Response to Revisor 1

Overall comments to the revisor

The authors would like to express their gratitude for the comments about the manuscript. Below you will find a detailed answer to all comments and how they were addressed in the text. We reproduce Revisor's comments and add our responses in *red italicized font*. In addition to those points, the authors would like to highlight that the changes throughout the text are in the uploaded marked-up version of the manuscript.

Answers to the major comments:

1) Physical consistency of the two clean reference states: the analyses relies on two different clean reference states: (1) a fixed irradiance reference (689.9 W m^{-2}) derived from 19 "cleanest days" across 2014–2015, and (2) a "seasonal" reference derived from two HALO flights (AC09 and AC18) in September 2014, with linearly interpolated irradiance between 689.9 and 767.9 W m^{-2} over the two years. One may wonder how representative they are. To justify the consistency between the two reference states, please consider providing more quantitative evidence demonstrating that the two reference constructions represent physically consistent "clean background". For example, it might be useful to show distributions (e.g. boxplots or PDFs) of LWP, r_eff, CBH, CTH, AOD, and β_sct for the two approaches. After that, you may discuss whether these two sets of days sample similar "natural" aerosol and cloud conditions, or whether the HALO days represent a specific regime (e.g., season, meteorological state, or cloud type).

R: *To assess the similarity of the two representative clean-scenario constructions, histograms, PDFs, and boxplots of the variables indicated by the reviewer were created. These graphs are presented in Figures 1 to 5 and cover the same time interval (between 06:00 and 18:00, local time, Manaus) used in the study. For the cloud-related variables (LWP, r_eff, CBH, and CTH), only those evaluated in the manuscript were considered, i.e., with CTH <= 3000 m and LWC_ground between 0.2 and 0.4 g/m³. Due to the marked absence of AOD measurements (as shown in Figures 1 and 2 of the manuscript), only β_sct was used to characterize the aerosols. The medians and PDFs of CBH and CTH show that in the second clean scenario, the analyzed clouds are relatively higher than in the first, indicating the first instance of the microphysical influence of more scattering aerosols on low-level liquid clouds. The presence of aerosols with higher β_sct in the second clean scenario (median of 16.8, compared to 4.8 in the first) is also linked to the higher medians of r_eff and LWP_ground in this second clean scenario, also indicating the first microphysical changes in clouds due to changes in aerosols. Despite the relative discrepancies between the two clean scenarios, Figure 6 shows that, compared to the days of August and September 2014 (historically the most polluted months in the Amazon), the second scenario can still be considered relatively clean. A sign of this is the PDF shifting even further towards higher β_sct values. This discussion was* *added to topic 4 of the revised version. At this point, it is important to highlight that flights AC09 and AC18 were conducted over two regions practically unaffected by biomass burning plumes from the deforestation arc, characterizing the cleanest possible atmosphere among all flights performed, as indicated by the average values of number concentration of aerosols (N_a) and CCN in Table 1. Despite the low number of samples (two cases), these flights are representative of the cleanest possible atmospheric regime within the polluted season in the region, which is why they constitute the best available samples to characterize the second clean scenario used in the study. This discussion was presented in subtopic 2.3 of the new version of the manuscript.*

