Author Responses to reviewers' comments

We thank both reviewers for carefully reading and providing insightful feedback to our paper, which helped us improve the manuscript. The major changes to the manuscript are listed below.

1. We modified the title from the broad scope of "Impacts of the Three-dimensional Radiative Effects on Cloud Droplet Number Concentration Retrieval and Aerosol Cloud Interaction (ACI) Analysis" to the more limited scope of "Impacts of the Three-dimensional Radiative Effects on Cloud Droplet Number Concentration Retrieval and Albedo Susceptibility Estimates."
2. For calculating albedo susceptibility, we now compute albedo (α) from hemispheric integration of upwelling radiance, so anisotropy is included explicitly. Thus, we no longer use broadband nadir Reflectance (R_{BB}) as a proxy for albedo.
3. The rest of the minor suggestions from the two reviewers were addressed accordingly throughout the manuscript.

Please find our point-by-point response to the referee's comments as well as changes made below (reviewer's comments are in black, authors' responses are in blue, *while changes made to the revised manuscript are in red italics*). We believe these revisions have adequately addressed the reviewer's comments and welcome any additional suggestions or comments from the referees and editor.

Reply to Reviewer 1:

General Comments: This study examines uncertainties in the most widely used remote sensing retrievals of droplet number concentrations and their implications of albedo susceptibility calculations, which are a core part of the analysis of aerosol cloud interactions and highly relevant to our understanding of climate. The authors make a valuable novel contribution in this study by using Large Eddy Simulations and 3D radiative transfer (3DRT) models and so can provide end-to-end analysis of errors. They also claim to have analysed albedo susceptibility but there are significant flaws in this analysis for which I recommend major revisions. I believe that once these are addressed that this article will make a valuable contribution to ACP.

Specific Comments

“ACI analysis” in the title & abstract etc. is far too broad for what is covered in this paper. This should be narrowed to “albedo susceptibility estimates”.

Reply: We agree with the reviewer that "ACI analysis" is too broad. Since “albedo susceptibility” is an important criterion for quantifying the radiative forcing component of ACI. We have modified the title of the manuscript to reflect this narrow scope. The updated title is thus *"Impacts of Three-dimensional Radiative Effects on Cloud Droplet Number Concentration Retrieval and Albedo Susceptibility Estimates."* We have also updated the text in the abstract and manuscript, where necessary, to reflect these changes.

1. Anisotropy & Flux:

Here, I will be blunt, if only to impress upon the editor the importance of the issues I raise. This study neglects the anisotropy of radiation while making claims about a radiative flux (and its susceptibility) in Section 2.4. There is also a lack of a reference calculation of albedo susceptibility (discussed further below). These omissions render the results in this section largely meaningless.

A large part of the 3DRT effect on TOA irradiances is changes in the anisotropy of radiation with optical depth. For example, for near overhead sun, nadir reflectance saturates more quickly for finite clouds (e.g., a cube) than the oblique views. In other words, the angular pattern of the susceptibility changes from 1DRT to 3DRT. The effects of 3DRT on albedo susceptibility cannot be estimated while neglecting the anisotropy of radiation. The burden of proof is on the authors to demonstrate that the anisotropy susceptibility is unimportant for their analysis.

This study presents several different estimates of albedo susceptibility and compares them in Section 2.4. For example, a 1DRT+regression estimate applied to measurements generated using 1DRT and measurements generated using 3DRT (at 100 m or 800 m resolution) are compared. None of these four estimates are the truth and so it is not possible to tell whether agreement between estimates indicates good or poor performance. Without a clear reference, this makes it easy for a reader to come to an incorrect conclusion about what these results imply about the accuracy of using 1DRT to estimate albedo susceptibility.

The derivative of TOA albedo (averaged over some reference area) with respect to optical depth in each column has an exact expression in terms of linearized 3DRT (Doicu and Efremenko, 2019) that should be used as the reference for evaluating all approximations of albedo susceptibility.

These elements need to be present or other supporting evidence included to justify the claims presented here. Either the claims about the efficacy of estimating albedo susceptibility using 1DRT need to be removed or appropriate evidence needs to be added.

I note that the calculation of reference-quality albedo susceptibility will require a greater application of technical effort and computational power than the entire contents of the manuscript, which is why I suggest that it be removed.

Reply: We agree that 3D effects on albedo susceptibility cannot be thoroughly assessed using a single-view reflectance proxy without accounting for anisotropy. Therefore, in the revised manuscript we therefore compute top of the domain (TOD) albedo directly from the radiative transfer solution by hemispherically integrating the upwelling radiance at the top of the RT domain for both 1D and 3D RT, so anisotropy is included explicitly by construction. This change is reflected around L228 to 247 in the revised manuscript.

"The CERES TOA flux is obtained from directional radiance measurements using scene-dependent anisotropy factors (often described via angular distribution models, ADMs), and therefore the CERES radiance to flux procedure explicitly accounts for the angular redistribution of radiation. In this study, rather than approximating fluxes from a single viewing direction or adopting an ADM assumption at LES scales, we compute the shortwave flux and albedo directly from radiative transfer by hemispherically integrating the upwelling radiance (I) at the top of the RT domain over a discrete angular grid. The integration is performed using radiance simulations at 37 viewing zenith angles (θ_v) spanning 0–90° in 2.5° increments and 16 viewing azimuth angles (ϕ_v) spanning 0–360° in 22.5° increments. The upwelling flux ($F^{\uparrow}$) at solar zenith angle θ_o is defined as:

$$F^{\uparrow}(\theta_o) = \int_0^{2\pi} \int_0^{\pi/2} I(\theta_v, \phi; \theta_o) \cos\theta_v \sin\theta_v d\theta_v d\phi, \quad (1)$$

where ϕ is the azimuth integration variable, $\phi = \phi_v - \phi_o$ is the relative azimuth angle, with ϕ_o and ϕ_v denoting the viewing and solar azimuth angles, respectively.

To ensure accurate flux and albedo estimates from the discrete angular sampling, the hemispheric integration was evaluated using finite-volume angular quadrature, in which each *viewing-zenith bin was weighted by the exact integral of $\cos\theta_v \sin\theta_v$ over its bin edges and each azimuth bin was weighted by its corresponding $\Delta\phi$. The resulting hemispherically integrated upwelling flux was then normalized by the incident solar flux at the top of the domain to obtain the broadband albedo (α), defined as:*

$$\alpha(\theta_o) = \frac{F^{\uparrow}(\theta_o)}{F_o \mu_o} \qquad (2)$$

where F_o is the incident solar flux at domain top and $\mu_o = \cos\theta_o$."

We also add an empirical reference susceptibility by computing a relative albedo susceptibility from the 3D-RT albedo field and LES microphysical fields, conditioned on LWP LES bins. While we acknowledge that exact column-wise Jacobians from linearized 3D RT (e.g., Doicu and Efremenko, 2019) would provide an even more formal RT-derivative reference, implementing a linearized/adjoint solver is beyond the scope of the present forward-modeling framework.

Other specific flux issues:

Line 211-212: CERES is an estimate of F3D, it is not measuring a fundamentally different quantity. Yes, there will be errors. By construction an empirical ADM method with a one-dimensional scene ID will produce a high-variance estimate of the scene's contribution to the flux. Note that the factor mapping from flux to radiance is the anisotropy factor, not the ADM. The ADM is a model for the anisotropy factor (or correction).

Reply: We agree with the reviewer that CERES aims to estimate the same physical quantity—TOA hemispheric flux—rather than a fundamentally different flux. In the revised manuscript we have removed wording implying that F_{3D} and the CERES flux represent different quantities, and we now clarify that the CERES procedure estimates TOA flux from measured directional radiance using an anisotropy factor, with the ADM serving as a model for that anisotropy factor. This can be found around L228 to 231 of the revised manuscript (highlighted in the reply to the previous comment above).

