

The authors are very grateful for the helpful comments provided by both reviewers. Our point-by-point responses are provided below in blue directly after the reviewers' comments.

To summarize, we made 5 main changes in the manuscript:

- 1) we now use the latest 2b-cldclass-lidar release (R05) in the case studies and statistics since the changes compared to R04 are substantial –although the conclusions remain unchanged–, which is not the case for the RL-GeoProf (reviewer 1).
- 2) we further justify future use of CASCCAD dataset for model evaluation in the conclusions (reviewer 1).
- 3) we now compare the CASCCAD datasets with the passive-sensor clustering datasets in addition to the COT-CTP technique (reviewer 1 and 2).
- 4) we further define the content of the CASCCAD dataset at the beginning of the dataset section (reviewer 2).
- 5) finally, we slightly modified the CASCCAD algorithm to keep track of the cloud type called Cu with stratiform outflow, which was accounted for as pure Cu before. As a result:
 - a. the algorithm chart flow (fig. 2) has been updated as well as the cartoon explaining the VCF and horizontal continuity test logic (Fig. 5).
 - b. the case studies have been updated although the changes are minor
 - c. the global statistics are slightly changed in RL-GeoProf CASCCAD (a little less/more Cu/transitioning cloud) while it is negligible in GOCCP-CASCCAD.
 - d. these changes do not impact the conclusions of the study

Reviewer 1

Summary

This paper describes a simple algorithm (CASCCAD) to classify low level clouds into either stratocumulus or cumulus based on cloud top height, horizontal cloud fraction, and vertical cloud fraction as measured by active space borne sensors (either lidar or lidar+radar). This algorithm is then compared against an existing radar/lidar algorithm (cldclass). The new dataset (CASCCAD) identifies far more cumulus than (cldclass), which agrees better with our a-priori assumption about the geographic distribution of these cloud types. An extremely simple classification based on the cloud optical thickness and cloud top pressure is also examined and found not to perform favorable compared to the CASCCAD dataset.

The paper is well written. It represents a novel contribution to the field. The paper should be published following some major revisions.

My main complaint with the paper is the comparisons with COT-CTP classification methodologies. For quite some time now people have been using more sophisticated k-means approaches that use a 2d joint histogram of COT-CTP to classify imager data into 'weather states'. If the authors want to compare their algorithm to the imager algorithms they should use the more recent approaches. I don't believe that the current comparison represents a fair assessment of how well imager data can be used to classify cloud regimes. Below is a limited list of papers that use MODIS based k-means regimes. 10.1002/2013JD021409, 10.1002/2016jd025193, 10.1002/2015jd024502, 10.1002/2016jd026120, 10.1007/s00382-016-3064-0, 10.1007/s00382-017-3806-7. I believe that the CASCCAD should be compared against the k-means approach ideally from MODIS but possibly also from ISCCP. The MODIS based regimes do have a clear advantage over CASCCAD; there is a believable way to compare model output to those regimes.

The author needs to seriously elaborate on how CASCCAD can be used to evaluate models or drop the statements about model evaluation. I suggest the latter as I don't think it is possible to do a meaningful comparison.

Regarding the reviewer's two main comments: (1) k-means approaches can only indirectly be compared to CASSCAD because they do not identify individual clouds, but we have now done what we feel we can and included it in the revised manuscript. (2) We feel that CASSCAD is actually a better dataset for evaluating models in a way that is meaningful for tracing errors to their source rather than just concluding that the model is wrong, which we now discuss in the revised paper. Since the reviewer also mentions his two main concerns in the minor concerns below, we address both of these points in more detail directly below.

Other than these two points, I only have minor comments which are listed below.

Comments:

-Page 3, line 23: should be 1.4x1.7 km (tanelli et al., 2008). 1.1 is the distance between adjacent pixels but they overlap somewhat.

-page 4, line 31. Same as above

Thank you for pointing this out. We corrected this in both places in the new manuscript.

-page 6, line 18: The RL-GeoProf release 5 (newer version) is now available past the 2011 anomaly.

The new version came out after we submitted the study. However, even with the additional R05 time period, a substantial amount of data are still missing, which would still limit the possibility of computing statistical relationships between cloud amount and environmental variables. We reproduced the case studies with the R05 version of RL-GeoProf and we found that the change is negligible (see new Fig. S2-3-4-5). For these reasons, we decided not to update the results with the R05 for the sake of computational, space and time resources. We now explain this at the end of section 3.1: *"Note that the release R05 of RL-GeoProf came out after the submission of the original manuscript (late May 2019), which includes data after April 2011. However, these are for daytime only and substantial periods of time are still missing (e.g., May 2011 through May 2012 and the whole year 2014), which makes it difficult to compute a consistent climatology and derive statistical relationships between clouds and environmental variables. Additionally, the differences between CASSCAD using RL-GeoProf R04 and R05 are very small in the case studies of section 4.1 (see Fig. S2-S3-S4-S5). There is a small decrease of the overall cloud fraction (76.3 vs. 75.8 %, 37.2 vs. 35.2 %, 15.9 vs. 16 % and 80.8 vs. 77.2 % for R04 and R05 respectively), which affects mostly the Cu cloud fraction. Since the change is almost negligible, we decided not to update the global statistics with the R05 version for the sake of computational, space and time resources."*

Concerning the 2BCCL product, the change between R04 and R05 is substantial. The overall excess of low-level cloud fraction over the tropical ocean is somewhat fixed in R05, likely due to the use of a different lidar product (LIDAR-AUX, Wang et al., 2013). However, the excess and lack of Sc and Cu, respectively, are still present, most likely because there is no or little change in the 2BCCL algorithm for Sc, St and Cu clouds in the low levels

(http://www.cloudsat.cira.colostate.edu/sites/default/files/products/files/2B-CLDCLASS-LIDAR_PDICD.P1_R05.rev0_.pdf).

It is mentioned in section 2.3 that we use R05 version and that the reader should refer to the original version of the manuscript for results using the older R04: *"Note that there are substantial differences between results using the R04 and R05 versions, which is why the reader should refer to the original manuscript published in May, 22nd 2019 on the ESSD discussion website for results using the older R04 version."*

-page 7, line 18: why are attenuate profiles excluded from HCF? They are certainly cloudy right?

The way attenuated pixels are treated is consistent with what is done in the GOCCP product for cloud fraction computations. Here, we exclude the profile when all pixels (= 4 pixels) below 1.92 km (which is where the HCF is computed) are fully attenuated pixels because it is not possible to know with certainty whether these pixels are clear or cloudy.

We modified this sentence to better explain why this is done: *“Finally, note that when all 480-m-pixels below 1.92 km are fully attenuated (4 pixels), the profile is excluded from the HCF computation – consistent with what is done in the GOCCP product for cloud fraction computations– because we do not know whether these pixels are clear or cloudy.”*

-page 7, section 3.4: can you explain here what influence multiple scattering has on the signal? It seems to me that when you see several bins of lidar return you are probably looking at a multiply scattered signal.

The CASCAD algorithm is applied to level 2 GOCCP and CloudSat-CALIPSO files, which already include orbital cloud mask profiles. Therefore, we do not use directly the attenuated backscattered signal from the lidar, we simply compute an along-track cloud fraction using the GOCCP or CloudSat-CALIPSO cloud mask, which both take into account the multiple scattering issues in their algorithms.

-Section 4.2.1: This COT-CTP method is quite simple. The k-means methods of Rossow et al. (2005) applied to ISCCP or more recently Oreopolous (<https://doi.org/10.1002/2013JD021409>) applied to MODIS probably do a much better job because they consider the joint histograms of COT-CTP over a large area. You should at least comment on this. There is a 1x1 degree daily gridded dataset that has maps of the regimes. If I were you I would get this data and plot his regimes next to yours. The aqua MODIS regimes are even coincident in time and space with yours.

The WS and CR approaches do not discriminate cloud type or even low cloud from middle and high clouds. They represent mixtures of cloud types although one cloud type is often prevalent. As such, it is not exactly consistent to directly compare them with the CASCAD datasets, however, we agree with the reviewers that showing the CR and WS more recent approaches would provide a broader context and additional information. For this reason, we added the ISCCP WSs and MODIS CRs observations in section 4 of the analysis although we already briefly mentioned these techniques in the first version of the manuscript (p11 line 9).

There is an additional figure (new Figure 15) with the ISCCP WSs and MODIS CRs geographical distributions and we added the zonal and global means to the existing Figure 14. In addition, we discuss the new results in section 4.2.1. In essence, the clustering-derived datasets show a better agreement with CALIPSO and CloudSat-CALIPSO CASCAD particularly in the tropics and for the Sc clouds. The Cu-regimes are better correlated with CALIPSO CASCAD in terms of geographical distribution although MODIS underestimates their fraction while ISCCP overestimates it (likely due to mid and high-level clouds contained in WS7 and WS8).

-Page 10, line 29: I would remove this statement unless you want to elaborate. I can't see how you could use this data to evaluate a GCM in anything but a qualitative way. There is generally nothing in the model called Sc and Cu to compare against. Also, there is no way to apply an instrument simulator to recreate the data using the model output because the categories depend on spatial continuity.

We disagree with the reviewer's comments. Models make Sc and Cu explicitly, based on specific but different physical mechanisms. We understand that some models are now using unified turbulence schemes of one type or another, but these schemes still have to produce Cu and Sc and the transition between them under appropriate conditions. We discuss in response to a later comment why CASCAD is valuable for the evaluation of such parameterizations. We also note that Reviewer 2 feels that CASCAD is useful for model evaluation.

It is true that it would be difficult to recreate the same Cu Sc diagnostic in the simulator as in the CASCAD algorithm, however:

- 1) The simulator sums up the convective and stratiform cloud fraction before computing the diagnostics and therefore one could just separate their contribution rather than summing them up.
- 2) A simulator is not necessarily needed for this particular type of comparison, because we identify the different cloud modes explicitly and we can select regimes in which lidar attenuation is negligible (e.g., $\omega_{500} > 0$ hPa/day, Cesana et al., 2019a)

We expanded the last paragraph of the conclusion to discuss this: *“Finally, one of the reasons we developed CASCAD is to provide an improved observational constraint for low-level cloud feedbacks in GCMs. Although the CASCAD DA cannot be implemented in a lidar simulator (Chepfer et al., 2008), it is still possible to use CASCAD datasets for model evaluation because i) both the convective and stratiform cloud fraction are provided as inputs to the lidar simulator and could be easily saved separately rather than summed up; and ii) a simulator is not necessarily needed for model-to-obs comparison of Cu and Sc clouds over the tropical oceans, because we identify the different cloud modes explicitly and we can select regimes in which lidar attenuation is negligible (e.g., $\omega_{500} > 0$ hPa/day, Cesana et al., 2019a).”*

-Figure 15: I’m concerned about the cloud fraction PDF’s (left column). What is the resolution over which the cloud fraction PDF’s are calculated? Is the resolution different for the different products? The PDF depends on the resolution so if the different products have different resolution the PDF’s will be different just based on that fact. For this reason, the nadir-only-sampling of the CALIPSO-CloudSat data will inherently be different than a 2D imager. I think you should remove these panels because of these concerns and I also don’t think they add much to the paper.

To answer the reviewer’s concerns, we removed these panels. We also averaged MISR and MODIS onto a 2.5x2.5 grid, similar to the other products, in the new version of this figure. We now mention this information in the figure’s caption.

-Page 123, line 20: Again, I don’t know how you use this to evaluate a model. Just because the model has some arbitrary distinction between a boundary layer parameterization and a shallow convection parameterization doesn’t mean that this is the same as stratocumulus and cumulus as you have defined them from the observations. Furthermore, newer parameterization are beginning to ‘unify’ these distinct regimes (e.g. CLUBB, EDMF). I think the dataset is interesting enough in its own right without having to sell it as a model evaluation tool.

Here we also disagree with the reviewer’s comment. As stated above, the models do actually simulate Sc and Cu based on physical mechanisms. Existing GCMs fall into two classes: (1) Those that use separate parameterizations for cumulus and stratocumulus clouds, for which CASCAD is directly applicable as an evaluation tool to determine whether biases are attributable to process assumptions made about one cloud type vs. the other. (2) Those that use unified turbulence parameterizations that are intended to represent the full spectrum of boundary layer clouds. Even for the latter class of parameterizations, though, the models are in effect capturing the two cloud types and the transition between them via different segments of the pdf and making assumptions about when each type of cloud is active, either in the physics itself or in the diagnosis of the model behavior. In both cases, the distinction between Cu and Sc is made crudely using criteria such as inversion strength or 500 mb vertical velocity as the “definition” of one vs. the other cloud type. For example, Koehler et al. (2011, QJRMS), in an EDMF scheme, use inversion strength as a “decoupling criterion” to determine when they should turn off the shallow convection (MF) component of EDMF. Bogenschutz et al. (2013, J. Climate) use inversion strength and vertical velocity to define “stratocumulus,” “transition,” and “cumulus” regimes for the purpose of understanding where CLUBB is performing well vs. poorly. We feel that a dataset that directly diagnoses the clouds themselves is ultimately a more reliable indicator than assumptions about cloud types based on

large scale environmental properties. CASSCAD in fact allows the relationship between Sc, Cu, and environmental state to be diagnosed directly rather than being assumed. Given the large impact that schemes like CLUBB have on climate sensitivity (Gettelman et al. 2019, GRL), there is an increasing need to use metrics beyond mean state biases to decide what is and is not realistic. Note that we did not explain this in the previous manuscript. We now specify this in the description of the algorithm, section 3.1: *“However, one cannot separate clouds according to the mechanisms that form them as GCMs do using different PBL and convective parameterizations, which is why we choose to use the morphology to discriminate cloud types in this study.”*

Based on our study, it seems that CASSCAD is doing better at identifying Cu-Sc clouds than the other methods –while it is not perfect– and therefore CASSCAD looks like a good candidate for model evaluation.

We now specify this in the last paragraph of the conclusion: *“Finally, ... to judge model realism and fidelity”*.

Reviewer 2

This paper describes a new global cloud dataset built from active sensor CALIPSO and CloudSat observations. The distinguishing feature of the dataset is that cloud profile data is used to discriminate between stratocumulus and shallow cumulus, rather than relying on cloud-top altitude and optical depth derived from passive sensors. The motivation for this effort is nicely summarized in the introduction. The paper is well written and the work is a useful contribution to the field. The paper deserves publication but would benefit from some additional discussion and clarification of a few points.

1. The algorithm used to identify and discriminate cumulus and stratocumulus is described well, but the paper needs some description of the actual CASCCAD data set:

Is this a Level 2 or Level 3 dataset? What is the spatial resolution of the product? What parameters are reported – are cloud base and top altitudes included?

Section 5 says CASCCAD dataset is formed using both GOCCP and RL-GeoProf.

How? Are these two datasets merged together somehow, or are they included as two separate elements of CASCCAD?

We acknowledge that the description of the dataset was somewhat ambiguous in the previous manuscript. The CASCCAD datasets are actually two distinct datasets using the same algorithm but applied to two different level 2 observational datasets, which are CALIPSO-GOCCP and RL-GeoProf. There are level 2 and level 3 CASCCAD files for both datasets. The level 2 files use the native resolution of the GOCCP and RL-GeoProf datasets (i.e., 333 m x 480 m for GOCCP and ~1.1 km x 240 m for RL-GeoProf). The level 3 files are global statistics over a 2.5°x2.5° grid over 40 480 m levels for each month between 2007 and 2016, which are then averaged over the time dimension. The monthly means will be made available once the manuscript is accepted. The parameters reported are 2D and 3D cloud fractions of low, Cu, Sc and transitioning clouds for level 3 files and cloud mask profiles of Cu, Sc and transitioning cloud types for the level 2 files. While the cloud top/base is not reported explicitly as a variable for each profile, it can be derived from the cloud mask contained in level 2 files.

We added a full paragraph at the beginning of section 2 –before section 2.1– to address the reviewer’s concerns and describe what we report above. Additionally, we make it clear throughout the manuscript that the GOCCP and RL-GeoProf CASCCAD datasets are two distinct datasets based on the same algorithm but using GOCCP native level 2 observations and the combined CloudSat-CALIPSO RL-GeoProf level 2 observations, respectively. For example, in the last paragraph of the introduction, the first of section 4.1 and the first paragraph of section 5, we now specify that the algorithm is apply “separately” to both GOCCP and RL-GeoProf observations. Finally, in the conclusion, when the CASCCAD datasets are mentioned we now refer to “distinct GOCCP and RL-GeoProf CASCCAD datasets” rather than only “CASCCAD datasets”.

