Earth and Space Science

RESEARCH ARTICLE
10.1029/2020EA001137

Evaluation of Cloud Microphysical Properties Derived From MODIS and Himawari-8 Using In Situ Aircraft Measurements Over the Southern Ocean

Lijun Zhao[1], Chuanfeng Zhao[1], Yang Wang[1], Yuan Wang[2], and Yikun Yang[1]

[1]State Key Laboratory of Earth Surface Processes and Resource Ecology, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China, [2]Division of Geology and Planetary Science, California Institute of Technology, Pasadena, CA, USA

Key Points:

- Himawari-8 (MODIS) cloud product overestimates r_e near cloud top by 33% (26–31%) over Southern Ocean
- Himawari8 r_e and N_d agree better with aircraft measurements when considering only the in situ data acquired in the upper portions of clouds
- The performance of MODIS retrievals depends on the wavelength, with larger biases at 1.6 μm than 2.1 and 3.7 μm

Correspondence to:
C. Zhao,
czhao@bnu.edu.cn

Citation:
Zhao, L., Zhao, C., Wang, Y., Wang, Y., & Yang, Y. (2020). Evaluation of cloud microphysical properties derived from MODIS and Himawari-8 using in situ aircraft measurements over the Southern Ocean. *Earth and Space Science*, *7*, e2020EA001137. https://doi.org/10.1029/2020EA001137

Received 12 FEB 2020
Accepted 25 MAR 2020
Accepted article online 14 APR 2020

Abstract Cloud microphysical properties from aircraft measurements during the Southern Ocean Clouds, Radiation, Aerosol Transport Experimental Study are used to evaluate the cloud products from the geostationary satellite Himawari-8 (H-8) and the polar-orbiting satellite the Moderate Resolution Imaging Spectroradiometer (MODIS). Compared to the in situ aircraft observations when aircraft flew horizontally near cloud tops, the cloud droplet effective radius (r_e) and number concentration (N_d) from H-8 (MODIS) are 33% (26%–31%) and 2% (9–13%) larger. Both the H-8 and MODIS retrievals behave similarly for liquid-only and mixed-phase low-level clouds, indicating the weak sensitivity of the satellite cloud retrieval performance to cloud phase. The r_e and N_d of the cloud profiles from aircraft measurements were also used to compare with the satellite product. It shows that H-8 r_e and N_d agree better with aircraft measurements when considering only the in situ data acquired in the upper portions (highest 20%) of the clouds. Roughly, the r_e overestimation by H-8 decreases from 18% to 3% when considering the upper portions of clouds compared to all cloud layer averages, except for one case with drizzles appeared. In addition, the performance of MODIS r_e and N_d is highly dependent on the wavelengths the retrieval method uses. The droplet r_e retrievals using wavelength of 1.6 μm have much larger biases than that using the other two channels. The potential effects of the cloud vertical variation and the photon penetration depth, the cloud heterogeneity, the cloud droplet size spectra, and the drizzle on satellite retrievals have also been discussed.

1. Introduction

Clouds play a significant role on the global radiative energy balance, which has a cooling effect on the surface by reflecting solar radiation back into space and has a warming effect by trapping longwave radiation within the atmosphere (Schneider, 1972; Ramanathan et al., 1989; Zhao & Garrett, 2015). Especially important are the low-level marine stratocumulus clouds which cover more than 50% of ocean regions and therefore enhance the planetary albedo over the oceans and regulate energy fluxes affecting Earth's climate system (Cess, 1976; Klein & Hartmann, 1993; Lohmann & Feichter, 2005). Microphysical properties of clouds, such as the droplet effective radius (r_e), and droplet number concentration (N_d), are critical in determining the cloud radiative properties (Albrecht, 1989; Garrett & Zhao, 2006).

Many studies have been carried out to obtain cloud properties from in situ aircraft measurements, ship-based or ground-based remote sensing, and satellite remote sensing. Due to the large spatial coverage and long-term continuous records, satellite remote sensing is widely used in statistical analysis of both regional and global cloud properties, along with the model evaluation studies, compared to other types of measurements (Bao et al., 2018; Jiang et al., 2012; Li et al., 2018; Platnick et al., 2003; Rossow & Schiffer, 1991; Stephens et al., 2002; Winker et al., 2009; Zhao, Chen, et al., 2019). However, large uncertainties could exist in the satellite remote sensing observations. In order to assess the uncertainties, satellite observations are generally evaluated with ground-based remote sensing measurements or in situ aircraft measurements (Garrett & Zhao, 2013; Mace et al., 2004; Protat et al., 2009). Since the ground-based observations of cloud properties are also from remote sensing instead of in situ observations, the inherent issues also exist (Zhao et al., 2012, 2014), which need further evaluation by aircraft measurements.

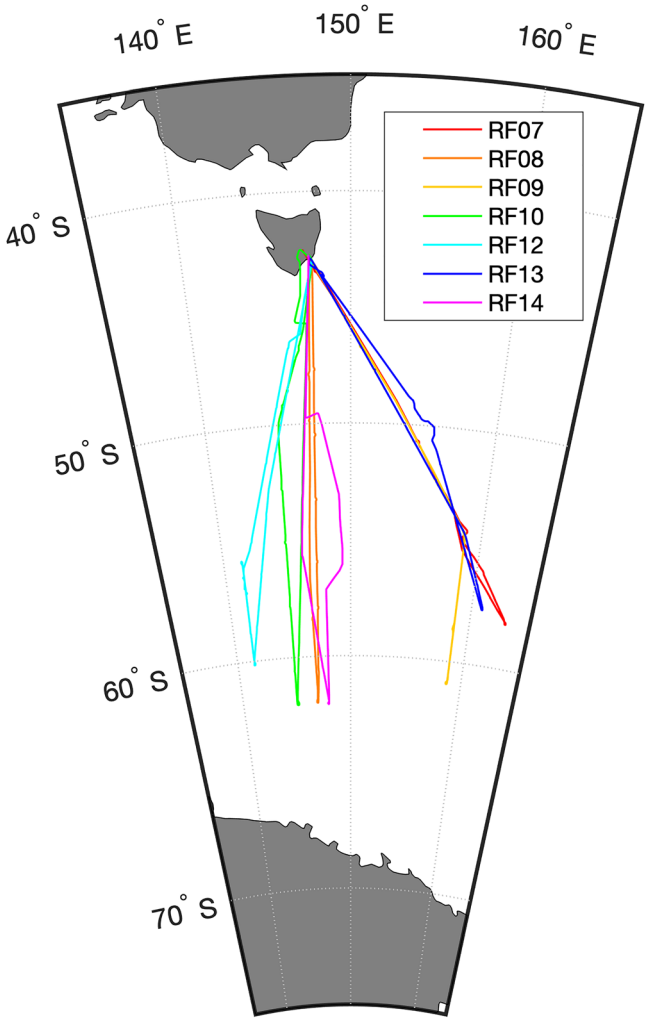

Figure 1. Flight tracks of seven research flights (RF) during SOCRATES. The gray color represents the land region, and the seven different colors represent tracks of seven flights over the Southern Ocean.

Many studies have evaluated the satellite cloud retrievals using in situ aircraft observations directly. Nakajima et al. (1991) showed that cloud droplet r_e obtained from Multispectral Cloud Radiometer was overestimated by about 2–3 μm compared to in situ observations for marine stratocumulus over southern California, while Nakajima and Nakajima (1995) showed that new method could make the satellite retrievals consistent with the in situ aircraft observations. Painemal and Zuidema (2011) found that the Moderate Resolution Imaging Spectroradiometer (MODIS) retrievals overestimated the cloud droplet r_e by about 15–20% or 2.1 μm compared to in situ aircraft observations for marine stratocumulus over the Southeast Pacific. Ahn et al. (2018) found that CALIPSO and MODIS underestimated mixed-phase cloud occurrence, and MODIS (2.1-μm channel) overestimated r_e by about 13 μm for nondrizzling clouds and by about 10 μm for heavy drizzling cases. Using different near-infrared bands, the satellite-based r_e could also vary significantly due to the different photon penetration depths. Nakajima et al. (2010a) further indicated that the existence of drizzle and small cloud droplets near cloud top may also affect the satellite retrievals. All of these previous studies suggested a potentially high uncertainty of cloud properties in satellite retrievals.