We also remove statements suggesting the anisotropy factor is constant over our LES domain. Instead of adopting a CERES-like ADM at sub-kilometer scales, we compute TOD upwelling flux and albedo directly from the RT solution by hemispherically integrating upwelling radiance, thereby avoiding assumptions about scene-dependent anisotropy corrections while still producing the same target quantity (TOD flux/albedo).

Line 213: This does not make sense to me. CERES flux is computed over on the order of 20 km FOVs, CERES does not measure at the scales of pixels in the LES simulations. Moreover, each CERES flux estimates a contribution to the flux at TOA from a localized region (the ~20 km FOV), the radiation from that region reaches TOA spread over a ~hundred km region. See, for example, the difference between the MISR restrictive and expansive albedos. F_{3D} is the flux at TOA, and so of course it is spatially smoothed as the upwelling flux at each horizontal location is sourced from a huge area on the surface, due to integration of upwelling radiance over angle.

Reply: We agree with the reviewer that CERES fluxes are reported for (20 km) footprints and that it is not appropriate to interpret CERES as measuring "pixel-scale" fluxes comparable to the 100 m LES grid. We have therefore removed the text implying a pixel-to-pixel CERES-like flux at LES resolution and the associated discussion contrasting F3D with FBBCERESF. In the revised manuscript we compute TOA albedo directly from hemispheric integration of upwelling radiance for both 1D and 3D RT, and we interpret these albedos as RT-consistent quantities defined on the LES domain rather than as emulations of CERES footprint flux retrievals. We also now state explicitly that α(3D RT) at a given horizontal location is inherently nonlocal ("radiatively smoothed") because hemispheric integration causes the upwelling flux at TOA to include contributions from surrounding columns via horizontal photon transport and multiple scattering. Our focus is therefore on how 3D radiative transport and this nonlocality affect the albedo–NdN_dNd relationship and susceptibility diagnostics within the modeled domain, rather than on reproducing the CERES observational sampling geometry.

Line 219-223: It sounds like the CERES-type flux referred to in this paper is a non-standard quantity that warrants a clear precise definition.

Reply: In the revised manuscript we have remove the CERES-type flux construct and do not attempt to emulate the CERES footprint retrieval procedure at LES scales. Instead, we compute TOA upwelling flux directly from the radiative transfer solution by hemispherically integrating the upwelling radiance at the top of the RT domain and define TOD albedo albedo.

Line 223-224: "The ADM is the same for all cloudy pixels." This is incorrect.

The ratio between flux and radiance (the anisotropy factor) varies everywhere in space and angle and is defined everywhere within the 3DRT domain. It is not constant over a scene or small LES domain. CERES approximates the anisotropy factor as constant over a large (order ~20 km) region. The definition for the CERES flux relationship (i.e., between a localized region in a FOV and a flux spread over the TOA) can be applied at higher spatial resolution. This is exemplified in the MISR restrictive albedo and local albedo products, the latter of which is a TOA flux product at 2.2 km resolution (though not broadband). The difficulties here are in the geometric definition of the reference levels for registering TOA flux contributions to, and the ability to produce a model for the anisotropy. This becomes increasingly difficult at high resolution (e.g., ~100 m) because there is no one-size-fits-all definition of a reference level for which we can construct useful low variance models for anisotropy. These issues are irrelevant when the cloud is specified and the albedo susceptibility can be calculated exactly using appropriate numerical modelling.

Reply: Thank you for the comment. This sentence has been removed from the revised manuscript.

Line 733-734: Its not flux, it's a broadband radiance. Given fixed profiles of absorbing gases (or their absence), it would be weird if broadband radiance and bi-spectral radiances weren't highly correlated.

Reply: We agree with the reviewer that our previous wording at lines 733–734 was imprecise: the quantity discussed there was a broadband reflectance at nadir, not flux. In the revised manuscript we no longer refer to nadir reflectance as "flux." Our updated methodology computes TOD upwelling flux and albedo directly by hemispherically integrating the upwelling radiance, our susceptibility analysis is now based on albedo (α) rather than a directional broadband radiance proxy.

Line 737-740: These statements are not justified based on the evidence presented either in the paper or with appropriate references.

Reply: We agree that the statements in lines 737–740 were not adequately supported in the original manuscript. In the revised version we have (1) revised the text to avoid claims that cannot be directly demonstrated from our analysis, and (2) added explicit supporting evidence where the conclusion remains. Specifically, susceptibility results are now computed using TOD albedo derived from hemispheric integration of upwelling radiance (rather than a directional reflectance proxy), and we include a clear benchmark ("truth") susceptibility computed from α(3D) and LES microphysical fields conditioned on the LES LWP.

Line 829-831: This is unjustified, again, as 3D effects on albedo susceptibility have not actually been calculated.

Reply: We agree that the original statement in lines 829–831 was not justified because the manuscript did not explicitly compute albedo susceptibility from 3D RT albedo. In the revised manuscript, we now calculate TOD albedo (α) directly from hemispheric integration of upwelling radiance for both 1D and 3D RT and then compute the relative albedo susceptibility using the simulated α and N_d for a constrained LWP within the LWP bins defined in Table 1. This update allows us to quantify 3D effects on albedo susceptibility directly by comparing susceptibility estimates based on α under 3D RT (α(3D RT)) and cloud properties retrieved from 3D RT reflectance directly with its 1D counterpart, and also by comparing the retrieval-based estimates against a "truth" susceptibility derived from α(3D RT) and LES microphysical fields. The new results still show that the disparity between relative albedo susceptibility from 1D and 3D RT at 100m resolution reduces at a more coarse footprint of 800 m.

Reference (Zuidema et al., 2008) and discuss.

1. Droplet Concentration

My challenge to the authors is to better explain their choice of definition for droplet concentration retrieved from bispectral retrievals, as it differs from how such retrievals are used in practice. So, for general readers to get useful information from the study, I think it would be beneficial to directly engage with this issue and address it clearly in the text.

Eq. 1: Why should anyone care about this definition of vertically-averaged Nd? As noted above this Equation in the text, the reference Nd for in situ studies is the average of Nd across the vertical extent of the cloud (or the Nd at cloud base). The vertical averaging of Nd is also the reference used for field validation of remote sensing retrievals of droplet number concentration (Gryspeerdt et al., 2022; Painemal and Zuidema, 2011). Most of the theory used in ACI analysis uses a simplified adiabatic framework in which Nd is constant with height (not accounting for expansion with height i.e., geometrically thin layers). The reasons for these definitions are simple; to isolate the effects of droplet concentrations from other effects (e.g., dynamical and thermodynamic), even if through an incomplete characterization of the 3D Nd field using assumptions that introduce errors. The isolation of droplet concentration is critical to ACI analysis. By using an Nd weighted by extinction coefficient as a reference, this utility is lost, so by choosing this reference there is no traceability to existing ACI analysis.

So, why should I worry whether retrievals agree with the extinction weighted Nd_LES, or not?

If there is a need to change the definition/interpretation of remotely-sensed droplet concentration from quasi-adiabatic to retrieving Eq. 1, then this is an important point that needs to be clearly communicated to the communities that use these retrievals. If this is the case, then I suggest making that argument be the focus of the paper.

For example, the argument might go: “Due to our inability to specify appropriate microphysical profile information, we are unable to usefully bound uncertainty in droplet concentration and therefore must restrict the definition to an extinction-weighted quantity and limit the theoretical insight that is available.”

If the authors are not taking a strong position on how droplet concentration ‘retrievals’ are interpreted, it would be useful for them to cleave to the existing definitions and thereby provide useful error characteristics for people using bispectral retrievals to infer droplet concentration. As is, the presented results have limited utility for refining the interpretation of existing studies. The same limitation holds for Kokkola et al. (2025). Gonzalez et al. (2025) take a different approach and should be discussed.

References: (Gonzalez et al., 2025; Kokkola et al., 2025) (see bottom for full citation).