2. As reported here, classifications based on observed 3D structure appear to be better than classifications based on passive retrievals of cloud altitude and optical depth, which are subject to ambiguities. It would be useful to cite Mace and Wrenn (2013) who discuss a comparison in the eastern Pacific

Mace and Wrenn, 2013: Evaluation of the Hydrometeor Layers in the East and West Pacific within ISCCP Cloud Top Pressure-Optical Depth Bins Using Merged CloudSat and CALIPSO Data, J. Climate, doi:10.1175/JCLI-D-12-00207.1

Thank you for pointing out this relevant reference. We added it to the manuscript at the beginning of section 4.2.1 where the comparison between passive and active is performed:

In addition, (Mace and Wrenn, 2013) showed that except for thin cirrus and Sc cloud types, the COT-derived cloud types are mostly mixtures of different cloud types in two regions of the Eastern Pacific.

3. Referee #1 dismisses the utility of CASCAD for model evaluation for reasons which are not clear. CASCAD appears quite useful for model evaluation to me. Some additional discussion of this would be useful, especially to clarify differences in using GOCCP (which is based on a simulator approach) versus RL-Geoprof (which is not).

We agree with reviewer 2 and we added some additional discussion –as requested by reviewer 2– to better explain how we believe CASCAD can be used for model evaluation in the conclusion. A point by point answer is given to reviewer 1 above. Based on the case studies, profile and zonal means presented in the paper, the biggest differences between CASCAD versions for the two datasets mentioned above are the tendency for the RL-GeoProf version to diagnose more transitioning clouds than the GOCCP version poleward of 40°, because i) RL-GeoProf is less affected by overlapping mid- and high-cloud attenuation and better captures the full vertical extent of low clouds and ii) RL-GeoProf contains larger cloud clusters than GOCCP, making the continuity test more efficient. We now mention this in the second paragraph of section 4.2.1.

4. Finally, I agree with Referee #1 that it would be interesting and useful for the authors to say something about the weaknesses and strengths of CASCAD vs. weather state approaches to analyzing passive observations.

The WS and CR approaches do not exactly classify clouds by cloud type or by low, middle and high layers although they provide CTP-tau histograms for each WS/CR. Therefore they represent mixtures of cloud types among which one cloud type is often prevalent. As such it is not exactly consistent to directly compare with the CASCAD datasets, however, we agree with the reviewers that showing the CR and WS more recent approaches would provide a broader context and additional information as their horizontal and time sampling and their time record are better than CloudSat-CALIPSO.

For this reason, we added the ISCCP WSs and MODIS CRs observations in section 4 of the analysis.

Minor points:

The last sentence of the first paragraph of the Introduction repeats an earlier sentence.

The first occurrence of this sentence has been removed.

To document limitations of CloudSat in sensing shallow clouds it would be useful to cite: Liu, D., Q. Liu, L. Qi and Y. Fu, 2016: “Oceanic single-layer warm clouds missed by the Cloud Profiling Radar as inferred from MODIS and CALIOP measurements”, JGR, doi:10.1002/2016JD025485.

Thank you for bringing this reference to our attention. We modified the sentence at L26 of P3 to include that reference: “*Although the radar-only product extends over a longer time-period (for daytime only, see section 2.2), the CPR is less sensitive to fractionated and thin shallow cumulus clouds than the CALIPSO lidar and its ground clutter prevents cloud detection below 1 km, which preclude the detection of a large amount of marine low-level clouds (mostly Sc and Cu clouds, Liu et al., 2016).*”.

“CALIPSO” is mis-spelled in the caption of Figure 6:

This was corrected

The Cumulus And Stratocumulus CloudSat-CALIPSO Dataset (CASCCAD)

Grégory Cesana^{1,2}, Anthony D. Del Genio² and Hélène Chepfer^{3,4}

¹ Columbia University, Center for Climate Systems Research, Earth Institute, New York, NY, USA

² NASA Goddard Institute for Space Studies, New York, NY, USA

³ LMD/IPSL, Sorbonne Université, Paris, France

⁴ LMD/IPSL, CNRS, École Polytechnique, Palaiseau, France

Correspondence to: Gregory V. Cesana (Gregory.cesana@columbia.edu)

To submit to Earth System Science Data (copernicus open discussion)

Abstract.

Low clouds continue to contribute greatly to the uncertainty in cloud feedback estimates. Depending on whether a region is dominated by cumulus (Cu) or stratocumulus (Sc) clouds, the interannual low-cloud feedback is somewhat different in both space-borne and large eddy simulation studies. Therefore, simulating the correct amount and variation of the Cu and Sc cloud distributions could be crucial to predict future cloud feedbacks. Here we document spatial distributions and profiles of Sc and Cu clouds derived from Cloud-Aerosols Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) and CloudSat measurements. For this purpose, we create a new dataset called the Cumulus And Stratocumulus CloudSat-CALIPSO Dataset (CASCCAD), which identifies Sc, broken Sc, Cu under Sc, Cu with stratiform outflow and Cu. To separate the Cu from Sc, we design an original method based on the cloud height, horizontal extent, vertical variability and horizontal continuity, which is separately applied to both CALIPSO and combined CloudSat-CALIPSO observations. First, the choice of parameters used in the discrimination algorithm is investigated and validated in selected Cu, Sc and Sc-Cu transition case studies. Then, the global statistics are compared against those from existing passive and active-sensor satellite observations. Our results indicate that the cloud optical thickness –as used in passive-sensor observations– is not a sufficient parameter to discriminate Cu from Sc clouds, in agreement with previous literature. Using clustering-derived datasets show better results although one cannot completely separate cloud types with such an approach. On the contrary, classifying Cu and Sc clouds and the transition between them based on their geometrical shape and spatial heterogeneity leads to spatial distributions consistent with prior knowledge of these clouds, from ground-based, ship-based and field campaigns. Furthermore, we show that our method improves existing Sc-Cu classifications by using additional information on cloud height and vertical cloud fraction variation. Finally, the CASCCAD datasets provide a basis to evaluate shallow convection and stratocumulus clouds on a global scale in climate models and potentially improve our understanding of low-level cloud feedbacks. The CASCCAD dataset (Cesana,

Deleted: and

Deleted: and

Deleted: , while

Deleted: u

Deleted: .

Deleted:

Deleted: (Cu)

Deleted: boundary layer (Sc)

2019, DOI: <http://doi.org/10.5281/zenodo.2667637>) is available on the Goddard Institute for Space Studies (GISS) website at <https://data.giss.nasa.gov/clouds/casccad/> and on the zenodo website at <https://zenodo.org/record/2667637>.

Formatted: English (US)

1 Introduction

Whether clouds will amplify or dampen global warming, referred to as cloud feedbacks, continues to be a dominant source of uncertainty in future climate projections, for which low clouds over the tropics and at mid-latitudes contribute up to 50 % in recent generations of the Coupled Model Intercomparison Project (CMIP) models (Zelinka et al., 2016). Low clouds (i.e., cloud top height lower than ~ 3 km) may be separated into two main categories: stratocumulus (Sc, including stratus clouds) and cumulus (Cu). Driven by radiative cooling, Sc clouds cap the planetary boundary layer over cool oceans in conditions with a strong cloud top inversion, mostly off the western coasts of continents (e.g., Klein and Hartmann, 2002). They are typically hundreds of meters thick with a large horizontal extent, which can be either homogeneous (in decks) or heterogenous (open and closed cells), and a stable cloud-top height. Due to their large cloud cover, these clouds strongly reflect shortwave (SW) radiation and contribute to surface cooling. As the ocean warms up further west in the trade-wind regions, the latent heat flux increases and the convection becomes surface driven, therefore favoring breaking up of Sc and the subsequent formation of Cu clouds (Albrecht et al., 2019; Wyant et al., 1997). These clouds are horizontally limited and scattered – i.e., with a modest cloud cover – and their tops can rise above the PBL. Since these low clouds are governed by distinct processes, they may respond differently to climate warming (e.g., Bretherton, 2015) and there is no fundamental reason to expect the two cloud types to exhibit the same feedback.

Moved down [1]: Since these clouds are governed by distinct processes, they are likely to respond differently to climate warming (i.e., Bretherton, 2015).

Idealized large-eddy simulation (LES) studies have partly supported the hypothesis of different cloud type responses to climate warming, i.e., a substantial decrease of Sc as opposed to a moderate decrease to no change for Cu, yet, the underlying processes remain poorly understood, particularly for trade-wind Cu (Bretherton, 2015). Furthermore, recent satellite-based studies have shown that the interannual low-cloud feedback in response to SST forcings is somewhat different depending on whether a region is dominated by Cu or Sc clouds (Cesana et al., 2019a; McCoy et al., 2017). Therefore, simulating the right amount of Cu and Sc clouds could be crucial for models to reproduce the overall interannual low-cloud feedback as observed from space (Cesana et al., 2019a) and to predict future cloud feedbacks (Klein and Hall, 2015). Yet the relatively scarce amount of observations that fundamentally distinguish Sc and Cu clouds (mostly field campaigns and ground-based sites, i.e. Zhou et al., 2015; Rémillard et al., 2012) limits our ability to study the present-day global distribution of these clouds and their response to surface warming, and hence to better constrain the climate models.

Moved (insertion) [1]

Deleted: Since these clouds are governed by distinct processes, they are likely to respond differently to climate warming (i.e., Bretherton, 2015).

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Deleted: i.

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Estimating the global radiative impact of clouds on past, present and future climate continues to be a challenging question that requires observations of cloud macrophysical (e.g., height, spatial extent) and microphysical (e.g., phase, effective radius) properties on a global scale. Knowledge of the cloud type provides only a little leverage on determining their radiative properties, which may explain why cloud-type classification has received far less attention in the past. The first global-scale cloud-type observations were collected visually from land stations and ships in the 1950s and were subsequently compiled to make a coarse digital database in the late 1980s (Hahn et al., 1988) and were updated a couple of decades after (Hahn and Warren, 2007). A few years later, the first global-scale cloud type climatology derived from passive-sensor satellites emerged, based on cloud top pressure (CTP) and cloud optical thickness (COT) (Rossow and Schiffer, 1991, 1999). While very useful because of its long-time record, large spatial cover and finer resolution than Hahn et al. (1988), such datasets suffer from both methodology and instrumental limitations that make it difficult to fully discriminate Sc from Cu clouds. The CTP-COT method does not exploit information on the spatial shape of an individual cloud, that is to say its horizontal and vertical extent, and thus is not always accurate (Hahn et al., 2001; Pincus et al., 1999). Additionally, passive-sensor satellites do not describe the entire profile but only the uppermost layers, integrated over a height that is moreover hard to quantify with less confidence over land and in regions of strong inversions (Garay et al., 2008; Marchand et al., 2010). Thus, low cloud cover in regions of extensive higher clouds is underestimated. Finally, some instruments may not be well-suited for Cu cloud detection (Marchand et al., 2010).

By combining collocated observations from active CloudSat Cloud Profiling Radar (CPR, Stephens et al., 2002) and Cloud-Aerosols Lidar and Infrared Pathfinder Satellite Observations (CALIPSO; Winker et al., 2010) lidar and passive Moderate Resolution Imaging Spectroradiometer (MODIS; King et al., 2013) spectrometer, the 2B-CLDCLASS-LIDAR product (Sassen and Wang, 2008; Wang et al., 2013) addresses most of the above caveats. It classifies clouds into eight types based on cloud vertical and horizontal extent, reflectivity, precipitating state, temperature, height and brightness (Huang et al., 2015). However, this combined dataset is only available for a relatively short period of time (about 4.5 years), which influences statistical correlations between environment variables and cloud fraction (e.g., Cesana et al., 2019a; Klein et al., 2017). Although the radar-only product extends over a longer time-period (for daytime only, see section 2.2), the CPR is less sensitive to fractionated and thin shallow cumulus clouds than the CALIPSO lidar and its ground clutter prevents cloud detection below 1 km, which preclude the detection of a large amount of marine low-level clouds (mostly Sc and Cu clouds, Liu et al., 2016). Furthermore, the CPR horizontal resolution ($\sim 1.4 \text{ km} \times 1.7 \text{ km}$) is not ideal for shallow cumulus detection, which are typically smaller than 1 km (Rodts et al., 2003; Zhang and Klein, 2013). Therefore, creating a specific Sc-Cu cloud classification product based on CALIPSO observations only would allow one to overcome these issues although confining the analysis to regions with optically thin or no overlying high-clouds.

Here we propose to document spatial distributions and profiles of Sc, Cu and transitioning (i.e., broken Sc, Cu under Sc and Cu with stratiform outflow) clouds derived from CALIPSO measurements. To achieve this goal, we create an original method

Deleted: the

Field Code Changed

Field Code Changed

Deleted: includes a large portion of

Deleted: (

Deleted: 1

Deleted: and

based on the cloud height, horizontal extent, vertical variability and horizontal continuity, which can be applied separately to both CALIPSO and combined CloudSat-CALIPSO observations, referred to as the Cumulus And Stratocumulus CloudSat-CALIPSO Dataset (CASCAD, Cesana, 2019; DOI: <http://doi.org/10.5281/zenodo.2667637>). The datasets are presented in section 2. The sensitivity of the CASCAD algorithm is assessed in section 3 and the global-scale results are then discussed and compared to a subset of existing cloud-type datasets in section 4. Finally, section 5 summarizes the results.

2 Datasets

In this section, we describe the CALIPSO (GOCCP, section 2.1) and the combined CloudSat-CALIPSO (RL-GeoProf, section 2.2) observations that we use with the CASCAD algorithm and the only other active-sensor cloud-type product available (2B-CLDCLASS-LIDAR, section 2.3). The CASCAD algorithm (see full description in section 3) is designed so that it can be applied to active-sensor level 2 observations (i.e., cloud mask at the orbital level) at their native resolution to generate level 2 Sc-Cu cloud mask profiles, which consist of along-track profiles of Sc, Cu and transitioning (i.e., broken Sc, Cu under Sc and Cu with stratiform outflow) cloud types. Here we separately apply CASCAD to GOCCP and CloudSat-CALIPSO RL-GeoProf. The resulting GOCCP- or CloudSat-CALIPSO-CASCAD Cu-Sc profiles are then accumulated onto a 2.5°x2.5° grid (2D cloud fraction) and over 40 480-m levels (3D cloud fraction) for each month to create level 3 files (i.e., cloud fraction on the global scale). A detailed description of the content of both the GOCCP and RL-GeoProf CASCAD level 3 files – averaged over the whole period of time, 2007 to 2016 for GOCCP and 2007 to 2010 for RL-GeoProf – are available on the CASCAD website here: <https://data.giss.nasa.gov/clouds/cascad/>. To summarize, the distinct GOCCP and RL-GeoProf CASCAD datasets contain 2D and 3D cloud fraction of all-low, Cu, Sc and transitioning (i.e., broken Sc, Cu under Sc and Cu with stratiform outflow) cloud type. Finally, the GOCCP and RL-GeoProf level 2 files (i.e., along-track Cu-Sc mask at native resolution for each granule/orbit) are not yet available on the website (as of August 2019) but can be made available upon request.

2.1 GOCCP

The instantaneous cloud mask of the GCM-oriented CALIPSO Cloud Product (GOCCP) version 3.1.2 (Guzman et al., 2017) is used by our discrimination algorithm (referred to as DA, described in section 3.2) to derive Sc-Cu cloud fraction statistics from 2007 to 2016 over a 2.5° grid and for 40 levels with 480 m spacing from 0 to 19.2 km. GOCCP (Chepfer et al., 2010) was developed to facilitate the evaluation of cloud properties in GCMs when combined with a lidar simulator (Chepfer et al., 2008) that uses the same cloud definitions, and ensures a consistent comparison between observations and simulations (Cesana and Waliser, 2016). The ratio of the Total Attenuation Backscatter signal (ATB) to the molecular ATB – so-called Scattering Ratio (SR) – is computed from the level 1B CALIPSO files for every 333 m-along-track-resolution near-nadir lidar profile for 480 m height intervals. This lidar-based quantity is a proxy of the presence of particulate matter in a layer. GOCCP uses a fixed SR threshold to detect clouds (SR>5), for either daytime or nighttime data, regardless of the vertical level. This threshold

Deleted:

Deleted: and

Formatted: English (UK), Not Highlight

Formatted: English (UK), Not Highlight

Formatted: English (UK), Not Highlight

Formatted: English (UK), Not Highlight

Formatted: English (UK), Not Highlight

Formatted: English (UK), Not Highlight

Formatted: English (UK), Not Highlight

Formatted: English (UK), Not Highlight

Formatted: English (UK), Not Highlight

Deleted: separately

Formatted: English (UK), Not Highlight

Formatted: English (UK), Not Highlight

Formatted: English (UK), Not Highlight

Formatted: English (UK), Not Highlight

Formatted: English (UK), Not Highlight

Formatted: English (UK), Not Highlight

Field Code Changed

Formatted: English (UK), Not Highlight

Deleted: July

Formatted: English (UK), Not Highlight

Formatted: Normal

allows the detection of thin cirrus cloud in the high-troposphere (McGill et al., 2007), hence the majority – if not all – of optically thicker PBL clouds except when masked by overlying high-clouds (e.g., the trade-wind regions). It also prevents most false detections of aerosol layers as being cloudy in the PBL (Chepfer et al., 2013). GOCCP has been validated against in situ (Cesana et al., 2016) and ground-based observations (Lacour et al., 2017). Caveats for this dataset are discussed in

5 Cesana et al. (2016) and in Cesana and Waliser (2016). Compared to GOCCP version 2.9 (Cesana et al., 2016), version 3.1.2 improves the detection of fully attenuated pixels by introducing a surface echo detection. When no lidar echo is detected, pixels below the lowest cloudy pixel are diagnosed as being fully attenuated and therefore not accounted for in the cloud fraction computation. This new feature reduces the underestimation of the cloud fraction underneath optically thick liquid-topped clouds in the lower troposphere (Cesana et al., 2016).

