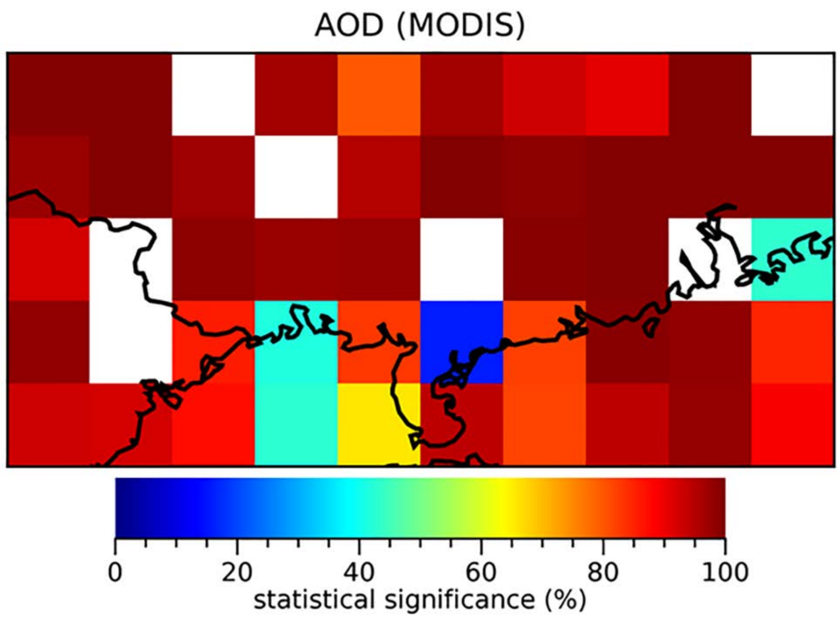
We thank the Anonymous Referee #2 for this second review. In contrast to the first review it gives concrete criticism, which we can respond to. Most comments are either based on misunderstanding and/or can be addressed by providing additional information. However, some comments are plainly wrong. Specifically, the referee mentions a wrong range of months to dismiss some of our findings (comment on page 1, lines 12-13 of the manuscript), describes wrongly the semi-direct effect for aerosols above clouds, thus expecting results opposite to our findings (comment on page 1, lines 15-17), and refers to irrelevant figures to disprove some of our results (comments on page 9, lines 1-4 and page 9, lines 10-11). Following are our point-by-point replies with the referee comments in italic.

Pg. 1, Lines 10-12: This statement in the abstract follows from Fig. 2 and Fig. 5 (panels above). It is important to see predictive statistics associated with these trend lines rather than just the percent changes. There's quite a bit of scatter in the data, some of which may be seasonal variability, some of which may be interannual variability, and then there's the uncertainty of the Level 3 data product itself. What are the slopes of the dotted lines (yr^{-1})? What are the p-values for the statistical test that one can reject the null hypothesis that the slope of the dotted line is zero?

Figures 2 and 5 provide information on the statistical significance of the plotted lines in the 95% confidence intervals, with corresponding percent changes highlighted in bold. This is explained in the caption of Fig. 2. It was omitted from Fig. 5 caption, but it should indeed be included for clarity. Regarding Fig. 5, it is also clearly stated (page 6, line 10) that in most of the cases the statistical significance level is below the 95% confidence interval. This is the reason why the analysis proceeds further to monthly changes, where the seasonal variability is removed. Reporting percent changes instead of (absolute) slopes was selected as a more intuitive measure of change. Regarding p-values, a table could also be added for completeness. The following table provides the information requested by the referee:

Parameter	Unit	CALIPSO	MODIS	CLARA-A2
		Change (%)/slope (<unit> yr^{-1})/p-value	Change (%)/slope (<unit> yr^{-1})/p-values	change (%)/slope (<unit> yr^{-1})/p-value
Total AOD	1	-23.3/-0.013/0.013	-17.6/-0.010/0.002	
Dust AOD	1	+8.4/0.0003/0.797		
Smoke AOD	1	-22.5/-0.006/0.071		
Polluted Dust AOD	1	-33.5/-0.008/0.003		
All-sky LWP	g m^{-2}		+12.4/0.837/0.204	+14.2/0.913/0.242
Liquid CFC	1		+6.8/0.003/0.219	+3.4/0.002/0.465
Liquid COT	1		+5.5/0.089/0.399	+3.6/0.058/0.607
Liquid REFF	μm		+1.6/0.018/0.239	+5.2/0.034/0.0003

The following figure depicts, on a pixel basis, the level of statistical significance for MODIS AOD changes (corresponding to Fig. 2a). For similar maps corresponding to the changes shown in Fig. 5a and 5b, the reviewer is referred to one of our later replies (page 11 of this document).



There's also quite a bit of day-to-day and sub-pixel variability that is not reflected in the Level 3 gridded monthly mean product as well as different numbers of measurements (i.e., samples) in each pixel that need to be considered and this is not really discussed in the manuscript. Some mention of area weighting of pixels is given on Pg. 3, Line 38, but is not described; how was this done? The monthly-averaged CALIPSO observations also have different numbers of observations that are averaged to yield the reported gridded mean and standard deviation – in addition to the area weighting, were these differences in number of samples accounted for when averaging across the region or across different months/years?

The area weighting mentioned by the referee concerns the differences in surface areas of grid boxes due to different latitudes. Because of the small size of the domain these differences are minor. This could be clarified in the statement of page 3, line 38. The different number of observations was accounted for by applying a threshold on the minimum number of days used in the monthly mean calculation (on a pixel basis) before estimating the spatial average (see also Section 2.4). In the case of CALIPSO, averages were weighted by the number of samples used, which is available in the level 3 data. Data sets from different sources will of course have different numbers of observations being averaged, as the referee mentions. The same concern led us to apply the thresholds described in Section 2.4, in order to minimize ensuing discrepancies. While we agree that sub-pixel variability is not reflected in the gridded monthly mean products, we consider that some rephrasing could answer the referee points previously mentioned.

