We thank both of the reviewers for taking the time to read and comment on the paper; your comments have helped to greatly improve the paper. The reviewer comments are repeated below in green italics and our responses are in black.

Responses to Reviewer 1

The manuscript describes a method that, in principle, corrects errors in adiabatic satellite cloud droplet number concentration (Nd) due to the inconsistency of utilizing satellite cloud effective radius (r_e) that represents values slightly below the cloud top, whereas satellite cloud optical depth (tau) fully captures the optical thickness of the clouds. To achieve this goal, the authors simulate a number of idealized cloud profiles with a 1D radiative transfer model, and then retrieve r_e and tau from the synthetic reflectances. Next, the authors derive an "effective" tau that corresponds to the optical thickness where the retrieved r_e and the synthetic r_e match each other (the vertical penetration effect). They use the difference between the retrieved and the effective tau (applying a fit to their theoretical calculations) to quantify the error in MODIS-based Nd that does not account for the fact that the satellite r_e is not exactly that at the cloud top due to the vertical photon penetration, which is in turn dependent on the sensor wavelength and the specific thickness of the cloud (and probably solar zenith angle and viewing geometry).

The manuscript makes an interesting use of the results in Platnick (2000), which shows that the retrieved r_e should differ from the observed r_e by a few um (or less) due to the photon penetration. The manuscript is concise and well-written, however when I first browsed the paper, I got confused about whether the authors wanted to show a real satellite bias in tau (and Nd) or a methodological bias (I realized it was the latter).

*My fundamental criticism of Grosvenor et al. is that, from a remote sensing point of view, the problem is not that the satellite tau should be reduced because r_e is not at the cloud top. Instead, it is that r_e is smaller than the observed r_e at the top due to vertical stratification, and probably r_e should be somehow increased (i.e. r_e drives the uncertainty in Nd). This is the correct interpretation, as it is well known from the early work by Nakajima (King and co-authors) that satellite tau is almost insensitive to the cloud vertical structure, and only r_e can be greatly affected by the vertical stratification. So, the Nd bias should be expressed in terms of a delta r_e. Another inconsistency (related to my previous comment) is with the use of the (pseudo) adiabatic model, which if I interpret correctly, it implies that the liquid water path (LWP) is proportional to r_e*tau.*
So, any error calculation applied to Nd has to be also valid for LWP.

However, if we apply equation (13) to LWP, i.e.:

LWP_uncorrected/LWP_corrected=(tau/(tau-dtau))

Using a dtau=4.5 for tau=10 (figure 1a), then LWP_uncorrected/LWP_corrected=10/6.5=1.54. A 54% overestimation in LWP is clearly a mathematical contradiction. On the other hand, if, for instance, we utilize the results in Platnick (2000) for a cloud top r_e =12 um, and a retrieved r_e= 10.7 um (2.1 um wavelength), we get:

LWP_uncorrected/LWP_corrected=r_uncorrected/r_corrected=10.7/12=0.89.

That is, the retrieved r_e yields an underestimation of LWP. Again, this result points to a main reasoning problem in the manuscript, which is, the error should not be expressed in terms of tau.

We were originally considering the special case of the Nd retrieval that makes the assumption that r_e is at cloud top and the application to LWP was not considered; the original method using the tau correction only applies to the Nd retrieval and not to the LWP calculation, which we did not make clear. We originally chose to use a tau correction since the correction for tau seemed simpler to parameterize and less prone to uncertainties in the parameterization. Further work following from your suggestions (and those of the other reviewer) has shown that dre/re can also be modelled fairly accurately with a fitted curve and so we now also present results using this formulation and make some estimates of the LWP bias too. The following new figures are included showing 2D histograms of the re correction (divided by re)vs the optical depth, along with the parameterized fit to the data (as previously shown for the tau correction) :-

(a) 2.1 μm

(b) 3.7 μm
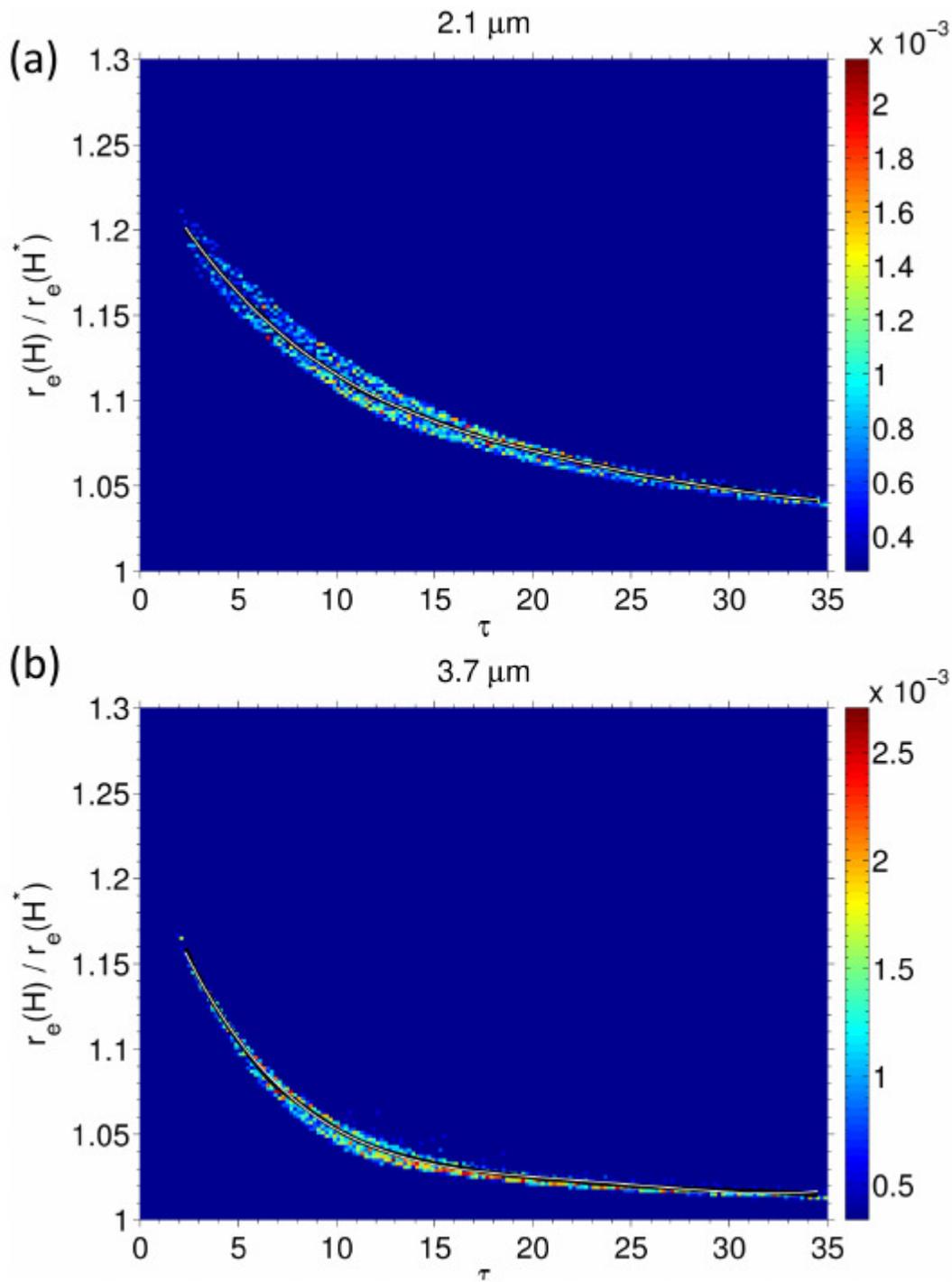
Lastly, the authors say that there are several other errors that can bias r_e and tau. This is a key statement, and a literature review will show that biases in r_e are not dominated by the cloud vertical stratification (I am not aware of any studies that actually show an adiabatic signature in the satellite r_e bias). For instance, If one calculates the difference between MODIS r_e at 2.1 um and 3.7 um, the difference is positive everywhere over the ocean (the difference can be larger than 5 um, see Fig. 10 in

Zhang and Platnick, 2011). This result suggests that the error discussed in Grosvenor et al. is negligible. So, I find it surprising that the authors found errors up to 50 % in Nd (Figure 6), which is very large. Since their results are only valid in a plane parallel world (sub-pixel variability is not accounted for) and with the use of idealized profiles, the validity of the correction cannot be demonstrated. The authors do discuss some of these issues but, unfortunately, the main concern remains, that is, it is unclear that the correction will yield an improved estimate of cloud droplet number concentration.

We agree that other r_e biases are important for Nd retrievals and are probably of equal or stronger magnitude than the changes in r_e due to the vertical profile changes. In the original paper we wrote (p. 16, line 10) :-

- "It is also clear that the suggested correction for the vertical penetration effect should only be applied to the retrievals of Nd with consideration of other bias sources. These other potential error sources are numerous and include re biases due to sub-pixel heterogeneity (Zhang and Platnick, 2011; Zhang et al., 2012, 2016); 3D radiative effects (Marshak et al., 2006);…" (etc.)

And on p.10, line 12:-

- "Thus, the application of the correction described in this paper in isolation has the potential to enhance any negative bias in Nd caused by a positive re bias."

we caution the reader that the bias correction should only be applied after other biases have been accounted for. We realize that this limits the current usefulness of the correction until the other biases have been quantified - we therefore have added the following text to the and of the above :-

- "; it is thus recommended that the bias correction is not applied until the other error sources have been fully characterized."

Also we have modified the following paragraph in the conclusions to make this point clearer and to recommend restriction to low heterogeneity situations (see later for the justification for this) :-

We caution that the correction for the N_d and LWP vertical penetration biases presented here should only be considered in combination with corrections for other biases that affect τ and r_e. Zhang et al. (2016) suggest a correction for the sub-pixel heterogeneity bias effect, but corrections may not currently exist for all biases and it is likely that some unidentified biases still exist. Therefore, we recommended that our correction is currently only applied to homogeneous cloud scenes in order to minimize possible entanglements with biases resulting from heterogeneity effects, which are not accounted for. Such conditions can be obtained by limiting retrievals to associated heterogeneity (H_σ) values (available in MODIS MYD06 Collection 6 products) to less than about 0.1. Otherwise N_d and LWP biases could be made worse; for example, in situations where the fortuitous cancellation of opposing errors leads to initially small N_d errors. The latter was suspected to have occurred for the comparison between MODIS N_d retrievals and in-situ aircraft observations as presented in Painemal and Zuidema (2011). We showed that the SE Pacific, which is the region examined in that study, had a mean vertical penetration depth error of 35% suggesting that another unidentified N_d bias may have been operating in order to give good agreement.

However, we feel that it is useful and important to quantify the vertical penetration bias nevertheless and to suggest ways to remedy it (albeit in the sense of an idealised retrieval with no other bias sources). The addition of a previously unconsidered underlying bias is important since, for example, it

would disrupt the cancellation of the other errors that led to the good agreement between aircraft and MODIS Nd seen in Painemal and Zuidema (2012), possibly suggesting that another unaccounted for error source exists.

The question is whether the vertical penetration effect occurs in addition to the other errors; e.g., whether (scenario A) the presence of cloud heterogeneity somehow prevents the effects of the vertical stratification from influencing the retrieved r_e and makes it irrelevant, or whether (scenario B) the vertical stratification is influencing r_e in the expected way (i.e. a tendency to cause re_37>re_21), but with a corresponding counter-influence in the opposite direction due to heterogeneity. We argue for scenario B, but it is hard to prove this within the scope of this study, since it would likely require computationally expensive 3D radiative transfer modelling of known cloud fields (e.g., from LES models), followed by r_e and tau retrievals.

Some evidence for scenario B is that it may explain why VOCALS aircraft measurements showed that re_21 and re_37 were very similar; it is possible that sub-pixel (or other) heterogeneity effects tended to increase re_21 relative to re_37, but that the vertical penetration effect has the opposite tendency, resulting in overall similar values. We also note that there are many situations when the expected vertical stratification of r_e does occur (i.e. re_37>re_21), as demonstrated in the following figure (included in the revised paper):-
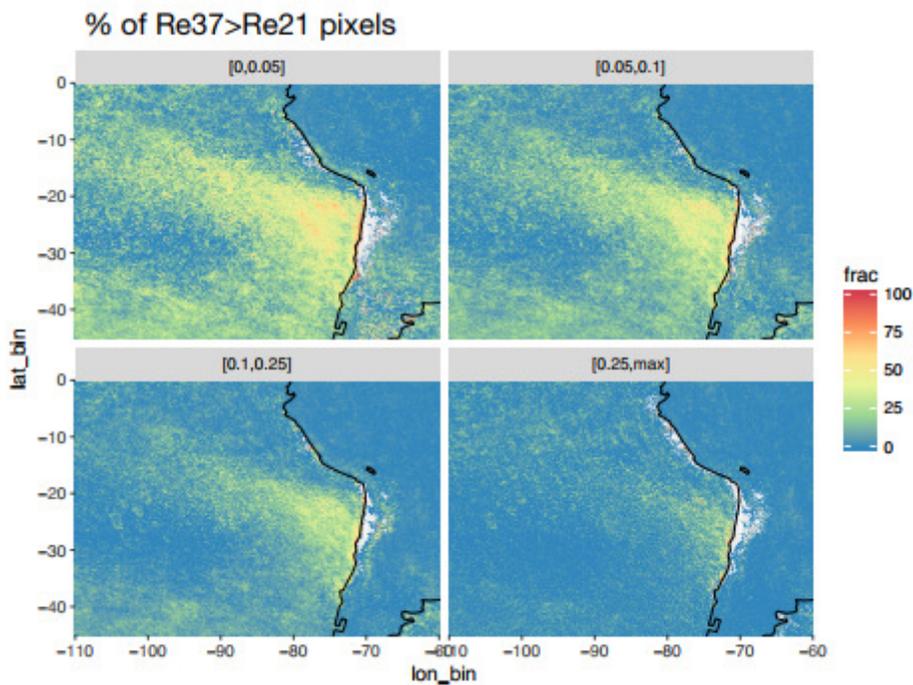
It shows the percentage of pixels where re_37>re_21 for 90 days (Jan, Feb, Mar) of 0.1° resolution Collection 6 MODIS observations (single layer liquid clouds only; filtered to exclude tau<5 and partially cloudy pixels). The four panels are for four different bins of the heterogeneity parameter (the standard deviation of the 250m resolution 0.86um reflectance divided by the mean reflectance) with bin ranges labelled above the panels. It is clear that for many regions the relative r_e values that are consistent

with an adiabatic profile occur more than 50% of the time, particularly when the cloud heterogeneity is low. Similarly, the Bennartz (2017) Nd dataset requires that re_37>re_21 in order for a datapoint to be included in the dataset indicating that there are a lot of times when this is the case.

We have also added this figure, which shows the ratio between re_37 and re_21 for an example MODIS scene :-
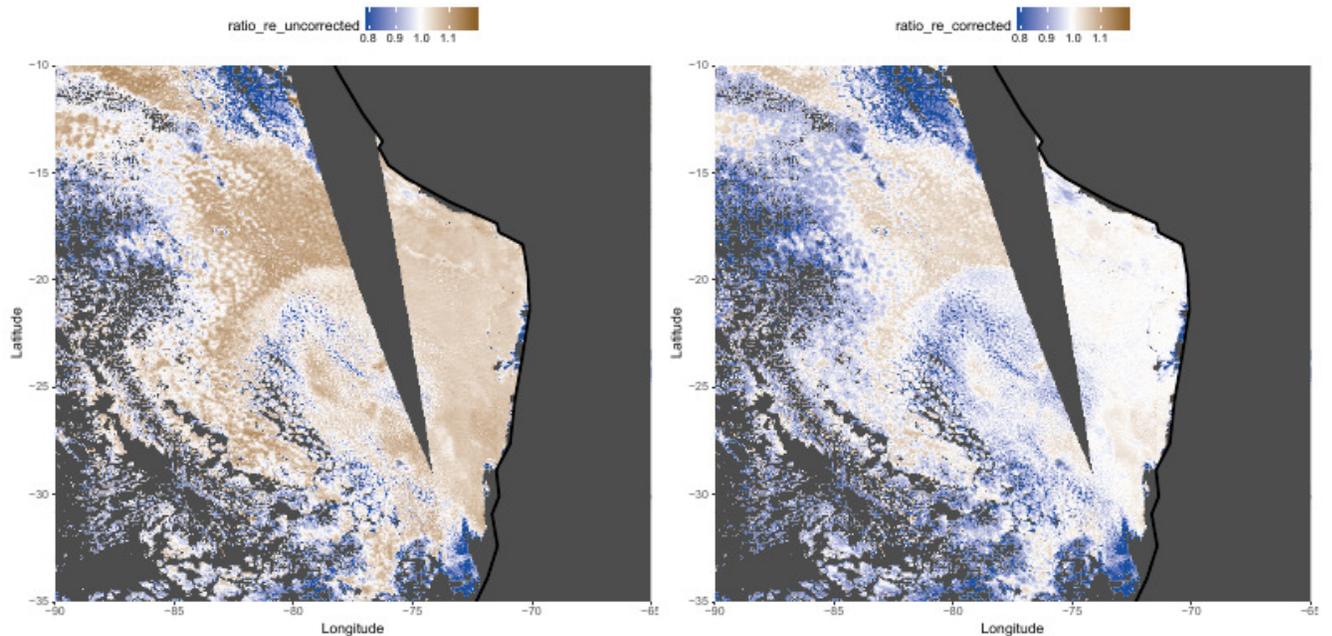
It shows that the ratio is generally larger than one for the standard uncorrected MODIS product values (left) for the overcast stratocumulus regions. The right plot shows the ratios after the effective radius correction has been applied (based on our parameterizations for both 2.1 and 3.7um). If our parameterization is working well and if the re_21 vs re_37 differences are caused by vertical stratification then the ratio should equal one. It can be seen that this is indeed the case for most of the overcast region with just a small amount of positive bias in the overcast clouds to the west and north. The more broken cloud regions should be ignored since Nd retrievals would ideally not be made for such clouds.