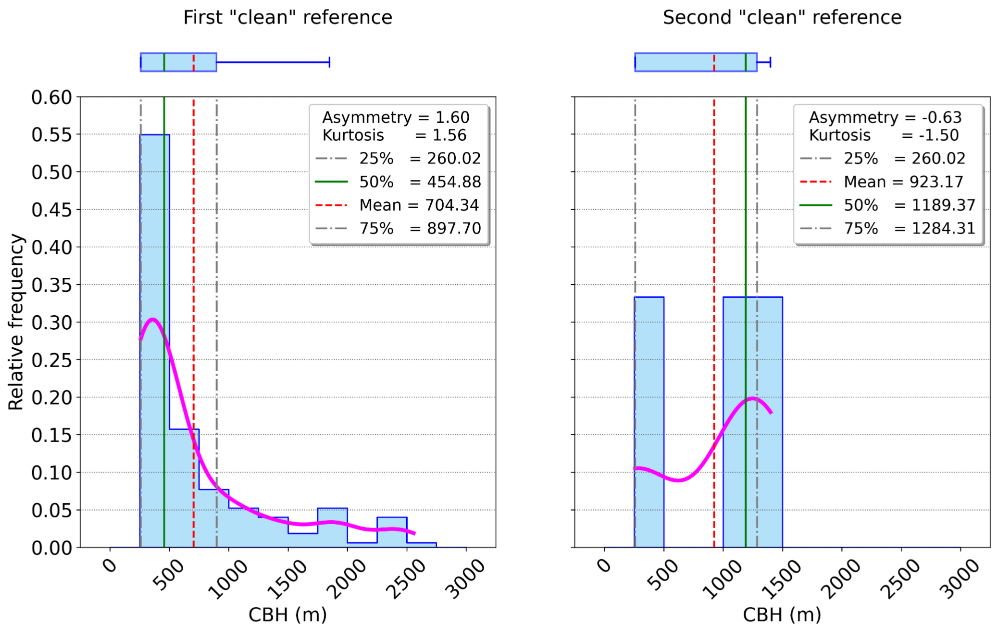

Figure 1: Histograms, PDFs, and statistics of CBH for the two clean scenarios.

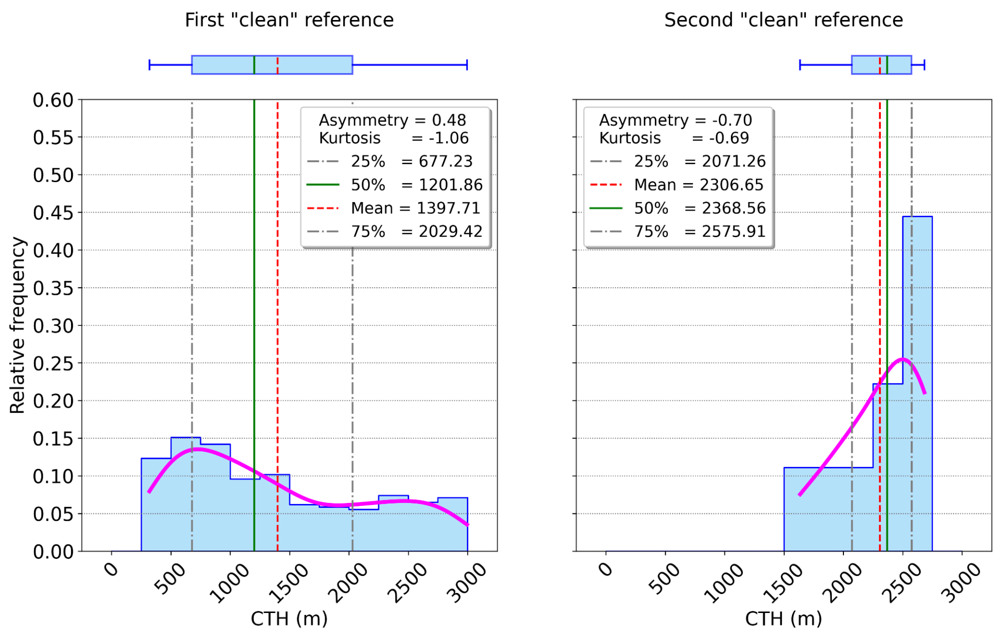

Figure 2: Histograms, PDFs, and statistics of CTH for the two clean scenarios.