10 2.2 CloudSat-CALIPSO RL-GeoProf

Additionally, the Sc-Cu DA is applied to combined CloudSat-CALIPSO profiles (using the radar-lidar geometrical profile product [RL-GeoProf], Mace and Zhang, 2014; version R04) from 2007 to 2010 over a 2.5° grid and for 80 levels with 240 m spacing from 0 to 19.2 km. From its launch to early April 2011, CloudSat flew approximately 15 s ahead of CALIPSO making

15 it possible to observe the same scene from a lidar-radar perspective when using both the 2B-GEOPROF and 2B-GEOPROF-LIDAR products (Mace and Zhang, 2014). CloudSat experienced a severe anomaly in April 2011, which forced the satellite to leave the A-Train constellation before coming back in June 2012 in “Daylight Only Operation mode” (DO-Op). As a result, CloudSat-CALIPSO combined observations are only available for 4.5 years. The CPR cloud mask provides confidence levels for the cloud detection (Marchand et al., 2009). Following Mace and Zhang (2014), we chose the confidence level of 20 and

20 higher to characterize the presence of a cloud from the radar cloud mask. The CPR has a coarser horizontal ($\sim 1.4 \text{ km} \times 1.7 \text{ km}$) and vertical resolution (240 m) than the CALIPSO lidar ($\Delta z = 30 \text{ m}$ below 8 km and 60 m above 8 km for $\sim 70 \text{ m}$ footprint every 333 m). As a result, several lidar profiles fall within the CloudSat radar footprint most of the time. These are used to compute the lidar cloud fraction (CF_{lidar} , from 0 to 1) based on the CALIPSO Vertical Feature Mask (VFM) of the level2 5km CALIPSO files (Vaughan et al., 2009), at the CPR resolution in the 2B-GEOPROF-LIDAR files. While the CF_{lidar} can

25 sometimes be lower than 0.5, we kept the 0.5 threshold to diagnose the presence of a cloud as in Mace and Zhang (2014) and Cesana et al. (2019b). Diagnosing a pixel as cloudy from values below 0.5 may result in an overestimate of the averaged cloud fraction when compared to ground-based measurements (Fig. S1 and Marchand et al., 2010).

2.3 2B-CLDCLASS-LIDAR

30 The 2B-CLDCLASS-LIDAR product [R05](#) (Sassen and Wang, 2008; [Wang et al., 2013](#); referred to as 2BCCL in the remainder of the manuscript) merges collocated observations from the CloudSat CPR, CALIPSO lidar and MODIS spectrometer to classify clouds into eight types based on several criteria: their vertical and horizontal extent, their precipitating state, their temperature, and their radiance. Although eight cloud types are available in this dataset (Deep convective, Cirrus, Nimbostratus, Altostratus, Altocumulus, Cumulus including fair-weather and congestus, Stratus and Stratocumulus), we only

Deleted: 1

Field Code Changed

focus on the Cu, stratus (St) and Sc cloud-types. The St and Sc cloud-types are combined into a single category referred to as Sc for consistency with the Sc-Cu discrimination algorithm, which does not differentiate these two categories. In addition, the Sc and St clouds are particularly difficult to distinguish in the 2BCCL product because of the ground clutter contamination in the radar signal (Sassen and Wang, 2008) as shown by Huang et al. (2015). In 2BCCL, the cloud base and top are given for up to ten cloudy layers, which is why we re-project these cloudy layers onto a 480 m vertical grid from 0 to 19.2 km to be consistent with the GOCCP and RL-GeoProf datasets described in sections 2.1 and 2.2. Those are then accumulated into a 2.5°x2.5° grid as for the two other products. Note that there are substantial differences between results using the R04 and R05 versions, which is why the reader should refer to the original manuscript published in May, 22nd 2019 on the ESSD discussion website for results using the older R04 version.

Deleted: versiond

10

3 Description of The Cumulus And Stratocumulus CloudSat-CALIPSO Dataset (CASCAD) discrimination algorithm

3.1 Why choose GOCCP and CloudSat-CALIPSO RL-GeoProf?

The main goal of this study is to document spatial distributions and profiles of Sc and Cu clouds on a global scale, with the desire to further analyze long-term relationships between Sc-Cu clouds and environmental parameters in future studies. For this purpose, we need to i) distinguish the two cloud types based on observable cloud-properties and ii) use datasets that are available for a time-period sufficiently long (~ 10 years) to compute statistically significant relationships (e.g., using 4 years of GOCCP rather than 10 may decrease the amplitude of the relationship between low clouds and SST anomalies by more than 15 %, Cesana et al., 2019a). Although both Sc and Cu clouds form within the planetary boundary layer (PBL), they have relatively different shapes as they are controlled by different physical mechanisms. However, one cannot separate clouds according to the mechanisms that form them as GCMs do using different PBL and convective parameterizations, which is why we choose to use the morphology to discriminate cloud types in this study. The Cu (Fig. 1, second to last column) can stretch up past the PBL into the lower free troposphere ($z \sim 3$ km) while they have typically a small horizontal extent (no more than a few km, e.g. Lamer et al., 2015; Nuijens et al., 2015b). On the contrary, the Sc (Fig. 1, first column) have a relatively small vertical extent (no more than a few hundred meters) and a cloud-top height (CTH) controlled by the PBL depth but spread out over tens to hundreds of kilometers (Wood, 2012) either homogeneously or heterogeneously (open-cell). In between, these two distinct regimes, various transitioning clouds may form (Albrecht et al., 2019; Teixeira et al., 2011) and the most frequent are: broken Sc and transition Sc-Cu, which is composed of Cu under Sc (Albrecht et al., 2019; Rauber et al., 2007) and Cu with stratiform outflow (Lamer et al., 2015; Nuijens et al., 2015b) (Fig. 1, second to fourth column, respectively). Clouds with a cloud base and a cloud top within and outside the lower free troposphere, respectively, are classified as deep Cu (Fig. 1, last column). Bearing the above facts in mind, we design the CASCAD DA based on cloud height, vertical and horizontal cloud fraction and horizontal continuity, which can be applied to both GOCCP instantaneous profiles and the CloudSat-CALIPSO level 2 geometrical profile product (referred to as RL-GeoProf).

Deleted: the

Formatted: English (UK)

Formatted: English (UK)

Formatted: English (UK)

Formatted: English (UK)

Deleted: ing

Deleted: and

GOCCP instantaneous profiles satisfy the criteria i) and ii) mentioned above. They use all 70-m-large lidar shots every 333 m along-track without horizontal averaging, which allows the detection of the geometrically sparsest shallow Cu, besides the more horizontally extended Sc, over a relatively extended time-period (from June 2006 to 2017, while CALIPSO is still operating as of April 2019). This decadal dataset makes it possible to analyze climatological values of Cu and Sc cloud fraction and their relationships to environmental parameters. However, as the lidar penetrates within cloudy layers, the signal eventually attenuates completely for optical thickness greater than 3 to 5. Therefore, it is not always possible to observe the full troposphere with a space-borne lidar, which may cause differences in satellite-based cloud climatologies obtained from different instruments (Kikuchi et al., 2017; Thorsen et al., 2013). In these instances –i.e., in deep convective clouds or in the storm tracks–, the CPR capability complements cloud profiles beneath the height at which the lidar attenuates, although the CPR clutter prevents using CloudSat data below ~ 1000 m. Unfortunately, the RL-GeoProf product is only available for a short period of time (~ 4.5 years) due to the severe anomaly of April 2011, which is why CloudSat-CALIPSO observations satisfy i) but only partially ii). Note that the release R05 of RL-GeoProf came out after the submission of the original manuscript (late May 2019), which includes data after April 2011. However, these are for daytime only and substantial periods of time are still missing (e.g., May 2011 through May 2012 and the whole year 2014), which makes it difficult to compute a consistent climatology and derive statistical relationships between clouds and environmental variables. Additionally, the differences between CASCAD using RL-GeoProf R04 and R05 are very small in the case studies of section 4.1 (see Fig. S2-S3-S4-S5). There is a small decrease of the overall cloud fraction (76.3 vs. 75.8 %, 37.2 vs. 35.2 %, 15.9 vs. 16 % and 80.8 vs. 77.2 % for R04 and R05 respectively), which affects mostly the Cu cloud fraction. Since the change is almost negligible, we decided not to update the global statistics with the R05 version for the sake of computational, space and time resources.

Deleted: amount

Deleted: ,

Deleted: t

3.2 First criterion of the discrimination algorithm: the cloud-top height

As mentioned above, four main criteria –represented by different colors in Fig. 2– are used in the DA to separate Sc from Cu clouds and to characterize the various transitioning clouds presented in Fig. 2: the CTH, the Horizontal Cloud Fraction (HCF), the Vertical profiles of Cloud Fraction (VCF) and the horizontal continuity test. Note that the sensitivity to these criteria is later tested in section 4.1. The first step of the DA depends on the height of the cloud top (Fig. 2, 1st column, in grey). Because trade Cu and Sc are low clouds, their CTH must be within the lower free troposphere, which is defined as 3.36 km in GOCCP (approximately equivalent to the 680-hPa definition of Rossow and Schiffer, 1999). Furthermore, since Sc clouds cap the PBL, their CTH is typically lower than the PBL height over the main Sc deck areas, i.e., ~ 2 km (Albrecht et al., 1995; Bretherton et al., 2010; Garay et al., 2008; Wood, 2012; Zhou et al., 2015; Zuidema et al., 2009). Therefore, all cloud layers –i.e., a vertically-contiguous group of cloudy 480-m-pixels– with a cloud top higher than 1.92 km, which is the closest 480-m GOCCP level to 2 km, are diagnosed as Cu type. We remind the reader that the sensitivity of the algorithm to this criterion, as well as the other criteria, is tested in section 4.1.

Deleted: three

Deleted: Sc-Cu

Deleted: and

Formatted: French

3.3 Second criterion: the horizontal cloud fraction

The remaining clouds are passed to the second step of the DA (Fig. 2, 2nd column, in orange), which computes the HCF either centered around the lidar profiles (CHCF) or using forward (FHCF) or backward profiles (BHCF) as shown in Fig. 3. These three HCFs ensure capturing the full horizontal extent of the cloud layer regardless of whether the lidar probes toward the edge or the center of the layer while remaining more computationally efficient than treating clouds by contiguous horizontal-clusters. For example, the CHCF is 100 % in the specific case of Fig. 3a whereas FHCF and BHCF are about 50 %. Should the first lidar profile be at the beginning of the cloud, the centered, forward and backward HCF would be about 53, 100 and 7 % (Fig. 3b). Additionally, the HCFs are computed over four different length scales, 10, 20, 40 and 80 km, to characterize various cloud scenarios: open-cell and closed-cell Sc (Wood, 2012), and different Cu organizations in Sc-Cu transition clouds (Albrecht et al., 2019; Lamer et al., 2015; Rauber et al., 2007; Teixeira et al., 2011). The larger 40-80-km scales permit a clear distinction between the two type of clouds since Sc clouds typically cover vast areas compared to more fractionated trade-Cu clouds. Figures 4c and 4d show the probability density function (PDF) of the 40-km and 80-km CHCFs, respectively, for typical Cu (light blue bars) and Sc (light red bars) cases extracted from day and night CALIPSO orbits over the tropics (35°S/N, eight orbit segments in total). These results confirm a rather clear separation, marked by purple lines, between the two populations for CHCFs. Although some slight overlap is visible, it disappears when the 40 and 80-km CHCFs are run together (not shown). However, in regions of open-cell Sc and Sc-Cu transition, the overlap may be larger (Fig. S6) and additional tests are needed to determine the type of clouds, i.e., Sc, broken Sc, Cu under Sc, Cu with stratiform outflow or Cu. In those instances, the 10 and 20-km CHCFs help further distinguish Sc from the other clouds (Fig. 4a and 4b). Finally, note that when all 480-m-pixels below 1.92 km are fully attenuated (4 pixels), the profile is excluded from the HCF computation –consistent with what is done in the GOCCP product for cloud fraction computations– because we do not know whether these pixels are clear or cloudy.

3.4 Third criterion: the vertical profiles of cloud fraction

The third step of the DA (which comes into play only if the first two criteria have ruled out a pure Cu cloud) utilizes the VCF to capture the Cu under Sc and broken Sc clouds (Fig. 2, 3rd column, in purple). In our study, the Vertical profile of Cloud Fraction (VCF, sometimes referred to as vertical distribution of cloud fraction or cloud fraction profiles in previous literature) corresponds to the cloud fraction computed over 80-km along track at a particular level, in other words, it represents how often a cloud is encountered over 80 km in the horizontal direction at a particular level. Since Sc clouds are relatively shallow – no more than a few hundred meters (Wood, 2012) – any cloud layer with a substantial VCF over 3 levels or more (vertical extension greater than 1.44 km) is corrected from Sc to Cu under Sc while the rest remain diagnosed as Sc (Fig. 5a). The VCF threshold (= 0.12) is defined as approximately two times the standard deviation of the PDF of the 40-km CHCF computed using each of the seven first levels (0 to 3.36 km) in typical Cu regions. Note that because the CloudSat-CALIPSO vertical cloud fraction is less affected by attenuation, we defined a larger VCF threshold (= 0.3) for the RL-GeoProf CASCAD

Deleted: ing

Field Code Changed

Deleted: (Albrecht et al., 2019; Teixeira et al., 2011) and different Cu organizations (Lamer et al., 2015; Rauber et al., 2007)

Deleted: 2

Deleted: transitioning Sc-Cu

Deleted: , computed 80-km along track,

Deleted: transitioning Sc-Cu

Deleted: s

Deleted: diagnosed

Deleted: as transitioning Sc-Cu

Deleted: are diagnosed

version. Finally, this VCF test is also applied to any cloud layer that passes the 40 and 80-km HCFs thresholds, regardless of their 10 and 20-km HCFs. In these cases, the cloud type is diagnosed as Cu under Sc if the VCF threshold is met, otherwise the cloud layer is diagnosed as broken Sc.

5 3.5 Fourth criterion: the horizontal continuity test

Once the cloud type is determined using the first three criteria, it is applied vertically to the whole cloudy layer. While the DA takes into account the horizontal extent of a cloud system via the computation of HCFs (including possible clear sky profiles), it does not track horizontally-contiguous clusters of clouds (without clear sky profiles). As a result, in Cu and transition Sc-Cu cases, the same horizontally-contiguous cloud layer may be diagnosed as both Sc-type (Sc and/or broken Sc) and Cu (Fig. 5b), which is more likely a Cu with a stratiform outflow (Lamer et al., 2015). To capture these particular clouds, we apply a horizontal continuity test (Fig. 2, 4th column, in green), which first detects a horizontally-contiguous cluster of clouds and then turns the non-Cu part (i.e., Sc, broken Sc or Cu under Sc) into a Cu with stratiform outflow if one third of the cluster is diagnosed as Cu type. We chose this arbitrary threshold because, on average, the fraction of the Cu that expands further aloft (geometrically thicker) is typically smaller than that near the lifting condensation level (Nuijens et al., 2015a) or that detrained near the trade-wind inversion (Nuijens et al., 2015a, Lamer et al., 2015).

4 Results

4.1 Case studies

To assess our CASCAD DA, we analyze a series of three typical case studies: trade cumulus, stratocumulus and stratocumulus-cumulus transitioning clouds. First, we investigate the sensitivity of the DA to some of the criteria presented in Section 3 using GOCCP observations: the HCF (more or less conservative), CTH (one level higher) and VCF (smaller threshold, divided by 2) thresholds and the horizontal continuity test (turned off). We then compare the results of the standard DA applied separately to GOCCP and RL-GeoProf against the 2BCCL cloud types –for the same case studies– and utilize the collocated MODIS reflectance to provide a broader context of the cloud scene.

