Crumeyrolle et al. are interested in quantifying the role of aerosol perturbations for cloud microphysics. They aim to do so by a statistical analysis of observational data, combining ground-based observations with passive microwave ground-based remote sensing and cloud retrievals from a geostationary satellite. The method is meant as a proof-of-concept study.

We would like to thank reviewer 1 for his/her comments and suggestions. We fully acknowledge all his/her comments and suggestions and hope to have improved the (revised) manuscript accordingly.

With this aim, the study may be ok for Atmos Chem Phys. The study itself, however, does not provide a substantial new scientific insight. There is a lot of data pre-processing, where some of the applied filters are poorly motivated and seem partly contradictory. Ultimately, a mere 188 data points remain for analysis. This severe lack of data precludes meaningful statistical analysis. The uncertainty quantification consequently shows ranges of almost ±100%.

This manuscript is currently submitted to *Atmospheric Measurement Techniques* (AMT), not to *Atmospheric Chemistry and Physics* (ACP). It was redirected to AMT following the recommendation of the ACP editor, who rightly considered that the main contribution of the work lies in the development and demonstration of a new methodology rather than in providing new atmospheric process understanding.

The primary objective of this study is to introduce and evaluate a novel methodology to investigate aerosol-cloud interactions using the combined analysis of ground-based and satellite observations over one specific site. At this stage, the manuscript should be regarded as a proof of concept, demonstrating the feasibility of the approach and highlighting its potential for future applications rather than providing definitive climatological conclusions. Beyond this initial demonstration, the ultimate goal is to apply this methodology to larger datasets collected across the ACTRIS network and to exploit the unique aerosol chemical composition measurements available at these stations. This will make it possible to investigate aerosol-cloud interactions as a function of aerosol type and composition, providing insights into the role of different aerosol species in cloud microphysical processes that are not accessible from satellite observations alone.

The limited size of the final dataset is a deliberate consequence of this proof-of-concept strategy. We intentionally restricted the analysis to a single ACTRIS site in order to simplify the initial implementation of the methodology and avoid introducing additional sources of variability related to differences in site characteristics, aerosol environments, and meteorological conditions. Once the methodology has been validated, it can be extended to multiple sites and significantly larger datasets. Note that following minor refinements to the filtering pipeline since the previous version, the final dataset now comprises 201 cases (previously 188), see revised Table 2.

The extensive data filtering is also intentional and scientifically motivated. Our objective is to isolate aerosol-cloud interactions while minimizing the influence of dynamical processes that could mask or bias the relationships under investigation. Because the methodology combines ground-based observations with satellite retrievals, several strict quality criteria are required. These include, for example, excluding cases affected by multilayer cloud systems or clouds not connected to the boundary layer where aerosol measurements are performed. Although these filtering steps substantially reduce the number of available cases, they are necessary to ensure that the remaining observations satisfy the assumptions underlying the proposed methodology.

We fully acknowledge in the manuscript that the resulting dataset is small and that this limits the statistical robustness of the derived aerosol-cloud susceptibilities. We explicitly state that future studies will require substantially larger datasets, including observations from additional

ACTRIS stations and longer observation periods, to obtain statistically robust estimates and better quantify the associated uncertainties.

Finally, we would like to emphasize that the large uncertainties reported in our study should also be viewed in the context of the existing literature. During the preparation of the manuscript and in response to the reviewer's comments, we compiled a broad range of aerosol-cloud susceptibility estimates obtained using different methodologies and for different liquid water path (LWP) regimes (Figures R1 and R2). This comparison shows that many published susceptibility estimates are not associated with a specific LWP range (white area) and often do not include uncertainty estimates (stars correspond to the study including uncertainties calculated directly in their paper). In several cases, the uncertainties shown in our comparison had to be reconstructed from the information provided in the original publications, mostly coming from observations performed over different region or different season. This highlights that uncertainty quantification remains a general challenge in aerosol-cloud interaction studies rather than being specific to our work. Figures R3 and R4 will be added to the manuscript to provide a better comparison to the literature results (see comment below).

If the editor decides the topic still fits Atmos Chem Phys, I would have a number of concerns, partly major, partly minor, to be addressed before further consideration.

L14 What exactly is meant by “spatially imprecise”?
Satellite observations provide spatially extensive measurements but at a coarse spatial resolution, representing averages over large footprints. In contrast, ground-based instruments sample local conditions at a specific location and can therefore resolve fine-scale spatial variability that is often smoothed out in satellite products. Additionally, satellite retrievals are affected by vertical averaging, retrieval assumptions, and temporal sampling constraints, which further complicate direct comparisons with point measurements.
Revised text: “*The effect of aerosols on cloud properties remains poorly constrained, in part because ground-based measurements are highly local while space-based retrievals are spatially imprecise, representing averages over large footprints.”*

L92 What are “total..coefficients”?
Nephelometers provide measurement of the total scattering coefficient and back scattering coefficient. The total scattering coefficient is then a measurement of the light scattered over the full range of angles available in the instrument (10-170°).

L129 It would be useful to briefly report the method to retrieve LWC here, and discuss its uncertainties.
We thank the reviewer for this suggestion. We have added a description of the LWC retrieval method and a discussion of its associated uncertainties. The microwave profiler (MWRP) does not directly measure the vertical distribution of liquid water content (LWC). Unlike temperature and humidity profiles, which are directly retrieved from microwave observations, LWC is derived assuming an adiabatic cloud profile following the approach of Karstens et al. (1994). In this method, the adiabatic LWC profile is empirically corrected to account for departures from adiabatic conditions, such as entrainment and mixing, and is constrained to be consistent with the retrieved liquid water path (LWP) through vertical integration.
The uncertainty of LWC profiles retrieved from microwave radiometry is difficult to quantify precisely because it depends not only on the uncertainty of the integrated LWP but also on the assumptions used to reconstruct the vertical profile. The uncertainty of MWR-derived LWP is typically estimated to be on the order of 20–30 g m^{-2} for liquid clouds (Marchand et al., 2003; Turner et al., 2007), with larger relative uncertainties for low-LWP conditions. However,

additional uncertainties affect the retrieved LWC profile, including cloud boundary determination, the assumption of adiabaticity, and the empirical correction applied for sub-adiabatic conditions. Therefore, the retrieved LWC profiles should be considered as approximate estimates of the liquid water distribution rather than direct measurements, and caution is required when using them for quantitative analyses.