Reply: We thank the reviewer for this important comment. We agree that an extinction-weighted N_{d_LES} reference is not equivalent to the vertically averaged or cloud-base N_d commonly used in in situ validation and ACI studies. In the revised manuscript, we now clarify that our use of the extinction weighted average as our N_{d_LES} is not intended to redefine satellite-retrieved N_d as a general microphysical quantity. Rather, it is used as a radiatively weighted LES benchmark that is more directly aligned with the passive retrievals and albedo-susceptibility calculations examined in this study. We also explicitly state that this choice limits the interpretation of our results: agreement with N_{d_LES} indicates consistency with the optically dominant portions of the cloud column, but does not necessarily imply agreement with vertically averaged or cloud-baseN_d. We have revised the text to make this distinction clear and to better connect our definition to the intended radiative application of the manuscript.

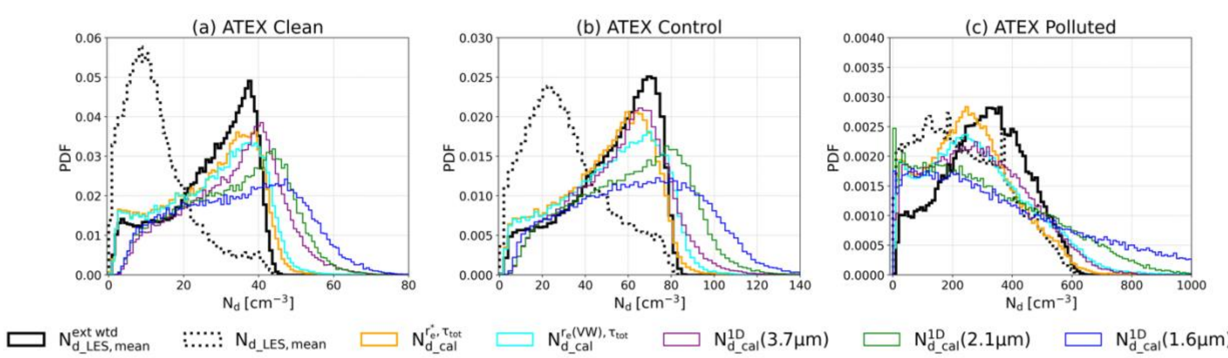

Figure R1. Probability Density Functions (PDFs) of obtained from the LES ($N_{d\_LES,mean}^{ext\ wtd}$, $N_{d\_LES,mean}$) and calculated from the adiabatic equation using 1D cloud properties retrieved from 1D RT simulated reflectance (for absorbing channels 1.6 , 2.1 and 3.7), and under different levels of controls across all time steps for the ATEX clean, control and polluted cases in (a), (b) and (c) respectively.

Figure R1 (not included in the revised manuscript) shows that the retrieval-derived N_d PDFs are more consistent with the extinction-weighted LES reference ($N_{d\_LES,mean}^{ext\ wtd}$) than with the unweighted layer-mean ($N_{d\_LES,mean}$), indicating that the retrievals are more strongly linked to the optically dominant portions of the cloud column. We therefore retain $N_{d\_LES,mean}^{ext\ wtd}$, (for simplicity written as N_{d_LES} in the manuscript) as the main radiative reference while explicitly acknowledging that it should not be interpreted as a replacement for cloud-base or vertically averaged N_{d_LES} used in process-level albedo susceptibility and ACI studies. A statement regarding this shown below has been put around L 368 to 376 of the revised manuscript.

"We emphasize that N_{d_LES} is not intended to replace the vertically averaged or cloud-base N_d definitions commonly used in in situ and field-validation studies. Rather, it is used here as a radiatively weighted LES reference that is more directly aligned with the passive retrievals and broadband albedo analyzed in this study. Therefore, agreement between retrieval-derived N_d and N_{d_LES} should be interpreted as agreement with a radiatively effective cloud-column quantity, not necessarily as agreement with the vertically averaged or cloud-base N_d. This choice limits the theoretical interpretation of the retrieved N_d as a purely microphysical quantity, but it provides a more appropriate reference for evaluating how retrieval errors propagate into albedo susceptibility estimates."

Line 159: Cumulus rising into stratocumulus is rather complex microphysically. Should this cloud field be considered representative? How can the veracity of the cloud microphysical properties in the simulations be evaluated? Some further supporting evidence or commentary on this aspect should be included.

Reply: Due to the microphysical complexity of cumulus rising into broken stratocumulus, the sentence "These LES are representative of a trade wind cumulus regime, characterized by scattered cumulus clouds rising into a thin, broken stratocumulus layer." has been removed from the revised manuscript.

Line 288: How is the CBH and CTH defined? i.e., how do you separate between and cloudy and clear volumes in the model? I ask because the order of magnitude difference between the extinction-weighted and pure droplet concentration averages in the appendix suggest a very strict criterion for clear sky. Commonly, these definitions are designed to pursue coherence between models and in-situ measurements and are based on the sensitivity of the in-situ measurements; e.g., 0.01 g/kg or 10#/cm^3 etc. as in other studies that examine microphysical effects on susceptibility equations (Feingold et al., 2022) (This reference should be discussed in the text). I suggest that this point be examined in more detail as the order-of-magnitude differences seem larger than the errors in Eq. 3 estimated from other LES or from in-situ measurements (Feingold et al., 2022; Grosvenor et al., 2018).

Reply: We thank the reviewer for this comment. The domain cloud boundaries were diagnosed from the liquid water content field using a threshold of 0.001 gm^{-3}. To obtain the domain cloud base and top, the maximum LWC at each height was first computed over all horizontal grid points, and the lowest and highest levels exceeding the threshold of 0.001 gm^{-3} were taken as the domain CBH and CTH, respectively.

The boundaries of the cloudy column were also diagnosed in a similar way, but applied independently to each column, with CBH defined as the lowest cloudy level and CTH as the highest cloudy level in that column.

The vertical average in the initial plot had used the average vertical average across the domain extent thus was responsible for the order of magnitude difference observed as stated by the reviewer. When we average correctly from our prescribed CTH to CBH as used in our study, the updated plot is below.

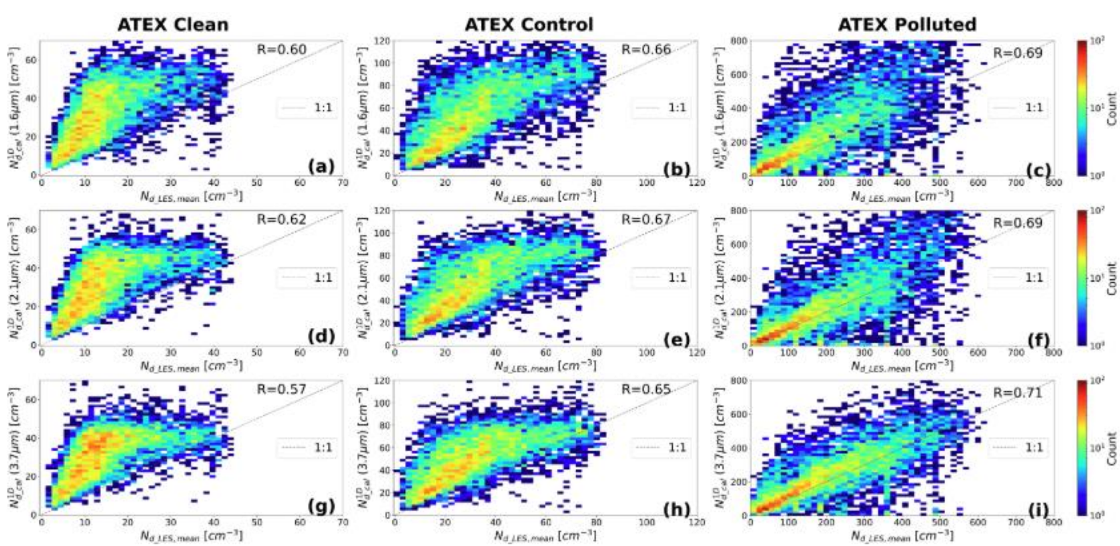

Figure showing comparison of N_d derived from bi-spectral retrievals based on 1D RT reflectance $(N^{1D}_{d\_cal}(\lambda_{abs}))$ versus N_d obtained from the LES $(N_{d_{LES}}, mean)$ for (a) ATEX clean, (b) ATEX control and (c) ATEX polluted for retrievals at absorbing channel, (λ_{abs}) of 1.6 μm, for (d) ATEX clean, (e) ATEX control and (f) ATEX polluted for retrievals at λ_{abs} of 2.1 μm and for (g) ATEX clean, (h) ATEX control and (i) ATEX polluted for retrievals at λ_{abs} of 3.7 μm at 4 h of simulation time. Note: 1D RT and bi−spectral retrievals utilized for this comparison were carried out at SZA 50^o. Dotted lines indicate 1:1 relationship between the derived and LES N_d.

