Low-level marine tropical clouds in six CMIP6 models are too few, too bright but also too compact and too homogeneous

Dimitra Konsta
LMD/IPSL, Sorbonne Université, Ecole Polytechnique, Ecole Normal Supérieur, CNRS, Paris, France

Jean-Louis Dufresne
LMD/IPSL, Sorbonne Université, Ecole Polytechnique, Ecole Normal Supérieur, CNRS, Paris, France

Hélène Chepfer
LMD/IPSL, Sorbonne Université, Ecole Polytechnique, Ecole Normal Supérieur, CNRS, Paris, France.

Jessica VIAL
LMD/IPSL, Sorbonne Université, Ecole Polytechnique, Ecole Normal Supérieur, CNRS, Paris, France.

Tsuyoshi Koshiro
Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan

Hideaki Kawai
Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan

Alejandro Bodas-Salcedo
Met Office Hadley Centre, Exeter, United Kingdom

Romain Roehrig
CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

Masahiro Watanabe
Atmosphere and Ocean Research Institute, University of Tokyo, Kashiwa, Japan

Tomoo Ogura
National Institute for Environmental Studies, Tsukuba, Japan

Abstract

Several studies have shown that most climate models underestimate cloud cover and overestimate cloud reflectivity, particularly for the tropical low-level clouds. Here we analyze the characteristics of low-level tropical marine clouds simulated by six climate models, which provided COSP output within the CMIP6 project. CALIPSO lidar observations and PARASOL mono-directional reflectance are used for model evaluation. It is found that the 'too few, too bright' bias is still present for these models. The reflectance is particularly overestimated when cloud cover is low. Models do not simulate any optically thin clouds. They fail to reproduce the increasing cloud optical depth with increasing lower tropospheric stability as observed. These results suggest that most models do not sufficiently account for the effect of the small-scale spatial heterogeneity in cloud properties or the variety of cloud types at the grid scale that is observed.

1. Introduction

Low-level clouds are ubiquitous in the tropics and play an important role in the Earth's radiative budget and climate radiative feedbacks. Low-level cloud feedback differences are a major source of spread in model estimates of climate sensitivity (e.g., Roeckner et al., 1987; Bony and Dufresne, 2005; Webb et al., 2006; Vial et al., 2013; Zelinka et al., 2020).

The cloud radiative effect in the SW (shortwave) primarily depends on the cloud cover, but also on cloud albedo. Several studies have shown that most climate models underestimate the cloud cover and overestimate the cloud albedo, a deficiency referred to as the 'too few too bright bias' (e.g., Webb et al., 2001; Zhang et al., 2005; Nam et al., 2012; Klein et al., 2013). The coupling between these two biases mainly results from the radiation budget tuning of coupled atmosphere-ocean climate models, needed to prevent any global temperature drift due to an unbalanced energy budget (e.g., Mauritsen et al., 2012; Hourdin et al., 2017). This deficiency particularly impacts tropical marine low-level clouds (Webb et al., 2001; Zhang et al., 2005; Nam et al., 2012; Klein et al., 2013). The goal of this study is to examine whether the "too few too bright" bias is still present in six models that recently participated to the sixth phase of the Coupled Model Intercomparison Project (CMIP6) (Eyring et al., 2016), and to examine whether it may have a common origin among different climate models.

The CMIP6 climate models, the satellite observations and the methodology used for the model evaluation are described in Sect. 2. The simulated cloud cover, reflectance and vertical distribution are analyzed section 3. Conclusions are given in Sect. 4.

2. Methodology

2.1 CMIP6 models and COSP simulator

Six general circulation models (GCMs) that participated in CMIP6 are considered (Table S1 in supporting information). We analyze the results of the AMIP experiment where atmospheric models are forced with observed sea surface temperatures and sea-ice cover. This AMIP model configuration, in which the interannual variability is rather consistent with the historical sequence, especially over the tropical ocean, allows us to use a shorter record for model-observation comparison than if coupled configuration was used. The simulated cloud properties are compared with observations over the 2007-2010 period using the Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package (COSP) (Bodas-Salcedo et al., 2011). More specifically, we use the CALIPSO (Chepfer et al., 2008) and PARASOL (Konsta et al., 2016) simulators that compute the cloud cover, the vertical profile of the cloud fraction and the cloud reflectance that may be directly compared with observations. The total reflectance observed by the instrument contains the clear sky contribution. The cloud reflectance CR, which excludes the contribution of the clear sky around clouds, is calculated for every grid cell and for each time step, according to the relation

$$CR = [\, R - (1 - CC) * CSR \,] / CC \qquad (2.1)$$

where R is the monodirectional total reflectance, CC is the cloud cover estimated by the lidar simulator and CSR is the clear-sky reflectance (Konsta et al., 2016).

The analysis of the instantaneous cloud properties gives a detailed view of how the parameterizations actually work, allowing a more demanding evaluation of their behaviors and possibly finding ways to improve them (Konsta et al., 2016). For that reason we use the highest possible temporal resolution, which is a daily resolution for the CMIP6 experiments analyzed here, meaning that Eq. 2.1 is calculated using the daily averages of CC and CR. Using multiple models (IPSL-CM6A, CNRM, MRI and HadGEM3) we verified that the analysis results shown here are consistent when using either daily outputs or outputs every 3 hours. Regarding the spatial resolution, we keep the native resolution of the models (Table S1), which is close to that of the observations (2°x2°).

2.2. Observational and reanalysis datasets

For each GCM, we compare the cloud cover and the cloud vertical distribution simulated by COSP with the GCM-Oriented CALIPSO Cloud Product (GOCCP), developed to be consistent with COSP (Chepfer et al., 2010). Here we use four years of observations (2007-2010) of daily statistics being representative of the cloud climatology over a 2°x2° grid and with a vertical resolution of 480 m. Clouds present at a pressure larger than 680 hPa are considered as low-level clouds following ISCCP definition. A more detailed description of the observational and reanalysis datasets is presented in Text S1.

The PARASOL satellite provides measurements of reflectance at 6x6 km^2 (Tanré et al., 2011). The monodirectional reflectance measurements are only kept for one viewing angle (Konsta et al., 2012) and are collocated to the CALIPSO trace. Then, in every 2°x2º grid box, the mean cloud reflectance is calculated from the values of the reflectance observed by PARASOL and the cloud cover observed by CALIPSO at the same time (Eq. 2.1) (Konsta et al., 2012). The directional cloud reflectance is chosen because it is less sensitive to cloud geometry and instrument viewing angle than the cloud albedo and is essentially dependent on the cloud optical depth (Konsta et al., 2016). Cloud optical depth increases with cloud reflectance, e.g. cloud reflectance of 0.1, 0.3 and 0.6 correspond to values of cloud optical depth of about 1.6, 5.5, and 16.5 respectively for homogeneous liquid water clouds composed of spherical droplets. Cloud albedo and cloud reflectance are closely related and the two can be merged if one wishes to retain only a general image (Fig. S1).

In order to analyze how cloud properties depend on their environment, we use the ERA-Interim atmospheric reanalysis (Dee et al., 2011). These data are interpolated on a 2°x2° grid at 13:30 local time, the approximate time of the CALIPSO/PARASOL daytime passing in the tropics. We will make use of the lower tropospheric stability (LTS), defined as the potential temperature difference $\Delta\theta$ between the 700 hPa level and the surface (e.g., Klein and Hartmann 1993).

3. Low-level tropical marine clouds in six CMIP6 models

3.1. The too few and too bright bias

We focus on the tropical ocean (30°S-30°N) and on situations where low-level clouds are the dominant clouds. To determine whether low-level clouds are dominant in a mesh, we use as a criterion that the fraction CClow of low-level clouds is larger than 90% of the total cloud cover (CClow > 0.9 * CC). Adding the criterion of excluding mid and high-level clouds (CCmid + CChigh < 0.1 * CC) did not significantly change the results (Konsta et al., 2012). We obtain that the relative frequency of occurrence of situations where low-level clouds are dominant in a 2°x2° grid cell over tropical oceans is 35% in observations and from 27% up to 40% in models (Fig. S2). This is consistent with the value of about 30% obtained by Oreopoulos et al. (2017). All the results presented in the rest of the paper concern these situations.

The multi-model mean low-level cloud cover presents a spatial pattern that corresponds globally to the observations with fairly low and uniform values in the trade wind regions, and higher values in the east of the ocean basins (Fig. 1a,c). However, observations show that the cloud cover is close to 1 along the east coast of the tropical oceans, while the model ensemble mean cloud cover is only about 0.7 above the same areas. Beyond the multi-model mean, this underestimation of cloud cover is present in all models except IPSL-CM6A (see Fig. S3 for individual models). This bias is not due to the too low occurrence of clouds with the right fraction but rather to lack of clouds with a high fraction (Fig. 1e). The frequency of occurrence of low-level clouds with a fraction close to 1 is small for all models due to parameterization problems in the stratocumulus clouds (Slingo, 1980; Kawai et al., 2019), except IPSL-CM6A, for which dedicated developments clearly improved their representation (Hourdin at al., 2019) but led to an overestimation of their occurrence. For cloud cover lower than 1, observations show a fairly flat statistical distribution with a maximum around 0.35, while almost all models show a more sharp and skewed distribution, with a maximum around 0.1-0.2 (Fig. 1e). This high frequency of occurrence of low cloud cover is found in the observations for small tropical cumulus clouds (Mieslinger et al., 2019). An exception is MIROC6 which shows a fairly flat statistical distribution, but a maximum around 0.6.