Some discussion on these issues has been added to the Discussion section of the revised paper :-

Our paper quantifies the vertical penetration bias in isolation to the other effects mentioned above. It should be questioned, though, whether the presence of cloud heterogeneity and other effects somehow prevent the effects of the vertical stratification from influencing the retrieved r_e, making it irrelevant. This could be a potential explanation for why it is often observed that $r_{e2.1}$ is larger than $r_{e3.7}$ (Zhang and Plantnick, 2011) in contrast to the direction expected from adiabatic clouds given the vertical penetration effect, since it is known that sub-pixel heterogeneity effects tend to cause a positive $r_{e2.1}$ bias relative to $r_{e3.7}$ (Zhang et al., 2012). We argue, though, that the vertical stratification effect occurs in addition to other effects (heterogeneity, etc.) with the latter cancelling out and often exceeding the former such that the positive difference between $r_{e2.1}$ and $r_{e3.7}$ would be even larger without the vertical stratfication effect. The cancellation of biases may also explain why VOCALS aircraft measurements (Painemal and Zuidema, 2011) tended to show that $r_{e2.1}$ and $r_{e3.7}$ were very similar. We also note that there are many situations when the expected result due to vertical stratification of r_e does occur (i.e. $r_{e3.7} > r_{e2.1}$), as demonstrated in Fig. 8. This shows ratios between $r_{e3.7}$ and $r_{e2.1}$ for an example MODIS scene in the SE Pacific stratocumulus region. Ratios using the uncorrected MODIS r_e values are shown, which shows that the ratio is exceeds one for most of the stratocumulus cloud region (the clouds that adjoin the coast) with ratios ranging from around 1.1 to 1.2. In the more broken clouds the ratio is less than one, which is likely a result of cloud heterogeneity. However, it would be expected that N_d retrievals would not be applied to such clouds. The figure also shows the ratios calculated using $r_{e3.7}$ and $r_{e2.1}$ values that have been corrected using the g_{re} factors. If the differences between $r_{e3.7}$ and $r_{e2.1}$ were caused by vertical stratification alone and if our parameterization was correctly predicting the cloud top r_e for both MODIS channles then this ratio should be equal to one. This is the case for the clouds close to the coast indicating that our parameterization is working well for these clouds. The ratios are a little higher than one further north and west (around 1.05-1.08) indicating that either our parameterization is not working correctly for these clouds, or that other factors are causing relative differences between $r_{e3.7}$ and $r_{e2.1}$ (e.g. sub-adiabaticity, cloud heterogeneity, etc.). Figure 9 shows the percentage of pixels for which $r_{e3.7} > r_{e2.1}$ for 90 days of MODIS SE Pacific observations divided into four different heterogeneity bins. Heterogeneity is characterized by the H_σ parameter (Liang et al., 2009), which is the standard deviation of the 250 m resolution 0.86 µm reflectance ($R_{0.86}$) divided by the mean $R_{0.86}$. It is clear that for many regions relative r_e values that are consistent with an adiabatic profile occur more than 50% of the time, particularly when the cloud heterogeneity is low. This suggests that it may be possible to use H_σ to determine the situations in which the bias correction is more applicable. However, it is hard to definitively prove our argument within the scope of this study, particularly for more heterogeneous regions, since it would likely require computationally expensive 3D radiative transfer modelling of known cloud fields (e.g., from LES models), followed by r_e and τ retrievals.

References

Bennartz, R. and Rausch, J.: Global and regional estimates of warm cloud droplet number concentration based on 13 years of AQUA-MODIS observations, Atmospheric Chemistry and Physics, 17, 9815–9836, doi:10.5194/acp-17-9815-2017, https://doi.org/10.5194/acp-17-9815-2017, 2017.

Responses to Reviewer 2

- *Summary:*

  This paper assesses a particular type of error in the satellite-based retrieval of cloud droplet number concentration (Nd) retrieval from passive sensors such as MODIS. The error stems from the fact the shortwave infrared band used in MODIS cloud droplet size retrieval (i.e., cloud droplet effective radius CER) does not correspond to the CER at the exact cloud top, but somewhere below the cloud top due to the penetration of the light into the cloud (termed as the "penetration depth bias" in this study), which leads to underestimation of CER and overestimation of Nd. This study investigates the size of this bias under different conditions and also provide a simple parameterization scheme to correct this bias in the observation. The topic of this paper is suitable for AMT. The paper is well-written, concise and easy to follow for the readers with the right background (but perhaps too technical for general readers). Overall, I recommend publication after some revision.

- *Comments/Suggestions:*

  My biggest concern/criticism for this study and many other studies on Nd retrieval is that most of them are based on highly idealized cloud model, namely, the perfect, 1D, plane-parallel, adiabatic cloud with linear LWC lapse rate and constant Nd. It seems to me that, the meaningfulness of this study depends pretty much on the validity of this ideal cloud model.

  In particular, it is well known that the entrainment process can significantly affect the cloud microphysics at cloud top and thereby deviates the cloud vertical structure from the classic model assumed in Nd retrieval. How may the cloud top entrainment process influence those equations in section 2? What is the typical vertical scale of cloud top entrainment in comparison with the penetration depth of the SWIR band? Do homogenous mixing and inhomogeneous mixing as a result of cloud top entrainment have a different or similar impact on cloud top CER structure and Nd retrieval?At least, these questions should be mentioned, discussed with some references.

We have now included some discussion on entrainment effects in the Discussion section. The evidence suggests that for stratocumulus clouds cloud top entrainment results in extreme inhomogeneous mixing, so that the CER remains constant - this is also backed up by the VOCALS aircraft measurements. The other effects due to the non-constant Nd profile and non-adiabatic liquid water content profiles are likely to be small since we estimate from aircraft observations that the entrainment region only contributes around 0.5 optical depths to the total optical depth. The added text is as follows :-

Evaporation effects related to entrainment also have the potential to reduce r_e, N_d and LWC near cloud top and hence negate some of the assumptions upon which the N_d retrievals rest. However, we argue that the entrainment effect upon r_e is likely to be minimal for two reasons: Firstly, the evidence suggests that for stratocumulus clouds extreme inhomogeneous mixing occurs at cloud top, which reduces the LWC and N_d, but does not change r_e (Burnet and Brenguier, 2007; Brenguier et al., 2011; Painemal and Zuidema, 2011). Secondly, the results of Painemal and Zuidema (2011) indicate that entrainment occurs within approximately the first 0.5 optical depths from cloud top on average; the penetration depths calculated here are considerably

<center>16</center>

larger than this for reasonably thick clouds (Fig. 2). The effect of the reduced N_d and LWC within the entrainment zone is not so clear-cut; this would negate the assumption of a vertically constant N_d and monotonically increasing LWC used to formulate the total τ. However, given the likely small τ contribution from the entrainment region relative to the total τ, this effect is likely to be small.

What is the COT (τ) used in the Nd retrieval? Note that in MODIS operational retrieval, clouds are assumed to be vertically homogeneous. Because of the "penetration depth bias", the retrieve CER is different from the CER at the cloud top. Another possible bias is that the retrieved COT is different from the true COT. This might be small but should be quantified.

The COT used in our Nd retrieval is that directly from the MODIS products and so may contain biases due to the non-uniform CER profile that is likely to occur in reality combined with the fact that MODIS assumes vertically uniform clouds. We have now included a figure that quantifies the percentage bias in the retrieved optical depth (relative to the model profile optical depth) :-
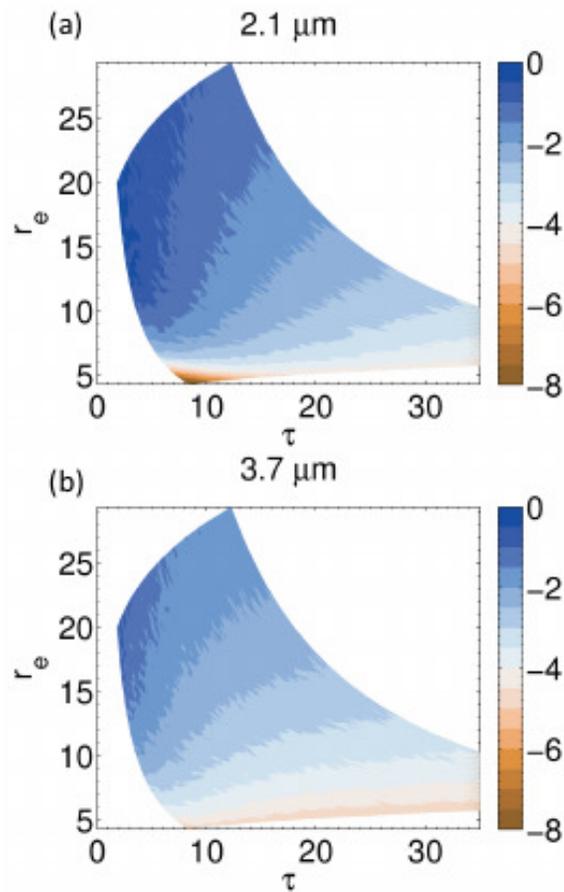
(a) 2.1 μm

(b) 3.7 μm

Figure 8. Percentage τ bias (retrieved minus actual value from the input model profile) as a function of τ and r_e.

The following text has also been added to the discussion :-

The MODIS retrieval uses reflectances from both a visible and a shortwave infra-red (SWIR) wavelength channel with the former being primarily determined by τ and the latter by r_e. However, a bi-spectral retrieval is used and so there is also some sensitivity of the retrieved τ to the SWIR reflectance, which will be representative of the r_e below cloud top due to the vertical penetration effect. This, combined with the fact that the MODIS forward retrieval model assumes a vertically uniform cloud, will result in the retrieved τ being biased relative to the real value (assuming the real cloud has an adiabatic profile). Figure 9 shows the difference between the retrieved and model profile τ; the bias is negative and smaller in magnitude than 5 % for the 3.7 μm retrieval. They are slightly larger for the 2.1 μm retrieval, but still lower in magnitude than 5 %, except at $r_e \lesssim 7$ μm. Although it should be noted that some of this bias may be due to other causes related to the inconsistencies between the vertically uniform and adiabatic models, rather than the r_e vertical penetration bias. Since the retrieved N_d is proportional to the square root of τ, this will lead to small N_d biases. Biases in LWP will be similar to those in τ since LWP is proportional to τ, but r_e biases are still likely to dominate (e.g., see Fig. 1). Thus, we have not pursued this further.

In this study, only the solar reflective part of the 3.7 μm band is considered. In reality, the radiance in this band is contributed by two parts during the daytime, the solar reflection and thermal emission. The emission part is "corrected" based on the 11 μm band radiance in the MODIS retrieval. This should be pointed out and if the correction process could somehow confound the results then some discussion is needed. This is especially important as the paper claims that 3.7 μm band is better for Nd retrieval (which I agree) than the 2.1 μm band.

We have added some discussion on this in the methods and now describe how this is dealt with in our retrievals :-

and 3.7μm wavelengths, which are hereafter referred to to as $r_{e2.1}$ and $r_{e3.7}$, respectively. The MODIS $r_{e3.7}$ retrieval requires a correction to account for the contribution to the observed radiance from thermal emission, which is based on the observed 11 μm radiance (Platnick and Valero, 1995; King et al., 2015; Platnick et al., 2017). We account for this in our retrievals by removing the thermal contribution during the RT calculation instead of via the 11 μm radiance, which should produce a consistent end result. The RT calculations were performed assuming a black surface, a clear atmosphere (i.e. gaseous absorption is neglected),

And we also add some discussion in the Discussion section :-

Finally, we note that the thermal emission correction for the MODIS $r_{e3.7}$ (see Section 3.1) retrieval has some uncertainty that should be considered; the uncertainty for this is included (combined with other uncertainties) in the MODIS Collection 6 pixel level uncertainty products (Platnick et al., 2017). It is possible that effects additional to those included, such as cloud heterogeneity, surface heterogeneity, etc., may further increase the uncertainty beyond that estimated in the MODIS products, but these are currently not well documented.

For 3.7 μm band, its weighting function is close to two-way transmittance. I'd like to encourage the author to try to come up with an analytical solution of CER if the weighting function follows the two-way transmittance.*
A paper that might be helpful Zhang et al. 2017 JGR (http://onlinelibrary.wiley.com/doi/10.1002/2016JD025763/full) (Equation 4)

We have now been able to parameterize the correction to CER in order to return the cloud top CER as a function of the retrieved optical depth and retrieved CER. This was also suggested by Reviewer 1 - please see the response given there for details.

Why is land always masked in Nd retrievals? Why or why not can the same method be applied to land?

We restricted the analysis to ocean retrievals since the Nd and LWP retrievals are most suited to stratocumulus and these types of clouds occur much more frequently over the oceans. Land retrievals also add additional complications from surface albedo uncertainties and cloud masking problems, which we also tried to avoid. We have added some discussion on these points :-

In this paper we have only considered retrievals over the ocean, although retrievals over land for τ and r_e are available for MODIS. MODIS surface albedo uncertainties are likely to be much higher over land than over the oceans (King et al., 2004; Rosenfeld et al., 2004; Bréon and Doutriaux-Boucher, 2005) since the surface albedo is much more variable over land. In addition, cloud masking is more difficult over land, particularly over non–vegetated surfaces, transitional areas between desert and vegetated surfaces and above high-altitude regions (Platnick et al., 2003). We have ignored land regions in order to avoid such complications and also because stratocumulus clouds are more prevalent over ocean regions (Klein and Hartmann, 1993; Wood, 2012). However, the results shown in this paper may still apply over land. The results of Rosenfeld et al. (2004) and Platnick et al. (2017, their Fig. 14) suggest that surface albedo uncertainties are more important at lower optical depths ($\lesssim 5$) and for the 2.1 µm retrieval (relative to the 3.7 µm one). Thus, for thicker clouds and the 3.7 µm retrieval land surface albedo issues may be less problematic.

List of changes

- A parameterization of the corrected effective radius (reff) as a function of the retrieved reff and optical depth has been added.
    - A new figure for this has been added, along with a new equation for the parameterized fit and a new table listing the fit coefficients. Also, a new equation showing the ratio of uncorrected to corrected droplet concentrations as a function of the correction parameter.
    - Additional text has been added to describe this in the Abstract, Section 2, Section 3.1 and the conclusions.
- The new parameterization is now used in place of the old one, which makes small differences to the results. Figure 4 (of new manuscript) is also updated using the new parameterization.
- The new parameterization allows LWP biases to also be estimated and so new text describing this has been added to the Abstract, Section 2, Section 4 (Results) and Conclusions
- The previous Table 2 has been split into two in order to separate the regional optical depth statistics from the biases. The LWP biases are now also listed.
- Following these changes the title has been changed to :-
    - "Parameterizing cloud top effective radii from satellite retrieved values, accounting for vertical photon transport: Quantification and correction of the resulting bias in droplet concentration and liquid water path retrievals."
- Information on the thermal emission correction for 3.7 micron MODIS reff retrievals has been added to Section 3.1.
- The data from Figure 3 now includes the tau>5 filtering for consistency with the rest of the paper.
- Several additional discussion points have been added to the Discussion section following the reviewer comments (see the responses for a point-by-point response and more details on the additions), including paragraphs on :-
    - The effects of entrainment on reff retrievals and our results.
    - Whether other effects (such as heterogeneity) prevent the vertical stratification from having an impact. Two new figures have been added to aid with this (Figs. 8 and 9) and are discussed. This leads to the recommendation to restrict the use of the bias correction to more homogeneous clouds in the conclusions.
    - How much optical depth retrievals are likely to be biased due to vertical penetration effects (including a new figure – Fig. 10 - to quantify this).
    - Additional biases that are likely to occur over land as justification for not including land retrievals.
    - The thermal correction for the 3.7 micron reff retrieval.
- A few grammatical changes have also been made.

~~Quantifying~~ Parameterizing cloud top effective radii from satellite retrieved values, accounting for vertical photon transport: Quantification and ~~correcting the effect~~ correction of ~~vertical penetration assumptions on~~ the resulting bias in droplet concentration and liquid water path retrievals~~from passive satellite instruments~~.