So far, few studies have been carried out for the cloud properties and processes over the Southern Ocean region, including both evaluation study and statistical study of cloud properties. Haynes et al. (2011) found that low-level clouds have important effect on the radiation balance over the top of atmosphere over the Southern Ocean based on model simulation study. Given the limited knowledge about cloud processes and properties over the Southern Ocean, many studies have shown the high uncertainties in climate model simulations of clouds over Southern Ocean, resulting in large SW radiation biases (Mace, 2010; Trenberth & Fasullo, 2010). Of course, the identification of cloud phase and supercooled liquid droplets is also an important issue (Hu et al., 2010; Li et al., 2017; Morrison et al., 2011). Thus, accurate information about cloud properties are needed over the Southern Ocean. However, due to the remote geographical location, few ground-based remote sensing or in situ aircraft measurements of cloud properties have been carried out, making it challenging to evaluate and then improve the performance of satellite-based cloud retrievals over the Southern Ocean.

Therefore, this study aims to evaluate the performance of MODIS and Himawari-8 cloud products using the aircraft observations from the Southern Ocean Clouds, Radiation, Aerosol Transport Experimental Study (SOCRATES). SOCRATES was conducted during the period from 3 January to 26 February 2018 in the region with latitude between Hobart, Australia (42°S) and 61°S and with longitude from 134°E to 163°E. The paper is organized as follows. Section 0 describes the data and method. Section 0 shows the results of comparison study. Section 0 analyzes the potential error sources in satellite retrievals. And section 0 presents the summary and conclusions.

2. Data and Method

2.1 In situ Measurements

There are totally 15 research flights by the Gulfstream-V High-performance Instrumented Airborne Platform for Environmental Research aircraft during the SOCRATES. By excluding the flights with additional layers of clouds above and flights with small size cumulus clouds, seven research flights are employed in this study. Figure 1 shows the seven flight tracks (denoted by different colors) in the study region with latitudes between 42.51°S and 62.08°S and longitudes between 141.35°E and 163.03°E. A suite of cloud probes was installed onboard the Gulfstream-V High-performance Instrumented Airborne Platform for

Table 1
In Situ Airborne Instruments, Along With the Company, Measured Variables, Measurement Range, Uncertainty, and the References for the Instruments and Their Uncertainties

Instrument	Company	Measurement	Range	Uncertainty	References
King hot-wire probe (King)	PMS, DMT	LWC	0.05–3 g m^{-3}	15%	King et al. (1978)
Cloud droplet probe (CDP)	DMT	LWC, N_t, PSDs	2–50 (μm)	20%	Lance et al. (2010)
Two-dimensional stereo probe (2DS)	SPEC	LWC, N_t, PSDs	10–1,260 (μm)	NF	Lawson et al. (2006), Lawson (2011)
Closed-path hygrometer (CLH-2)	University of Colorado	TWC	0.005–1 g m^{-3}	15%	Davis et al. (2007)
Particle Habit Imager and Polar Nephelometer (PHIPS)	KIT	Image cloud particles	20–700 (μm)	NF	Abdelmonem et al. (2016)

Note. NF in "uncertainty" denotes that the uncertainty information is not found in literature.

Environmental Research to measure cloud properties. Details about the probes are listed in Table 1 and described below.

The Cloud Droplet Probe (CDP) is an optical probe which measures cloud droplet size distribution in 30 size bins ranging from 2 to 50 μm in diameter (D) (Lance et al., 2010) with an uncertainty of 20% (Faber et al., 2018). The Two-Dimensional Stereo Particle Imaging Probe (2DS) is an optical array probe which uses images to get cloud particle size distribution with D ranging from 10 to 1,260 μm (Lawson, 2011; Lawson et al., 2006).

The King hot-wire probe (King) is designed for measuring the ice water content and liquid water content (LWC) with an uncertainty of 15% (King et al., 1978). It works with relatively large uncertainties during collision with ice particles due to the residual effect of ice (Korolev et al., 1998), which makes hot wire observations likely unreliable when clouds are mixed or ice phase. King probe has smaller LWC than CDP (Painemal & Zuidema, 2011), and the King probe gets unsensitive to droplets with sizes larger than 40 μm (Biter et al., 1987). Here, the LWC derived from King was used to be compared with the LWC integrated from the CDP droplet size distributions. As shown in Figure 2, the LWCs from these two probes agree well with each other when they are less than 0.3 g m^{-3}, particularly for the liquid phase. For cases with LWC larger than 0.3 g m^{-3}, the agreement becomes worse, implying the decreasing reliability of the King probe. Figure 2 also suggests that the LWCs from the two probes agree much better for liquid-phase clouds than for mixed-phase clouds, simply because the ice in the mixed-phase clouds may affect the King probe by possible contamination (Faber et al., 2018). To avoid this potential contamination error, the LWC from CDP is used in this study. The condensed total water content (TWC) was measured by a Closed-Path Hygrometer (CLH-2) with an uncertainty of 15% (Davis et al., 2007).

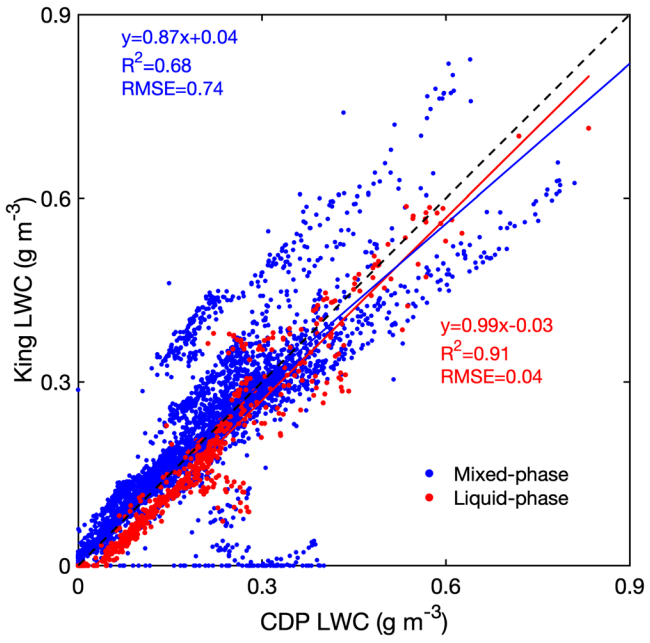

Figure 2. The comparison of LWCs between the CDP and the King hot-wire for both liquid-phase (red color) and mixed-phase (blue color) clouds, measured by the seven research flights over the Southern Ocean region in January and February 2018.

The Particle Habit Imaging and Polar Scattering probe (Abdelmonem et al., 2016) data were used to discriminate ice particles and droplets based on angular light scattering measurements (Schnaiter et al., 2018). The appearance of liquid droplets in clouds was defined when the LWC is greater than 0.005 g m^{-3} in a previous study by Baumgardner and Rodi (1989). Slightly different, the existence of cloud is identified when the TWC is greater than 0.01 g m^{-3} in this study. The fraction of LWC to TWC is used to distinguish mixed-phase and liquid-phase clouds (Dorsi et al., 2015; Field et al., 2004; McFarquhar et al., 2007). When the fraction is larger than 0.85, the cloud is assumed to be liquid phase; otherwise, it is assumed to be mixed phase.