Line 350-352: N_{d_LES} is not the reference point for adiabatic behaviour so disagreement between N_{1D} and N_{d_LES} are not necessarily evidence of deviations from adiabatic conditions. If the cloud is adiabatic then N_{1D} will agree with both the vertical average of the droplet concentration (no extinction weighting) and N_{d_LES}. However, agreement between N_{1D} and N_{d_LES} does not necessarily mean the cloud is adiabatic. This is just one of the complications of the chosen definition.

Reply: We thank the reviewer for this clarification. We agree that N_{d_LES}, particularly when defined using extinction weighting, should not be interpreted as a direct reference point for diagnosing adiabatic behavior. In the revised manuscript, we have removed the statement that disagreement between $N^{1D}_{d\_cal}(\lambda_{abs})$ and N_{d_LES} directly indicates deviations from adiabatic conditions. We now clarify that such differences reflect a combination of factors, including vertical microphysical inhomogeneity, the extinction-weighted nature of the LES reference, and the assumptions embedded in the adiabatic N_d retrieval equation. A sentence regarding this shown below has been added to the revised manuscript around L456 to 460

"The disparities observed between $N^{1D}_{d\_cal}(\lambda_{abs})$ and N_{d_LES} reflect the combined influence of vertical microphysical inhomogeneity, extinction weighting, and the assumptions embedded in the adiabatic N_d retrieval equation (Equation 6). Therefore, agreement with N_{d_LES} indicates consistency with the optically dominant portions of the cloud column sampled by the passive retrieval. "

Several figures (e.g., Fig. 4&5) show comparisons of retrievals but not the reference (regardless of its definition). Please include the reference/truth in the figures so we can understand which errors compensate.

Reply: Thank you for this suggestion. The reference (N_{d_LES}) has been added to the plots in the revised manuscript.

There are exact calculations for the partial derivatives (whatever the definition of N_d). Regression across columns and a cloud field does not provide an exact reference because it can include unwanted covariates, e.g., a correlation between mean droplet concentration and droplet profile shape, which will bias the power-law relationship. Please compare the estimates to the exact calculation for each column.

Reply: We agree that an exact expression for $\partial\alpha/\partial\tau$ can be obtained using linearized 3D RT (e.g., Doicu and Efremenko, 2019). Our SHDOM configuration is forward-only and we do not have access to a linearized/adjoint implementation; computing column-wise Jacobians would therefore require an impractically large number of additional 3D RT reruns. Instead, to provide a clear and physically consistent reference for the susceptibility diagnostic used in ACI studies, we introduce a "truth" benchmark computed directly from the 3D-RT TOD albedo field α(3D RT) (hemispheric flux-derived) and LES microphysical fields. Specifically, within each LES LWP bin we regress α(3D RT) against $\ln N_{d\_LES}$ to obtain the reference $S_{\alpha,rel}$, and we evaluate all retrieval-based susceptibility estimates relative to this explicit reference. This benchmark uses true albedo (not a reflectance proxy), accounts for anisotropy by construction, and avoids ambiguity associated with comparing multiple approximations without a reference.

Technical Comments.

Line 53: Perhaps remote sensing retrievals rather than instruments.

Reply: The sentence has been rewritten for clarity in the revised manuscript. It was modified from "Operational satellite remote sensing instruments such as the Moderate-resolution Imaging Spectroradiometer (MODIS) do not directly retrieve N_d" to

"Operational passive satellite instruments such as the Moderate-resolution Imaging Spectroradiometer (MODIS) do not directly retrieve N_d."

Line 59: Twomey & Seton 1980 is the earlier paper where bispectral sensitivity is introduced.

Reply: Thank you for this suggestion, "Twomey & Seton 1980" has been added as a reference for bi-spectral retrievals technique in the revised manuscript.

Line 70: Should "Most" operational retrievals be "All"?

Reply: "Most" have been changed to "All" in the sentence in revised manuscript.

Line 81: I suggest adding Zinner et al. 2010; Cornet & Davies

Reply: References Zinner et al. 2010 and Cornet & Davies 2008 have been added to as reference for the 3D effects in the revised manuscript.

Line 90: Várnai and Davies (1999) did not analyse retrieval accuracy.

Reply: Thank you for this point. Várnai and Davies (1999) have been removed from the reference in the sentence.

Line 86-93: At which wavelength are changes in cloud reflectance being used to define “brightening” and “darkening”. Changes in the relative amount of brightening and darkening across visible and absorbing wavelengths affects retrieval errors.

Reply: Brightening and darkening occur in both VNIR and SWIR bands, although more predominant in the VNIR band. When Reflectance under 3D RT > Reflectance under 1D RT is brightening and vice versal for darkening effect.

Line 93-99: I’m not sure what the message is. With the “Although XXX, previous studies have shown YYY” construction, is it saying that out of the several factors (variability in cloud top height, solar and view geometry, and photons leaking from optically thick to optically thin cloudy regions) that only solar-viewing geometry is significant? I suggest changing “photons leaking …” to “light scattering from optically thick to optically thin regions”.

Reply: We thank the reviewer for this comment. We agree that the original sentence was unclear and could be interpreted as suggesting that solar-viewing geometry is the only important factor. We have revised the text to clarify that cloud-top variability, cloud horizontal/vertical heterogeneity, and light scattering from optically thick to optically thin regions all contribute to 3D radiative effects, while solar and viewing geometry strongly modulate the sign and magnitude of their impact on bi-spectral retrievals. We have also replaced “photons leaking from optically thick to optically thin cloudy regions” with “light scattering from optically thick to optically thin regions,” as suggested.

“Several factors can contribute to 3D radiative effects in broken cloud fields, including variability in cloud-top height, cloud horizontal and vertical heterogeneity, solar and viewing geometry, and light scattering from optically thick to optically thin regions. Previous studies have shown that solar and observation geometry strongly modulate how these 3D effects appear in bi-spectral cloud-property retrievals, influencing both the sign and magnitude of retrieval errors in τ and r_e (Kato and Marshak, 2009; Marshak et al., 2006; Zhang et al., 2012), with these errors being generally more pronounced in broken cloud fields.”

Line 111: Gryspeerdt et al. 2019 and associated sentence does not seem to mix with the rest of the paragraph, which are about retrieval errors, not ACI in general.

Reply: The sentence about Gryspeerdt et al. 2019 has been removed from the paragraph in the updated manuscript.

Line 167: What is the range of the droplet size bins?

Reply: The range of the droplet size bins is from approximately **0.874** to **281.76**.

Line 196: Does this mean there is integration of multiple monochromatic reflectances over the band? Averaging of optical properties over the band? Or just a single monochromatic calculation at each nominal wavelength.

Reply: Optical properties over MODIS spectral response function in each MODIS band was averaged as a representative of the nominal/central wavelength.

Line 191-205: What is the domain top of the RT calculations? Is it the LES top of 3 km? or based on a particular atmospheric profile?

Reply: The domain top of the RT calculations were at the global cloud top for each cloud domain. The global domain cloud boundaries were diagnosed from the liquid water content (LWC) field using a threshold of 0.001 gm^{-3}. To obtain the domain cloud base and top, the maximum LWC at each height was first computed over all horizontal grid points, and the lowest and highest levels exceeding the threshold of 0.001 gm^{-3} were taken as the domain bottom and top, respectively.

Line 315-316: No need to cite Loveridge & Di Girolamo (2024) here as that study just employs simplified versions of microphysics based on other studies.