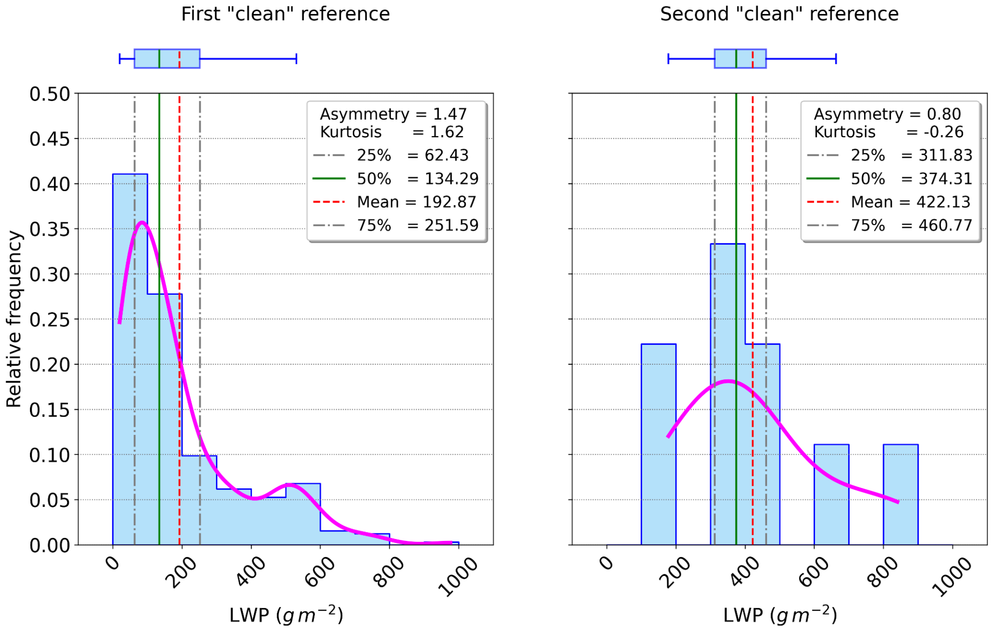

Figure 3: Histograms, PDFs, and statistics of LWP for the two clean scenarios.

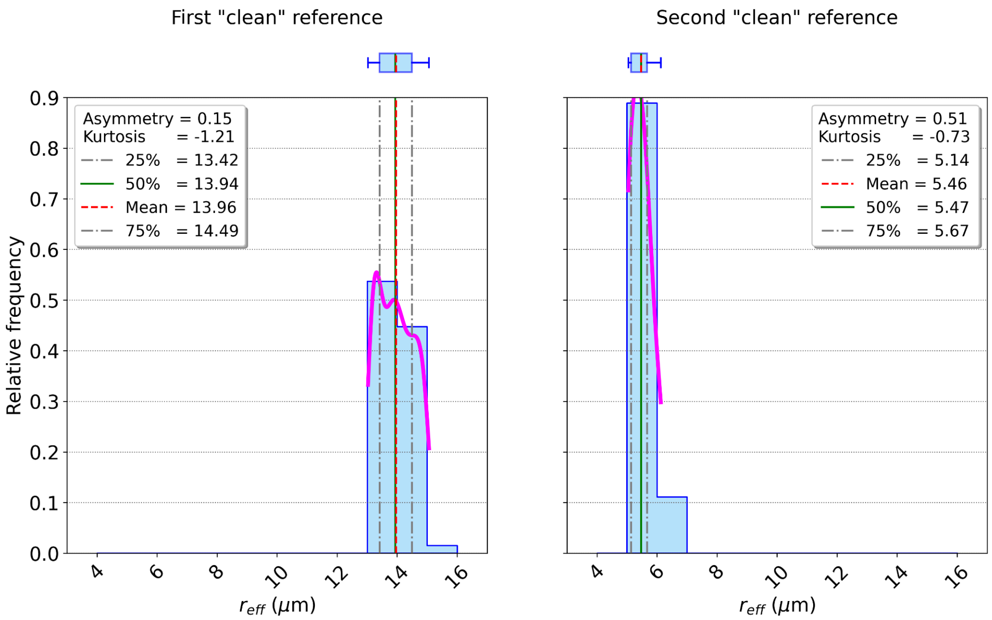

Figure 4: Histograms, PDFs, and statistics of r_eff for the two clean scenarios.