The cloud droplets r_e and N_d are often derived based on the spectral number concentration measurements as follows:

$$r_e = \frac{\int r^3 n(r)\mathrm{d}r}{\int r^2 n(r)\mathrm{d}r} = \frac{\sum N_{ci} r_i^3}{\sum N_{ci}\, r_i^2} \quad (1)$$

$$N_d = \int n(r)\mathrm{d}r = \sum N_{ci} \quad (2)$$

where $n(r)$ is cloud droplet number concentration at radius r and N_{ci} and r_i are the cloud droplet number concentration and radius at the ith bin of the instrument, respectively. In order to investigate the impacts of large droplets such as drizzle, we have also calculated the cloud droplet r_e and N_d by using the combination of CDP from 2–50 μm and 2DS from 50–200 μm. Note that the first two bins from CDP and the first four bins from 2DS have been neglected due to the large potential uncertainties, such as the uncertainties in the probe's depth of field for small size particles in 2DS (Baumgardner & Korolev, 1997). By combining the CDP and 2DS, we can get the cloud droplet spectra with diameters ranging from 2 to 1,280 μm.

2.2 Satellite Measurements

Himawari-8 (H-8) was launched on 7 October 2014 and operated at 140° east (Bessho et al., 2016; Letu et al., 2020). The Advanced Himawari Imager has 16 observational bands, from the visible to thermal infrared spectral regions with a nadir spatial resolution of 0.5 km (1 band), 1 km (2 bands), and 2 km (13 bands). The Advanced Himawari Imager has been used for the retrieval of cloud properties, including the cloud optical depth (COD), r_e, cloud top temperature, and cloud top height (Iwabuchi et al., 2018; Letu et al., 2018). The H-8 beta data, which were announced to have serious retrieval issue by its official team, are the only available one from the official website during the period of SOCRATES. Thus, we use the H-8 cloud retrieval data set that was produced specifically to support SOCRATES by the Satellite ClOud and Radiative Property Retrieval System (SatCORPS) at NASA Langley Research Center (hereafter referred as SatCORPS H-8). The Level-2 (L2) cloud properties in the SatCORPS H-8 product are provided at a nearly 2-km × 2-km spatial resolution and a 10-min temporal resolution, which include cloud droplet r_e, COD, cloud top height, cloud type, and cloud top temperature.

MODIS instrument was launched on Terra satellite in 1999 and on Aqua in 2002. MODIS has 36 spectral bands measuring visible and infrared radiation based on which properties including aerosol, water vapor, cloud, atmospheric profile, and cloud mask have been retrieved (Platnick et al., 2003). MODIS cloud retrieval properties from MYD06 L2 product with a spatial resolution of 1 km × 1 km are evaluated in this study, including the cloud r_e and COD at three channels of 1.6, 2.1, and 3.7 μm. Note that only the MYD06 L2 product onboard Aqua is used here simply because the aircraft cases in this study occurred in the afternoon.

Cloud N_d is not directly provided by the SatCORPS H-8 and MODIS cloud products, so we calculate N_d using equation 3 (Painemal & Zuidema, 2011) following the correcting method from Albrecht et al. (1990),

$$N_d[r_e, \tau] = 1.4067 \times 10^{-6} \cdot \frac{\tau^{\frac{1}{2}}}{r_e^{\frac{5}{2}}} \tag{3}$$

where τ and r_e are COD and effective radius from satellite retrievals, respectively.

2.3 Methods

2.3.1 Case Selection

In order to reliably evaluate the satellite observations, aircraft measurements (usually with homogeneous low cloud scenes) have been selected with following two criteria, making only seven flights used in this study. First, all cases with extra layers of clouds above the one aircraft measured were removed to exclude the contamination of upper clouds based on the camera observations onboard aircraft. Second, two kinds of cases are investigated in this study, which are cloud measurements at the top of clouds when aircraft flew horizontally within around 100–200 m below cloud tops and cloud measurements when aircraft flew vertically within clouds. After selecting the aircraft observation cases, we average the variables obtained every 10 s, simply because the average aircraft flying speed is 130 m s^{-1} and thus the flight distance in 10 s nearly matches the satellites' spatial resolutions.

2.3.2 Colocation of Aircraft and Satellites Observations

To evaluate the satellite performance with in situ aircraft observations, it is necessary to match the observations from the two systems in both location and time. Therefore, the satellite observations within 5 km and 10 min centered at each aircraft measurement were selected and used in this analysis. Figure 3 shows two examples illustrating how we do the matching of observations between satellite and aircraft observations. Figures 3a and 3b show the matching process between the SatCORPS H-8 cloud r_e and the aircraft (rf07) measurements at 04:55:00–4:59:50 UTC on 31 January 2018, and Figures 3c and 3d show that between MODIS r_e at 2.1 μm and the aircraft (rf12) measurements at 04:36:00–04:38:30 UTC on 17 February 2018.

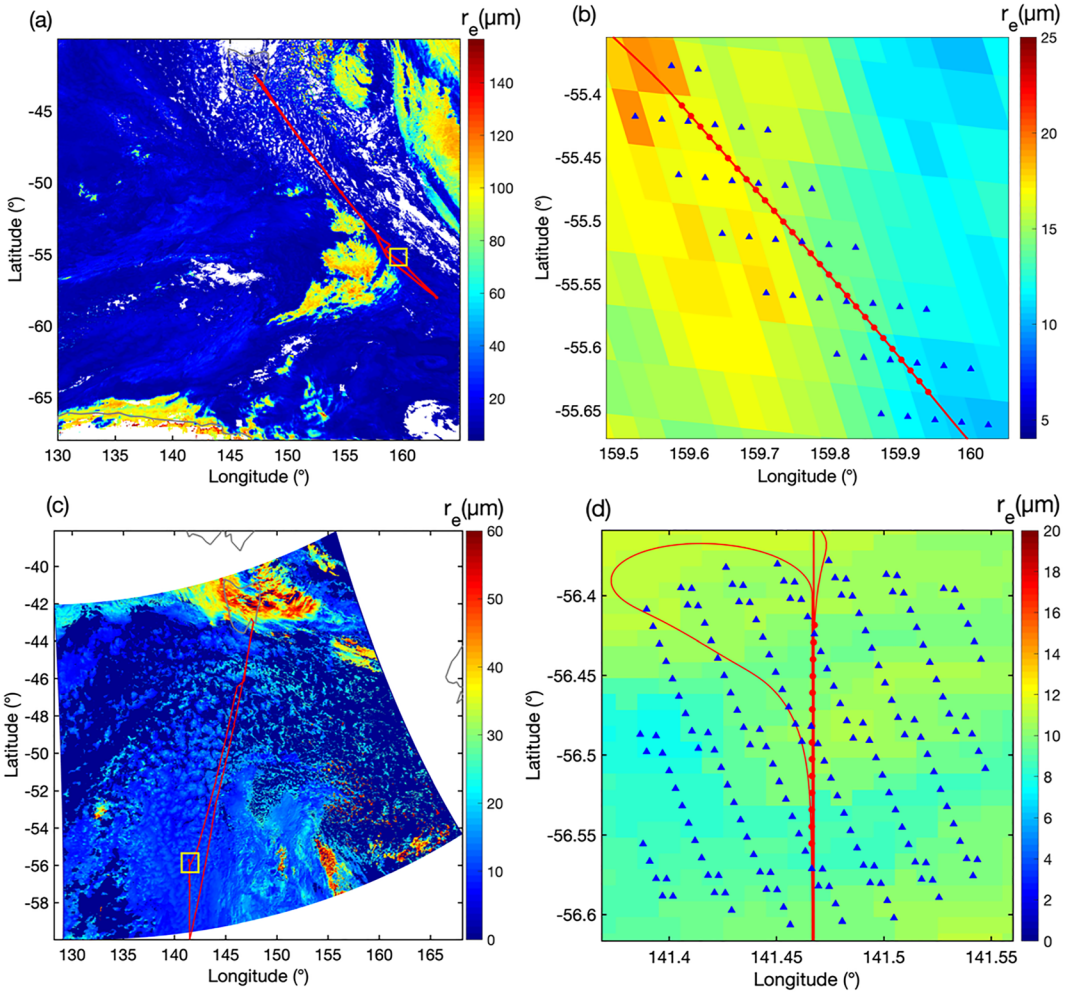

Figure 3. Two cases depicting the matching strategy for comparing satellite and aircraft observations. (a, b) For MODIS satellite and aircraft flight rf07 conducted at 04:55:00–5:59:50 UTC on 31 January 2018. (c, d) For Himawari-8 satellite and aircraft flight rf12 conducted at 04:36:00–04:38:30 UTC on 17 February 2018. The images depict the cloud effective radius derived from (a) MODDIS and (c) Himawari-8 measurements. The yellow squares are the study domain selected for the two cases. Figures 3b and 3d further illustrate the matching strategy, in which the red line represents the G-V flight tracks, the red circles represent cases of aircraft locations collocated with satellite overpass, and the blue triangles are the collocated satellite pixels, respectively.