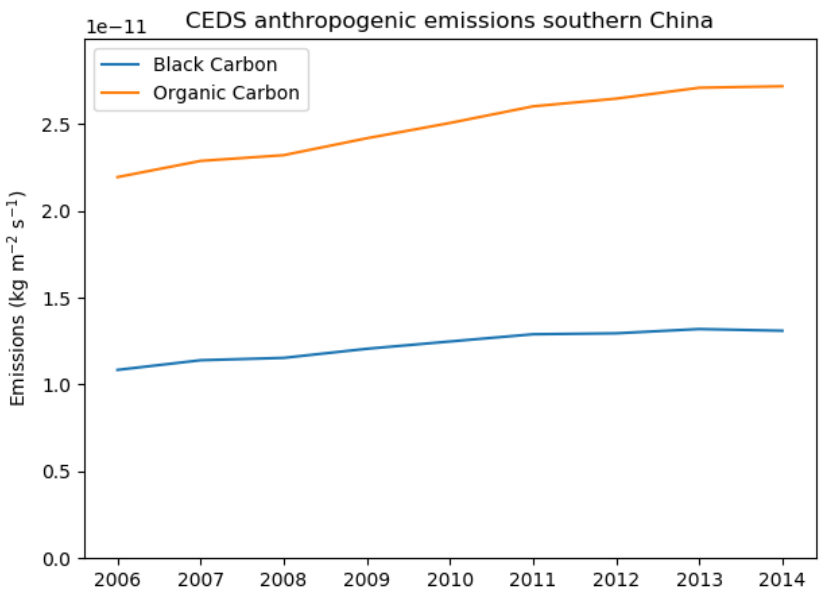
It would helpful for the reader to see the Level 3 standard deviations on these trendline graphs as error bars or as a shaded region. How is this additional sub-month, sub-grid-cell variability being captured in the statistical tests to assess whether or not there is a trend? Assuming there are indeed, statistically-significant trends (which I don't think hasn't been discussed very extensively at all) is the trend in AOD is related to the trend in CFC or LWP or are they coincidental? The italics statement above from the abstract implies that there is a non-coincidental relationship.

The requested information could be added in the graphs. While we considered that the statistical significance of trends was adequately discussed, this discussion could also be extended, with an addition of a relevant table of p-values, as shown before. However, based on the present discussion in the paper, it is explicitly mentioned that some changes are statistically significant and some are not, as was also explained in our first reply. Hence, we don't understand why the referee would still "assume" that "there are indeed, statistically significant trends". Regarding the question on the relation or coincidence of changes in aerosols and clouds, it is one of the main science questions of this study, as described in the Introduction (page 2, lines 11-12), and we attempt to address it based on the analysis described in Section 3.3. Our results imply that there is indeed a non-coincidental relationship.

Pg. 1, Line 12-13: The fundamental flaw with this conclusion is that it is not clear that all aerosol types have been captured, so one cannot say that the "main driver" of AOD trends over the last decade is biomass burning or continental pollution or marine aerosol or other types, because only the three aerosol types are included in the Level 3 CALIPSO product (dust, smoke, and polluted dust), and critical information about the trends of these other aerosol types is lacking. One must also ask the question, are the CALIPSO aerosol types sufficient to answer the question that's being posed, or does one need more specificity with regard to aerosol composition (e.g., sulfate, organics, dust, black carbon) that must be obtained from a model? Therefore, I would characterize this conclusion as unsupported by the underlying data and highly speculative. One way to address this criticism would be to not use the Level 3 data, but rather to use the Level 2 data that has more aerosol type classifications. Another approach would be to use model data products to explore this research question. Of course there would be uncertainties associated with any aerosol type classification scheme that would make it difficult to compare across different data sets – for example, the CALIPSO smoke aerosol probably is not only associated with biomass burning and also includes the contribution of other anthropogenic combustion sources. Another advantage of using the Level 2 data products is that they are not gridded and temporally averaged, so they capture a truer range of measured variability. This helps avoid biases, because the mean of the means is not always the same as the mean of the population if sample sizes are not constant and this unequal weighting is not accounted for properly.

It is true that the lack of the full set of aerosol subtypes is prohibitive for the attribution of their overall decrease to a "main driver". This expression should be corrected accordingly. Our results, however, show a statistically significant decrease in total AOD and in an aerosol subtype. This may indeed not be the main driver, but the lacking information on how some other subtypes change is not "critical" for further analyzing this specific subtype. Regarding the sufficiency of the CALIPSO data used "to answer the question that's being posed", and based on the referee's suggestions (level 2 data, model data) and some later comments, it seems that there is a misunderstanding on the reason of using CALIPSO in this study, which should be clarified also in the manuscript: it is neither to attribute aerosol sources, nor to unveil ACIs in their process level. It is to include information on the vertical distribution of aerosols, and its possible changes, which is critical for their position relative to clouds. This is why, while the CALIPSO subtypes are analyzed in terms of changes and compared to other data sets (e.g. GFED), the ambiguity regarding their origin is repeatedly stressed throughout the manuscript.

Following a suggestion by referee #1 we have collected CEDS anthropogenic emission estimates compiled for CMIP6. As shown in the figure below, these estimates suggest that both the organic carbon and black carbon emissions cannot explain the decrease in AOT in southern China, because both show slight increases in the 2006-2014 period.



Regarding the referee's remark on the unequal weighting in averaging, as explained in our previous reply, the unequal number of observations is actually accounted for.

Putting aside the major flaw of the missing aerosol types, it is also hard for me to see the trends in the data that the authors are using as a basis for saying that AOD changes occur in late Autumn and early Spring. From Fig. 1a and Fig.3 (shown at right), the peak in biomass burning in GFED is apparent between Nov.-Mar., while the ΔAOD traces vary quite a bit but don't really peak in this period. There is some decrease (the traces are below zero), but there is also a good bit of scatter in the data.

The combination of plots provided here by the referee to support the above criticism actually compares the seasonal variation in biomass burning emissions from GFED with the *changes* in the seasonal variation of AOD and biomass burning emissions. It is not obvious why the maximum change in a parameter should also coincide with its maximum average value, as the referee seems to expect. It is also worth noting that the original statement, cited by the referee (page 1, lines 12-13) reads "changes occurred mainly in late autumn and early winter months", not "*early Spring*". In fact, nowhere in the manuscript is said that “*AOD changes occur in late Autumn and early Spring”*. The sentence cited by the referee mentions “early winter”, and actually refers to AOD from MODIS and CALIPSO, and their changes in October, November and December (Figs 3a and 3b). We would be happy to rephrase this part in order to clarify it. However, the referee has drawn here a red box ranging from November to March, to show that (indeed) ΔAOD does not peak in this period. The same mistake is repeated in a later comment (page 9 in this document). Hence, this criticism is rather superficial and obviously unsupported.