Reply: Thank you for this suggestion. The manuscript has been modified to suit this.

Line 531-532: Aren't PPHA biases caused by sub-pixel heterogeneity?

Reply: We thank the reviewer for pointing this out. We agree that PPHA biases are primarily caused by sub-pixel horizontal heterogeneity within the satellite footprint, rather than inter-pixel heterogeneity. We have revised the text to correct this wording and now distinguish more clearly between sub-pixel heterogeneity, which contributes to PPHA bias, and inter-pixel horizontal transport, which is more directly related to 3D radiative effects between neighboring columns. These changes (shown below) have been added to the revised manuscript around L729 to 736 .

"However, at the coarser spatial resolutions common in global satellite imager retrievals (e.g., MODIS Visible Infrared Imaging Radiometer Suite (VIIRS)), cloud-property retrievals are also affected by sub-pixel horizontal heterogeneity. This can introduce plane-parallel homogeneous approximation (PPHA) biases because the retrieval assumes each pixel is horizontally homogeneous, whereas the actual coarse pixel may contain unresolved variability in cloud optical thickness, cloud fraction, and cloud structure. Thus, at coarse resolution, the retrieved N_d can be influenced by both 3D radiative effects between neighboring cloudy regions and PPHA biases associated with unresolved sub-pixel heterogeneity."

Line 533-534: It shouldn't be regardless. There should be logic motivating the analysis that shouldn't be waved aside.

Reply: We agree with the reviewer. The word "Regardless" weakened the logic motivating the coarse-resolution analysis. We have revised the transition to clarify that the analysis is needed because operational passive imagers such as MODIS observe clouds at coarser spatial scales than the LES native resolution, and spatial aggregation can alter the balance between 3D radiative effects, plane-parallel bias, and subpixel heterogeneity. Therefore, extending the analysis to satellite-like resolution is necessary for assessing whether the native-resolution retrieval errors persist under more realistic satellite-observation conditions. The sentence below has been added to the revised manuscript to reflect this.

"Therefore, extending the analysis of derived N_d to coarser satellite-like resolutions is necessary to determine whether the 3D-RT-induced retrieval errors observed at the native LES resolution persist after spatial aggregation and under resolutions more comparable to operational passive imager observations."

Line 547: And plane-parallel optical depth bias contribution?

Reply: plane-parallel optical depth bias plays a role in the observed behavior, thus has been included in the revised manuscript.

Line 839-840: Sinclair et al. 2019 does not use active instruments.

Reply: Thank you for this comment. The reference "Sinclair et al. 2019" has been removed from the sentence in the revised manuscript.

Fig. 6&7: The dashed and solid lines for the PDF and the mean do not correspond to each other in the legend, which is confusing.

Reply: The legends of the plot has been amended in the revised manuscript.

Reply to Reviewer 2:

The manuscript adds some useful analysis to the already published state of knowledge of 3D radiative transfer and inhomogeneity effects influencing passive satellite retrievals. The authors do not simply repeat earlier studies of 3D effects on cloud property retrievals, but follow the impact on retrievals of droplet number concentration Nd from synthetic cloud model truth to far reaching conclusions based on such satellite retrievals on aerosol-cloud effects in other publications (Twomey effect and the connected cloud albedo susceptibility). Although the approach is not new (LES model as basis of synthetic measurements), demonstration of the effect leads to some clear conclusions, I have not seen summed up so nicely anywhere else before.

The paper is largely well written. At some places I see potential for some tightening of the text and removal of unnecessary detail.

I suggest publication after minor revision.

--- Whenever I ask for clarification or explanation in the following, I usually ask for it to be included in the manuscript, not in an answer to the reviewer. ---

General points:

1. Underestimation/Overestimation: You use this term to compare 3D RT based results vs 1D RT based results. That does not sound right. To my feeling these comparative terms are used to compare to the truth - which you almost nowhere do. In addition, if you use them for 3D vs 1D, 3D is definitely closer to being "truth". Example: "3D RT based underestimate 1D RT based" sounds awkward. It rather is "1D based overestimates 3D based" or "3D based is larger than 1D based". E.g. l.508/509, l.730, l.783, l.797, ...

Reply: Thank you for this suggestion. As suggested, the initial way of comparing the 1D versus 3D RT based results have been revised, this time by utilizing 1D based results to either overestimate or underestimate the 3D based results. These changes have been incorporated in the revised manuscript.

2. Connected to point 1. At several points I'm missing the comparison to the "truth". Do you not show it on purpose?

E.g. l.551+556, Fig.6+7 - Why not show N_{d_LES} PDF here?

Reply: We thank the reviewer for this comment. The N_{d_LES} have been included in the plots in the revised manuscript.

l.562-571 - Whether retrievals are better or worse cannot be checked as you do not show the true N_d.

Reply: The N_{d_LES} has been shown in the plots and discussions have been made on them in the revised manuscript.

Details:

l.31 - "plane-parallel effects": Only 3D experts know what you're talking about. Rather use "averaging effects" or "the use of the plane-parallel averaging assumption".

Reply: Thank you for this comment, the use of the plane-parallel averaging assumption has been added to the revised abstract.

l.33-35 - Please already mention here why this is true - "if albedo-Nd sensitivity is derived in a radiatively consistent way".

Reply: This sentence has been modified in the revised manuscript. The changes is shown below.

"Furthermore, for albedo susceptibility, 3D RT modifies the local α–N_d relationship at native LES resolution, especially under more oblique solar geometry, but the differences between 1D- and 3D-based susceptibility estimates are substantially reduced at coarser resolution. These results indicate that although 3D RT can strongly affect susceptibility at the LES resolution, its impact is reduced at satellite-like spatial resolution."

l.121 - "few have": Please add or repeat citations of papers who have.

Reply: The reference Arola et al. 2022 has been added as a reference to the reference to support this sentence.

l.210 - Acronym I3RC has not been introduced.

Reply: Since albedo is now computed from hemispherical integration of broadband radiance, the acronym I3RC has been removed from the revised manuscript.

l.212 - The "CERES algorithm" needs a citation or extra explanation please.

Reply: The sentence containing CERES algorithm has been removed from the revised manuscript since albedo is now computed from hemispherical integration of broadband radiance

l.237 - Please give reference to the original publications on such retrievals by Nakajima and King (1990) and Platnick et al (2003) showing the facts you refer to here.

Reply: Thank you for this suggestion, Nakajima and King (1990) has been added to the text as a reference for the bi-spectral retrieval technique.

l.300 - What is the fixed fraction f_{ad}? Is it one number? Can you give it here?

Reply: f_{ad} represents cloud subadiabaticity. While it is sometimes treated as one fixed scalar, we do not prescribe a single value here because $f_{ad}C_w$ is obtained directly from the LES and it varies for different cloud columns. This information was provided in the manuscript around L405 to 412.

"To avoid these unconstrained constants driving our understanding of retrieval behavior, throughout this work we have calculated both k, and the quasi-adiabatic lapse rate ($f_{ad}c_w$) in each cloudy column directly from the LES cloud field."

l.304 - Brenguier et al.(2000) or even earlier authors already set up the adiabatic cloud model. Grosvenor nicely summarized it. Please mention the more original publication too.

Reply: Thank you for this suggestion. Earlier studies such as Brenguier et al. (2000), Quaas et al. (2006), Boers et al. (2006) have been added to the paragraph to indicate that Grovesnor et al., (2018) builds on them.

l.339, section 3.1: This a 1D RT experiment here, excluding all 3D RT. Please, clarify early. I needed a while to get it.

Reply: In the revised manuscript, the section 3.1. title has been modified from "LES N_d and Derived N_d compariosn" to "Comparison of LES N_d and N_d Derived from Cloud Properties Retrieved from 1D RT Simulated Reflectance" additionally, 1D RT has been included in the first paragraph to make it easy for readers to get early enough that it is a 1D closure experiment.