In Figures 3a and 3c, the yellow boxes represent the specific case study regions used to illustrate the matching process, and the red lines represent the flight tracks. Figures 3b and 3d show the matching results for two cases. The red points represent the centered aircraft locations with the matched satellite blue points scattered around. Considering the cloud homogeneity and potential error of matching, satellite measurements at all collocated pixels are averaged and compared with the aircraft observations at the matching points.

2.3.3 Evaluation Metrics

Two evaluation metrics are defined to evaluate the performance of satellite retrievals, which are bias and relative mean bias (RMB). These two metrics are calculated as follows:

$$bias = \frac{1}{n}\sum_{i=0}^{n}\left(x_s^i - x_a^i\right) \tag{4}$$

$$RMB = \bar{x_s}/\bar{x_a} \tag{5}$$

where x represents cloud properties such as cloud droplet r_e and N_d, x_s indicates cloud properties retrieved by satellite, x_a indicates cloud properties observed by aircraft, i means the ith sample with total number of

Table 2
Details About the Mean Cloud Macrophysical and Microphysical Properties for All 19 Aircraft Observation Cases Selected in This Study

RF	Phase	T (°C)	H (km)	CT (km)	r_{e_CDP} (μm)	$r_{e\_CDP+2DS}$ (μm)	N_d (cm^{-3})	LWC$_{_CDP}$ (g cm^{-3})	LWC$_{_2DS}$ (g cm^{-3})	Drizzle (%)
07T	Liquid	−8.4	1.7	—	8.26	8.58	58	0.13	0.01	5
07A	Mixed	−7.5	1.6	0.6	11.84	13.44	42	0.20	0.05	39
08A	Liquid	−14.5	1.8	0.4	7.21	7.22	195	0.23	0	0
08D	Mixed	−13.8	1.8	0.6	6.60	6.63	255	0.26	0	0
09T	Mixed	−19.5	2.6	—	6.23	6.34	114	0.11	0.02	10
09A	Liquid	−8.8	1.7	0.7	7.36	7.87	189	0.28	0.03	18
10T	Mixed	−3.1	1.3	—	7.63	7.71	48	0.08	0.03	0
12A	Mixed	−6.2	1.3	0.5	7.91	8.45	138	0.24	0.03	2
12T	Mixed	−6.8	1.4	0.2	7.34	7.58	160	0.22	0.03	7
12T	Mixed	−6.5	1.4	—	8.48	8.58	186	0.40	0.01	7
12T	Mixed	−6.9	1.4	—	7.08	7.17	164	0.19	0.01	2
12T	Liquid	−6.8	1.4	—	8.07	8.18	156	0.28	0.01	10
12T	Mixed	−6.9	1.3	—	7.56	7.58	158	0.24	0	0
12D	Mixed	−4.1	1	0.7	7.42	10.48	125	0.13	0.03	39
13A	Mixed	−0.5	0.7	0.3	8.36	8.71	128	0.28	0.04	0
13D	Mixed	1.9	0.4	0.4	9.05	9.75	159	0.19	0.03	15
13A	Mixed	−1.8	1	0.4	9.14	9.33	154	0.50	0.05	0
13T	Mixed	0.1	1	—	7.14	7.32	206	0.27	0.02	0
14T	Mixed	−2.5	1.1	—	8.62	8.80	56	0.09	0.01	0

Note. The RF column shows the research flight ID, in which the ending letter A or D indicates ascending or descending profiles, respectively, and the T denotes cases near the cloud top. T and H denote the average temperature and height for all aircraft measurements within a cloud case, and CT denotes the cloud thickness. Drizzle appearance is represented as percentage. r_e, N_d, and LWC represent cloud droplet effective radius, number concentration, and liquid water content, respectively. Two types of r_e and LWC have been calculated based on different instrument measurements.

samples denoted by n, and $\bar{}$ represents the mean value. The bias indicates the average deviation of satellite retrievals compared to the in situ measurements. RMB indicates the average uncertainty estimation of the satellite retrievals relative to the aircraft estimates, where RMB > 1 and RMB < 1 represent overestimation and underestimation, respectively. In addition to these two metrics, we also make linear fitting regression analysis with coefficient of determination (R^2) between satellite and in situ observations.

3. Evaluation Analysis of Satellite Retrievals

We evaluate the performance of SatCORPS H-8 and MODIS cloud product based on the in situ observations from the seven flights. Since the satellite passive remote sensing retrievals often obtain cloud properties more representative of those within tops of clouds, we will analyze the aircraft observations for both vertical profiles of cloud properties and cloud properties near cloud tops. Table 2 lists the details of cloud properties for 19 aircraft cases selected in this study. Note that the cloud microphysical properties listed in Table 2 are all averaged values for the corresponding cases. The cloud base and top are defined as the lowest and highest levels with TWC from CDP larger than 0.01 g m^{-3} for each case.

Table 2 shows that most clouds observed during the SOCRATES are mixed-phase clouds, with only four cases as pure liquid-phase clouds. Moreover, there are totally 10 cases with horizontal aircraft observations near cloud tops and 9 cases constructed for vertical profiles through clouds among which 6 flight profiles were ascending and 3 flight profiles were descending. For most cases, the cloud temperature measured on the aircraft is between −20 and 0 °C. All clouds belong to low-level clouds as mentioned above (at heights below 3 km). The cloud depths of these clouds are generally below 1.0 km, most of which are stratocumulus. The cloud droplet r_e from CDP and 2DS mainly lies between 6 and 10 μm, with only two cases with r_e larger than 10 μm. Associated with the relatively clean background, the average cloud N_d is generally below 200 cm^{-3}, with most values slightly above 100 cm^{-3}. The cloud LWC from the CDP is between 0.1 and 0.5 g m^{-3}, which is much larger than that from 2DS, suggesting the dominance of small cloud droplets. Based on the temperature and liquid water content observations, supercooled liquid water below freezing temperature exists in 14 case observations. With these cloud property observations, we next evaluate the performance of satellite retrievals.