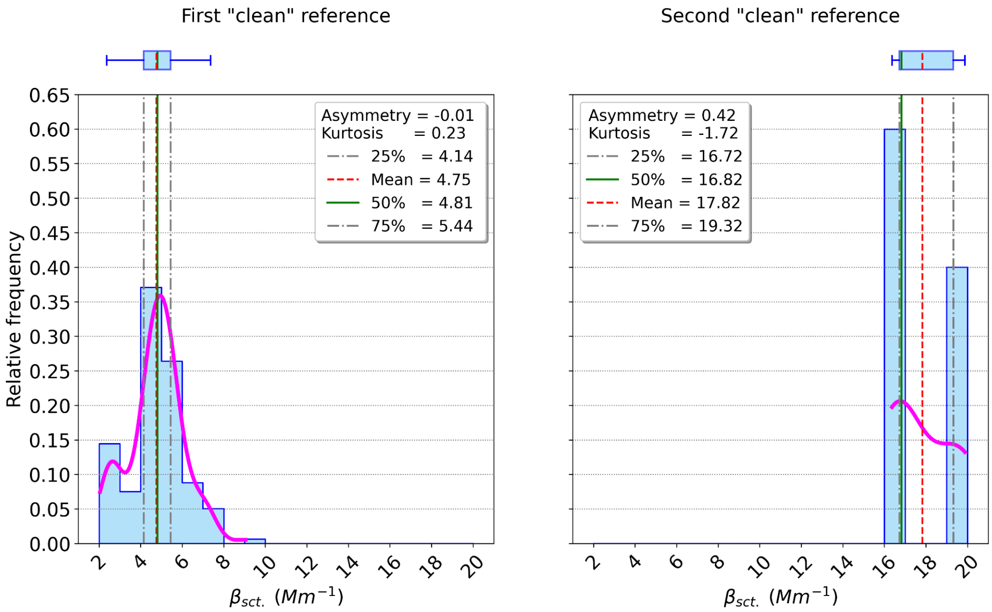

Figure 5: Histograms, PDFs, and statistics of β_sct for the two clean scenarios.

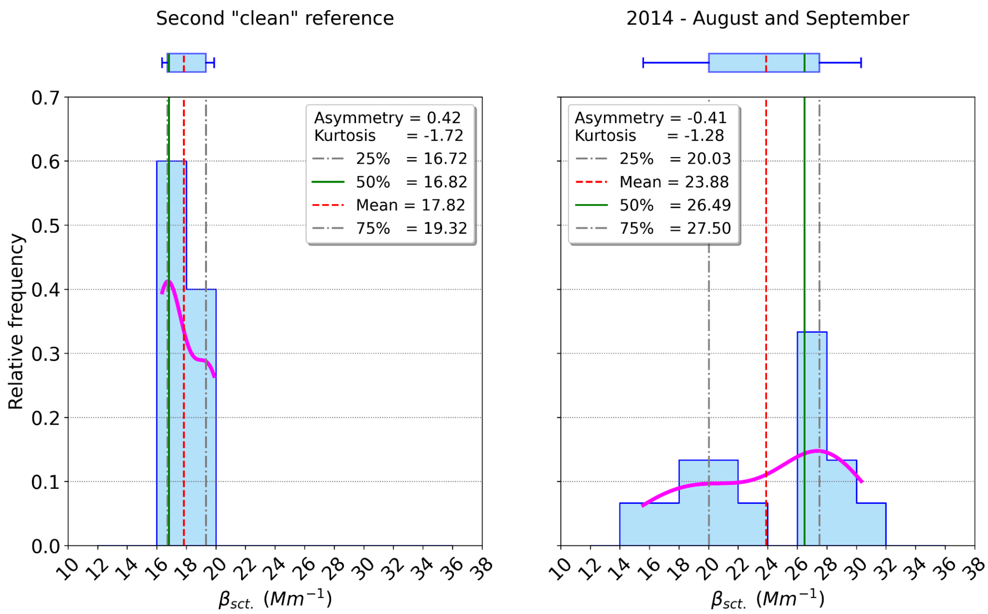

Figure 6: Histograms, PDFs, and statistics of β_sct for the second clean scenario and for days of August and September, 2014.