"3.1. Comparison of LES N_d and N_d Derived from Cloud Properties Retrieved from 1D RT Simulated Reflectance

We first perform a RT closure comparison between the reference obtained directly from the LES ($\mathbf{N_d}$) and the derived from Eq. (6) using τ and retrieved from homogeneous plane parallel RT-based simulated reflectance (i.e., 1D RT) from the 0.86 channel paired with an absorbing wavelength channel , (λ_{abs}). "

l.344: Do you see a way to simplify your notation of e.g. "Nd_cal^1D(lambda_abs)". Isn't it enough to have the litte "1D" and "3D" in the superscript as soon as it is RT based (skip "cal")?

Reply: Thank you for this comment. The use of “cal” in the subscript is deliberate because Nd is not directly retrieved in operational passive satellites but **calculated** from cloud retrievals. This naming convention has not been changed in the revised manuscript.

l.355, Fig.2, "1D RT bi-spectral retrievals" - This is not the most precise wording, is it? Each retrieval in this manuscript is "1D" or "1D RT based", because it doesn't use an 3D RT inversion or correction. It should rather read "retrievals based on 1D RT tabulated reflectance". Please clarify wording throughout the manuscript for these cases. It can be quite confusing in the beginning otherwise. (similar e.g. in l.613 or l.634/635)

Reply: Thank you for this suggestion. The wordings have been adjusted to "bi-spectral retrievals based on 1D RT reflectance" and "bi-spectral retrievals based on 3D RT reflectance" where they occur. These changes have been made throughout the revised manuscript.

l.375 - Please give a summary sentence on the VW method by Miller et al.
Reply: A summary sentence has been provided in the revised manuscript around L469 to L473.

“First, we utilize the vertically weighted (VW) r_e (denoted as r_e(VW)), computed using the two-way-transmittance-weighted effective radius relationship derived from the optically weighted droplet size distribution defined by Miller et al. (2016) (check their Eq. (14)), and optical thickness (τ_{tot}) from the LES into Eq. (6) to derive N_d (denoted as $N_{d\_cal}^{r_e(VW),\tau_{tot}}$).”

l.380 - I don't understand the motivation behind the no drizzle test. It can be assumed that the drizzle drops > 30mu don't have a large impact on radiation (small numbers tail of the size distribution, large forward scatter) most of the time. However, if there would be many droplets>30mu they will have radiative impact and you would need to consider them, wouldn't you? Anyway, the figure shows that this test is not relevant too. I would suggest to only mention your result (and state "not shown"). I would also remove them from Fig. 3 where the difference is hardly visible.

Reply: We agree with the reviewer that the no-drizzle test is not central to the analysis and that the figure showed negligible differences. We have therefore removed the corresponding curves from Figure. 3 in the revised manuscript and now mention the test only briefly in the text, stating that excluding droplets with r > 30 µm had negligible impact on the Nd PDFs and did not affect the conclusions.

“A sensitivity test was also performed where Nd is calculated using r_e^* and cloud optical depth computed from the LES that excludes cloud optical depth contributions from drizzle drop sizes (r) by

isolating r > 30 μm from the droplet size distributions. The resulting PDFs were mostly unchanged from the PDFs of $N_{d\_cal}^{r_e^*,\tau_{tot}}$ *(not shown). This was expected because drizzle-size drops typically occur in small numbers and contribute weakly to , although their radiative impact could become important in cases with substantial drizzle populations."*

l.386 - I do see three Nd PDF derived from LES "cal" and three retrieved at different channels, but no further "derived from retrievals under the different levels of constraints". Please clarify.

Reply: The three different levels of constraints were mentioned earlier in the manuscript (L372 to L384 in the initial manuscript). They are (1) we utilize the vertically weighted (VW), and optical thickness from the LES into the Grosvernor et al. (2018) pseudo-adiabatic equation to derive N_d. (2) we use droplet effective radius of the layer where the LWC is maximum as a representative of the adiabatic cloud top effective radius and optical thickness from the LES as inputs to the pseudo-adiabatic equation to derive N_d. (3) we derive N_d, by using r_e^* and cloud optical depth computed from the LES that excludes cloud optical depth contributions from drizzle drops sizes and cloud droplet effective radius into the pseudo-adiabatic equation.

How ever, since the plot of the drizzle test has been removed from Figure 3 in the revised manuscript, we only show two levels of constraints. We clarify this by explicitly stating that we refer to when the optical thickness and droplet effective radius are constrained as earlier discussed in the previous sentence in the manuscript. These changes are around L475 to 500 in the revised manuscript.

"The idea of testing with r_e^* *is because it is representative of the droplet radius profile that follows the adiabatic assumption of the retrieval — absent significant entrainment modification at cloud top or model resolution artifacts. PDFs of* N_d *obtained from the LES, those inferred from the retrievals obtained from 1D RT simulated reflectance (at SZA* 50°*) when* λ_{abs} *is* $1.6,\ 2.1$ *and* $3.7\ \mu m$*, and those derived from retrievals when the optical thickness and droplet effective radius are constrained as earlier discussed, for the ATEX clean, ATEX control, and ATEX polluted cases are presented in Figure 3."*

l.395, Fig.3 - Here the sub and superscripts get too small and unreadable. Please improve.
Reply: The size of the subscripts and superscripts labels have been increased in the Figure in the revised manuscript.

l.454 - You cannot expect that $r_e(VW)$ as an average provides a better agreement than the cloud top leaning r_e^*, if your N_d calculation assumes a cloud top value, can you? Please clarify.

Reply: The purpose of the r_e(VW) test was not to imply that a vertically weighted average effective radius used in calculating N_{d_cal} should provide better agreement with N_{d_LES} than when r_e at maximum LWC (r_e^*) is utilized. Rather, it was included as a diagnostic to examine how sensitive the derived N_d is to the vertical weighting of r_e (which is commonly used as a reference for r_e in satellite remote sensing studies), given that passive retrievals are radiatively weighted and do not sample the cloud column uniformly. We have revised the text in the manuscript to clarify that r_e, taken at maximum LWC, is more consistent with the cloud-top-leaning/adibatic assumption embedded in the N_d equation, whereas r_e(VW) is used only to test the impact of using a vertically weighted effective radius. We have also modified the interpretation accordingly and no longer suggest that r_e(VW) is expected to improve agreement relative to r_e^*. This changes are made around L563 to 572 in the revised manuscript as shown below.

"When we consider the PDF of calculated N_d when τ and r_e are constrained to various degrees using values derived directly from the LES cloud field, the $N_{d\_cal}^{r_e(VW),\tau\ tot}$ PDF shows substantial deviations from the PDF of N_{d_LES}. A reason for this is that $r_e(VW)$ represents a weighted vertical average which is highly sensitive to the cloud microphysics and vertical structure of r_e, especially if the optical extinction in the cloud entrainment region is large enough to impact the $r_e(VW)$. The agreement between the PDF's of N_d derived under the τ and r_e constraints (i.e., $N_{d\_cal}^{r_e^*,\tau\ tot}$ and $N_{d\_cal}^{r_e(VW),\tau\ tot}$) and N_{d_LES} reduces as the adiabatic realism of the input properties decreases; $N_{d\_cal}^{r_e^*,\tau\ tot}$ showing better agreement with N_{d_LES} than $N_{d\_cal}^{r_e(VW),\tau\ tot}$ and N_{d_LES} comparison because $N_{d\_cal}^{r_e^*,\tau\ tot}$ utilizes r_e^* which is most representative of an adiabatic cloud top r_e value which is a key requirement for formulation of Equation 6"

l.467 - What is "almost similar"? Sounds like "very different".

Reply: The sentence has been modified from "almost similar" to "similar" in the revised manuscript. This is to indicate that both PDF's are similar.

l.470 - Are peaks smaller in frequency? Or in Nd?

Reply: The $N_{d\_cal}^{3D}(3.7\ \mu m)$ peak around slightly smaller N_d values than $N_{d\_cal}^{1D}(3.7\ \mu m)$.

l.467-477 - Please consider shortening here. Wording is rather un-precise. And I'm missing the rough reasons (3D reflectance smaller due to "channeling" or "diffusion" from large to small optical thickness, leading to smaller tau and larger r_e, leading to smaller N_d).