There are no metrics of statistical variability included in this graph (Fig. 3) – only the means – so it is hard for me to assess the statistical significance of the data. The authors say they did t-tests, but on what? Where are these statistical results presented? Statements are made in multiple places that trends are statistically significant but no p-values are provided. Where variables are thought to be correlated (as in the case of biomass burning emissions and AOD), there are no correlation coefficients provided. I see this lack of scientific rigor as a major flaw in this study. It also led me to comment that most of the correlations suggested are determined by whether one or more variables trend up/down together over time, which I guess is determined visually. Having some numbers here related to the statistics, I think, is very important.

We appreciate the referee’s request for more details on the metrics of statistical variability. There are indeed statements in multiple places that changes are statistically significant, with corresponding p-values not provided, and a concentrated report of corresponding metrics could help. However, describing the method used for the assessment of statistical significance, and then reporting results individually, is a rather common practice in similar cases and does not constitute “lack of scientific rigor”. In fact, it is a matter of a simple revision for these metrics to be provided. However, the “visual determination” of correlations, mentioned here by the referee, does not qualify as a valid scientific approach. Hence it is surprising that this is the referee’s “guess” regarding our methodology, and unfortunately it is not even accompanied by the “benefit of the doubt”. In our opinion, such a serious statement in a review process should at least give to the authors the opportunity to disprove it, instead of leading directly to such a negative judgement. Nevertheless, we take the opportunity here and provide the requested metrics:

- Page 5, line 6: p-value=0.002 for MODIS, 0.013 for CALIPSO.
- Page 6, line 4: please refer to the maps provided here in page 11.
- Page 6, line 10: please refer to the table provided here in page 1.
- Page 6, lines 33-34: please refer to the table provided here in pages 11-12.
- Page 8, line 9: the referred decrease is not “significant” in a statistical sense (95% confidence interval). The term should be replaced to avoid misunderstandings.
- Page 9, line 9: p-value=0.03. This decrease refers to polluted dust aerosols, should be replaced for clarification.

The following table shows the Pearson’s correlation coefficients, on a monthly basis, of each CALIPSO aerosol subtype with GFED emissions. Please note that nowhere in the text is the total AOD correlated with GFED, as mentioned by the referee. Please also note the correlation between GFED and polluted dust AOD in November, which led to the discussion in page 5, lines 24-29 of the manuscript.

	Dust AOD	Smoke AOD	Polluted dust AOD
January	-0.27	-0.06	-0.10
February	-0.28	0.37	0.18
March	0.14	0.59	0.16
April	-0.38	-0.50	0.30
May	0.42	-0.07	0.27
June	-0.37	0.24	-0.07
July	-0.31	0.16	-0.23
August	0.28	0.00	0.06
September	-0.44	-0.37	-0.37
October	-0.09	-0.07	0.37
November	-0.43	-0.05	0.74
December	-0.30	0.62	0.25

Pg. 1, Line 13: The panels from Figs. 3 and 7 shown at right on this page indicate the seasonal variation in changes of AOD (top) and cloud properties (bottom three panels). There is a very clear and distinct change in cloud properties in Nov.-Dec. that does not appear to be related to the changes in AOD during this period. I don't understand the basis for the italicized statement made in the abstract that changes in AOD "coincided with changes in cloud properties".

The term "late autumn and early winter months" should probably be replaced by "November and December" to clarify the issue. It should be obvious also from the red box that the referee has drawn that in these months there are changes in AOD and cloud properties that coincide. It is not clear what the referee means by "*does not appear to be related to the changes in AOD during this period*". The term "related" was not used in the statement that the referee cites, and no "relation" was established based on the plots that the referee has compiled. The fact that in other months (e.g. October) AOD and cloud property changes do not coincide, does not negate our statement. It is actually the main finding of our paper that the AOD changes in October occur at a higher level and thus have different effects on clouds.

Pg. 1, Line 15-17: The semi-direct posits that solar heating of above-cloud absorbing aerosol layers changes the temperature profile of the atmosphere, reducing buoyancy, and ultimately cloud cover and liquid water path. To be consistent, then, with the semi-direct effect, I would expect to see an inverse correlation between absorbing aerosols above cloud and these cloud properties. What is shown in Figure 8 are monthly-averaged differences in the vertical profile of aerosol extinction as well as the vertical profile of cloud extinction. First, extinction is not absorption. Even relatively close to fires, the scattering-to-extinction ratio is > 0.8 (e.g., Yokelson et al., Atmos. Chem. Phys., 2009; https://doi.org/10.5194/acp-

9-5785-2009), and it is known that the ratio is much higher as the smoke plumes age. No data is being presented regarding smoke age, whether or not the smoke is from urban pollution or biomass burning, or that there is or isn't any trend in absorbing aerosols over this region.

Second, the CALIPSO Level 3 data typing algorithm identifies smoke only when the layer is elevated – by definition! Therefore, it is not appropriate to use the positioning of this smoke product to suggest that there is some sort of vertical relationship with cloud. The smoke classification type shares many similar features to the polluted continental classification type, except that the latter is at the surface and not elevated. The polluted continental classification type has not been considered in the present analysis, which is a major gap in the analysis. Finally, I don't understand the relevance of the ISCCP classification types to this discussion – this classification scheme seems much too coarse to be meaningful. In sum, I see no conclusive evidence that aerosol changes are altering the temperature profile of the atmosphere to effect changes in clouds. Consequently, I don't think that it's appropriate for the authors to suggest that the semi-direct effect is a causal mechanism for the observed, 5-13% increase in LWP and cloud fraction from 2006-2015.

The referee's statement that the semi-direct "posits that solar heating of above-cloud absorbing aerosol layers changes the temperature profile of the atmosphere, reducing buoyancy, and ultimately cloud cover and liquid water path" contradicts the widely described semi-direct effect mechanism for absorbing aerosols above stratocumulus clouds (e.g., Koch and Del Genio, 2010). The buoyancy is indeed reduced but this will lead to less entrainment at the cloud top and consequently an **increase** in cloud cover / liquid water path. Hence, a decrease in absorbing aerosols above clouds would be consistent with a corresponding decrease in stratocumulus clouds below. This is exactly what is shown in Figs. 8b and 9b.