Table 1: Average concentrations of N_a and CCN measured during some HALO flights during 2014. Adapted from CECCHINI et al., 2017.

Flight	*N_a (cm^{-3})*	*CCN (cm^{-3})*
AC18	*744*	*408*
AC09	*821*	*372*
AC07	*2498*	*1579*
AC11	*2691*	*1297*
AC12	*3057*	*2017*
AC13	*4093*	*2263*

2) Seasonal reference and linear interpolation (Fig. 4): the seasonal reference is constructed by linearly varying the clean reference irradiance between 689.9 and 767.9 W m^{-2} as a function of calendar day (Fig. 4). This assumes a linear evolution of "natural" clean-state TOA flux through the year. Please justify the linear interpolation more explicitly. For example, can you show that when only the cleanest days are considered, the observed F_corr or r_eff displays an approximately linear seasonal evolution?

R: *The linear variation of radiation flux in the TOA for constructing the seasonal reference (according to a clean atmosphere scenario) constitutes a first approximation to consider the natural variation of the microphysical characteristics of the clouds evaluated in the study. For this purpose, the diurnal medians of r_eff of these clouds (CTH <= 3000m and LWC between 0.2 and 0.4 g/m³) were plotted for the cleanest days within the clean and within the polluted periods of 2014. These days are represented by the red dots in Figure 7, while the blue dots refer to the transition days. Although not showing a perfectly linear variation between the cleanest days, Figure 7 shows that the natural variability of r_eff occurs in an approximately linear manner throughout 2014, justifying the linear approximation used. This discussion was included in section 2.8 of the revised manuscript.*