Revised text: *"The microwave profiler RPG-HATPRO G5 (MWRP G5) measures the radiation in the 20-29 GHz and 50-59 GHz ranges emitted by the troposphere. Using a neural network (Rose et al., 2005; Löhnert and Crewell, 2003; Crewell and Löhnert, 2007) adapted to the radiometer geographic situation and trained by around 25000 profiles, the MWRP G5 provides retrievals of tropospheric vertical profiles (0 − 10 km) of absolute humidity and temperature. From these data, relative humidity profiles can be calculated, as well as other thermodynamic quantities (e.g. potential temperature, liquid water content, etc). Unlike temperature and humidity, LWC is not directly retrieved from the microwave observations but is derived assuming an adiabatic cloud and following the approach of Karstens et al. (1994). In this method, the adiabatic LWC profile is empirically corrected to account for departures from adiabatic conditions, such as entrainment and mixing. The resulting profile is then constrained to be consistent with the retrieved liquid water path (LWP) through vertical integration. The uncertainty of LWP retrievals from microwave radiometry is typically estimated to be on the order of 20–30 g m^{-2} for liquid clouds (Marchand et al., 2003; Turner et al., 2007), although larger relative uncertainties are expected for low-LWP conditions. The uncertainty of the derived LWC profiles is more difficult to quantify, as it additionally depends on the assumptions used in the retrieval, including cloud boundary determination, adiabaticity, and the empirical correction applied to represent sub-adiabatic conditions."*

L140 This line reads like VIS channels would have 3 km resolution and NIR, 1 km, but that is not true. Only the HRV channel has 1km.

That is correct, only one channel has a resolution of 1 km (at the nadir), and the other 11 channels have a resolution of 3 km.

Revised text: "*The Spinning Enhanced Visible and Infrared Imager (SEVIRI), on board the geostationary METEOSAT Second Generation satellites, has a temporal resolution of 15 minutes and measures in four visible and near-infrared channels and eight infrared channels, with a spatial resolution of 1 km for the High-Resolution Visible (HRV) channel (broadband, 0.4–1 µm) and 3 km for the others (Schmetz et al., 2002).*"

L155 say the "idealised" involves the assumption of adiabaticity

In line with the reviewer's comment, we propose to change the text.

Revised text: "*...we are considering an idealized stratiform boundary-layer cloud following adiabaticity assumptions, with a fixed dispersion of the assumed cloud droplet size distribution (Bennartz & Rausch, 2003).*"

L159 Sentence unclear. The comparison with CALIPSO reveals biases of 30%. But isn't that a function of the instrument sensitivity? See the GEWEX assessment, Stubenrauch et al. Surveys in Geophysics 2024. The sentence continues with "is less than 10% above the ATOLL station". What is meant? The bias vs. CALIPSO? Or the liquid fraction itself?

That is correct, it depends on the instrument sensitivity. The difference of 30% is from the article Benas et al. (2023) and represents a maximum bias, based on the year 2013, averaged on a regular 1.5˚×1.5˚ grid. For the value above the ATOLL station, it refers to the liquid fraction itself — not the bias versus CALIPSO — extracted from Fig. 6c of Benas et al. (2023). The comparison is indeed only valid according to the sensitivity of the instruments, and biases might vary depending on the cloud type (Stubenrauch et al., 2024). Nevertheless, in Benas et al. (2023), CALIPSO is used as the reference for cloud phase, but this should be considered with caution.

Revised text : "*A comparison with CALIPSO on a 1.5°x1.5° grid for 2013 reveals differences in liquid cloud fraction of up to 30%, but lower than 10% over the ATOLL station. The difference is even smaller if upper clouds with* τ *less than 0.1 are excluded from CALIOP measurements. Note that CALIPSO serves as the reference in Benas et al. (2023), and that the reported bias encompasses all cloud types – not exclusively low-level clouds – so type-specific biases may vary due to the sensitivity of the instruments (Stubenrauch et al., 2024).*"

L162 - AMSR2 and CLAAS LWP are not directly comparable. Microwave LWP is more like an all-sky retrieval, NIR one, in-cloud. See Seethala and Horváth JGR 2010
The comparison was performed by Benas et al. (2023), and it is not the aim of the present study to re-evaluate the dataset. We thank the reviewer for pointing out this publication, and we therefore suggest a **revised version**: "*Cloud optical depth, re, and Nd were compared with AMSR2, MODIS, and in situ measurements. A correlation coefficient of 0.79 was retrieved between the liquid water path from CLAAS-3 and from AMSR-2 measurements.*"

L163 Why jump between r and r^2?
As before, all the comparisons are from Benas et al. (2023), and we simply reported the conclusions of that study in the paragraph. Unfortunately, we cannot re-analyse the data to provide all the coefficients of determination.
Revised text: "*A comprehensive evaluation of CLAAS-3 is given in Benas et al. (2023), and we report here a brief summary of it.*"

L161-164 Please report confidence intervals for the CLAAS retrievals, from the comparison to the reference data, rather than correlation coefficients.

Unfortunately, we cannot retrieve the confidence intervals for the entire CLAAS retrievals, since we do not have access to them. Nevertheless, we can provide the confidence intervals for our own dataset. Over our entire dataset, we retrieve an uncertainty ranging from 0.43 to 31.94 μm for r_e and from 0.3 to 3276.6 cm^{-3} for n_d, with 90% of the dataset showing an uncertainty lower than 3.54 μm and 221.0 cm^{-3} for r_e and n_d, respectively.
We propose to revise it into: "*From our filtered dataset, over the ATOLL station, the uncertainty ranges from 0.56 to 17.67 μm for* r_e *and from 3.6 to 3276.6* cm^{-3} *for* n_d*, with 90% of the dataset showing an uncertainty lower than 3.54 μm and 221.0* cm^{-3} *for* r_e *and* n_d*, respectively.*"

L165 Typically, the statistical error is considered as uncertainty, while if a bias is known, it should be corrected for.
We thank the reviewer for this pertinent point. In the case of statistical error, we cannot correct since it depends on the cases. We propose to **change the text and to remove the notion of bias that can be misleading**: "*Lastly, it is important to acknowledge that several hypotheses and a priori assumptions in the retrieval methods could introduce uncertainties.*"

L173 (Table 1) Why not LWP from CLAAS?
We chose to use the LWP retrieved from the RPG microwave radiometer because it provides a continuous, ground-based measurement of the cloud liquid water path directly above the station, making it particularly well suited for classifying the cloud cases according to their local LWP. In contrast, the CLAAS-3 LWP represents the average value over the satellite pixel, whereas

the aerosol observations are representative of the air mass sampled at the station. Using the RPG LWP therefore provides a classification that is more consistent with the local observational framework adopted in this proof-of-concept study.