Daniel P. Grosvenor[1,2], Odran Sourdeval[3], and Robert Wood[4]

[1]School of Earth and Environment, University of Leeds, Leeds, LS2 9JT, UK
[2]National Centre for Atmospheric Science (NCAS), University of Leeds, Leeds, LS2 9JT, UK
[3]Leipzig Institute for Meteorology, Universität Leipzig, Germany
[4]Department of Atmospheric Sciences, University of Washington, Seattle, USA

Correspondence to: D. P. Grosvenor
(daniel.p.grosvenor@gmail.com)

Abstract. Droplet concentration (N_d) and liquid water path (LWP) retrievals from passive satellite retrievals of cloud optical depth (τ) and effective radius (r_e) usually assume the model of an idealised cloud in which the liquid water content (LWC) increases linearly between cloud base and cloud top (i.e., at a fixed fraction of the adiabatic LWC)~~with a constant N_d profile~~. Generally it is assumed that the retrieved r_e value is that at the top of the cloud. In reality, barring r_e retrieval biases due to cloud heterogeneity, etc., the retrieved r_e is representative of ~~that~~ smaller values that occur lower down in the cloud due to the vertical penetration of photons at the shortwave infra-red wavelengths used to retrieve r_e. This inconsistency will cause an overestimate of N_d and an underestimate of LWP (referred to here as the "penetration depth bias"), which this paper quantifies via a parameterization of the cloud top r_e as a function of the retrieved r_e and τ. Here we estimate ~~penetration depths in terms of optical depth below cloud top ($d\tau$)~~ the relative r_e underestimate for a range of idealised modelled adiabatic clouds using bispectral retrievals and plane-parallel radiative transfer. We find a tight relationship between ~~$d\tau$~~ $g_{re} = r_e^{cloudtop}/r_e^{retrieved}$ and τ and that a 1-D relationship approximates the modelled data well. Using this relationship we find that ~~$d\tau$~~ g_{re} values and hence N_d and LWP biases are higher for the 2.1 μm channel r_e retrieval ($r_{e2.1}$) compared to the 3.7 μm one ($r_{e3.7}$). The theoretical bias in the retrieved N_d is ~~likely to be~~ very large for optically thin clouds, ~~nominally approaching infinity for clouds whose τ is close to the penetration depth. The relative N_d bias~~ but rapidly reduces as cloud thickness increases~~, although still~~. However, it remains above 20 % for $\tau < 19.8$ and $\tau < 7.7$ for $r_{e2.1}$ and $r_{e3.7}$, respectively. We also provide a parameterization of penetration depth in terms of the optical depth below cloud top ($d\tau$) for which the retrieved r_e is likely to be representative.

The magnitude of the N_d ~~bias upon climatological N_d~~ and LWP biases for climatological data sets is estimated globally using one year of daily MODIS (MODerate Imaging Spectroradiometer) data. Screening criteria are applied that are consistent with those required to help ensure accurate N_d and LWP retrievals. The results show that the SE Atlantic, SE Pacific (where

the VOCALS field campaign took place) and Californian stratocumulus regions produce fairly large overestimates due to the penetration depth bias with mean biases of ~~35–38~~32–35 % for $r_{e2.1}$ and ~~17–20~~15–17 % for $r_{e3.7}$. For the other stratocumulus regions examined the errors are smaller (~~25–30~~24–28 % for $r_{e2.1}$ and ~~11–14~~10–12 % for $r_{e3.7}$). Significant time variability in the percentage errors is also found with regional mean standard deviations of ~~20–40~~19–37 % of the regional mean percentage error for $r_{e2.1}$ and ~~40–60~~32–56 % for $r_{e3.7}$. This shows that it is important to apply a daily correction to N_d for the penetration depth error rather than a time–mean correction when examining daily data. We also examine the seasonal variation of the bias and find that the biases in the SE Atlantic, SE Pacific and Californian stratocumulus regions exhibit the most seasonality with the largest errors occurring in the December, January, February (DJF) season. LWP biases are smaller in magnitude than those for N_d (-8 to -11 % for $r_{e2.1}$ and -3.6 to -6.1 % for $r_{e3.7}$).

~~We show that this effect can be corrected for by simply removing $d\tau$ from the observed τ and provide a function to allow the calculation of $d\tau$ from τ. However, in reality~~ this In reality, and especially for more heterogeneous clouds, the vertical penetration error will be combined with a number of other errors that affect both the r_e and τ, which are potentially larger and may compensate or enhance the bias due to vertical penetration depth. Therefore caution is required when applying the bias corrections; we suggest that they are only used for more homogeneous clouds.

1 Introduction

Clouds have a major impact on Earth's radiative balance (Hartmann et al., 1992) and small changes in their properties are predicted to have large radiative impacts (e.g., Latham et al., 2008). The amount of shortwave (SW) flux reflected by fully overcast warm (liquid water) clouds for a given sun and scattering angle, or the reflectance of a cloud, is primarily determined by the cloud optical depth (τ), which in turn can often be characterized by the liquid water path (LWP, the vertical integral of liquid water content) and the cloud droplet number concentration (N_d). For a given cloud updraft, N_d is determined by the number concentration and physico-chemical properties of aerosols. Thus, couching cloud reflectance in terms of N_d links the cloud albedo to aerosol and microphysical effects via the Twomey (1974) effect making N_d a very useful quantity to determine observationally. N_d can also influence cloud macrophysical feedbacks via its control on rain formation (Albrecht, 1989; Stevens et al., 1998; Ackerman et al., 2004; Berner et al., 2013; Feingold et al., 2015) and stratocumulus cloud top entrainment (Ackerman et al., 2004; Bretherton et al., 2007; Hill et al., 2009).

Satellite observations of clouds and N_d are immensely useful for studying clouds, cloud–aerosol interactions and for model evaluation since they afford large ~~spatial-temporal~~ spatial and temporal coverage. A method to obtain N_d from passive satellite observations ~~(e.g., from MODIS (MODerate Imaging Spectroradiometer); Salomonson et al. (1998))~~ (e.g., from the MODerate Imaging Spectroradiometer, MODIS; Salomonson et al., 1998) of τ and the cloud droplet effective radius (r_e) for stratiform liquid clouds has been previously demonstrated (Han et al., 1998; Brenguier et al., 2000; Nakajima et al., 2001; Szczodrak et al., 2001; Boers et al., 2006; Quaas et al., 2006; Bennartz, 2007; Grosvenor and Wood, 2014; Bennartz and Rausch, 2017) and is described further below. For more details see the Grosvenor et al. (2018) review paper on this technique, which also describes the known sources of error. In cloudy environments, aerosol optical ~~depths~~ depth cannot be

retrieved from satellite making cloud property observations such as N_d and the cloud droplet effective radius (r_e) the only useful indicator of the influence of aerosol on clouds. An advantage to using N_d rather than r_e to study cloud–aerosol interactions is that r_e is also determined by the cloud water content and thus is a function of cloud macrophysical properties. N_d on the other hand is only weakly controlled by cloud macrophysics allowing some separation of microphysical and macrophysical effects.

5   However, retrievals of N_d from space are still somewhat experimental and there is a lack of comprehensive validation of the retrievals and the assumptions required. There is a need to characterize and quantify the associated errors; in this paper we focus on doing this for one source of N_d error using a one–year N_d data set for stratocumulus clouds from MODIS.

2 The ~~vertical penetration depth N_d bias and the~~ adiabatic N_d and LWP retrieval model, and the vertical penetration depth bias

10 N_d ~~is~~ and LWP are retrieved from passive satellite retrievals of r_e and τ using an adiabatic cloud model that is described below. However, as shown in ~~Platnick (2000); Bennartz and Rausch (2017)~~ Platnick (2000) and Bennartz and Rausch (2017), for a retrieval free from other error sources (e.g. those due to cloud heterogeneity), the retrieved r_e is representative of the r_e value lower down in the cloud due to the vertical penetration of photons at the shortwave infra-red wavelengths used to retrieve r_e. In contrast, the retrieved τ is comprised of contributions from the extinction coefficient $\beta_{ext}(h)$, where h represents height

15 from cloud base, throughout the whole cloud profile :-

$$\tau = \int\limits_0^H \beta_{\mathrm{ext}}(h)dh. \tag{1}$$

Here h=0 represents cloud base and h=H is cloud top.

$\beta_{ext}(h)$ is defined as :-

$$\beta_{\mathrm{ext}}(h) = \pi \int\limits_0^\infty Q_{\mathrm{ext}}(r)r^2 n(r)dr, \tag{2}$$

20 where r is the droplet radius and $n(r)$ is the droplet size number distribution within a cloud unit volume such that $N_d = \int_0^\infty n(r)dr$. $Q_{\mathrm{ext}}(r)$ represents the ratio between the extinction and the geometric cross ~~sections~~ section of a given droplet and can be approximated by its asymptotic value of 2 (van de Hulst, 1957) since droplet radii are generally much larger than the wavelength of light concerned (typically 0.6 to 0.85 μm) such that the geometric optics limit is almost reached.

r_e and liquid water content LWC at a given height are respectively defined as:

25 $$r_e(h) = \frac{\int_0^\infty r^3 n(r)dr}{\int_0^\infty r^2 n(r)dr} \tag{3}$$

and

$$LWC(h) = \frac{4\pi \rho_w}{3} \int\limits_0^\infty r^3 n(r) dr, \tag{4}$$

where ρ_w is the density of liquid water. Combining Eqns. 3 and 4 and inserting into Eqn. 2 gives :-

$$\beta_{\text{ext}}(h) = \frac{3Q_{\text{ext}}}{4\rho_w} \frac{LWC(h)}{r_e(h)} \tag{5}$$

5   To determine the form of $r_e(h)$ in the above equation in terms of $L(h)$ and $N_d(h)$ we can utilize the fact that the "k" value,

$$k = \left(\frac{r_v}{r_e}\right)^3, \tag{6}$$

which is a measure of the width of the droplet size distribution (lower values indicate wider distributions), has been shown to be approximately constant in stratocumulus clouds (Martin et al., 1994; Pawlowska et al., 2006; Painemal and Zuidema, 2011). In this study we adopt a value of ~~$k = 0.8$~~$k = 0.72$, which is the ~~most commonly used value in previous studies~~
10 ~~that perform N_d retrievals (e.g., Bennartz, 2007; Painemal and Zuidema, 2011) in stratocumulus~~value assumed by the MODIS retrieval (Zhang, 2013). r_v is the volume radius, defined as :-

$$r_v(h)^3 = \frac{1}{N_d(h)} \int\limits_0^\infty r^3 n(r) dr = \frac{3LWC(h)}{4\pi \rho_w N_d(h)} = kr_e(h)^3, \tag{7}$$

where we have used Eqn. 4 to insert LWC and Eqn. 6 to write r_v as a function of k and r_e. Now we utilize the assumptions that $N_d(h)$ is constant with height and that LWC(h) is a constant fraction, f_{ad}, of adiabatic. The latter equates to :-

15 $$LWC(h) = f_{ad} c_w h, \tag{8}$$

where c_w is the rate of increase of LWC with height ($dLWC/dz$, with units ~~kgm^{-4}~~kgm^{-4}) for a moist adiabatic ascent and is referred to as the "condensation rate" in Brenguier et al. (2000), or the "water content lapse rate" in Painemal and Zuidema (2011). See Ahmad et al. (2013) for a derivation. c_w is a constant for a given temperature and pressure. Allowing these assumptions, using Eqn. 7 to substitute for r_e in Eqn. 5 and combining with Eqns. 1 and 8 we can write :-

$$\tau^* = \int\limits_0^{H^*} Q_{\text{ext}} \left(\frac{3f_{ad}c_w}{4\rho_w}\right)^{2/3} (N_d \pi k)^{1/3} h^{2/3} dh$$

20 $$= \frac{3Q_{\text{ext}}}{5} \left(\frac{3f_{ad}c_w}{4\rho_w}\right)^{2/3} (N_d \pi k)^{1/3} H^{*5/3} \tag{9}$$

4

At this stage, H^* is any arbitrary height above cloud base and τ^* is thus the optical depth between the cloud base and that height. H^* can be expressed as a function of $r_e(H^*)$, k, N_d and some constants by using Eqns. 7 and 8. Then, given $r_e(H^*)$ and τ^*, N_d can be calculated as follows :-

$$N_d = \frac{\sqrt{5}}{2\pi k}\left(\frac{f_{ad}c_w\tau^*}{Q_{ext}\rho_w r_e(H^*)^5}\right)^{1/2} \tag{10}$$

5   Generally, when retrieving N_d it is assumed that the r_e obtained from satellite is representative of that from cloud top, i.e., $r_e(H^*)=r_e(H)$ (e.g. Bennartz, 2007; Painemal and Zuidema, 2011). This would then mean that τ^* is the full cloud optical depth (τ) as retrieved by the satellite and thus could be used in Eqn. 10 above to obtain N_d. However, since the r_e obtained by satellite is actually equal to $r_e(H^*)$ then $\tau^* < \tau$ and thus τ^* should be used in Eqn. 10 instead of the retrieved τ; the problem lies in the fact that τ^* is unknown. However, in this paper we fit a simple function for τ^* as a function of τ based on radiative

10   transfer modelling of a variety of idealised clouds.

Alternatively, Eqn. 10 can be formulated using the retrieved τ over the full cloud depth (setting $\tau^* = \tau$) and the cloud top r_e (setting $r_e(H^*) = r_e(H)$). The problem then becomes one of estimating $r_e(H)$ from the retrieved $r_e(H^*)$. Here we formulate a parameterization of $r_e(H)$ / $r_e(H^*)$ as a function of τ. Note, that either the τ or r_e corrections should be applied to correct N_d, but not both together.

15   Then we estimate the error introduced in N_d retrievals for one year of MODIS data due to the usual assumptions of $r_e(H^*)=r_e(H)$ and $\tau^*=\tau$, on the assumption that there are no other biases affecting the r_e retrieval. We label this bias the "vertical penetration bias".

The method of correcting r_e has the advantage over the τ correction since it also allows a correction to the retrieval of LWP. LWP can be estimated (see e.g., Szczodrak et al., 2001) using :

$$LWP = \frac{5}{9}\rho_w r_e(H)\tau. \tag{11}$$

For a corrected LWP the cloud top r_e and the retrieved (total) τ values should be used. Since the retrieved $r_e(H^*)$ is likely to be underestimated due to the vertical penetration depth bias, LWP would otherwise be underestimated and the correct value can be obtained by using the parameterized $r_e(H)$ instead.

3   Data and Methods

3.1   Calculation of ~~$d\tau$~~ τ and r_e corrections

In order to calculate

$$g_{re} = \frac{r_e(H)}{r_e(H^*)} \tag{12}$$

and

$$dτ = τ - τ^* \tag{13}$$

we have performed r_e retrievals on idealised clouds using a similar algorithm to that used for MODIS retrievals. We produced idealised clouds that span a large range of stratocumlus–like clouds as represented by combinations of N_d and LWP.

We chose 41 values between N_d=10 and 1000 cm^{-3} that were equally spaced in log space and 91 values between LWP=20 and 200 gm^{-2} spaced equally in linear space. All of the possible combinations from this sampling were used to sample the 2-D (N_d, LWP) phase space. For each combination, discretized adiabatic model profiles following the form of those described in Section 1 (i.e., with a vertically constant N_d and LWC that increases linearly with height) were generated using c_w=1.81 ~~$\times 10^{-6}\ kg\ m^{-4}$~~ $\times 10^{-6}\ \text{kg m}^{-4}$, f_{ad}=0.8 and using a vertical spacing of 1 m. The droplet size distributions at each height were represented by a modified gamma distribution with a k value of 0.72, i.e. representative of an effective variance of 0.1. 1-D radiative transfer (RT) calculations, assuming plane-parallel clouds, were performed on these profiles using the DISORT (Discrete Ordinates Radiative Transfer Program; Stamnes et al., 1988) radiation code in order to simulate reflectances at wavelengths of 0.86, 2.1 and 3.7 μm, matching those measured by MODIS to retrieve $τ$ and r_e ~~. The RT~~ over an ocean surface. Note that MODIS provides r_e retrievals using both 2.1 μm and 3.7 μm wavelengths, which are hereafter referred to to as $r_{e2.1}$ and $r_{e3.7}$, respectively. The MODIS $r_{e3.7}$ retrieval requires a correction to account for the contribution to the observed radiance from thermal emission, which is based on the observed 11 μm radiance (Platnick and Valero, 1995; King et al., 2015; Platnick et al., 2017). We account for this in our retrievals by removing the thermal contribution during the RT calculation instead of via the 11 μm radiance, which should produce a consistent end result. The RT calculations were performed assuming a black surface, a clear atmosphere (i.e. gaseous absorption is neglected), using a solar zenith angle of 20$^\text{o}$ and a nadir viewing angle.

These reflectances were then used to retrieve $τ$ and r_e values using the Nakajima and King (1990) bi-spectral method, as operationally used by MODIS. To do so, a lookup table was built from reflectances ~~similarly~~ calculated for a range of clouds that were assumed to be plane-parallel in nature, as assumed for the operational MODIS retrievals; i.e. these clouds were uniform in the vertical and horizontal with infinite horizontal extent. Again, a black surface and a k value of 0.72 were assumed along with the same viewing geometry as for the RT calculations on the adiabatic clouds. A fixed depth of 1 km was assumed with cloud base at an altitude of 1 km and cloud top at 2 km, although the cloud depth has no major effect on the reflectances generated ~~.~~ for a given $τ$ and r_e. g_{re} was then calculated using the retrieved and model top r_e values. $dτ$ was ~~then~~ calculated by choosing the value ~~of~~ from the model profile of $τ$, as measured from cloud top downwards, that corresponded to the ~~value of~~ location where the model profile ~~of~~ r_e ~~that~~ matched the retrieved r_e.