For the cloud reflectance, the difference between observations and models is even more dramatic (Fig.1b,d) (see Fig. S3 for individual models). The observed reflectance PDF is highly skewed and peaks at a low value of 0.12 (Fig.1f). The most frequent low-level clouds have a low reflectance. The PDF of the models' reflectance is in contrast almost symmetric, centered at a much higher value. The median of the cloud reflectance is about 0.15 for the observations. It is much larger for the models, going from 0.25 (for IPSL-CM6A) up to 0.4 (for HadGEM3) with a mean value of about 0.35.

3.2. Relationship between low-level cloud cover and brightness

We now analyze the covariation between cloud fraction and cloud reflectance. Two separate cloud populations appear clearly in the observations (Fig. 2a): one population with a small or intermediate cover (CC<60%) and a small reflectance (CR < 0.3) corresponding to cumulus clouds with cumulus cloud regime covering most of the ocean, and another population with a large reflectance (0.2 < CR < 0.7) and a cloud cover close to one corresponding to stratocumulus clouds mainly on the east side of the ocean basins (Konsta et al., 2016). This is consistent with what is already shown in Figure 1 but emphasizes that, for cumulus clouds, their reflectance is low when their cover is low, and it increases with increasing cloud cover. A synthesis view is shown in Fig. 2h, where cloud reflectance has been averaged in each cloud cover bin. This is consistent with the results of Leahy et al. (2012) who show that the relative fraction of optically thin clouds increases with decreasing low-level cloud cover.

Models show a very different picture. As already noted in Fig. 1, only two models (IPSL-CM6A and MRI) simulate the two distinct cloud populations. But what is clear here is the inability of the models to simulate clouds with low fraction and low reflectance (i.e. low optical thickness, low water content). Instead of showing an increase in cloud reflectance with increasing cloud cover, several models show an opposite relationship, especially when the cloud cover is low. In these models (HadGEM3, IPSL-CM6A and to a lesser degree GFDL), the smaller the cloud fraction, the larger the cloud reflectance. This behavior was also noted in the IPSL-CM5 model family (Konsta et al., 2016). However, several of them (IPSL, CNRM, HadGEM3, MIROC6 and GFDL) show a positive relationship between cloud fraction and cloud reflectance when CC > 0.4. MIROC6 simulates the increase in cloud reflectance with the cloud fraction, but it fails to simulate enough cloud with low fraction and clouds with small reflectance. MRI simulates the increase in cloud reflectance with cloud fraction for the cumulus clouds only, but cumulus cloud reflectance is too high and CR for high cloud fraction is too low. The difficulty of the models to reproduce the increase of cloud reflectance with increasing cloud cover is evident in Fig. 2-h, and none of the models simulate the low values of cloud reflectance when the cloud cover is low.

3.3. Sensitivity of the low-level cloud properties to their environment

As is well recognized (e.g., Klein and Hartmann 1993, Wood and Bretherton 2006), the cloud cover increases when the LTS increases (Fig. 3a) in long-term observations. This feature is examined here for instantaneous model/observations pairs and is shown to be reproduced by the models, with a slope consistent with observations but with a bias that can be large. In models, the LTS when low-level clouds are dominant are too low compared to those in the reanalysis, except for GFDL (Fig. 3c).

Observations show that the cloud reflectance increases with the LTS (Fig. 3b). But all models simulate a decrease in cloud reflectance with increased LTS (Fig. 3b), i.e. a variation opposite to that observed. This problem is consistent with the large difference between observations and models in how cloud reflectance varies with cloud cover (Fig. 2).

Observations also show the increase of the cloud cover when the near-surface wind speed increases (Fig. 3d) as explained in Nuijens et al. (2015) and already mentioned in previous analyses (Mieslinger et al., 2019, Scott et al., 2020). In contrast, the models simulate no dependence, they only exhibit a similar cloud cover – wind relationship for low wind speeds (except for MIROC) but not when the surface wind speed exceeds about 5m/s. The cloud reflectance shows no dependence on the surface wind speed both for the observations and the models (not shown).

3.4. Vertical structure of low-level cloud properties

The vertical structure of low-level clouds is critical as it may significantly impact low-level cloud feedbacks (Brient et al. 2016). In observations, the low-level cloud fraction over ocean exceeds 10% from slightly above the surface up to 2.5 km with a maximum of about 20% near 1.25 km (Fig. 4a). Our sample of CMIP6 models do not show the strong bias present in most of the CMIP5 models for which the cloud layer was confined within the first kilometer (Nam et al., 2012). However, the models differ significantly from one to another; while HadGEM3, MRI and MIROC6 simulate the maximum cloud fraction at a height close to that in the observations, other models simulate it at a much lower (750 m in CNRM and GFDL) or higher altitude (2.2 km in IPSL). There is also a large inter-model spread in the cloud fraction maximum, ranging from about 15 % (for CNRM, MRI and GFDL) to about 30% for MIROC6, HadGEM3 being the closest to the observed value (~22%). It should be noted that the 480 m vertical resolution of the data from the GOCCP observations and the COSP simulator smooths the cloud profiles and therefore limits a detailed analysis along the vertical.

CALPISO lidar permits observations of optically-thin low-level clouds (CR < 0.2, i.e. optical thickness < 3) throughout their depth (Chepfer et al., 2008). As shown in section 3.1, these clouds are dominant in observations but not in models. As compared to the overall cloud profile (Fig. 4a), optically-thin clouds tend to be shallower on average (maximum peaks at 750 m) with reduced cloudiness throughout cloud depth (Fig. 4b). All models (except IPSL) also simulate shallower optically-thin clouds with a maximum cloud fraction at around 750 m. But, unlike in the observations where optically-thin clouds can be found up to 2.5 km, in models these clouds remain exclusively confined within the lowest atmospheric levels. The IPSL model is the only one to simulate these clouds, yet with a strong overestimation of the amplitude and height of the cloud fraction maximum.

In observations, optically thick low-level clouds (CR > 0.4, i.e. optical thickness > 8) exhibit a greater vertical extension and a significantly larger maximum fraction than optically-thin clouds. Note that the sharp decrease in cloud fraction below the cloud peak height may be partially due to the attenuation of the lidar beam as it passes through thick clouds. Thus, the cloud fractions at low levels are strongly affected by the cloud top height in both models and observations.

To provide a more complete view, we show on Figure 4d how cloud top altitude varies with cloud reflectance. In observations, the cloud-top height is at about 1.5 km, in good agreement with Lu et al. (2021), and increases only slightly with the cloud reflectance. In contrast, this increase is substantially stronger in the models, especially for optically-thin clouds (CR below 0.4-0.5). This is also visible when the mean cloud-top altitude is shown as a function of both the cloud cover and the cloud reflectance (Fig. S4). A hypothesis to explain this difference is that at the scale of a 2°x2° mesh some optically-thin veil clouds, commonly observed beneath the trade inversion in stratocumulus-to-cumulus transition zones, but also more broadly over the tropical oceans (Kuang-Ting et al., 2018; Wood et al. 2018), could be missing in models. Results shown on Figures 3 and 4 are not significantly changed when removing situations where stratocumulus type clouds are dominant (cloud fraction above 0.9), which suggests that this discrepancy between observations and models concerns primarily cumulus-type of clouds.

4. Discussion and Conclusions

The "too few and too bright" bias of low-level clouds is still present in the subset of CMIP6 models we analyzed. The distribution of the observed daily cloud cover shows a broad maximum of cloud fraction at around 0.35, and a sharp secondary maximum near 1 corresponding to stratocumulus clouds over the eastern part of the ocean basins. For most of the models, this distribution has a marked main mode for low values of cloud cover and a missing or very limited secondary maximum for cloud cover near 1, except for IPSL-CM6A for which this secondary maximum is large and for MIROC6 for which this distribution is flat and symmetrical. The errors on the daily cloud reflectance are very different. The distribution is almost symmetrical for all models, while for the observations the distribution is concentrated around the low values with a long tail towards the high reflectance. This frequent occurrence of optically-thin low-level clouds is also found by Leahy et al. (2012) and Mieslinger et al. (2021).

The co-variations of cloud cover and cloud reflectance also exhibit very different behaviors between models and observations. While in observations the cloud reflectance increases as the cloud fraction increases, models show either an inverse dependence or no dependence at all. The cloud optical thickness in models is much too large when the cloud cover is low. A consequence of this problem emerges when analyzing the dependence of cloud properties on cloud environmental conditions. In particular, while the cloud fraction increases with the lower tropospheric stability in both observations and models, the reflectance increases with the LTS in observations but not in models.

The vertical profile of cloud fraction in this sample of CMIP6 models better agrees with that of the observations than did the CMIP5 models (Nam et al., 2012). However, the cloud-top height is too low for optically-thin clouds. Cloud-top height increases much faster with cloud optical thickness in these CMIP6 models than in observations.

These results may reflect the fact that outside the stratocumulus region on the eastern part of the oceans, the models simulate small cumulus clouds that are “too compact”, i.e. low cloud cover, high reflectance. This could arise if the models’ representation of clouds does not sufficiently account for (if at all) the sub-grid scale heterogeneities of cloud properties. As noted by Del Genio et al. (1996), GCM cloud schemes assume that the cloud fractions by area and by volume are equal, i.e. clouds occupy the entire depth of individual model layers over the cloud fraction of that layer, whereas in observations (Brooks et al., 2005) and LES models (Neggers et al., 2011) the former is much larger than the later. Accounting for sub-grid scale heterogeneity in the geometry of clouds influences the cloud radiative properties, by increasing the fraction and reducing the reflectance (Jouhaud et al., 2018). In addition, accounting for sub-grid scale heterogeneity in the autoconversion rate reduces the cloud water content (Hotta, et al., 2020), and thus the cloud reflectance.