25 4.1.1 Stratocumulus case

Figure 6 shows the sensitivity of the GOCCP Sc-Cu Mask (Fig. 6b) to the different parameters used in the DA (Fig. 6c-d-e-f-g) for a daytime orbit segment off the coast of California during summer (see Fig. S7 for the exact location). The refined CALIPSO Science-Team (CALIPSO-ST) cloud mask ($\Delta z = 30$ m; Vaughan et al., 2009) helps us get a better sense of the cloud geometrical thickness where the lidar is not fully attenuated (Fig. 6a, black color). Additionally, the MODIS true reflectance image confirms the presence of stratiform layers of clouds throughout the CALIPSO-CloudSat path (Fig. 7d). Except for the lower VCF parameters, which reduce the along-track Sc CF (HCF_{Sc}) by 5.6 % (absolute value, Fig. 6c), the HCF_{Sc} is quite insensitive to the DA parameters ($HCF_{Sc} = 69.4 \pm 0.3 / -1.3 \%$). Reducing the VCF (Fig. 6c) turns the edges of

Formatted: Not Highlight
Deleted: transitioning Sc-Cu

Deleted: H
Formatted: Not Highlight

Deleted: ing

Deleted: take
Deleted: To avoid this “slicing” issue
Deleted: it
Deleted: homogeneous

Deleted: d
Deleted: b-
Deleted: 3
Deleted: 6.4
Deleted: 6b
Deleted: 67
Deleted: 2.2
Deleted: 4
Deleted: 6b

the Sc decks into transitioning Sc - Cu clouds (around 16°N and 20°N). However, changing the HCF thresholds have limited effect on the HCF_{Sc} in this particular case (Fig. 6f and 6g).

GOCCP and RL-GeoProf have a similar HCF_{Sc} while that of the 2BCCL product is ~~smaller~~ along with its total HCF (Fig. 7). The substantial difference between RL-GeoProf and 2BCCL total HCFs is mostly due to differences in lidar cloud fraction treatment. The lidar cloud fraction from RL-GeoProf comes from the CALIPSO 5km VFM mask, whereas that of 2BCCL comes from the Lidar-AUX product (Wang et al., 2013).

4.1.2 Cumulus case

As for the Sc case, the along-track Cu CF (HCF_{Cu}) is weakly sensitive to variations of the different DA parameters (Fig. 8, $HCF_{Cu} = 23.6 \pm 1.1/-1.5\%$). The most sensitive parameter is the horizontal continuity (Fig. 8e). When activated, it ~~converts~~ most of the large Cu between 20°S and 18°S ~~from Sc to Cu with stratiform outflow~~ although its southernmost edge remains likely incorrectly diagnosed as Sc. Most of the other cloudy features are diagnosed as Cu except for the cloudy layer located between 31° and 29°S, which seems to be stratiform judging from its geometrical thickness (Fig. 8a) and reflectance (Fig. 9d). Here again, both GOCCP and RL-GeoProf diagnose similar Sc and Cu HCFs although the DA fully captures the aforementioned large Cu only when it is used with RL-GeoProf observations (Fig. 9b). Unlike the Sc case, GOCCP and RL-GeoProf disagree substantially with 2BCCL. For example, several cloud clusters are diagnosed as Sc by the 2BCCL algorithm although their geometrical thickness is larger than 1.5 km (Fig. 9c, around 34°S, 19°S and 12°S), making it very unlikely that these clusters are actual Sc. As a result, the HCF_{Sc} (~~17.2%~~) is ~~as large as~~ the HCF_{Cu} (~~18.3%~~) and approximately ~~three~~ times larger than that of GOCCP and RL-GeoProf. An additional Cu case, in the trade-dominated NW Atlantic, confirms the ability of the DA to correctly diagnose a field of purely Cu clouds with no Sc (Fig. 10a and 10b). ~~Contrary to~~ Fig. 9c, the 2BCCL product ~~does not detect Sc-type clouds~~ (Fig. 10c, $HCF_{Sc} = 0\%$) and ~~even underestimates the Cu cloud fraction~~ ($HCF_{Cu} = 10\%$) compared to GOCCP and RL-GeoProf ($HCF_{Cu} = 19.5\%$ and 15.9% , respectively). ~~This case also highlights one of two notable differences between the GOCCP and RL-GeoProf CASSCAD datasets: the latter misses a significant amount of cloud below 1 km altitude in scattered cumulus environments, presumably due to a combination of the radar surface clutter limitation and the lidar CF_{lidar} threshold used to diagnose a pixel as cloudy (CF_{lidar} = 0.5, cf section 2.2).~~

4.1.3 Cumulus and open-cell stratocumulus case

The last case study extends from the subtropics to the extra-tropics. Such location allows us to characterize transitioning Sc-Cu cases, which includes Sc, open-Sc and ~~the different~~ Cu-type clouds (Fig. 11). A visual inspection of the CTH variation from the CALIPSO VFM (Fig. 11a) suggests that this orbit segment contains three distinct clusters of clouds: Sc from 55° to 43°S and from 25° to 20°S and Cu in between. These three distinct layers are quite well captured by the DA although the DA is more sensitive to changes in the parameters than in the other cases. The most sensitive parameters are the VCF and HCF thresholds. Reducing the VCF threshold (Fig. ~~11c~~) turns ~~7.6%~~ of the Sc into Cu (absolute value) mostly poleward of 43°S because the BL height decreases, causing multiple levels to be cloudy and subsequently diagnosed as Cu. Choosing smaller

Deleted: larger

Deleted: captures

Deleted: 27

Deleted: .9

Deleted: larger

Deleted: than

Deleted: 22

Deleted: 6

Deleted: five

Deleted: In addition, it is important to note that the 2BCCL total HCF is ~35 % larger than that of RL-GeoProf for two reasons. Firstly, the 2BCCL algorithm is derived from a different version of the collocated lidar product (Wang et al., 2013). Secondly, no CF_{lidar} threshold is used, which may cause a slight overestimate of the cloud fraction, particularly in the low levels albeit to a smaller extent than the differences between the two lidar products. As a result, the low-level cloud fraction of the 2BCCL product may be largely overestimated over the tropical and subtropical oceans. This overestimation is supported by a comparison with ground-based data and other satellite products over the Barbados (more than 2 times larger, Fig. S1), a region dominated by trade cumulus clouds (Nuijens et al., 2015b).

Deleted: As in

Deleted: classifies a non-negligible amount of the clouds as Sc

Deleted: 6.8

Deleted: the

Deleted: overestimates the

Deleted: t

Deleted: of

Deleted:

Deleted: s

Deleted: T

Deleted:

Deleted: 11b

Deleted: 9

HCF thresholds (Fig. 11g) increases the Sc amount by 5.2% (absolute value). This converts the few Cu poleward 43°S into Sc as well as some Cu around 24°S, which could very well be “true” Sc.

The three clusters of clouds are also well captured in the RL-GeoProf dataset, which detects somewhat a little more Cu and Sc than GOCCP making the total HCF larger as well (Fig. 12), possibly due to the CF_{lidar} threshold, which may cause an overestimation of the cloud fraction. Also note that the additional Cu-type clouds detected by RL-GeoProf product are mostly classified as Cu with stratiform outflow. On the contrary, the 2BCCL product diagnoses nearly two times more Sc than the CASCAD datasets and two to three times less Cu while its total HCF is in between the two datasets.

4.2 Statistical analysis

4.2.1 Maps

In this section, we analyze climatological geographical distributions of Sc and Cu clouds for the three products presented before as well as for a subset of passive-sensor observations. For the CASCAD datasets, we treat transitioning clouds as distinct cloud types and thus do not include them in the Sc and Cu cloud fractions to be discussed below, although one could easily include the broken Sc in the Sc cloud fraction and the Cu under Sc as well as the Cu with stratiform outflow in the Cu cloud fraction. The passive-sensor observations include the International Satellite Cloud Climatology Project (ISCCP, Rossow and Schiffer, 1999), the Moderate Resolution Imaging Spectroradiometer (MODIS, King et al., 2013) and the Multi-Angle Imaging Spectroradiometer (MISR, Marchand et al., 2010) observations. In the passive remote sensing datasets, Sc and Cu are separated using a cloud top pressure (CTP) - cloud optical thickness (COT) diagram introduced by Rossow and Schiffer (1999): CTP must be larger than 680hPa for each type and COT smaller or larger than 3.6 for Cu and Sc clouds, respectively. There are advantages and disadvantages to each approach. In general, passive remote sensing instruments provide better temporal and spatial sampling than active sensors and the data products exist for a longer period of time. On the other hand, passive sensors can have difficulty identifying cloud top altitude, especially in multi-layer situations, relative to active sensors. For our purposes, we note that the COT-based method has been shown to mis-classify Cu and Sc that have moderate optical thickness (e.g., Pincus et al., 1999). In addition, Mace and Wrenn (2013) showed that except for thin cirrus and Sc cloud types, the COT-derived cloud types are mostly mixtures of different cloud types in two regions of the Eastern Pacific. However, passive-sensor estimates of Sc and Cu provide a broader context and help us emphasize the added value of new Sc-Cu discrimination methods based on active-sensor satellites.

Overall, all products identify quite well the large cloud fraction in the tropical and subtropical stratocumulus areas, off the west coast of the continents (Fig. 13, top row). RL-GeoProf and 2BCCL products detect the largest low-level cloud fraction (Fig. 14, zonal and global mean, top row) although they might overestimate the fractionated clouds (e.g., Cu). Unlike the passive-sensor products –based solely on the CTP and COT– the active-sensor products do not detect a significant amount of Cu off the west coasts of the continents compared to the large Sc cloud fraction, which ranges from 50 % off the coast of Australia up to 85 % in the heart of the deck off the coast of Peru in both GOCCP and RL-GeoProf (Fig. 13, third row). These results are somewhat different from a previous analysis in which the Sc cloud fraction ranges from 40 to 60 % over the Sc

Deleted: 3

Deleted: classifies open cells preferentially as Cu, while the GOCCP product more often identifies them as Sc-Cu transition.

Deleted: do not account for any

Deleted: se

Deleted: general

Deleted:

Deleted: Note that this

Deleted: (

Deleted: ,

Deleted: P

Deleted: , in particular 2BCCL –as shown in the case study analysis– because it uses a different version of the collocated lidar product...

Deleted: ,

Deleted: ern

deck areas (Wood, 2012). On the contrary, the CASCCAD products place a substantial amount of Cu clouds west of the Sc decks (up to 40 %) and in the trade-wind regions (between 20 and 30 %), similar to MISR observations (Fig. 13, third row), which is more sensitive to fractionated clouds than ISCCP and MODIS, while 2BCCL only classifies a small amount of Cu clouds (between 10 and 15 %) in these regions. Here again our findings somewhat contradict earlier results retrieving Sc clouds 20 % of the time in the trade-wind regions (Wood, 2012). Besides the Cu and Sc categories, the CASCCAD products have a third category referred to as transitional clouds (Fig. 13, fourth row), which is supposed to capture regions of transition between Cu and Sc clouds. As expected, these clouds are located between Sc decks and trade-wind regions and in the extra-tropics, where one could expect the two types of clouds to co-exist. However, they represent a small part of the total low-cloud fraction in the tropics while it is larger in the extra-tropics (between 10 and 20 %, Fig. 14, fourth row), where both Sc and Cu types have a similar and substantial cloud fraction (15 to 30 %). As mentioned in the cumulus and open-cell stratocumulus case, RL-GeoProf CASCCAD classifies more transitioning clouds than its GOCCP counterpart. This is mainly due to the fact that RL-GeoProf contains larger cloud clusters than GOCCP, making the continuity test more efficient (e.g., the large Cu cloud cluster around 19°S in Fig. 9 and the Cu part of the open-cell and Cu case, Fig. 12) and also because RL-GeoProf is less affected by overlapping mid- and high-cloud attenuation and better capture the full vertical extent of low clouds. Finally, the ratio of Sc clouds to Sc and Cu clouds (transitioning clouds being excluded) document the regions dominated by each type of clouds (Fig. 13, bottom row). Such information could be very useful to evaluate GCMs, which struggle to reproduce the Sc-Cu transition (e.g., Teixeira et al., 2011). The CASCCAD products robustly describe the tropical open ocean regions being almost exclusively dominated by Cu clouds while the oceans off the west coasts of the continents are mostly covered by Sc clouds. Such a picture is consistent with previous results from field campaigns, e.g., along the Global Energy and Water Experiment (GEWEX) Cloud System Studies (GCSS) Pacific Cross-Section Intercomparison (GPCI) transect (e.g., Zhou et al., 2015), or ship-based observations in the southeastern Pacific (Garay et al., 2008) and ground-based data, e.g., over Barbados (e.g., Nuijens et al., 2015). Finally, the CASCCAD products classify less Sc than the other products (Fig. 13 and 14, second row) in the extra-tropics and polar regions (poleward of 35°).

In the 2BCCL, MODIS and ISCCP products, most of the globe is dominated by Sc clouds, as opposed to the CASCCAD products and MISR, which depict the tropics dominated by Cu clouds and the extra-tropics by Sc clouds. The passive-sensor results suggest that the optical thickness does not permit a clear distinction between Cu and Sc clouds, in agreement with earlier studies (e.g., Pincus et al., 1999), although clustering analysis from passive-sensor observations may represent a better alternative (Tselioudis et al., 2013, their Fig. 3). To further investigate this statement, we classify GOCCP low clouds as a function of their opacity. The low-cloud containing profiles that are diagnosed as opaque (i.e., no surface echo retrieved) have an optical thickness larger than approximately 3 (Guzman et al., 2017). This optical thickness is further used as a threshold to separate Sc (opaque clouds with COT > 3) from Cu clouds (thin clouds with COT < 3), which is about the same optical thickness used in the passive sensor (i.e., COT = 3.6) to distinguish Cu and Sc clouds (i.e., COT = 3.6). As for the passive-sensor satellite observations, this method does not allow a clear separation between the two cloud populations (Fig. S8) although the ratio of Sc to Sc and Cu of derived from the two methods are well-correlated (~ 0.65). However, it confirms that

Deleted: Additionally,

Deleted:

Deleted: (Fig. 14, fourth row)

Deleted: and

Deleted: ern

Deleted:

Deleted: the

Deleted:), where the cloud feedbacks due to a change in low-cloud cover is somewhat less important –yet non-negligible– than in the tropics (Zelinka et al., 2016)

Deleted: Unsurprisingly,

Deleted: the transitioning Sc-Cu cloud fraction is also the largest in the extratropics where both types have a similar and substantial cloud fraction (30 to 40 %).

Deleted: all the passive satellite products but MISR

Deleted: with higher frequency of large ratio (> 50%, Fig. 14, bottom row)

Deleted: a more continuous distribution in

Deleted: 4

trade-wind regions have smaller opacity and therefore have a different radiative impact on surface and TOA fluxes than more opaque Sc-dominated regions.

The newer clustering approaches, such as the ISCCP weather states (ISCCP-WS, Tselioudis et al., 2013) and MODIS cloud regimes (MODIS-CR, Oreopoulos et al., 2014), do not discriminate cloud types or even low cloud from middle and high clouds. Instead, they represent mixtures of cloud types although one cloud type is often prevalent in a given cluster. As such, it is not straightforward to directly compare them with the CASCAD datasets. Yet, we believe that showing the results from the CR and WS approaches provides a broader context and additional information. For this reason, we analyze the ISCCP WS and MODIS CR observations against CASCAD datasets in Fig. 15. The ISCCP-WS cloud fractions are obtained by multiplying the monthly relative frequency of occurrence (RFO) of WS7-8 for Cu and WS9-10-11 for Sc by the monthly total cloud fraction, based on Tselioudis et al. (2013). A similar method is applied for MODIS-CR cloud fractions with CR11 for Cu and CR7-8-9-10 for Sc, based on Oreopoulos et al. (2014). The clustering-derived datasets show an overall better agreement with CALIPSO and CloudSat-CALIPSO CASCAD, particularly for the Sc regimes (Fig. 14 and Fig. 15). In the tropics, the Cu regimes are better correlated with CALIPSO CASCAD in terms of geographical distribution (from $r = 0.26$ to 0.63 for ISCCP and $r = -0.02$ to 0.35 for MODIS) although MODIS underestimates their fraction while ISCCP overestimates it (likely due to mid and high-level clouds contained in WS7 and WS8). Note that CR12 is currently investigated to be further decomposed in CR12a-b-c and some of these sub-CRs could correspond to Cu-type regimes (Oreopoulos et al., personal communication).