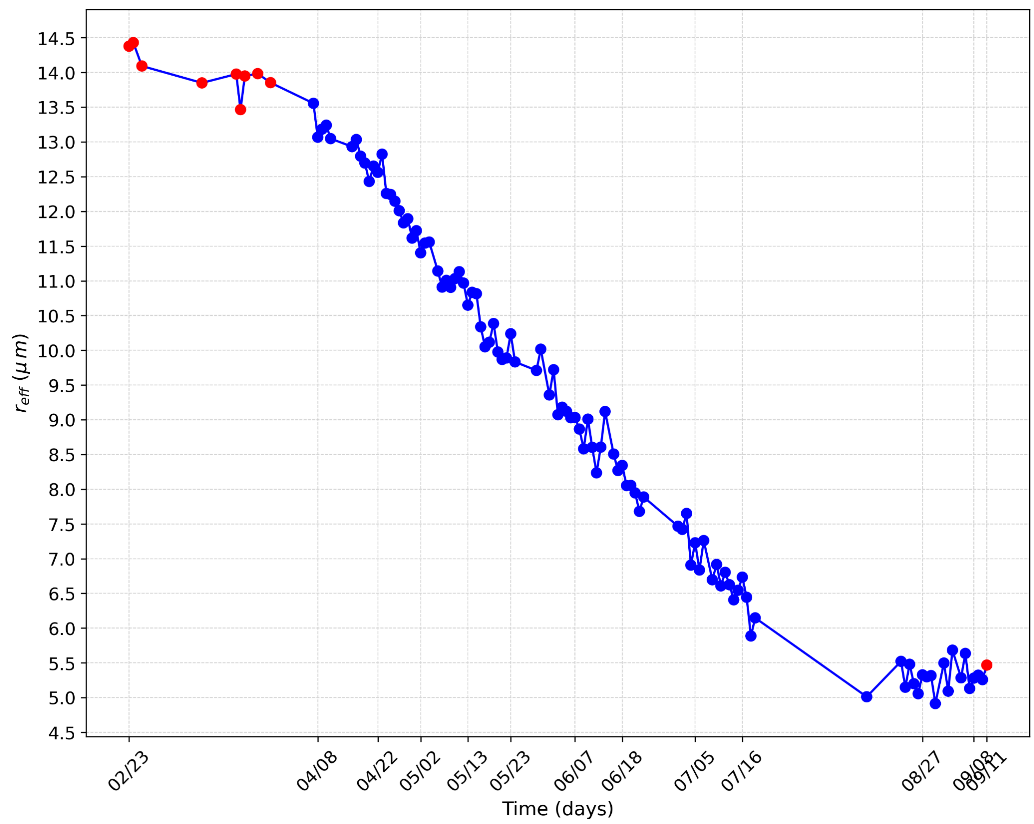

Figure 7: Diurnal medians of r_eff during 2014. The cleanest days within the clean and polluted seasons are represented by the red dots, while the bluish dots correspond to the transition days.

3) Comparability with IPCC Twomey ERFaci estimate: you mention that the campaign-mean IRFaci for the second reference (−1.3 W m^{-2} with IQR −5.8 to 0.59 W m^{-2}) is close to the IPCC AR6 estimate of Twomey ERFaci (−0.7 ± 0.5 W m^{-2}). Please clarify that your quantity is a regional, top-of-atmosphere, daytime-only IRFaci for warm clouds, whereas the IPCC value is a global, all-sky, annual-mean effective radiative forcing, including adjustments (changes in cloud fraction, lifetime, etc.). I'd recommend rephrasing the comparison to emphasize that the agreement is in order of magnitude and that your results are best viewed as observation-based benchmarks for models in the Amazon rather than direct constraints on the global ERFaci.

R: *The estimate reported by the IPCC and mentioned in the text (−0.7 ± 0.5 W m^{-2}) does not refer to the effective forcing (ERFaci), but rather to the average value of the instantaneous forcing, as in section 7.3.3.2.1 of the AR6 WGI. The value including the rapid cloud adjustments (ERFaci) reported by AR6 is −1.0 ± 0.7 W m^{-2}. However, all comparative sections between the obtained result and the IPCC's IRFaci estimate have been rewritten. As suggested, this was done to emphasize and clarify that the result referring to the second clear atmosphere reference is regional, for the top of the atmosphere, for low-level liquid clouds, and only for the daytime period, while the IPCC result refers to a global average. The phrase "order of magnitude" was also included in the text changes, as suggested. All these changes can be seen in the marked-up version.*

4) Daily slab cloud representation and sub-daily variability: warm clouds are represented as a single "daily" two-layer slab, with base/center/top heights derived from CBH/CTH percentiles and LWC/r_eff percentiles. All radiative transfer is computed at local noon assuming full cloud cover (fc = 1), and later scaled by the daily mean fc. Please discuss more explicitly that this is a first-order approximation that neglects sub-daily variability in cloud fraction, LWP, and r_eff, as well as differences between cumulus and stratiform elements.

R: *We agree that the use of statistics such as mean and percentiles constitutes a low-dimensional statistical description. Such statistics are adequate and were used to characterize the central tendency and dispersion of the variables mentioned, although they are insufficient to explicitly capture intraday variability or transient processes. This observation was included in section 2.6 of the revised version of the manuscript. Regarding cumuliform and stratiform clouds, we also agree that, despite the attempt to include both types in the terms "warm clouds," "low warm clouds," and "low-level warm clouds," a clear separation between these elements was not made. Due to this, we standardized all the previously mentioned terms to "low-level liquid clouds" and made it clear that this term includes both shallow cumulus and low stratiform cloud types. These modifications are present throughout the updated version of the manuscript, as well as in the title, where "warm" was replaced with "low-level liquid."*