The purpose of Fig. 8b and c is to indicate at which height the changes in aerosol occur. The CALIPSO extinction profile suits that purpose. The fact that "extinction is not absorption" is irrelevant for the conclusions regarding the vertical location of the aerosol changes. We acknowledge in the manuscript that the CALISPO smoke classification is accompanied with some uncertainty. As a result we do not know for sure how absorbing these elevated aerosols are. However, for the polluted dust aerosols, showing largest decreases in November, we have solid indications that they are strongly absorbing because their decrease goes together with a decrease in GFED biomass burning emissions (Fig. 3c) while anthropogenic emissions did not show a decrease (see figure on page 4 of this reply).

The ISCCP classification was included to highlight the changes in low clouds for October and November. We don't see why this classification, which has been used widely in the past, does not serve this purpose.

Pg. 4, Line 32-33: Why was it not possible to explore this discrepancy? How would further investigation be carried out? This is a very shallow approach to analyzing the data.

Our statement reads "…*it was not possible to pinpoint specific reasons for the March-April differences based on the data sets used here".* Contrary to the referee's understanding, this statement denotes that this discrepancy was actually explored, based on the data sets used here, but no explanation was found. Hence, further investigation would require further analysis of additional data sets, focusing on these months. This would extend beyond the scope of this study, which focuses on the October and November months.

Pg. 4, Line 35-36: Biomass burning aerosols do contribute to smoke layers, but so do other sources of combustion. Similarly, biomass burning, urban pollution, and fossil fuel combustion aerosols contribute to the polluted continental aerosol type (which is not accounted for in this study). A key difference between the CALIPSO smoke and polluted continental aerosol types is whether or not the layer is at the surface or elevated. Since the aerosol classification types are based on aerosol intensive and extensive parameters, there can be misclassification and some ambiguity across aerosol types, particularly for categories dominated by smoke and urban pollution because both types of aerosol are dominated by relatively small, non-depolarizing aerosols. The polluted dust category isn't necessarily a mix of biomass burning and dust – it represents the middle part of the continuum between smoke/continental-pollution (small and weakly depolarizing) and dust (large and strongly depolarizing). The satellite aerosol-typing products are very useful, but they are not unambiguous. This statement is too strong and not supported by the data.

We thank the referee for these clarifications, which could be added in the relevant discussion along with the appropriate references. We also agree that the aerosol-typing products are useful, but they are not unambiguous. In fact, the ambiguity regarding especially the smoke aerosol type is explicitly stated in page 5, lines 31-32. This statement could also be rephrased in accordance with these ambiguities.

Pg. 4, Lines 36-37: I agree with the authors' statement here, and yet, Figure 1 and Figure 3 attempt to make precisely this comparison.

This statement, along with the next sentence (page 4, lines 37-39), explain why biomass burning emissions and satellite-based AOD are expected to differ (i.e. not being directly comparable). However, including them as subplots in Figs. 1 and 3 is useful, in our opinion, since the former can help explain the latter.

Pg. 5, Lines 7-8: It is true that fitted lines to both polluted dust and smoke aerosols trend down during this period along with the overall AOD. However, it is unclear what the trend in continental pollution or marine aerosols are for this period because they have not been considered by this study. Certainly, decreases in polluted dust and smoke contribute to the decrease in AOD, but I don't think that the *authors can "attribute" the change to only these two aerosol types when there are other types that are not being considered.*

We understand that the term "can be attributed" may be misunderstood as rather definitive, when other possibilities are not excluded. However, the total AOD from the three CALIPSO categories matches quite closely with the total AOD from MODIS, so other categories appear to play a minor role. Moreover, anthropogenic emission estimates (including continental pollution) do not show a decrease over the investigated time period (see page 4 of this reply).

Pg. 5, Lines 14-15: What evidence is there that the smoke and polluted dust aerosol types are dominated by biomass burning aerosol versus other sources of combustion or pollution aerosols? The CALIPSO aerosol type is not specific to biomass burning. Consequently, for the authors to make this conclusion, they need to provide some other evidence. Since no such evidence is apparent in this manuscript, this seems highly speculative.

One piece of evidence that biomass burning emissions play a major role, already provided in the paper, is GFED. Another piece of evidence are the CMIP emission inventories, mentioned here in page 4, which could be added to the manuscript.

Pg. 5, Line 15-18: No data on residential energy sources are provided or discussed in this manuscript, so this statement is entirely speculative, and, frankly irrelevant to the present study. The previous studies cited in the next sentence are also not sufficient to support this statement, as they are not recent enough be cover the 2006-2015 time period in this study. Even if there was a decrease in residential biomass burning emissions starting in the 1990s, such a decrease does not necessarily extend to present day. This conclusion is unfounded.

Information on the seasonal peak in residential energy sources is provided in He et al., (2011). While this reference should be added here, it should also be clear that this is not a conclusion of the present study, since no relevant data are provided here. The way the referee connects this sentence with the next ones is rather arbitrary. We hope it is clear now that these previous studies are not provided here to support the referred statement, but in a more general discussion on how previous findings relate to ours, which is a rather common and necessary practice.

Pg. 5, Line 20-: Again, showing the same figure as before at right, it can be seen that there is no agreement between the change in AOD and the change in C emissions (delta-AOD even becomes positive in January, while delta-C is fairly constant). I'm not sure I understand what is being meant by the term, "partially agrees". It appears that during the seasons where delta-C reaches a local minimum and is fairly stable that both MODIS and CALIPSO delta-AOD are quite variable and not at a local minimum or maximum. Finally, is it even appropriate to be trying to establish this comparison, as it was already *stated on Pg. 4, Lines 36-37: "Biomass burning emissions and satellite-based AOD are not directly comparable"?*

Biomass burning emissions constitute part of the total aerosol load. Hence, they are expected to agree better with the total aerosol load concentration (or changes) when they dominate compared to other sources, rather than when other sources (or their corresponding changes) dominate the total aerosol load. This is the intended meaning of the term "partial agreement". ΔC actually reaches a minimum in November, when MODIS and CALIPSO ΔAOD also exhibit large decreases, but not necessarily their minimum values (since they are not expected to always agree). It is obvious that the term "not directly comparable" causes a misunderstanding, and could be replaced by "not expected to always agree", which is probably more appropriate for the intended meaning. Regarding the appropriateness of the comparison between biomass burning emissions and satellite-based AOD, the referee is referred to our previous reply (page 8 of this document). In short: yes, comparing a total with its part gives insights on the relative contribution of the latter to the former.

