Author/s:
Ahn, E;Huang, Y;Siems, ST;Manton, MJ

Title:
A Comparison of Cloud Microphysical Properties Derived From MODIS and CALIPSO With In Situ Measurements Over the Wintertime Southern Ocean

Date:
2018-10-16

Citation:
Ahn, E., Huang, Y., Siems, S. T. & Manton, M. J. (2018). A Comparison of Cloud Microphysical Properties Derived From MODIS and CALIPSO With In Situ Measurements Over the Wintertime Southern Ocean. Journal of Geophysical Research Atmospheres, 123 (19), pp.11-140. https://doi.org/10.1029/2018JD028535.

Persistent Link:
https://hdl.handle.net/11343/285006

Ahn Jenny (Orcid ID: 0000-0002-4648-3656)

A comparison of cloud microphysical properties derived from MODIS and CALIPSO with *in-situ* measurements over the wintertime Southern Ocean

Eunmi Ahn[1,2], Yi Huang[1,2], Steven T. Siems[1,2], and Michael J. Manton[1]

---

[1] School of Earth, Atmosphere and Environment, Monash University, Melbourne, Victoria, Australia

[2] Australian Research Council (ARC) Centre of Excellence for Climate System Science, Monash University, Melbourne, Victoria, Australia

Corresponding author: Eunmi Ahn, (jenny.ahn@monash.edu)

† School of Earth, Atmosphere and Environment, 9 Rainforest Walk, Monash University, Melbourne, Victoria, Australia

Key Points:

- The CALIOP and MODIS cloud phase products underestimated the occurrence of mixed-phase clouds compared to the *in-situ* observations.
- In comparison to the *in-situ* observations the MODIS effective radius is overestimated for non-drizzling clouds and, to a lesser extent, for lightly drizzling cases.
- Sub-pixel heterogeneity, cloud droplet size variability, bimodal distribution, and solar zenith angle may have measureable impacts on the MODIS r_{eff} retrieval bias.

Abstract

In-situ observations of cloud effective radius (r_{eff}), droplet number concentration (N_d) and thermodynamic phase from eleven wintertime flights over the Southern Ocean (SO) (43–45°S, 145–148°E) are compared to products from MODerate-resolution Imaging Spectroradiometer (MODIS) and Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP). The *in-situ* observations were in close alignment with A-train overpasses for a 30-minute window. For open mesoscale cellular convection (MCC), which was predominantly observed, clouds were commonly found to be intermittently drizzling, patchy and mixed-phase. Compared to the *in-situ* observations of the cloud thermodynamic phase, the CALIOP and MODIS cloud phase optical property products consistently underestimated the occurrence of mixed-phase clouds, whereas the MODIS infrared-based phase product showed a better qualitative agreement despite a frequent classification of uncertainty. The MODIS $r_{eff\_2.1}$ overestimated the *in-situ* r_{eff} for non-drizzling clouds (by ~13μm on average) and, to a lesser extent, for lightly drizzling cases. Conversely, MODIS $r_{eff\_2.1}$ underestimated the *in-situ* r_{eff} for heavily drizzling cases by ~10μm on average. The overestimation of r_{eff} is much greater than that for the stratocumulus over the Southeast Pacific shown in other studies. An examination on sub-pixel heterogeneity, droplet size variability, a bimodal distribution, and solar zenith angle suggests that all of these factors have measurable impacts on the MODIS r_{eff} bias. The MODIS N_d is largely consistent with the *in-situ* observations. However, the N_d of the two high N_d cases (closed MCC) are highly underestimated. An error analysis suggests that the N_d biases are likely a result of a compensating error effect.

1. Introduction

The low-altitude clouds over the Southern Ocean (SO) remain poorly represented in coupled global climate models and reanalysis products. These clouds are extensive (e.g. Huang *et al.*, 2012a; 2012b) and have an immediate impact on the SO energy (e.g. Haynes *et al.*, 2011) and water budget. A persistent bias in the energy budget over the SO (Trenberth & Fasullo, 2010) not only limits our ability to model the SO sea surface temperature (Bodas-Salcedo *et al.*, 2016; Kay *et al.*, 2014) and ocean-atmosphere interactions, but also limits our ability to simulate the Antarctic climate (Zelinka & Hartmann, 2012) and wider climate phenomena. Accordingly, the U.S. National Science Foundation (NSF) Advisory Committee for Geosciences has identified the need to better understand the cloud processes over the SO under one of their current research priorities (NSF, 2014).

Over the past decade, observations from the A-train satellite constellation (Stephens *et al.*, 2002) which includes Aqua, CloudSat (Im *et al*. 2005), and Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO; Winker *et al*. 2009) have been widely employed to better understand the unique macrophysics and microphysics of SO clouds and their precipitation (e.g. Bennartz 2007; Hu *et al.*, 2010; Huang *et al.*, 2012a; Muhlbauer *et al.*, 2014). The unique nature of SO clouds and the lack of verification studies over this region, however, limit our confidence in these products. Huang *et al.* (2015a) illustrated substantial differences in the cloud-top thermodynamic phase classifications from the MODIS, CALIPSO and a merged radar-lidar product (DARDAR Mask; Delanoë & Hogan, 2010) over the remote SO. The uncertainty between these different products can be exacerbated during the winter months, if the observations, such as those from the MODIS, are sensitive to the solar zenith angle (SZA) (Wood *et al.*, 2012; Grosvenor & Wood, 2014; Di Girolamo *et al*., 2010). When the SZA is small the clouds may appear smoother due to "radiative smoothing" from the net horizontal transport of sunlight from thicker clouds to thinner clouds (e.g. Zuidema and Evans, 1998) and when the SZA is large, clouds may appear rougher due to the side illumination/shadowing of cloud top bumps (e.g. Várnai and Marshak, 2003).

Few *in-situ* observations of clouds exist over the SO to help evaluate these satellite products, primarily because of the harsh environment of the Southern Ocean storm track and its remote location. Past field campaigns such as the Southern Ocean Cloud Experiments (SOCEX, Boers *et al.*, 1996; 1998) and the Aerosol Characterisation Experiment (ACE-1, Bates *et al.*, 1998) pre-date the A-train observations.

More recently, Ahn *et al.* (2017, hereafter A17) detailed the *in-situ* observations of twenty flights made over the SO to the west and south of Tasmania during the winter season (June - October) over a three-year period (2013 - 2015) to investigate the low-altitude cloud microphysical properties. Consistent with the early findings of SOCEX I, A17 found the pristine liquid-phase clouds commonly have very low cloud droplet number concentration, N_d (10 - 40 cm^{-3}) and large effective radius, r_{eff} (11 - 16 μm) with an exception when overcast closed mesoscale cellular convection (MCC) (Wood & Hartmann, 2006) was encountered. For all in-cloud observations from the 20 flights, ice crystals were frequently encountered (38.5%), primarily in a mixed-phase regime (36.7%). Near 50% of all liquid cloud samples were drizzling intermittently; solid patches of either drizzle or non-drizzle with a duration greater than 10s (~ 100m) were rarely observed.

The presence of a drizzle mode together with small cloud droplets near cloud top, presumably caused by evaporation, has previously been found to affect the MODIS spectral retrievals by introducing vertical inhomogeneity in the cloud droplet size distribution (e.g., Nakajima *et al.*, 2010a, 2010b). Using MODIS, Bennartz (2007) estimated that 89% of the clouds over this region are likely to be drizzling. Wang *et al.* (2015) found that even CloudSat underestimated the frequency of light precipitation over the SO. The frequent presence of ice in these clouds adds further uncertainty to the retrievals. Huang *et al.* (2017) examined one of the A17 flights in detail, exploring the potential for secondary ice production (specifically, the Hallett-Mossop process) to account for the widespread presence of mixed-phase clouds in open MCC.

Beyond the microphysics, the structure of these open mesoscale cellular clouds may also present challenges in employing remote sensing observations with confidence. Previous work has shown that the cloud r_{eff} for broken clouds overlaying a dark surface is overestimated because of the reduced radiance at shortwave/midwave infrared wavelengths due to the clear-sky contribution and to the use of plane-parallel forward models (e.g., Zhang & Platnick 2011; Zhang *et al*., 2012).

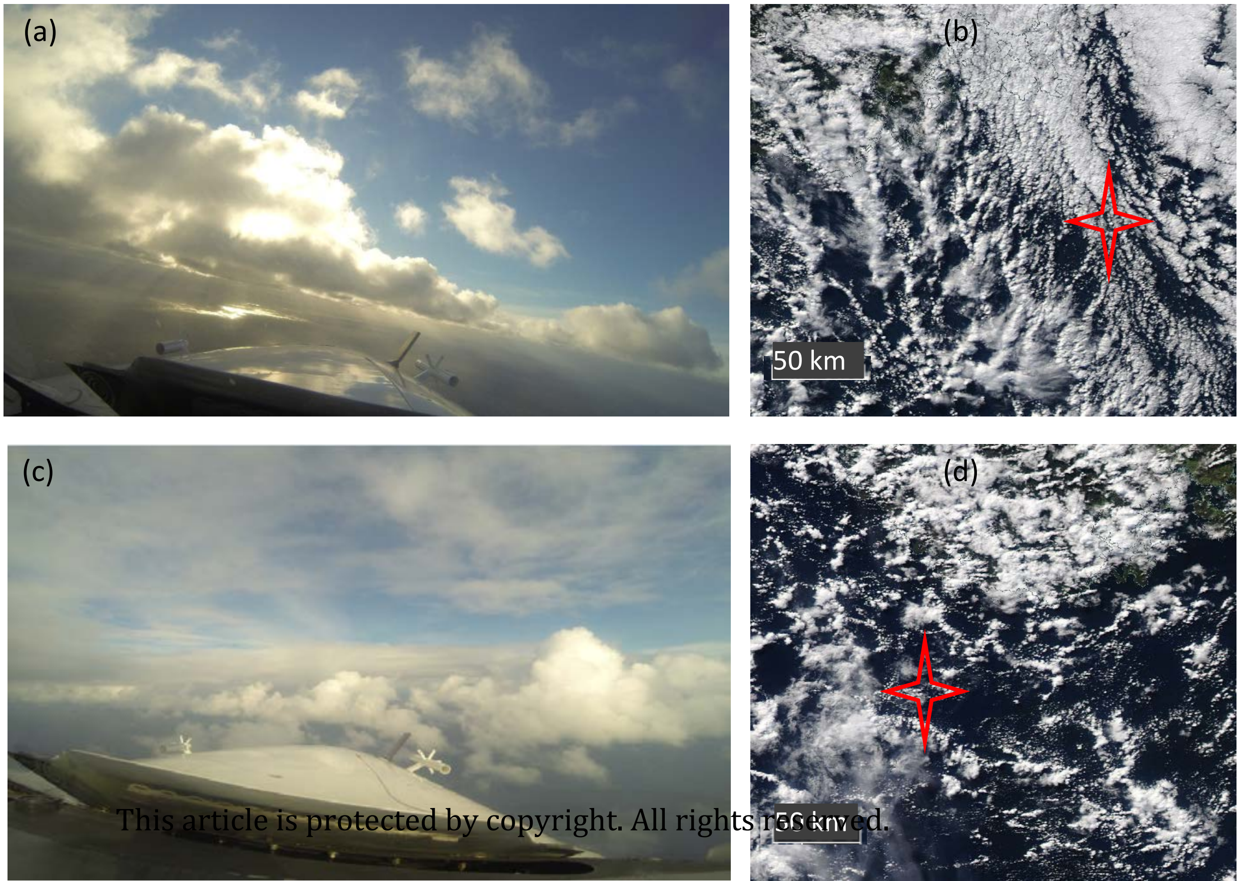
Wolters *et al.* (2010) reported that the overestimation of MODIS r_{eff} for a broken cloud regime over the Atlantic Ocean is up to 2.5 times higher than for inhomogeneous overcast clouds. In contrast, Coakley *et al.* (2005) found that r_{eff} for clouds in partly cloudy pixels is generally smaller than that for clouds in nearby overcast pixels. Platnick *et al.* (2015) also showed that for a certain month (April 2005), the Aqua MODIS (collection 6; hereafter C6) global monthly mean 1° gridded r_{eff} for the partly cloudy pixels was significantly smaller than that for the overcast pixels. The confidence of satellite retrievals of effective radius and cloud droplet number concentration over open MCC needs to be further examined.

Given the unique nature of SO clouds, it is a necessary exercise to directly compare/evaluate satellite retrievals against *in-situ* observations, as has been done elsewhere. For example, Painemal & Zuidema (2011) and Min *et al.* (2012) employed *in-situ* observations made over the Southeast Pacific during the Variability of the American Monsoon Systems (VAMOS) Ocean-Cloud-Atmosphere-Land Study Regional Experiment (VOCALS-REx) to evaluate MODIS retrievals products of effective radius (r_{eff}) and cloud droplet number concentration (N_d), as well as liquid water path and cloud optical thickness (COT). These passive retrieval products are frequently used for broader climate studies given their large spatial and temporal coverages. The r_{eff} derived from the MODIS retrieval algorithm has been found to overestimate *in-situ* observations by 15 - 20% (Painemal & Zuidema 2011), by up to 1.75 μm (Zheng *et al.*, 2011), by up to 2.3 μm (Min *et al.*, 2012) and by 13% (King *et al.*, 2013). The MODIS-derived N_d, on the other hand, has been found to agree reasonably well with *in-situ* observations (Painemal &

Zuidema, 2011; Min *et al.*, 2012). It is noted, however, that the clouds examined in all these abovementioned studies are semi-permanent, stratified stratus or stratocumulus, detected under moderately small SZA (15^o - 50^o). In contrast, the clouds examined in our study are typically fast-moving, broken, and highly inhomogeneous, commonly detected under high SZA (50^o – 74^o) (Figure 1 and Table 1). Boeke *et al*. (2016) revealed that the overcast assumption used in the MODIS cloud property retrievals contributed to the overestimate of the effective radius and cloud top temperatures (Coakley *et al*., 2005).

Eleven of the twenty flights detailed in A17 were designed to closely align with A-Train overpasses. This research directly compares *in-situ* cloud thermodynamic and microphysical properties (i.e. cloud thermodynamic phase, $r_{eff,}$ and N_d) against MODIS and CALIOP products over the remote Southern Ocean. Further, we examine the sensitivity of the retrieved products to operational constraints (i.e. a high SZA) and cloud regimes (i.e. drizzle, mixed-phase clouds and partially cloudy pixels). Finally, we contrast our findings against similar studies over the Southeast Pacific (i.e. Min *et al.*, 2012; King *et al.*, 2013) to better understand the skills and limitations of these products over the SO.

Figure 1. Cloud images taken by a GoPro camera onboard the Cessna Conquest aircraft and observed by the MODIS on 14 June 2013 (a & b) and 03 Sep 2014 (c & d). Images (a) and (c)



were taken just before the A-Train overpass when the aircraft was closely aligned with the CALIPSO ground track (04:20 and 04:19 UTC, respectively). The corresponding MODIS visible images, (b) and (d), indicate the location of the aircraft with a red cross.

2. Data and methodology

2.1 Flight overview and aircraft measurements

A full description of the meteorology and microphysical observations of the twenty wintertime flights over the Southern Ocean (43 – 45°S, 145 – 148°E) is detailed in A17. The eleven research flights employed in this study (Table 1) were designed to closely align with A-Train overpasses. In addition to having a reasonable forecast cloud amount, the research flights were conducted under the 'baseline' conditions where the clouds had not been influenced by any upwind terrestrial sources of aerosols for a minimum of 3 days. These conditions were verified by the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model (Draxler and Hess, 1998) with the employment of the Global Data Assimilation System (GDAS) reanalysis (A17).

The ideal sampling module of a flight mission consisted of two soundings (20 - 30 min each before and after the A-Train overpass), one above-surface leg (8-10 min), one below-cloud leg (8-10 min), one in-cloud leg (8-10 min), and one above-cloud leg (8-10 min). More complicated modules also include conducting 'sawtooths' near cloud tops and/or multiple in-cloud legs for multilayered clouds. In practice, however, the cloud fields encountered were highly complex; patchy, fast-moving, inhomogeneous, and more often than not, multilayered cloud fields were most common (Figure 1). As such, it was necessary to make real-time altitude adjustments to maximize the number of samples in larger-size clouds along the designated flight track. The adjusted flight path could therefore be a 'consistent-level' pattern (9 flights) or a 'v-line' pattern (2 flights), as illustrated in Figure S1 (a) and (b), respectively.

All flight data were recorded and analysed at a temporal resolution of 1 Hz, which corresponds to a spatial resolution of approximately 100 m based on the typical aircraft true air speed. A number of these flights have previously been analysed and employed to evaluate numerical simulations of SO clouds (Huang *et al.*, 2014; 2015b). The N_d and r_{eff} of liquid water clouds were derived from measurements with a Droplet Measurement Technologies (DMT) Cloud, Aerosol and Precipitation Spectrometer probe (CAPS, Baumgardner *et al.*, 2001), which consists of a hot-wire liquid water sensor (LWC-100), a Cloud Aerosol Spectrometer (CAS) that measures particles from 0.5 to 50 μm, and a Cloud Imaging Probe (CIP) that is used for larger particles from 25 μm to 1.55 mm. r_{eff} (CAS + CIP) indicates that both the CAS and CIP are employed to calculate r_{eff}. For all calculations, we omit the measurements from the first three bins of the CAS (bin sizes less than ~ 0.68 μm) and the first two bins of the CIP (bin sizes less than ~ 62.5 μm), given the high measurement uncertainties in these size ranges. The 'System for Optical Array Probe Data Analysis (SODA)' software developed at the National Center for Atmospheric Research (NCAR) was used to analyze the microphysical properties measured by the CIP. Further details of the CIP data processing are described in the supporting information of A17 and Huang *et al*. (2017).