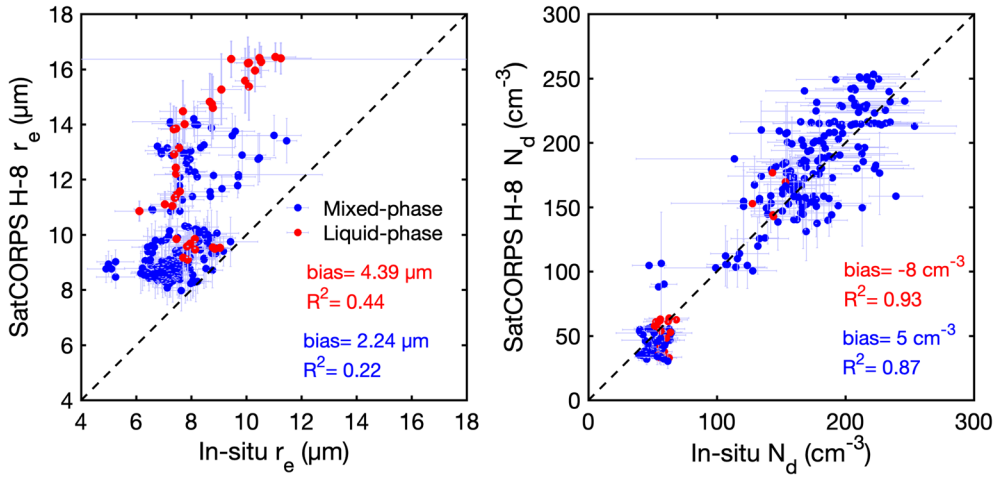

Figure 4. Scattered plot of cloud droplet effective radius (r_e) and droplet number concentration (N_d) between SatCORPS H-8 cloud retrieval product and in situ measurements near cloud tops, in which blue colors are for mixed-phase clouds and red colors for liquid-phase clouds. The dots and error bars represent the means and standard deviations (10 s) of in situ aircraft and satellite-observed cloud properties. The dashed line represents the 1:1 line.

3.1 Evaluation of the SatCORPS H-8 Cloud Product

3.1.1 Using Aircraft Observations Near Cloud Tops

We first evaluate the performance of the SatCORPS H-8 cloud products using the in situ aircraft observations near cloud tops from CDP. Figure 4 shows scatter plots of cloud r_e and N_d between the SatCORPS H-8 product and in situ aircraft observations near cloud tops. Note that the dots and error bars represent the means and standard deviations of the data. There is a total of 252 samples used for this comparison. The data in Figure 4 show that the in situ averaged r_e is 7.74 μm with a standard deviation of 0.50 μm, and the in situ average N_d is 137 cm^{-3} with a standard deviation of 19 cm^{-3} near cloud tops. It also shows that the in situ aircraft observed r_e (N_d) ranges from 4.91 μm (39 cm^{-3}) to 11.47 μm (254 cm^{-3}). In comparison, SatCORPS H-8 r_e (N_d) ranges from 7.98 μm (30 cm^{-3}) to 16.44 μm (254 cm^{-3}). These results show that SatCORPS H-8 gives relatively broad ranges of cloud r_e than that from aircraft observations but gives similar range of N_d. In general, the in situ cloud droplet r_e and N_d are smaller than that retrieved from SatCORPS H-8 cloud product, particularly for cloud droplet r_e. Quantitatively, the SatCORPS H-8 r_e retrieval shows an overestimation of 33% (RMB = 1.33) and the N_d retrieval shows an overestimation of only 2% (RMB = 1.02), relative to the in situ data. Note that the slight overestimation of N_d by SatCORPS H-8 is not significant.

We further investigate the performance of SatCORPS H-8 retrievals for pure liquid and mixed phase clouds, respectively. For liquid-phase clouds, SatCORPS H-8 cloud r_e shows a high overestimation of 52% (RMB = 1.52) with $R^2 = 0.44$, and N_d shows a very slight underestimation with RMB = 0.91 and $R^2 = 0.93$. In contrast, for mixed-phase clouds, SatCORPS H-8 cloud droplet r_e shows an overestimation of 29% (RMB = 1.29) with $R^2 = 0.22$ and N_d shows a very slight overestimation of 3% (RMB = 1.03) with $R^2 = 0.87$.

3.1.2 Using Vertical Profiles of Aircraft Observations

We also make a comparison study of the cloud microphysical properties between the in situ aircraft observations and satellite retrievals for cases when aircraft flew through the vertical extent of the clouds. Figure 5 shows the vertical profiles of average cloud microphysical properties including LWC, r_e, and N_d. Note that the altitude has been normalized by setting cloud bases as 0 and cloud tops as 1. Similar to the findings over various locations by previous studies (Qiu et al., 2017; Yang et al., 2019; Zhao et al., 2018; Zhao, Zhao, et al., 2019), both cloud LWC and r_e increase with height within clouds due to the adiabatic or near-adiabatic cooling. Moreover, cloud droplet N_d varies little with heights within clouds. It is found that LWC and N_d decrease slightly near the cloud top, which could be associated with entrainment near the cloud top. In contrast, cloud droplet r_e still increases with height near cloud top, which remains difficult to explain and requires further study in future. The in situ average r_e (N_d) in cloud is 8.6 μm (166 cm^{-3}) with a standard deviation of 1.4 μm (50 cm^{-3}).

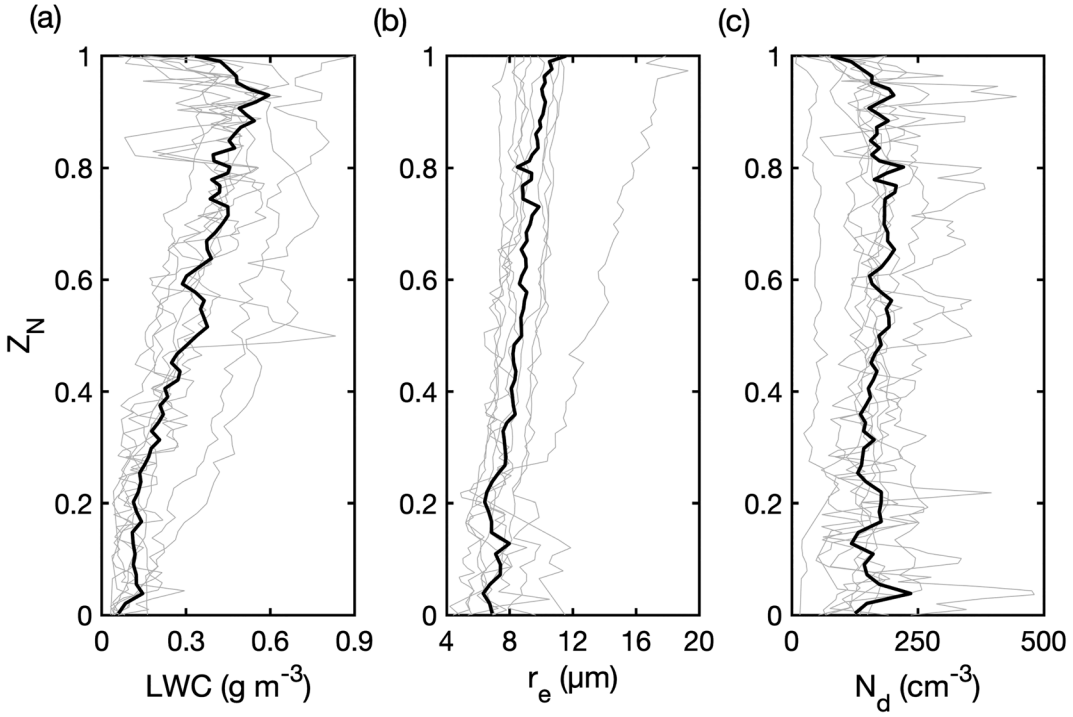

Figure 5. Normalized profiles of nine cases: (a) LWC, (b) r_e, and (c) N_d. Black lines indicate the mean profiles. $Z_N = 0$ indicates the cloud base, whereas $Z_N = 1$ indicates the cloud top.

Figure 6 shows the intercomparison of cloud droplet r_e and N_d between SatCORPS H-8 cloud product and aircraft observations for cases when aircraft flew through the clouds vertically, including both liquid-phase clouds and mixed-phase clouds. The measurements have been averaged within the whole vertical profile. The SatCORPS H-8 cloud droplet r_e is about 1.18 μm larger than that from aircraft observations, and the SatCORPS H-8 cloud droplet N_d is about 60 cm^{-3} more than that from aircraft observations. In other words, the SatCORPS H-8 cloud droplet r_e shows an overestimation of 18% (RMB = 1.18) with $R^2 = 0.59$ and N_d shows a slight underestimation with RMB = 0.99 and $R^2 = 0.70$. Figure 6 further implies that the comparison results seem to have very weak sensitivity to cloud phases.