Reply: Thank you for this comment. The main point here is that the generally similar PDFs of $N_{d\_cal}^{1D}$

to and $N_{d\_cal}^{3D}$ comes from similar reflectance under 1D and 3D RT which yield similar τ and r_e bi-spectral retrievals if other factors are largely unchanged. I.e., for the $N_{d\_cal}^{1D}$ to be similar to $N_{d\_cal}^{3D}$, then the input τ and r_e used in the calculation are similar. Not necessarily 3D reflectance smaller than the 1D reflectance leading to leading to smaller τ and larger r_e, leading to smaller N_d. We have shortened this paragraph and revised the interpretation to be clearer. The changes made to the revised manuscript is around L613 to 618 and shown below.

"When the sun is high, the PDFs of $N_{d\_cal}^{3D}(\lambda_{abs})$ are generally similar to their $N_{d\_cal}^{1D}(\lambda_{abs})$ counterparts (where $\lambda_{abs} = 1.6$, 2.1, and $3.7\ \mu m$), especially for the $\lambda_{abs} = 2.1\ \mu m$ retrieval (Figures 4a, d, and g). A reason for this similarity is the minimal 3D effects (lack of extreme darkening and brightening in the simulated 3D RT reflectance) when the sun is high, leading to the reflectance utilized for the cloud property retrieval under 3D RT at high sun comparable to its 1D RT counterpart. This in turn yields comparable derived N_d. "

l.475 - I would not call these "narrow peaks". To me this is a misconception of the lower boundary value "0". You should also cut the x axis at N_d=0 $cm^{-3}$and not at N_d=-5 cm^{-3} in Figs. 4, 6, 7.

Reply: Thank you for this comment. We have revised the wording to avoid imprecise descriptions of PDF peak behavior. Additionally, the x axis has also been set to N_d=0 in Figures 4, 5, and 6 in the revised manuscript as suggested. Changes made to the text of the revised manuscript is around L618 TO 623 as shown below.

"Results for SZA 30^o, show that the PDF of $N_{d\_cal}^{3D}(\lambda_{abs})$ broadens slightly at both ends (Figure 4b, e and h) compared to the high sun case (Figures 4a, d, and g). This slight broadening of the $N_{d\_cal}^{3D}(\lambda_{abs})$ PDFs occurs because 3D-induced darkening and brightening effects occurring

simultaneously. The darkening effects makes typically lead to retrieval of smaller τ and larger r_e*,* *which together produce smaller* $N_{d\_cal}^{3D}(\lambda_{abs})$ *values compared to* $N_{d\_cal}^{1D}(\lambda_{abs})$*."*

l.512-517 - Please check. It's either my confusion or you're mixing up things here. Differences between spectral bands are driven by absorption differences, yes. But 3D effects are not. The effect that leads to darkening for high sun is enhanced (macrophysical) forward scatter from exposed tops and regions of large optical thickness due to cloud geometry (downward "channeling"). It is not enhanced absorption. The same enhanced macrophysical forward scattering leads to average brightening for low sun.

Reply: We thank the reviewer for this comment. The physical impact of 3D radiative effects changes with solar geometry, while the way these effects propagate into the retrieval also depends on spectral band. The 3D effects themselves are primarily controlled by cloud geometry, horizontal photon transport, cloud-top structure, forward scattering, and side illumination/channeling, rather than by absorption alone. Under high-sun conditions, enhanced macrophysical forward scattering and downward channeling through exposed cloud tops and optically thicker regions can reduce the reflected signal locally, producing darkening. At more oblique SZAs, side illumination and horizontal transport across cloud boundaries can instead enhance the reflected signal, producing average brightening. However, the retrieval response to these 3D-affected reflectances differs among the absorbing channels because the 1.6, 2.1, and 3.7 μm bands have different liquid-water absorption strengths and vertical penetration depths. Thus, when we state that darkening or brightening effects dominate for a given channel, we refer to their statistical impact on the retrieved τ, r_e, and $N_{d\_cal}^{3D}(\lambda_{abs})$, not to absorption being the fundamental cause of the 3D effect.

l.531. Acronym PPHA never defined.

Reply: The acronym PPHA has previously been defined as plane parallel homogeneous approximation (PPHA) in Section 2.3 L246 before being used in L531.

l.573, Fig.8 and l.581-588 - This might be information you could summarize and say "(not shown)".

Reply: Figure 8 has been taken out of the revised manuscript and the information summarized as suggested. These changes are around L 744 to 752 in the revised manuscript and shown below.

"We also examined the sensitivity of $N^{3D}_{d\_cal}(\lambda_{abs})$ to spatial aggregation by comparing native LES-resolution (100 m) results with coarse satellite-like resolution (800 m) results for $\lambda_{abs} = 1.6\ and\ 3.7\ \mu m$ (not shown). In general, the coarse-resolution results based on overcast 800 m pixels are more consistent with the native-resolution mean $N^{3D}_{d\_cal}$ than results based on all cloudy 800 m pixels. This behavior is found for both absorbing channels, although for λ_{abs}=3.7 μm at SZA =10°, the all-cloudy and overcast-only estimates are both comparable to the native-resolution result. These tests indicate that the treatment of partially cloudy coarse pixels can influence aggregated N_d, but the main resolution-dependent behavior remains robust across the two absorbing channels considered."

l.590, section 3.4 - Why don't you show the real Twomey effect in this section - the truth? LWP from LES, N_{d_LES} and R_{BB} from 3D RT?

Reply: Thank you for this suggestion, an empirical truth comprising of LWP from LES, N_{d_LES} and albedo (α) from 3D RT has been introduced and discussed in section 3.4 of the revised manuscript as suggested.

l.615, Tab.1 - Please comment on the awkward threshold values. How did you find them?

We thank the reviewer for pointing this out. The LWP bin limits were derived from percentiles of the valid retrieved LWP distribution at 10% intervals. The first two percentile ranges consisting of small LWP ranges were excluded, leaving eight LWP bins numbered (B1–B8) for the susceptibility related calculation. This percentile-based procedure explains why the values in Table 1 are non-rounded, and we have clarified this in the revised manuscript and table caption.

"The bin edges were derived from percentile boundaries of valid retrieved LWP distribution computed at 10% intervals; however, the first two percentile ranges were excluded, leaving the eight retained bins. "

l.664 - Please check equation. It does not seem correct unit-wise.

Reply: The relative susceptibility is unitless, so LHS of the equation is unitless. For the RHS of the equation, the units of N_{d_cal} (which is cm^{-3}) cancel out leaving the unitless albedo as the units for the RHS making it equal to LHS.

l.666 - An "alpha" instead of an "alb".

Reply: Thank you for this suggestion, the revised manuscript has been modified to utilize alpha (α) for the albedo and in the subscript of the susceptibility.

l.678 - Please remove the binning from the Fig. caption.

Reply: The binning has been removed from Figure 10 in the revised manuscript.

l.691 - Fig. 10+11, It would be nice to relate Fig 10 and Fig 11 by showing S_alb,rel values (and according line fits through R_BB and N ?) in Fig 10. Then the reader could compare the SZA=50 result to results over all SZA in Fig 11.

Second point - Why not use colored lines here on top of using symbols? You used a lot of color in all other plots.

Reply: We thank the reviewer for this comment, plotting line fits through α vs. N_d plots will be clumsy due to the closeness of the LWP bins. However, using this idea, two diagnostics of SZA 20 and 50 were introduced and discussed in the revised manuscript, including a separate plot of their relative susceptibility. Additionally, colored lines have been utilized in the updated plot in the revised manuscript.

l.725 - I don't see a "more pronounced decrease at larger LWP" above 100.
Reply: This has been reexplained in a more consistent way in the revised manuscript.

l737 - "concerns about 3D effects" - Isn't it the 3D effects which (do not) impact and not the "concerns about"?

Reply: This sentence has been removed in the revised manuscript l745 - You put the importance of Nd accuracy first here, but I'm missing the clear conclusions on a "best-practice" here.