Table 1 Summary of the eleven Flights

Flight Date	Coincided Start [UTC]	END	Lat (mean)	Lon (mean)	MCC[1)] type	SZA [°]	Drizzle Fraction	Rain Rate [mmhr^{-1}]	Ice Fraction	*In-situ* r_{eff} (std.) [μm]	$r_{eff\_1.6}$ (std.) [μm]	$r_{eff\_2.1}$ (std.) [μm]	$r_{eff\_3.7}$ (std.) [μm]	*In-situ* r_{eff} variations [μm]	*In-situ* N_d [cm^{-3}]	MODIS N_d [cm^{-3}]	CM_SPI[2)] band1/ band2 [%]	r_{eff} Retrieval Failure Rate
20130614	04:20	04:50	-43.66	148.3	open MCC	74	0%	0.001	1.3%	9.6 (1.5)	15.6 (7.4)	18.5 (4.9)	16.1 (3.3)	16	31	24	15/19	8%
20130628	04:19	04:49	-43.89	145.2	open MCC	74	86%	0.375	8%	23.3 (10.5)	17.1 (7.5)	23.4 (4.6)	18.5 (3.2)	45	23	21	12/17	27%
20130707	04:13	04:43	-44.59	146.98	open MCC	73	0%	0.001	85%	16.4 (2.7)	16.1 (7.2)	20.4 (5.2)	15.7 (3.3)	16	18	45	26/39	32%
20130723	04:12	04:42	-44.05	146.94	closed MCC	71	0%	0.000	0%	8.3 (1.1)	14 (6.2)	12.9 (3.1)	12.6 (2)	13	91	66	12/15	7%
20130806	04:30	05:00	-42.64	143.63	open MCC	66	39%	0.530	12%	16.9 (17)	15.2 (7.7)	20.2 (5.4)	17.1 (3.9)	101	33	24	23/32	35%
20130815	04:19	04:49	-42.9	144.64	disorganised MCC	64	8%	0.004	43%	12.7 (3)	22.1 (6.3)	25.1 (3.2)	20.8 (2.6)	24	10	7	14/20	12%
20131011	04:12	04:42	-44.58	147.08	disorganised MCC	49	1%	0.002	54%	9.3 (2.9)	18.7 (6.4)	22.2 (4.4)	18 (2.9)	31	21	14	15/19	13%
20140903	04:19	04:49	-44.01	145.28	open MCC	60	14%	0.097	24%	17.8 (18.2)	18.2 (7.3)	22.6 (4)	18.3 (2.9)	102	14	11	13/17	13%
20140912	04:27	04:57	-44.86	147.04	open MCC	58	40%	0.043	18%	16.9 (6.4)	17.9 (6.9)	20.9 (4.4)	16.1 (3.3)	38	15	13	7/9	16%
20150830	04:11	04:41	-44.2	146.78	open MCC	61	88%	0.980	3.1%	35.4 (14.5)	16.4 (7.2)	25.8 (3)	23 (3)	41	19	9	13/17	37%
20151001	04:12	04:42	-44.3	146.46	closed MCC	52	0%	0.001	0.8%	7.5 (4.9)	15.1 (8)	20.3 (5)	12.7 (1.6)	65	74	20	15/18	19%

Note: the eleven flights from 2013 – 2015 were closely aligned with A-Train overpasses (i.e. MODIS and CALIPSO ground tracks). The average cloud properties are for liquid only (except ice fraction) according to both the *in-situ* and MODIS observations. The MODIS r_{eff} is retrieved from both the overcast and PCL pixels. The standard deviation of r_{eff} is indicated in parenthesis. Drizzling flights are highlighted in light blue.

Cloud was defined by a total water content (TWC) threshold of 0.01 gm^{-3} (e.g. Korolev *et al.*, 2003; Wood & Field, 2011; Boutle *et al.*, 2014) as measured by the Science Engineering Associates Inc. (SEA) WCM-2000 Multi Element Water Content System (Lilie *et al.*, 2005). Liquid water clouds were separated from mixed-phase and ice clouds when the ice-water fraction, μ_{ice}, was less than 0.1. μ_{ice} is defined as W_{ice} / (W_{liq} + W_{ice}) with liquid water content (W_{liq}) and ice water content (W_{ice}) derived from the measured liquid water content (LWC) and TWC from the WCM-2000 hot-wire elements (2 mm diameter and 4 mm scoop-shaped, respectively). In this study, a drizzling cloud sample is defined when the LWC in liquid cloud (TWC > 0.01 gm^{-3}) from droplets (drop size greater than 112.5 μm) as measured by the CIP is greater than 0.005 gm^{-3} (A17). Rain rate is calculated using the fall speed of those droplets (Pruppacher & Klett, 2010). Details of the rain rate for the flights are found in Table 1.

2.2 MODIS products

MODIS level-2 (L2) cloud top and optical product (earth science data set MYD06 from Aqua) from Collection 6 (C6) is employed for retrieval cloud properties with a 1 km x 1 km spatial resolution at nadir. Notable changes to the C6 cloud property retrieval algorithms (Platnick *et al.*, 2017) pertinent to this study include: 1) improved shortwave-derived cloud thermodynamic phase; 2) separate cloud effective radius retrieval data sets from the 1.6, 2.1, and 3.7 μm channels (different from C5 where only the differences with respect to the 2.1 μm channel are reported); 3) processing partly cloudy pixels and cloud edges; 4) a new cloud radiative transfer look-up table including for more optically thin clouds; and 5) retrieval failure metrics that provide diagnostic information for the failure. A comprehensive description of the collection 6 reprocessing for Aqua MODIS can be found in Platnick *et al.* (2015; 2017).

It is necessary to determine the cloud-top thermodynamic phase initially to remove any mixed-phase and glaciated cloud pixels before moving on to analyze N_d and r_{eff}. MODIS C6 provides two cloud phase products: a tri-spectral infrared (IR) based product (Baum *et al.*, 2012), which reports both daytime and night-time cloud-top phase using brightness temperature and emissivity ratios from three different band pairs (8.5, 11, and 12 μm, Baum *et al.*, 2012), as well as a daytime-only product (cloud phase optical property or CPOP) that is based on a combination of visible, shortwave IR and IR channels (Marchant *et al.*, 2016). Compared to C5, the CPOP algorithm has been completely rewritten to improve the phase discrimination skill for a variety of cloudy scenes (e.g., thin/thick clouds, over ocean/land/desert/snow/ice surface, etc.) (Marchant *et al.*, 2016). The CPOP algorithm uses a simple voting methodology that takes into account all available phase-related information such as cloud-top temperature, r_{eff}, 1.38-μm cloud mask test, and IR phase decision and provides the final cloud phase decision (Marchant *et al.*, 2016).

The MODIS clear sky restoral (CSR) algorithm provides four categories for cloud mask; overcast (CSR = 0), partially cloud-covered (CSR = 1), cloud edge (CSR = 3), and clear sky (CSR = 2). C6 has a newly defined modifier "partly cloudy" (PCL) which indicates the pixel being either 'CSR = 1' or 'CSR = 3'. Earlier versions did not distinguish between cloud edge and partially cloud-covered. Cloud edge pixels were excluded in C5 MODIS retrievals, while C6 attempts to retrieve cloud optical and microphysical properties on these pixels (Platnick *et al.*, 2017).

Cloud top and optical properties examined in this study include cloud-top temperature (CTT), cloud-top pressure (CTP), and cloud-top height (CTH), and cloud optical thickness (COT). Cloud effective radius (r_{eff}) is the major cloud optical and microphysical property examined. Cloud phase is investigated with the cloud phase infrared (IR) and cloud phase optical property (CPOP) products. The science dataset (SDS) provides critical information for analysis including solar zenith angle (SZA), cloud multi-layer flag (CMLF), and cloud mask sub-pixel heterogeneity index (CM_SPI) that is used to examine cloud heterogeneity and retrieval failure metrics. The Level-1 geolocation products (MYD03) are also used to determine the longitude and latitude for MODIS pixels at a 1 km x 1 km spatial resolution.

Equation (1) (Grosvenor & Wood, 2014; Huang *et al.*, 2016) is employed to calculate N_d from MODIS observations. This equation has been modified from the original one used in Boers *et al.* (2006) and Bennartz (2007). The calculation assumes that clouds are stratiform marine boundary layer clouds and developed adiabatically.

$$N_d = \frac{2\sqrt{10}}{\mathrm{k}\pi Q^3}\left(\frac{c(T,P)COT}{\rho_l {r_{eff}}^5}\right)^{\frac{1}{2}} \quad (1)$$

$$k = \left(\frac{r_v}{r_{eff}}\right)^3$$

where r_{eff} and r_v are the cloud top effective and volume mean radius, respectively. Q is scattering efficiency and 2 is taken as droplet radii are much larger compared to the wavelength of light (Bennartz, 2007). A constant 0.8 is adopted for k as in Bennartz (2007), Grosvenor and Wood (2014) and Huang *et al.* (2016). ρ_l is the liquid water density (1 kg m^{-3}), and c (in kg m^{-4}) is the temperature- and pressure-dependent condensation rate. c is set to 80% (constant), meaning 80% of the adiabatic value is taken. The $c^{1/2}$ curve in the supplement material of Huang *et al.* (2016) is used. A comparison of $c^{1/2}$ used in this study has been compared with values derived from other methods, with negligible differences found for temperatures below 283 K (Huang *et al.*, 2016). Therefore, potential errors associated with $c^{1/2}$ are expected to be small.

2.3 CALIOP products

The CALIOP payload aboard CALIPSO was launched in April 2006 and consists of a near-nadir viewing two wavelength polarization-sensitive lidar. It measures the backscatter at 532 and 1064 nm and linear depolarization ratio at 532 nm. The depolarization of the lidar backscatter signal is found to be highly efficient in discriminating ice and liquid phase clouds due to their different depolarization behaviors (Hu *et al.*, 2009). This study uses CALIPSO version 4.1 (V4) lidar level 2 cloud layer products (Vaughan *et al.*, 2017). Layers in the 1-km horizontally-resolved product are identified as clouds by the Cloud Aerosol Discrimination (CAD) algorithm. Layers identified as clouds are further classified into water, randomly-oriented ice (ROI), horizontally-oriented ice (HOI) or unknown phase categories. Among the improvements in V4 over the previous version are cloud subtyping, ice-water phase determination, as well as the elimination of known retrieval artifacts. The data quality summary is found in https://www-calipso.larc.nasa.gov/resources/calipso_users_guide/.

2.4 Collocation and comparison methodology

The A-train satellites move at a far greater speed than the aircraft and are temporally collocated for only an instant. Further, unlike those semi-permanent, stratified stratus or stratocumulus sampled by VOCALS-Rex, the dynamic, fast-moving, patchy, and highly inhomogeneous clouds commonly encountered during our flights made a space-time collocation extremely difficult. Moreover, due to the different sensitivities and signal penetration depths of the space-borne sensors, it was practically impossible to sample the exact clouds that were directly detected by different satellite sensors within any reasonable timeframe (detailed in Section 2.1). Therefore, despite some obvious caveats, our collocations and comparisons are only performed in a statistical sense, with the assumption that the *in-situ* samples are representative of the range of cloud characteristics at different stages of a cloud lifetime for a given synoptic setting. Likewise, the satellite observations taken from a broader sample area are assumed to represent the main characteristics of the cloud field sampled by the aircraft. To make the comparison statistically robust, both the satellite and *in-situ* observations are considered over an extended period. Aircraft observations are extended to a 30-minute window centered at the time of the A-train overpass. This 30-minute window is, typically, a much shorter period of time than the full records used in A17. Given the complexity of the SO open MCC, we limited our time window to 30 minutes to minimize any cloud evolution arising from the high wind speeds (22±8.7 ms^{-1}) and strong vertical wind shear (e.g. Hande *et al.*, 2012).

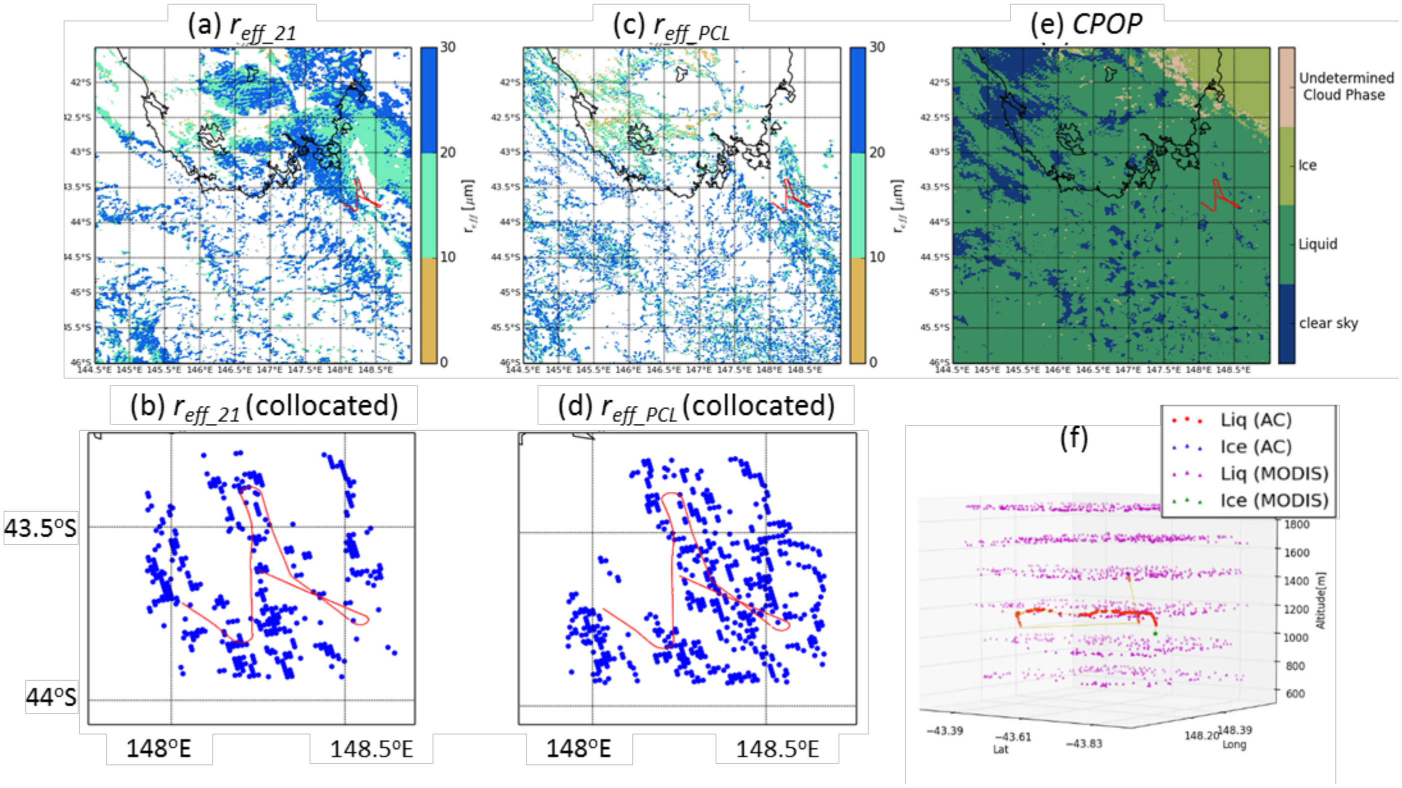
Horizontally, MODIS pixels up to 10 km away from the aircraft are considered for comparison with the *in-situ* observations (Figure 2 (b)). Vertically, the MODIS pixels within ±50 hPa and ±500 m of the 30-minute aircraft track are considered to be collocated. Menzel *et al.* (2008)

estimate an error in CTH/CTP of up to 1 km (or 50 hPa) using airborne lidar, CALIOP and High Resolution Infrared Radiometer Sounder (HIRS, Wylie & Menzel, 1999), Holz *et al.* (2008) estimated the error in CTH to be 1.4 ± 2.9 km using CALIOP. The spatiotemporally coincident images of a flight track and MODIS retrievals are shown in Figure 2 (b), (d) and (f) with MODIS 'overcast', 'PCL' and 'overcast + PCL' pixels. Qualitatively, our results were found to be largely insensitive to the ±500 m vertical distance window.

For a comparison of cloud thermodynamic phase with CALIPSO, the spatial collocation is calculated based on the ~ 135s overpassing time lag of CALIPSO against that of the MODIS. The collocation calculation is limited to ±15 min from the CALIPSO overpass time. Because the width of CALIOP footprint is only ~1 km, the sample size of the CALIOP is much smaller than that of the MODIS.

Figure 2

Composite images of an aircraft flight track overlaid on Aqua MODIS retrievals for the 20130628 case. The flight track is identified in red on the maps. (a) and (c) are MODIS retrievals of r_{eff_21} for overcast and r_{eff_21} PCL pixels, respectively. (b) and (d) are the spatiotemporally collocated MODIS pixels and aircraft ground track. Blue dots are the MODIS cloud pixels. (e) The MODIS cloud phase optical properties (CPOP) with aircraft track. (f) The 3-D image of the thermodynamic phase from the *in-situ* observations and MODIS CPOP (overcast + PCL) retrievals.



3. Cloud thermodynamic phase

Table 2 summarises the compositions of cloud thermodynamic phase from aircraft measurements and the retrievals from the CALIOP, MODIS IR and MODIS CPOP products for each flight. Although the table is presented quantitatively, differences in the sampling volume and measurement platforms limit any direct comparison. A qualitative analysis may be more appropriate. The differences in the sampling volume for all eleven flights are provided in the 'supporting information' (Figure S2.)

As in A17, μ_{ice} is employed to define the level of glaciation within a cloud sample. Following Korolev *et al.* (2003), liquid clouds and mixed-phase/ice clouds are sorted by $\mu_{ice} \leq 0.1$ and $\mu_{ice} > 0.1$, respectively. The frequency of heavily glaciated cloud samples ($\mu_{ice} > 0.5$) is also noted. The two closed MCC flights record little, if any, ice. Conversely, three of the eleven flights (20130707, 20130815 and 20131011) contain ice throughout with over 40% of the cloud samples recording $\mu_{ice} > 0.1$. These flights are nominally called 'widely iced.' A fourth flight (20140903) was observed to be heavily glaciated 12.5% of the time. The CIP images where ice particles are recorded for these four flights are illustrated in Figure S3 in the 'supporting information'.

3.1 CALIOP

The CALIOP cloud phase product was available for only nine of the eleven flights as indicated, and the CALIOP vertical cross-sections for these nine flights are presented in Figure 3. Both the full profile and a zoomed profile of the targ*et* altitudes are presented. Focussing on the lower altitude clouds (between the red lines indicated in Figure 3) that were sampled by the aircraft, it is evident that liquid water was the primary cloud phase reported by the CALIOP product for all nine flights. For two of the nine flights (20130723 and 20131011), liquid water was the only cloud phase reported. This is fully consistent with the *in-situ* observations for flight 20130723, where the clouds were at low altitude and the temperature was above freezing throughout the cloud, as detailed in A17. The sampled clouds were observed as closed MCC. For flight

20131011, however, multi-layered clouds were present. CALIOP observed cloud top heights at altitudes of 4000 to 4500 m. These clouds were at temperatures far below freezing (-15 ºC, Figure S_Vreff) and widespread ice was recorded by *in-situ* observations often with μ_{ice} greater than 0.5 at this altitude. The majority of the *in-situ* observations were actually made at a much lower elevation (~1500 m).

