS AOD and CALIPSO AOD are different in Mach and April. The difference of MODIS AOD and CALIPSO total AOD in March and April is 0.3, which is comparable to the seasonal variation of CALIPSO total AOD.
Reply:
The difference in total AOD between the two data sets in March and April is indeed remarkable. However, we could not identify possible reasons for these discrepancies based on the data sets analyzed in this study. While this feature certainly merits further investigation, we decided not to provide hypotheses that cannot be supported by the present analysis. Since a possible explanation requires analyzing data from additional sources, we consider it as a significant digression from the present analysis, also considering that the main conclusions of this study were not drawn from features appearing in these two months. We plan, however, to highlight the need for further investigation in the revised manuscript.

P7 line 23: The CALIPSO lidar is unable to detect aerosols and clouds underneath optically dense cloud layers; therefore, the extinction coefficient of low-level cloud in Figure 8a is underestimated.
Reply:
We thank the Referee for this clarification, which supports our conclusion that low clouds prevail over the area during this season. We plan to include it in the revised manuscript.

P8 line 5: "less absorbing aerosols above stratocumulus clouds would lead : : :". The "polluted continental" type of CALIPSO aerosol models (Omar et al., 2009) is also regarded as one of absorbing aerosols. Why the "polluted continental" type is not included in the analysis?

Reply:

Please refer to our reply in the second specific comment.