5) Robustness of IRF sensibility analysis: the manuscript concludes that low warm clouds in the Amazon may be more sensitive to aerosol perturbations under clean conditions than under already polluted conditions. This conclusion is primarily obtained based on Table 5 which relates |mean IRFdaily| to percentage changes in AOD for clean vs. polluted periods. While it aligns with recent reports that ERFaci uncertainty is dominated by clean conditions, the metric mixes absolute |IRF| in one regime with relative AOD changes between regimes, and the denominator (ΔAOD%) ties to arbitrary period definitions. Moreover, there are no uncertainty bars or significance tests provided along with the sensitivity values (e.g., 12.12 vs 4.08 W m^{-2}/ΔAOD% in 2014). Given relatively small numbers of days and the presence of positive IRFdaily tails, it is not clear that clean vs. polluted sensitivities are statistically distinguishable.

R: *We thank the reviewer for this question, as it made us realize that the entire analysis regarding the sensitivity of daily instantaneous forcings to AOD should be redone with greater statistical rigor. Considering this, we present here the step-by-step process of the new analysis, which is found in subtopic 3.3 of the revised version of the manuscript: First, we calculated the daily medians of AOD for the clean and polluted scenarios of each year, and for each of these scenarios we selected the 10 days with the lowest medians. All days of a scenario were used to calculate the average AOD (for the clean scenario of 2014, for example, we called this average "T_clean_14"), while another average was calculated for the 10 cleanest days selected previously ("t_clean_14" for the clean scenario of 2014, for example). Next, we used the equation ΔAOD% = 100 * (T - t) / t to obtain the percentage variation of AOD within each scenario, taking t as the "clean" reference and seeking to determine by what percentage the average T is greater than t. In order to test the hypothesis ΔAOD%_clean < ΔAOD%_polluted for a given year, we applied a bootstrap test and obtained the statistical significance (at the 5% level) of the percentage variations for each year. The results, also present in the revised version of the manuscript, show that the hypothesis was confirmed only for 2014. As with AOD, the average IRFdaily values for each scenario and for the 10 cleanest days within each scenario were calculated (T_IRF_polluted_14, t_IRF_polluted_14, for example). The difference between these averages (ΔIRFdaily = T_IRF - t_IRF) was used for the new sensitivity calculation S = ΔIRFdaily/ΔAOD% (units of $Wm^{-2}\%^{-1}$), for each scenario of each year. The last step consisted of performing a new statistical significance test, considering the hypothesis that S in the clean scenario is greater than in the polluted one. The results (also at the 5% level) showed that this only occurred in 2014. A table with these new results replaces Table 5 in the new version of the manuscript, as well as new discussions in sections 3.3 and 4.*

Answers to the specific comments:

1- L85: Have you cross-validated the LWP derived via the algorithm proposed by Turner et al. (2007)? Would the relatively smooth vertical gradients in the temperature and moisture profiles obtained by the method impact the quality of LWP retrievals?

R: *We thank the reviewer for this question, which was also pointed out by the other reviewer. The LWP data measured by the MWR radiometer could, in principle, be validated from in-situ LWC measurements in clouds made during flights over site T3, when co-located spatially and temporally with cloud radar and lidar (WACR + MPL) measurements, used to estimate CBH and CTH. However, there were no cases of simultaneous measurements from the three instruments at both IOPs, and for this reason we could not perform cross-validation of the LWP data. This observation was included in the second paragraph of section 2.2 of the new version of the manuscript. Regarding the second question, the temperature and humidity profiles used in the Turner et al. (2007)'s method do not interfere with the quality of the LWP retrievals, because they are used in independent steps of the method. One observation regarding this was also included in the second paragraph of section 2.2.*

2- L118: Could you clarify if only samples taken just above the T3 site were used in the study?