A complementary hypothesis is that the models simulate too often, or even almost exclusively, small cumulus clouds at low levels (i.e. near the lifting condensation level). In models, the distribution of cloud fraction resembles that of the observed active cumuli and the reflectance increases with the cloud-top altitude, as expected for this type of cloud. These clouds do not leave such a marked signature in the observations that we use here. This might be explained by recent analyses showing that thin layers of clouds are often present beneath the trade inversion, and generally mixed with other cloud types when looking at a scale of a few hundred kilometers (Wood et al. 2018, Bony et al., 2020, Stevens et al., 2020). In the observations that we use, which are on a 2°x2° grid, close to that of the models, the probability of observing only small cumulus clouds is low, they are almost always mixed with other cloud types. Another way to phrase our hypothesis is that the models do not manage to simulate, in the same atmospheric column, the variety of low-level cloud types that is present in nature.

Acknowledgements

DK was supported through the French Government Scholarships for High level research scientific visit. JLD and JV 7 acknowledge funding by CNES. JLD, JV and RR have received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 820829.
We also thank L. Silvers for his useful comments.
Thanks are due to CNES and NASA for the PARASOL and CALIPSO data and to AERIS/ICARE Data and Services Center for the colocalization of PARASOL data to CALIPSO trace.
We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP. We thank the climate modeling groups (listed in Table S1) for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF.

Open Research

Data Availability Statement

The CALIPSO-GOCCP data used for cloud properties in this study are available online through the GOCCP website (https://climserv.ipsl.polytechnique.fr/cfmip-obs/Calipso_goccp.html).

The POLDER/PARASOL Level-1 data were originally provided by CNES. The Level-1 PARASOL normalized radiance at CALIPSO/CALIOP subtrack are produced and distributed by ICARE Data and Services Center (https://www.icare.univ-lille.fr/parasol/products/).
The original CMIP6 data can be accessed through the ESGF data portal via https://esgf-node.llnl.gov/search/cmip6/

References

Belmonte Rivas, M. and Stoffelen, A.: Characterizing ERA-Interim and ERA5 surface wind biases using ASCAT, Ocean Sci., 15, 831–852, https://doi.org/10.5194/os-15-831-2019, 2019.

Bodas-Salcedo, A., and Coauthors, 2011: COSP: Satellite simulation software for model assessment. Bull. Amer. Meteor. Soc., 92, 1023–1043, https://doi.org/10.1175/2011BAMS2856.1.

Bony, S., and J. Dufresne (2005), Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models, Geophys. Res. Lett., 32, L20806, doi:10.1029/2005GL023851.

Bony, S., Schulz, H., Vial, J. & Stevens, B.(2020). Sugar, gravel, fish, and flowers: Dependence of mesoscale patterns of trade-wind clouds on environmental conditions. *Geophysical Research Letters*, *47*: e2019GL085988. doi:10.1029/2019GL085988

Boucher, O., et al. (2020). Presentation and evaluation of the IPSL-CM6A-LR climate model. J. Adv. Model. Earth Syst., 12(7), e2019MS002010. https://doi.org/10.1029/2019MS002010

Brient, F., Schneider, T., Tan, Z. *et al.* Shallowness of tropical low clouds as a predictor of climate models' response to warming. *Clim Dyn* **47,** 433–449 (2016). https://doi.org/10.1007/s00382-015-2846-0

Brooks, M. E., Hogan, R. J., & Illingworth, A. J. (2005). Parameterizing the difference in cloud fraction defined by area and by volume as observed with radar and lidar. Journal of the atmospheric sciences, 62(7), 2248–2260.

Chepfer, H., S. Bony, D. Winker, M. Chiriaco, J. Dufresne, and G. Sèze (2008), Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model, Geophys. Res. Lett., 35, L15704, doi:10.1029/2008GL034207.

Chepfer, H., S. Bony, D. Winker, G. Cesana, J. Dufresne, P. Minnis, C. Stubenrauch, and S. Zeng (2010), The GCM-Oriented CALIPSO Cloud Product (CALIPSO-GOCCP), J. Geophys. Res., 115, D00H16, doi:10.1029/2009JD012251.

Dee, D., et al. (2011), The ERA-interim reanalysis: Configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137, 553–597.

Del Genio, A. D., Yao, M.-S., Kovari, W., & Lo, K. K. (1996). A prognostic cloud water parameterization for global climate models. Journal of Climate, 9(2), 270–304.

Deschamps P-Y, Bréon F-M, Leroy M, Podaire A, Brickaud A, Buriez J-C, Sèze G (1994) The POLDER mission: instrument characteristics and scientific objectives. IEEE 32:598–615

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937-1958, doi:10.5194/gmd-9-1937-2016, 2016.

Fougnie B, Bracco G, Lafrance B, Ruffel C, Hagolle O, Tinel C (2007) PARASOL in-flight calibration and performance. Appl Opt 46:5435–5451

Guan, B., D. E. Waliser, and F. M. Ralph, 2018: An intercomparison between reanalysis and dropsonde observations of the total water vapor transport in individual atmospheric rivers. *J. Hydrometeor.*, 19, 321–337, https://doi.org/10.1175/JHM-D-17-0114.1.

Hourdin, F., et al., 2017: The Art and Science of Climate Model Tuning. *Bull. Amer. Meteor. Soc.*, **98**, 589–602, https://doi.org/10.1175/BAMS-D-15-00135.1.

Hourdin, F., Jam, A., Rio, C., Couvreux, F., Sandu, I., Lefebvre, M.-P., et al. (2019). Unified parameterization of convective boundary layer transport and clouds with the thermal plume model. *Journal of Advances in Modeling Earth Systems*, 11, 2910– 2933. https://doi.org/10.1029/2019MS001666

Hourdin, F., Rio, C., Grandpeix, J.-Y., Madeleine, J.-B., Cheruy, F., Rochetin, N., et al. (2020). LMDZ6A: The atmospheric component of the IPSL climate model with improved and better tuned physics. Journal of Advances in Modeling Earth Systems, 12, e2019MS001892. https://doi.org/10.1029/2019MS001892

Hotta, H., Suzuki, K., Goto, D., & Lebsock, M. (2020). Climate Impact of Cloud Water Inhomogeneity through Microphysical Processes in a Global Climate Model, *Journal of Climate*, *33*(12), 5195-5212, doi: 10.1175/JCLI-D-19-0772.1

Jouhaud J., J.-L. Dufresne, J.-B. Madeleine, Frédéric Hourdin, F. Couvreux, N. Villefranque, A. Jam (2018), Accounting for vertical subgrid-scale heterogeneity in low-level cloud fraction parameterizations, JAMES, https://doi.org/10.1029/2018MS001379

Kawai, H., Yukimoto, S., Koshiro, T., Oshima, N., Tanaka, T., Yoshimura, H., and Nagasawa, R.: Significant improvement of cloud representation in the global climate model MRI-ESM2, Geosci. Model Dev., 12, 2875–2897, https://doi.org/10.5194/gmd-12-2875-2019, 2019.

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

Klein, S. A., Zhang, Y., Zelinka, M. D., Pincus, R., Boyle, J., and Gleckler, P. J. (2013), Are climate model simulations of clouds improving? An evaluation using the ISCCP simulator, *J. Geophys. Res. Atmos.*, 118, 1329– 1342, doi:10.1002/jgrd.50141.

Konsta, D., H. Chepfer and J-L. Dufresne, 2012 : A process oriented characterization of tropical oceanic clouds for climate model evaluation, based on a statistical analysis of daytime A-train observations. Clim Dyn, 39, 2091–2108, doi:10.1007/s00382-012-1533-7.

Konsta, D., J-L. Dufresne, H. Chepfer, A. Idelkadi and G. Cesana, 2016 : Use of A-train satellite observations (CALIPSO-PARASOL) to evaluate tropical cloud properties in the LMDZ5 GCM, Clim Dyn, 47: 1263-1284, doi:10.1007/s00382-015-2900-y.

Kuan-Ting, O, Wood, R., Tseng, H.-H. (2018). Deeper, precipitating PBLs associated with optically thin veil clouds in the Sc-Cu transition. *Geophysical Research Letters*, 45, 5177– 5184. https://doi.org/10.1029/2018GL077084

Leahy, L. V., R. Wood, R. J. Charlson, C. A. Hostetler, R. R. Rogers, M. A. Vaughan, and D. M. Winker (2012), On the nature and extent of optically thin marine low clouds, J. Geophys. Res., 117, D22201, doi:10.1029/2012JD017929

Lu, X., Mao, F., Rosenfeld, D., Zhu, Y., Pan, Z., and Gong, W., 2021 : Satellite retrieval of cloud base height and geometric thickness of low-level cloud based on CALIPSO, Atmos. Chem. Phys., 15: 11979-12003, doi: 10.5194/acp-21-11979-2021

Luo, B., Minnett, P. J., Szczodrak, M., Nalli, N. R., & Morris, V. R. (2020). Accuracy Assessment of MERRA-2 and ERA-Interim Sea Surface Temperature, Air Temperature, and Humidity Profiles over the Atlantic Ocean Using AEROSE Measurements, *Journal of Climate*, *33*(16), 6889-6909.

Mauritsen, T., et al. (2012), Tuning the climate of a global model, J. Adv. Model. Earth Syst., 4, M00A01, doi:10.1029/2012MS000154

Medeiros, B., B. Stevens, I. Held, M. Zhao, D. Williamson, J. Olson, and C. Bretherton (2008), Aquaplanets, climate sensitivity, and low clouds, J. Clim., 21, 4974–4991.