Pg. 5, Lines 23-26: What evidence is there to assert that the aerosols or smoke observed over this region is transported from neighboring regions such as Indochina (versus long-range transport or local emissions)? No data on fire activity in neighboring regions is presented, nor is any information on air mass back trajectories. What about the confounding influences of local, urban pollution and non-biomass combustion aerosols on the CALIPSO types? This italicized statement seems highly speculative.

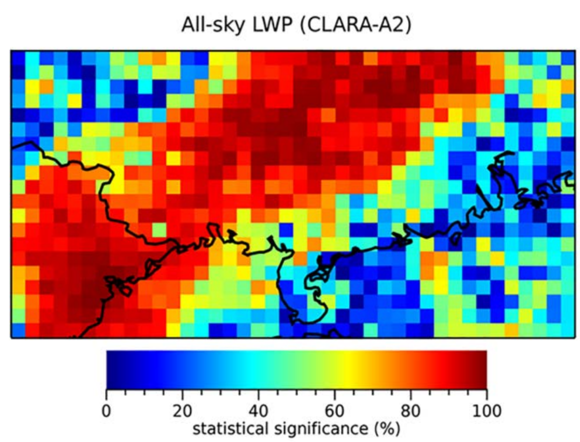
This statement is not an "assertion" based on "evidence". It is a "suggestion", a possible explanation for the reported results, not excluding other possibilities. We agree that there are alternative possibilities, and we thank the referee for suggesting methods that could lead to an assertion based on evidence.

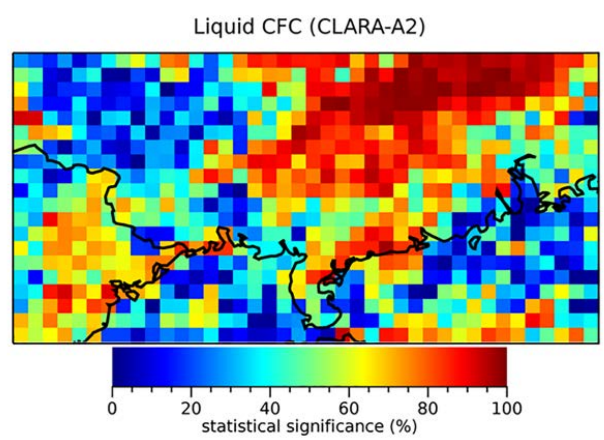
Pg. 5, Lines 29-33: I agree 100% with this statement. The problem is that this ambiguity undercuts many of the conclusions put forward in this manuscript. This is made even more problematic in that the continental pollution aerosol type is not included in this analysis. Given these uncertainties, a fundamental question that must be asked is, is this satellite-based AOD aerosol type data set appropriate to address source attribution? Given that some major aerosol types are missing, I think the answer is that it is not appropriate.

We agree that there are serious limitations in using this data set to address source attribution (see page 8, lines 40-42). As stated in page 2, lines 26-27, this data set was not used here for this purpose. It could also be (further) clarified that possible relations of CALIPSO aerosol types with aerosol sources were based only on previous studies (page 4, lines 35-36, page 5, lines 28-29), and do not constitute "strong" conclusions of the present study. This should be obvious from page 8, line 41, but it could be further emphasized.

Pg. 6, Lines 8-10: I'm struggling to interpret the meaning of this statement. Spatial distributions of the change in LWP and cloud fraction are presented that demonstrate large increases in LWP and cloud fraction over land and decreases over water. No indication is given in Figure 5 of the areas where these changes are or are not statistically significant (often I have seen this done with speckling overlaid on the statistically significant portions). It also sounds like from this statement that most of the cases in Fig. 5 are not statistically significant. Yet, these are the numbers that are being quoted in the abstract for changes in liquid cloud cover and liquid water path of "5% and 13%, respectively" (Pg. 1, Line 12) and for drawing other conclusions later on. Are these numbers statistically significant?

The information requested by the referee can be easily provided. In fact, the following two maps show the level of statistical significance, on a pixel basis, of all-sky LWP (left) and liquid CFC (right):





In the paper, these results were summarized in the statement that the referee mentions, but they could also be included, to avoid struggles with interpretation. We think, however, that the statement "*reduces this significance to levels below 95% in most cases of Fig. 5.*" is clear. The referee is also referred to the table in the first page of this reply. It is not clear, however, what the referee means by the term "drawing other conclusions later on". After stating that the statistical significance is reduced to levels below 95% in most cases of Fig. 5 due to averaging (page 6, lines 9-10), the 34-year CLARA-A2 time series is analyzed (page 6, lines 13-26), and then changes are examined on a seasonal basis. In page 8, lines 33-34 it is stated which of the cloud properties exhibit statistically significant changes in which months. To clarify any further misunderstandings, we include here a detailed table with corresponding levels of statistical significance (in %) for every cloud property examined and every month:

	All-sky LWP		Liquid CFC		In-cloud LWP		Liquid COT		Liquid REFF	
	CLARA	MODIS	CLARA	MODIS	CLARA	MODIS	CLARA	MODIS	CLARA	MODIS
Jan	14.13	13.49	54.18	39.26	6.15	5.13	3.59	6.41	14.13	56.52
Feb	10.79	3.68	44.67	39.57	10.60	15.01	5.28	1.13	14.10	52.25
Mar	5.52	15.19	1.47	12.82	12.85	7.65	12.19	11.08	47.77	45.10
Apr	33.05	42.88	28.30	4.58	58.39	61.04	50.12	54.49	73.79	72.08
May	79.19	94.20	61.31	30.53	96.01	99.22	85.28	95.12	86.93	48.38
Jun	13.69	26.27	85.18	49.95	39.64	62.92	36.34	55.18	61.09	59.25
Jul	22.71	39.33	2.57	56.52	26.96	26.94	23.06	2.27	90.11	93.16

Aug	4.06	77.65	57.83	42.97	74.50	29.39	34.42	1.00	86.12	97.35
Sep	36.62	27.02	55.97	37.12	30.88	33.54	21.79	22.90	92.01	46.53
Oct	46.82	69.47	74.85	81.51	72.35	70.81	59.18	66.37	79.20	50.22
Nov	90.07	90.99	99.26	99.55	65.37	86.26	48.39	78.23	89.59	88.76
Dec	97.05	99.33	35.75	54.15	99.43	99.91	96.45	98.57	97.46	67.40

In fact, contrary to the referee's statement, no conclusion is reported based on the changes in the entire time series of Fig. 5. The reason is that "most of the cases in Fig. 5 are not statistically significant". We hope this is clearer now.