R: *For G1, only measurements taken within a 0.2° x 0.3° area (60.75° W, 3.3° S, 60.45° W, 3.1° S) centered on T3 were used. For the HALO aircraft, all measurements from flights AC09 and AC18, which occurred over densely forested areas north and northwest of Manaus, were considered. Although not located exactly above T3, these regions represented the "cleanest" data source available during IOP 2 over the forest. This explanation was included at the beginning of the second paragraph of section 2.6 of the revised version.*

3- L265: Related to major comment #4, Is there any justification for this assumption? Moreover, what's the sensitivity of simulated clouds to the value of fc?

R: *The assumption of fc=1 in all simulations was made because, considering that the simulations require long computation periods, adding variable values of fc would cause new sets of simulations to be performed, greatly increasing the time to obtain the results. The attempt to include the sensitivity of the simulated clouds in relation to the average variation of fc observed over T3 is already part of the results shown in section 3 of the manuscript.*

4- Figure 5: Might be useful to again mark the periods of clean and polluted periods.

R: *The suggested change was made to Figures 5 and 10 to maintain standardization throughout the text. Available in the revised version of the manuscript.*

5- L333: There might be a typo in the title of Section 3.2 ("Seasonal")

R: *Done. Available in the revised version of the manuscript.*

6- L374: It's good to mention the interannual variability in IRFaci here. However, one would wonder what the reasons could be causing the distinct variation. Please consider elaborating it more as this may help generalize the approaches used in this study for other periods.

R: As suggested, the interannual variabilities of the IRFaci averages for the two clear atmosphere references were mentioned at the end of the second paragraph of section 4. This interannual variability of IRFaci depends on the aerosol load in the atmosphere, the occurrence of which varies significantly each year, depending on meteorological factors such as the occurrence of droughts, or social and economic pressures, and also on public policy decisions. For the generalization of the approach used in this work to other periods, it is therefor fundamental to adequately record the daily and seasonal variations in the properties of aerosols and clouds, as illustrated in Figs. 1 and 2 of the manuscript. This discussion was included in the fourth paragraph of section 5.

7- L375: Can you elaborate more on what do you mean by the microphysical structure of clouds here? may vary based on the existing statistics?

R: The term "microphysical structure" for the modeled clouds refers to the representative statistics of the daily r_eff and LWC distributions used, and does not vary according to the clean atmosphere references used for the forcing calculations. Considering this, this section of the text has been rewritten more accurately.

8- L385: This suggests the assumption in irradiance and algorithm for derivation of reference state can significantly modulate the IRF values. Therefore, as given in the major comments, it is critical to ensure the robustness of the proposed algorithms.

R: As answered in major question 2), the robustness of the references for a clean atmosphere is analyzed using Figure 6 presented above. This figure shows that, compared to the historically most polluted months in the Amazon, the second scenario can still be considered relatively clean. This discussion was added to topic 4 in the revised version.

9- L415 and L423: I guess fc=1 is identical with fc=100%. If so, please use the consistent way of expression throughout the manuscript to avoid potential confusion.

R: *Yes, "fc=1" is identical to "fc=100%". All the occurrences of "fc=1" throughout the text have been changed to "fc=100%".*

10- L430: It's essentially above and near the T3 site. In terms of the region over the Amazon, most likely there is large spatial variability which is not addressed in the study. One may wonder how the spatial variability may alter the overall conclusion made here. Please clarify.

R: *We agree with the reviewer. We modified the excerpt "The results obtained stem from an effort to estimate instantaneous radiative forcings entirely from in situ and ground-based remote sensing data collected over the Amazon" to "The results obtained stem from an effort to estimate instantaneous radiative forcings entirely from in situ and ground-based remote sensing data collected over a specific region of the Brazilian Amazon".*

11- L435: Please list all the applicable datasets specifically in the "Data availability" section.

R: *Done. A link (https://doi.org/10.5281/zenodo.20610494) to a compiled folder containing all the datasets used has been made available in the "Data availability" section of the revised version.*