Mieslinger, T., Horváth, Á., Buehler, S. A., & Sakradzija, M. (2019). The dependence of shallow cumulus macrophysical properties on large-scale meteorology as observed in ASTER imagery. *Journal of Geophysical Research: Atmospheres*, 124, 11477– 11505. https://doi.org/10.1029/2019JD030768

Mieslinger, T., Stevens, B., Kölling, T., Brath, M., Wirth, M., and Buehler, S. A.: Optically thin clouds in the trades, Atmos. Chem. Phys. Discuss. [preprint], https://doi.org/10.5194/acp-2021-453, in review, 2021.

Mülmenstädt, J., Sourdeval, O., Henderson, D. S., L'Ecuyer, T. S., Unglaub, C., Jungandreas, L., Böhm, C., Russell, L. M., and Quaas, J.: Using CALIOP to estimate cloud-field base height and its uncertainty: the Cloud Base Altitude Spatial Extrapolator (CBASE) algorithm and dataset, Earth Syst. Sci. Data, 10, 2279–2293, https://doi.org/10.5194/essd-10-2279-2018, 2018.

Nam, C., Bony, S., Dufresne, J.-L., and Chepfer, H. (2012), The 'too few, too bright' tropical low-cloud problem in CMIP5 models, *Geophys. Res. Lett.*, 39, L21801, doi:10.1029/2012GL053421.

Neggers, R. A. J., Heus, T., & Siebesma, A. P. (2011). Overlap statistics of cumuliform boundary-layer cloud fields in large-eddy simulations. Journal of Geophysical Research, 116, D21202. https://doi/org/10.1029/2011JD015650

Nuijens, L., Medeiros, B., Sandu, I., and Ahlgrimm, M. (2015), Observed and modeled patterns of covariability between low-level cloudiness and the structure of the trade-wind layer, J. Adv. Model. Earth Syst., 7, 1741– 1764, doi:10.1002/2015MS000483.

Oreopoulos, L., N. Cho, and D. Lee (2017), New insights about cloud vertical structure from CloudSat and CALIPSO observations, J. Geophys. Res. Atmos., 122, 9280–9300, doi:10.1002/2017JD026629

Pfahl, S., E. Madonna, M. Boettcher, H. Joos, and H. Wernli (2014), Warm conveyor belts in the ERA-Interim data set (1979–2010). Part II: Moisture origin and relevance for precipitation, J. Clim., 27, 27– 40, doi:10.1175/JCLI-D-13-00223.1

Pfeifroth, U., Mueller, R., and Ahrens, B.: Evaluation of Satellite-Based and Reanalysis Precipitation Data in the Tropical Pacific. J. Appl. Meteor. Climatol., 52, 634–644. doi: http://dx.doi.org/10.1175/JAMC-D-12-049.1, 2013

Roeckner E, Schlese U, Biercamp J, Leowe P (1987) Cloud optical depth feedbacks and climate modelling. Nature 329: 138±140

Roehrig, R., Beau, I., Saint-Martin, D., Alias, A., Decharme, B., Guérémy, J.-F., et al. (2020). The CNRM global atmosphere model ARPEGE-Climat 6.3: Description and evaluation. Journal of Advances in Modeling Earth Systems, 12, e2020MS002075. https://doi.org/10.1029/2020MS002075

Scott, R. C., Myers, T. A., Norris, J. R., Zelinka, M. D., Klein, S. A., Sun, M., & Doelling, D. R. (2020). Observed sensitivity of low-cloud radiative effects to meteorological perturbations over the global oceans. Journal of Climate, 33(18), 7717– 7734. https://doi.org/10.1175/jcli-d-19-1028.1

Silvers, Levi; Blanton, Chris; McHugh, Colleen; John, Jasmin G; Radhakrishnan, Aparna; Rand, Kristopher; Balaji, V; Dupuis, Christopher; Durachta, Jeff; Guo, Huan; Hemler, Richard; Lin, Pu; Nikonov, Serguei; Paynter, David J; Ploshay, Jeffrey; Vahlenkamp, Hans; Wilson, Chandin; Wyman, Bruce; Robinson, Thomas; Zeng, Yujin; Zhao, Ming (2018). NOAA-GFDL GFDL-CM4 model output prepared for CMIP6 CFMIP. Version 20190916.Earth System Grid

Slingo, J. M. (1980) A cloud parametrization scheme derived from GATE data for use with a numerical model, Q. J. Roy. Meteor. Soc., 106, 727-770 , https://doi.org/10.1002/qj.49710645008.

Smalley, K. M., and A. D. Rapp, 2020: The Role of Cloud Size and Environmental Moisture in Shallow Cumulus Precipitation. *J. Appl. Meteor. Climatol.*, **59**, 535–550, https://doi.org/10.1175/JAMC-D-19-0145.1.

Stevens, B., Bony, S., Brogniez, H., Hentgen, L., Hohenegger, C., Kiemle, C., L'Ecuyer, T., Naumann, A., Schulz, H., Siebesma, P., Vial , J., Winker, D. & Zuidema, P. (2020). Sugar, gravel, fish, and flowers: Mesoscale cloud patterns in the tradewinds. *Quarterly Journal of the Royal Meteorological Society*, *146*, 141-152. doi:10.1002/qj.3662

Tatebe, H., et al., 2019, Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6, Geosci. Model Dev., 12, 2727–2765, https://doi.org/10.5194/gmd-12-2727-2019

Tanré, D., Bréon, F. M., Deuzé, J. L., Dubovik, O., Ducos, F., François, P., Goloub, P., Herman, M., Lifermann, A., and Waquet, F.: Remote sensing of aerosols by using polarized, directional and spectral measurements within the A-Train: the PARASOL mission, Atmos. Meas. Tech., 4, 1383–1395, https://doi.org/10.5194/amt-4-1383-2011, 2011.

Vergados, P., A. J. Mannucci, and C. O. Ao, 2014: Assessing the performance of GPS radio occultation measurements in retrieving tropospheric humidity in cloudiness: A comparison study with radiosondes, ERA-Interim, and AIRS data sets. *J. Geophys. Res. Atmos.*, 119, 7718–7731, https://doi.org/10.1002/2013JD021398.

Vial, J. , J-L Dufresne, S. Bony (2013). On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates; Clim. Dynamics, 41(11-12), pp 3339-3362, doi: 10.1007/s00382-013-1725-9

Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Carslaw, K., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Jones, C., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whitall, M., Williams, K., and Zerroukat, M.: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, Geosci. Model Dev., 12, 1909–1963, https://doi.org/10.5194/gmd-12-1909-2019, 2019.

Webb, M., C. Senior, S. Bony, and J.-J. Morcrette (2001), Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models, Clim. Dyn., 17(12), 905–922.

Wood, R., and C. S. Bretherton, 2006: On the Relationship between Stratiform Low Cloud Cover and Lower-Tropospheric Stability. *J. Climate*, **19**, 6425–6432, https://doi.org/10.1175/JCLI3988.1.

Yukimoto S. et al. 2019: The Meteorological Research Institute Earth System Model version 2.0, MRI-ESM2.0: Description and basic evaluation of the physical component. *J. Meteor. Soc.* Japan, 97, 931–965, doi:10.2151/jmsj.2019-051.

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., et al. (2020). Causes of higher climate sensitivity in CMIP6 models. Geophysical Research Letters, 47, e2019GL085782. https://doi.org/10.1029/2019GL085782

Zhang, M. H., et al. (2005), Comparing clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements, *J. Geophys. Res.*, 110, D15S02, doi:10.1029/2004JD005021.

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., Chen, J.-H., Chen, X., Donner, L. J., Dunne, J. P., Dunne, K., Durachta, J., Fan, S.-M., Freidenreich, S. M., Garner, S. T., Ginoux, P., Harris, L. M., Horowitz, L. W., Krasting, J. P., Langenhorst, A. R., Liang, Z., Lin, P., Lin, S.-J., Malyshev, S. L., Mason, E., Milly, P. C. D., Ming, Y., Naik, V., Paulot, F., Paynter, D., Phillipps, P., Radhakrishnan, A., Ramaswamy, V., Robinson, T., Schwarzkopf, D., Seman, C. J., Shevliakova, E., Shen, Z., Shin, H., Silvers, L. G., Wilson, J. R., Winton, M., Wittenberg, A. T., Wyman, B., & Xiang, B. (2018a). The GFDL global atmosphere and land model AM4.0/LM4.0: 1. Simulation characteristics with prescribed SSTs. Journal of Advances in Modeling Earth Systems, 10, 691– 734. https://doi.org/10.1002/2017MS001208

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., Chen, J.-H., Chen, X., Donner, L. J., Dunne, J. P., against the style, K., Durachta, J., Fan, S.-M., Freidenreich, S. M., Garner, S. T., Ginoux, P., Harris, L. M., Horowitz, L. W., Krasting, J. P., Langenhorst, A. R., Liang, Z., Lin, P., Lin, S.-J., Malyshev, S. L., Mason, E., Milly, P. C. D., Ming, Y., Naik, V., Paulot, F., Paynter, D., Phillipps, P., Radhakrishnan, A., Ramaswamy, V., Robinson, T., Schwarzkopf, D., Seman, C. J., Shevliakova, E., Shen, Z., Shin, H., Silvers, L. G., Wilson, J. R., Winton, M., Wittenberg, A. T., Wyman, B., & Xiang, B. (2018b). The GFDL global atmosphere and land model AM4.0/LM4.0: 2. Model description, sensitivity studies, and tuning strategies. Journal of Advances in Modeling Earth Systems, 10, 735– 769. https://doi.org/10.1002/2017MS001209

Figures

Figure 1: For situations over the tropical ocean where low-level clouds are dominant, geographical distribution of the a) observed (CALIPSO-GOCCP) and c) multi-models mean (IPSL-CM6A, CNR-CM6, HadGEM3, MRI-ESM2, MIRCO6 and GFDL-CM4) total cloud cover, and of the b) observed (CALIPSO-GOCCP, PARASOL) and d) multi-models mean cloud reflectance. For the same situations, Probability Distribution Function of e) the cloud cover and f) the cloud reflectance observed with CALIPSO-GOCCP and PARASOL (black line) and simulated by the models (colored lines). All the data are daily for the 4 years period 2007 – 2010.