Figure ~~2~~1a shows a 2-D histogram of ~~$dτ$~~ g_{re} values as a function of $τ$ for the 2.1 μm retrieval. It shows that when plotted in this way ~~$dτ$~~ g_{re} forms a fairly tight relationship with $τ$ so that for a given $τ$ only a small range of ~~$dτ$~~ g_{re} values are possible. This suggests that the relationship can be parameterized based upon a 1-D relationship fitted to this data with little loss of accuracy. The ~~mean~~ median value of each $τ$ bin is also plotted (after smoothing over $τ$ windows of 0.2) and this is the relationship used in this paper. ~~$dτ$~~ g_{re} is seen to ~~increase~~ decrease with $τ$ with a gradient that decreases with $τ$. ~~A~~ Similarly, Figure 2 shows $dτ$

6

vs τ, which also shows a tight relationship that is suited to a 1-D parameterization. 4th order polynomial ~~curve~~ curves can be fitted (using the least squares method) to ~~this mean value relationship that takes~~ the median value relationships that take the form:-

$$d\tau g_{re} = a_4\tau^4 + a_3\tau^3 + a_2\tau^2 + a_1\tau + a_0 \tag{14}$$

and

$$d\tau = b_4\tau^4 + b_3\tau^3 + b_2\tau^2 + b_1\tau + b_0 \tag{15}$$

The coefficients of ~~this fit and for the 3.7 μm retrieval~~ these fits are given in ~~Table~~ Tables 1 and 2 along with the maximum ~~absolute τ error~~ errors for the fit (relative to the mean or median line) for the range shown. The ~~curve (white line in Fig.~~ curves (white lines in Figs. 1a and 2a) ~~fits~~ fit the mean data well with ~~a maximum absolute difference of~~ maximum absolute differences of 0.001 and 0.09, ~~although~~ respectively, for the g_{re} and $d\tau$ curves. However, there will be some error when using this relationship (or the mean value relationship) due to the spread in the g_{re} and $d\tau$ values seen in the underlying ~~histogram~~ histograms.

Table 1. Coefficients for the fitting curve (Eqn. ~~15~~14) to estimate the ~~mean $d\tau$~~ median g_{re} value as a function of τ. The maximum absolute error ~~from~~ between the fit and the median line is also shown.

Retrieval wavelength	a_4	a_3	a_2	a_1	a_0	Max abs g_{re} error
2.1 μm	2.413e-07	-2.467e-05	9.883e-04	-0.02049	1.244	0.001
3.7 μm	5.367e-07	-5.179e-05	0.00186	-0.03038	1.217	0.003

Table 2. Coefficients for the fitting curve (Eqn. 15) to estimate the mean $d\tau$ value as a function of τ. The maximum absolute error between the fit and the mean line is also shown.

Retrieval wavelength	b_4	b_3	b_2	b_1	b_0	Max abs τ error
2.1 μm	-3.174e-06	3.931e-04	-0.021	0.5754	0.3216	0.09
3.7 μm	-1.281e-05	1.099e-03	-0.03304	0.4168	0.6005	0.14

~~Figure 2b shows~~ Figures 1b and 2b show the same results for the 3.7 μm retrieval. Again ~~a~~ tight 1-D ~~relationship is suggested~~~~with a $d\tau$ value that increases with τ~~relationships are suggested. Here, though, ~~the curve flattens~~ g_{re} and $d\tau$ values are lower for a given τ and the curves are steeper at lower τ values, but flatten off much more rapidly~~, so that by~~. By τ=7.5 there is little dependence of $d\tau$ on τ and $d\tau$ saturates at a mean value of ~2.6. The fit ~~estimate for the mean curve (Eqn. 15 and Table~~estimates for the curves (Eqns. 14 and 15 and Tables 1 and 2) again ~~matches the actual curve~~match the actual curves closely with a maximum absolute error in g_{re} and $d\tau$ of 0.003 and 0.14~~.~~, respectively.

3.2 MODIS data

For the MODIS data we use one year (2008) of MODIS Aqua data and follow a similar methodology to that used in Grosvenor and Wood (2014) in order to create a data set akin to the MODIS Level-3 (L3) product (King et al., 1997; Oreopoulos, 2005). We processed MODIS collection 5.1 ~~joint-L2~~joint Level-2 (L2) swaths into $1^o \times 1^o$ grid boxes. Joint-L2 swaths are sub-sampled versions of the full L2 swaths (sampling every 5th 1 km pixel) that also contains fewer parameters. We process the data from L2 to L3 in order to allow the filtering out of data at high solar zenith angles and to provide ~~r_e retrievals from both the 2.1μm and 3.7μm MODIS channels, hereafter referred to to as~~ both $r_{e2.1}$ and $r_{e3.7}$ ~~, respectively~~retrievals.

For this work we relax the screening methodology slightly from that used in Grosvenor and Wood (2014) since here we are interested in the effects of the vertical penetration N_d bias upon a more general global data set. We applied the following restrictions to each $1^o \times 1^o$ sample that goes into the daily average (since multiple overpasses per day are possible) in order to attempt to remove some artifacts that may cause biases:

1. At least 50 joint-L2 1 km resolution pixels from the MODIS swath that did not suffer from sunglint were required to have been sampled within each gridbox.

2. At least 80 % of the available (non-sunglint) pixels were required to be of liquid phase based upon the "primary cloud retrieval phase flag". Analysis was only performed on these pixels. A high cloud fraction helps to ensure that the clouds are not broken, since broken clouds are known to cause biases in retrieved optical properties due to photon scattering through the sides of clouds. Often retrievals of N_d are restricted to high cloud fraction fields for this reason ~~(B07; PZ11)~~ (Bennartz, 2007; Painemal and Zuidema, 2011) and so we focus on such datapoints here.

3. Only the pixels remaining after (2) for which the "cloud mask status" indicated that the cloud mask could be determined, the "cloud mask cloudiness flag" was set to "confident cloudy", successful simultaneous retrievals of both τ and r_e for the 2.1μm channel were performed and the cloud water path confidence from the MODIS L2 quality flags was designated as "very good confidence" (the highest level possible) were used. This is a little different from the official MODIS L3 product where a set of cloud products are provided that are weighted using the quality assurance flags. Rather than weighting our L3-like product with the QA flags we have simply restricted our analysis to pixels with the highest confidence for water path.
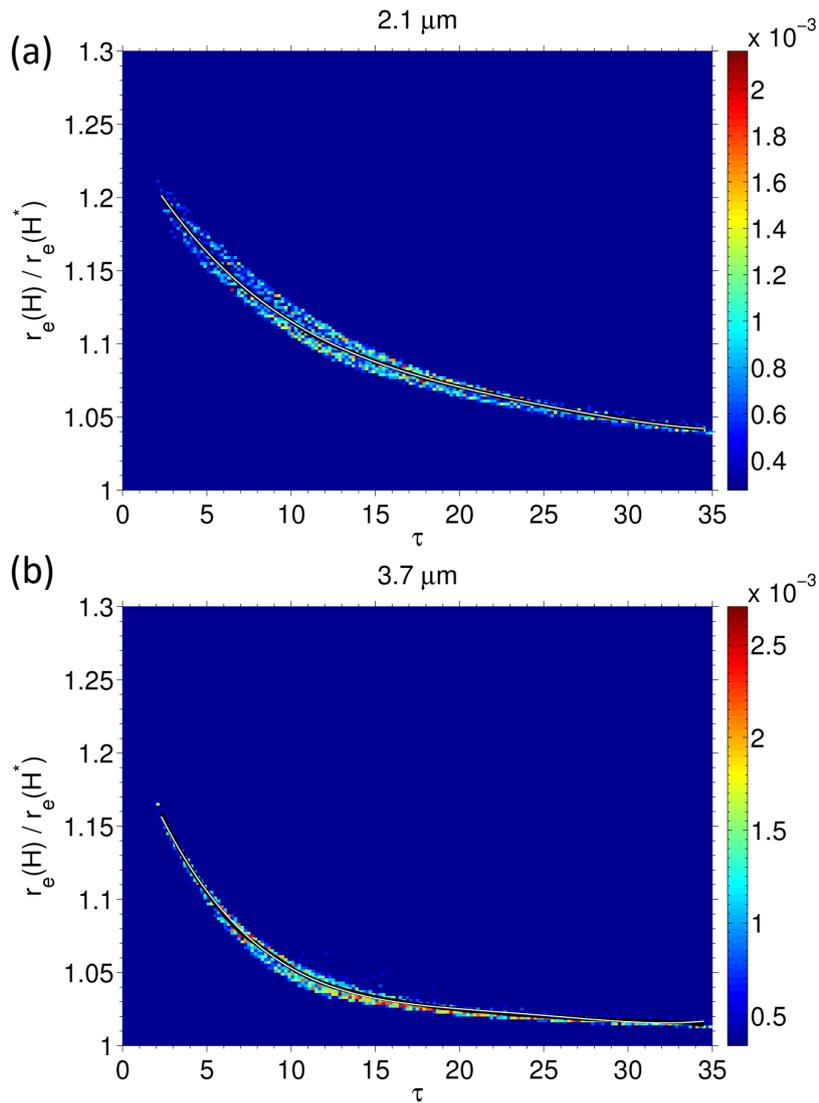
Figure 1. 2D histogram of g_{re} as a function of τ for a range of clouds (see text) for the 2.1 μm r_e retrieval (a) and the 3.7 μm retrieval (b). The black line is the median g_{re} in each τ bin after smoothing over τ interval windows of 0.2. The white line is the fit to the mean curve using Eqn. 14.

4. The mean $1^\circ \times 1^\circ$ cloud top height (CTH) is restricted to values lower than 3.2 km. This is done both to avoid deeper clouds for which N_d retrievals are likely to be problematic due to the increased likelihood of a breakdown of the assumptions required to estimate N_d, such as a constant fraction of the adiabatic value for LWC and vertically constant N_d,
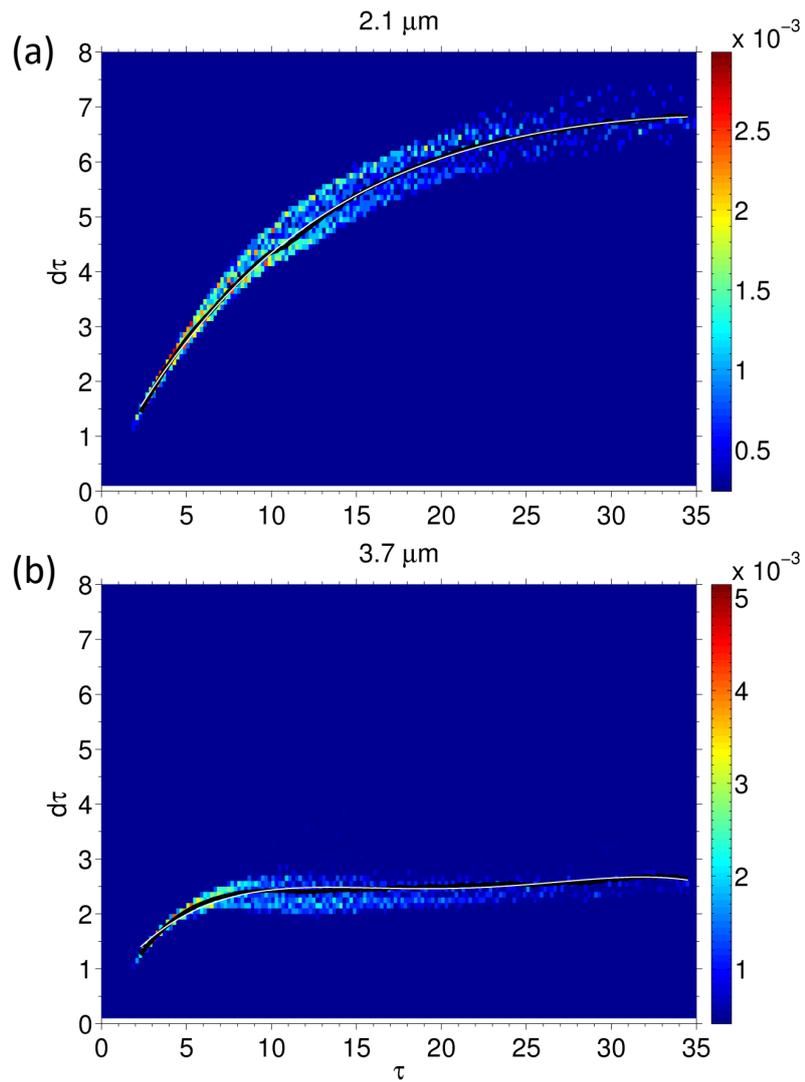
Figure 2. ~~2D histogram of~~ As for Fig. 1 except for $d\tau$ as a function of τ ~~for a range of clouds (see text) for the 2.1 µm r_e retrieval (a)~~ and ~~the 3.7 µm retrieval (b). The black line is the mean $d\tau$ in each τ bin after smoothing over τ interval windows of 0.2. The white line is the fit to the mean curve~~ using ~~using~~ Eqn. 15 for the white line.

as well increased retrieval issues due to cloud heterogeneity. CTH is calculated from the MODIS $1^\circ \times 1^\circ$ mean cloud top temperature (CTT) and the sea surface temperature (SST) using the method of Zuidema et al. (2009). SST data was obtained from the v2 of the NOAA Optimum Interpolation (OI) Sea Surface Temperature data set (NOAA_OI_SST_V2)

that provides weekly SST data at $1^o \times 1^o$ resolution. This was interpolated to daily data on the assumption that SST does not vary significantly over sub-weekly timescales.

5. The mean $1^o \times 1^o$ solar zenith angle (SZA) was restricted to $\leq 65^o$ following the identification of biases in the retrieved τ, r_e and N_d at high SZAs (Grosvenor and Wood, 2014).

6. $1^o \times 1^o$ grid-boxes were rejected if the maximum sea-ice areal coverage over a moving two week window exceeded 0.001 %. The sea-ice data used was the daily $1^o \times 1^o$ version of the "Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data, Version 1" data set (Cavalieri et al., 1996).

7. Only ~~pixels with~~ $1^o \times 1^o$ gridpoints with mean $\tau > 5$ were considered for the N_d data set due to larger uncertainties from instrument error and other sources of reflectance error for τ and r_e retrievals at low τ (Zhang and Plantnick, 2011; Sourdeval et al., 2016).

Following this screening, the $1^o \times 1^o$ gridboxes associated with each MODIS Aqua overpass were averaged into daily mean values for ocean covered surfaces only. Figure 3a shows the number of days from the year of data examined in this study (year 2008) that fulfilled the above criteria and thus are likely to produce a good N_d retrieval. Regions with high numbers of days where useful N_d retrievals can be made have been selected for closer examination in this study; they are listed in Table ~~4~~ 3 along with information on the mean and maximum numbers of days of good data ~~and some statistics on the N_d biases that will be described later~~. The permanent marine stratocumulus decks are among those selected, namely those: in the SE Pacific off the western coast of S. America (Region #1); in the SE Atlantic off the western coast of southern Africa (Region #2); off the coast of California and the Baja Peninsula (Region #3); in the Bering Sea off the SW coast of Alaska (Region #6); and in the Barents Sea to the north of Scandinavia (Region #8). These regions are where the highest numbers of selected days occur with values ranging up to a maximum of 141 days (for the Bering Sea region). The Barents Sea region has the lowest maximum number of days out of this group, reflecting the fact that N_d retrievals cannot be made during a lot of the winter season in this region due to a lack of sunlight. The Southern Ocean (Region #5) and the NW Atlantic (Region #7) regions frequently produce stratocumulus, although it is often associated with the cold sectors of cyclones and so its location from day to day is more transient. These regions are also affected by high solar zenith angles in the winter seasons, which also restricts the number of retrievals possible there. The East China Sea region (Region #4) produces the lowest mean and maximum numbers of days since the stratocumulus areas are mostly restricted to near the coast and occur mostly in the winter season.