Considering that the satellite retrievals may be more representative of cloud properties near cloud tops, the results shown above might be not surprising. We further analyzed the cloud top data from the in situ profiles by calculating the r_e and N_d using the in situ measurements in the highest 20% of the profile (i.e., $r_{e\text{-top}}$ and $N_{d\text{-top}}$) with the corresponding SatCORPS values. As shown in Figures 6c and 6d, the SatCORPS H-8 cloud retrievals are more consistent with the in situ aircraft observations found near cloud tops. Quantitatively, SatCORPS H-8 cloud $r_{e\text{-top}}$ was overestimated only about 3% (RMB = 1.03) with $R^2 = 0.17$ and SatCORPS H-8 cloud N_d was just slightly overestimated about 0.5% (RMB = 1.005) and $R^2 = 0.27$, compared with in situ aircraft measurements. R^2 could increase to 0.90 without the point squared by orange dotted line, which is one case occurring at 4:44:22–4:45:35 UTC on 31 January 2018 with much larger droplets in cloud. The much larger droplets than satellite retrievals found by the aircraft at 4:44:22–4:45:35 UTC on 31 January 2018 could be related to their sample difference: aircraft observations for small volume of cloud samples while satellite retrievals for a large area of clouds. This implies the importance of space/time mismatch for the intercomparison of cloud properties between satellite retrievals and aircraft observations.

3.2 Evaluation of MODIS Cloud Product

3.2.1 Using Aircraft Observations Near Cloud Tops

Same as the evaluation study of SatCORPS H-8 product, we first evaluate the performance of MODIS retrievals using the in situ observations when aircrafts flew horizontally within 100–200 m below cloud tops. Due

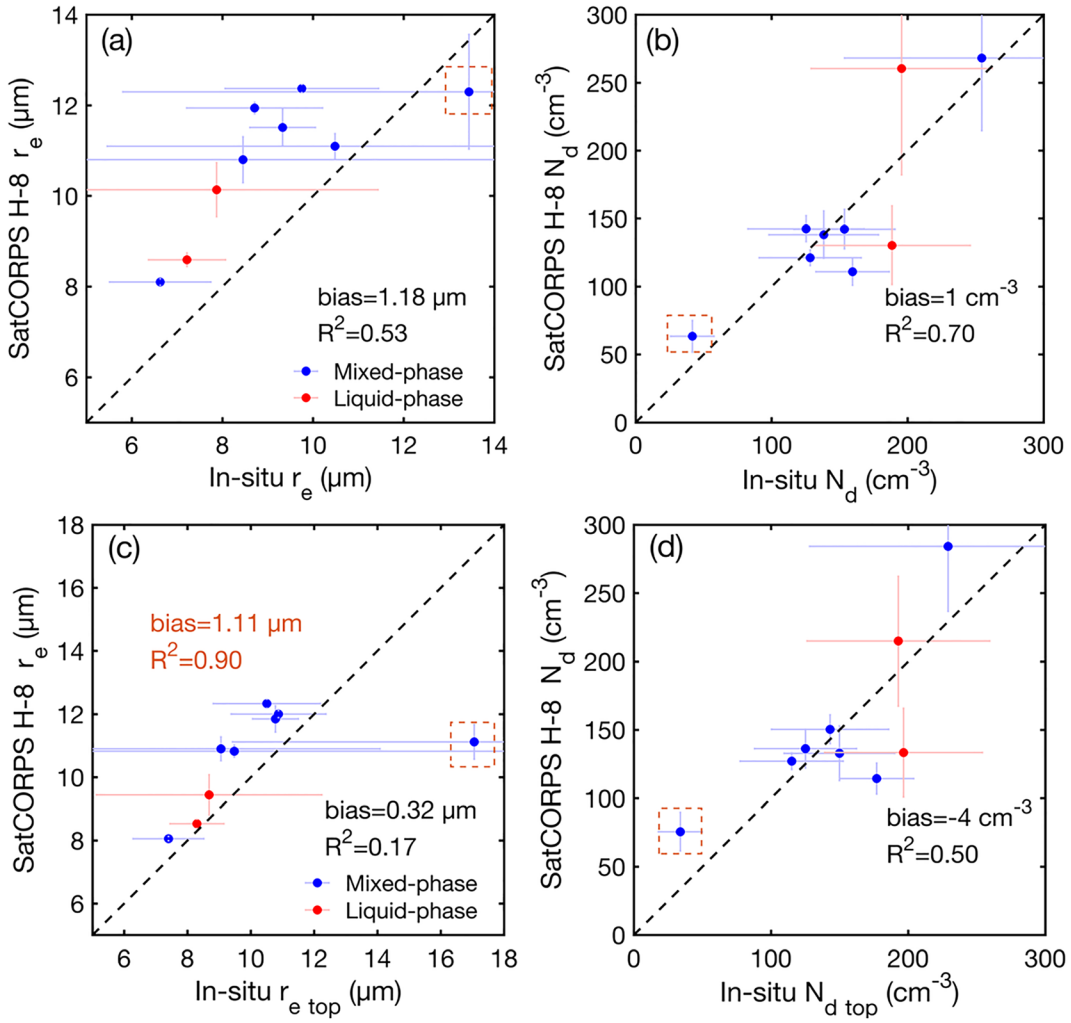

Figure 6. (a–d) Scatter plots of in-cloud r_e, in-cloud $r_{e\text{-top}}$ (averaged r_e for the top 20% of the cloud) in cloud N_d, and in-cloud $N_{d\text{-top}}$ (averaged N_d for the top 20% of the cloud) between in situ measurements and SatCORPS H-8 cloud products. The horizontal error bars represent the 10-s standard deviations of in situ r_e and N_d, and vertical error bars represent the standard deviation of collocated SatCORPS H-8 r_e and N_d.

to the limited revisiting times during a day by the MODIS, only four horizontal flight cases of rf12 are identified with aircraft observations collocated with MODIS when it overpassed.

Figure 7 shows the comparison of cloud droplet r_e and N_d near cloud tops between aircraft observations and MODIS retrievals using 1.6-, 2.1-, and 3.7-μm channels. The dots and error bars represent the 10-s means and standard deviations of in situ aircraft and satellite observations of cloud r_e and N_d. The in situ aircraft observations show that cloud droplet r_e varies between 6.08 and 9.43 μm with average and standard deviation values of 7.66 and 0.46 μm near cloud tops, respectively, and cloud N_d varies between 120 and 243 cm^{-3} with average and standard deviation values of 164 and 22 cm^{-3}, respectively. The MODIS cloud r_e shows similar ranges for retrievals based on different channels of 1.6 μm ($r_{e1.6}$), 2.1 μm ($r_{e2.1}$), and 3.7 μm ($r_{e3.7}$), which are 8.51–11.69, 8.61–10.98, and 8.67–11.01 μm, respectively, while they are larger than that measured by the aircraft. Differently, MODIS cloud N_d retrievals at 1.6 μm ($N_{d1.6}$) range from 127 to 234 cm^{-3}, slightly less than that at 2.1 μm ($N_{d2.1}$) and 3.7 μm ($N_{d3.7}$), which are 158–214 and 150–212 cm^{-3}, respectively. MODIS cloud droplet $N_{d1.6}$ is roughly consistent with the in situ aircraft observations, while $N_{d2.1}$ and $N_{d3.7}$ are slightly larger (not significant) than the in situ aircraft observations. Differently, the MODIS cloud droplet r_e is