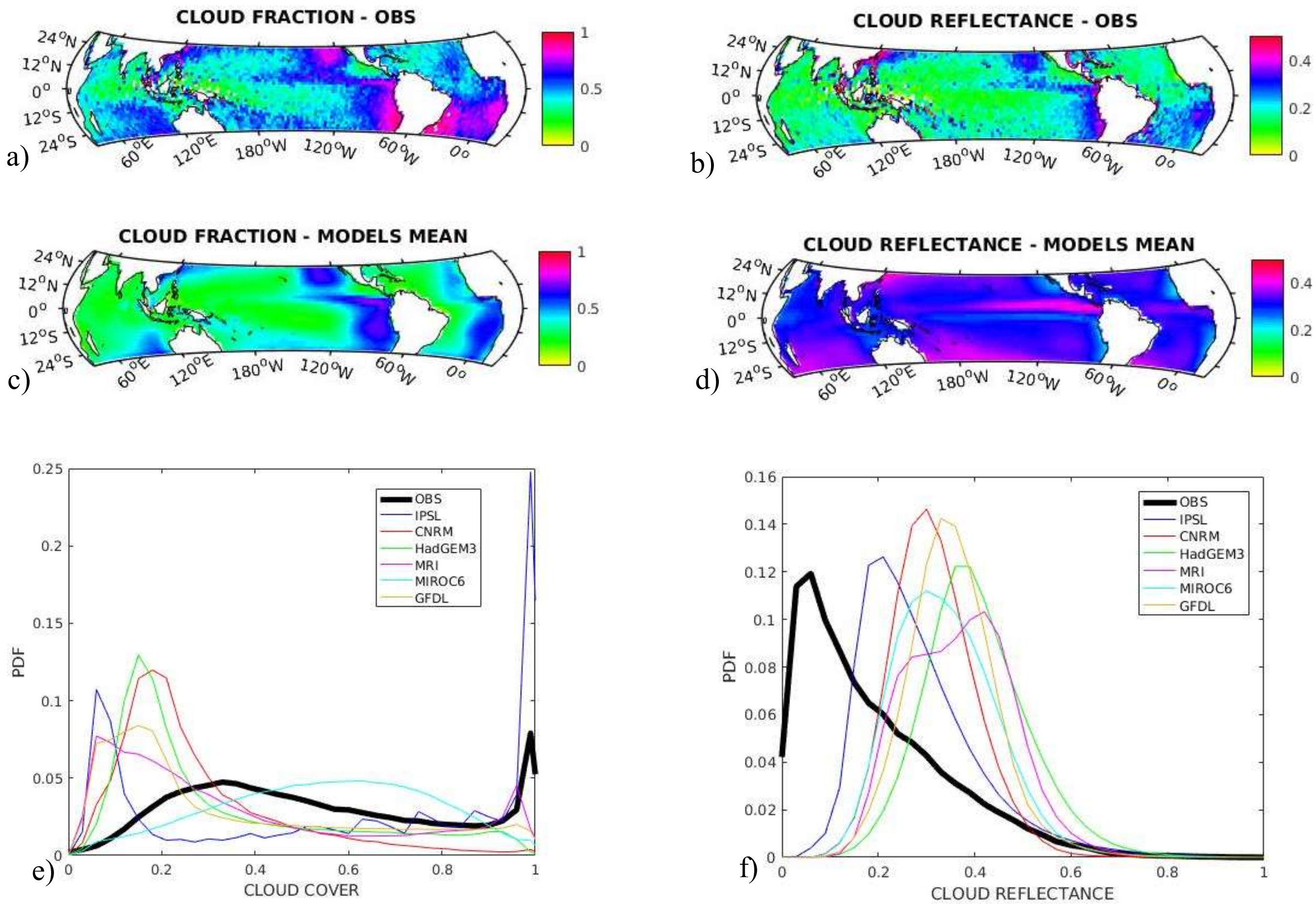
Figure 2: 2D histograms of cloud reflectance and cloud cover a) observed (CALIPSO-GOCCP, PARASOL) and simulated by b) IPSL, c) CNRM, d) HadGEM3, e) MRI, f) MIROC6, and g) GFDL models, and h) mean cloud reflectance for each cloud cover bin of 0.03 observed with CALIPSO-GOCCP and PARASOL (black line) and simulated by the models (colored lines). The error bars mark the standard error of the mean cloud reflectance within each cloud cover bin. All the data are daily values over the tropical ocean, when low-level clouds are dominant and for the period 2007-2010. The colorbar gives the number of points at each grid cell (cloud cover – cloud reflectance) divided by the total number of points.

Figure 3: a) Cloud cover, b) Cloud reflectance as a function of the LTS, c) PDF of LTS and d) Cloud Cover as a function of the surface wind speed. The black lines correspond to observation and ERA Interim reanalysis, the colored lines to models results. All the data are daily values taken over the tropical ocean, when low-level clouds are dominant and for the period 2007-2010. The standard error of the mean is below 0.01 % (Text S2) and not shown in the Figure for the sake of clarity.

Figure 4: Vertical profile of the cloud fraction (CF3D) a) for all low-level clouds, b) for optically thin low-level clouds (CR<0.2), c) for optically thick low-level clouds (CR>0.4), and d) mean cloud

top altitude as a function of cloud reflectance, for the observations (CALIPSO-GOCCP, PARASOL, black lines) and models (lidar and PARASOL simulator, colored lines). Cloud top altitude is defined as the highest level of low-level clouds where the sum of the cloud fraction (CF3D) from the top is greater than 10% of the cloud cover ($sumCF3D_{(from\ top)} > 10\%CC$). All the data are daily values over the tropical ocean, when low-level clouds are dominant and for the period 2007-2010. The standard error of the mean is below 0.01 % (Text S2) and not shown in the Figure for the sake of clarity.

Figure 1.



a)
CLOUD FRACTION - OBS
b)
CLOUD REFLECTANCE - OBS
c)
CLOUD FRACTION - MODELS MEAN
d)
CLOUD REFLECTANCE - MODELS MEAN
24°N
12°N
0°
12°S
24°S
60°E
120°E
180°W
120°W
60°W
0°
e)
PDF
CLOUD COVER
f)
CLOUD REFLECTANCE
OBS
IPSL
CNRM
HadGEM3
MRI
MIROC6
GFDL

Figure 2.

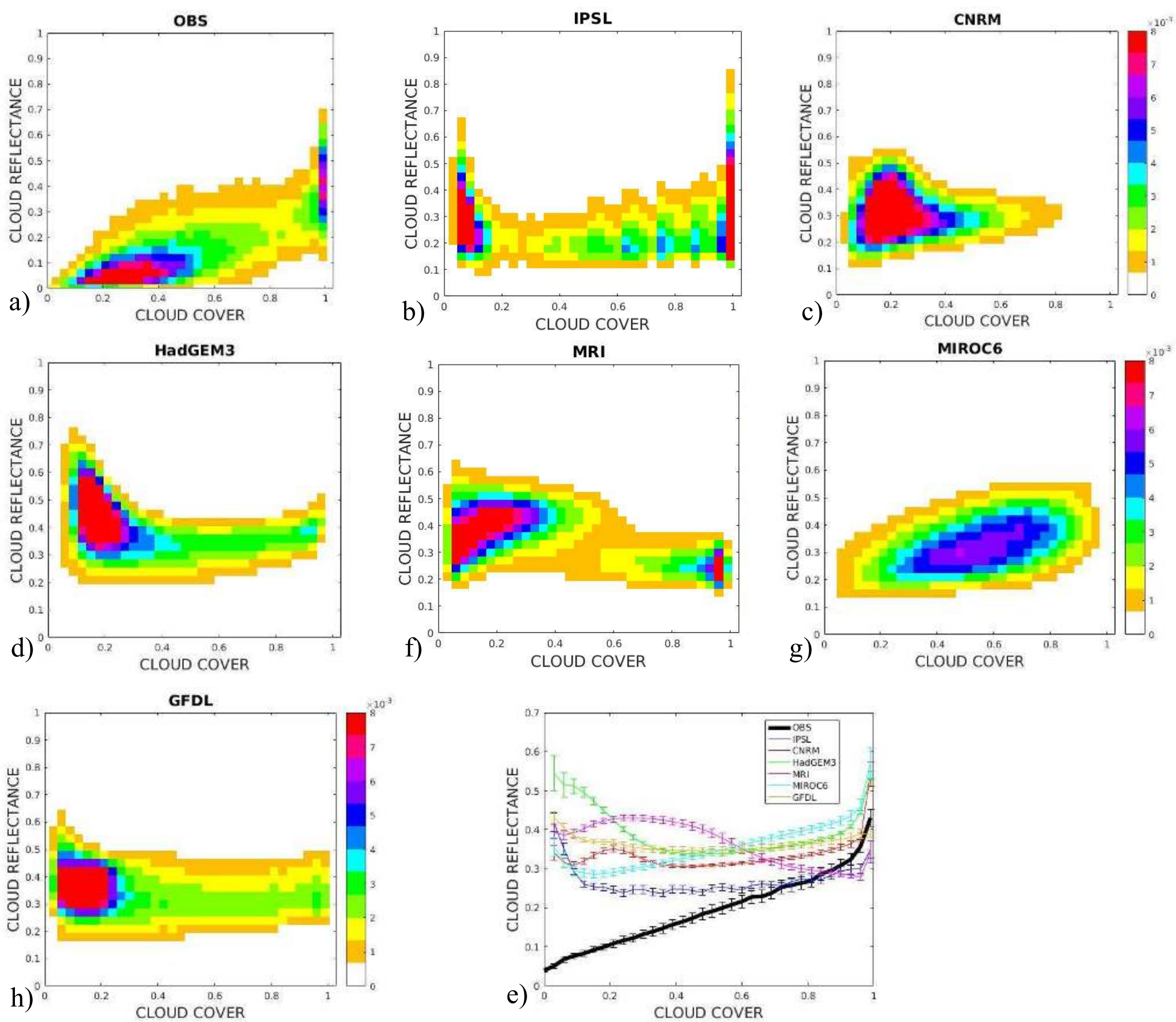

OBS
IPSL
CNRM
HadGEM3
MRI
MIROC6
GFDL
CLOUD REFLECTANCE
CLOUD COVER
a)
b)
c)
d)
f)
g)
h)
e)
OBS
IPSL
CNRM
HadGEM3
MRI
MIROC6
GFDL

Figure 3.