Table 2. Compositions of Cloud Phase for the Eleven Flights

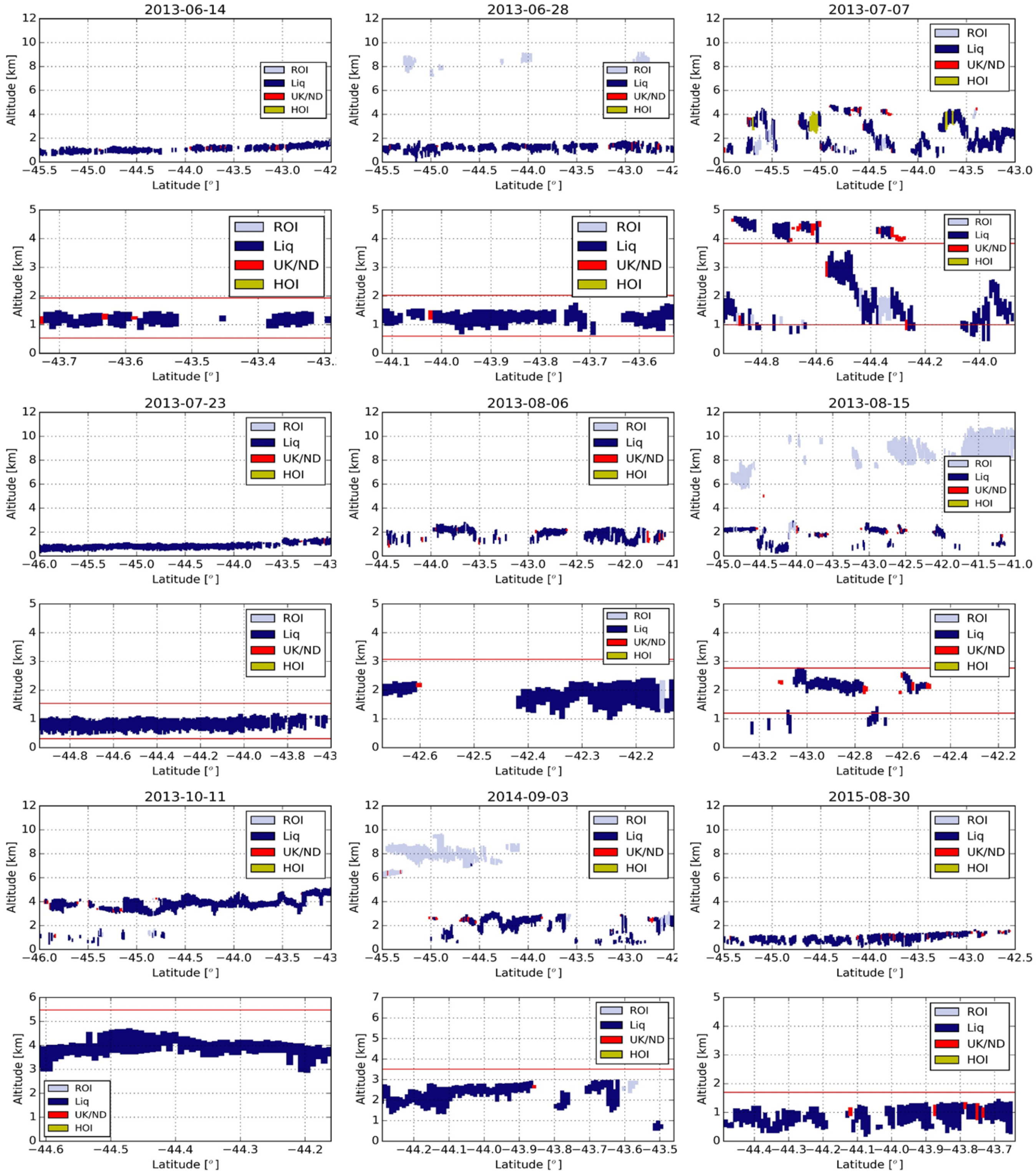
Flight Date	AC			CALIOP			Single layer						All clouds (single layer + multi layer)					
							IR			CPOP			IR			CPOP		
	samples	Ice [%] $\mu_3 > 0.5$	Ice [%] $\mu_3 > 0.1$	PP	Ice [%]	UD [%]	PP	Ice [%]	UD [%]	PP	Ice [%]	UD [%]	PP	Ice [%]	UD [%]	PP	Ice [%]	UD [%]
20130614	451	0.0	1.3	34	0.0	5.9	1302	4.5	0.0	1302	0.0	0.0	1567	10.3	0.0	1567	0.0	0.0
20130628	1274	0.3	8.0	56	0.0	1.8	533	8.3	0.1	533	2.1	0.0	1351	10.4	0.1	1351	0.9	1.2
20130707	873	56.0	85.2	51	13.7	2.0	965	31.6	24.8	965	10.7	3.4	3132	43.4	15.0	3131	15.3	7.1
20130723	538	0.0	0.0	142	0.0	0.0	1411	2.1	0.0	1411	0.0	0.0	1464	5.3	0.0	1464	0.0	0.0
20130806	387	1.0	11.6	41	2.4	2.4	680	21.3	9.9	680	6.2	0.6	1193	26.2	9.9	1185	6.3	4.4
20130815	851	9.3	42.7	51	0.0	17.6	1468	81.2	5.4	1468	44.0	3.2	2518	77.3	5.4	2518	25.8	8.0
20131011	291	28.5	54.0	51	0.0	0.0	1423	25.3	42.0	1423	3.5	0.2	4091	63.9	17.8	4091	1.8	3.1
20140903	392	12.5	24.0	71	5.6	1.4	2072	23.3	8.2	2072	2.9	0.4	4107	26.8	5.6	4037	2.0	2.2
20140912	533	0.9	17.6	-	-	-	795	2.9	0.0	795	0.0	0.0	1174	2.3	0.0	1150	0.0	0.0
20150830	893	0.1	3.1	84	0.0	7.1	249	0.0	0.0	249	0.0	0.0	1898	0.1	0.0	1898	0.0	0.0
20151001	261	0.0	0.8	-	-	-	1859	4.2	4.7	1859	0.0	0.0	3437	4.5	3.6	3437	0.0	0.0
CTT [^{o}C]			-5					-2.5			-3			-6.1			-5.4	
CTH [m]			1941					2071			2100			2698			2469	

μ_3: Cloud Phase Coefficient, PP (pixel population), UD (undetermined)

Note: the composition of cloud phase was discriminated by aircraft measurements, CALIOP retrievals, and MODIS IR and CPOP product. The MODIS cloud phase is separated with single-layer cloud and all clouds including multi-layered clouds determined by MODIS “Cloud Multi-layer Flag” science data set (SDS). Cloud top temperature (CTT) and cloud top height (CTH) are the average values for the ice phase clouds for all cases. The three ‘widely iced’ cases are highlighted.

Figure 3

Vertical profiles of the CALIOP cloud phase for nine flights from 2013 – 2015. Two CALIOP images are shown for each flight; the first one with high altitude and the second zoomed in to target the coincident area. The red lines on the second image indicate the coincident altitude range. ROI, HOI and UK/ND indicate “randomly-oriented ice”, “horizontally-oriented ice”, and “unknown / not determined”.



Flight 20130707 was another ‘widely iced’ flight and has similarities with 20131011. This flight also observed clouds at ~2000m and at temperatures below freezing. Once again, liquid cloud phase is dominant in CALIOP retrieval although a few randomly oriented ice (ROI) and unknown/not determined (UK/ND) classes were recorded. It is interesting to note that a section of ROI is found directly underneath the overlying clouds. Huang *et al* (2012a) noted that the CALIOP cloud phase classification of low level clouds over the SO was sensitive to the presence of overlying ice clouds. The third ‘widely iced’ flight, 20130815, is once again predominantly seen as a liquid by CALIOP, although some UK/ND is observed. No ROI was reported by CALIOP, contrary to the *in-situ* observations. It is evident that even when ice is widely present through the low-level clouds over the SO, CALIOP predominantly returns liquid as the cloud phase. The remaining five flights have a relatively low occurrence of mixed-phase cloud samples ($\mu_{ice} > 0.1$). The fraction of ice (either ROI or HOI) and UK/ND for all nine flights are provided in Table 2. As discussed, liquid is dominant. In general, the CALIOP retrievals do not agree well with aircraft phase observations for the three ‘widely iced’ flights when they are compared either quantitatively or qualitatively. CALIOP records ice in cloud for only one out of these three flights. Cesana *et al* (2016) evaluated three cloud phase products derived from CALIPSO against *in-situ* measurements with three flights at mid-latitudes and two flights in the Arctic. They showed the CALIPSO cloud phase agreed ~95% of the time with in situ for the homogeneous high ice clouds at mid-latitudes, but disagreed for the artic clouds for which cloud phase vertical variability is larger. One of the three products, CALIPSO-General Circulation Models-Oriented Cloud Product reported the high amount of undefined-phase clouds, which likely correspond to mixed-phase clouds. In this research, while the sample size is quite small, there is no evidence of a strong relationship between the presence of *in-situ* ice particles and the CALIOP retrieval of cloud thermodynamic phase. The comparisons also illustrate that in a multi-layer cloud scenario, low-lying clouds over the SO may readily be missed by CALIOP due to the full attenuation/extinction of the lidar signal through optically thick liquid clouds at higher altitudes (Huang 2012b; Chan and Comiso, 2011).

It is noted that CALIOP may readily see SLW at cloud tops (e.g., Rauber and Tokay, 1991), which we cannot adequately observe with the aircraft. However, there is still a substantial amount of ice and mixed-phase within these shallow SO clouds, although it may not reside at cloud top. The fact that the CALIOP algorithm only returns one phase to the entire layer detected also contributes to differences with the *in-situ* observations. Given that lidar is by nature more sensitive to liquid water drops, the retrieval is inherently biased towards liquid phase. Our finding is consistent with that from previous studies (e.g. Huang *et al.* 2015; Cesana *et al.* 2016) when comparing the CALIOP product with *in-situ* observations and/or other satellite retrievals.

3.2 MODIS

The MODIS cloud-top thermodynamic phase products (IR and CPOP) are compared with the *in-situ* observations. A further distinction between single cloud layer pixels and multiple cloud layer pixels is made as defined by the Cloud_Multi_Layer_Flag (CMLF) (Pavolonis and Heidinger, 2004; Wind *et al.*, 2010; Joiner *et al.*, 2010). CMLF = 1 indicates a single-layer cloud, CMLF >= 2 indicates multi-layer clouds, and CMLF = 0 indicates no cloud. The pixel population (PP) for single-layer clouds versus all clouds (single and multi-layer) reveals that multi-layer pixels are, on average, more common for these 11 flights, although this varies considerably from flight to flight. MODIS multilayer retrievals cannot accurately estimate the number of layers of clouds (Wind *et al.*, 2010). Wang *et al.* (2016) showed that the MODIS C6 multilayer detection is underestimated by a factor of two in comparison with the collocated CloudSat-CALIPSO cloud vertical structures in the global distribution of clouds.

For the IR single-layer observations, a rough qualitative agreement is found with the *in-situ* observations. The three ‘widely iced’ flights have the greatest percentage of ice and undetermined (UD) pixels for the single-layer IR product. When combined, ice and UD pixels account for a minimum of 56% of the PP for each of these flights. Similarly, the fourth flight with substantial icing (20140903) also has 31.5% of the PP as ice and UD. Flights with less *in-situ* mixed-phase cloud samples tend to have a smaller percentage of the PP recorded as ice and

UD for the single-layer IR product.

Next, for the CPOP single-layer retrievals, a much lower PP percentage of ice and UD pixels is observed compared to that of the IR product. For the three "widely iced" flights, the ice and UD PP percentage drops from 56.4 to 14.1%, 86.6 to 47.2% and 67.3 to 3.7%, respectively. Quantitatively, it did not agree well with *in-situ* observations although it showed a certain level of consistency, qualitatively. That is, the ice fraction of the CPOP for the 'widely iced' flights is higher than that for the other flights. Qualitatively, the single-layer CPOP is in better agreement with the CALIOP cloud phase rather than with the *in-situ* observations. This is understandable as CALIOP observations have been employed in the development of the C6 MODIS CPOP product (Platnick *et al.*, 2015; 2017).

Among the 'widely-iced' cases and the substantial icing case, all three have a relatively small optical thickness, compared to the rest of the flights. For two cases with relatively thin clouds (20131011, COT ~18 and 20140903 COT ~10 in Figure 6), CPOP underestimated the aircraft ice-phase observations, but the other case (20130815 COT ~8) agreed well with the aircraft measurements. It is noted that the CTT from the MODIS is always higher than that from the aircraft. These CPOP cloud phase results show that the algorithm still encounters a problem of determining cloud phase when there are relatively thin clouds over a warm surface.

The analysis is now expanded to consider all cloud pixels (single-layer + multi-layer). Focussing first on the IR product, the percentage of ice pixels increases for nine of the ten flights. In general, the percentage of UD pixels decreases when multi-layer pixels are included. Overall, this is still a rough agreement between the *in-situ* observations and the IR phase product. The three 'widely iced' flights have the largest percentage of ice pixels. For the CPOP product, the extension to include multi-level pixels produces an increase in the UD pixel percentage. In summary, regardless of the CMLF flag, the thermodynamic phase from the IR product is closer to the *in-situ* observations, while those from the CPOP product is closer to the CALIOP product results.

4. Cloud microphysics

The cloud microphysics properties (r_{eff} and N_d) are next considered, employing MODIS cloud property retrievals. To evaluate MODIS retrieved r_{eff} and N_d of the liquid cloud over the SO, it is worthwhile to first investigate the frequency and causes of failed MODIS cloud property retrievals (Cho *et al.*, 2015, hereafter Cho15). The cloud observations are filtered for liquid only observations for both *in-situ* measurements and MODIS retrievals using μ_{ice} and CPOP, respectively. For MODIS property retrievals, only single-layer clouds are included to minimize the retrieval bias (Wind *et al.*, 2010; Huang *et al.*, 2016; Platnick *et al.*, 2017).

The single-layer liquid cloud pixel population detected by MODIS and the retrieval failure rates on r_{eff} are detailed in Table 3 for the 2.1 and 3.7 μm channels. The 1.6 μm channel is not detailed as its pixel population is relatively small (~25% of the 3.7 μm pixel population). Also, observations from the 1.6 μm channel on Aqua MODIS are reconstructed (some of the detectors on band 6 of the Aqua MODIS were severely damaged) so it should be treated with greater caution. The MODIS 1.6 μm channel reconstruction or restoration has been made by employing a quantitative image restoration (QIR) algorithm (Gladkova *et al.*, 2012) to restore the lost data based on the information in a spatial-spectral integrated window in which the other spectral bands provide information to estimate the missing data. Beyond the photon vertical penetration depth, the pixel population (PP) for each channel may vary due to a number of channel-dependent sensitivities. The cloud inhomogeneity/sub-pixel variability (Zhang & Platnick, 2011; Zhang *et al.*, 2012), 3-D radiative effects (Davis & Marshak, 2010) and atmospheric transmittance corrections (Zhang & Platnick, 2011; Zhang *et al.*, 2012) may all affect the failure rate (Cho15).

Table 3 MODIS Effective Radius Retrieval Failure Rates

Channel	Drizzling/ Cloud regime	Cloud PP (Liq)	Failure Rate				r_{eff} retrieval success PP
			r_{eff} too large	r_{eff} too small	COT failure	Total	
2.1 μm channel	Drizzling Flights	4291	19.1%	1.1%	0.3%	20.5%	3412
	Non-drizzling Flights	7702	13.1%	1.0%	0.3%	14.4%	6592
	Total	11993	15.2%	1.1%	0.3%	16.6%	10004
	Overcast	8218				15.0%	6984
	PCL	3775				19.8%	3027
3.7 μm channel	Drizzling Flights	4033	18.2%	1.3%	0.3%	19.8%	3235
	Non-drizzling Flights	7476	9.7%	1.1%	0.3%	11.1%	6648
	Total	11509	12.7%	1.1%	0.3%	14.1%	9882
	Overcast	7855				12.6%	6862
	PCL	3654				17.3%	3021

Note: the failure rates were produced for drizzling/non-drizzling flights and cloud regime (overcast and PC) spatiotemporally aligned with *in-situ* observations from eleven flights from 2013-2015. The failure rate is broken down by causes of failure. All calculations are for single-layer liquid clouds determined by the MODIS CPOP and CMLF products.

The failure rate is defined as the ratio of the number of failed pixels retrievals to the cloud PP. Overall, the failure rate is 16.6% for the 2.1 μm channel and 14.1% for the 3.7 μm channel. The retrieval failure metric SDS details the causes of retrieval failures such as "r_{eff} is too large" or "too small" or "COT retrieval failure". The majority of the failures were attributed to r_{eff} being too large, similar to Cho15. As the existence of drizzle mode droplets could significantly affect the microphysical properties, such as retrieval differences in each MODIS channel and its pixel population, (Nakajima *et al.*, 2010a, Lebsock *et al.*, 2011; Zhang *et al.*, 2011; Zheng *et al.*, 2012; Zhang & Platnick, 2011; Painemal *et al.*, 2013), flights were identified as drizzling or non-drizzling by the drizzle fraction and rain rate (Table 1). There are 6 non-drizzling and 5 drizzling flights. Note that as only liquid clouds are examined in this analysis, drizzle does not include potential frozen or mixed-phase precipitation, which may occur in both non-drizzling and drizzling cases. For both the 2.1 and 3.7 μm channels, the failure rate for drizzling flights is substantially greater than for non-drizzling flights. As drizzle arises from large droplets, it follows directly that the increased MODIS retrieval failure is attributed to r_{eff} being too large.

It is also readily possible to identify the sensitivity of the failure rate to cloud regime by sorting the cloud pixels into 'overcast' pixels and partly cloudy (PCL) pixels. Consistent with Cho15, the failure rate is greater for PCL pixels, although to a much lesser extent than their analysis over global marine liquid clouds. Our analysis finds the failure rate for PCL pixels is ~1.3 times greater than overcast pixels (19.8/15, Table 3).

The failure rate of the MODIS algorithms has also been found to be sensitive to the cloud mask sub-pixel heterogeneity (Cho15), which is recorded in the C6 collection as CM_SPI. Formally, CM_SPI is defined as a ratio of the standard deviation to the mean of the measured reflectances of the sixteen 250 m resolution sub-pixels within the 1 km MODIS footprint (Liang *et al.* 2009). The average values of CM_SPI from band 1 (0.66 μm) and band 2 (0.87 μm) are listed in Table 1. The greatest SPIs are recorded for flights 20130707, 20130806 and 20150830, all open MCC flights. These flights also have higher retrieval failure rates than the rest of the flights (Table 1).