Nevertheless, to assess the sensitivity of our results to this choice, we repeated the susceptibility calculations using the CLAAS-3 LWP. As shown in Table R1, the resulting values of both S_n and S_{re} are very similar and agree within their uncertainties. This demonstrates that the main conclusions of the study are not sensitive to the choice of LWP dataset, although we retained the RPG-derived LWP because it is more consistent with the local, ground-based observational framework of the proposed methodology.

Table R1 : S_n and S_{re} retrieved using LWP range calculated by CLAAS-3 and MWP.

	20 - 100 $g.m^{-2}$	100-200 $g.m^{-2}$
S_n - LWP CLAAS3	0.33 +/- 0.39	0.38 +/- 0.25
S_n - LWP MWP	0.40 +/- 0.30	0.33 +/- 0.27
S_{re} - LWP CLAAS3	0.09 +/- 0.12	0.11 +/- 0.08
S_{re} - LWP MWP	0.12 +/- 0.09	0.08 +/- 0.09

L200 This is an awkward definition of stratiform clouds. It should rather be defined by some homogeneity criterion, such as temporally constant LWP and cloud base altitude over some time from the ground-based measurement (or distance from the satellite).
We thank the reviewer for this comment. We agree that the previous wording could be interpreted as providing an incomplete definition of stratiform clouds. In the revised manuscript, we have clarified that no single criterion is intended to define stratiform clouds. Instead, we now state that a combination of complementary filtering criteria is applied to isolate shallow boundary-layer stratiform clouds suitable for aerosol-cloud interaction studies. Specifically, the LWP threshold is used to exclude optically thick cloud systems, the cloud-base criterion ensures that the cloud is connected to the boundary layer while excluding fog, and the cloud geometric thickness criterion removes deep convective clouds. We have revised the text accordingly to make this definition clearer.

Revised text: *"To isolate shallow boundary-layer stratiform clouds, we applied a set of complementary filtering criteria. Only clouds with a liquid water path (LWP) below 200 g m^{-2} (Pruppacher and Klett, 2010), a cloud base located above 100 m and below the estimated boundary-layer height, and a cloud geometric thickness smaller than 1 km were retained. These criteria are applied jointly to select shallow stratiform clouds while minimizing the influence of fog, free-tropospheric clouds and deep convective systems."*

L201 How does a cloud base height threshold filter out rain? There should be a better way to define precipitation from the ground-based measurements.
We thank the reviewer for this insightful comment. Our original rationale for applying the cloud-base-height threshold was based on the assumption that precipitation could significantly contaminate the RPG-HATPRO retrieved LWC profiles and that an additional filtering was therefore required. We acknowledge that this statement was misleading. The cloud-base criterion is now only used to exclude fog and to ensure that the cloud is connected to the

boundary layer where the aerosol measurements are performed. Following the reviewer's suggestion, we have removed the justification related to precipitation and revised the manuscript accordingly.

Revised text: *"To isolate shallow boundary-layer stratiform clouds, we applied a set of complementary filtering criteria. Only clouds with a liquid water path (LWP) below 200 g m* $^{-2}$ *(Pruppacher and Klett, 2010), a cloud base located above 100 m and below the estimated boundary-layer height, and a cloud geometric thickness smaller than 1 km were retained. These criteria are applied jointly to select shallow stratiform clouds while minimizing the influence of fog, free-tropospheric clouds and deep convective systems."*

L214 This is now a second, redundant, definition of stratiform clouds. Why would that criterion be "effective"?
We thank the reviewer for pointing out this redundancy. We agree that the previous version of the manuscript gave the impression that cloud geometric thickness alone was used to define stratiform clouds. This was not our intention. In the revised manuscript, the cloud thickness criterion is no longer presented as an independent definition but as one of the complementary filters used to isolate the targeted cloud population. The geometric thickness threshold (< 1 km) is introduced solely to exclude deep convective clouds, thereby reducing the influence of convective mixing and dynamical lifting on the retrieved aerosol-cloud interaction relationships. The manuscript has been revised to make this distinction explicit.

Revised text: *"To isolate shallow boundary-layer stratiform clouds, we applied a set of complementary filtering criteria. Only clouds with a liquid water path (LWP) below 200 g m* $^{-2}$ *(Pruppacher and Klett, 2010), a cloud base located above 100 m and below the estimated boundary-layer height, and a cloud geometric thickness smaller than 1 km were retained. These criteria are applied jointly to select shallow stratiform clouds while minimizing the influence of fog, free-tropospheric clouds and deep convective systems."*

L215 Why do the authors want to minimise convective mixing? Just a few lines earlier (L201) the aim was, on the contrary, to ascertain the aerosol in the boundary layer mixes into the clouds.

This approach minimises the influence of strong convective activity and dynamic lifting on cloud development. The objective is to retain cases in which boundary-layer aerosols are representative of the cloud-forming air mass while avoiding unstable conditions where cloud evolution is primarily controlled by dynamical processes rather than aerosol perturbations.

L216 Why this lower bound on the BLH? Just a little earlier (L208) the authors explained they want to limit themselves to stable conditions. Here it now seems that on the contrary, convectice situations are sought.
We agree that the formulation was unclear and could be misinterpreted as favouring convective conditions. The lower bound of 500 m was not introduced to select convective cases, but rather to ensure a physically meaningful estimation of the boundary layer height from the retrieval/diagnostic method. Very shallow BLH values (below ~500 m) are often associated with retrieval uncertainties, nocturnal surface-layer artefacts, or situations where the boundary layer structure cannot be reliably resolved.

Our objective remains to focus on stable conditions and low-level cloud situations. The BLH criterion was therefore applied as a quality control threshold rather than a dynamical classification of atmospheric regime.
Revised text: *"Moreover, we only retained cases with a boundary layer height (BLH) larger than 500 m to ensure a robust and reliable estimation of the boundary layer structure, avoiding situations where very shallow BLH retrievals are uncertain, this is a data-quality threshold, not a dynamical selection criterion."*

L216 Didn't the earlier filtering criteria already reduce the data to 18% of its original size (L203)?
When giving these reduction estimations, all filters are applied independently. So, the 3.1 filter is reducing the original data set by 19% while the filter 3.2 is reducing the original data set by 40%.