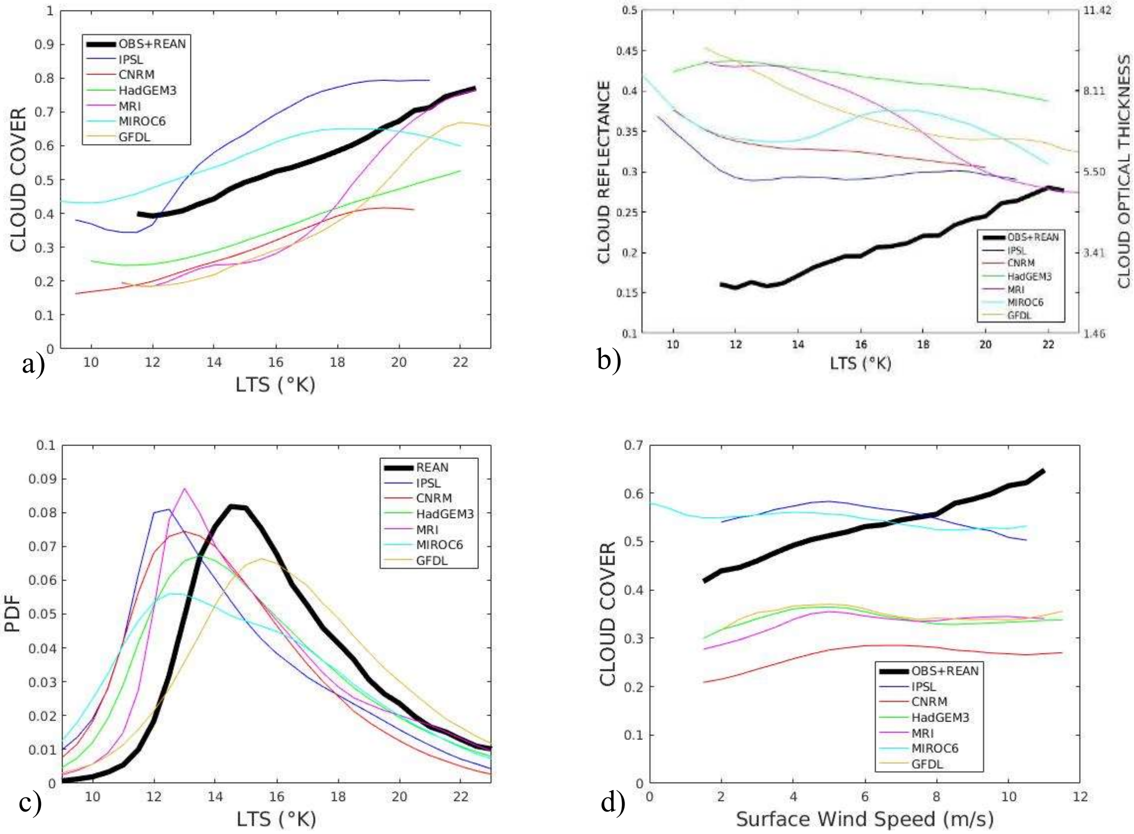
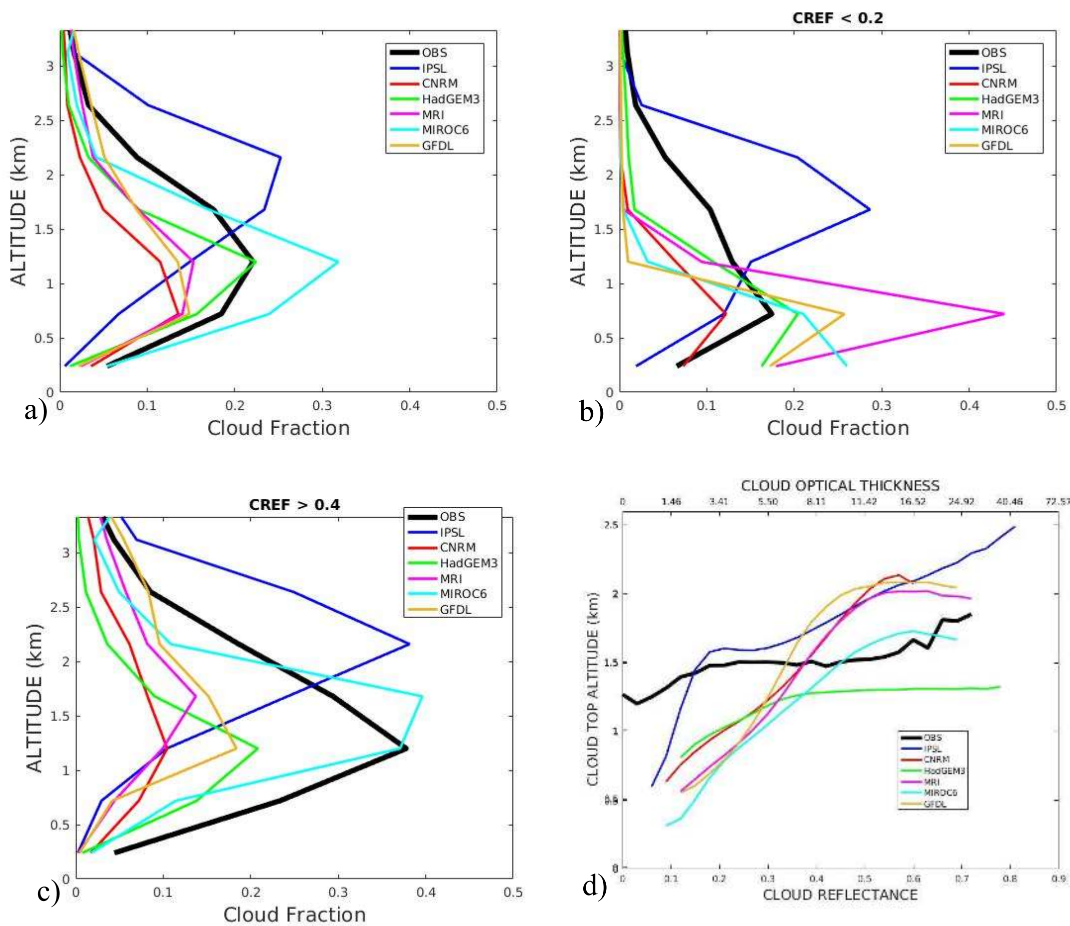

a)
CLOUD COVER
LTS (°K)
OBS+REAN
IPSL
CNRM
HadGEM3
MRI
MIROC6
GFDL
b)
CLOUD REFLECTANCE
CLOUD OPTICAL THICKNESS
LTS (°K)
OBS+REAN
IPSL
CNRM
HadGEM3
MRI
MIROC6
GFDL
c)
PDF
LTS (°K)
REAN
IPSL
CNRM
HadGEM3
MRI
MIROC6
GFDL
d)
CLOUD COVER
Surface Wind Speed (m/s)
OBS+REAN
IPSL
CNRM
HadGEM3
MRI
MIROC6
GFDL

Figure 4.


a)
ALTITUDE (km)
Cloud Fraction
OBS
IPSL
CNRM
HadGEM3
MRI
MIROC6
GFDL
b)
CREF < 0.2
ALTITUDE (km)
Cloud Fraction
OBS
IPSL
CNRM
HadGEM3
MRI
MIROC6
GFDL
c)
CREF > 0.4
ALTITUDE (km)
Cloud Fraction
OBS
IPSL
CNRM
HadGEM3
MRI
MIROC6
GFDL
d)
CLOUD OPTICAL THICKNESS
CLOUD TOP ALTITUDE (km)
CLOUD REFLECTANCE
OBS
IPSL
CNRM
HadGEM3
MRI
MIROC6
GFDL

Geophysical Research Letters

Supporting Information for

Low-level marine tropical clouds in six CMIP6 models are too few, too bright but also too compact and too homogeneous

Dimitra Konsta[1], Jean-Louis Dufresne[1], Hélène Chepfer[1], Jessica Vial[1], Tsuyoshi Koshiro[2], Hideaki Kawai[2], Alejandro Bodas-Salcedo[3], Romain Roehrig[4], Masahiro Watanabe[5], Tomoo Ogura[6]

[1] LMD/IPSL, Sorbonne Université, Ecole Polytechnique, Ecole Normal Supérieur, CNRS, Paris, France
[2] Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan
[3] Met Office Hadley Centre, Exeter, United Kingdom
[4] CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France
[5] Atmosphere and Ocean Research Institute, University of Tokyo, Kashiwa, Japan
[6] National Institute for Environmental Studies, Tsukuba, Japan

Contents of this file

Text S1: Observational properties

The observational datasets used in the study are the GCM-Oriented CALIPSO Cloud Product (GOCCP) that provides the total cloud cover (CC), the cloud cover in three layers (Low, Mid, High following ISCCP definition – CClow, CCmid, CChigh) as well as the cloud fraction profile (3D Cloud Fraction - CF3D), and the PARASOL visible directional reflectance, which is a surrogate for the cloud optical depth.