4.1.1 Comparison of Cloud r_{eff}

Here we compare *in-situ* and MODIS r_{eff} only for observed / retrieved liquid clouds ~~Limiting~~ (Table 1 and Figure 4). Overcast and PCL pixels are combined for the MODIS retrievals. For each flight, the MODIS mean r_{eff} is calculated for the 1.6, 2.1 and 3.7 μm channels (Figure 4). In general, the *in-situ* r_{eff} is smaller than the MODIS products, although there are some exceptions: for 10 of the 11 flights, the *in-situ* r_{eff} is smaller than the $r_{eff\_2.1}$, for 7 of the 11 flights, the *in-situ* r_{eff} is smaller than $r_{eff\_3.7}$ and $r_{eff\_1.6}$ (Figure 4 (a) and (b)). The range of the average *in-situ* r_{eff} over the 11 flights is from 7.5 to 35.4 μm, larger than the range of the MODIS r_{eff} (~12.6 to 25.8 μm). In all, except for one flight (20130723), $r_{eff\_2.1}$ is greater than $r_{eff\_1.6}$ or $r_{eff\_3.7}$. The $r_{eff\_1.6}$ and $r_{eff\_3.7}$ are broadly consistent, being within 20% of each other for 10 of the 11 flights. The mean biases of r_{eff} (MODIS - *in-situ*) are 1.1, 5.3 and 1.4 µm for 1.6, 2.1 and 3.7 µm, respectively. The square root mean biases are 8, 8.2 and 6.1 µm, respectively.

Figure 4 (c) shows $r_{eff\_2.1}$ is larger than $r_{eff\_1.6}$ or $r_{eff\_3.7}$ for most cases. It is ~~appreciated~~reported in the literature that the photon penetration depth of the different channels is sensitive to the size of the droplets, which underpins differences between $r_{eff\_2.1}$ and $r_{eff\_3.7}$ (Nakajima *et al.*, 2010a; Suzuki *et al.*, 2010; Zhang and Platnick, 2011), especially for marine liquid clouds (Nakajima *et al.*, 2010b). Specifically, $r_{eff\_3.7}$ has been reported to be smaller than either $r_{eff\_2.1}$ or $r_{eff\_1.6}$ (e.g. Liang *et al*., 2015), as the 3.7- µm channel is more sensitive to the small cloud droplets that are commonly located near cloud-top. The r_{eff} differences between the MODIS channels can be associated with the cloud's spatial inhomogeneity and clear-sky contamination (Painemal *et al.*, 2013), which will be investigated further in Section 4.1.2. MODIS r_{eff} retrieved from the overcast pixels tends to be larger than that from the PCL pixels (Figure 4(d)).

Drizzling flights are indicated with crosses in each plot. In general, non-drizzling flights are clustered in the left side of the diagram and above the one-to-one line (Figure 4(a)) for the 2.1 and 3.7 µm channels. On average, the difference between the *in-situ* r_{eff} and the MODIS r_{eff} is larger for the non-drizzling flights. The r_{eff} (CAS + CIP) for non-drizzling cases ranges from 7.5 – 16.4 µm and for drizzling cases from 17 – 35 µm, while the MODIS $r_{eff\_2.1}$ ranges from 12.9 – 25.1 µm and 20.2 – 25.8 µm for non-drizzling cases and drizzling cases, respectively. That is, the range of MODIS r_{eff} for drizzling cases is relatively smaller than that of the *in-situ* observations whereas the range of MODIS r_{eff} for non-drizzling cases is larger than that of the *in-situ* observations. Note that the MODIS look-up table limits the r_{eff} to at most 30 µm for liquid phase clouds, meaning that there is a built-in limitation to the r_{eff} that MODIS can retrieve. This upper limit also helps explain the higher retrieval failure rate for drizzling clouds (Table 3).

Figure 4

(a) Comparison of liquid cloud effective radius (r_{eff}) observed from *in-situ* measurements (CAS + CIP) and retrieved from Aqua MODIS with the combined pixels (overcast + PCL). (b) r_{eff} differences between *in-situ* observations and MODIS retrievals (overcast + PCL) from 3 channels. (c) The MODIS r_{eff} (overcast + PCL) differences between the channels. (d) The MODIS r_{eff} differences between overcast and PCL pixels for the 2.1 µm and 3.7 µm channels. Drizzling flights are indicated as crosses.

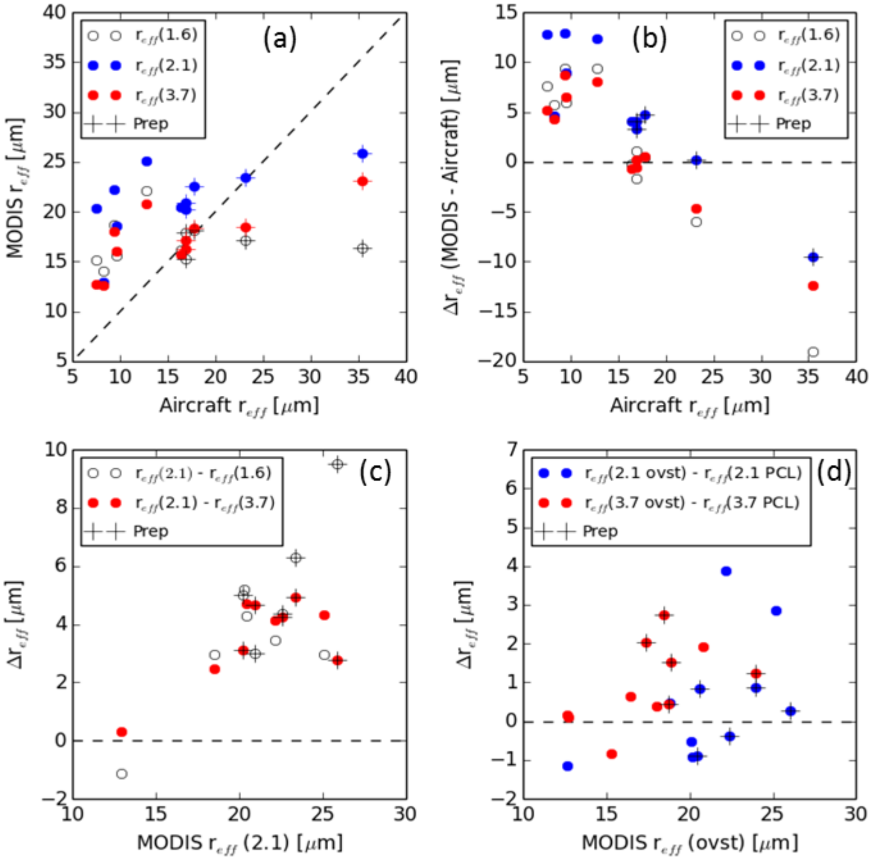
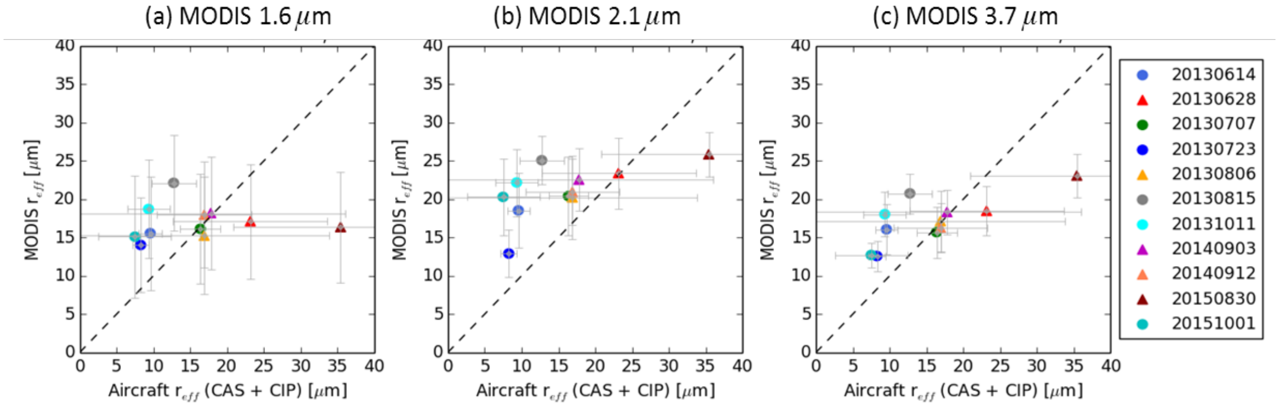


Figure 5 presents the comparison of the flight-mean liquid water r_{eff} from *in-situ* and MODIS retrievals. The largest differences between the *in-situ* and MODIS r_{eff} occur on flights 20130614, 20130815, 20131011 and 20151001. The *in-situ* r_{eff} is smaller than 13 µm for all four flights, and none of them experienced significant drizzle. Flight 20130723 also has a small *in-situ* r_{eff} and does not experience drizzle, but the *in-situ* and MODIS r_{eff} values are in relatively good agreement. Flights 20151001 and 20130723 are the two flights that encountered closed MCC.

Figure 5

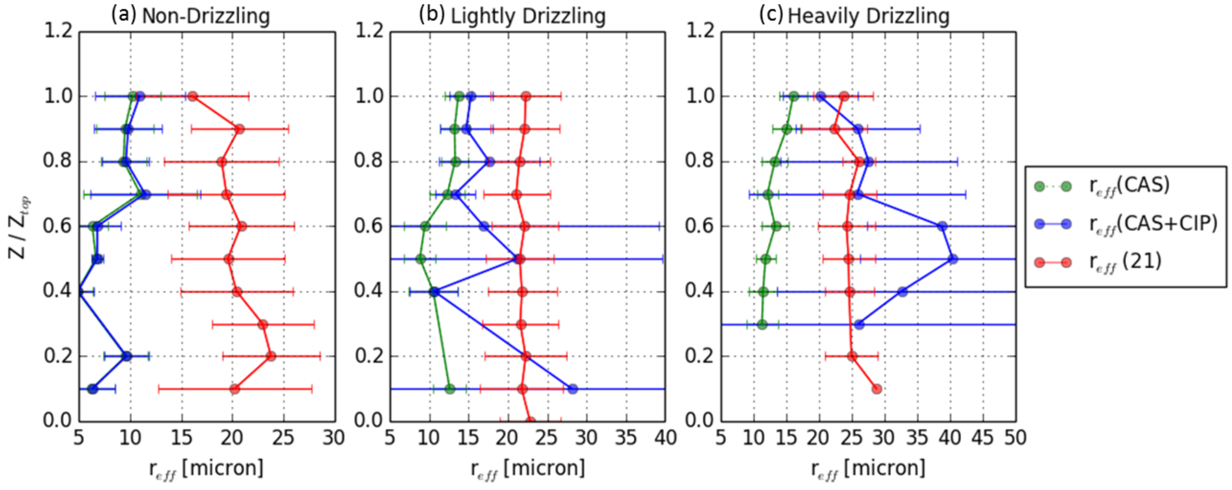
Comparison of the flight-mean liquid cloud effective radius (r_{eff}) between *in-situ* measurements from CAS + CIP and Aqua MODIS retrievals from the cloudy (overcast + PCL) pixels for (a) 1.6, (b) 2.1 and (c) 3.7 µm channels. The horizontal and vertical bars represent standard deviations of cloud r_{eff} from CAS+CIP and MODIS, respectively. The dashed lines represent 1:1 lines. Non-drizzling flights and drizzling flights are indicated as circles and triangles, respectively.



An issue that potentially impedes a direct, quantitative comparison of the *in-situ* observations and MODIS retrievals is that the observed cloud droplet size may be spatiotemporally variable and is sensitive to the aircraft position within the cloud field. While the MODIS (and CALIPSO) products tend to be more sensitive to the properties near cloud top, the aircraft was not constrained to fly only near cloud top. In Figure 6, the r_{eff} as a function of normalised cloud altitude is detailed for the 11 flights. Variations in the *in-situ* r_{eff} with altitude are identified for non-drizzling flights, which feature higher discrepancies with MODIS r_{eff}, while the magnitude of the *in-situ* r_{eff} is consistently smaller than that of MODIS $r_{eff\_2.1}$ across all the in-cloud sampling altitudes. For lightly drizzling cases, the discrepancies become smaller but the uncertainties in the *in-situ* samples increase towards the normalised cloud base, which might be due to the smaller sample sizes. The overall mean *in-situ* r_{eff} tends to be larger than that of MODIS $r_{eff\_2.1}$ for the heavily drizzling cases, while a better agreement is found towards the normalised cloud top. We acknowledge that observing potentially different clouds may explain some of the differences between the *in-situ* and MODIS r_{eff} (Platnick & Valero, 1995; Breon & Doutriaux-Boucher, 2005; Nakajima *et al.*, 2010a). We also note that the *in-situ* r_{eff} is smaller at the higher altitudes in a number of flights, possibly a consequence of entrainment across cloud-top.

Figure 6

Vertical profiles of r_{eff} with normalised cloud altitude (Z / Z_{top}) for non-drizzling (a), lightly drizzling (b) and heavily drizzling (c) flights from *in-situ* measurements and the MODIS retrievals with 2.1 channel for overcast + PCL pixels. Z indicates the cloud height and Z_{top} indicates the cloud top height.



As discussed in A17, the *in-situ* r_{eff} for the nine open/disorganised MCC flights were consistent with the wintertime observations of SOCEX I (Boers *et al.*, 1998) and the two closed MCC flights were highly consistent with the summertime observations of SOCEX II (Boers *et al.*, 1998). The same behavior is found in the MODIS r_{eff}, too, particularly for the 3.7 μm channel.

4.1.2 MODIS r_{eff} error analysis

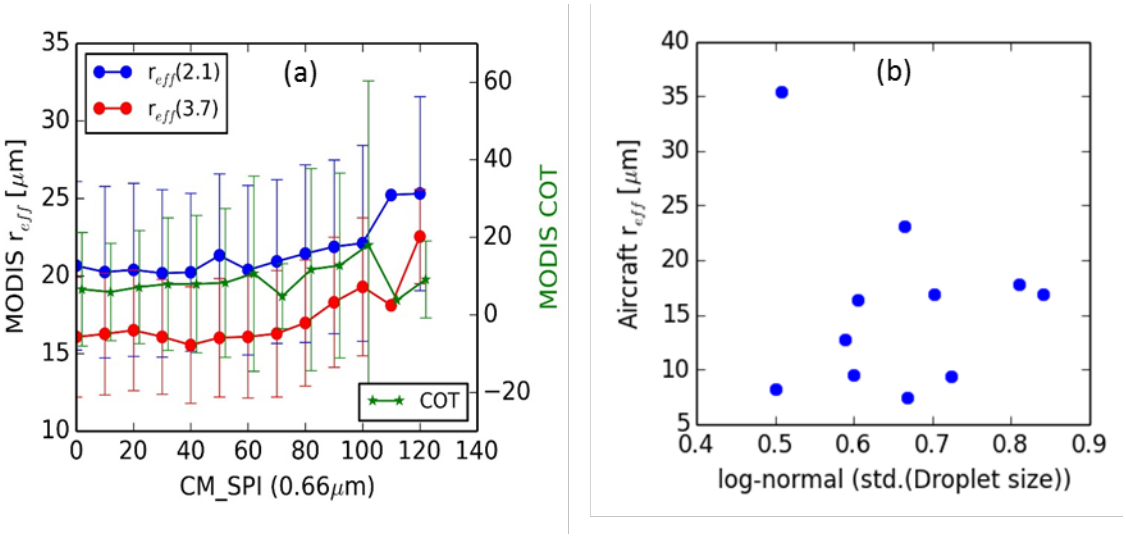
In this section, several hypothesized sources of error within the MODIS retrievals are further investigated: the sub-pixel inhomogeneity, the droplet spectra variability (cloud mode), a bimodal distribution, and the effects of SZA. Although the aim is to isolate and quantify each of these effects, in reality they are commonly intermixed making it virtually impossible to single out the effects using the limited samples from only eleven flights.

Effect of Sub-pixel Inhomogeneity

Cloud mask sub-pixel inhomogeneity (CM_SPI) can impact the MODIS r_{eff} channel differences (Zhang and Platnick, 2011; Zhang *et al.*, 2012; Painemal *et al.*, 2013). Using MODIS retrievals from the Southeast Pacific where SZA was moderately small (20º-35º) and drizzle contamination was minimal, Painemal *et al.* (2013) found a close match between $r_{eff\_2.1}$ and $r_{eff\_3.8}$ for homogeneous cases and a greater increase of $r_{eff\_2.1}$ with CM_SPI. They suggested a ~ 4 μm (maximum) channel difference could be attributed to the sub-pixel heterogeneity. Following their method, the relationships between MODIS r_{eff}, cloud optical thickness (COT) and CM_SPI are investigated in Figure 7 (a). Given that only the "collocated" MODIS pixels (11,750 pixels) are compared, any potential errors from *in-situ* measurements and sampling methods (as discussed in Section 2.1 and 2.4) are precluded. Interestingly, different from the finding in Painemal *et al.* (2013), the difference between $r_{eff\_2.1}$ and $r_{eff\_3.7}$ in our analysis shows little dependence on CM_SPI. Instead, the mean difference is relatively uniform at ~ -5 μm, showing no convergence towards homogeneous scenes (i.e. low CM_SPI). The 5 μm mean channel difference together with the lack of convergence for the most homogeneous scenes suggest that, in addition to sub-pixel heterogeneity itself, other sources of error (e.g. large SZA, vertical variability, etc.) are very likely to have a measureable impact on MODIS r_{eff} retrieval, although these error sources may not be independent (e.g. the correlation coefficients between CM_SPI and SZA are found to be 0.223 and 0.282 for band 1 and band 2, respectively). It is also interesting to note that the COT in our analysis is found to increase slightly with CM_SPI, as opposed to that in Painemal *et al.* (2013). This likely indicates the distinctiveness of the heterogeneous cloud regimes over these two geographical regions: (weakly drizzling) stratus over the Southeast Pacific and precipitating MCC over the SO.

Figure 7

(a) The MODIS $r_{eff\_2.1}$ and $r_{eff\_3.7}$ and COT binned by CM_SPI for all 1km × 1km collocated cloudy (overcast + PCL) pixels. (b) The *in-situ* r_{eff} variations against log-normal standard deviation of drop size distribution.