L222 I find this criterion puzzling. How is the 60% motivated? This doesn't seem to me similar at all, if stratiform, and therefore spatially homogeneous clouds, are considered. If the point is to avoid overlying clouds, aren't there better criteria to filter from the satellite retrievals, e.g. selecting scenes with liquid clouds only?
The initial wording may have implied that the 60% threshold is a physically motivated criterion which it is not. Instead, this threshold serves as an empirical consistency check between ground-based (MWRP) and satellite-derived (CLAAS-3) liquid water path (LWP) retrievals. Its purpose is to mitigate discrepancies arising when the two instruments sample fundamentally different cloud scenes — such as those caused by horizontal displacement, sub-pixel inhomogeneity, or vertically complex/multilayer cloud structures. The 60% value was selected as a pragmatic balance between preserving a sufficiently large dataset and ensuring a reasonable degree of agreement between the two independent measurements. We explicitly acknowledge its empirical nature in the revised manuscript and have removed any suggestion that it directly infers cloud morphology or stratification.
We have clarified this point in the revised manuscript and removed any implication that this criterion directly identifies cloud morphology or stratification: *"To ensure consistency between ground-based (MWRP) and satellite (CLAAS-3) liquid water path (LWP) retrievals, we applied a 60% relative difference threshold as an empirical filter. This threshold was selected pragmatically to retain 45% of the available cases while reducing observations for which the two independent measurements disagree substantially, likely because they sample different cloud scenes (e.g., due to horizontal displacement or vertically heterogeneous cloud structures). When applied independently, this filter removes approximately 45% of the original dataset."*

L223 Now another number for how much data is left is provided, 45%. Why this? What with the 18% (L203) and the 40% (L216)?
These percentages do not correspond to a sequential application of filters on the same progressively reduced dataset. Instead, each filter was applied independently to the full dataset in order to evaluate its individual impact and identify the most restrictive criteria. The aim was not to construct a cumulative filtering chain, but rather to assess how each physical constraint (cloud-type selection, boundary layer height, and satellite-ground consistency) independently reduces the available sample. This independent filtering approach was chosen specifically to avoid conflating the effects of different physical assumptions and to provide transparency regarding the sensitivity of the dataset to each criterion.
We have revised the text to explicitly clarify that all filters were applied independently, and we now report the corresponding retained fractions in a more structured way to avoid ambiguity.

L228 This is true for well-mixed boundary layers, but weren't criteria applied that tried to avoid these?

A well-mixed boundary layer would imply neutrally stable conditions on the threshold of being convectively unstable, and therefore unsuitable for comparisons with stratiform clouds. We filter out all such cases using the atmospheric stability criterion we employ. When the boundary layer is more stable, it is reasonable to anticipate a difference between aerosol concentrations measured at the surface and those at the cloud base. Cloud susceptibilities to aerosol burden are typically quantified, as we have done, in terms of relative changes in cloud quantities and aerosol qualities. Thus, susceptibility calculations are expected to be relatively insensitive to absolute differences in aerosol concentration. Accordingly, changes in surface aerosol concentrations are assumed to scale proportionately with those at cloud base. For example, a 10% increase in surface aerosol concentrations would correspond with at 10% increase in cloud base aerosol concentrations, even if the atmospheric profile were unstable.

Revised text : "*Aerosol properties measured at the surface are used as a proxy for the aerosol population interacting with the cloud base. This assumption is made for practical reasons, as direct observations at cloud base are not available. When the boundary layer is more stable, it is reasonable to anticipate a difference between aerosol concentrations measured at the surface and those at the cloud base. However, since cloud susceptibilities to aerosol burden is quantified in terms of relative changes, the metric is relatively insensitive to absolute differences in aerosol concentration: changes in surface aerosol concentrations are assumed to scale proportionately with changes at cloud base. »*

L232 How are these constrained if no information about CCN is available?

The CCN-related assumption is indeed not based on direct measurements of cloud condensation nuclei, which are not available in ATOLL nor at most stations of the ACTRIS network. Instead, we use an aerosol size-based criterion as a first-order proxy for CCN activity. Specifically, particles with effective radii smaller than 39 nm are assumed not to act as CCN under typical atmospheric supersaturation conditions encountered in stratiform cloud environments.

This assumption is consistent with Köhler theory and numerous observational studies showing that particles in the Aitken mode generally require higher supersaturations than those typically present in boundary-layer stratiform clouds to activate into cloud droplets, whereas CCN activity becomes significant for larger accumulation-mode particles (typically above 50 nm depending on aerosol composition and hygroscopicity, Seinfeld and Pandis, 2016; Dusek et al., 2006; Petters and Kreidenweis, 2007; Kulmala et al., 2004).

L239 (Table 2) Why this very low number of observations in the initial database? The upper bound would be 4 years * 365 days/year * 12 hours daylight / day * 4 15-min steps / hour = 70.000 time steps. Why only an order of magnitude fewer data points?

That would be true if each instrument was providing data 24 hours a day and 7 days a week all year long. But instruments may suffer some technical issues independently from each other. The size of the data set is determined from data availability of all instruments on each platform simultaneously. Unfortunately, it decreases the original dataset.

L245 My own experience with analysing clouds and sensitivities of microphysical quantities is that one needs a few thousands to obtain robust results. The remaining 188 data points are too few to yield meaningful insights.

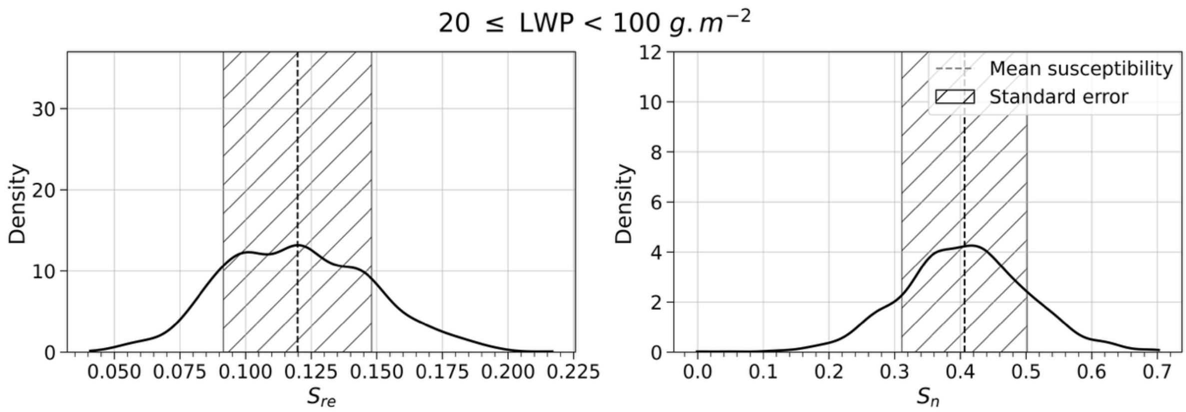
The authors fully agree that a larger number of samples is generally required for robust statistical analyses of cloud microphysical relationships. We would like to clarify that the primary objective of this study is not to derive quantitative climatological conclusions, but rather to demonstrate the feasibility and potential of the proposed combined ground-based and satellite-based methodology for investigating low-cloud microphysical sensitivities to aerosol burden.