The GOCCP product consists in applying Scattering Ratio (SR) thresholds values to the 532 nm lidar SR signal to detect the presence of clouds (Chepfer et al., 2010). The cloud detection (0 or 1) is done at the original horizontal Level 1 CALIOP resolution (330 m along track and 75 m cross track of the satellite orbit), and on a lower vertical resolution (40 equidistant vertical levels of 480 m height). Layered cloud fractions are also computed for three atmospheric layers: upper levels (between 50 and 440 hPa) middle levels (between 440 and 680 hPa) and low levels (altitudes below the 680 hPa level). In case of overlapping clouds CClow+CCmid+CChigh cloud be > 1. The cloud fraction is then interpolated on a 2° x 2° latitude/longitude grid to provide the final cloud product used in the analysis. To ensure that the values are statistically significant, only grid boxes containing more than 30 % of the maximum possible number of measurements (based on the satellite overflights) are considered in the analysis (Konsta et al., 2012).

The PARASOL instrument [POLDER-like, (Deschamps et al., 1994)] has a multi-viewing angle capability, allowing for the estimation of instantaneous monodirectional reflectance of clouds. The calibration of PARASOL is described by Fougnie et al., (2007). The calibration accuracy is within 1.5% for the 865 nm channel. Over the ocean surface, the visible directional reflectance is mostly sensitive to the solar zenith angle, to the viewing direction and to the cloud optical depth. In this analysis the reflectance observed in a single viewing direction has been selected (Konsta et al., 2015), so that it is mostly sensitive to the cloud optical depth and less to other parameters. After avoiding directions less sensitive to the optical depth (e.g. directions sensitive to glitter reflection, the backscatter and the nadir direction), the one at 865 nm which is most frequently observed by PARASOL was selected. All directional reflectance values measured by PARASOL in this direction have a spatial resolution of 6 x 6 km^2. They are then projected onto a 2° x 2° grid. The representativeness of sampling the PARASOL pixels (6 x 6 km^2) collocated with CALIOP pixels (330 m x 75 m) and averaged on a 2 ° x 2 ° grid is presented in detail in Appendix 1 of Konsta et al. (2015).

The ERA-Interim reanalysis is used in this study to estimate the lower tropospheric stability (LTS). ERA-Interim reanalysis performance is initially discussed in Dee et al., 2011. Since then, several studies have investigated the performance of ERA-Interim against radiosonde measurements for a variety of applications (Luo et al., 2020; Guan et al., 2018; Vergados et al., 2014). ERA-Interim is the predecessor of ERA5 and has been used extensively by the climate community (e.g. Pfeifroth et al., 2013; Pfahl et al., 2014, Bony et al., 2020), while biases on the surface wind speed have been identified (Belmonte Rivas and Stoffelen, 2019).

Text S2: Statistical error calculation

As it is commonly done, we defined the standard error of the mean of random variable M as the standard deviation σ of M divided by the square root of the number of degrees of freedom N'

$$\frac{\sigma}{\sqrt{N'}}.$$

Here we assumed that the spatio-temporal auto-correlations of the variables are sufficiently low so that N' is eaual to the number of samples to compute the mean. The calculations are performed over 4 years of daily values for every grid cell. No time average takes place nor the seasonal cycle is removed.

Model's short name	Modeling Center	CMIP6 Model	Resolution	Number of vertical layers below 680 hPa	Key references
IPSL	IPSL, France	IPSL-CM6A-LR	2.5º x 1.25º	29	Boucher et al., 2020 Hourdin et al., 2020
CNRM	CNRM-CERFACS, France	CNRM-CM6-1	1.4º x 1.4º	21	Roehrig et al., 2020
HadGEM3	Hadley Centre, UK	HadGEM3-GC31-LL	1.875º x 1.25º	20	Walters et al., 2019
MRI	MRI, Japan	MRI-ESM2-0	1.125º x 1.1214º	17	Yukimoto, 2019; Kawai et al. 2019
MIROC6	MIROC, Japan	MIROC6	1.4º x 1.4º	13	Tatebe et al., 2019
GFDL	NOAA-GFDL, USA	GFDL-CM4	2.5º x 2.0º	13	Zhao et al., 2018a, 2018b; Silvers et al., 2018

Table S1. CMIP6 models used in this study

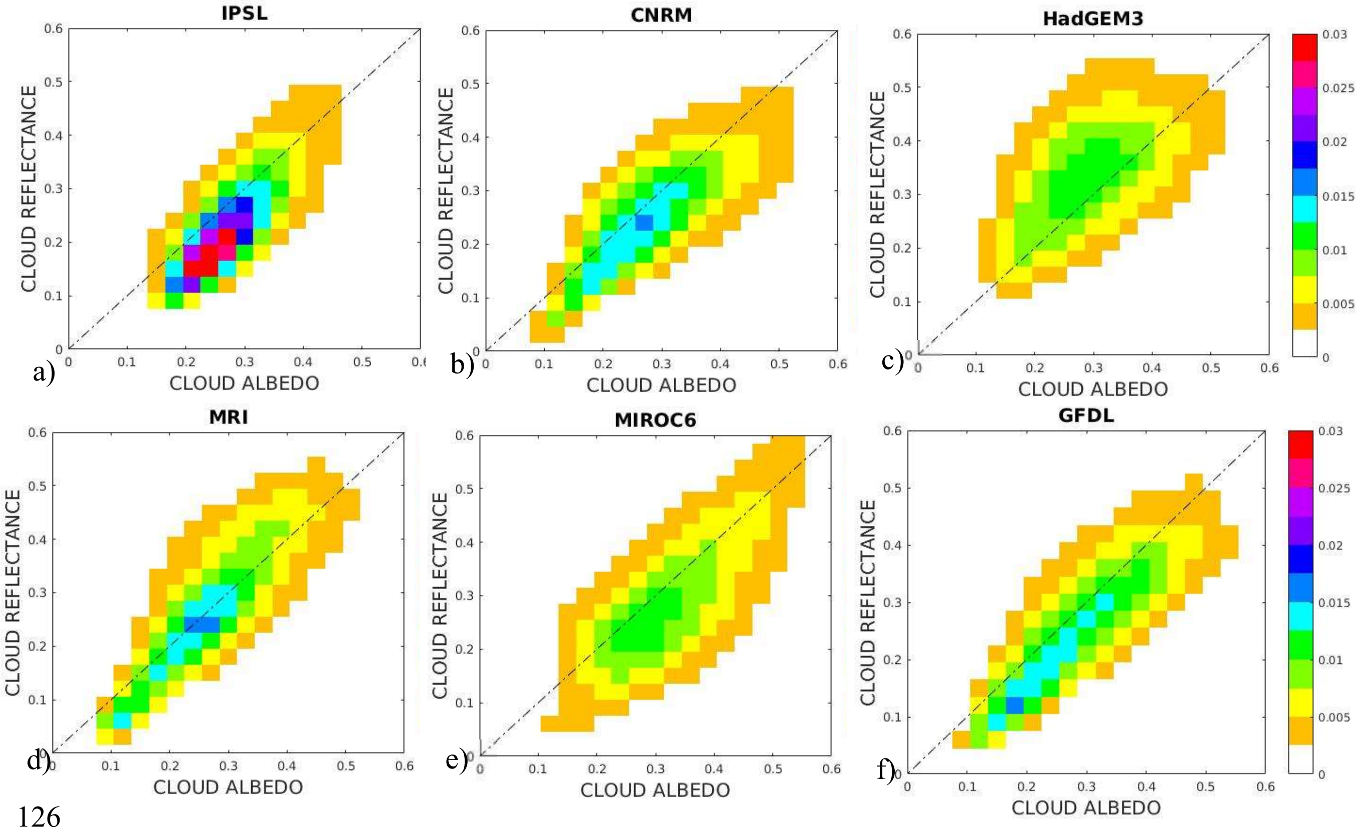


Figure S1: Relationship between daily mean Cloud Albedo and daily mean cloud reflectance over the tropical oceans, over the period 2007-2010 simulated with a) IPSL, b) CNRM, c) HadGEM3, d) MRI, e) MIROC6 and f) GFDL. Cloud Albedo is calculated from the difference of the upward SW radiation between clear sky and all sky conditions divided with the downward SW radiation for the cloudy part.