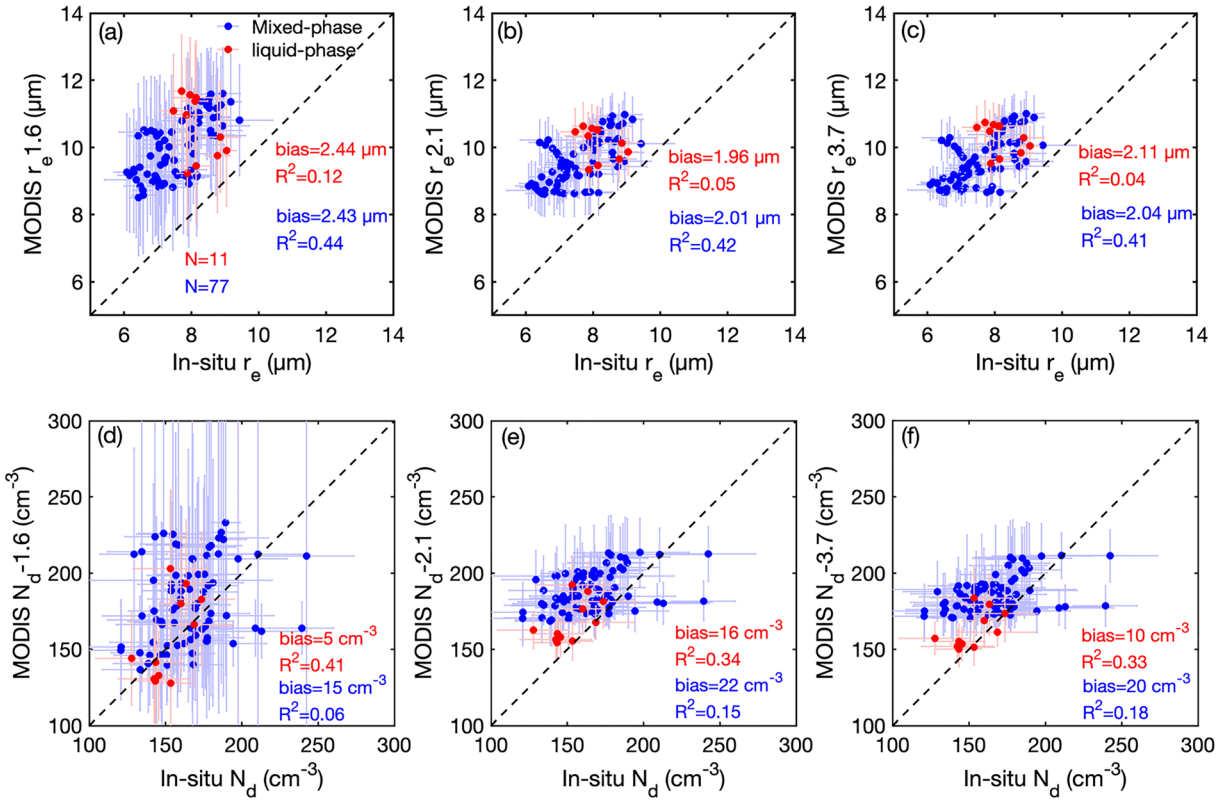

Figure 7. Comparison of mean cloud droplet r_e and N_d between in situ aircraft CDP and 2DS measurements near cloud top and MODIS retrievals using 1.6, 2.1, and 3.7 μm channels. Blue color is for mixed-phase clouds, and red color is for liquid-phase clouds. The dots and error bars represent the 10-s means and standard deviations of in situ aircraft and satellite observations.

significantly larger than that measured by aircraft, which is about 2.1–2.6 μm overestimated. Specifically, compared to the in situ aircraft observations, the MODIS $r_{e1.6}$ is about 31% overestimated with a positive bias of 2.44 μm, $r_{e2.1}$ about 26% overestimated with a positive bias of 2.00 μm, and $r_{e3.7}$ about 27% overestimated with a positive bias of 2.05 μm. The MODIS $N_{d1.6}$ is about 9% overestimated with a positive bias of 14 cm^{-3}, $N_{d2.1}$ about 13% overestimated with a positive bias of 22 cm^{-3}, and $N_{d3.7}$ about 12% overestimated with a positive bias of 19 cm^{-3}.

We also investigated the impact of cloud phase on the performance of MODIS cloud retrievals in Figure 7. It is clear that cloud phase plays weak impact on the performance of MODIS cloud retrievals of r_e and N_d. MODIS cloud droplet $r_{e1.6}$ is overestimated nearly 30% (32%) with a positive bias of 2.44 μm (2.43 μm) for liquid-phase (mixed-phase) clouds compared to in situ aircraft observations. Similar to the findings shown above, the biases are larger than that retrieved at 2.1-μm channel (24% for liquid-phase cloud and 26% for mixed-phase cloud) and 3.7-μm channel (26% for liquid-phase cloud and 27% for mixed-phase cloud). In contrast, MODIS cloud droplet $N_{d1.6}$ is overestimated nearly 3% (9%) with a positive bias of 5 cm^{-3} (15 cm^{-3}) for liquid-phase (mixed-phase) clouds compared to in situ aircraft observations. Similar to the findings shown above, the biases are larger than that retrieved at 2.1-μm channel (11% for liquid-phase cloud and 14% for mixed-phase cloud) and 3.7-μm channel (6% for liquid-phase cloud and 12% for mixed-phase cloud).

We further examined the differences of MODIS retrieved cloud droplet r_e at three channels of 1.6, 2.1, and 3.7 μm by using $r_{e1.6}$–$r_{e2.1}$ and $r_{e3.7}$–$r_{e2.1}$, which are shown in Figure 8. It is clear that $r_{e1.6}$ is much larger than $r_{e2.1}$, while $r_{e3.7}$ is close to $r_{e2.1}$. Similar results have been reported in earlier studies that the MODIS $r_{e1.6}$ is larger than $r_{e2.1}$ and $r_{e3.7}$. For example, Haney (2013) and Liang et al. (2015) found that $r_{e1.6}$ is much larger than $r_{e3.7}$, while $r_{e1.6}$ was found to be similar to $r_{e2.1}$ in trade wind clouds of the tropical western Atlantic. In contrast, Zhang and Platnick (2011) found that $r_{e3.7}$ is similar to $r_{e2.1}$ for the most homogeneous clouds,