The final sample size of 201 cases reflects the application of a set of physically motivated filtering criteria designed to ensure consistency between instruments and to minimize retrieval uncertainties. While this inevitably reduces the available statistics, it allows us to focus on a well-constrained subset of observations. We acknowledge that this dataset is not sufficient for deriving fully robust statistical relationships, and we have clarified this point in the revised manuscript. Future work will focus on extending this analysis to larger datasets, including multiple ACTRIS sites and ARM observations, in order to increase the sample size and enable more comprehensive scientific analyses.

L252 The authors should provide some evidence for this statement. For example, they could assess what happens to the conclusions if they further reduce the number of data points. If the results remain unchanged for a reduction by a further quarter or third, that would be convincing.

To address this comment, we performed a sensitivity analysis by randomly excluding 25% of the final dataset 1,000 times and recalculating S_n and S_{re} for each realization. The resulting probability density functions (PDFs) of S_n and S_{re} for both low and high LWP ranges are shown in Figure R1.

For S_{re}, the PDFs are centred at 0.12 and 0.08 for the low and high LWP ranges, respectively, whereas the values obtained from the complete dataset are 0.12 and 0.08. These results indicate that the retrieved S_{re} values are only weakly affected by the random removal of one quarter of the data, demonstrating the robustness of the analysis despite the limited dataset. Similar conclusion can be drawn for S_n. The PDFs are centred at 0.41 and 0.33 for the low and high LWP ranges, respectively, which are identical to the values retrieved from the complete dataset. This further supports the robustness of the estimated S_n values with respect to random reductions in the dataset. The Figure R1 and part of this discussion is now included in the supplementary material.



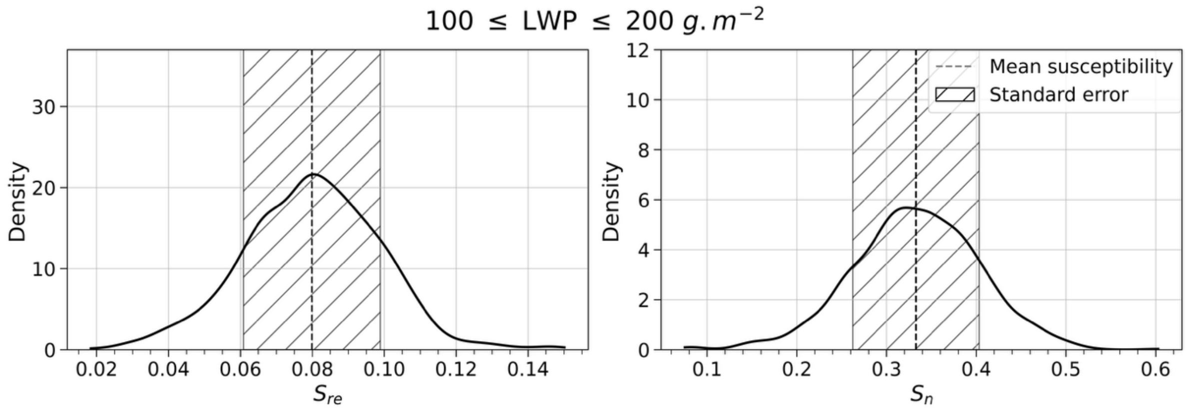


Figure R1: **Probability density functions of S_n and S_{re} obtained from 1,000 random realizations in which the final dataset was artificially reduced by 25%. S_n and S_{re} were recalculated for each realization.**

Revised text: "*To evaluate the robustness of our results, we performed a sensitivity analysis by randomly excluding 25% of the dataset 1,000 times and recalculating S_n and S_{re} for each realization. The mean retrieved S_{re} values were 0.12 and 0.08 for the low and high LWP ranges, respectively, while the corresponding mean S_n values were 0.41 and 0.33. These results indicate that our analysis is robust to random reductions in the dataset, despite its limited size.*"

L263 I agree that Panel (f) is useful, to demonstrate there is very little data from Aug – Oct but otherwise an approximately equal seasonal distribution. What about the distribution between the years? For (a) – (e) it would be more useful to show the overall histrograms, it seems not to make sense to show the seasonal cycle. There is simply not enough data for meaningful conclusions about this.
We agree with the reviewer that the distribution over the years is interesting, to ensure that no single year dominates the dataset. The figure can be found below. We observe that there is a large amount of data from 2023 (~45% of the filtered data). The unfiltered dataset (not shown) also shows more data in 2023, but the difference is less pronounced. We observed that all the other parameters do not show strong variations over the years. Stability does change slightly, but the impact is limited on the dataset, since the cases are still considered as stable according to our filtering method.