4.2.2 Profiles

Figure 16 shows global zonal profiles of cloud fraction for Cu, Sc, transitioning and all low-level clouds as observed by GOCCP, RL-GeoProf and 2BCL, for the first time. As in Section 4.2.1, the transitioning clouds are not accounted for in the Sc and Cu cloud fractions of the CASCAD datasets. Consistent with the map analysis, 2BCL observations retrieve clouds in the low-levels more frequently than GOCCP (Fig. 16, top row) while the difference with RL-GeoProf is rather small compared to that with GOCCP (approximately two times larger). The large difference between GOCCP and RL-GeoProf cloud fractions in the low levels mostly comes from mid and high-level topped clouds (e.g., frontal clouds in the extra-tropics and cumulonimbus and congestus clouds in the tropics, Fig. S9), which typically obscure CALIPSO vision by attenuating the lidar beam before it reaches the low levels. When separated into cloud types, GOCCP and RL-GeoProf observations agree quite well for Sc and transitioning clouds globally (both in terms of pattern and amount) and for Cu clouds in the deep tropics (15°S/N) and down to 2 km in the subtropics and the extra-tropics (Fig. 16, 2nd row). Between 2 km and 1 km, RL-GeoProf diagnoses more Cu than GOCCP. Such a difference is due to the different sensitivities of the lidar and radar instruments to clouds. The lidar signal becomes quickly attenuated by the optically and geometrically thick Cu—besides the attenuation from overlapping mid or high-clouds—whereas the CloudSat radar continues detecting clouds down to 1 km. Below 1 km the surface clutter and the lidar attenuation make it difficult to retrieve a reliable cloud fraction.

- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Deleted: exactly consistent
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Deleted: more recent
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Deleted: (
- Deleted: ,
- Deleted: (
- Deleted: ,
- Formatted: English (UK)
- Deleted: 5
- Deleted: case studies and
- Deleted: and RL-GeoProf
- Deleted: 5
- Deleted: although
- Deleted: 5
- Deleted: 5

Consistent with the case study and geographical analysis (Section 4.2.1), the 2BCCL product diagnoses far more (less) Sc (Cu) than the CASCAD products, making the ratio of Sc to Sc-and-Cu clouds largely dominated by Sc (Fig. 16, bottom row). Furthermore, the vertical distribution of Sc and Cu clouds substantially differs from that of the CASCAD products. 2BCCL Sc clouds may extend up to 3 km while the most part of Cu clouds are concentrated around 1 km and almost exclusively in the tropics. This lack of Cu and excess of Sc clouds above 1 km in tropical subsidence regimes (Fig. 16, right column) is in disagreement with the CASCAD products but also with previous studies focused on Sc regions (Cesana et al., 2019a, their Fig. 6) and Cu regions (Nuijens et al., 2015a, their Fig. 2, see also Fig. S1) (Nuijens et al., 2015a). Therefore, one may think that 2BCCL overestimates (underestimates) Sc (Cu) clouds by mis-diagnosing some Cu clouds into Sc clouds. On the contrary, the CASCAD products better match typical profiles of Sc and Cu clouds.

Finally, these global scale profiles are consistent with the physical processes controlling each type of cloud and measurements from previous literature. The Sc clouds –driven by radiative cooling– cap the PBL, a little higher than 1 km in the tropics and lower toward the poles, and their geometrical thickness is smaller than 1 km. Averaged over the tropical subsidence regimes (and in the extra-tropics, Fig. S10), the Sc vertical cloud fraction peaks between 10 and 12 %, depending on the dataset, whereas it is larger for Sc deck areas only (not shown). On the contrary, Cu cloud base –forced by surface fluxes– mostly form below 1 km, near the LCL, and vertically extend further aloft (around 2.5 km). Although the Cu map cloud fraction is smaller than that of Sc in the tropics (Fig. 13, compare the second and third rows), their vertical cloud fraction is about the same, around 8%, because they cover a larger domain (Fig. 16, right column). While the Sc vertical cloud fraction remains unchanged in the extra-tropics, its Cu and transitioning counterparts appear slightly larger, around 10 % and between 6 and 10% (Fig. S10), respectively.

5 Conclusion

In this paper, we document spatial distributions and profiles of stratocumulus (Sc), cumulus (Cu) and transitioning (i.e., broken Sc, Cu under Sc and Cu with stratiform outflow) clouds on a global scale. To this end, we design a discrimination algorithm (DA; Section 3) that distinguishes Sc and Cu based on four observable cloud-properties: cloud top height (CTH), horizontal cloud fraction (HCF), vertical cloud fraction variability (VCF) and horizontal continuity. These simple criteria are sufficient to characterize the distinctive shape of Cu, which have a limited horizontal extent and highly variable CTH as opposed to Sc, which cover larger areas and have a small and stable geometrical thickness. The DA is utilized separately on instantaneous profiles of active-sensor CALIPSO-GOCCP (Guzman et al., 2017) and CloudSat-CALIPSO combined observations (RL-GeoProf; Cesana et al., 2019b; Mace and Zhang, 2014) to create the distinct GOCCP and RL-GeoProf Cumulus And Stratocumulus CloudSat-Calipso Datasets (CASCAD).

The choice of DA parameters is then investigated in Cu, Sc and Sc-Cu transition case studies (Section 4.1), supported by additional viewings from MODIS true reflectance and full resolution CALIPSO VFM ($\Delta z = 30\text{m}$). The results show that the

Deleted: 5

Deleted: 5

Field Code Changed

Deleted: s

Deleted: 6

Deleted: 9

Deleted: between

Deleted: and 10

Deleted: 5

Deleted: s

Deleted: up

Deleted: to

Deleted: 15

Deleted: for RL-GeoProf

Deleted: 6

Deleted: and

Deleted: ¶

Deleted: three

Deleted: and

Deleted: ing

DA robustly captures Sc, Cu and Sc-Cu transition clouds although the choice of VCF and HCF thresholds may slightly affect the Sc-Cu partitioning in open-Sc and Cu regions.

The **distinct GOCCP and RL-GeoProf** CASCAD global-scale statistics (Section 4.2) are then compared to a subset of passive-sensor satellite datasets and to the only existing CloudSat-CALIPSO cloud-type climatology 2B-CLDCLASS-LIDAR (2BCCL, Sassen and Wang, 2008). In passive-sensor satellite observations, Sc and Cu co-exist everywhere and no region is fully dominated by a particular type of cloud (Fig. 12, bottom row) **when distinguished only based on their cloud optical thickness while Sc are mostly confined to specific regions (off the west coast of continents) with the more recent clustering approach**. On the contrary, Sc clouds largely dominate the global statistics from the 2BCCL point of view, which may misdiagnose a substantial portion of Cu clouds as Sc clouds in the trade-wind and extra-tropical regions. Interestingly, the CASCAD observations depict tropical oceans being almost exclusively dominated by Cu clouds (**around 20 % on average, and up to 40 %**) while the oceans off the west coasts of the continents are mostly covered by Sc clouds (50 to 85 %), with transitioning clouds in between (10 to 15 %). Our results provide a broader context to earlier findings from ground-based and field campaigns (Albrecht et al., 2019, 1995; Bretherton et al., 2010; Comstock et al., 2004; Garay et al., 2008; Wood, 2012; Zhou et al., 2015; Zuidema et al., 2009). For example, our globally-averaged profiles of Cu cloud fraction over the tropical oceans are almost identical to that found by Nuijens et al. (2015a) over the Barbados, in terms of shape (cloud base below 1 km and cloud top above 2 km) and frequency of occurrence (~ 10 %). Another interesting result concerns the distribution and magnitude of Sc cloud fraction. Our results indicate that the Sc clouds occur up to 85 % of the time over Sc deck areas compared to 60 % in earlier studies (i.e., Wood, 2012) and that their presence in trade-wind regions is negligible as opposed to a 20 % cloud frequency (i.e., Wood, 2012). Furthermore, our analysis indicates that the optical thickness, albeit useful, is not a sufficient parameter to discriminate Cu from Sc clouds, in agreement with previous literature (e.g., Pincus et al., 1999).

Finally, one of the reasons we developed CASCAD is to provide an improved observational constraint for low-level cloud feedbacks in GCMs. Although the CASCAD DA cannot be implemented in a lidar or radar simulator (Chepfer et al., 2008; Marchand et al., 2008), **it is still possible to use CASCAD datasets for model evaluation because i) both the convective and stratiform cloud fraction are provided as inputs to the lidar simulator and could be easily saved separately rather than summed up; and ii) a simulator is not necessarily needed for model-to-observation comparison of Cu and Sc clouds over the tropical oceans, because we identify the different cloud modes explicitly and we can select regimes in which lidar attenuation is negligible (e.g., $p_{500} > 0$ hPa/day, Cesana et al., 2019a). Therefore, the two CASCAD datasets make it possible to evaluate the shallow convection (Cu type) and boundary layer (Sc type) clouds in state-of-the art climate models, which are typically generated by distinct parametrizations (i.e., Cesana et al., 2019a), but also in unified-scheme turbulence models, in which Sc and Cu are still generated by distinct physical mechanisms (eddy diffusion for Sc and mass flux for Cu, e.g., Köhler et al., 2011). In such parameterizations, the existence of shallow convection is determined by a large-scale environmental index of inversion strength. In other unified turbulence schemes, the diagnosis of model success is achieved by defining Sc, Sc-Cu, and Cu regimes in terms of inversion strength** (Bogenschutz and Krueger, 2013). **We suggest that a dataset such as CASCAD that directly identifies these physically different cloud types is a better metric to use to judge model realism and fidelity.** By

- Deleted: ing
- Deleted: which distinguish Sc and Cu only based on their cloud optical thickness,
- Deleted: d
- Formatted: English (UK)
- Formatted: English (UK)
- Deleted: the
- Formatted: English (UK)
- Formatted: English (UK)
- Formatted: English (UK)
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto
- Deleted: by documenting the global geographical distribution of Sc and Cu clouds for the first time
- Formatted: Font: 10 pt, Font color: Auto
- Formatted: Font: 10 pt, Font color: Auto, Subscript
- Formatted: Font: 10 pt, Font color: Auto
- Deleted: (
- Deleted: C
- Deleted: (
- Deleted: Bogenschutz et al., 2013)
- Deleted: S

doing so, one could also assess the radiative contribution of Sc and Cu clouds to climate and potentially improve our understanding of low-level cloud feedbacks.

Data availability

5 The distinct GOCCP and RL-GeoProf CASCAD statistical datasets (Cesana, 2019) can be downloaded on the GISS website (<https://data.giss.nasa.gov/clouds/cascad/>) and on the zenodo website (<https://zenodo.org/record/2667637>; DOI: <http://doi.org/10.5281/zenodo.2667637>).

GOCCP instantaneous profiles used to produce the CASCAD dataset were downloaded from the CFMIP-Obs website (http://climserv.ipsl.polytechnique.fr/cfmip-obs/Calipso_goccp.html). The CloudSat-CALIPSO data used to produce the

10 CASCAD dataset (i.e., 2B-GEOPROF and 2B-GEOPROF-LIDAR) and the 2B-CLDCLASS-LIDAR product were obtained from the CloudSat Data Processing Center (<http://www.cloudsat.cira.colostate.edu/data-products/level-2b>).

The ISCCP WS data were obtained on the official ISCCP website (<https://isccp.giss.nasa.gov/wstates/gcluster.html>) and the matching total cloud fraction from the CFMIP-OBS website (<https://climserv.ipsl.polytechnique.fr/cfmip-obs/Isccp.html>). The MODIS CR and matching total cloud fraction were provided by Lazaros Oreopoulos and Nayeong Cho who can make them

15 available upon request by contacting them at Lazaros.Oreopoulos@nasa.gov.

Author contributions

GC designed the study, developed the products and carried out the analysis with inputs from AD and HC. GC wrote the manuscript with contributions from AD and HC.

20 **Competing interests**

The authors declare that they have no conflict of interest.

Acknowledgements

GC and AD were supported by a CloudSat-CALIPSO RTOP at the NASA Goddard Institute for Space Studies. The authors would like to thank NASA and CNES for giving access to CALIPSO and CloudSat observations, and Climserv for giving

25 access to CALIPSO-GOCCP observations and for providing computing resources. The authors also thank Lazaros Oreopoulos and George Tselioudis for their help in processing the MODIS CR and ISCCP WS datasets, respectively. GC thanks Zhien

Formatted: English (UK)

Formatted: Font color: Auto

Formatted: Hyperlink

Formatted: Font: Times New Roman, 10 pt, Font color: Auto

Formatted: Font: Times New Roman, 10 pt, Font color: Blue

Formatted: Font: Times New Roman, 10 pt, Font color: Auto

Formatted

Formatted: Justified, Line spacing: 1.5 lines

Deleted: ¶

Wang for his help in understanding the 2B-CLDCLASS-LIDAR dataset and Andy Ackerman, Ann Fridlind, Max Kelley and Greg Elsaesser for helpful comments and discussion.

References

- Albrecht, B., Ghate, V., Mohrmann, J., Wood, R., Zuidema, P., Bretherton, C., Schwartz, C., Eloranta, E., Glienke, S., Donaher, S., Sarkar, M., McGibbon, J., Nugent, A. D., Shaw, R. A., Fugal, J., Minnis, P., Paliknoda, R., Lussier, L., Jensen, J., Vivekanandan, J., Ellis, S., Tsai, P., Rilling, R., Haggerty, J., Campos, T., Stell, M., Reeves, M., Beaton, S., Allison, J., Stossmeister, G., Hall, S. and Schmidt, S.: Cloud system evolution in the trades (CSET) following the evolution of boundary layer cloud systems with the NSF-NCAR GV, *Bull. Am. Meteorol. Soc.*, 100(1), 93–121, doi:10.1175/BAMS-D-17-0180.1, 2019.
- 5 Albrecht, B. A., Bretherton, C. S., Johnson, D., Scubert, W. H. and Frisch, A. S.: The Atlantic Stratocumulus Transition Experiment—ASTEX, *Bull. Am. Meteorol. Soc.*, 76(6), 889–904, doi:10.1175/1520-0477(1995)076<0889:TASTE>2.0.CO;2, 1995.
- Bogenschutz, P. A. and Krueger, S. K.: A simplified PDF parameterization of subgrid-scale clouds and turbulence for cloud-resolving models, *J. Adv. Model. Earth Syst.*, 5(2), 195–211, doi:10.1002/jame.20018, 2013.
- 15 Bretherton, C. S.: Insights into low-latitude cloud feedbacks from high-resolution models, *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, 373(2054), doi:10.1098/rsta.2014.0415, 2015.
- Bretherton, C. S., Wood, R., George, R. C., Leon, D., Allen, G. and Zheng, X.: Southeast Pacific stratocumulus clouds, precipitation and boundary layer structure sampled along 20° S during VOCALS-REx, *Atmos. Chem. Phys.*, 10(21), 10639–10654, doi:10.5194/acp-10-10639-2010, 2010.
- 20 Cesana, G. and Waliser, D. E.: Characterizing and understanding systematic biases in the vertical structure of clouds in CMIP5/CFMIP2 models, *Geophys. Res. Lett.*, 43(19), 10,538–10,546, doi:10.1002/2016GL070515, 2016.
- Cesana, G., Chepfer, H., Winker, D., Getzewich, B., Cai, X., Jourdan, O., Mioche, G., Okamoto, H., Hagihara, Y., Noel, V. and Reverdy, M.: Using in situ airborne measurements to evaluate three cloud phase products derived from CALIPSO, *J. Geophys. Res.*, 121(10), 5788–5808, doi:10.1002/2015JD024334, 2016.
- 25 Cesana, G., Del Genio, A. D., Ackerman, A. S., Kelley, M., Elsaesser, G., Fridlind, A. M., Cheng, Y. and Yao, M.-S.: Evaluating models’ response of tropical low clouds to SST forcings using CALIPSO observations, *Atmos. Chem. Phys.*, 19(5), 2813–2832, doi:10.5194/acp-19-2813-2019, 2019a.
- Cesana, G., Waliser, D. E., Henderson, D., L’Ecuyer, T. S., Jiang, X. and Li, J.-L. F.: The Vertical Structure of Radiative Heating Rates: A Multimodel Evaluation Using A-Train Satellite Observations, *J. Clim.*, 32(5), 1573–1590, doi:10.1175/JCLI-D-17-0136.1, 2019b.
- 30 Chepfer, H., Bony, S., Winker, D., Chiriacco, M., Dufresne, J.-L. and Sèze, G.: Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model, *Geophys. Res. Lett.*, 35(15), L15704, doi:10.1029/2008GL034207, 2008.