~~Following this screening, the $1^o \times 1^o$ gridboxes associated with each MODIS Aqua overpass were averaged into daily mean values.~~ N_d was ~~then calculated~~ calculated for both $r_{e2.1}$ and $r_{e3.7}$ using Eqn. 10 from the $1^o \times 1^o$ daily mean τ, r_e and CTT. ~~N_d was calculated for both $r_{e2.1}$ and $r_{e3.7}$.~~ This was done ~~both~~ by using the retrieved τ value ~~instead of τ^*~~ in Eqn. 10 along with both the retrieved r_e value (i.e., assuming that ~~re(H^*)~~$r_e(H^*)$=r_e(H) as is often assumed for N_d retrievals) and by estimating ~~τ^*~~$r_e(H^*)$ using the retrieved ~~τ~~r_e along with the ~~$d\tau$~~g_{re} values that were calculated as described above. This therefore gives N_d data sets for the "standard" method and a corrected method, allowing the differences between the two to be examined. A similar process was applied for the LWP retrieval.
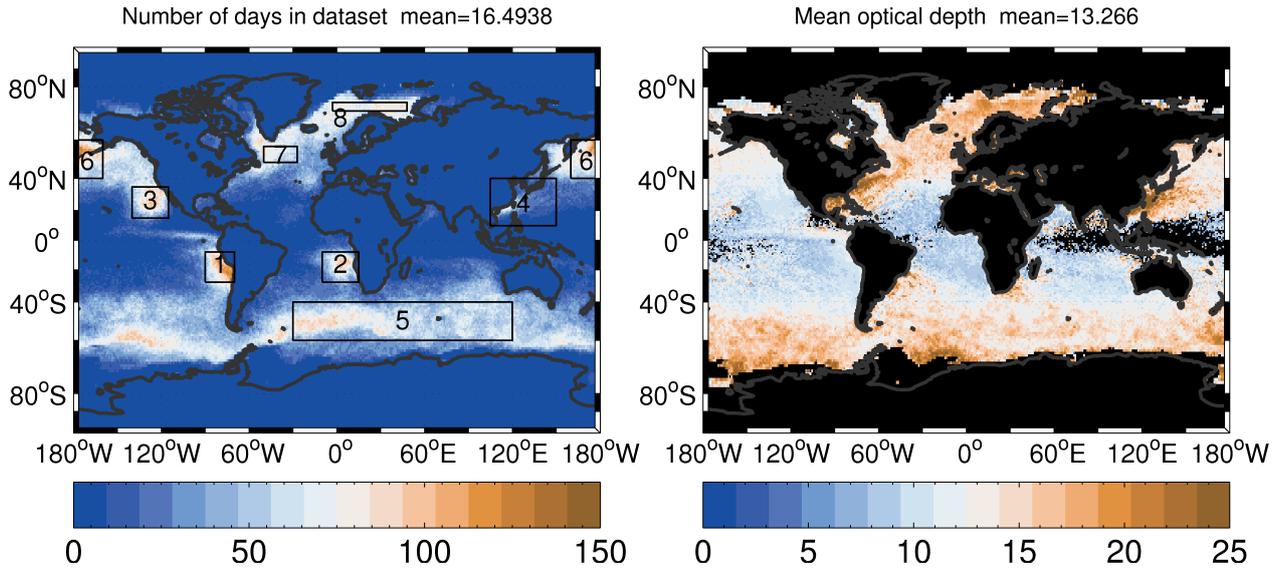
Figure 3. Left: Number of days in 2008 that fulfilled the criteria required to be counted as a valid N_d retrieval. See the text for details on the criteria. Various regions of interest are also denoted by the boxes and numbers. Right: Mean optical depth for data set with ~~only~~ filtering criteria ~~1-6~~ 1-7 applied (see text)~~; i. e., the $\tau > 5$ restriction was not applied.~~

4 Results

Following Eqn 10, the ratio between the uncorrected and corrected N_d values can be shown to be :-

$$\frac{N_{d(uncorrected)}}{N_{d(corrected)}} = \left(\frac{r_e(H)}{r_e(H^*)}\right)^{5/2} = g_{re}^{5/2} \tag{16}$$

5 ~~The equation shows that, for a constant $d\tau$, the relative N_d bias due to an uncorrected τ value would increase with decreasing τ as τ-$d\tau$ approaches zero.~~ Figure 4 shows how the relative N_d bias varies as a function of retrieved τ when using ~~the $d\tau$ values based on those~~ an r_e that has been corrected using the g_{re} from Fig. ~~2~~ 1 (mean curve, black line).

At τ=5 the relative error is ~~50~~46 % for the $r_{e2.1}$ retrieval and ~~30~~28 % for the $r_{e3.7}$ retrieval. At higher τ the errors reduce rapidly, but remain above 10% for the $r_{e2.1}$ retrieval over the τ range shown. For the $r_{e3.7}$ retrieval the relative error drops below 10% for ~~$\tau > \sim 15$~~ $\tau \gtrsim \sim 13$. Thus, the overall degree of error due to this effect will be determined by the distribution of

10 τ for the regions of interest, which we take into consideration here using MODIS data for a representative N_d data set.

Alternatively, if the correction is formulated in terms of a correction to τ we obtain :-

$$\frac{N_{d(uncorrected)}}{N_{d(corrected)}} = \left(\frac{\tau}{\tau^*}\right)^{1/2} = \left(\frac{\tau}{\tau - d\tau}\right)^{1/2} \tag{17}$$

The equation shows that, for a constant $d\tau$, the relative N_d bias due to an uncorrected τ value would increase with decreasing τ as τ-$d\tau$ approaches zero.
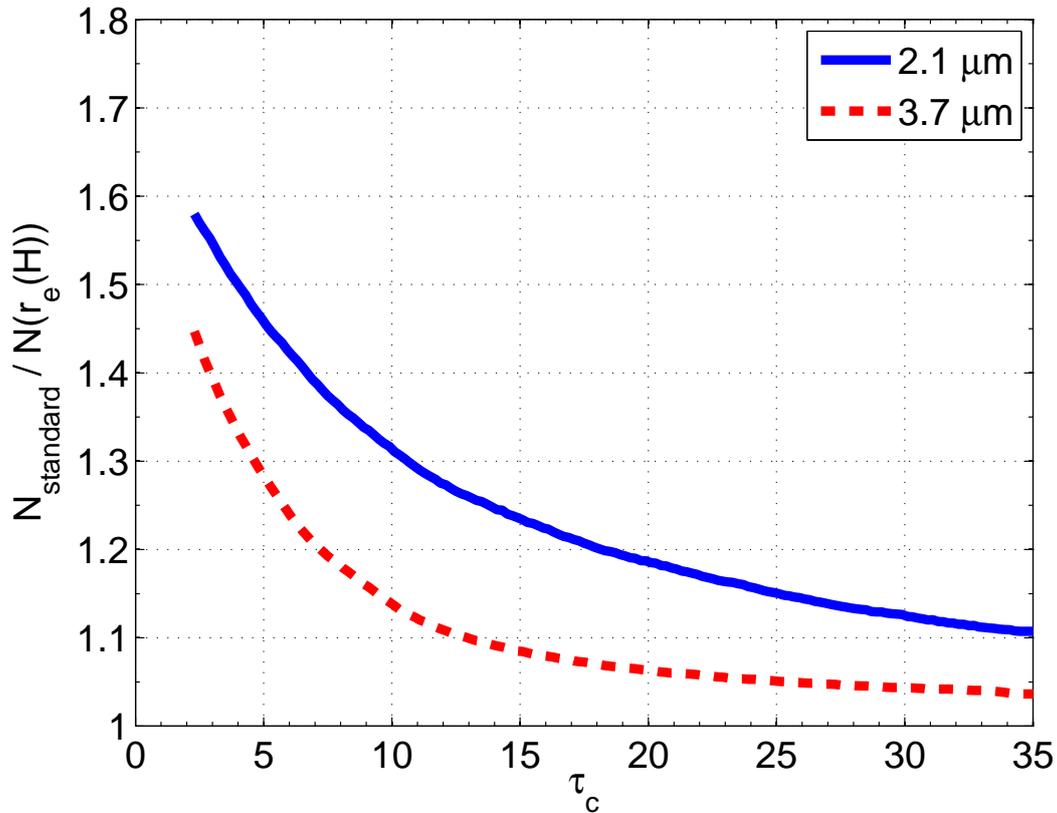

Figure 4. The ratios of N_d values from the standard MODIS calculation (using $\tau^*=\tau$, see the retrieved r_e for r_e^* in Eqn. 10) to those from the corrected calculation (using $\tau^*=\tau$-$d\tau$ the corrected r_e for r_e^* as calculated from the retrieved value and g_{re}; the $d\tau$ g_{re} values used are those shown by the black line in Fig. 2 1) vs retrieved τ.

Figure 3b shows the time–mean τ for the data set as filtered by criteria 1-6 1-7 above; i.e. without the τ >5 restriction to replicate the type of filtering that would likely be performed for N_d retrievals. Table 4 3 lists the regional means of these time–mean values along with the regional means of the standard deviations of τ over time. It shows that the mean τ values of the tropical and sub-tropical regions are generally lower than those at higher latitudes. The East China Sea, Barents Sea, NW Atlantic and Southern Ocean regions exhibit the highest mean τ values out of those examined and so should be expected to show the lowest N_d biases due to the vertical penetration effect. The SE Atlantic region (and the region to the west of Africa in general) show low τ and can be expected to give high N_d biases. Table 4 3 also lists the fraction of days for which $\tau \leq 10$ ($f_{\tau \leq 10}$). τ=10 is the value above which N_d biases drop below 33 31% for the 2.1 μm retrieval and below 15 14 % for $r_{e3.7}$

according to Fig. 4. Thus $f_{\tau \leq 10}$ indicates the fraction of days for which daily N_d biases will be greater than ~~33~~31% for that channel. The values in the table indicate that even in the least affected region (Barents Sea) this will occur for ~~23~~21% of the days. For the SE Atlantic and SE Pacific region the percentages rise to ~~72% and 55~~69% and 53% of the days, respectively. Thus, the vertical penetration depth N_d bias is prevalent in all regions for which N_d data sets are likely to be used, and particularly so in the sub–tropical stratocumulus regions where N_d retrievals have been widely used and studied.

The overall bias is now estimated using one year of actual MODIS data in order to obtain a realistic distribution of τ values. However, it should be noted that the data set used is deliberately filtered in order to only retain datapoints that are likely to give useful N_d data, namely low liquid clouds with extensive $1 \times 1°$ cloud fractions; i.e., predominately stratocumulus. This is done in order to assess biases for the types of clouds that N_d data sets will typically be used to study.

Table 3. Regional statistics for the various marine stratocumulus regions shown in Fig. 3. Shown are the mean and maximum number of days that fulfill the screening criteria in order to be considered as useful N_d retrievals; the regional means and standard deviations (σ) of the time-averaged optical depths (τ) for the screened data set; and the regional mean of the fraction of days for which $\tau \leq 10$ ($f_{\tau \leq 10}$), which is calculated using only data from gridpoints for which the number of days with N_d data was $\geq$ ~~15; regional means of the predicted time-mean percentage biases in N_d due to the vertical penetration depth error; and regional means of the relative (percentage) standard deviations (over time) of the percentage N_d biases (i.e., regional means of the values in Fig. 6). Bias results are shown for both the 2.1 µm and the 3.7 µm r_e retrievals.~~ 15.

#	Region name	Mean no. days	Max no. days	Mean τ	σ_τ	$f_{\tau \leq 10}$
1	SE Pacific	68.3	132	10.5	3.81	0.53
2	SE Atlantic	52.6	107	9.1	3.12	0.69
3	California	62.4	114	10.5	4.06	0.54
4	East China Sea	12.9	77	18.3	10.13	0.24
5	Southern Ocean	58.2	101	14.2	7.58	0.35
6	Bering Sea	73.4	141	13.6	6.84	0.36
7	NW Atlantic	64.3	90	15.9	9.22	0.29
8	Barents Sea	74.9	88	18.0	9.87	0.21

Fig. 5 shows a map of the mean percentage biases and Table 4 gives the regional means of the values in the map. Considering firstly the biases for the $r_{e2.1}$ retrieval, the biases are highest in the tropics and sub-tropics. The regional mean bias is ~~38~~34.5 % for the SE Atlantic region (Region #2), which is the stratocumulus region that seems to suffer the most. The biases are a little

14

Table 4. Regional means of the predicted time-mean percentage biases in N_d and LWP due to the vertical penetration depth error; and regional means of the relative (percentage) standard deviations (over time) of the percentage N_d and LWP biases (i.e., regional means of the values in Fig. 6 and the equivalent for LWP). Bias results are shown for both the 2.1 μm and the 3.7 μm r_e retrievals.

#	Region name	2.1 μm % N_d biases Mean bias (%)	2.1 μm % N_d biases σ (%)	3.7 μm % N_d biases Mean bias (%)	3.7 μm % N_d biases σ (%)	2.1 μm % LWP biases Mean bias (%)	2.1 μm % LWP biases σ (%)	3.7 μm % LWP biases Mean bias (%)	3.7 μm % LWP biases σ (%)
1	SE Pacific	~~72.1 138~~	~~10.2~~	~~3.95 0.55~~	37.0	~~34.6 24.0~~	18.2	~~17.0 40.5~~	33.6
		31.9	21.7	15.0		-10.4		-5.4	
2	SE Atlantic	~~59.7 119~~	~~8.5~~ 18.6	~~3.34 0.72~~	32.0	~~38.4 22.3~~	15.5	~~20.2 37.7~~	29.0
		34.5		17.1		-11.1		-6.1	
3	California	~~68.0 120~~	~~10.0~~	~~4.27 0.58~~	37.5	~~35.4 25.9~~	18.8	~~17.8 43.9~~	34.2
		32.0	22.3	15.1		-10.4		-5.4	
4	East China Sea	~~13.6 86~~ -24.6	~~17.5~~	~~10.2 0.28~~	53.5	~~27.4 40.5~~	31.3	~~12.7 60.3~~	49.5
			36.1	10.7		-8.3		-3.9	
5	Southern Ocean	~~60.6 103~~	~~13.8~~	~~7.68 0.37~~	49.0	~~29.9 34.0~~	27.2	~~13.9 53.7~~	45.1
		27.5	31.6	12.1		-9.1		-4.4	
6	Bering Sea	~~76.0 146~~	~~13.3~~	~~6.94 0.38~~	46.6	~~30.4 31.7~~	25.3	~~14.1 50.8~~	42.8
		28.0	29.5	12.4		-9.3		-4.5	
7	NW Atlantic	~~66.0 91~~ -25.9	~~15.6~~	~~9.29 0.31~~	52.1	~~27.9 36.5~~	29.7	~~12.6 56.3~~	48.0
			34.3	11.2		-8.7		-4.1	
8	Barents Sea	~~76.2 89~~ -23.7	~~17.8~~	~~9.94 0.23~~ -9.9	56.0	~~25.4 39.6~~	32.4	~~11.0 60.1~~	51.6
			37.4			-8.0		-3.6	

lower for the other major stratocumulus regions; e.g. for the SE Pacific region (Region #1) and the Californian region (Region #3) the mean biases are 3532%, although the biases increase further west where the dominant cloud regime tends to shift towards trade cumulus clouds. The remaining regions all have mean biases of 25–3024–28%. The Barents Sea region (Region #8) has a value of only 2523.7%, representing the stratocumulus region with lowest mean bias. These results indicate higher

5   τ values for the clouds in the East China Sea, Southern Ocean, Bering Sea, NW Atlantic and Barents Sea regions relative to the Californian and S.E. Pacific stratocumulus regions, with the SE Atlantic region exhibiting the lowest τ values. This is confirmed by the mean τ values shown in Table 43. The biases for the $r_{e3.7}$ retrieval display the same spatial patterns as for $r_{e2.1}$, but are significantly lower; the mean value in the region with the maximum bias (SE Atlantic, Region #2) is 2017% and that in the region with the lowest bias (Barents Sea, Region #8) is 1110%.

10   The regional mean LWP biases are also listed in Table 4. They are negative since an r_e underestimate from the vertical penetration effect leads to an LWP underestimate (see Eqn. 11). The biases are also smaller in magnitude than for N_d due to the smaller sensitivity of LWP to r_e inherent in the latter equation. They are anti-correlated with the N_d biases such that the region with largest N_d bias (SE Atlantic) has the largest negative LWP bias of -11.1 %. The smallest magnitude bias occurs in that Barents Sea region (-8 %).
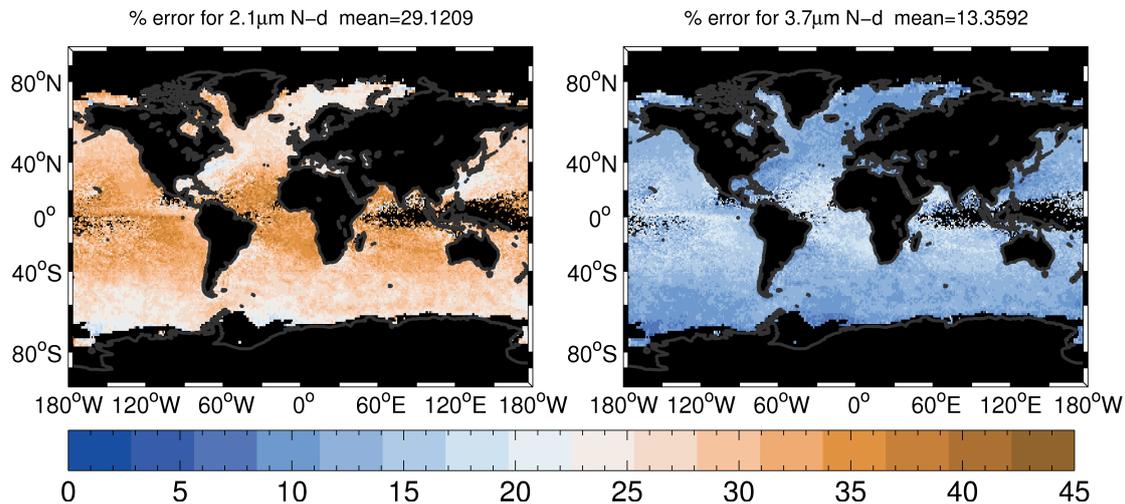

Figure 5. Maps of the annual mean percentage error for uncorrected N_d retrievals using a year (2008) of daily MODIS data that has been filtered to select data points in which N_d retrievals are favourable and therefore most likely to be used for N_d data sets (see text for details). The left plot shows the results for the $r_{e2.1}$ retrieval and the right for the $r_{e3.7}$ one.

15   It is also useful to know how variable the biases are from day to day for a given point in space since this will determine how useful the application of a single offset bias correction might be for correcting N_d biases for daily data. Figure 6 shows the time variability of the bias in the form of the relative standard deviations (over time) of the percentage N_d biases. It reveals that the percentage bias in N_d generally has a larger relative standard deviation at latitudes above around 40° with values typically

ranging up to around 30–50% (of the mean percentage N_d bias) for the 2.1 μm retrieval. Relative variability is greater for the 3.7 μm retrieval, perhaps due to the much lower mean percentage errors. Some of the selected regions show more variability than others, in particular the Barents Sea and East China Sea regions.