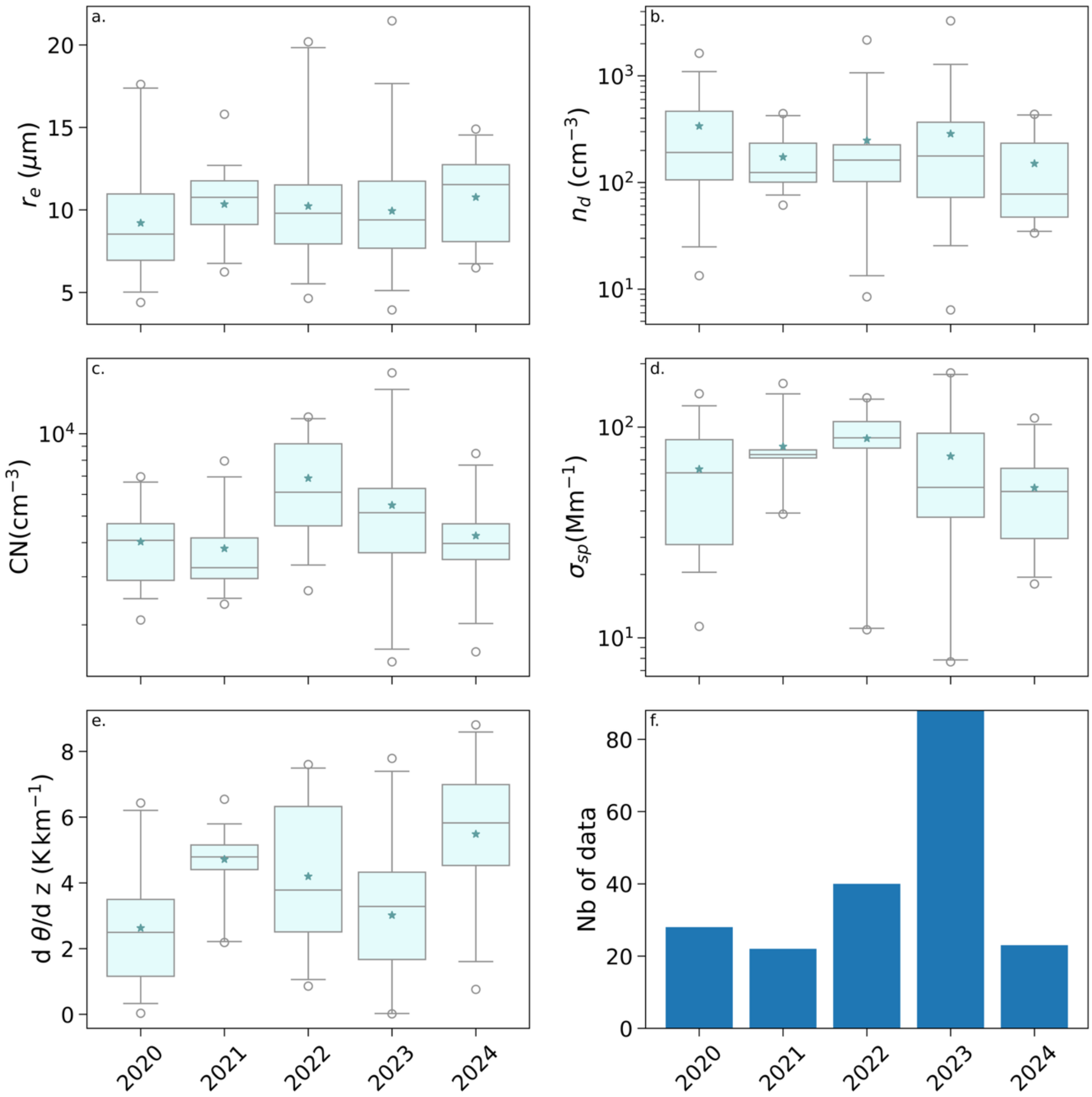


Fig R2: **Interannual variation of (a) cloud droplet effective radius (r_e), (b) cloud droplet number concentration (n_d), (c) particle concentration (*CN*), (d) dried aerosol light scattering ($\sigma_{sp,525nm}$), (e) potential temperature vertical gradient (*dθ/dz*), and (f.) the number of observations (Nb of data) over the ATOLL station when all filters described in section 3 are applied. For subfigures a to e, the box represents the interquartile range (25th–75th percentiles), with the median indicated by a horizontal line and the mean by the star. Whiskers extend to the 1st and 99th percentiles and outliers of the whiskers are shown as individual points.**

Revised text: "*The interannual distribution of the dataset is provided in the supporting information (Figure S2). While the year 2023 accounts for a larger fraction of the filtered dataset (~40%), most retrieved parameters do not show strong variations across years, with the exception of a slight increase in droplet size in 2023, whose origin remains unclear.*"

L271 It is not understandable why the anyway very small dataset is further split into two. What is the motivation?

We agree that the final sample size is limited, and that splitting the dataset further reduces the number of observations available in each subset. However, liquid water path (LWP) is a key controlling variable in aerosol-cloud interaction studies. The classical Twomey and Albrecht effects are generally evaluated under approximately constant LWP conditions to reduce the influence of cloud macrophysical variability on cloud microphysical properties. Furthermore, several studies have shown that aerosol cloud susceptibilities exhibit a strong dependence on LWP. In particular, Zheng et al. (2020) demonstrated systematic variations in cloud susceptibility as a function of LWP.

For this reason, we chose to separate the dataset into LWP ranges commonly used in the literature. Although the limited sample size prevents robust statistical conclusions, this stratification allows a more meaningful comparison with previous studies and avoids mixing cloud regimes characterized by substantially different liquid water contents.

We have clarified this motivation in the revised manuscript and explicitly acknowledge the limitations associated with the reduced number of observations in each bin.

Revised text : *"In this section, the cloud droplet number and effective radius susceptibilities to aerosol loadings will be estimated. Although dividing the dataset into liquid water path (LWP) classes reduces the number of observations available in each subset, this stratification is essential in aerosol-cloud interaction studies. LWP is a key controlling variable, and the classical Twomey and Albrecht effects are generally evaluated under approximately constant LWP conditions to minimize the influence of cloud macrophysical variability on cloud microphysical properties. Moreover, several studies have shown that aerosol-cloud susceptibilities strongly depend on LWP. In particular, Zheng et al. (2020) reported systematic variations in cloud susceptibility across different LWP regimes. Consequently, separating the analysis by LWP provides a more physically meaningful interpretation of the retrieved susceptibilities, despite the reduction in sample size."*

L276 "higher scattering coefficients are often associated with increased aerosol concentrations" why is that a meaningful statement? Are there situations where the opposite could be true?

The original statement was too general and has been revised. Aerosol scattering coefficient can be influenced both by aerosol abundance and by particle hygroscopic growth at elevated relative humidity. In the present study, however, measurements were performed following ACTRIS protocols, which require aerosol sampling under dry conditions (RH < 40%). Under these conditions, variations in the scattering coefficient primarily reflect changes in aerosol loading and/or aerosol size distribution rather than changes induced by water uptake.

We have clarified this point in the manuscript.

« Under the ACTRIS dry-sampling conditions (RH < 40%), higher scattering coefficients generally reflect increased aerosol loading and/or changes in aerosol size distribution. »

L278 "enhanced scattering is indicative of larger particle sizes or higher particle number concentrations" same question

The aerosol scattering coefficient depends on both particle number concentration and particle size. Because scattering efficiency increases strongly with particle diameter, a relatively small increase in the concentration of coarse particles can produce a substantial increase in the measured scattering coefficient. Conversely, changes in the concentration of smaller particles generally require much larger number increases to produce a comparable scattering signal.

Therefore, under dry conditions (RH<40%), an increase in scattering coefficient may result either from an increase in aerosol number concentration, from a shift toward larger particle sizes, or from a combination of both effects. We have revised the manuscript to clarify this interpretation.

« Enhanced scattering can result from either increased aerosol number concentrations, a shift toward larger particle sizes, or a combination of both, since aerosol scattering depends on both particle abundance and particle size distribution. »

L298 It may be useful to clarify that a hard upper bound for S_n is 1 (every aerosol activates into a droplet), unless mistakes in the definition are made (e.g. 1.59 is physically nonsense).