Pg. 6, Lines 33-34: Do I understand this statement correctly to say that there is only one month where each of the cloud property changes is statistically different from zero, and that only cloud fraction is statistically different from zero in November? Why would there be a trend in cloud properties in only a single month? Reading between the lines here, does this mean that the overall change in cloud property changes is not statistically significant, or is the effect size of this single month enough to drive the entire trend?

Yes, this is exactly what this sentence states. Regarding the referee's second question, a possible reason could be "differences in the seasonal (or monthly) characteristics and changes in factors affecting cloud properties". Such a factor is aerosols, and a large part of the present study investigates exactly this question. This single month is not enough to drive the entire trend, hence most of the cases in Fig. 5 are not statistically significant, as stated in page 6, lines 9-10. Our previous reply also provides further clarifications.

Pg. 6, Lines 37-38: I don't think this statement is correct. Liquid clouds increased only in November on a cloud fraction basis and only in December on a LWP or COT basis (based on the statistical significance discussion in the previous comment). The seasonal pattern of the AOD changes and the cloud properties changes are not similar (either correlated or anti-correlated). This statement suggests that they are anti-correlated, which is not true.

This statement was phrased carefully, to avoid misinterpretations. The referee correctly states that liquid clouds increased only in November on a cloud fraction basis and only in December on a LWP or COT basis, based on the statistical significance results (see also the relevant table in pages 11-12 of this document). We summarized these changes using the phrase "liquid clouds increased mainly in late autumn and early winter", but this could be rephrased, based on the referee's statement, to be more clear. We also agree that the seasonal pattern of the AOD changes and the cloud properties changes are not similar. In fact, we don't see why they should be similar (either correlated or anti-correlated) throughout the year. However, it is one of our findings that some statistically significant changes in properties of clouds and aerosols occurred in the same months, and they are indeed anti-correlated (see also the table in page 16 of this document).

Pg. 7, Line 10-11: First, statistical significance of the cloud changes is unclear (discussed previously). There does appear to be a decreasing trend in AOD, which the authors assert is statistically significant. That surface pressure and geopotential height do not show a statistically significant trend is insufficient to rule out meteorological drivers. The atmospheric temperature profile, moisture, and lower tropospheric stability are also important variables that do not appear to have been considered. Even if these variables fail to demonstrate a statistically significant trend that does not in and of itself rule out the existence of such a trend. All it means is that the available data are insufficient to reject the null hypothesis, but there may indeed be a trend that might be uncovered by additional data and/or a longer timeseries. The italicized statement is not demonstrated conclusively by the data presented, which are rather superficial.

We hope the statistical significance of cloud and aerosol changes in now clearer, based on our previous replies. Apart from surface pressure and geopotential height, other parameters could indeed be considered, as the referee suggests. It is not clear, however, what the referee means by the following: "*Even if these variables fail to demonstrate a statistically significant trend that does not in and of itself rule out the existence of such a trend. All it means is that the available data are insufficient to reject the null hypothesis, but there may indeed be a trend that might be uncovered by additional data and/or a longer timeseries"*. This statement seems to suggest that there is no way of excluding meteorological variability as a factor of cloud and aerosol changes, since there might always be a trend waiting to be uncovered by additional data and/or longer time series. This contradicts common practices followed in similar studies.

Pgs. 7-8, Section 3.3.2: I think that this paragraph is not at all supported by the underlying data, which until this point has focused on trends and changes over time. In this paragraph, process-level explanations are invoked, but are done at a highly-averaged level spanning months and 5 x 10 degree area. These are not the scales at which aerosol-cloud interactions would be expected to be evident (e.g., McComiskey and Feingold, ACP, 2012, https://www.atmos-chem-phys.net/12/1031/2012/), so the failure to see ACI effects in the trend data is not surprising. Saying that the authors' findings are "inconsistent" with the first and second indirect effects is too strong a statement. These effects may very be visible in this region if a more appropriate data set is used (e.g., aircraft, balloon, surface remote sensor scales measuring clouds over minutes to hours or a model with better space and time resolution). One cannot know. The same is true for the discussion on the semi-direct effect, with the additional comments that have been described above in this review that the attribution of the particles to biomass burning, as absorbing particles, and that these particles are at cloud level are all not established by this data set. In fact, the CALIPSO smoke type is not unambiguous as a marker for biomass burning. That the smoke type often occurs near cloud level is unsurprising given that the layer must be elevated by definition of that aerosol type in the CALIPSO scheme. Similarly, the continental pollution aerosol type is very similar to the smoke type, but it is not in an elevated layer. Since the data set used in this manuscript and its analyses is so highly averaged in space and time, it is of little utility for discussing ACI effects. Consequently, the conclusions as stated are not definitive and this entire paragraph should be removed.

This paragraph examines what can (and cannot) be deduced from the previous analysis regarding ACI and clearly states what can (and cannot) be supported by the data. It is not true that results of ACI cannot be evident at scales larger than their process scale, as the referee suggests. The problem with the temporal and spatial scales used here, is one of quantification of ACI, as clearly stated in the study cited by the referee, and acknowledged also in our study (page 1, lines 33-35). However, no ACI quantification was attempted in the present study, so the referee's criticism is rather unsupported. In the same sense, if a particular mechanism dominated over a large area and period, one would expect to see the consequences in a data set covering this area and period. If the analyzed data sets show changes in a direction opposite to the one expected, then they are "inconsistent" with the previous assumption. We don't see why such a term would be "strong" or "unsupported" by the data.