Effect of Droplet Spectra Variability of the Cloud Mode

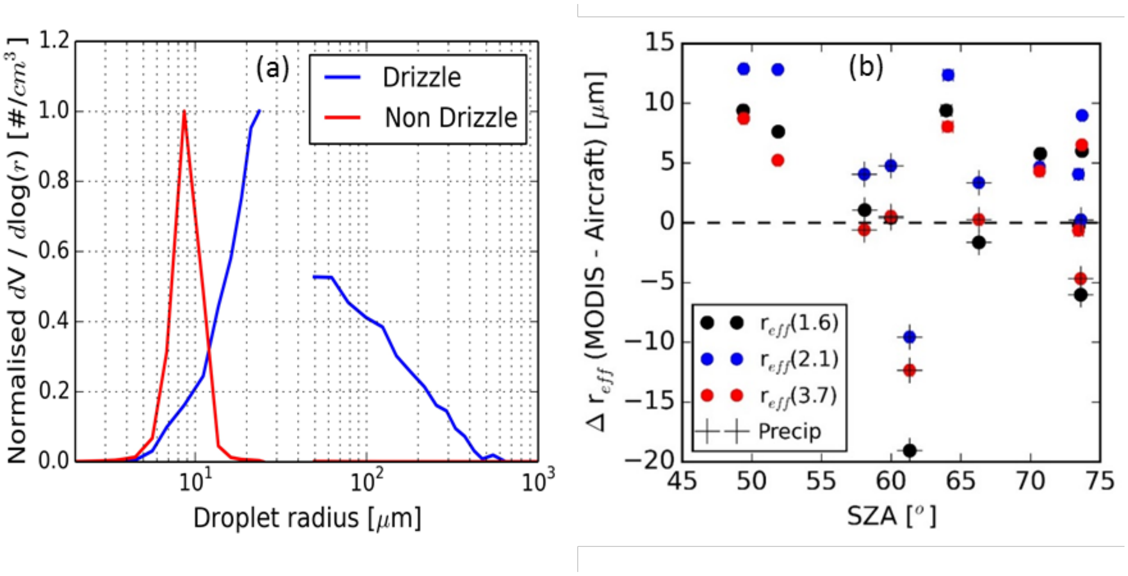
The variability of droplet spectra on the cloud mode may also impact the MODIS retrievals (Platnick and Valero, 1995). As suggested in Hansen and Travis (1974), the variability in the droplet size distribution can be assessed by the standard deviation of a lognormal droplet size distribution (*s*) (e.g. Painemal and Zuidema (2011); King *et al.*, (2013)). The MODIS algorithm uses a constant $s = 0.35$ to construct the r_{eff} look-up table. With *in-situ* data from the Southeast Pacific and radiative transfer calculations, Painemal and Zuidema (2011) showed that when the actual $s < 0.35$, the retrieved r_{eff} was larger. This bias was found to increase with r_{eff} up to 11 μm and reach a relatively constant value for $r_{eff} > 13$ μm. Using our aircraft data, however, the *in-situ* *s* is found to range from 0.5 and 0.85, with a mean value of 0.66 (Figure 7 (b)). All the drizzle flights have a relatively large *s*, while only the two non-drizzle flights have $s < 0.35$. This, once again, is vastly different from the *in-situ s* found in Painemal and Zuidema (2011), where most of their flights had an $s < 0.35$. The broader droplet size spectra (likely caused by precipitation) sampled over the SO is likely to introduce further uncertainties beyond what has been reported in the literature. An in-depth understanding of these uncertainties would require comprehensive radiative transfer simulations, which is beyond the scope of this research.

Effect of the Bimodal Distribution

Given the frequent presence of drizzle in the *in-situ* samples, an effect of a drizzle mode on the MODIS r_{eff} bias is expected. To quantify the significance of the impact, we plotted the normalized volume spectra as in Painemal and Zuidema (2011) to examine one non-drizzle flight (20130723) and two heavily drizzling flights (20150830 and 20130628). Distinct patterns are found in Figure 8 (a) where the non-drizzle flight is clearly dominated by the cloud mode whereas the drizzle flights are featured by a ratio of 1:0.5 between the largest cloud and drizzle mode. According to the numerical analysis in Nakajima *et al*. (2010b), the r_{eff} bias associated with the bimodal distribution for the selected drizzle flights is ~ 2.5μm.

Figure 8

(a) Normalized volume spectra for one non-drizzle flight (20130723) and two heavily drizzling flights (20150830 and 20130628). (b) The mean r_{eff} bias as a function of solar zenith angle (SZA).



Effect of Solar Zenith Angle

SZA can become relatively large during the Austral winter. Five of the eleven flights have the SZA in excess of 65°, a threshold defined in Grosvenor and Wood (2014). The estimated mean r_{eff} bias associated with large SZA in Grosvenor and Wood (2014) is ~1 μm, which is much smaller than the mean biases found in our study. Overall, no systematic biases on r_{eff} from the SZA are evident from these eleven flights, although the sample size is small with many confounding processes. The correlation coefficients between the mean SZA and the mean r_{eff} biases (Figure 8 (b)) are found to be -0.214, -0.032 and -0.209 for 1.6, 2.1 and 3.7 μm channels, respectively. The six flights having SZA < 65° still have high biases in r_{eff} (mean of ~ 9.5 μm), most likely due to other error sources as discussed previously. To gain a more quantitative understanding of the SZA impact on the biases, flight 20130723 is examined again. This is the only flight among the eleven where single-layer, homogenous, non-drizzle, warm boundary layer stratocumulus clouds were observed, except that the SZA (71°) is considered large. The mean r_{eff}

biases for the 1.6, 2.1 and 3.7 μm channels are 5.7, 4.6, and 4.3 μm, respectively (Table 1). Taking a 20% CAS sizing error into account (Ahn *et al*. 2017), the minimum mean r_{eff} bias should be ~ 2.7 μm (at 3.7 μm channel). Because the mean CM_SPI of this flight is small (12 and 15 for band 1 and band 2), the expected CM_SPI-induced bias should be negligible (according to Painemal *et al.* (2013)). Also, the mean *in-situ* *s* for the flight is 0.5, leading to an estimated *s*–induced error of less than 1 μm (based on the results of Painemal and Zuidema (2011)). As a result, the residual of the minimum bias (2.7 μm -2.7 μm = 1.7 μm) is most likely caused by the large SZA. This estimate is comparable to the finding in Grosvenor and Wood (2014).

4.2 Comparison of Cloud N_d

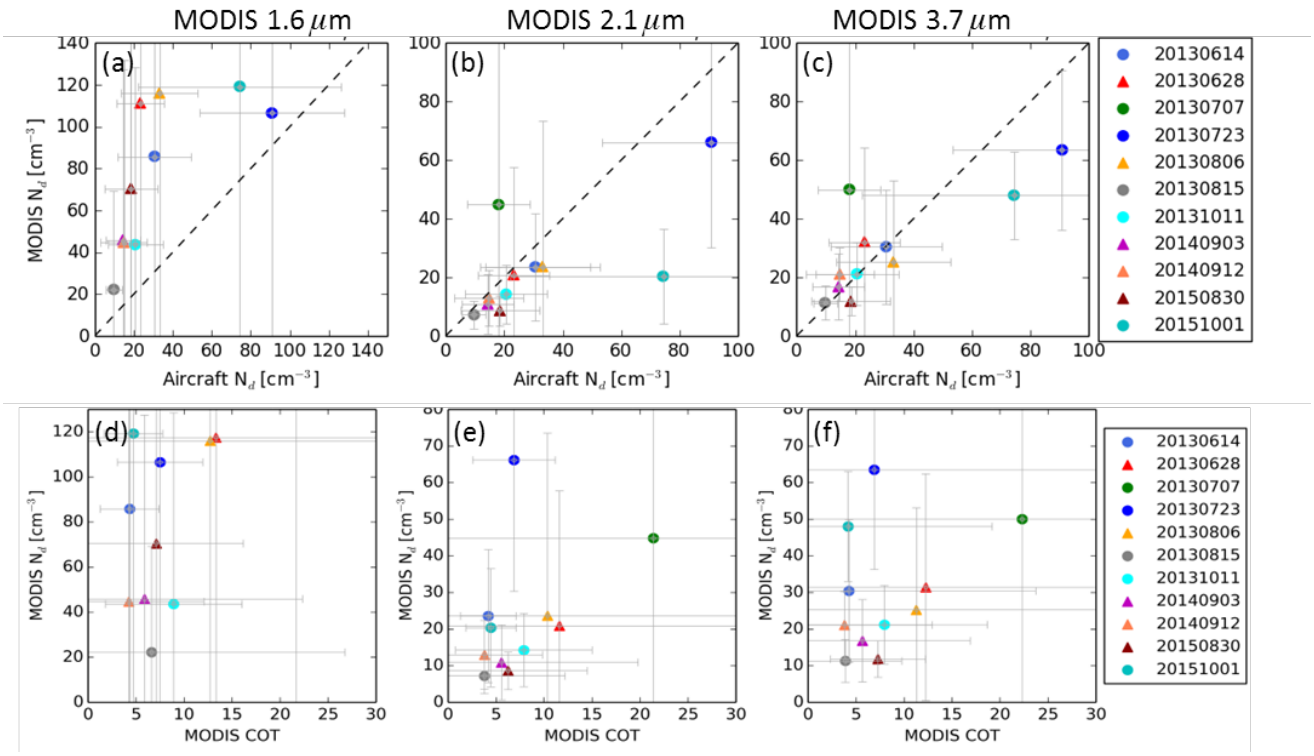
As with the r_{eff}, the MODIS-derived N_d can be compared against the *in-situ* observations with all cloudy (overcast and PCL) pixels (Figure 9). As discussed in A17, the majority of the clouds sampled in this region have low N_d with the large r_{eff}. For nine of the 11 flights in this study, *in-situ* calculations of N_d were less than 35 cm^{-3}. Five of the 11 flights had N_d less than 20 cm^{-3}. The two high-concentration outliers are the two closed MCC cases (flights 20130723 and 20151011) with concentrations of 89 and 80 cm^{-3}, respectively. As discussed in A17, these *in-situ* observations of low N_d are consistent with those of SOCEX experiments from 20 years earlier (Boers *et al.*, 1996; 1998).

The MODIS N_d calculated with the r_{eff} from the 2.1 and 3.7 μm channels are largely consistent with the *in-situ* N_d. Larger differences are found for the 1.6 μm channel, which can be associated with the faulty operations on this channel. As Painemal & Zuidema (2011) reported, a good agreement between the MODIS (2.1 and 3.7 μm channel) and *in-situ* N_d is remarkable given the systematic bias found in MODIS r_{eff}.

For eight of the 11 flights, the MODIS calculations more closely follow a one-to-one relationship with the *in-situ* observations. However, there are three outliers in the MODIS 2.1 and 3.7 μm channel plots: 20130707, 20130723, and 10251001 cases, all non-drizzling. Overall, the 3.7 μm N_d has a closer agreement with the *in-situ* observations, due to a better agreement for the three outlier flights. The mean biases are -53, 9, and 1 cm^{-3} for 1.6, 2.1 and 3.7 μm channels, respectively. The correlation coefficients for them are 0.44, 0.68 and 0.80, respectively. The t-test results show the correlation coefficient with the 3.7 μm channel is statistically significant: the p-values for the three channels are 0.002, 0.379 and 0.876, respectively.

Figure 9

(a), (b) and (c): same as Figure 5 but for N_d. N_d is derived by the MODIS retrievals of r_{eff} and cloud optical thickness (COT) from equation (1). (d), (e) and (f): N_d as a function of COT for the combined (overcast + PCL) pixels retrieved from Aqua MODIS. The horizontal bars and vertical bars represent standard deviations of N_d and COT, respectively. Non-drizzling flights and drizzling flights are indicated as circles and triangles, respectively.



The overall good agreement between the *in-situ* and MODIS derived N_d is curious, given the large bias in r_{eff} as discussed before. This result is consistent with the findings in Painemal & Zuidema (2011), where the good agreement in N_d is attributed to an error compensation because the "c" and "k" parameters in Equation (1) differ from those in real clouds. An attempt has also been made here to understand the error budget using the *in-situ* data from our flights, following the Painemal & Zuidema (2011) method. In equation (1), when the constant values are excluded, the potential error in the MODIS retrievals can be explained by

$$\alpha = \frac{c^{1/2} \cdot COT^{1/2}}{k \cdot {r_{eff}}^{5/2}} \qquad (2).$$

Painemal & Zuidema (2011) assumed the COT of *in-situ* observations and MODIS retrievals matched very well so they did not include the COT term as a potential error source. We keep that term in our analysis because the potential bias in COT in the SO region cannot be ruled out using our limited *in-situ* data. Due to the highly inhomogeneous nature of the clouds, even calculating c and k is very challenging. We will use several cases (20130614, 20132723, 20140912 and 20151001) where clouds developed close to adiabatically to help provide further insights for the error investigation. For the case of 20130614 where only single layer clouds were sampled, MODIS N_d is 0.77 times of *in-situ* N_d. Based on the estimated cloud base and top heights, c, the condensation rate [g/m^3/km] is calculated using the *in-situ* LWC measurements during the sampling periods. The *in-situ* c is found to be 0.664 g/m^3/km, whereas the c value used for the MODIS N_d calculation overestimates the *in-situ* c by 2.7 times. k at the cloud top for *in-situ* observations is 0.95, 0.842 times larger than the constant value 0.8 used in Equation (1). In this case the MODIS r_{eff} overestimates the *in-situ* r_{eff} by 1.9 times. When these numbers were put together in equation (2), α becomes

$$\alpha = \frac{(2.7c_{in-situ})^{\frac{1}{2}} \cdot COT^{\frac{1}{2}}}{0.842k_{in-situ} \cdot (1.9r_{eff_{in}-situ})^{\frac{5}{2}}} = 0.38 \cdot COT^{\frac{1}{2}} \qquad (3)$$

This result suggests that the overestimation of r_{eff} is compensated by the overestimated c and underestimated k (Painemal & Zuidema, 2011) to a certain extent, but this compensation alone is still not sufficient to explain the 0.77 N_d underestimation. The uncertainty of COT, presumably overestimated may also have contributed to the compensation. The other three cases had similar results, although the overestimated r_{eff} of MODIS played a major role in the underestimation of N_d for the 20130723 and 20151001 cases.

5. Discussions and Summary

The aim of this research has been to employ the *in-situ* observations to better appreciate the skill and limitations of satellite derived products for thermodynamic phase (MODIS and CALIPSO), r_{eff} (MODIS) and N_d (MODIS) over the Southern Ocean. We consider eleven research flights taken over three cold seasons (May - Oct, 2013-2015) that were timed/designed to pass under A-train observations.

While many of the observations made during VOCALS-REx were of warm, boundary layer clouds defined by closed MCC, the environment over the Southern Ocean is far different. The clouds observed in these flights over winter time have more open or unorganised MCC (Muhlbauer *et al*. 2014) and have a higher (~32%) portion of PCL pixels than global marine clouds (Cho15), meaning that the assumption of horizontal homogeneity commonly employed in satellite-based algorithms is often invalid. Further, these clouds were frequently observed to contain ice, complicating the underlying assumptions. This was evident in the relatively higher (42% with 3.7 μm channel) failure rate for the MODIS cloud optical property retrievals, compared to the global marine clouds (Cho15), although many of these failures were actually attributed to the droplet radius being too large.

Firstly, the thermodynamic phase of the MODIS CPOP and IR products was examined, as well as CALIOP. Due to the complexity of the sampled cloud fields and operational constraints, our collocations and comparisons are only performed in a statistical sense. The CALIOP product primarily returns liquid phase layers, even at altitudes of 4 km and greater, and at relatively cold cloud-top temperatures (-15 ºC). Despite the limited *in-situ* samples that impede a more quantitative evaluation, our finding is consistent with that in Cesana *et al*. (2016), where three different CALIOP products are found to overestimate the liquid phase fraction in mixed-phase clouds identified by the airborne *in-situ* data over the Arctic.

The MODIS CPOP phase product, which has been influenced by CALIOP phase discrimination in the C6 data set, behaves comparably; liquid phase is predominantly recorded at cloud-top. The CPOP product phase did not agree well with *in-situ* observations quantitatively. The MODIS IR phase product, however, qualitatively agrees more closely with the *in-situ* observations, although it overestimated the *in-situ* ice fractions by a factor of 2 – 8. However, the C6 MODIS IR and CPOP products showed an ‘internal’ qualitative consistency within the IR and CPOP data. That is, the ice fraction of the IR or CPOP for the iced flights is higher than that from the other flights.

Next, the *in-situ* observations were employed to evaluate the r_{eff} and N_d retrieved from the 1.6, 2.1 and 3.7 μm channels on MODIS. The *in-situ* observations revealed that five of the eleven flights were drizzling, two heavily, and the remaining six flights were not observed to be drizzling significantly. As would be expected, r_{eff} is greater and N_d is smaller for drizzling clouds, overall.

Over the eleven flights, the MODIS r_{eff} has a much smaller range than the *in-situ* observations. It underestimated r_{eff} for heavily drizzling clouds and highly overestimated r_{eff} for non-drizzling clouds. The MODIS r_{eff} bias was not observed to be particularly sensitive to the solar zenith angle for these flights, i.e. the natural variability of r_{eff} of the observed clouds was greater than the potential bias identified in Grosvenor and Wood (2014.)

For the six non-drizzling flights, each of the MODIS r_{eff} products overestimates the *in-situ* r_{eff} with the r_{eff} of the 3.7 and 1.6 μm channels performing marginally better than that of the 2.1 μm channel. This is largely consistent with the observations made over the Southeast Pacific during VOCALS-REx (e.g. Zheng *et al.*, 2011; Painemal & Zuidema, 2011; Min *et al.*, 2012; King *et al.*, 2013), although the magnitude of the overestimate (4 – 13 μm) is greater (67%) for the observations over the Southern Ocean. In addition to this systematic overestimation, part of the issue is assumed to be due to the ice contamination (e.g. Zheng *et al.*, 2011).

For the three lightly drizzling flights, the MODIS r_{eff} roughly agrees with the *in-situ* observations with high correlation, while for the two heavily drizzling flights, the MODIS r_{eff} underestimates the *in-situ* observations. For these two heavily drizzling flights, droplets with an r_{eff} larger than 30 μm are present, which is an upper limit for the MODIS r_{eff} algorithm. Accordingly, large MODIS failure rates were noted for these two flights. These heavily drizzling events are likely to be problematic for the MODIS algorithms for a number of reasons. Moreover, the in-situ observations have included samples relatively far below cloud-top, where precipitation is likely to have more fully developed. MODIS retrievals may not have penetrated to such depths.

Lastly, the MODIS derived N_d, is largely consistent with the *in-situ* observations in spite of the limitation in r_{eff}. Remarkably, skill was still evident for heavily drizzling, horizontally inhomogeneous, mixed-phase clouds. The two closed MCC flights had the largest values of N_d, but still underestimated the *in-situ* observations. Overall the N_d of the 3.7 µm channel was in better agreement with the observations than that of the 2.1 µm channel. An error analysis using *in-situ* data suggests that the better agreement in N_d is likely a result of compensating effects, although the impact of each error source may vary from flight to flight.

Given the unique aspects of these open MCC clouds over the SO, it is not surprising that products of microphysical properties derived from satellite observations often do not compare well with *in-situ* observations. Ultimately, though, this analysis suffers from very limited *in-situ* observations. Eleven flights are simply too few to establish differences with a high degree of confidence, l*et al*one isolate and quantify the factors governing these differences.