Reply: We thank the reviewer for this comment. We agree that the original conclusion emphasized the importance of N_d accuracy but did not clearly state the practical implications of our results. We have revised the conclusion to make the "best-practice" message more explicit. Specifically, we now state that retrieval-derived N_d should be interpreted with caution at native high resolution and large SZA, where 3D RT effects strongly perturb the retrieved τ and r_e. For albedo-susceptibility applications, we recommend evaluating Sα,rel using flux-derived albedo within fixed LWP regimes and comparing retrieval-based estimates at the same spatial resolution. We also clarify that although the 3.7µm retrieval generally agrees better with the LES N_d reference in our simulations, its stronger cloud-top weighting must be considered when interpreting it as a cloud-layer N_d.

l.767, "reference LES Nd" - Why not stick to "N_{d_LES}"?

Reply: The N_{d_LES} has been used to describe our reference N_d in most parts of the revised manuscript.

l.789-799 - I think a more compact summary and conclusions would be nice after a complex manuscript. Here, e.g., I would avoid to start talking about PDF distributions again.

Reply: The We agree with the reviewer that this section should not be described in terms PDF descriptions in the conclusion. We have shortened this paragraph and revised it to summarize the physical interpretation instead: at larger SZA, simultaneous 3D brightening and darkening effects introduce broader variability in $N_{d\_cal}^{3D}(\lambda_{abs})$, while the domain-averaged response depends on which effect dominates for a given SZA, absorbing channel, and cloud scene. The paragraph has therefore been modified in the revised manuscript around L1195 to 1200 as shown below.

"When the sun is low at SZA 50°, 3D radiative effects become stronger and produce larger variability in the retrieval-derived N_d. This occurs because both brightening and darkening effects occur in significant amount at the low sun angle. Brightening effects tend to produce larger τ and smaller r_e, leading to larger $N_{d\_cal}^{3D}(\lambda_{abs})$ compared to corresponding $N_{d\_cal}^{1D}(\lambda_{abs})$ values, while darkening effects tend to produce smaller τ and larger r_e, leading to smaller $N_{d\_cal}^{3D}(\lambda_{abs})$

compared to $N_{d\_cal}^{1D}(\lambda_{abs})$ values. At the domain scale, the relative magnitude of these effects changes with SZA and absorbing channel used for retrievals. For larger SZAs, the brightening effect becomes more dominant in several cases, causing the domain-averaged $N_{d\_cal}^{1D}(\lambda_{abs})$ to exceed its 1D RT counterpart, although the SZA at which this transition occurs depends on the absorbing channel and cloud scene.''

l.800-810 - Why not generalize the conclusion for all absorbing channels? Above you stated that there is no difference.

Reply: We agree with the reviewer. We have revised this paragraph to generalize the conclusion across all three absorbing-channel retrievals rather than emphasizing only the 2.1 µm case. The revised text now states that the same behavior is observed for the 1.6, 2.1, and 3.7 µm retrievals, supporting the conclusion that the resolution-dependent behavior is not specific to a single absorbing channel. The text below has been added to the revised manuscript around L1227 to 1235

"When all valid cloudy pixels are included, including partially cloudy pixels, the N_{d_cal} distributions tend to shift toward smaller values, indicating that partially cloudy coarse pixels can influence the retrieved N_d. However, when only confident overcast cloudy pixels are retained, the mean N_{d_cal} values at the native LES and coarse satellite-like resolutions become more comparable. This behavior is observed consistently for the 1.6, 2.1, and 3.7 µm absorbing-channel retrievals, which suggests that it is not just restricted to a specific band. Therefore, satellite-like coarse-resolution retrievals can provide comparable mean N_d estimates to the native-resolution results when partly cloudy pixels are excluded, and only confidently cloudy pixels are utilized."

l.842 - I'm not sure that you need the appendix. I would prefer a simple statement in the main text with something like: "Column averaging is not useful. Optically irrelevant, i.e., optically thin parts at cloud edges would dominate averages or make arbitrary cut-offs of small LWC necessary".

Reply: The appendix has been removed in the revised manuscript

References (New added references are highlighted)

Brenguier, J.-L., Pawlowska, H., Schüller, L., Preusker, R., Fischer, J., & Fouquart, Y. Radiative properties of boundary layer clouds: Droplet effective radius versus number concentration. *Journal of the Atmospheric Sciences*, **57**, 803–821. **https://doi.org/10.1175/1520-0469(2000)057<0803:RPOBLC>2.0.CO;2**, 2000.

Doicu, A. and Efremenko, D. S.: Linearizations of the Spherical Harmonic Discrete Ordinate Method (SHDOM), Atmosphere, 10, 292, https://doi.org/10.3390/atmos10060292, 2019.

Feingold, G., Goren, T., and Yamaguchi, T.: Quantifying albedo susceptibility biases in shallow clouds, Atmospheric Chemistry and Physics, 22, 3303–3319, https://doi.org/10.5194/acp-22-3303-2022, 2022.

Gonzalez, J., Dipu, S., Jimenez, G., Camps-Valls, G., and Quaas, J.: Machine Learning-Based Retrieval of Cloud Droplet Number Concentration and Liquid Water Path From Satellite Spectral Data, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 18, 21910–21922, https://doi.org/10.1109/JSTARS.2025.3601981, 2025.

Gryspeerdt, E., McCoy, D. T., Crosbie, E., Moore, R. H., Nott, G. J., Painemal, D., Small-Griswold, J., Sorooshian, A., and Ziemba, L.: The impact of sampling strategy on the cloud droplet number concentration estimated from satellite data, Atmospheric Measurement Techniques, 15, 3875–3892, https://doi.org/10.5194/amt-15-3875-2022, 2022.

Painemal, D. and Zuidema, P.: Assessment of MODIS cloud effective radius and optical thickness retrievals over the Southeast Pacific with VOCALS-REx in situ measurements, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2011JD016155, 2011.

Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., Alexandrov, M. D., Bennartz, R., Boers, R., Cairns, B., Chiu, J. C., Christensen, M., Deneke, H., Diamond, M., Feingold, G., Fridlind, A., Hünerbein, A., Knist, C., Kollias, P., Marshak, A., McCoy, D. T., Merk, D., Painemal, D., Rausch, J., Rosenfeld, D., Russchenberg, H., Seifert, P., Sinclair, K., Stier, P., van Diedenhoven, B., Wendisch, M., Werner, F., Wood, R., Zhang, Z., and Quaas, J.: Remote Sensing of Droplet Number Concentration in Warm Clouds: A Review of the Current State of Knowledge and Perspectives, Reviews of Geophysics, 56, 409–453, https://doi.org/10.1029/2017RG000593, 2018.

Kokkola, H., Tonttila, J., Calderón, S. M., Romakkaniemi, S., Lipponen, A., Peräkorpi, A., Mielonen, T., Gryspeerdt, E., Virtanen, T. H., Kolmonen, P., and Arola, A.: Model analysis of biases in the satellite-diagnosed aerosol effect on the cloud liquid water path, Atmospheric Chemistry and Physics, 25, 1533–1543, https://doi.org/10.5194/acp-25-1533-2025, 2025.

Zuidema, P., Xue, H., and Feingold, G.: Shortwave Radiative Impacts from Aerosol Effects on Marine Shallow Cumuli, Journal of the Atmospheric Sciences, 65, 1979–1990, https://doi.org/10.1175/2007JAS2447.1, 2008.

Cornet, C., & Davies, R. Use of MISR measurements to study the radiative transfer of an isolated convective cloud: Implications for cloud optical thickness retrieval. *Journal of Geophysical Research*, 113(D4). https://doi.org/10.1029/2007JD008921, 2008.

Zinner, T., Wind, G., Platnick, S., and Ackerman, A. S.: Testing remote sensing on artificial observations: impact of drizzle and 3-D cloud structure on effective radius retrievals, Atmos. Chem. Phys., 10, 9535–9549, https://doi.org/10.5194/acp-10-9535-2010, 2010.