We agree with the referee that ACI may be visible based on the suggested data and resolutions. However, the purpose of this study is not to provide evidence of ACI on their process level. It is (among others) to examine if their consequences can explain observed changes in a larger scale.

We acknowledge the referee's concerns on the limitations of the CALIPSO aerosol types based on their definitions. In fact, their limitations are acknowledged in several parts of the manuscript (page 5, lines 29-33, page 8, lines 18-20, page 8, lines 40-42). However, this section examines *changes* in their profiles, not concentrations. The referee finds "unsurprising" that "the smoke type often occurs near cloud level". This is a rather confusing statement, since it is nowhere made in the manuscript.

Pg. 8, Line 10: Are the aerosol and cloud profiles shown in Figure 8 an average profile or an individual, typical profile for each? What is meant by "autumn" or "Fall" for the cloud extinction profile – both October and November? If they are averaged profiles, how was that averaging carried out (e.g., was a weighted average of the sample numbers in each pixel used)? Are there meaningful differences in the profiles across the spatial area? A single set of averaged profiles over the entire spatial domain seem difficult to meaningfully interpret to me, as I would expect these profile changes to be very different over land and over water. How should the reader interpret these profiles with regard to representativeness? Are the changes in Fig. 8b and 8c statistically significant at all height levels? The commentary on Pg. 8, Line 21-23 suggests that only certain layers are statistically significant and in different months (e.g., 1-1.5 km altitude for smoke in October and 0.7-1.2km for polluted dust in November). It would appear that the CALIPSO smoke aerosol change is not statistically significant in November when the cloud fraction change is statistically significant. Conversely, the smoke change is statistically significant in October when the cloud fraction change is not statistically significant. What about the December profiles where the other cloud property changes are statistically significant? It is very difficult to unravel what is being presented here, but it certainly does not seem to be suggestive of an aerosol-cloud semi-direct effect (as is stated on Pg. 1, Lines 15-18).

The cloud profile shown in Figure 8 is spatially averaged over the study area and autumn (fall) months (September, October, November), based on measurements from 2007 to 2011 (see also Amiridis et al. 2015, for details on the LIVAS data set). The aerosol profiles of Figure 8 actually show changes, calculated, for each profile level, based on the method described in Section 2.4. The averaging was indeed carried out using the numbers of averaged samples, also provided in the data set, as weights. It is not clear what the referee means by "meaningful differences". It is true that differences should be expected, especially over land and sea, and selecting two of the four pixels covering the study region would probably be more representative of the land profiles. A separate analysis could be performed to answer the question on representativeness.

The referee interprets correctly the statement in page 8, lines 21-23. To clarify further: in October, smoke changes are statistically significant between 1-1.5 km, liquid cloud changes are not; in November, polluted dust changes are statistically significant between 0.7-1.2km and liquid CFC change is also statistically significant; in December, all liquid cloud properties changes examined, except for CFC, are statistically significant, and aerosol changes are not. While some rephrasing might help, we consider reporting the results on statistical significance along with corresponding changes really crucial, hence some difficulty in unravelling the findings should be expected.

Figure 9 and related discussion: Why is necessary to break out the cloud optical thickness data by cloud type? It has already been established from Figure 7 and associated discussion that delta-COT is not statistically significant in October or November. This is mentioned in passing on Pg. 8, Line 28. Yet, there is then extensive discussion on the coincidence between decreased biomass burning, increased liquid cloud fraction and water content in November and a decrease in smoke aerosols in October (Pg. 8, Lines 28-32). I find this discussion very confusing, but much of it appears to be based on source attribution that has already been discussed in this review as being speculative. The ISCCP cloud type classifications does not bring any additional clarity or information to the major flaws in the prior conclusions.

It is not clear what the referee means here. Establishing a non-significant change in a parameter of liquid clouds, does not necessarily exclude the same parameter from changing significantly in a cloud sub-type. Similarly, establishing a non-significant change in a time series does not exclude significance on a monthly basis, as was shown in this study. As the referee mentions in a previous comment, establishing non-significance does not rule out the existence of a significant change that might be uncovered by additional data and/or longer time series.

The part of the discussion mentioned by the referee seems indeed confusing and should be rephrased, since it is not based on source attribution, as the referee claims, but rather on the position of the aerosols relative to clouds. The notion that a non-significant change in COT should prevent an analysis and discussion of changes in biomass burning and smoke aerosols, coinciding with changes in liquid cloud fraction and water content, is also unsupported.

Pg. 8, Lines 40 – Pg. 9, Line 1: I fundamentally disagree with second part of this statement. The manuscript is saying that smoke aerosols are from biomass burning and are absorbing, and therefore, an association between smoke aerosol height and cloud height in Figure 8 is somehow related to the semi-direct effect. If the aerosols are misclassified or smoke dominated urban aerosols that are weakly *absorbing then this would indeed call the conclusions of the manuscript into question. The missing CALIPSO aerosol types and the compositional ambiguity provided by this typing method make this particular dataset less capable for addressing the types of science questions and drawing the types of conclusions that are sought in this manuscript. This limitation is a significant one that cannot be overcome without new data and new analyses.*

The referee keeps repeating the same argument. We acknowledge that CALIPSO does not give unambiguous information about certain aerosol types, in particular biomass burning smoke and urban pollution. However, as stated before, the GFED dataset demonstrates that (absorbing) biomass burning aerosol emissions have markedly decreased over the decade studied. At the same time, anthropogenic emissions have not decreased (see Figure on page 4 of this reply). These pieces of information give strong additional indications that the aerosols are not largely misclassified.

Pg. 9, Lines 1-4: Here are the monthly timeseries from Fig. 1 and Fig. 4. I think that it's hard to make the case as is done here that there is a meaningful anti-correlation between polluted dust AOD and liquid CFC only in November and December (circled regions) and that such an anti-correlation can be used to draw a conclusion. There is just too much variability (and that's before the additional requested information of standard deviation error bars or shaded regions are added to the graph). The decreasing pattern in the polluted dust aerosol is present for the smoke aerosol and does occur during other months. Given the scale, it's difficult to discern the trend for the dust trace.

The following table shows the Pearson's coefficients of monthly liquid CFC from CLARA-A2 and MODIS, and AOD from CALIPSO. Results are shown separately for total, dust, smoke and polluted dust AOD.