Notation

AC GEO	Advisory Committee for Geosciences
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization
CALIPSO	Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations
CAS	Cloud Aerosol Spectrometer
CIP	Cloud Imaging Probe
CMLF	Cloud multi-layer flag
CM_SPI	Cloud mask sub-pixel heterogeneity index
COT	Cloud optical thickness
CPOP	Cloud phase optical property
CSR	Clear sky restoral
CTH	Cloud top height

CTP	Cloud-top pressure
CTT	Cloud top temperature
HOI	Horizontally-oriented ice
IR	Infrared
MCC	Mesoscale cellular convection
MODIS	MODerate-resolution Imaging Spectroradiometer
N_d	Cloud droplet number concentration
NSF	National Science Foundation
PCL	Partly cloudy
PP	Pixel population
r_{eff}	Cloud effective radius
ROI	Randomly-oriented ice
SDS	Science data set
SOCEX	Southern Ocean Cloud Experiment
SZA	Solar zenith angle
SO	Southern Ocean
UD	Undetermined
UN/ND	Unknown/not determined
VOCALS-REx	Variability of the American Monsoon Systems (VAMOS) Ocean-Cloud-Atmosphere-Land Study Regional Experiment

Acknowledgements

This work is supported by Australian Research Council Linkage Project DP150102894. The research flights were conducted in cooperation with Hydro Tasmania. We deeply thank to the cloud-seeding team from Hydro Tasmania for their kind and consistent assistance with those research flights. The aircraft data used is listed in the references. We thank the three anonymous reviewers for their insightful advice to improve this manuscript. MODIS data were obtained from the NASA Goddard Distributed Active Archive Center (DAAC) data archive. CALIOP

data were obtained from the NASA the Atmospheric Science Data Center (ASDC). All the data used are listed in the references or archived in https://ladsweb.modaps.eosdis.nasa.gov/search/ repository.

References

Ahn E., Huang Y., Chubb T. H., Baumgardner D., Isaac P., Hoog M. d., Siems S.T., & Manton M. J. (2017), In situ observations of wintertime low-altitude clouds over the Southern Ocean, *Quart. J. Roy. Meteor. Soc.*, v**143**, i704, 1381–1394, doi: 10.1002/qj.3011.

Baum B. A., Soulen P. F., Strabala K. I., King M. D., Ackerman S. A., Menzel W. P., & Yang P. (2000), Remote sensing of cloud properties using MODIS Airborne Simulator imagery during SUCCESS: 2. Cloud thermodynamic phase, *J. Geophys. Res.*, **105**, 11,781–11,792.

Baum B. A., Menzel W. P., Frey R. A., Tobin D. C., Holz R. E., Ackerman S. A., & coauthors. (2012), MODIS cloud-top property refinements for Collection 6. *J. Appl. Meteor. Climatol.*, **51**, 1145–1163, doi:10.1175/JAMC-D-11-0203.1.

Baumgardner D, Jonsson H, & Dawson W. (2001), The cloud aerosol and precipitation spectrometer: A new instrument for cloud investigators. *Atmos. Res.* **59–60**: 251–264, doi: 10.1016/S0169-8095(01)00119-3.

Bennartz, R. (2007), Global assessment of marine boundary layer cloud droplet number concentration from satellite. *J. Geophys. Res.*, **112**, D02201, doi:10.1029/2006JD007547.

Bodas-Salcedo A., Hill P. G., Furtado K., Williams K. D., Field P. R., Manners J. C., & coauthors. (2016), Large contribution of supercooled liquid clouds to the solar radiation budget of the Southern Ocean. *J. Climate*, **29**, 4213–4228, doi:10.1175/JCLI-D-15-0564.1.

Boers R., Jensen J. B., & Krummel P. B. (1998), Microphysical and short-wave radiative structure of marine stratocumulus clouds over the Southern Ocean: Summer results and seasonal differences. *Quart. J. Roy. Meteor. Soc.*, **124**, 151–168, doi:10.1002/qj.49712454507.

Boers R., Acarreta J. R., & Gras J. L. (2006), Satellite monitoring of the first indirect aerosol effect: retrieval of the droplet concentration of water clouds, *J. Geophys. Res.-Atmos.*, **111**, D22208 doi:10.1029/2005jd006838.

Boutle I. A., Abel S.J., Hill P. G., & Morcrette C. J. (2014), Spatial variability of liquid cloud and rain: Observations and microphysical effects. *Q. J. R. Meteorol Soc.* **140**: 583–594, doi: 10.1002/qj.2140.

Breon F. M. & Doutriaux-Boucher M. (2005), "A comparison of cloud droplet radii measured from space," *in IEEE Transactions on Geoscience and Remote Sensing*, vol. **43**, no. 8, pp. 1796-1805, Aug. doi: 10.1109/TGRS.2005.852838

Chan M. A. & Comiso J. C. (2011), Cloud features detected by MODIS but not by CloudSat and CALIOP, *Geophys. Res. Lett.*, **38**, L24813, doi:10.1029/2011GL050063.

Cho H. M. & coauthors. (2015), Frequency and causes of failed MODIS cloud property retrievals for liquid phase clouds over global oceans. *J. Geophys. Res. Atmos.*, **120**, 4132–4154, doi:10.1002/2015JD023161.

Coakley J. A., Friedman M. A., & Tahnk W. R. (2005), Retrieval of cloud properties for partly cloudy imager pixels, *J. Atmos. Oceanic Technol.*, **22**, 3–17. https://doi.org/10.1175/JTECH-1681.1

Davis A. & Marshak A. (2010), 3D transport of solar radiation in clouds. *Rep. Prog. Phys.*, **73**, 026801, doi:10.1088/0034-4885/73/2/026801.

Delanoë, J. & R. J. Hogan (2010), Combined CloudSat-CALIPSOMODIS retrievals of the properties of ice clouds, *J. Geophys. Res.*, **115**, D00H29, doi:10.1029/2009JD012346.

Dong, X., Minnis, P., Mace, G. G., Smith Jr., W. L., Poellot, M., Marchand, R. T., & Rapp, A. D. (2002. Comparison of stratus cloud properties deduced from surface, GOES, and aircraft data during the March 2000 ARM Cloud IOP, *J. Atmos. Sci.*, **59**, 3256–3284.

Gladkova I., M. D. Grossberg, F. Shahriar, G. Bonev and P. Romanov (2012), "Quantitative

Restoration for MODIS Band 6 on Aqua," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. **50**, no. 6, pp. 2409-2416, doi: 10.1109/TGRS.2011.2173499

Grosvenor D. P. & Wood R. (2014), On the effect of solar zenith angle on MODIS cloud optical and microphysical retrievals. *Atmos. Chem. Phys.*, **14**, 7291–7321, doi:10.5194/acp-14-7291-2014.

Hande, L. B., S. T. Siems, M. J. Manton *et al*. (2012), Observations of wind shear over the Southern Ocean, *J. Geophys. Res.*, **117**, D12206.

Hansen, J., and L. Travis (1974), Light scattering in planetary atmospheres, *Space Sci. Rev.*, **16**, 527–610, doi:10.1007/BF00168069.

Haynes J. M., C. Jakob W. B. Rossow, G. Tselioudis, & Brown J. (2011), Major characteristics of Southern Ocean cloud regimes and their effects on the energy budget. *J. Climate*, **24**, 5061–5080, doi:10.1175/2011JCLI4052.1.

Holtz R. E., Ackerman S. A., & Nagle F. W. (2008), Global Moderate Resolution Imaging Spectroradiometer (MODIS) cloud detection and height evaluation using CALIOP, *J. Geophys. Res.*, **113**, D00A19, doi:10.1029/2008JD009837.

Hu Y., & Coauthors. (2009), CALIPSO/CALIOP cloud phase discrimination algorithm. *J. Atmos. Oceanic Technol.*, **26**, 2293–2309, doi:10.1175/2009JTECHA1280.1.

Hu Y., S. Rodier, K. Xu, W. Sun, J. Huang, B. Lin, P. Zhai, & Josset D. (2010), Occurrence, liquid water content, and fraction of supercooled water clouds from combined CALIOP/IIR/MODIS measurements. *J. Geophys. Res.*, **115**, D00H34, doi:10.1029/2009JD012384.

Huang Y., Siems S. T., Manton M. J., Hande L. B., & Haynes J. M. (2012a), The structure of low-altitude clouds over the Southern Ocean as seen by CloudSat. *J. Climate*, **25**, 2535–2546, doi:10.1175/JCLI-D-11-00131.1.

Huang Y., Siems S. T., Manton M. J., A. Protat, & Delanoë J. (2012b), A study on the low-altitude clouds over the Southern Ocean using the DARDAR-MASK. *J. Geophys. Res.*, **117**, D18204, doi:10.1029/2012JD017800.

Huang Y, Siems ST, Manton MJ, & Thompson G. (2014), An evaluation of WRF simulations of clouds over the Southern Ocean with A-Train observations. *Mon. Weather Rev*. **142**: 647–667, doi: 10.1175/MWR-D-13-00128.1.

Huang Y, Protat A, Siems S. T., & Manton M. J. (2015a), A-Train observations of maritime mid latitude storm-track cloud systems: Comparing the Southern Ocean against the North Atlantic. *J. Clim*. **28**: 1920–1939, doi: 10.1175/JCLID- 14-00169.1.

Huang Y, Franklin C. N., Siems S. T., Manton M. J., Chubb T. H., Lock A., Alexander S., & Klekociuk A. (2015b), Evaluation of boundary-layer cloud forecasts over the Southern Ocean in a limited-area numerical weather prediction system using in situ, space-borne and ground-based observations. *Q. J. R. Meteorol. Soc*. **141**: 2259–2276, doi: 10.1002/qj.2519.

Huang Y, Siems S. T., Manton M. J., Rosenfeld D., Marchand R., McFarquhar G. M., & Protat A. (2016), What is the role of sea surface temperature in modulating cloud and precipitation properties over the Southern Ocean? *J. Clim*. **29**:7453–7476. https://doi.org/10.1175/JCLI-D-15-0768.1.

Huang Y, Chubb T. H., Baumgardner D., Hoog M. d, Siems S. T., & Manton M. J. (2017), Evidence for secondary ice production in Southern Ocean open cellular convection, *Q. J. R. Meteorol. Soc.* **143**: 1685–1703, DOI: 10.1002/qj.3041.

Im E., C. Wu, & S. L. Durden. (2005), Cloud profiling radar for the CloudSat mission, IEEE Trans., *Aerosp. Electron. Syst.*, **20**, 15-18.

Joiner, J. A. P. Vasilkov, P. K, Bhartia, G. Wind, S. Platnick, & W. P. Menzel, (2010), Detection of multi-layer and vertically-extended clouds using A-Train sensors. *Atmos. Meas. Tech.*, **3**, 233-

247.

Kay, J. E., B. Medeiros, Y.-T. Hwang, A. Gettelman, J. Perket, & M. G. Flanner, (2014), Processes controlling Southern Ocean shortwave climate feedbacks in CESM, *Geophys. Res. Lett*., **41**, doi:10.1002/2013GL058315. doi:10.1002/2013GL058315.

King N. J., Bower, K. N., Crosier, J., & Crawford, I. (2013), Evaluating MODIS cloud retrievals with in situ observations from VOCALS-REx, *Atmos. Chem. Phys.*, **13**, 191-209, doi:10.5194/acp-13-191-2013.

Korolev A. V., Isaac G. A., Cober S. G., Strapp J. W., & Hallett J. (2003), Microphysical characterization of mixed-phase clouds. *Q. J. R. Meteorol. Soc.* **129**: 39–65.

Lebsock M. D., T. S. L'Ecuyer & Stephens G. L. (2011), Detecting the ratio of rain and cloud water in low-latitude shallow marine clouds. *J. Appl. Meteorol. Clim.*, **50**, 419-432.

Liang, L., L. Di Girolamo, & S. Platnick. (2009), View-angle consistency in reflectance, optical depth, and spherical albedo of marine water clouds off the coast of California through MISR-MODIS fusion. *Geophys. Res. Lett.*, **36**, L09811, doi: 10.1029/2008GL037124.

Lilie L, Emery E, Strapp J. W., & Emery J. (2005), 'A multiwire hot-wire device for measurement of icing severity, total water content, liquid water content, and droplet diameter'. 43rd *AIAA Aerospace Sciences Meeting and Exhibit*, 10–13 January 2005. Aerospace Sciences Meetings. Reno, NV.

Marchant B., S. Platnick, K. Meyer, G. T. Arnold, & Riedi J. (2016), MODIS Collection 6 shortwave-derived cloud phase classification algorithm and comparisons with CALIOP. *Atmos. Meas. Tech.*, **9**, 1587–1599, doi:10.5194/amt-9-1587-2016.

Menzel, W. P., R. A. Frey, H. Zhang, D. P. Wylie, C. C. Moeller, R. E. Holz, B. Maddux, B. A. Baum, K. I. Strabala, & L. E. Gumley. (2008), MODIS global cloud-top pressure and amount estimation: Algorithm description and results. *J. Appl. Meteor. Climatol.*, **47**, 1175–1198, doi:10.1175/2007JAMC1705.1.

Min Q., Joseph, E., Lin, Y., Min, L., Yin, B., Daum, P. H., Kleinman, L. I., Wang, J., & Lee, Y.-N. (2012), Comparison of MODIS cloud microphysical properties with *in-situ* measurements over the Southeast Pacific, *Atmos. Chem. Phys.*, **12**, 11261-11273, doi:10.5194/acp-12-11261-2012.

Muhlbauer A, McCoy I. L., & Wood R. (2014), Climatology of stratocumulus cloud morphologies: Microphysical properties and radiative effects. *Atmos. Chem. Phys*. **14**: 6695–6716, doi: 10.5194/acp-14-6695-2014.

Nakajima T. Y., K. Suzuki, & Stephens G. L. (2010a), Droplet growth in warm water clouds observed by the A-train. Part I: Sensitivity analysis of the MODIS-derived cloud droplet sizes, *J. Atmos. Sci.*, **67**(6), 1884–1896, doi:10.1175/2009JAS3280.1.

Nakajima, T. Y., Suzuki K., & Stephens G. L. (2010b), Droplet Growth in Warm Water Clouds Observed by the A-Train. Part II: A Multisensor View, *Journal of the Atmospheric Sciences*, vol. **67**, issue 6, pp. 1897-1907.

National Science Foundation (NSF), (2014), Investing in Science, Engineering, and Education for the Nation’s Future, Strategic Plan for 2014 – 2018.

Painemal, D. & Zuidema, P. (2011), Assessment of MODIS cloud effective radius and optical thickness retrievals over the Southeast Pacific with VOCALS-REx in situ measurements, *J. Geophys. Res.*, **116**, D24206, doi:10.1029/2011JD016155.

Painemal, D., P. Minnis, and S. Sun-Mack. (2013), The impact of horizontal heterogeneities, cloud fraction, and liquid water path on warm cloud effective radii from CERES-like Aqua MODIS retrievals. *Atmos. Chem. Phys*., **13**, 9997–10 003, doi:https://doi.org/10.5194/acp-13-9997-2013.

Pavolonis, M. J., & A. K. Heidinger. (2004), Daytime cloud overlap detection from AVHRR and VIIRS. *J. Appl. Meteor.*, **43**, 762-778.

Platnick S., & Valero F. P. J. (1995), A validation study of a satellite cloud retrieval during ASTEX. *J. Atmo. Sci.*, **52**, 2985-3001.

Platnick S. (2000), Vertical photon transport in cloud remote sensing problems. *J. Geophys. Res.*, **105**, 22919-22935.

Platnick, S., & coauthors (2015), MODIS Cloud Optical Properties: User Guide for the Collection 6 Level-2 MOD06/MYD06 product and associated level-3 datasets. NASA Goddard Space Flight Center, 34 pp.

Platnick. S., K. G. Meyer, M. D. King, G. Wind, N. Amarasinghe, B. Marchant, G., & coauthors. (2017), "The MODIS Cloud Optical and Microphysical Products: Collection 6 Updates and Examples From Terra and Aqua" in *IEEE Transactions on Geoscience and Remote Sensing*, vol. **55**, no. 1, pp. 502-525, doi: 10.1109/TGRS.2016.2610522.

Powell K., & Coauthors (2010), The CALIOP 532-nm channel daytime calibration: Version 3 algorithm. Proc. 25th *Int. Laser Radar Conf.*, St. Petersburg, Russia, Intl. Radiation Commission, 1367–1370.

Pruppacher H. R., & Klett J. D. (2010), Microphysics of Clouds and Precipitation, Chapter 10. Springer, doi: 10.1007/978-0-306-48100-0: Houten, Netherlands.

Seethala C., & Horváth Á. (2010), Global assessment of AMSR-E and MODIS cloud liquid water path retrievals in warm oceanic clouds, *J. Geophys. Res.*, **115**, D13202, doi:10.1029/2009JD012662.

Stephens, G. L. & coauthors. (2002), The CloudSat mission and the A-Train: Anew dimension of space-based observations of clouds and precipitation. Bull. *Amer. Meteor. Soc.*, **83**, 1771–1790, doi:10.1175/BAMS-83-12-1771.

Suzuki K., T. Y. Nakajima, & Stephens G. L. (2010), Particle growth and drop collection efficiency of warm clouds as inferred from joint CloudSat and MODIS observations, *J. Atmos. Sci.*, **67**(9), 3019–3032, doi:10.1175/2010JAS3463.1.

Rauber, R.M. and A. Tokay, 1991: An Explanation for the Existence of Supercooled Water at the Top of Cold Clouds. *J. Atmos. Sci.*, **48**, 1005–1023, https://doi.org/10.1175/1520-0469(1991)048<1005:AEFTEO>2.0.CO;2.

The aircraft dataset used in this analysis was produced and maintained by Mark de Hoog and Christina Nebel in Hydro Tasmania Ltd.

Trenberth KE, & Fasullo JT. (2010), Simulation of present-day and twenty-firstcentury energy budgets of the Southern Oceans. *J. Clim.* **23**: 440–454, doi: 10.1175/2009JCLI3152.1.

Vaughan M, Pitts M, Trepte C, Winker D. & Coauthors. (2017), Cloud – Aerosol LIDAR Infrared Pathfinder Satellite Observations (CALIPSO) Data Management System Data Products Catalog Release 4.10 Document No: PC-SCI-503

Wang Z., S. T. Siems, D. Belusic´, M. J. Manton, & Huang Y. (2015), A climatology of the precipitation over the Southern Ocean as observed at Macquarie Island. *J. Appl. Meteor. Climatol.*, **54**, 2321–2337, doi:10.1175/JAMC-D-14-0211.1.

Wentz F. J. and T. Meissner (2000), AMSR algorithm theoretical basis document, version **2**, *Remote Sens. Syst.*, Santa Rosa, Calif.