Deleted: ¶

- Chepfer, H., Bony, S., Winker, D., Cesana, G., Dufresne, J. L., Minnis, P., Stubenrauch, C. J. and Zeng, S.: The GCM-Oriented CALIPSO Cloud Product (CALIPSO-GOCCP), *J. Geophys. Res.*, 115(5), D00H16, doi:10.1029/2009JD012251, 2010.
- Chepfer, H., Cesana, G., Winker, D., Getzewich, B., Vaughan, M. and Liu, Z.: Comparison of Two Different Cloud Climatologies Derived from CALIOP-Attenuated Backscattered Measurements (Level 1): The CALIPSO -ST and the CALIPSO -GOCCP, *J. Atmos. Ocean. Technol.*, 30(4), 725–744, doi:10.1175/JTECH-D-12-00057.1, 2013.
- Comstock, K. K., Wood, R., Yuter, S. E. and Bretherton, C. S.: Reflectivity and rain rate in and below drizzling stratocumulus, *Q. J. R. Meteorol. Soc.*, 130(603), 2891–2918, doi:10.1256/qj.03.187, 2004.
- Garay, M. J., Szoeké, S. P. De and Moroney, C. M.: Comparison of marine stratocumulus cloud top heights in the southeastern Pacific retrieved from satellites with coincident ship-based observations, , 113, 1–13, doi:10.1029/2008JD009975, 2008.
- Guzman, R., Chepfer, H., Noel, V., de Guélis, T. V., Kay, J. E., Raberanto, P., Cesana, G., Vaughan, M. A. and Winker, D. M.: Direct atmosphere opacity observations from CALIPSO provide new constraints on cloud-radiation interactions, *J. Geophys. Res.*, 122(2), 1066–1085, doi:10.1002/2016JD025946, 2017.
- Hahn, C. J. and Warren, S. G.: A GRIDDED CLIMATOLOGY OF CLOUDS, Analysis, doi:10.3334/CDIAC/cli.ndp026e, 2007.
- Hahn, C. J., Rossow, W. B. and Warren, S. G.: ISCCP Cloud Properties Associated with Standard Cloud Types Identified in Individual Surface Observations, *J. Clim.*, 14(1), 11–28, doi:10.1175/1520-0442(2001)014<0011:ICPAWS>2.0.CO;2, 2001.
- Huang, L., Jiang, J. H., Wang, Z., Su, H., Deng, M. and Massie, S.: Climatology of cloud water content associated with different cloud types observed by A-Train satellites, *J. Geophys. Res. Atmos.*, 120(9), 4196–4212, doi:10.1002/2014JD022779, 2015.
- King, M. D., Platnick, S., Menzel, W. P., Ackerman, S. A. and Hubanks, P. A.: Spatial and Temporal Distribution of Clouds Observed by MODIS Onboard the Terra and Aqua Satellites, *IEEE Trans. Geosci. Remote Sens.*, 51(7), 3826–3852, doi:10.1109/TGRS.2012.2227333, 2013.
- Klein, S. A. and Hall, A.: Emergent Constraints for Cloud Feedbacks, *Curr. Clim. Chang. Reports*, 1(4), 276–287, doi:10.1007/s40641-015-0027-1, 2015.
- Klein, S. A., Hall, A., Norris, J. R. and Pincus, R.: Low-Cloud Feedbacks from Cloud-Controlling Factors: A Review, *Surv. Geophys.*, 38(6), 1307–1329, doi:10.1007/s10712-017-9433-3, 2017.
- Köhler, M., Ahlgrimm, M. and Beljaars, A.: Unified treatment of dry convective and stratocumulus-topped boundary layers in the ECMWF model, *Q. J. R. Meteorol. Soc.*, 137(654), 43–57, doi:10.1002/qj.713, 2011.
- Lacour, A., Chepfer, H., Shupe, M. D., Miller, N. B., Noel, V., Kay, J., Turner, D. D. and Guzman, R.: Greenland Clouds Observed in CALIPSO -GOCCP: Comparison with Ground-Based Summit Observations, *J. Clim.*, 30(15), 6065–6083, doi:10.1175/JCLI-D-16-0552.1, 2017.
- Lamer, K., Kollias, P. and Nuijens, L.: Observations of the variability of shallow trade wind cumulus cloudiness and mass flux, *J. Geophys. Res. Atmos.*, 120(12), 6161–6178, doi:10.1002/2014JD022950, 2015.

- Liu, D., Liu, Q., Qi, L. and Fu, Y.: Oceanic single-layer warm clouds missed by the cloud profiling radar as inferred from MODIS and CALIOP measurements, *J. Geophys. Res.*, 121(21), 12,947–12,965, doi:10.1002/2016JD025485, 2016.
- Mace, G. G. and Wrenn, F. J.: Evaluation of the hydrometeor layers in the East and West Pacific within ISCCP cloud-top pressure-optical depth bins using merged CloudSat and CALIPSO data, *J. Clim.*, 26(23), 9429–9444, doi:10.1175/JCLI-D-12-00207.1, 2013.
- 5 Mace, G. G. and Zhang, Q.: The CloudSat radar-lidar geometrical profile product (RL-GeoProf): Updates, improvements, and selected results, *J. Geophys. Res. Atmos.*, 119(15), 9441–9462, doi:10.1002/2013JD021374, 2014.
- Marchand, R., Mace, G. G., Ackerman, T. and Stephens, G.: Hydrometeor detection using Cloudsat - An earth-orbiting 94-GHz cloud radar, *J. Atmos. Ocean. Technol.*, 25(4), 519–533, doi:10.1175/2007JTECHA1006.1, 2008.
- 10 Marchand, R., Haynes, J., Mace, G. G., Ackerman, T. and Stephens, G.: A comparison of simulated cloud radar output from the multiscale modeling framework global climate model with CloudSat cloud radar observations, *J. Geophys. Res.*, 114(8), D00A20, doi:10.1029/2008JD009790, 2009.
- Marchand, R., Ackerman, T., Smyth, M. and Rossow, W. B.: A review of cloud top height and optical depth histograms from MISR, ISCCP, and MODIS, *J. Geophys. Res. Atmos.*, 115(16), 1–25, doi:10.1029/2009JD013422, 2010.
- 15 McCoy, D. T., Eastman, R., Hartmann, D. L. and Wood, R.: The change in low cloud cover in a warmed climate inferred from AIRS, MODIS, and ERA-interim, *J. Clim.*, 30(10), 3609–3620, doi:10.1175/JCLI-D-15-0734.1, 2017.
- Mcgill, M. J., Vaughan, M. A., Trepte, C. R., Hart, W. D., Hlavka, D. L., Winker, D. M. and Kuehn, R.: Airborne validation of spatial properties measured by the CALIPSO lidar, , 112, 1–8, doi:10.1029/2007JD008768, 2007.
- Nuijens, L., Medeiros, B., Sandu, I. and Ahlgrimm, M.: Observed and modeled patterns of covariability between low-level cloudiness and the structure of the trade-wind layer, *J. Adv. Model. Earth Syst.*, 7(4), 1741–1764, doi:10.1002/2015MS000483, 2015a.
- 20 Nuijens, L., Medeiros, B., Sandu, I. and Ahlgrimm, M.: The behavior of trade-wind cloudiness in observations and models: The major cloud components and their variability, *J. Adv. Model. Earth Syst.*, 7(2), 600–616, doi:10.1002/2014MS000390, 2015b.
- 25 Oreopoulos, L., Cho, N., Lee, D., Kato, S. and Huffman, G. J.: An examination of the nature of global MODIS cloud regimes, *J. Geophys. Res. Atmos.*, 119(13), 8362–8383, doi:10.1002/2013JD021409, 2014.
- Pincus, R., McFarlane, S. A. and Klein, S. A.: Albedo bias and the horizontal variability of clouds in subtropical marine boundary layers: Observations from ships and satellites, *J. Geophys. Res. Atmos.*, 104(D6), 6183–6191, doi:10.1029/1998JD200125, 1999.
- 30 Rauber, R. M., Stevens, B., Ochs, H. T., Knight, C., Albrecht, B. a., Blyth, a. M., Fairall, C. W. and Jensen, J. B.: Over the ocean: The RICO campaign, *Bull. Am. Meteorol. Soc.*, (December 2007), 1912–1928, doi:10.1175/BAMS-88-12-1912, 2007.
- Rodts, S. M. A., Duynkerke, P. G. and Jonker, H. J. J.: Size Distributions and Dynamical Properties of Shallow Cumulus Clouds from Aircraft Observations and Satellite Data, *J. Atmos. Sci.*, 60(16), 1895–1912, doi:10.1175/1520-

- 0469(2003)060<1895:SDADPO>2.0.CO;2, 2003.
- Rossow, W. B. and Schiffer, R. A.: ISCCP Cloud Data Products, *Bull. Am. Meteorol. Soc.*, 72(1), 2–20, doi:10.1175/1520-0477(1991)072<0002:ICDP>2.0.CO;2, 1991.
- Rossow, W. B. and Schiffer, R. A.: Advances in Understanding Clouds from ISCCP, *Bull. Am. Meteorol. Soc.*, 80(11), 2261–2287, doi:10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2, 1999.
- 5 Sassen, K. and Wang, Z.: Classifying clouds around the globe with the CloudSat radar : 1-year of results, , 35(January), 1–5, doi:10.1029/2007GL032591, 2008.
- Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., Illingworth, A. J., O’connor, E. J., Rossow, W. B., Durden, S. L., Miller, S. D., Austin, R. T., Benedetti, A. and Mitrescu, C.: THE CLOUDSAT MISSION AND THE A-TRAIN, *Bull. Am. Meteorol. Soc.*, 83(12), 1771–1790, doi:10.1175/BAMS-83-12-1771, 2002.
- 10 Takahashi, N., Oki, R., Kikuchi, M., Cesana, G., Hayasaka, T., Sato, K., Hagihara, Y., Okamoto, H. and Suzuki, K.: Development of Algorithm for Discriminating Hydrometeor Particle Types With a Synergistic Use of CloudSat and CALIPSO, *J. Geophys. Res. Atmos.*, 122(20), 11,022–11,044, doi:10.1002/2017jd027113, 2017.
- Teixeira, J., Cardoso, S., Bonazzola, M., Cole, J., DelGenio, A., DeMott, C., Franklin, C., Hannay, C., Jakob, C., Jiao, Y., Karlsson, J., Kitagawa, H., Köhler, M., Kuwano-Yoshida, A., LeDrian, C., Li, J., Lock, A., Miller, M. J., Marquet, P., Martins, J., Mechoso, C. R., Meijgaard, E. v., Meinke, I., Miranda, P. M. A., Mironov, D., Neggers, R., Pan, H. L., Randall, D. A., Rasch, P. J., Rockel, B., Rossow, W. B., Ritter, B., Siebesma, A. P., Soares, P. M. M., Turk, F. J., Vaillancourt, P. A., Von Engeln, A. and Zhao, M.: Tropical and Subtropical Cloud Transitions in Weather and Climate Prediction Models: The GCSS/WGNE Pacific Cross-Section Intercomparison (GPCI), *J. Clim.*, 24(20), 5223–5256, doi:10.1175/2011JCLI3672.1, 2011.
- 20 Thorsen, T. J., Fu, Q. and Comstock, J. M.: Cloud effects on radiative heating rate profiles over Darwin using ARM and A-train radar/lidar observations, *J. Geophys. Res. Atmos.*, 118(11), 5637–5654, doi:10.1002/jgrd.50476, 2013.
- Tselioudis, G., Rossow, W., Zhang, Y. and Konsta, D.: Global weather states and their properties from passive and active satellite cloud retrievals, *J. Clim.*, 26(19), 7734–7746, doi:10.1175/JCLI-D-13-00024.1, 2013.
- 25 Vaughan, M. A., Powell, K. A., Winker, D. M., Hostetler, C. A., Kuehn, R. E., Hunt, W. H., Getzewich, B. J., Young, S. A., Liu, Z. and McGill, M. J.: Fully Automated Detection of Cloud and Aerosol Layers in the CALIPSO Lidar Measurements, *J. Atmos. Ocean. Technol.*, 26(10), 2034–2050, doi:10.1175/2009JTECHA1228.1, 2009.
- Wang, Z., Vane, D., Stephens, G. and Reinke, D.: Level 2 Combined Radar and Lidar Cloud Scenario Classification Product Process Description and Interface Control document, CloudSat Proj. Jet Propuls. Lab. Pasadena ..., (D), 2013.
- 30 Winker, D. M., Pelon, J., Coakley, J. A., Ackerman, S. A., Charlson, R. J., Colarco, P. R., Flamant, P., Fu, Q., Hoff, R. M., Kittaka, C., Kubar, T. L., Le Treut, H., McCormick, M. P., Mégie, G., Poole, L., Powell, K., Trepte, C., Vaughan, M. A. and Wielicki, B. A.: The CALIPSO Mission, *Bull. Am. Meteorol. Soc.*, 91(9), 1211–1230, doi:10.1175/2010BAMS3009.1, 2010.
- Wood, R.: Stratocumulus Clouds, *Mon. Weather Rev.*, 140(8), 2373–2423, doi:10.1175/MWR-D-11-00121.1, 2012.

- Wyant, M. C., Bretherton, C. S., Rand, H. A. and Stevens, D. E.: Numerical Simulations and a Conceptual Model of the Stratocumulus to Trade Cumulus Transition, *J. Atmos. Sci.*, 54(1), 168–192, doi:10.1175/1520-0469(1997)054<0168:NSAACM>2.0.CO;2, 1997.
- Zelinka, M. D., Zhou, C. and Klein, S. A.: Insights from a refined decomposition of cloud feedbacks, *Geophys. Res. Lett.*, 5 43(17), 9259–9269, doi:10.1002/2016GL069917, 2016.
- Zhang, Y. and Klein, S. A.: Factors Controlling the Vertical Extent of Fair-Weather Shallow Cumulus Clouds over Land: Investigation of Diurnal-Cycle Observations Collected at the ARM Southern Great Plains Site, *J. Atmos. Sci.*, 70(4), 1297–1315, doi:10.1175/JAS-D-12-0131.1, 2013.
- Zhou, X., Kollias, P. and Lewis, E. R.: Clouds, precipitation, and marine boundary layer structure during the MAGIC field 10 campaign, *J. Clim.*, 28(6), 2420–2442, doi:10.1175/JCLI-D-14-00320.1, 2015.
- Zuidema, P., Painemal, D., De Szoeki, S. and Fairall, C.: Stratocumulus cloud-top height estimates and their climatic implications, *J. Clim.*, 22(17), 4652–4666, doi:10.1175/2009JCLI2708.1, 2009.

Figures

Figure 1: Cartoon representing the different cloud morphologies used in this study. The blue dashed line denotes the upper limit of the low-level clouds following GOCCP definition (3.36 km). Note that the Deep Cu type is not a low cloud, but either a mid- or high-level cloud.

5

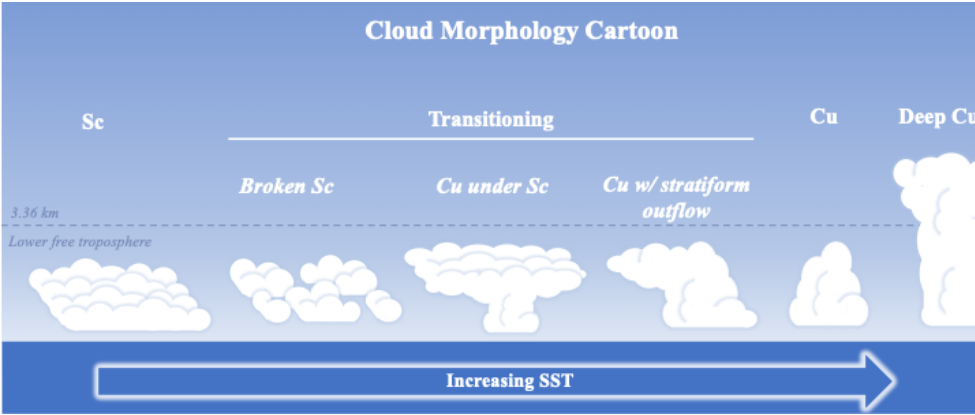
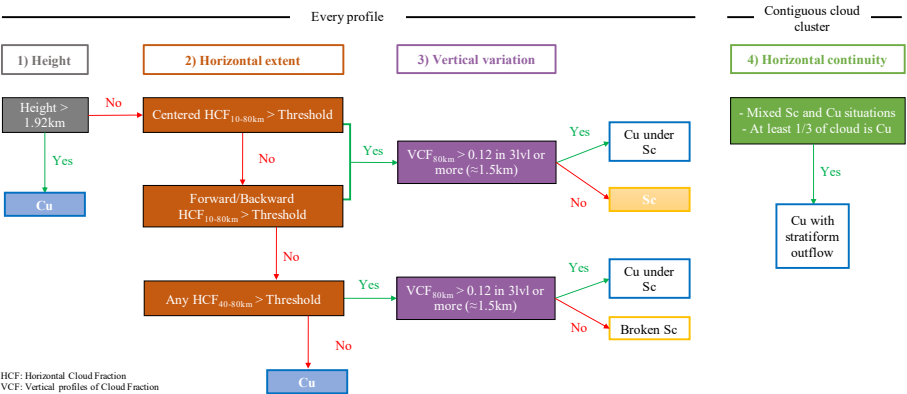


Figure 2: Flow diagram of the algorithm that performs the Sc-Cu discrimination. The grey, orange, purple and green colors represent the different steps of the discrimination algorithm to separate Sc, broken Sc, Cu, Cu under Sc and Cu with stratiform outflow clouds: height, horizontal extent, vertical extent and horizontal continuity, respectively. The choice of HCF thresholds is discussed in section 3.3 and their values are shown in Fig. 4. The VCF threshold (0.12) and the horizontal continuity test are also explained in section 3.4 and further used in Fig. 5. Note that the steps 1 to 3 are applied to every profile while the step 4 is only applied to contiguous cloud clusters.