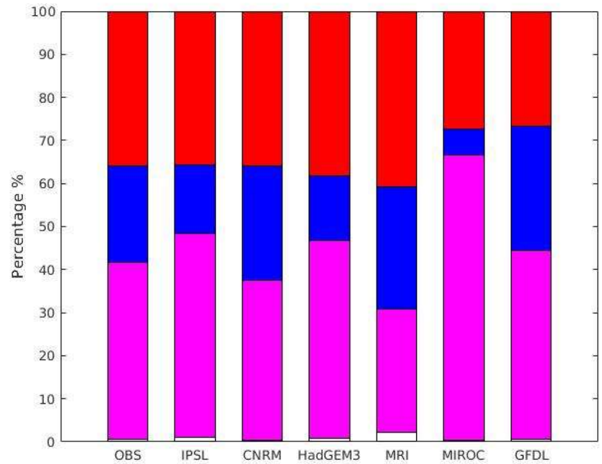


Figure S2: Probability that the sky in a 2ºx2º grid cell corresponds to one of the following situations: clear sky (CC=0, white bar), mix clouds (magenta bar), mid or high dominant clouds (blue bar) and low dominant clouds (red bar) for observations (CALIPSO-GOCCP, left column) and the models (with COSP lidar simulator) for daily values on a 4-year period (2007- 2010). Low dominant clouds are defined by using the criterion CClow > 0.9 *CC, similarly for mid and high dominant clouds.

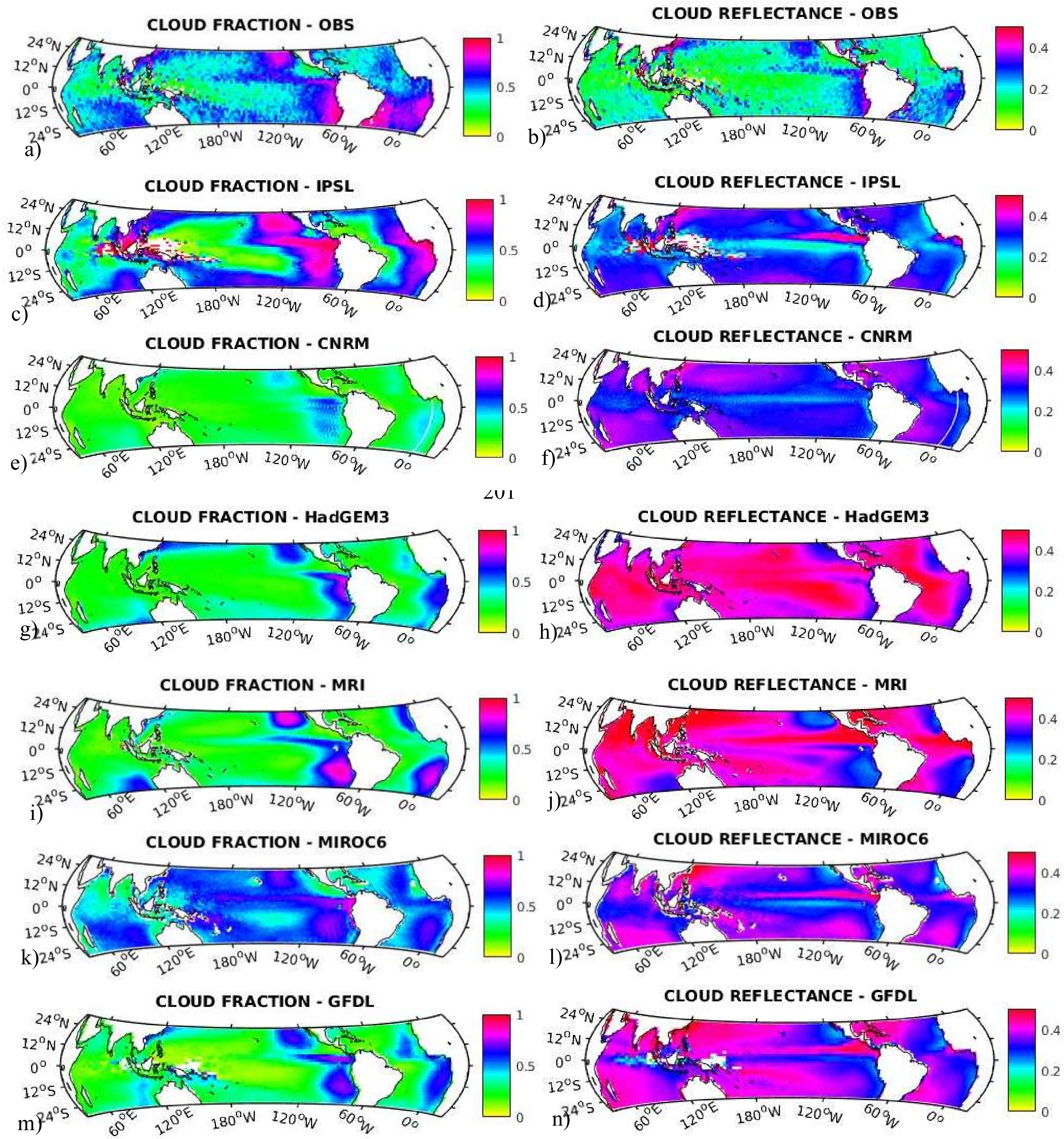


Figure S3: For situations where low-level clouds are dominant, geographical distribution of the cloud fraction (left column) a) observed (CALIPSO-GOCCP) and simulated with c) IPSL, e) CNRM, g) HadGEM3, i) MRI, k) MIRCO6 and m) GFDL, and cloud reflectance (right column) b) observed (CALIPSO-GOCCP, PARASOL) and simulated with d) IPSL, f) CNRM, g) HadGEM3, j) MRI, l) MIRCO6 and n) GFDL.

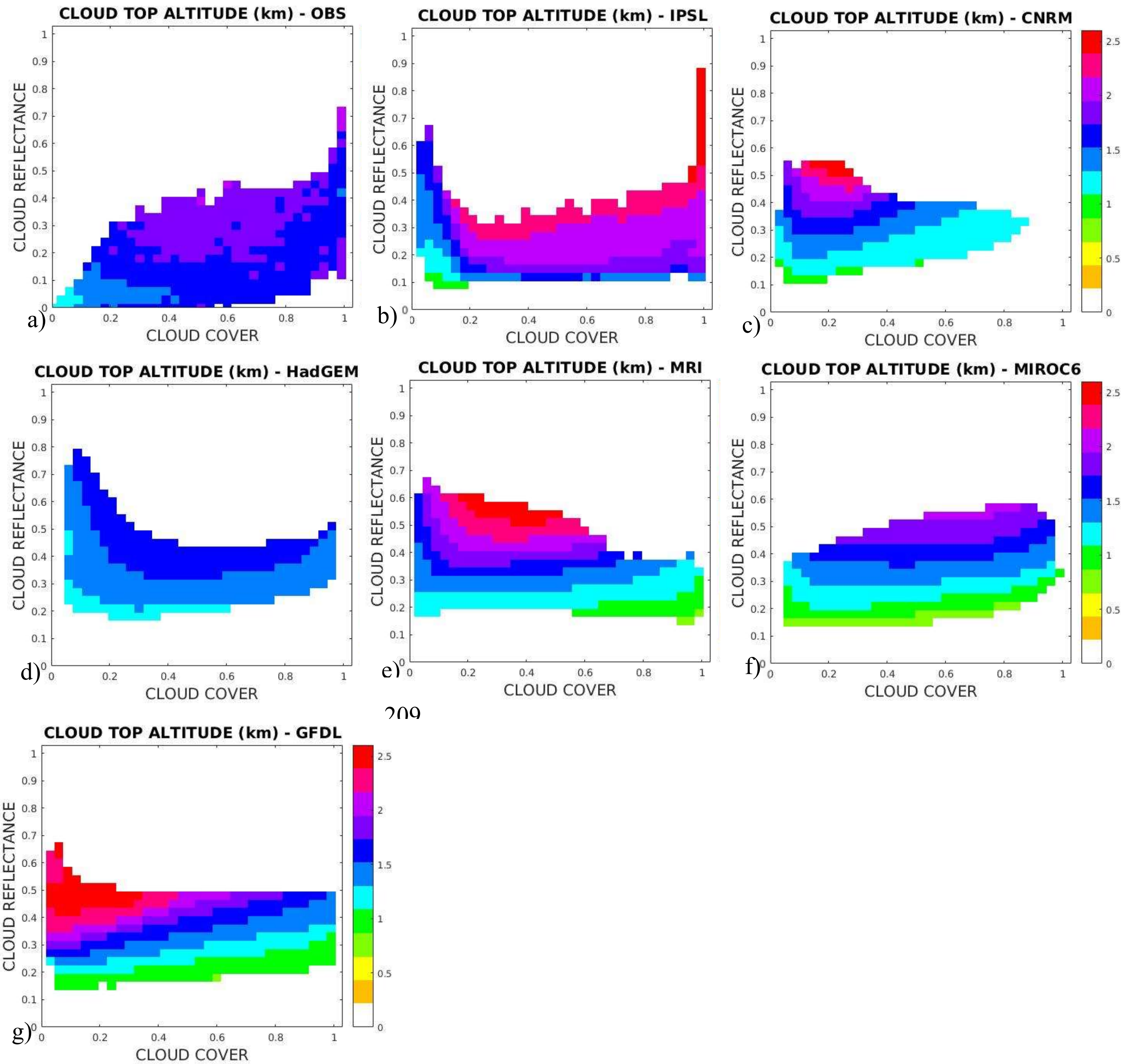


Figure S4: Color maps of cloud reflectance and cloud cover over the tropical ocean for dominant low level clouds a) observed and simulated by b) IPSL, c) CNRM, d) HadGEM3, e) MRI, f) MIROC6, and g) GFDL, where the color bar presents the mean value of cloud top altitude (in km) at each grid cell (cloud cover – cloud reflectance). Cloud top altitude is defined as the highest point over the 2°x2° grid where the sum of the cloud fraction (CF3D) from the top is greater than 10% of the cloud cover ($sumCF3D_{(from\ top)} > 10\%CC$).