Wind G. S., Platnick M. D., King P. A., Hubanks M. J., Pavolonis A. K., Heidinger P. Yang & B. A. Baum. (2010), Multilayer cloud detection with the MODIS near-infrared water vapor absorption band. *J. Appl. Meteor. Climatol.*, **49**, 2315-2333.

Wolters E. L. A., H. M. Deneke, B. J. J. M. van den Hurk, J. F. Meirink, & Roebeling R. A.

(2010), Broken and inhomogeneous cloud impact on satellite cloud particle effective radius and cloudphase retrievals. *J. Geophys. Res.*, **115**, D10214, doi:10.1029/2009JD012205.

Wood R. (2012), Stratocumulus clouds, *Mon. Wea. Rev.*, **140**, 2373-2423.

Wood R, & Field P. R. (2011), The distribution of cloud horizontal sizes. *J. Clim.* **24**: 4800–4816, doi: 10.1175/2011JCLI4056.1.

Wood R., & Hartmann D. L. (2006), Spatial variability of liquid water path in marine boundary layer clouds: The importance of mesoscale cellular convection. *J. Climate*. **19**: 1748–1764. doi: http://dx.doi.org/10.1175/JCLI3702.1.

Wylie, D. P., & W. P. Menzel. (1999), Eight years of high cloud statistics using HIRS. *J. Climate*, **12**, 170–184.

Zelinka M. D., & Hartmann D. L. (2012), Climate feedbacks and their implications for poleward energy flux changes in a warming climate, *J. Clim*, **25**, 608-624.

Zhang Z., & Platnick S. (2011), An assessment of differences between cloud effective particle radius retrievals for marine water clouds from three MODIS spectral bands. *J. Geophys. Res.*, **116**, D20215, doi:10.1029/2011JD016216.

Zhang, Z., Ackerman A. S., Feingold G., Platnick S., R. Pincus, & Xue H. (2012), Effects of cloud horizontal inhomogeneity and drizzle on remote sensing of cloud droplet effective radius: Case studies based on large-eddy simulations. *J. Geophys. Res.*, **117**, D19208, doi:10.1029/2012JD017655.

Zheng X., Albrecht, B., Jonsson, H. H., Khelif, D., Feingold, G., Minnis, P., Ayers, K., Chuang, P., Donaher, S., Rossiter, D., Ghate, V., Ruiz-Plancarte, J., & Sun-Mack, S. (2011), Observations of the boundary layer, cloud, and aerosol variability in the southeast Pacific near-coastal marine stratocumulus during VOCALSREx, *Atmos. Chem. Phys.*, **11**, 9943–9959, doi:10.5194/acp-11-

9943-201.

Table Captions

Table 1 Summary of the 11 Flights

Note: the eleven flights from 2013 – 2015 were closely aligned with A-Train overpasses (i.e. MODIS and CALIPSO ground tracks). The average cloud properties are for liquid only (except ice fraction) by *in-situ* and MODIS observations. The MODIS r_{eff} is retrieved from the combined pixels with overcast and PCL. The standard deviation of r_{eff} is indicated in parenthesis. Drizzling flights are highlighted in light blue.

Table 2. Compositions of Cloud Phase for the Eleven Flights

Note: the composition of cloud phase was discriminated by aircraft measurements, CALIOP retrievals, and MODIS IR and CPOP product. The MODIS cloud phase is separated with single-layer cloud and all clouds including multi-layered clouds determined by MODIS "Cloud Multi-layer Flag" science data set (SDS). Cloud top temperature (CTT) and cloud top height (CTH) are the average values for the ice phase clouds for all cases. The three 'widely iced' cases are highlighted.

Table 3 MODIS Effective Radius Retrieval Failure Rates

Note: the failure rates were produced for drizzling/non-drizzling flights and cloud regime (overcast and PC) spatiotemporally aligned with *in-situ* observations from eleven flights from 2013-2015. The failure rate is broken down by causes of failure. All calculations are for single-layer liquid clouds determined by the MODIS CPOP and CMLF products.

Figure Captions

Figure 1.

Cloud images taken by a GoPro camera onboard the Cessna Conquest aircraft and observed by the MODIS on 14 June 2013 (a & b) and 03 Sep 2014 (c & d). Images (a) and (c) were taken just before the A-Train overpass when the aircraft was closely aligned with the CALIPSO ground track (04:20 and 04:19 UTC, respectively). The corresponding MODIS visible images, (b) and (d), indicate the location of the aircraft with a red cross.

Figure 2

Composite images of an aircraft flight track overlaid on Aqua MODIS retrievals for the 20130628 case. The flight track is identified in red on the maps. (a) and (c) are MODIS retrievals of r_{eff_21} for overcast and r_{eff_21} PCL pixels, respectively. (b) and (d) are the spatiotemporally collocated MODIS pixels and aircraft ground track. Blue dots are the MODIS cloud pixels. (e) The MODIS cloud phase optical properties (CPOP) with aircraft track. (f) The 3-D image of the thermodynamic phase from the *in-situ* observations and MODIS CPOP (overcast + PCL) retrievals.

Figure N_Vreff.

Figure 3

Vertical profiles of the CALIOP cloud phase for nine flights from 2013 – 2015. Two CALIOP images are shown for each flight; the first one with high altitude and the second zoomed in to target the coincident area. The red lines on the second image indicate the coincident altitude range. ROI, HOI and UK/ND indicate “randomly-oriented ice”, “horizontally-oriented ice”, and “unknown / not determined”.

Figure 4

(a) Comparison of liquid cloud effective radius (r_{eff}) observed from *in-situ* measurements (CAS + CIP) and retrieved from Aqua MODIS with the combined pixels (overcast + PCL). (b) r_{eff} differences between *in-situ* observations and MODIS retrievals (overcast + PCL) from 3 channels. (c) The MODIS r_{eff} (overcast + PCL) differences between the channels. (d) The

MODIS r_{eff} differences between overcast and PCL pixels for the 2.1 µm and 3.7 µm channels. Drizzling flights are indicated as crosses.

Figure 5

Comparison of the flight-mean liquid cloud effective radius (r_{eff}) between *in-situ* measurements from CAS + CIP and Aqua MODIS retrievals from the cloudy (overcast + PCL) pixels for (a) 1.6, (b) 2.1 and (c) 3.7 µm channels. The horizontal and vertical bars represent standard deviations of cloud r_{eff} from CAS+CIP and MODIS, respectively. The dashed lines represent 1:1 lines. Non-drizzling flights and drizzling flights are indicated as circles and triangles, respectively.

Figure 6

Vertical profiles of r_{eff} with normalised cloud altitude (Z / Z_{top}) for non-drizzling (a), lightly drizzling (b) and heavily drizzling (c) flights from *in-situ* measurements and the MODIS retrievals with 2.1 channel for overcast + PCL pixels. Z indicates the cloud height and Z_{top} indicates the cloud top height.

Figure 7

(a) The MODIS $r_{eff\_2.1}$ and $r_{eff\_3.7}$ and COT binned by CM_SPI for all 1km × 1km collocated cloudy (overcast + PCL) pixels. (b) The *in-situ* r_{eff} variations against log-normal standard deviation of drop size distribution.

Figure 8

(a) Normalized volume spectra for one non-drizzle flight (20130723) and two heavily drizzling flights (20150830 and 20130628). (b) The mean r_{eff} bias as a function of solar zenith angle (SZA).

Figure 9

(a), (b) and (c): same as Figure 5 but for N_d. N_d is derived by the MODIS retrievals of r_{eff} and cloud optical thickness (COT) from equation (1). (d), (e) and (f): N_d as a function of COT for the combined (overcast + PCL) pixels retrieved from Aqua MODIS. The horizontal bars and vertical bars represent standard deviations of N_d and COT, respectively. Non-drizzling flights and drizzling flights are indicated as circles and triangles, respectively.

Figure S1

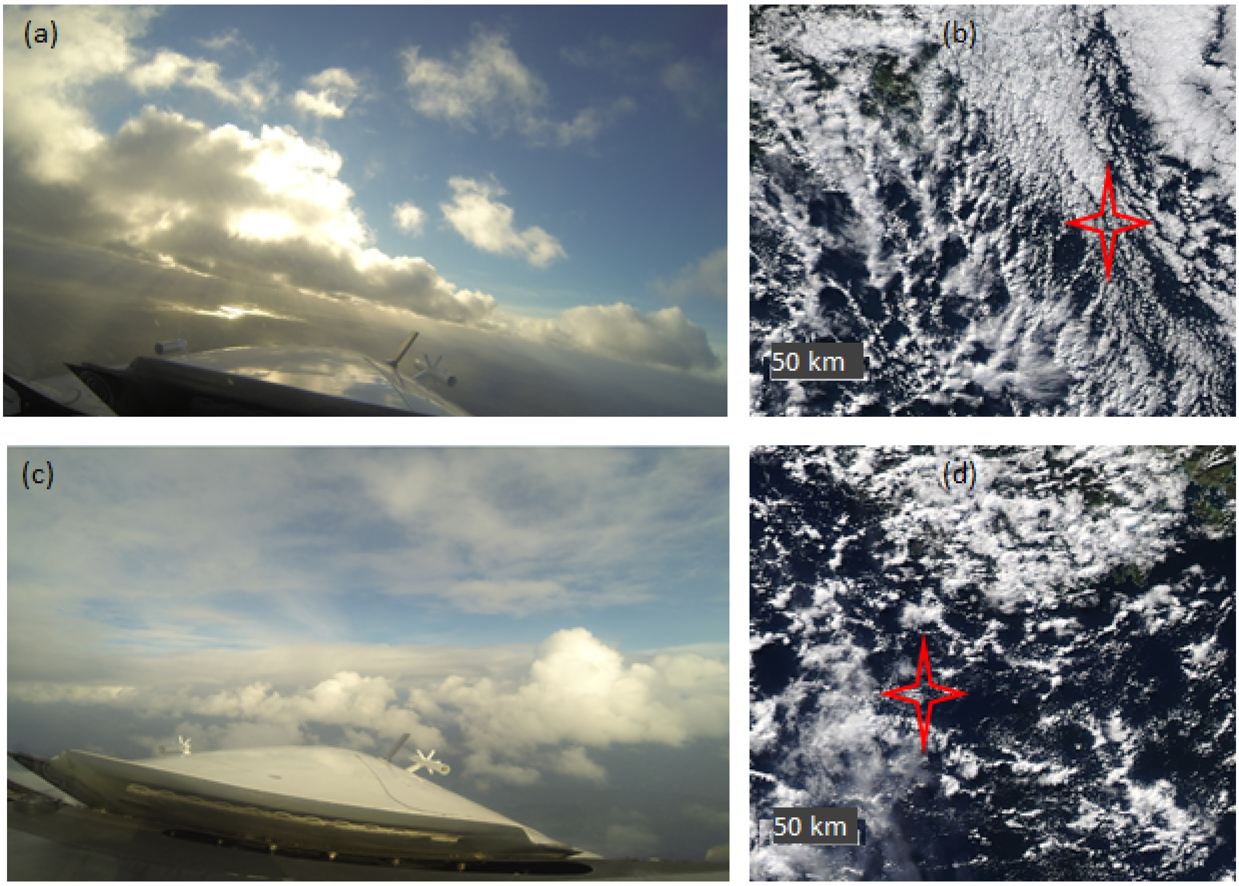
A conceptual diagram of the flight path: (a) consistent-level path (b) v-line path. The altitude ranges are indicated as ~0.5 km and ~2.5 km for (a) and (b), respectively. Nine flights largely followed the ‘consistent-level’ pattern, and two flights (20131011 and 20140903) followed the ‘v-line’ pattern, during the A-train overpass time window. The range of vertical adjustments during the consistent-level and v-line patterns were ~0.5 km and ~2.5km, respectively.

Figure S2

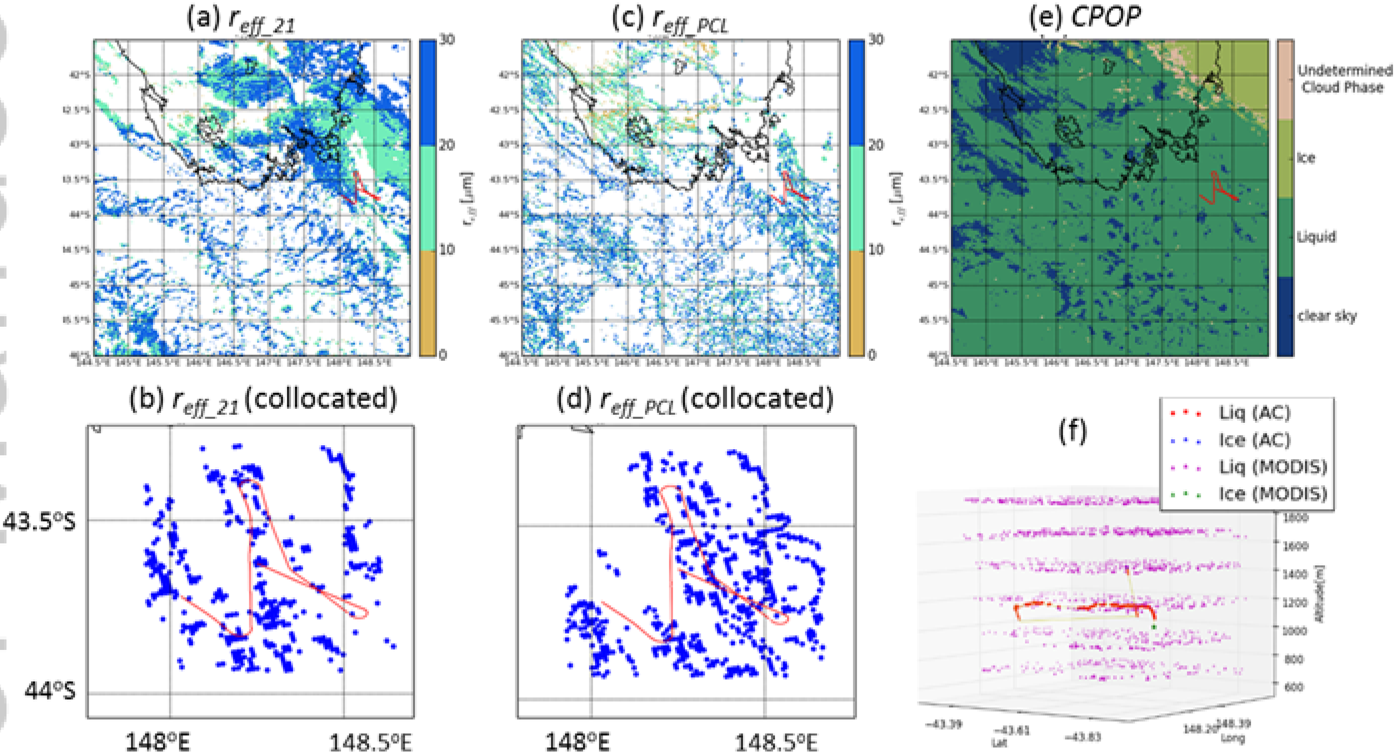
Composite 3-D images of aircraft track and MODIS cloud phase optical properties (CPOP) retrieval pixels for each flight which are spatiotemporally coincident. The aircraft track is plotted with yellow line with green star for starting and purple star for ending point for the collocation. Cloud phase determined by aircraft measurements and MODIS CPOP product is indicated with different colours.

Figure S3

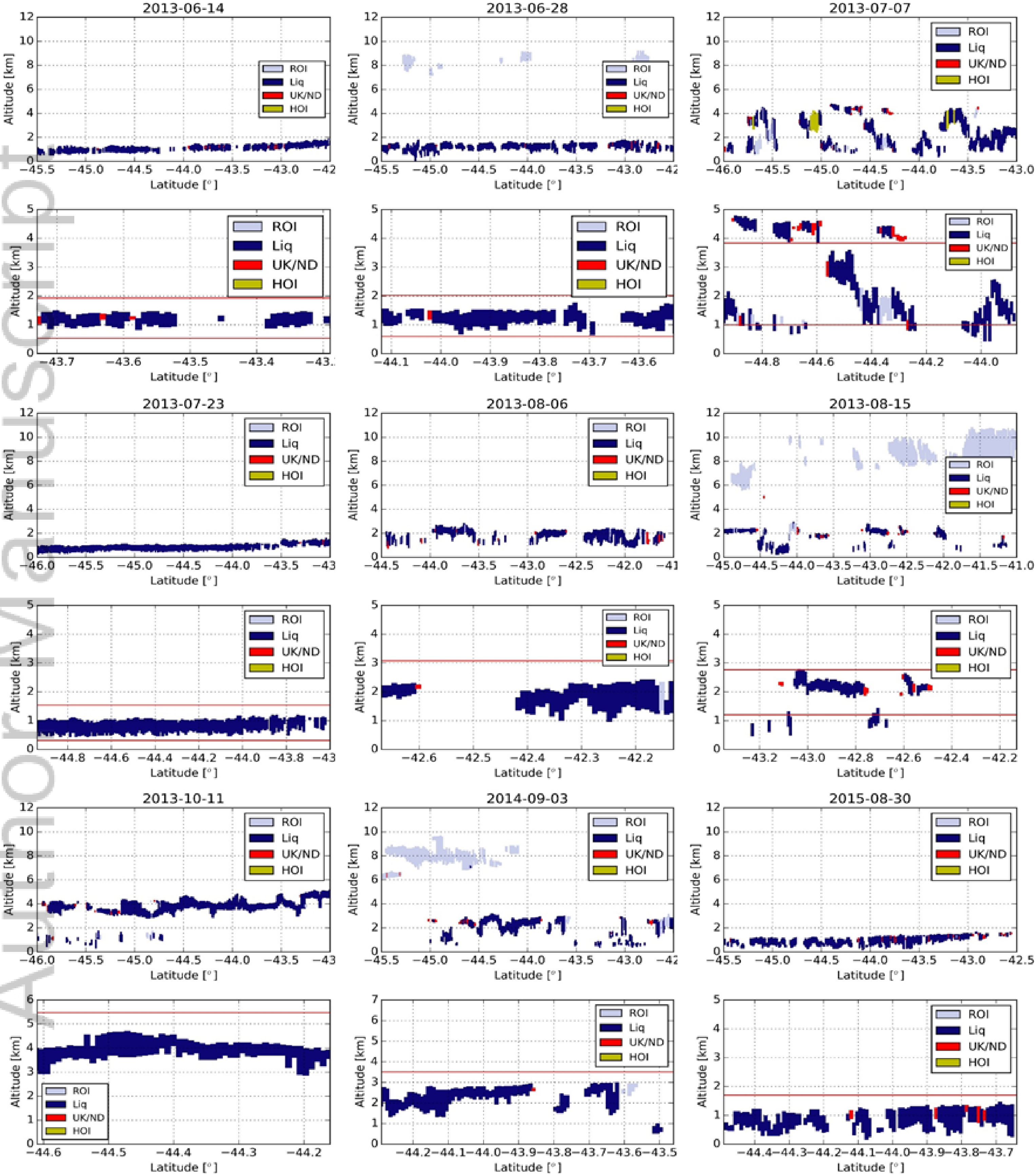
The CIP images of mixed/ice phase clouds for four flights which heavily or lightly experienced of ice.