Flow Diagram of the Cu – Sc discrimination



Deleted: and

Deleted: and Sc-Cu transitioning

Deleted: and

Deleted: is

Figure 3: Cartoon of the horizontal along-track averaging that is used by the DA to compute the horizontal cloud fraction (HCF): a) the lidar beam is in the middle of the cloud and b) the lidar beam is at the edge of the cloud. The dark-green circles correspond to the profiles used in the different horizontal averaging scenarios (centered, forward or backward).



Figure 4: PDF of the different along-track CHCFs for Cu (blue) and Sc (light red) typical regions computed from eight orbit segments. The thresholds are represented in purple, at approximately the overlap of the two distributions.

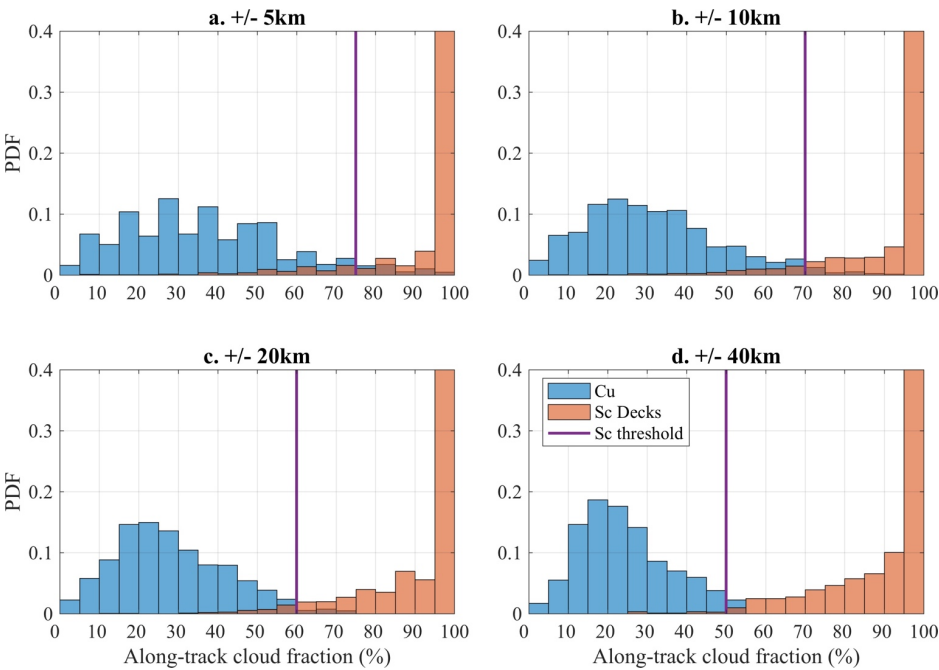
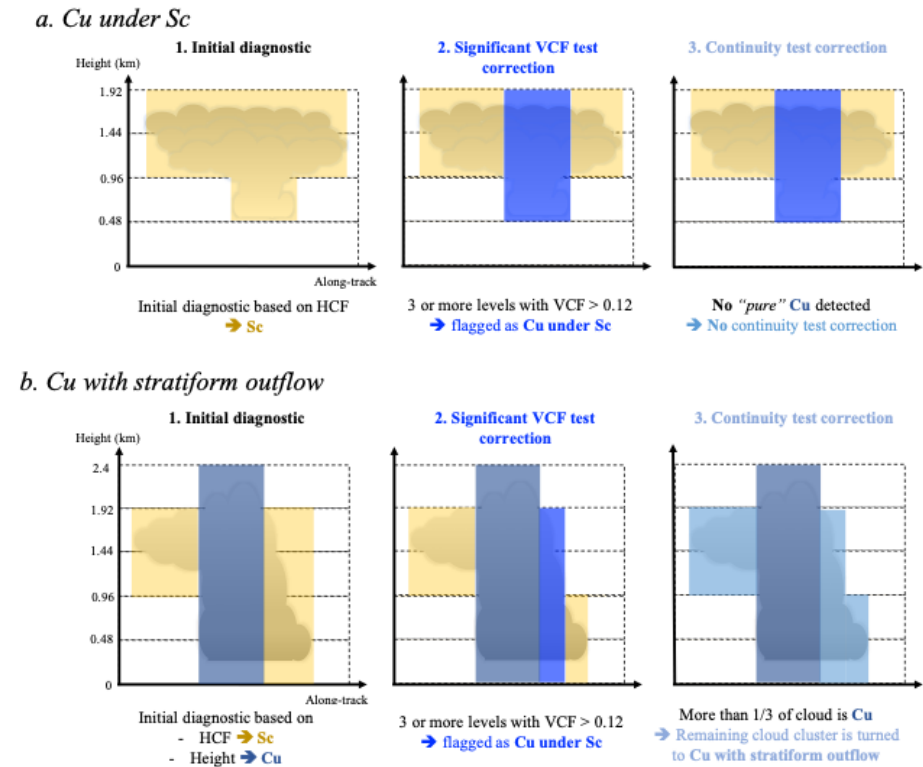


Figure 5: Cartoon of the **Cu and Sc mixed situations using** additional diagnostics in the DA: a) **Cu under Sc** and b) **Cu with stratiform outflow**. After the DA diagnoses (1) an initial cloud type, the additional (2) “significant VCF variability” test and (3) “horizontal continuity” correction are applied **to contiguous cloud clusters** (See details in section 3.4 and 3.5).



Deleted: a

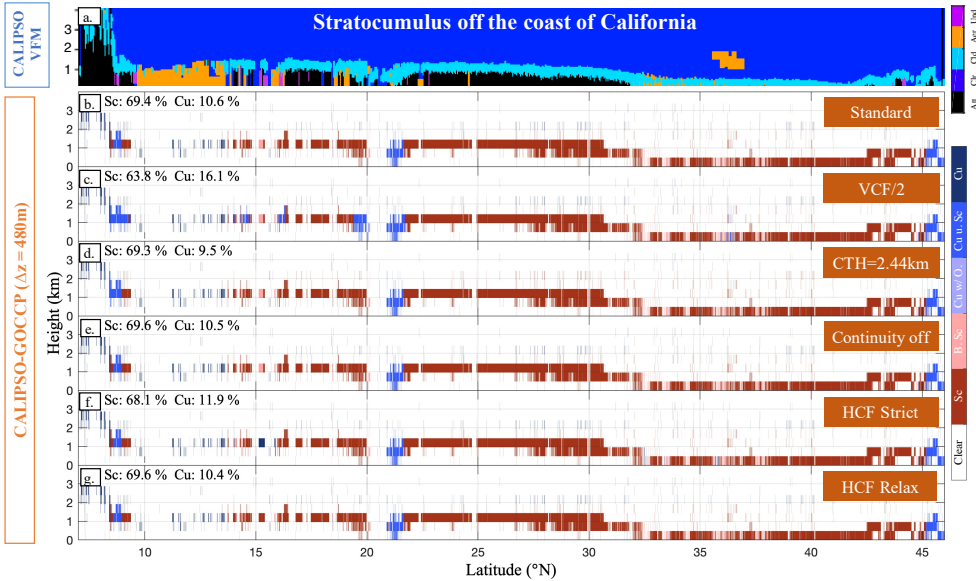
Deleted: b

Deleted: c

Deleted: test

Formatted: English (UK)

Figure 6: Daytime orbit segment off the coast of California (~ 8°N to 46°N, 2008-07-08 21:47:30) showing a typical stratocumulus case. (a) VFM CALIPSO-ST at 30-m vertical resolution, (b) to (g) sensitivity of the GOCCP Sc-Cu mask to different DA parameters: (b) the VCF threshold is divided by 2, (c) the CTH threshold is elevated from 1.92 km to 2.4 km, (d) standard DA parameters, (e) the continuity test is turned off, (f) the 10-km and 20-km HCF thresholds are more conservative (i.e., increased by 0.1, absolute value) and (g) the 10-km and 20-km HCF thresholds are less conservative i.e., (decreased by 0.1, absolute value). The HCF of Sc-dominated (including broken Sc) and Cu-dominated (including Cu under Sc and Cu with stratiform outflow) clouds are given in each subplot (top left corner). Reddish and bluish pixels correspond to Sc-dominated and Cu-dominated type of clouds, respectively. From the top to the bottom, the VFM color bar's labels correspond to undetermined, aerosol, cloudy, clear and fully attenuated pixels. From the top to the bottom, the GOCCP Sc-Cu mask color bar's labels correspond to cumulus, Cu under Sc, Cu with stratiform outflow, broken Sc, stratocumulus, and clear pixels.



Deleted: S

Deleted: Sc-Cu transitioning

Deleted: Sc-Cu transitioning

Deleted: edge of Sc,

Figure 7: Same stratocumulus case study as in Fig. 5 but for (a) the standard GOCCP Sc-Cu mask (same as Fig. 5b), (b) the RL-GeoProf Sc-Cu mask, (c) 2BCCL cloud type mask and (d) MODIS true reflectance. The red line corresponds to the CALIPSO-CloudSat overpass.

Deleted: correspond

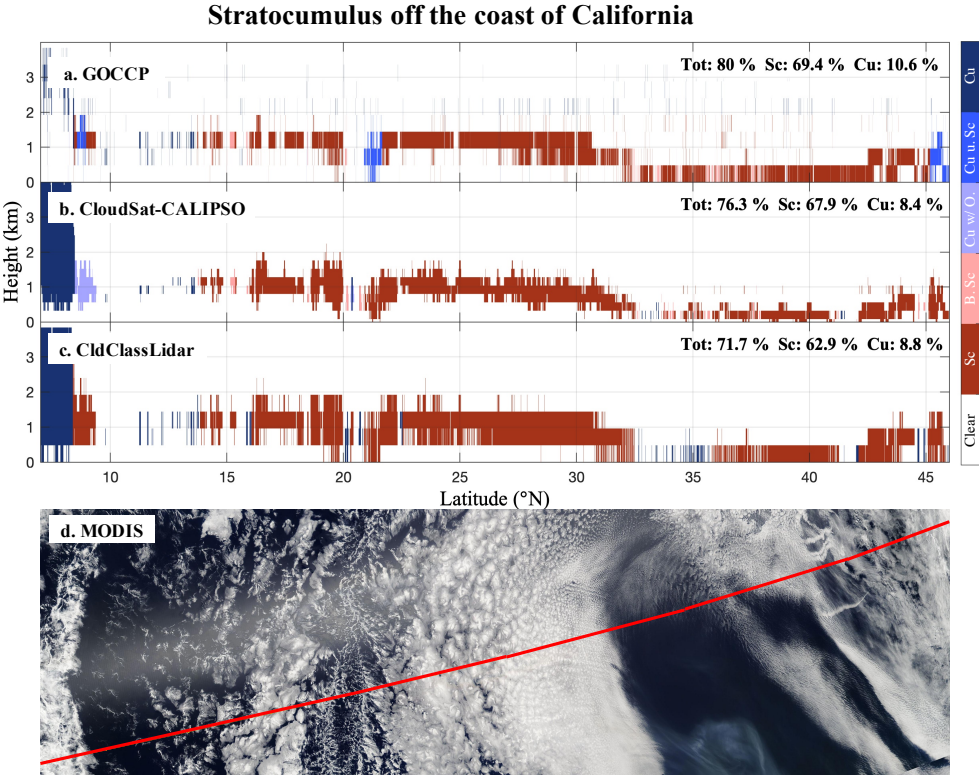


Figure 9: Same as Fig. 6 but for a typical cumulus case in the south-east Pacific (~ 37°S to 8°S, 2008-07-08 21:47:30, daytime).

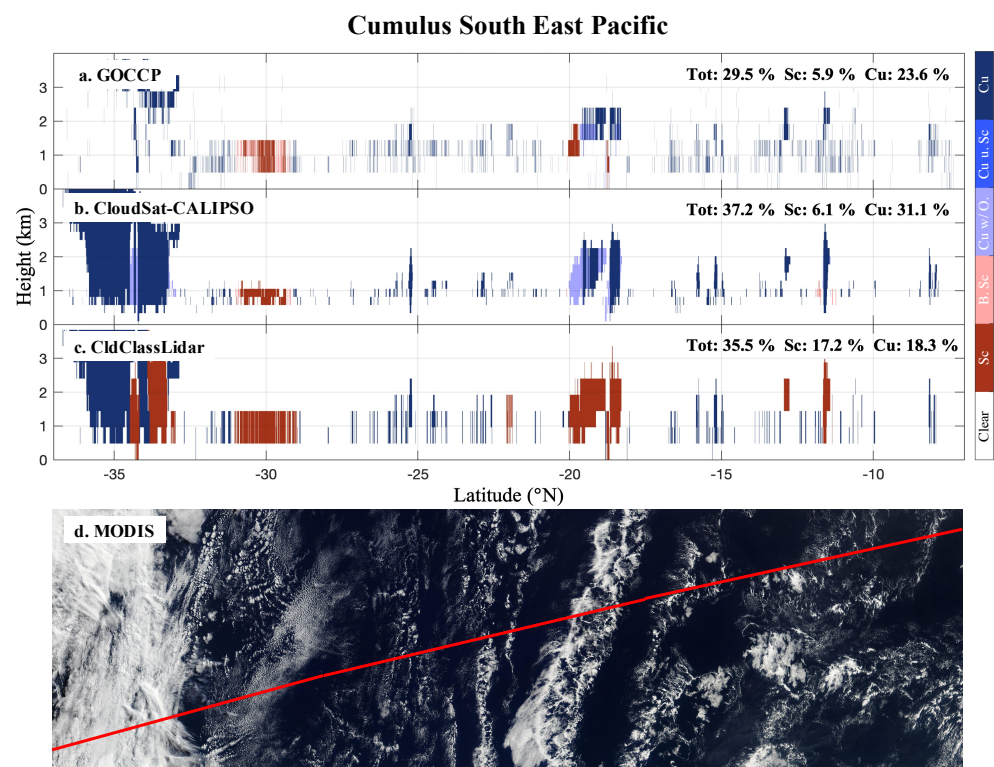


Figure 10: Same as Fig. 8 but for another cumulus case overpassing the Barbados (~ 10°N to 35°N, 2008-07-09 17:34:08 daytime)

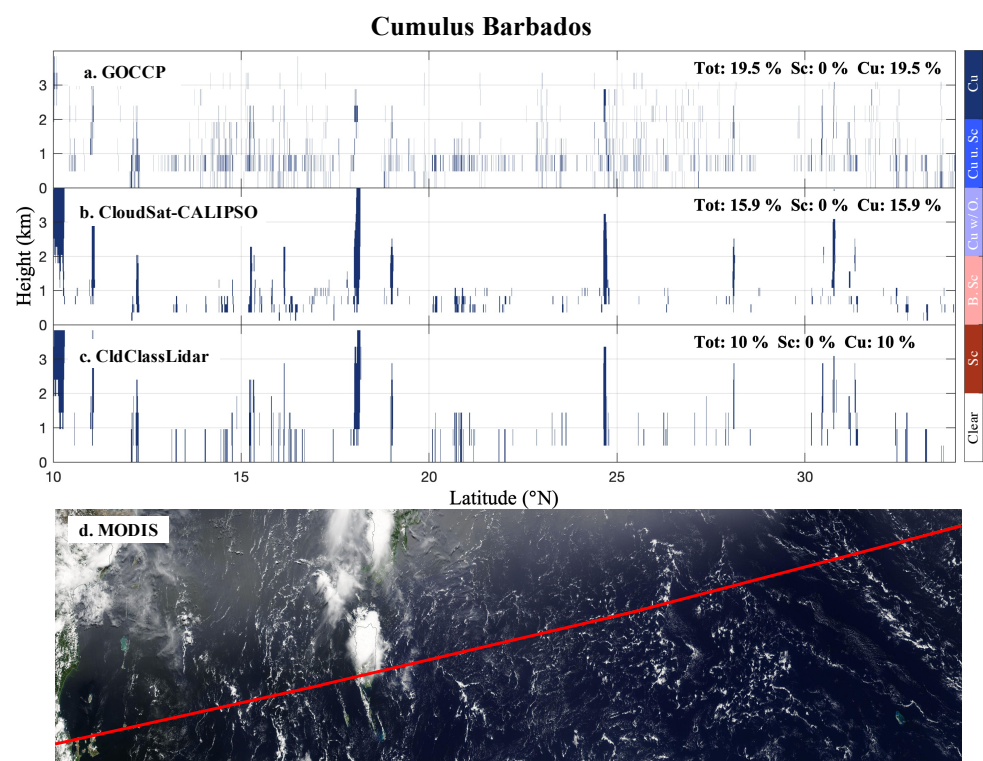


Figure 11: Same as Fig. 5 but for a typical Open-cell Se to Cu transitioning case overlapping the south-east Pacific and the Southern Ocean (~ 55°S to 20°S, 2008-07-14 21:10:19, daytime).