	Total		Dust		Smoke		Polluted dust	
	CLARA-A2	MODIS	CLARA-A2	MODIS	CLARA-A2	MODIS	CLARA-A2	MODIS
January	-0.21	-0.27	0.26	0.18	-0.23	-0.26	-0.11	-0.17
February	-0.32	-0.23	-0.10	0.05	-0.45	-0.44	-0.15	-0.07
March	-0.42	-0.29	0.01	<0.01	-0.60	-0.53	-0.09	0.07
April	0.02	0.05	0.30	0.35	0.35	0.43	-0.48	-0.51
May	0.49	0.55	-0.13	-0.34	0.39	0.53	0.34	0.36
June	0.15	0.42	-0.06	0.17	-0.07	-0.17	0.11	0.38
July	0.36	0.26	-0.24	-0.53	0.21	0.41	0.47	0.12
August	0.79	0.74	-0.04	-0.15	0.69	0.75	0.41	0.26
September	0.02	0.18	0.18	0.37	-0.31	-0.18	0.08	0.26
October	<0.01	-0.29	0.42	0.01	0.02	-0.32	0.05	-0.15
November	-0.53	-0.50	0.21	0.21	0.08	0.10	-0.73	-0.69
December	-0.83	-0.81	0.22	0.14	-0.69	-0.66	-0.80	-0.79

A quick inspection of the table makes our statement obvious: *the liquid CFC and the polluted dust AOD are anti-correlated in November and December, with correlation coefficients around -0.7 to -0.8.* This statement is true for both CLARA-A2 and MODIS liquid CFC. Inclusion of this table would clarify this issue. However, the referee inadequately uses the plots showing the monthly averages of these variables and falsely calls these plots “monthly time series”, to disprove a statement that was not based on those plots. This is an unsettling level of misunderstanding.

Pg. 9, Line 8-10: No data on aerosol absorption are presented in this study. Some CALIPSO aerosol types are missing, and those that are included cannot be unambiguously attributed to biomass burning emissions. This conclusion is not true.

This same argument has been made many times, and we refer to our reply on page 16.

Pg. 9, Line 10-11: Technically this statement just says that cloud fraction and thickness both change across months. The changes are neither correlated nor anti-correlated (Figure 4a and 4c). At best, this sentence doesn’t reach a meaningful conclusion, and at worst it misleads the reader into thinking that concurrent changes somehow track each other. I recommend that this sentence be removed.

This sentence should indeed be rephrased. It is apparent that the referee is confused and makes an inadequate judgement, since Figure 4a and 4c do not show changes in liquid cloud fraction and optical thickness, but monthly averages of these parameters during the period examined. The sentence refers to changes concurrent with aerosol changes, in different months.

Pg. 9, Line 11-13: No data actually relating the position of aerosols to clouds is presented. Instead, the vertical distribution of cloud extinction and the vertical distribution of aerosol temporal change are presented in Figure 8. These are not the same thing. It is not shown that the sign of cloud changes is determined by the position of aerosol relative to clouds. This statement is not true.

We thank the referee for this remark. Indeed, the data provided actually relates the position of aerosol *changes* relative to clouds, and this statement should read “*Further analysis of vertical profiles of aerosol changes and clouds showed that the signs of cloud changes depended on the position of aerosol changes relative to clouds…”* in order to be true.

Pg. 9, Lines 17-18: It is very difficult to relate the trend analysis changes for the aerosol AOD and cloud changes to a process-level causal mechanism. Even comparing the trend changes of aerosol and cloud in this manuscript are difficult because they appear to vary differently across monthly and with different (and poorly explained) levels of statistical significance. It is not true to suggest that the data show a “high level of consistency” with the semi-direct or any ACI effect. This is because the data used here are highly averaged in both space and time, and are, therefore, less ideal for tackling these sorts of science questions.

Acknowledging the limitations of the data and concerns similar to the ones of the referee, we did not attempt to establish any process-level causal mechanism, as was emphatically noted in the statement right after the one cited here by the referee. We also tried to establish consistency in space and time (see Section 2.4) before attempting any comparison of changes. We hope that the levels of statistical significance are better explained now based on our previous replies. However, we consider that suggesting "*a high level of consistency with predictions*" of the semi-direct effect is a modest statement, which takes into account the limits of our analysis. The reason invoked by the referee to dismiss this statement, namely that the data used are highly averaged in both space and time, is indeed prohibitive for quantifying any ACI effect. However, this is not what we do here. And there is no physical reason limiting the *consequences* of any ACI on their process-level space and time scales only.

Pg. 9, Lines 18-20: This statement succinctly highlights the lack of depth of the analyses in this manuscript. There are a large number of critical limitations associated with the use of these data to try to draw these sorts of conclusions. The limitations are acknowledged in the manuscript, but there is no real attempt to overcome them. The model analyses discussed in this last sentence hold promise for being able to attribute aerosol to sources as well as to link those aerosol to clouds. I recommend that those tools be brought to bear on the questions being tackled here, perhaps with some context being provided by more complete set of aerosol types in the Level 2 version of the satellite data that is discussed here.

It is unfortunate that the acknowledgement of limitations in a data set and/or method of analysis is characterized as "lack of depth", especially when the referee refers to another kind of study and conclusions. Specifically, in the previous comment, the referee reasons that the data used here are "*less ideal for tackling these sorts of science questions"*, because they are "*highly averaged in both space and time"*. Indeed, these data are inadequate to establish an ACI cause and effect mechanism. This is clearly acknowledged as a limitation, but it was never described as a science question to be tackled in this study. Similarly, using level 2 data which are closer to the ACI process-level in terms of both spatial and temporal resolution, and/or model analyses (especially the latter), could probably lead to robust conclusions on aerosol sources and links with clouds, again, in the process level. In a previous comment, the referee invokes the study by McComiskey and Feingold (2012) to justify the inadequacy of the scales used here and a predefined failure to "*see ACI effects*". That study, however, tackles the question of quantifying the ACIs. This was never a goal in the present study, exactly because this limitation was acknowledged. The same holds for the suggestion, by the referee, of a "*more appropriate data set … (e.g., aircraft, balloon, surface remote sensor scales measuring clouds over minutes to hours or a model with better space and time resolution)*".