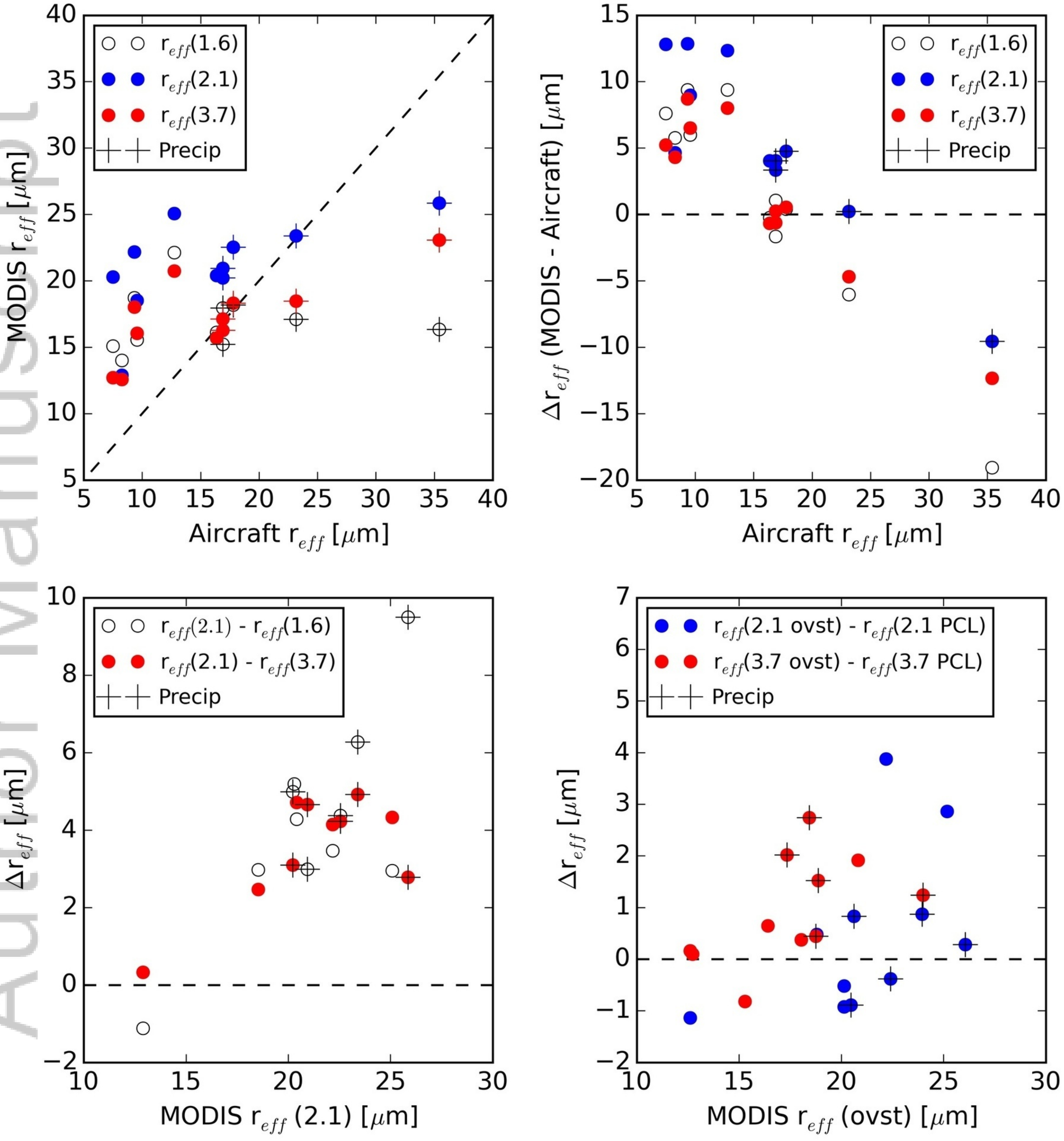
2018JD028535-f01-z-.tif



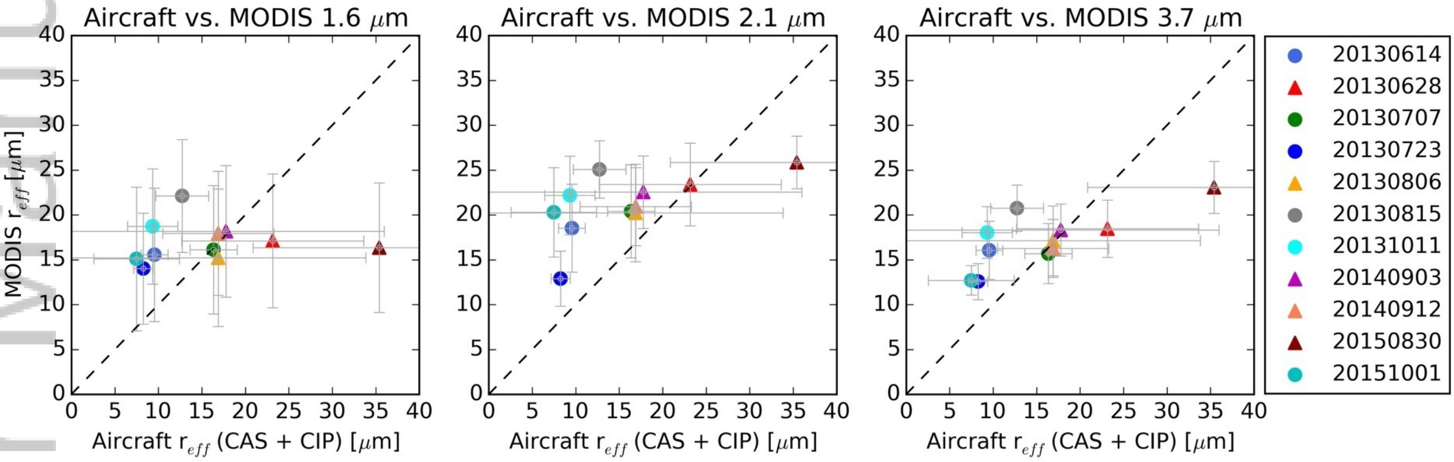
2018JD028535-f02-z-.tif



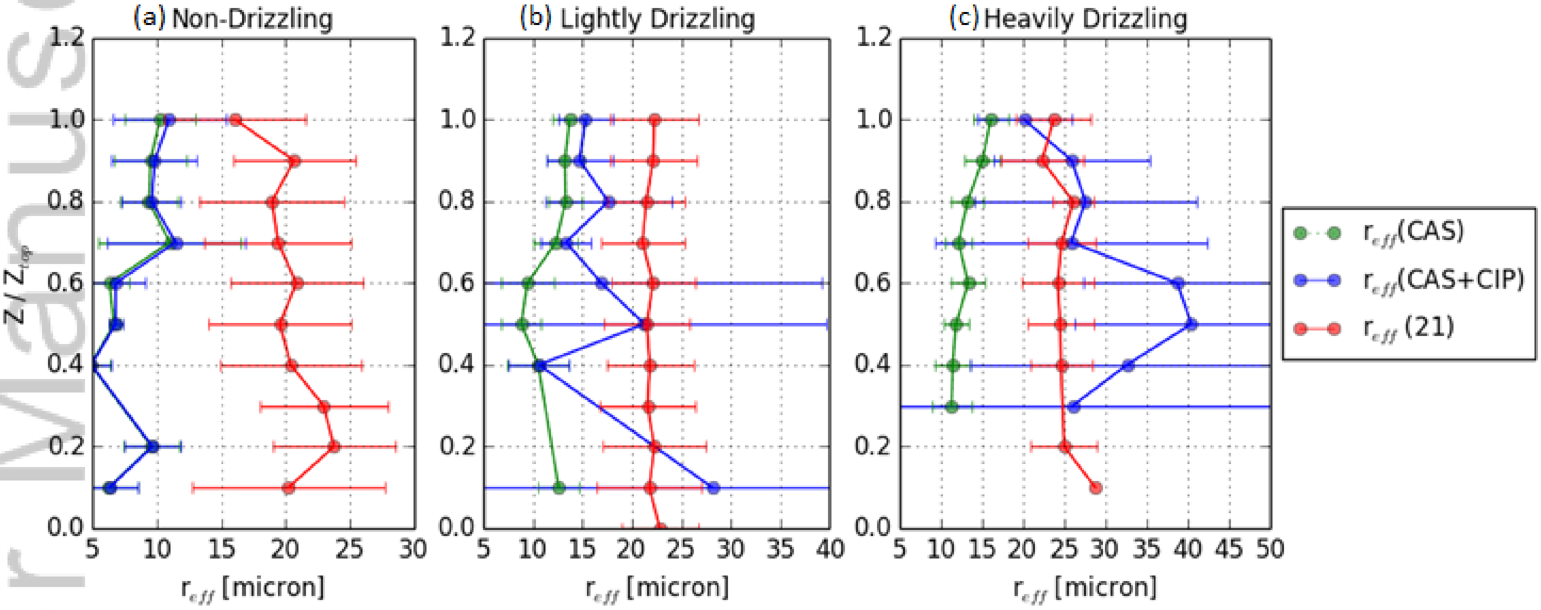
2018JD028535-f03-z-.tif



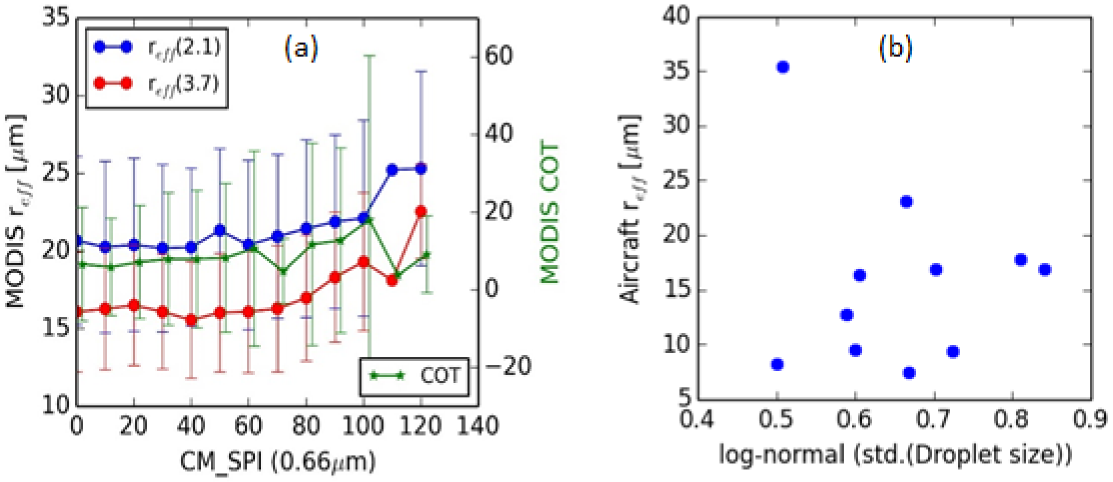
2018JD028535-f04-z-.jpg



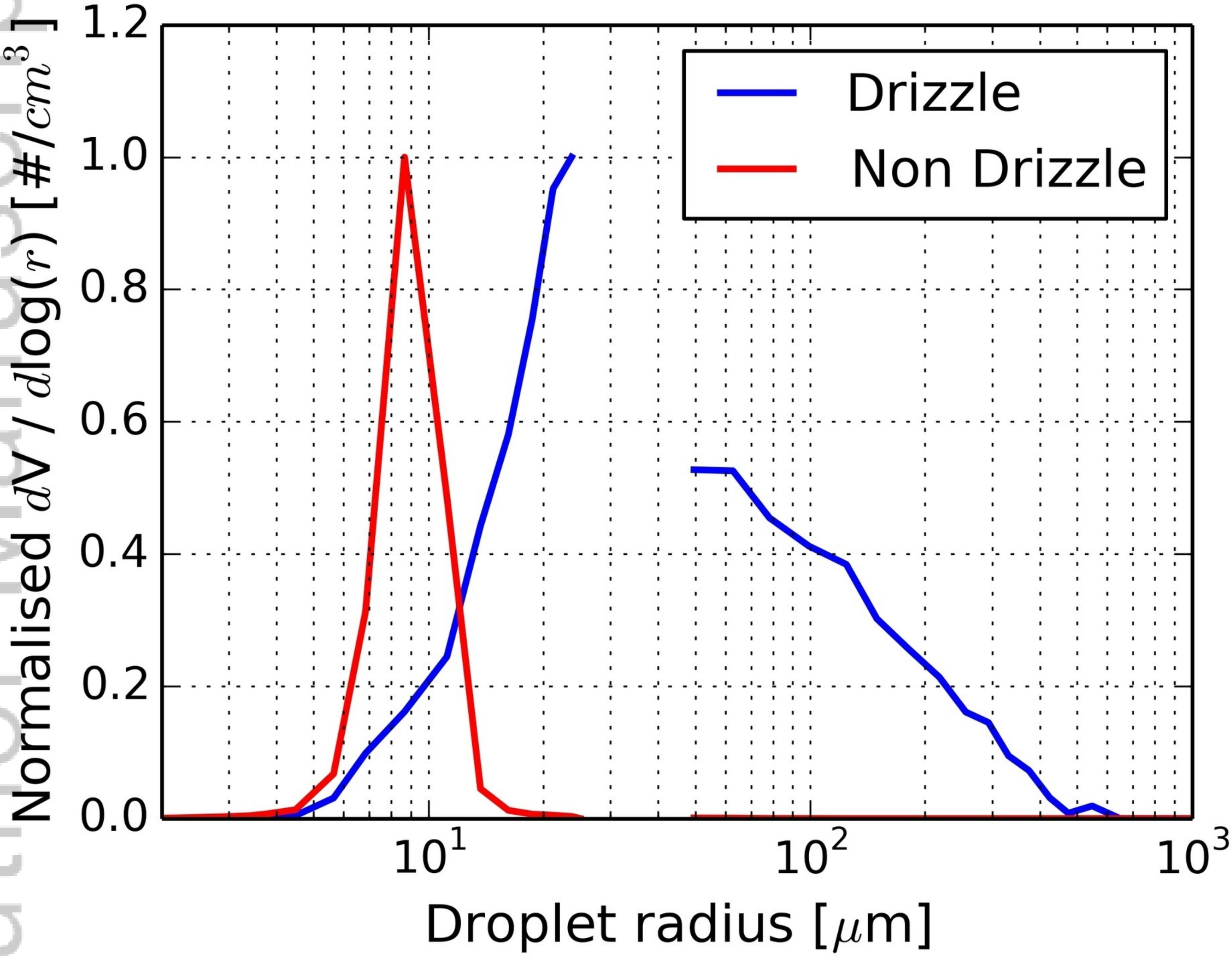
2018JD028535-f05-z-.jpg



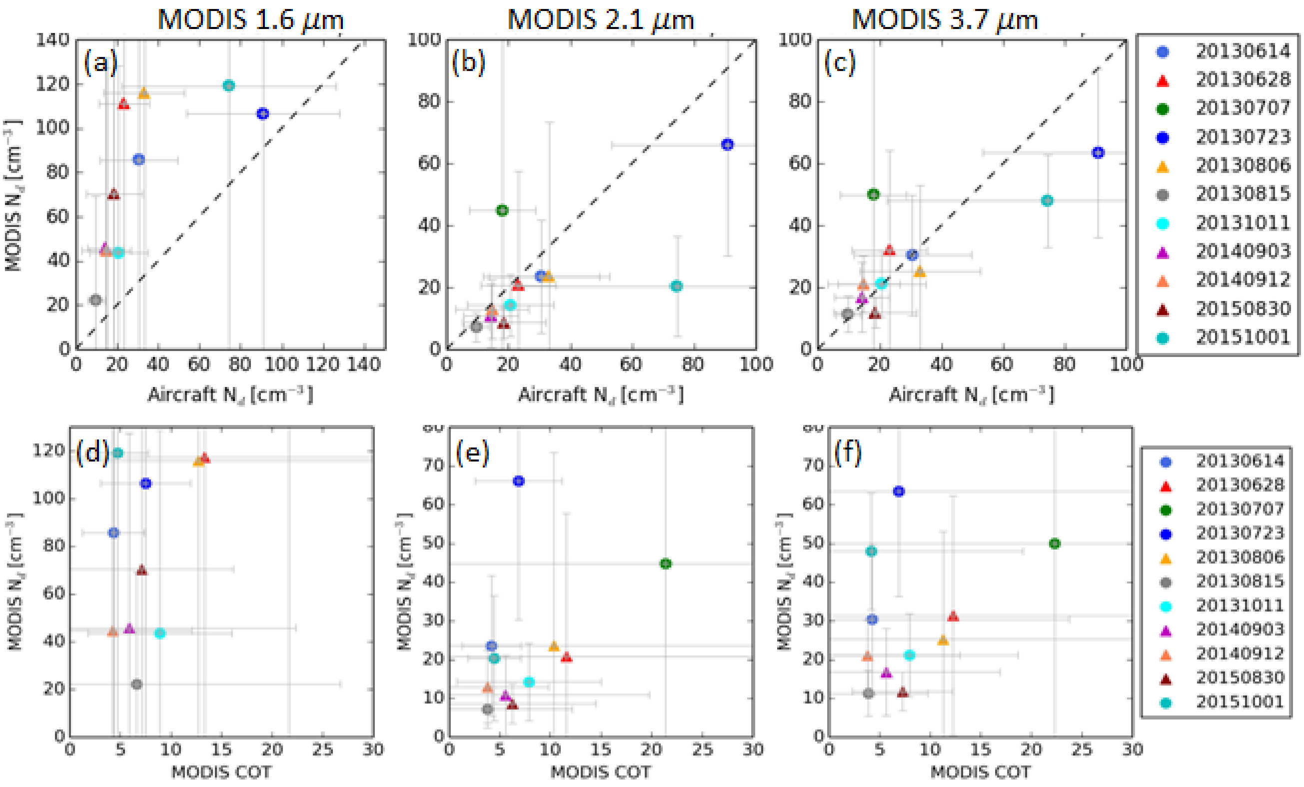
2018JD028535-f06-z-.tif



2018JD028535-f07-z-.tif



2018JD028535-f08-z-.jpg



2018JD028535-f09-z-.tif