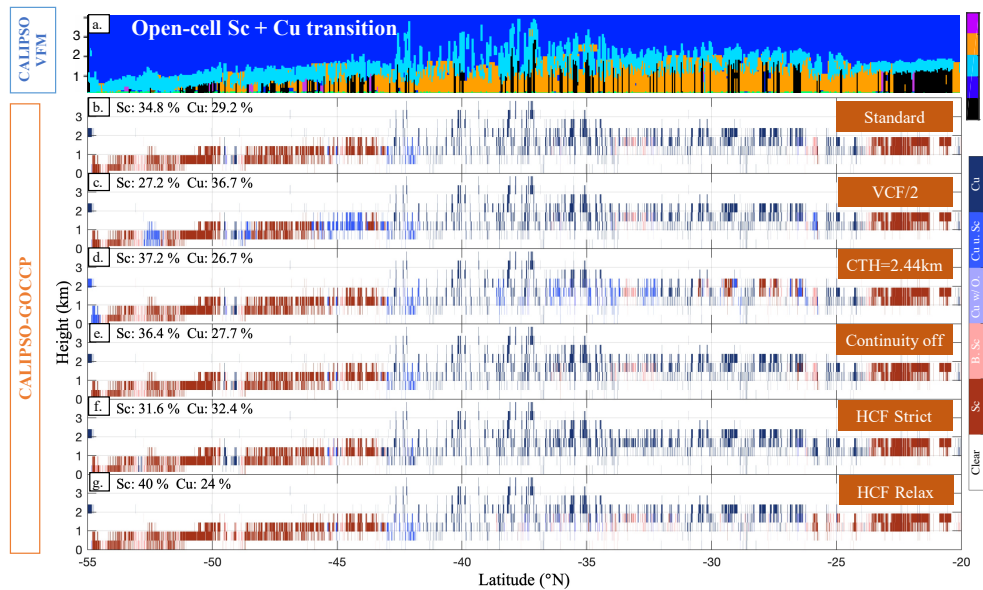


Figure 12: Same as Fig. 6 but for a typical Open-cell Sc to Cu transitioning case overlapping the south-east Pacific and the Southern Ocean (~ 55°S to 20°S, 2008-07-14 21:10:19, daytime).

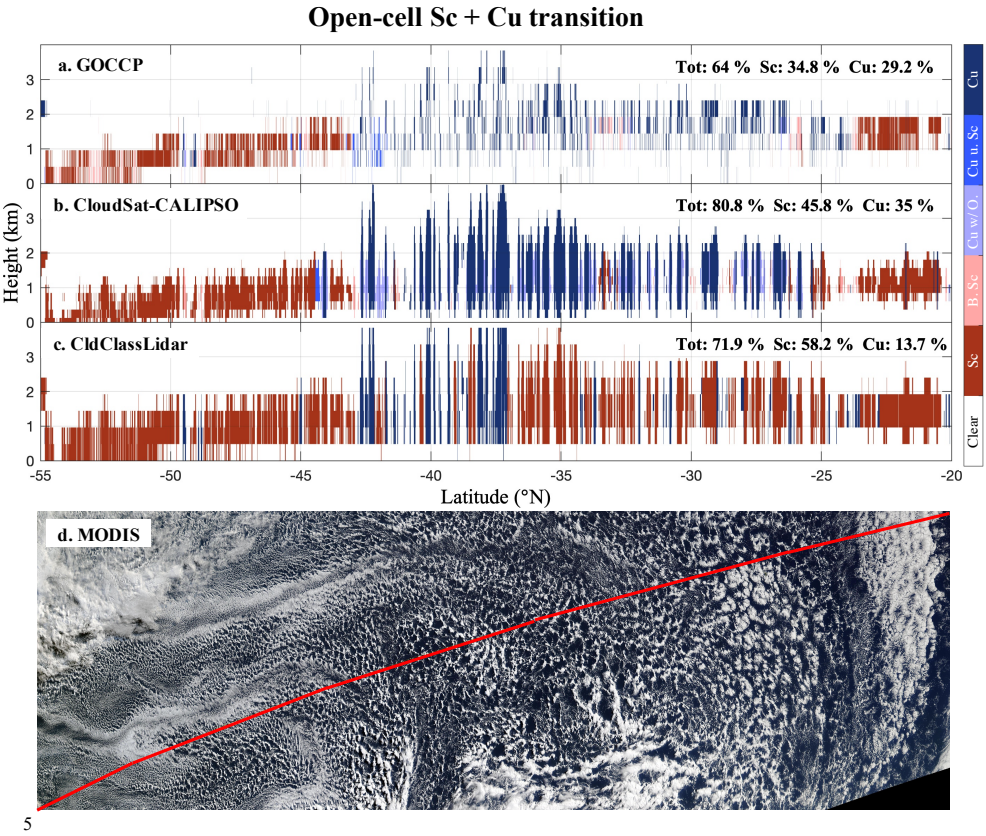


Figure 13: Maps (x axis, longitude [$^{\circ}$ E]; y axis, latitude [$^{\circ}$ N]) of (top to bottom) low, Sc, Cu, transitioning (i.e., broken Sc, Cu under Sc and Cu with stratiform outflow) and the ratio Sc to Sc-and-Cu cloud fraction (%) for (left to right) GOCCP (2007-2016), RL-GeoProf (2007-2010), 2BCCL (2007-2010), MISR (2003-2012), MODIS Terra and Aqua (2003-2015) and ISCCP (1983-2008). Different color bars are used to better separate each type of cloud. Note that for active-sensor satellites, the low category accounts for all clouds present below 3.36 km regardless of their cloud top height, hence the sum of Sc, Cu and transitioning cloud fraction can be smaller than the low cloud fraction. All datasets are averaged onto a $2.5^{\circ} \times 2.5^{\circ}$ grid.

Deleted: Sc-Cu

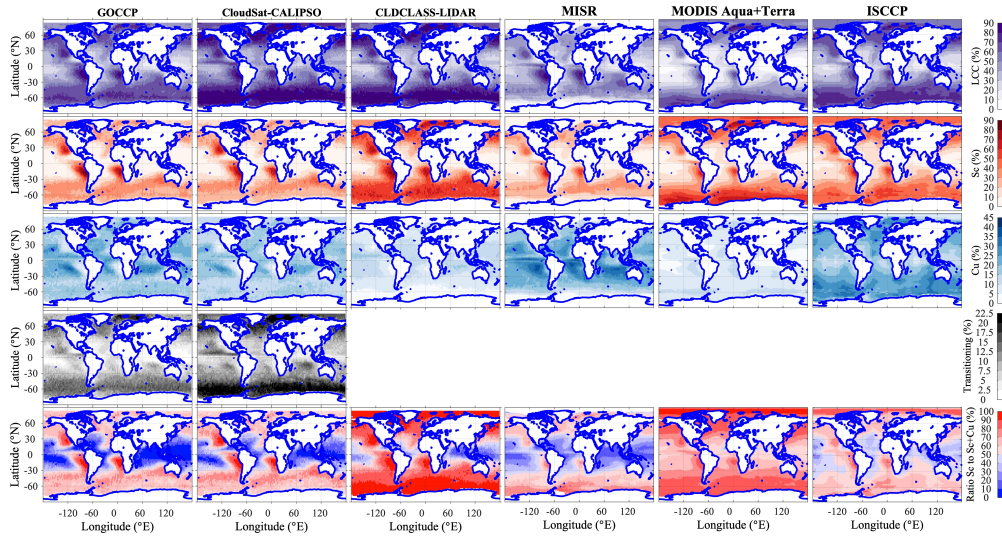
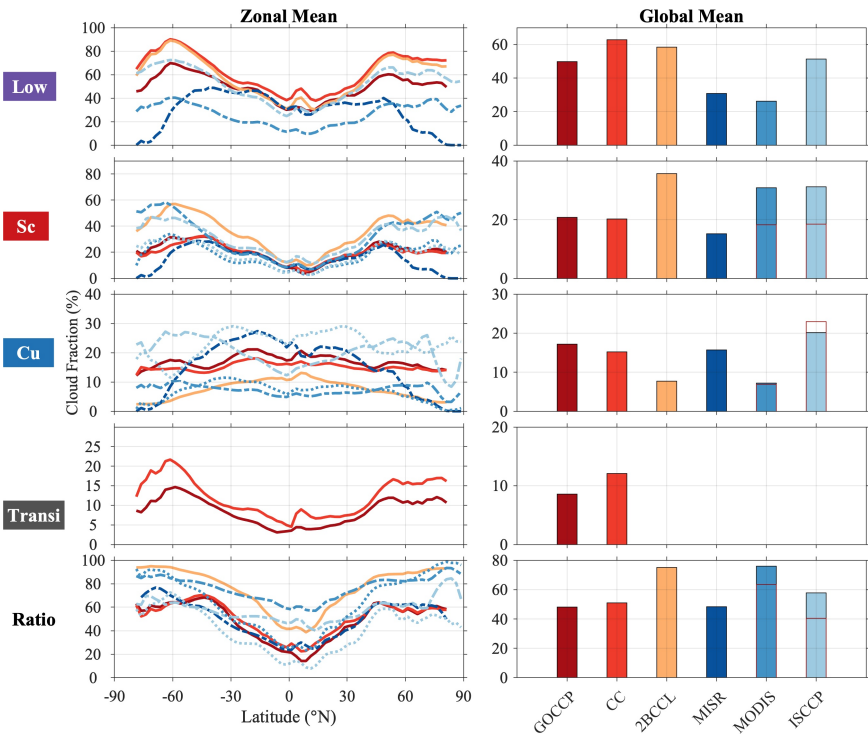


Figure 14: Global zonal mean (% left) and global area-weighted mean (% right) of low, Sc, Cu, transitioning (i.e., broken Sc, Cu under Sc and Cu with stratiform outflow) and the ratio Sc to Sc-Cu clouds –from the top to the bottom– for all the satellite datasets using the same time period as in Fig. 13. Active-sensor observations are represented in solid lines as opposed to dashed lines for COT-CTP passive-sensor observations and dotted lines for clustering passive-sensor observations. The clustering area-weighted means are represented by transparent bars with red borders. Note that the y-axis are different in every subplots. All datasets are averaged onto a 2.5°x2.5° grid.

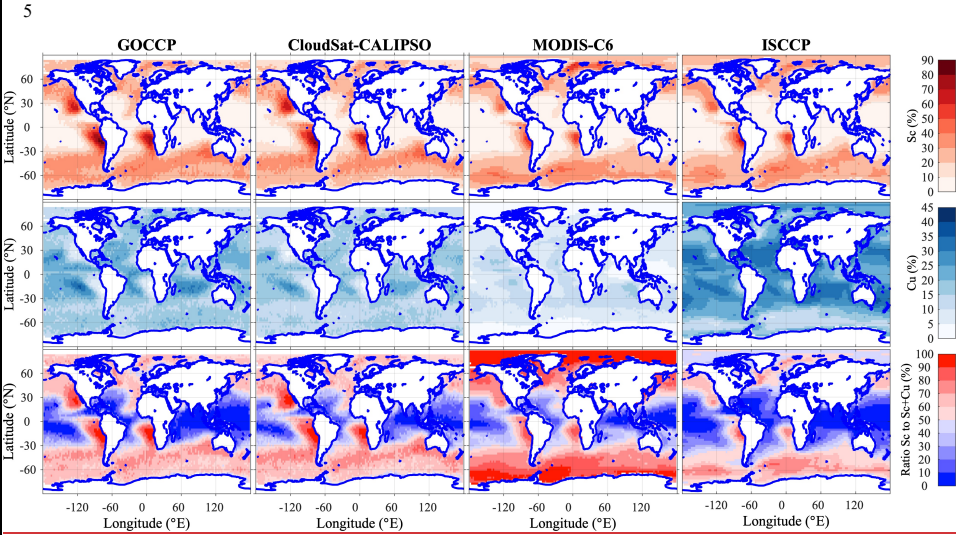


Deleted: PDF (left),

Deleted: middle

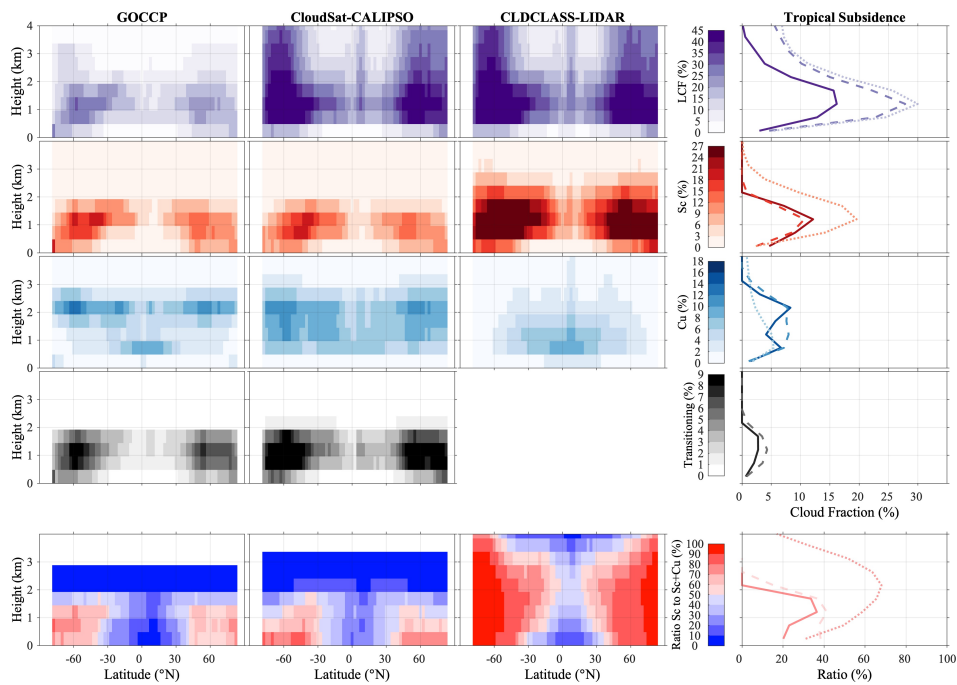
Deleted: Sc-Cu

Figure 15: Similar to Fig. 13, maps (x axis, longitude [°E]; y axis, latitude [°N]) of (top to bottom) Sc, Cu and the ratio Sc to Sc-and-Cu cloud fraction (%) for (left to right) GOCCP (2007-2016), RL-GeoProf (2007-2010), MODIS Terra and Aqua Cloud Regime (2003-2015) and ISCCP Weather State (2001-2008).



Formatted: Indent: Left: -0.38"

Figure 16: Zonal profiles (x axis, latitude [°N]; y axis, height [km]) of low (first row), Sc (second row), Cu (third row) and transitioning (i.e., broken Sc, Cu under Sc and Cu with stratiform outflow; fourth row) cloud fraction (%) for GOCCP (first column), RL-GeoProf (second column) and 2BCCL (third column). As in Fig. 13, different color bars are used to better separate each type of cloud. The corresponding profiles for tropical subsidence regimes ($\omega_{500} > 10$ hPa/day between 35°S/N following Cesana et al., 2019a) are represented in the fourth column for GOCCP (solid line), RL-GeoProf (dashed line) and 2BCCL (dotted line). Note that extratropical profiles are shown in Fig. S10.



releted: 5

releted: Sc-Cu

releted: (

releted: 6