Table 4 gives the regional means of the relative standard deviations revealing values that range from approximately 20 to 40% of the mean percentage biases for the 2.1 μm retrieval and 40–6030–60% for the 3.7 μm one. This shows that the application of a single annual mean offset bias correction is likely to lead to fairly large biases for the N_d estimates for individual days for regions where the mean N_d errors are significant. If daily data is used to determine relationships between cloud properties and N_d without correcting for the biases examined here then significant variability in N_d might be introduced that may affect those relationships via non-linear effects.
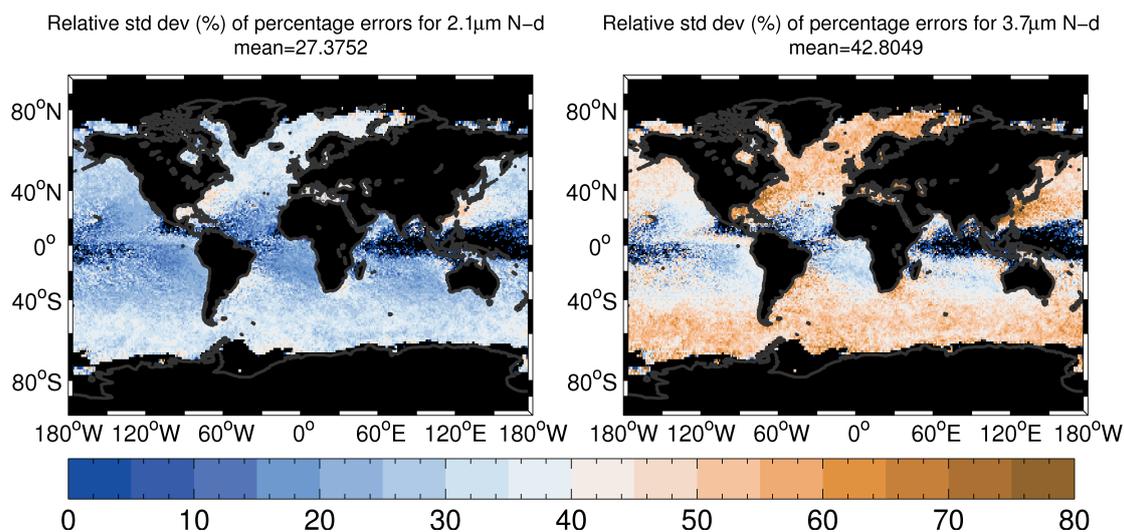
Figure 6. As for Fig. 5 except showing the relative (as a percentage) standard deviation of the percentage N_d bias over time.

Fig. 7 shows how the percentage N_d biases change with season for the $r_{e2.1}$ retrieval only. Interestingly, the highest biases tend to occur in the DJF season for the SE Pacific and SE Atlantic stratocumulus regions, indicating that τ values are lower in DJF for those seasons. The SON season also generally produces higher biases than MAM and JJA for those regions, particularly for SE Pacific. For the East China Sea region the biases are lower in SON and DJF seasons than in the other seasons. We note that there is little data in this region for JJA since there are few low-altitude clouds with large regional liquid cloud fraction fractions there in this season. The other regions either do not show a large amount of seasonal variability, or N_d data is only available for part of the year due to a lack of sunlight in the winter months.
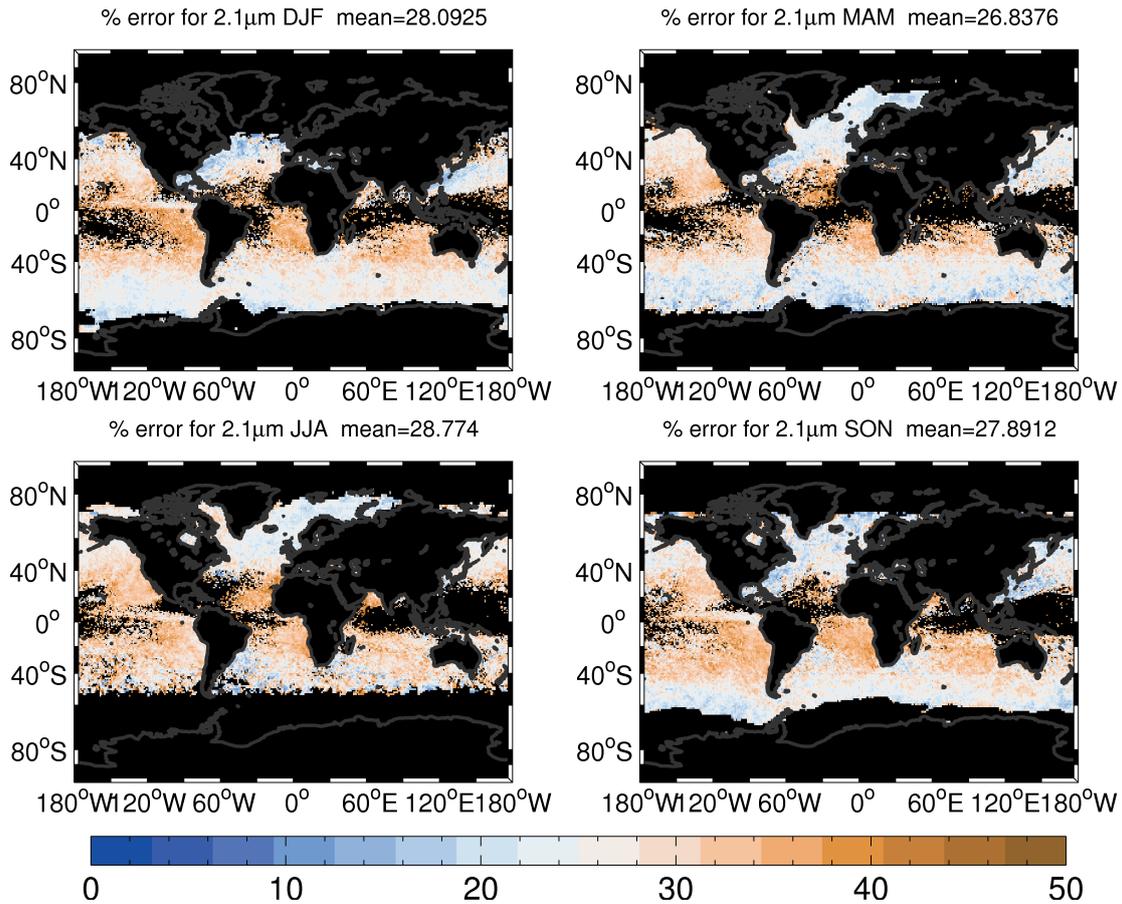
Figure 7. Seasonal mean percentage N_d biases for the $r_{e2.1}$ retrieval only.

5 Discussion

There are some caveats to the results that we presented here that we now discuss. We have shown that, theoretically, the effect of retrieving a lower r_e than the cloud top r_e ~~that is assumed in the N_d retrieval~~ can be corrected for by simply <u>replacing r_e</u> <u>with the parameterized cloud top version, or by</u> removing $d\tau$ from the observed τ. However, this rests upon ~~$d\tau$ having the~~

5   ~~dependence upon τ shown in Fig. 4~~ <u>the parameterizations being valid</u> across all of the cloud types relevant for the N_d ~~data set.~~ ~~This relationship was~~ <u>and LWP data sets. The relationships are</u> based on the retrieved r_e for a range of clouds, although only for a nadir viewing angle and a solar zenith angle of 20°. Platnick (2000) showed that $d\tau$ has some dependence on viewing geometry and so the consideration of a wider range of view and solar zenith angles should ideally be made.

    The modelling of the idealised clouds and the correction rests on the assumption that r_e increases monotonically with height

10   within the cloud (following the adiabatic assumption), but there is some suggestion that the development of precipitation-sized

droplets might lead to larger droplets being preferentially found below cloud top (Chang and Li, 2002; Nakajima et al., 2010a, b; Suzuki et al., 2010) . However, Zhang et al. (2012) found that MODIS retrievals of r_e performed on model generated clouds were not significantly affected by the presence of precipitation. Also, during the VOCALS field campaign in the SE Pacific region, aircraft observations showed that r_e generally did increase with height up to cloud top (Painemal and Zuidema, 2011), indicating that this is not a problem at least for the near-coastal clouds tested. Further offshore the likelihood of precipitation increases as clouds become more cumulus-like and so for those clouds the issue may be greater and hence more caution should be exercised when interpreting the results presented here for such regions.

Evaporation effects related to entrainment also have the potential to reduce r_e, N_d and LWC near cloud top and hence negate some of the assumptions upon which the N_d retrievals rest. However, we argue that the entrainment effect upon r_e is likely to be minimal for two reasons: Firstly, the evidence suggests that for stratocumulus clouds extreme inhomogeneous mixing occurs at cloud top, which reduces the LWC and N_d, but does not change r_e (Burnet and Brenguier, 2007; Brenguier et al., 2011; Painemal and Zuidema, 2011). Secondly, the results of Painemal and Zuidema (2011) indicate that entrainment occurs within approximately the first 0.5 optical depths from cloud top on average; the penetration depths calculated here are considerably larger than this for reasonably thick clouds (Fig. 2). The effect of the reduced N_d and LWC within the entrainment zone is not so clear-cut; this would negate the assumption of a vertically constant N_d and monotonically increasing LWC used to formulate the total τ. However, given the likely small τ contribution from the entrainment region relative to the total τ, this effect is likely to be small.

It is also clear that the suggested correction for the vertical penetration effect should only be applied to the retrievals of N_d with consideration of other bias sources. These other potential error sources are numerous and include r_e biases due to sub-pixel heterogeneity (Zhang and Plantnick, 2011; Zhang et al., 2012, 2016); 3D radiative effects (Marshak et al., 2006); assumptions regarding the degree of cloud adiabaticity (f_{ad} in Eqn. 10; Janssen et al., 2011; Merk et al., 2016); the choice of k value (assumed constant; Brenguier et al., 2011; Merk et al., 2016); the assumption of a vertically uniform N_d; the assumed droplet size distribution shape and width (Zhang, 2013); viewing geometry effects (Várnai and Davies, 1999; Horváth, 2004; Varnai and Marshak, 2007; Kato and Marshak, 2009; Liang et al., 2009; Di Girolamo et al., 2010; Maddux et al., 2010; Liang and Girolamo, 2013; Grosvenor and Wood, 2014; Liang et al., 2015; Bennartz and Rausch, 2017); upper level cloud and aerosol layers (Haywood et al., 2004; Bennartz and Harshvardhan, 2007; Davis et al., 2009; Meyer et al., 2013; Adebiyi et al., 2015; Sourdeval et al., 2013, 2016), etc. These errors have the potential to bias N_d in a way that opposes the positive bias expected from the vertical penetration effect such that the overall biases may cancel out. Indeed, the largest source of error in N_d is likely that from r_e biases given the sensitivity of N_d to r_e in Eqn. 10. MODIS r_e has generally been shown to be biased positively compared to aircraft observations (Painemal and Zuidema, 2011; King et al., 2013), which would lead to a negative N_d error when taken alone. Thus, the application of the correction described in this paper in isolation has the potential to enhance any negative bias in N_d caused by a positive r_e bias.

Our paper quantifies the vertical penetration bias in isolation to the other effects mentioned above. It should be questioned, though, whether the presence of cloud heterogeneity and other effects somehow prevent the effects of the vertical stratification from influencing the retrieved r_e, making it irrelevant. This could be a potential explanation for why it is often observed that

$r_{e2.1}$ is larger than $r_{e3.7}$ (Zhang and Plantnick, 2011) in contrast to the direction expected from adiabatic clouds given the vertical penetration effect, since it is known that sub-pixel heterogeneity effects tend to cause a positive $r_{e2.1}$ bias relative to $r_{e3.7}$ (Zhang et al., 2012). We argue, though, that the vertical stratification effect occurs in addition to other effects (heterogeneity, etc.) with the latter cancelling out and often exceeding the former such that the positive difference between $r_{e2.1}$ and $r_{e3.7}$ would

5   be even larger without the vertical stratification effect. The cancellation of biases may also explain why VOCALS aircraft measurements (Painemal and Zuidema, 2011) tended to show that $r_{e2.1}$ and $r_{e3.7}$ were very similar. We also note that there are many situations when the expected result due to vertical stratification of r_e does occur (i.e. $r_{e3.7} > r_{e2.1}$), as demonstrated in Fig. 8. This shows ratios between $r_{e3.7}$ and $r_{e2.1}$ for an example MODIS scene in the SE Pacific stratocumulus region. Ratios using the uncorrected MODIS r_e values are shown, which shows that the ratio is exceeds one for most of the stratocumulus

10   cloud region (the clouds that adjoin the coast) with ratios ranging from around 1.1 to 1.2. In the more broken clouds the ratio is less than one, which is likely a result of cloud heterogeneity. However, it would be expected that N_d retrievals would not be applied to such clouds. The figure also shows the ratios calculated using $r_{e3.7}$ and $r_{e2.1}$ values that have been corrected using the g_{re} factors. If the differences between $r_{e3.7}$ and $r_{e2.1}$ were caused by vertical stratification alone and if our parameterization was correctly predicting the cloud top r_e for both MODIS channles then this ratio should be equal to one. This is the case for

15   the clouds close to the coast indicating that our parameterization is working well for these clouds. The ratios are a little higher than one further north and west (around 1.05-1.08) indicating that either our parameterization is not working correctly for these clouds, or that other factors are causing relative differences between $r_{e3.7}$ and $r_{e2.1}$ (e.g. sub-adiabaticity, cloud heterogeneity, etc.). Figure 9 shows the percentage of pixels for which $r_{e3.7} > r_{e2.1}$ for 90 days of MODIS SE Pacific observations divided into four different heterogeneity bins. Heterogeneity is characterized by the H_σ parameter (Liang et al., 2009), which is the

20   standard deviation of the 250 m resolution 0.86 µm reflectance ($R_{0.86}$) divided by the mean $R_{0.86}$. It is clear that for many regions relative r_e values that are consistent with an adiabatic profile occur more than 50% of the time, particularly when the cloud heterogeneity is low. This suggests that it may be possible to use H_σ to determine the situations in which the bias correction is more applicable. However, it is hard to definitively prove our argument within the scope of this study, particularly for more heterogeneous regions, since it would likely require computationally expensive 3D radiative transfer modelling of

25   known cloud fields (e.g., from LES models), followed by r_e and τ retrievals.

Painemal and Zuidema (2011) actually demonstrated that MODIS N_d agreed rather well with N_d from aircraft for the SE Pacific region despite ~~the fairly large~~ a fairly large positive r_e bias; this was thought to be due to the fortuitous cancellation of (for N_d) the r_e bias with biases in the k parameter and f_{ad}. However, the agreement between aircraft and MODIS N_d seen in Painemal and Zuidema (2011) would deteriorate if a correction for the N_d bias due to the penetration depth effect discussed

30   here was also applied. Table 4 indicates that the result would be a MODIS N_d underestimate of around ~~35~~32 % (average for SE Pacific, region#1) for the 2.1 µm retrieval, assuming perfect initial agreement. This indicates that another N_d bias may have been operating in order to give the good observed agreement.

The MODIS retrieval uses reflectances from both a visible and a shortwave infra-red (SWIR) wavelength channel with the former being primarily determined by τ and the latter by r_e. However, a bi-spectral retrieval is used and so there is also some

35   sensitivity of the retrieved τ to the SWIR reflectance, which will be representative of the r_e below cloud top due to the vertical
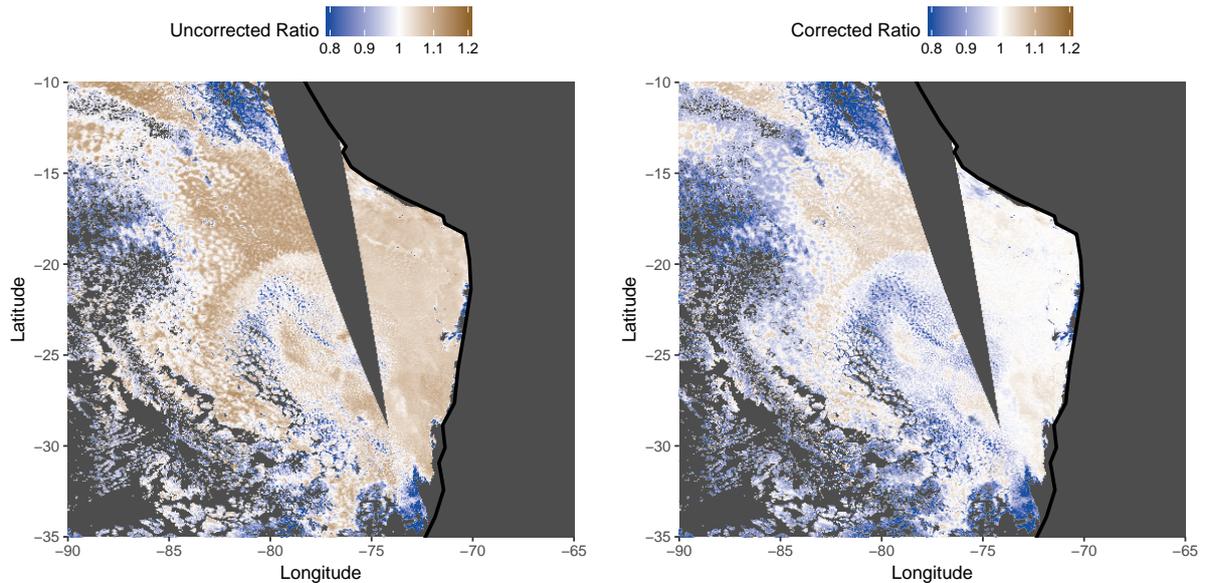
20

Figure 8. Ratios of $r_{e3.7}$ to $r_{e2.1}$ for a MODIS snapshot scene from the Southeast Pacific stratocumulus region from 16th June, 2015. Left: using uncorrected r_e values. Right: using r_e values that have been corrected using the parameterizations for g_{re} (for both $r_{e3.7}$ and $r_{e2.1}$). A ratio of one is expected for the plot on the right if the relative differences between $r_{e3.7}$ and $r_{e2.1}$ are caused by vertical stratification alone and if the parameterization is correctly predicting the relative differences.

penetration effect. This, combined with the fact that the MODIS forward retrieval model assumes a vertically uniform cloud, will result in the retrieved τ being biased relative to the real value (assuming the real cloud has an adiabatic profile). Figure 10 shows the difference between the retrieved and model profile τ; the bias is negative and smaller in magnitude than 5 % for the 3.7 μm retrieval. They are slightly larger for the 2.1 μm retrieval, but still lower in magnitude than 5 %, except at $r_e \lesssim 7$ μm.