We agree that when S_n is defined as the ratio of cloud droplet number concentration to CCN number concentration, its physical upper bound is indeed 1, since the number of activated droplets cannot exceed the available CCN. However, in this study, aerosol concentration is not represented by CCN number concentration but by the aerosol scattering coefficient, which is an optical proxy for aerosol loading. Because the denominator is not a particle number concentration, S_n is not bounded by 1, and values greater than 1 are physically possible and do not indicate an error in the calculation.

L300 Schmidt et al. ACP 2015 (doi:10.5194/acp-15-10687-2015) compiled a list of S_n, maybe it is useful to refer to this here.

We thank the reviewer for this suggestion, part of the missing references were added here using figure R3 and R4.

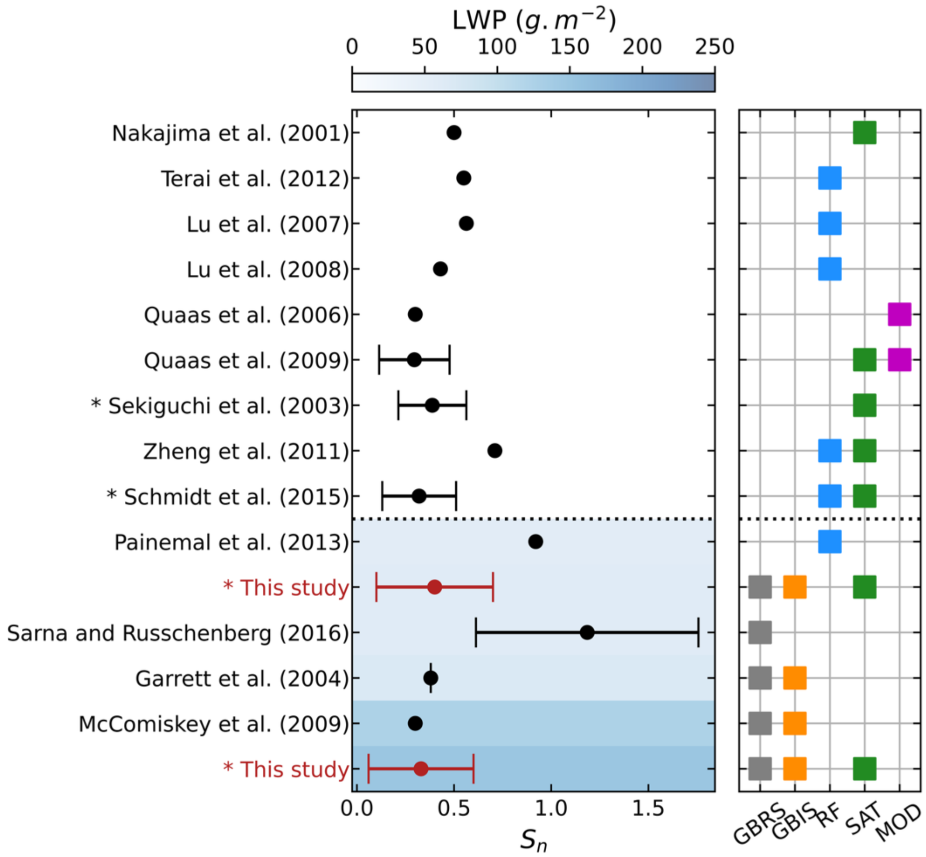
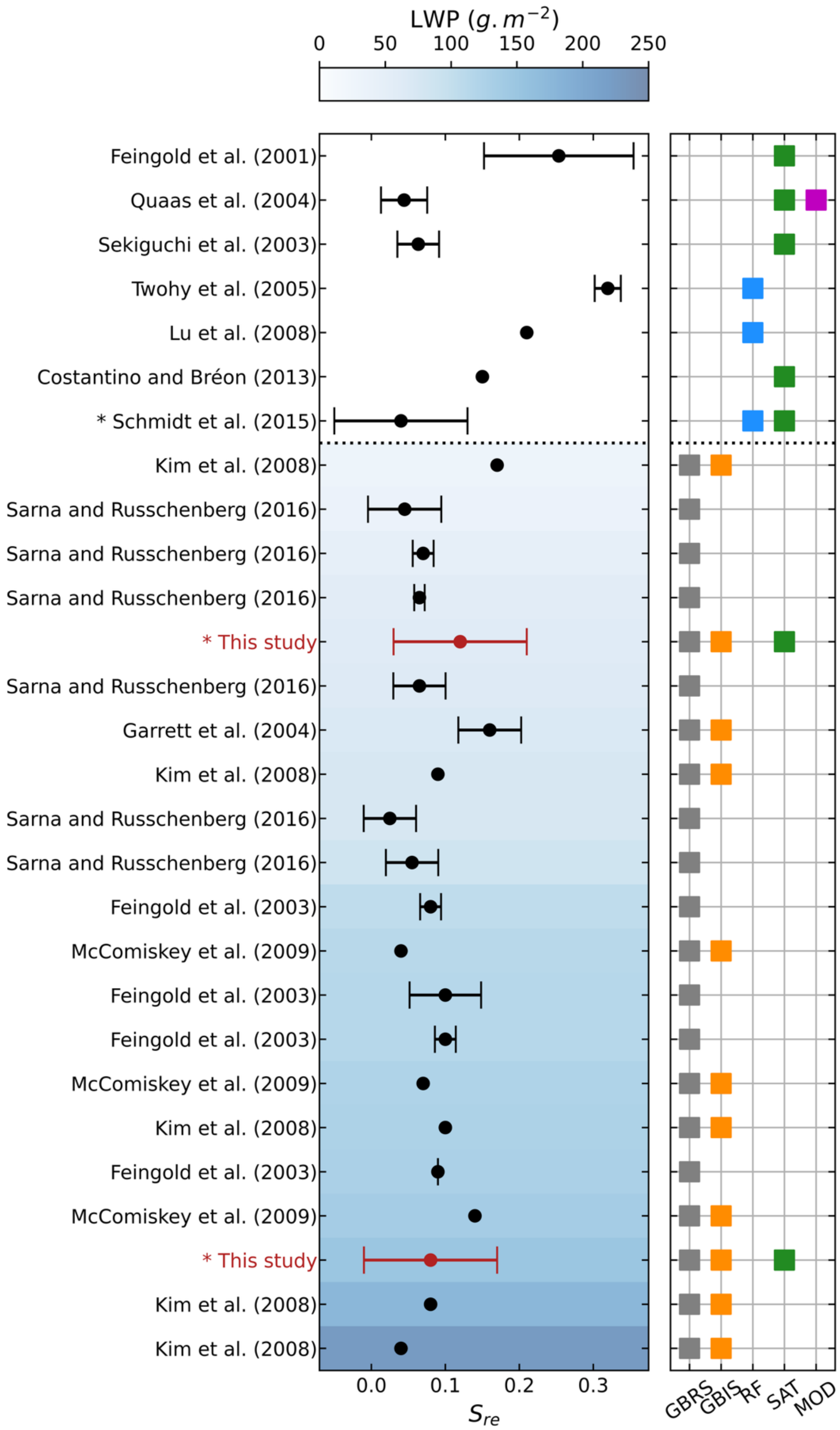