5   Although it should be noted that some of this bias may be due to other causes related to the inconsistencies between the vertically uniform and adiabatic models, rather than the r_e vertical penetration bias. Since the retrieved N_d is proportional to the square root of τ, this will lead to small N_d biases. Biases in LWP will be similar to those in τ since LWP is proportional to τ, but r_e biases are still likely to dominate (e.g., see Fig. 1). Thus, we have not pursued this further.

In this paper we have only considered retrievals over the ocean, although retrievals over land for τ and r_e are available for

10 MODIS. MODIS surface albedo uncertainties are likely to be much higher over land than over the oceans (King et al., 2004; Rosenfeld et al., 2004; Bréon and Doutriaux-Boucher, 2005) since the surface albedo is much more variable over land. In addition, cloud masking is more difficult over land, particularly over non–vegetated surfaces, transitional areas between desert and vegetated surfaces and above high-altitude regions (Platnick et al., 2003). We have ignored land regions in order to avoid such complications and also because stratocumulus clouds are more prevalent over ocean regions

15 (Klein and Hartmann, 1993; Wood, 2012). However, the results shown in this paper may still apply over land. The results of Rosenfeld et al. (2004) and Platnick et al. (2017, their Fig. 14) suggest that surface albedo uncertainties are more important at
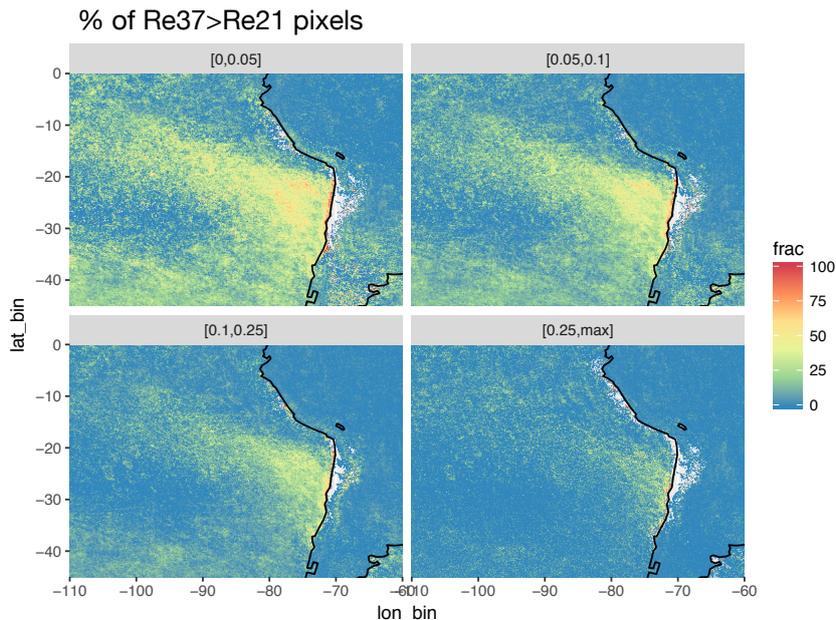
% of Re37>Re21 pixels

Figure 9. The percentage of pixels for which $r_{e3.7} > r_{e2.1}$ for 90 days (January, February and March of 2008) of 0.1° resolution MODIS Collection 6 observations for the SE Pacific stratocumulus region. Only single layer liquid clouds are included and datapoints have been filtered to exclude $\tau < 5$ and partially cloudy pixels. The four panels are for four different bins of the heterogeneity parameter (the standard deviation of the 250 m resolution 0.86 μm reflectance divided by the mean reflectance) with bin ranges labelled above the panels.

lower optical depths ($\lesssim 5$) and for the 2.1 μm retrieval (relative to the 3.7 μm one). Thus, for thicker clouds and the 3.7 μm retrieval land surface albedo issues may be less problematic.

Finally, we note that the thermal emission correction for the MODIS $r_{e3.7}$ (see Section 3.1) retrieval has some uncertainty that should be considered; the uncertainty for this is included (combined with other uncertainties) in the MODIS Collection 6 pixel level uncertainty products (Platnick et al., 2017). It is possible that effects additional to those included, such as cloud heterogeneity, surface heterogeneity, etc., may further increase the uncertainty beyond that estimated in the MODIS products, but these are currently not well documented.

6   Conclusions

We have described and quantified a positive bias in satellite retrievals of cloud droplet concentration (N_d) and Liquid Water Path (LWP) that make use of the adiabatic cloud assumption to estimate N_d these quantities from satellite observed cloud optical depth (τ), effective radius (r_e) and cloud top temperature. We term the N_d bias the "vertical penetration bias". The bias is specific to the methodology of the N_d retrieval, as opposed to being a bias in the underlying observations. It arises due to the
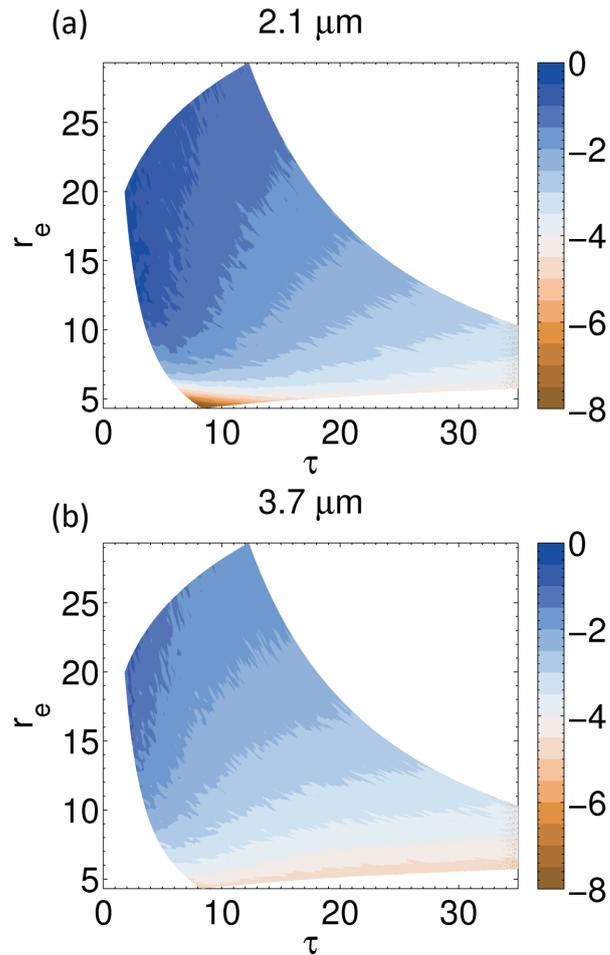
Figure 10. Percentage τ bias (retrieved minus actual value from the input model profile) as a function of τ and r_e.

well–documented vertical penetration of photons with wavelengths in the shortwave-infrared range into the upper regions of clouds, so that r_e retrievals are representative of values some distance below cloud top (Platnick, 2000; Bennartz and Rausch, 2017) rather than being those at cloud top as assumed by the N_d ~~retrieval~~and LWP retrievals. Here we quantified the optical depth as measured from cloud top downwards, $d\tau$, at which the retrieved r_e equaled the actual r_e for adiabatic clouds covering a large range of total cloud optical depths and N_d values. We showed that knowledge of $d\tau$ allows a corrected N_d to be calculated by subtracting $d\tau$ from the observed τ and using that in the N_d retrieval instead of τ. We characterised $d\tau$ as functions of τ for the 2.1 and 3.7 µm r_e retrievals ($r_{e2.1}$ and $r_{e1.6}$, respectively) and found that a 1-D relationship approximates the modelled data well. $d\tau$ -increases with τ and is larger for $r_{e2.1}$ than for $r_{e3.7}$ and so the vertical penetration N_d bias affects retrievals based on

$r_{e2.1}$ more than those using $r_{e3.7}$. Similarly, we also parameterized the true cloud top effective radius ($r_e(H)$) as a function of the retrieved r_e and τ, allowing both a corrected N_d and LWP to be calculated by using $r_e(H)$ instead of the retrieved value. Both the $d\tau$ and r_e correction methods give similar results for the N_d retrievals suggesting that the latter is preferable since it also allows for a correction to LWP. However, for some applications it may be useful to be able to parameterize $d\tau$.

5   We quantified the vertical penetration N_d bias for one–year N_d and LWP data sets. The corrections presented here suggest that N_d and LWP errors will increase as the τ value of the cloud scene gets lower. For many regions that are considered trustworthy for N_d and LWP retrievals (typically stratocumulus regions), there are high frequencies of low τ values and the N_d biases are significant. For example, for the SE Pacific and SE Atlantic regions clouds with $\tau \leq 10$ (for which N_d errors are expected to be ≥ 31% for $r_{e2.1}$ and ≥ 15% for $r_{e3.7}$) occur, respectively,

10   53 and 69% of the time on average. The mean $r_{e2.1}$ vertical penetration N_d biases for these regions were 32 and 35%, respectively. Out of the stratocumulus regions examined, these two were the worst affected. For $r_{e3.7}$ the N_d biases were much smaller; for example, mean biases for the SE Pacific and SE Atlantic regions were 15 and 17%, respectively. N_d biases were predicted to be worse for the tropical and sub-tropical regions than for higher latitudes. The time-variability of the biases were also examined and were shown to be significant (regional mean standard deviations

15   of 19–37% and 32–56% for $r_{e2.1}$ and $r_{e3.7}$, respectively). This indicates that long term averages of the vertical penetration N_d bias corrections are not useful for correcting N_d data over short timescales (e.g. daily N_d data). We also examined the seasonality of the N_d biases and showed that, for the stratocumulus regions, generally the DJF season was worst affected, followed by SON.

    LWP biases were of a lower magnitude than those for N_d and were negative. The largest biases were again for the SE Atlantic

20   region where the mean bias was -11.1 % and the smallest for the Barents Sea region (-8 %). Biases were also lower when using $r_{e3.7}$ with a maximum (most negative) bias of -6.1 %.

    We caution that the correction for the N_d and LWP vertical penetration biases presented here should only be considered in combination with corrections for other biases that affect τ and r_e. Zhang et al. (2016) suggest a correction for the sub-pixel heterogeneity bias effect, but corrections may not currently

25   exist for all biases and it is likely that some unidentified biases still exist. Therefore, we recommended that our correction is currently only applied to homogeneous cloud scenes in order to minimize possible entanglements with biases resulting from heterogeneity effects, which are not accounted for. Such conditions can be obtained by limiting retrievals to associated heterogeneity (H_σ) values (available in MODIS MYD06 Collection 6 products) to less than about 0.1. Otherwise N_d and LWP biases could be made worse; for example, in situations where the fortuitous cancellation of opposing errors leads

30   to initially small N_d errors. The latter was suspected to have occurred for the comparison between MODIS N_d retrievals and in-situ aircraft observations as presented in Painemal and Zuidema (2011). We showed that the SE Pacific, which is the region examined in that study, had a mean vertical penetration depth error of 32% suggesting that another unidentified N_d bias may have been operating in order to give good agreement.

Previous studies have shown that $r_{e3.7}$ is less prone to biases due to sub-pixel averaging (Zhang and Plantnick, 2011; Zhang et al., 2012, 2016). Thus, combined with the work presented here, this supports the conclusion that $r_{e3.7}$ likely represents a better choice for use in N_d and LWP retrievals.

For future work, it is recommended that additional characterization of $d\tau$ and g_{re} is performed for a range of viewing geometries in order to ensure that the results presented here are robust for all cloud retrievals. The use of 3D radiative transfer calculations and simulated retrievals upon known LES model fields would also be useful for investigating how heterogenity effects might interact with the vertical penetration effects. Further investigation into how the presence of precipitation affects our assumptions and results is also warranted.

7   Data availability

The data set is built from publically available MODIS Level-2 data (Collection 5). The MODIS data were obtained from NASA's Level 1 and Atmosphere Archive and Distribution System(LAADS http://ladsweb.nascom.nasa.gov/). NOAA_OI_SST_V2 data (SST data) was provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their website at http://www.esrl.noaa.gov/psd.

References

Ackerman, A. S., Kirkpatrick, M. P., Stevens, D. E., and Toon, O. B.: The impact of humidity above stratiform clouds on indirect aerosol climate forcing, Nature, 432, 1014–1017, doi:10.1038/nature03174, http://dx.doi.org/10.1038/nature03174, 2004.

Adebiyi, A., Zuidema, P., and Abel, S.: The convolution of dynamics and moisture with the presence of shortwave absorbing aerosols over the southeast Atlantic, J. Clim., 28, 1997–2024, doi:10.1175/JCLI-D-14-00352.1, 2015.

5 Ahmad, I., Mielonen, T., Grosvenor, D. P., Portin, H. J., Arola, A., Mikkonen, S., Kühn, T., Leskinen, A., Joutsensaari, J., Komppula, M., Lehtinen, K. E. J., Laaksonen, A., and Romakkaniemi, S.: Long-term measurements of cloud droplet concentrations and aerosol–cloud interactions in continental boundary layer clouds, Tellus B, 65, doi:10.3402/tellusb.v65i0.20138, http://dx.doi.org/10.3402/tellusb.v65i0.20138, 2013.

Albrecht, B. A.: Aerosols, Cloud Microphysics, and Fractional Cloudiness, Science, 245, 1227–1230, http://www.jstor.org/stable/1704234, 10 1989.

Bennartz, R.: Global assessment of marine boundary layer cloud droplet number concentration from satellite, JOURNAL OF GEOPHYSI-CAL RESEARCH-ATMOSPHERES, 112, doi:10.1029/2006JD007547, 2007.

Bennartz, R.: Global assessment of marine boundary layer cloud droplet number concentration from satellite, J. Geophys. Res., 112, D02 201, doi:10.1029/2006JD007547, 2007.

15 Bennartz, R. and Harshvardhan: Correction to "Global assessment of marine boundary layer cloud droplet number concentration from satellite", J. Geophys. Res., 112, D02 201, doi:10.1029/2007JD008841, 2007.

Bennartz, R. and Rausch, J.: Global and regional estimates of warm cloud droplet number concentration based on 13 years of AQUA-MODIS observations, Atmospheric Chemistry and Physics, 17, 9815–9836, doi:10.5194/acp-17-9815-2017, https://doi.org/10.5194/acp-17-9815-2017, 2017.

20 Berner, A. H., Bretherton, C. S., Wood, R., and Muhlbauer, A.: Marine boundary layer cloud regimes and POC formation in a CRM coupled to a bulk aerosol scheme, Atmospheric Chemistry and Physics, 13, 12 549–12 572, doi:10.5194/acp-13-12549-2013, http://dx.doi.org/10.5194/acp-13-12549-2013, 2013.

Boers, R., Acarreta, J. R., and Gras, J. L.: Satellite monitoring of the first indirect aerosol effect: Retrieval of the droplet concentration of water clouds, J. Geophys. Res., 111, D22 208, doi:10.1029/2005JD006838, 2006.

25 Brenguier, J., Burnet, F., and Geoffroy, O.: Cloud optical thickness and liquid water path-does the k coefficient vary with droplet concentration?, Atmos. Chem. Phys., 11, 9771–9786, doi:10.5194/acp-11-9771-2011, 2011.

Brenguier, J.-L., Pawlowska, H., Schuller, L., Preusker, R., Fischer, J., and Fouquart, Y.: Radiative properties of boundary layer clouds: Droplet effective radius versus number concentration, J. Atmos. Sci., 57, 803–821, doi:10.1175/1520-0469(2000)057<0803:RPOBLC>2.0.CO;2, 2000.

30 Bréon, F. M. and Doutriaux-Boucher, M.: A comparison of cloud droplet radii measured from space, IEEE Transact. Geosci. Remote Sens., 43, 1796–1805, doi:10.1109/TGRS.2005.852838, 2005.

Bretherton, C. S., Blossey, P. N., and Uchida, J.: Cloud droplet sedimentation, entrainment efficiency, and subtropical stratocumulus albedo, Geophysical Research Letters, 34, n/a–n/a, doi:10.1029/2006GL027648, http://dx.doi.org/10.1029/2006GL027648, 2007.

Burnet, F. and Brenguier, J.-L.: Observational study of the entrainment-mixing process in warm convective clouds, J. Atmos. Sci., 64, 1995–35 2011, doi:10.1175/JAS3928.1, 2007.

Cavalieri, D. J., Parkinson, C. L., Gloersen, P., and Zwally, H. J.: Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data, Version 1. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. Accessed 22nd December, 2016., doi:10.5067/8gq8lzqvl0vl, 1996.

Chang, F.-L. and Li, Z.: Estimating the vertical variation of cloud droplet effective radius using multispectral near-infrared satellite measurements, Journal of Geophysical Research: Atmospheres, 107, AAC 7–1–AAC 7–12, doi:10.1029/2001JD000766, http://dx.doi.org/10.1029/2001JD000766, 2002.

Davis, S. M., Avallone, L. M., Kahn, B. H., Meyer, K. G., and Baumgardner, D.: Comparison of airborne in situ measurements and Moderate Resolution Imaging Spectroradiometer (MODIS) retrievals of cirrus cloud optical and microphysical properties during the Midlatitude Cirrus Experiment (MidCiX), J. Geophys. Res. Atmos., 114, D02 203, doi:10.1029/2008JD010284, 2009.

Di Girolamo, L., Liang, L., and Platnick, S.: A global view of one-dimensional solar radiative transfer through oceanic water clouds, GEO-PHYSICAL RESEARCH LETTERS, 37, doi:10.1029/2010GL044094, 2010.

Feingold, G., Koren, I., Yamaguchi, T., and Kazil, J.: On the reversibility of transitions between closed and open cellular convection, Atmospheric Chemistry and Physics, 15, 7351–7367, doi:10.5194/acp-15-7351-2015, http://dx.doi.org/10.5194/acp-15-7351-2015, 2015.

Grosvenor, D. P. and Wood, R.: The effect of solar zenith angle on MODIS cloud optical and microphysical retrievals within marine liquid water clouds, Atmospheric Chemistry and Physics, 14, 7291–7321, doi:10.5194/acp-14-7291-2014, http://www.atmos-chem-phys.net/14/7291/2014/, 2014.

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. M., Diamond, M. S., Feingold, G., Fridlind, A., Hünerbein, A., Knist, C. L., Kollias, P., Marshak, A., McCoy, D., 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, doi:10.1029/2017rg000593, https://doi.org/10.1029/2017rg000593, 2018.

Han, Q., Rossow, W., Chou, J., and Welch, R.: Global variations of column droplet concentration in low-level clouds, Geophys. Res. Lett., 25, 1419–1422, doi:10.1029/98GL01095, 1998.

Hartmann, D. L., Ockert-Bell, M. E., and Michelsen, M. L.: The Effect of Cloud Type on Earth's Energy Balance: Global Analysis, Journal of Climate, 5, 1281–1304, doi:10.1175/1520-0442(1992)005<1281:teocto>2.0.co;2, http://dx.doi.org/10.1175/1520-0442(1992)005<1281:TEOCTO>2.0.CO;2, 1992.

Haywood, J. M., Osborne, S. R., and Abel, S. J.: The effect of overlying absorbing aerosol layers on remote sensing retrievals of cloud effective radius and cloud optical depth, Q. J. R. Meteorol. Soc., 130, 779–800, doi:10.1256/qj.03.100, 2004.

Hill, A. A., Feingold, G., and Jiang, H.: The Influence of Entrainment and Mixing Assumption on Aerosol–Cloud Interactions in Marine Stratocumulus, Journal of the Atmospheric Sciences, 66, 1450–1464, doi:10.1175/2008jas2909.1, http://dx.doi.org/10.1175/2008JAS2909.1, 2009.

Horváth, Á.: Anisotropy of water cloud reflectance: A comparison of measurements and 1D theory, Geophysical Research Letters, 31, doi:10.1029/2003gl018386, https://doi.org/10.1029/2003gl018386, 2004.

Janssen, R. H. H., Ganzeveld, L. N., Kabat, P., Kulmala, M., Nieminen, T., and Roebeling, R. A.: Estimating seasonal variations in cloud droplet number concentration over the boreal forest from satellite observations, Atmos. Chem. Phys., 11, 7701–7713, doi:10.5194/acp-11-7701-2011, 2011.

Kato, S. and Marshak, A.: Solar zenith and viewing geometry-dependent errors in satellite retrieved cloud optical thickness: Marine stratocumulus case, J. Geophys. Res., 114, D01 202, doi:10.1029/2008JD010579, 2009.

King, M. D., Platnick, S., Yang, P., Arnold, G. T., Gray, M. A., Riedi, J. C., Ackerman, S. A., and Liou, K.-N.: Remote Sensing of Liquid Water and Ice Cloud Optical Thickness and Effective Radius in the Arctic: Application of Airborne Multispectral MAS Data, Journal of Atmospheric and Oceanic Technology, 21, 857–875, doi:10.1175/1520-0426(2004)021<0857:rsolwa>2.0.co;2, https://doi.org/10.1175/1520-0426(2004)021<0857:rsolwa>2.0.co;2, 2004.

King, M. D., Tsay, S.-C., Platnick, S. E., Wang, M., and Liou, K. N.: Cloud retrieval algorithms for MODIS. Optical thickness, effective particle radius, and thermodynamic phase., NASA, MODIS Algorithm Theoretical Basis document No. ATBD-MOD-05, 1997.

King, M. D., Tsay, S.-C., Platnick, S. E., Wang, M., and Liou, K. N.: MODIS Cloud Optical Properties: User Guide for the Collection 6 Level-2 MOD06/MYD06 Product and Associated Level-3 Datasets, Version 1.0, NASA, https://modis-atmos.gsfc.nasa.gov/sites/default/files/ModAtmo/C6MOD06OPUserGuide.pdf, 2015.

King, N. J., Bower, K. N., Crosier, J., and Crawford, I.: Evaluating MODIS cloud retrievals with in situ observations from VOCALS-REx, Atmospheric Chemistry and Physics, 13, 191–209, doi:10.5194/acp-13-191-2013, https://doi.org/10.5194/acp-13-191-2013, 2013.

Klein, S. A. and Hartmann, D. L.: The Seasonal Cycle of Low Stratiform Clouds, Journal of Climate, 6, 1587–1606, doi:10.1175/1520-0442(1993)006<1587:tscols>2.0.co;2, https://doi.org/10.1175/1520-0442(1993)006<1587:tscols>2.0.co;2, 1993.

Latham, J., Rasch, P., Chen, C.-C., Kettles, L., Gadian, A., Gettelman, A., Morrison, H., Bower, K., and Choularton, T.: Global temperature stabilization via controlled albedo enhancement of low-level maritime clouds, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 366, 3969–3987, doi:10.1098/rsta.2008.0137, http://dx.doi.org/10.1098/rsta.2008.0137, 2008.

Liang, L. and Girolamo, L. D.: A global analysis on the view-angle dependence of plane-parallel oceanic liquid water cloud optical thickness using data synergy from MISR and MODIS, Journal of Geophysical Research: Atmospheres, pp. n/a–n/a, doi:10.1029/2012JD018201, http://dx.doi.org/10.1029/2012JD018201, 2013.

Liang, L., Di Girolamo, L., and Platnick, S.: View-angle consistency in reflectance, optical thickness and spherical albedo of marine water-clouds over the northeastern Pacific through MISR-MODIS fusion, Geophys. Res. Lett., 36, doi:10.1029/2008GL037124, 2009.

Liang, L., Girolamo, L. D., and Sun, W.: Bias in MODIS cloud drop effective radius for oceanic water clouds as deduced from optical thickness variability across scattering angles, Journal of Geophysical Research: Atmospheres, 120, 7661–7681, doi:10.1002/2015jd023256, https://doi.org/10.1002/2015jd023256, 2015.

Liang, L., Di Girolamo, L., and Platnick, S.: View-angle consistency in reflectance, optical thickness and spherical albedo of marine water-clouds over the northeastern Pacific through MISR-MODIS fusion, GEOPHYSICAL RESEARCH LETTERS, 36, doi:10.1029/2008GL037124, 2009.

Maddux, B. C., Ackerman, S. A., and Platnick, S.: Viewing Geometry Dependencies in MODIS Cloud Products, JOURNAL OF ATMOSPHERIC AND OCEANIC TECHNOLOGY, 27, 1519–1528, 2010.

Marshak, A., Platnick, S., Varnai, T., Wen, G., and Cahalan, R.: Impact of three-dimensional radiative effects on satellite retrievals of cloud droplet sizes, JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES, 111, doi:10.1029/2005JD006686, 2006.

Martin, G., JOHNSON, D., and SPICE, A.: THE MEASUREMENT AND PARAMETERIZATION OF EFFECTIVE RADIUS OF DROPLETS IN WARM STRATOCUMULUS CLOUDS, JOURNAL OF THE ATMOSPHERIC SCIENCES, 51, 1823–1842, doi:10.1175/1520-0469(1994)051<1823:TMAPOE>2.0.CO;2, 1994.

Merk, D., Deneke, H., Pospichal, B., and Seifert, P.: Investigation of the adiabatic assumption for estimating cloud micro- and macrophysical properties from satellite and ground observations, Atmos. Chem. Phys., 16, 933–952, doi:10.5194/acp-16-933-2016, 2016.

Meyer, K., Platnick, S., Oreopoulos, L., and Lee, D.: Estimating the direct radiative effect of absorbing aerosols overlying marine boundary layer clouds in the southeast Atlantic using MODIS and CALIOP, J. Geophys. Res., 118, 4801–4815, doi:10.1002/jgrd.50449, 2013.

Nakajima, T., Higurashi, A., Kawamoto, K., and Penner, J. E.: A possible correlation between satellite-derived cloud and aerosol microphysical parameters, Geophys. Res. Lett., 28, 1171–1174, doi:10.1029/2000GL012186, 2001.

Nakajima, T. Y., Suzuki, K., and Stephens, G. L.: Droplet Growth in Warm Water Clouds Observed by the A-Train. Part I: Sensitivity Analysis of the MODIS-Derived Cloud Droplet Sizes, JOURNAL OF THE ATMOSPHERIC SCIENCES, 67, 1884–1896, doi:10.1175/2009JAS3280.1, 2010a.

Nakajima, T. Y., Suzuki, K., and Stephens, G. L.: Droplet Growth in Warm Water Clouds Observed by the A-Train. Part II: A Multisensor View, JOURNAL OF THE ATMOSPHERIC SCIENCES, 67, 1897–1907, doi:10.1175/2010JAS3276.1, 2010b.

Oreopoulos, L.: The impact of subsampling on MODIS Level-3 statistics of cloud optical thickness and effective radius, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 43, 366–373, doi:10.1109/TGRS.2004.841247, 2005.

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, doi:10.1029/2011JD016155, 2011.

Pawlowska, H., Grabowski, W. W., and Brenguier, J.-L.: Observations of the width of cloud droplet spectra in stratocumulus, Geophysical Research Letters, 33, doi:10.1029/2006gl026841, https://doi.org/10.1029%2F2006gl026841, 2006.

Platnick, S.: Vertical photon transport in cloud remote sensing problems, JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES, 105, 22 919–22 935, doi:10.1029/2000JD900333, 2000.

Platnick, S. and Valero, F. P. J.: A Validation of a Satellite Cloud Retrieval during ASTEX, Journal of the Atmospheric Sciences, 52, 2985–3001, doi:10.1175/1520-0469(1995)052<2985:avoasc>2.0.co;2, https://doi.org/10.1175/1520-0469(1995)052<2985:avoasc>2.0.co;2, 1995.

Platnick, S., Meyer, K. G., D., K. M., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Holz, R. E., Yang, P., Ridgway, W. L., and Riedi, J.: The MODIS cloud optical and microphysical products: Collection 6 updates and examples from Terra and Aqua, IEEE Transact. Geosci. Remote Sens., 55, 502–525, doi:10.1109/TGRS.2016.2610522, 2017.

Platnick, S., King, M., Ackerman, S., Menzel, W., Baum, B., Riedi, J., and Frey, R.: The MODIS cloud products: Algorithms and examples from Terra, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 41, 459–473, doi:10.1109/TGRS.2002.808301, 2003.

Quaas, J., Boucher, O., and Lohmann, U.: Constraining the total aerosol indirect effect in the LMDZ and ECHAM4 GCMs using MODIS satellite data, Atmos. Chem. Phys., 6, 947–955, doi:10.5194/acp-6-947-2006, 2006.

Rosenfeld, D., Cattani, E., Melani, S., and Levizzani, V.: Considerations on daylight operation of 1.6-versus 3.7-um channel on NOAA and METEOP satellites, Bulletin of the American Meteorological Society, 85, 873–882, doi:10.1175/BAMS-85-6-873, https://doi.org/10.1175/BAMS-85-6-873, 2004.

Salomonson, V. V., Barnes, W. L., Maymon, P. W., Montgomery, H. E., and Ostrow, H.: MODIS: advanced facility instrument for studies of the Earth as a system, IEEE Transact. Geosci. Remote Sens., 27, 145–153, doi:10.1109/36.20292, 1998.

Sourdeval, O., Labonnote, L. C., Brogniez, G., Jourdan, O., Pelon, J., and Garnier, A.: A variational approach for retrieving ice cloud properties from infrared measurements: application in the context of two IIR validation campaigns, Atmos. Chem. Phys., 13, 8229–8244, doi:10.5194/acp-13-8229-2013, 2013.

Sourdeval, O., C. Labonnote, L., Baran, A. J., Mülmenstädt, J., and Brogniez, G.: A methodology for simultaneous retrieval of ice and liquid water cloud properties. Part 2: Near-global retrievals and evaluation against A-Train products, Quarterly Journal of the Royal Meteorological Society, 142, 3063–3081, doi:10.1002/qj.2889, 2016.

Stamnes, K., Tsay, S. C., Wiscombe, W., and Jayaweera, K.: Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media, Appl. Opt., 27, 2502–2509, 1988.

5 Stevens, B., Cotton, W. R., Feingold, G., and Moeng, C.-H.: Large-Eddy Simulations of Strongly Precipitating, Shallow, Stratocumulus-Topped Boundary Layers, J. Atmos. Sci., 55, 3616–3638, doi:10.1175/1520-0469(1998)055<3616:lesosp>2.0.co;2, http://dx.doi.org/10.1175/1520-0469(1998)055<3616:LESOSP>2.0.CO;2, 1998.

Suzuki, K., Nakajima, T. Y., and Stephens, G. L.: Particle Growth and Drop Collection Efficiency of Warm Clouds as Inferred from Joint-CloudSatand MODIS Observations, Journal of the Atmospheric Sciences, 67, 3019–3032, doi:10.1175/2010jas3463.1, https://doi.org/10.1175/2010jas3463.1, 2010.

10 Szczodrak, M., Austin, P., and Krummel, P.: Variability of optical depth and effective radius in marine stratocumulus clouds, J. Atmos. Sci., 58, 2912–2926, doi:10.1175/1520-0469(2001)058<2912:VOODAE>2.0.CO;2, 2001.

Twomey, S.: Pollution and the planetary albedo, Atmos. Environ., 8, 1251–1258, doi:10.1016/0004-6981(74)90004-3, 1974.

van de Hulst, H.: Light Scattering by small particles, Dover Publications, New York, 1957.

15 Várnai, T. and Davies, R.: Effects of Cloud Heterogeneities on Shortwave Radiation: Comparison of Cloud-Top Variability and Internal Heterogeneity, J. Atmos. Sci., 56, 4206–4224, doi:10.1175/1520-0469(1999)056<4206:EOCHOS>2.0.CO;2, 1999.

Varnai, T. and Marshak, A.: View angle dependence of cloud optical thicknesses retrieved by Moderate Resolution Imaging Spectroradiometer (MODIS), JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES, 112, doi:10.1029/2005JD006912, 2007.

Wood, R.: Stratocumulus Clouds, MONTHLY WEATHER REVIEW, 140, 2373–2423, doi:10.1175/MWR-D-11-00121.1, 2012.

20 Zhang, Z.: On the sensitivity of cloud effective radius retrieval based on spectral method to bi-modal droplet size distribution: A semi-analytical model, Journal of Quantitative Spectroscopy and Radiative Transfer, 129, 79–88, doi:10.1016/j.jqsrt.2013.05.033, http://dx.doi.org/10.1016/j.jqsrt.2013.05.033, 2013.

Zhang, Z. and Plantnick, S.: An assessment of differences between cloud effective particle radius retrievals for marine water clouds from three MODIS spectral bands , JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES, 116, 2011.

25 Zhang, Z., Werner, F., Cho, H. M., Wind, G., Platnick, S., Ackerman, A. S., Di Girolamo, L., Marshak, A., and Meyer, K.: A framework based on 2-D Taylor expansion for quantifying the impacts of subpixel reflectance variance and covariance on cloud optical thickness and effective radius retrievals based on the bispectral method, J. Geophys. Res. Atmos., 121, 7007–7025, doi:10.1002/2016JD024837, 2016.

Zhang, Z., Ackerman, A. S., Feingold, G., Platnick, S., Pincus, R., and Xue, H.: Effects of cloud horizontal inhomogeneity and drizzle on remote sensing of cloud droplet effective radius: Case studies based on large-eddy simulations, JOURNAL OF GEOPHYSICAL 30 RESEARCH-ATMOSPHERES, 117, doi:10.1029/2012JD017655, 2012.

Zuidema, P., Painemal, D., de Szoeke, S., and Fairall, C.: Stratocumulus Cloud-Top Height Estimates and Their Climatic Implications, Journal of Climate, 22, 4652–4666, doi:10.1175/2009jcli2708.1, http://dx.doi.org/10.1175/2009JCLI2708.1, 2009.