Figure R3: S_n derived from a combination of either ground-based measurements (in-situ: orange squares; remote sensing: grey squares), airborne measurements (blue squares), satellite retrievals, and/or simulation results (purple square). The white-to-blue shaded area represents the mean liquid water path (LWP) range covered across all studies. Stars indicate studies that directly reported uncertainties in their analysis, while for the remaining studies, uncertainties were recalculated using author provided datasets for different regions or seasons. The vertical dashed lines correspond to the physical upper bound for S_n (all aerosols activate into a droplet).



LWP $(g.m^{-2})$
0
50
100
150
200
250
Feingold et al. (2001)
Quaas et al. (2004)
Sekiguchi et al. (2003)
Twohy et al. (2005)
Lu et al. (2008)
Costantino and Bréon (2013)
* Schmidt et al. (2015)
Kim et al. (2008)
Sarna and Russchenberg (2016)
Sarna and Russchenberg (2016)
Sarna and Russchenberg (2016)
* This study
Sarna and Russchenberg (2016)
Garrett et al. (2004)
Kim et al. (2008)
Sarna and Russchenberg (2016)
Sarna and Russchenberg (2016)
Feingold et al. (2003)
McComiskey et al. (2009)
Feingold et al. (2003)
Feingold et al. (2003)
McComiskey et al. (2009)
Kim et al. (2008)
Feingold et al. (2003)
McComiskey et al. (2009)
* This study
Kim et al. (2008)
Kim et al. (2008)
0.0
0.1
0.2
0.3
S_{re}
GBRS
GBIS
RF
SAT
MOD

Figure R4: S_{re} derived from a combination of either ground-based measurements (in-situ: orange squares; remote sensing: grey squares), airborne measurements (blue squares), satellite retrievals, and/or simulation results (purple square). The white-to-blue shaded area represents the mean liquid water path (LWP) range covered across all studies. Stars indicate studies that directly reported uncertainties in their analysis, while for the remaining studies, uncertainties were recalculated using author provided datasets for different regions or seasons. The vertical dashed lines correspond to the physical upper bound for S_{re} (all aerosols activate into a droplet).

Revised text for S_n: *"These results are comparable in magnitude, though somewhat lower, than those reported in the literature (Figure 5). At least 9 studies reported in Fig. 5 does not provide LWP range using a combination of either airborne measurement, satellite and simulation (Nakajima et al., 2001; Terai et al., 2012; Lu et al., 2008, 2009; Quaas et al., 2006, 2009; Sekiguchi et al., 2003; Zheng et al., 2011, Schmidt et al., 2015). They reported Sn values from 0.078 and 0.787. Sarna and Russchenberg (2016), using ground-based remote sensing retrievals of droplet number concentration and integrated attenuated backscatter, found S_n values between 0.78 and 1.59 (above the upper S_n physical bound) for two case studies with LWP between 30 and 90 g m^{-2}. Furthermore, Painemal et al. (2013) reported a Sn of 0.92 in a range of LWP centred at 55 g m^{-2} using onboard in-situ instruments. In a larger statistical analysis, McComiskey et al. (2009) reported substantially lower S_n values, ranging from 0.51 to 0.64 for LWP between 107 and 143 g m^{-2}, using MWRP-derived n_d and aerosol light scattering from a nephelometer."*

Revised text for S_{re}: *"Figure 7 presents previously reported values of S_{re} derived from a combination of ground-based observations (in situ: orange squares; remote sensing: grey squares), airborne measurements (blue squares), satellite retrievals, and/or numerical simulations (purple square). The results from this study (red symbol) are comparable in magnitude to those reported in the literature. Higher S_{re} values (0.15–0.33) are generally found in studies where the LWP range was not constrained (Feingold et al., 2001; Twohy et al., 2005; Lu et al., 2008; Costantino and Bréon, 2013). Feingold et al. (2001) and Twohy et al. (2005) reported the highest retrieved S_{re} values, ranging from 0.12–0.38 and 0.307–0.332, respectively. The elevated S_{re} values reported by Feingold et al. (2001) may be attributed to smoke-affected conditions over Brazil, whereas those reported by Twohy et al. (2005) may reflect the limited sample size (only 16 observations) and the use of aerosol number concentration measurements rather than aerosol optical properties. For low LWP conditions (below 100 g m^{-2}), Sarna and Russchenberg (2016) reported S_{re} values ranging from 0.00 to 0.09, in good agreement with our findings. In contrast, Garrett et al. (2004) retrieved a higher S_{re} value of 0.17 over the Arctic, suggesting that aerosol-cloud interactions may vary depending on the environmental conditions in which the observations are made. For higher LWP conditions (above 150 g m^{-2}), reported S_{re} values typically range between 0.02 and 0.17 (Feingold et al., 2003; Kim et al., 2008; McComiskey et al., 2009), which encompasses the values obtained in the present study. Interestingly, based on cloud-radar estimates of r_e and Raman lidar aerosol extinction, Feingold (2003) derived S_{re} values of 0.03–0.16 from seven cases observed at the DOE ARM Southern Great Plains site in Oklahoma. Using the same methodology at the same site, Kim et al. (2008) reported comparable S_{re} values (0.04–0.17) from a much larger, three-year dataset of continental stratus clouds. This close agreement suggests that, when the retrieval methodology and observational setting are consistent, the number of observations alone may not substantially influence the retrieved S_{re} values. »*

L302 / Section on r_e: Is this not entirely redundant with the analysis of n_d? i.e., isn't n_d computed from the retrieved r_e?
The reviewer is correct that n_d is derived from the retrieved r_e, so the analyses are not fully independent. Our primary interest, however, is in quantifying the susceptibility of cloud properties to aerosol perturbations. Since the method provides consistent results for r_e, it is expected to yield consistent results for n_d as well. Nevertheless, we chose to present both analyses because the cloud droplet effective radius susceptibility S_{re} is the quantity most commonly reported in the literature. Including both r_e and n_d therefore allows a direct comparison with previous studies while also demonstrating the consistency of our retrieval-based approach.

L331 A "case study" would be for one case only (but in-depth), here there are 188 cases statistically analysed.
This has been removed