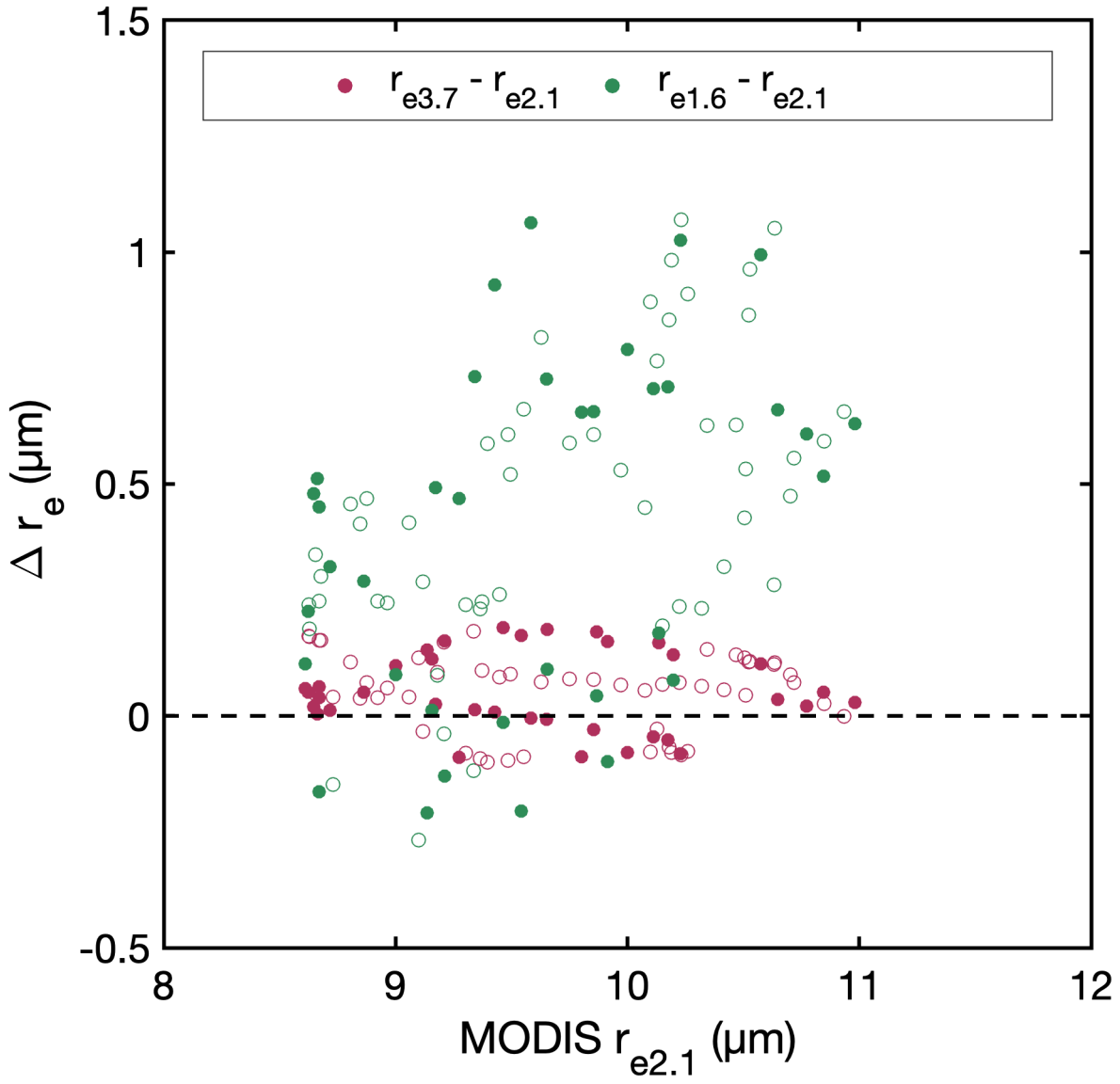

Figure 8. MODIS retrieved cloud droplet r_e differences among three channels. The green color is the difference between $r_{e1.6}$ and $r_{e2.1}$; the amaranth color is the difference of $r_{e3.7}$ and $r_{e2.1}$. Filled circles represent cloud samples with drizzle appearance. Hollow circles are samples without drizzle happening.

while $r_{e2.1}$ is larger than $r_{e3.7}$ up to ~10 μm for the most heterogeneous clouds. The finding that $r_{e2.1}$ and $r_{e3.7}$ are similar for many of the SOCRATES comparisons might indicate that the clouds studied here are more prone to be homogeneous. Since the cloud absorbs the thermal radiation the least at 1.6-μm channel, $r_{e1.6}$ represents the deepest layer from cloud top. About 3.7 μm is the most absorptive channel among the three by clouds, and thus, $r_{e3.7}$ represents the shallowest layer from cloud top (Chang & Li, 2002). We have shown some evidences for entrainment effects near cloud tops (e.g., Figure 4). Below the entrainment layer, cloud LWC and r_e generally increases with altitudes. This vertical distribution of cloud properties along with the entrainment effect could make $r_{e3.7}$ smaller, similar, or even larger than $r_{e2.1}$ or $r_{e1.6}$ depending on the effective photon penetration depth below the cloud tops. The results shown in Figure 8 might imply that $r_{e3.7}$ and $r_{e2.1}$ are more representative of the upper portion clouds with a significant contribution from the entrainment layer, while $r_{e1.6}$ is more representative of a deeper layer below the entrainment layer.

The influence of drizzle on the changes of cloud liquid droplet r_e with wavelength channels has been investigated in Figure 8. Note that the drizzle is defined as those liquid droplets with diameters larger than

100 μm. It seems that the performance differences of cloud liquid property retrievals among three channels have weak sensitivity to the existence of drizzle. Since the MODIS retrieval is based on the spectral radiation in near-infrared channels, the contribution of drizzle to near-infrared radiation is much less than that from the smaller and more cloud droplets due to their limited amount, suggesting the possibility that in some cases the cloud droplet r_e and N_d retrievals may be weakly dependent on the existence of drizzles. Instead, it is more dependent on the vertical structure of cloud droplets.

3.2.2 Using Vertical Profiles of Aircraft Observations

There are only two flight cases (rf07 20180131 4:44:20–4:45:47 and rf09 20180205 4:57:00–4:59:10) that aircraft flew vertically through the clouds when MODIS satellite overpassed the same location. Figure 9 shows the vertical profiles of cloud droplet r_e, N_d, and LWC from the aircraft observations for these two cases (a-c for rf07 and d-f for rf09), along with the colocated MODIS average retrieval results. Note that two types of in situ aircraft observations of cloud droplet r_e are shown in Figure 9, one is based on the CDP and the other is based on the combination of CDP and 2DS. Roughly, r_e from CDP represents the cloud droplet r_e without the impact of drizzle, and r_e from CDP and 2DS represents the cloud r_e with the impact of drizzle. As expected, the cloud droplet r_e from CDP + 2DS is larger than that from CDP. However, the difference is very small, indicating that the large drizzle droplets are much less than the cloud droplets. The rf07 case is for mixed-phase cloud which is dominated by relatively small number of large droplets with mean r_e of 14.92 μm and mean N_d of 36 cm^{-3}. In contrast, the rf09 case is for pure liquid-phase cloud which has smaller cloud droplets with mean r_e of 8.81 μm and mean N_d of 178 cm^{-3}. Note that the impacts of drizzle are not for all layers, with more impacts to low and middle parts of clouds.

For the mixed-phase cloud measured by rf07, MODIS cloud droplet r_e agrees well with the in situ aircraft observations using any of the three channels at normalized heights higher than 0.8, especially for channel of 3.7 and 2.1 μm, while it is obviously larger than the in situ aircraft observations at low and middle parts of clouds. As indicated earlier, the satellite observations are more representative of cloud properties near cloud tops. Thus, MODIS cloud r_e seems reasonable for the mixed-phase cloud with large droplets. In contrast, the MODIS cloud N_d is larger than the aircraft observations near cloud top, while roughly agrees with the aircraft observations within clouds. For the pure liquid-phase cloud measured by rf09, the MODIS cloud droplet r_e retrieved using any wavelength channel is clearly overestimated compared to the in situ aircraft observations at all altitudes. By average, the mean cloud droplets r_e from MODIS at three channels are 11.23, 11.18, and 10.63 μm, while the average cloud r_e from in situ aircraft observation is only 8.81 μm. Differently, the MODIS cloud droplet N_d agrees well with the in situ aircraft observations. However, it is challenging to identify the exact error sources for the mismatch of cloud properties. In addition to the impacts from the vertical structure of cloud properties, the horizontal spatial variability of cloud properties as indicated in Figure 3 could be another potential influential factor considering the mismatch in locations between satellite and aircraft observations.

4. Error Analysis for Satellite Retrieval Performance

For satellite passive retrieval, it is generally assumed that the cloud is vertically homogenous and the retrieval variables represent the single bulk value of the whole cloud. However, there are considerable variations of cloud vertical structure and the shape of droplet size distribution in realistic clouds (Miles et al., 2000; Nakajima et al., 2010a). In other words, the satellite passive retrievals are heavily weighted toward cloud tops. In contrast, in situ aircraft measurements can provide detailed information of vertical variations of cloud properties (Miles et al., 2000). Various influential factors could affect the evaluation of satellite retrievals, such as the cloud homogeneity, the droplet size spectra, the appearance of drizzle, and the solar zenith angle. Here, we briefly discuss the possible influences on the retrieved r_e from the cloud homogeneity, the droplet size spectra, and the appearance of drizzle.

4.1 Cloud Heterogeneity

Liang et al. (2009) indicated that the cloud heterogeneity could play important roles to the MODIS cloud retrievals. We here investigate the impact of cloud heterogeneity (H_σ), which is defined as follows:

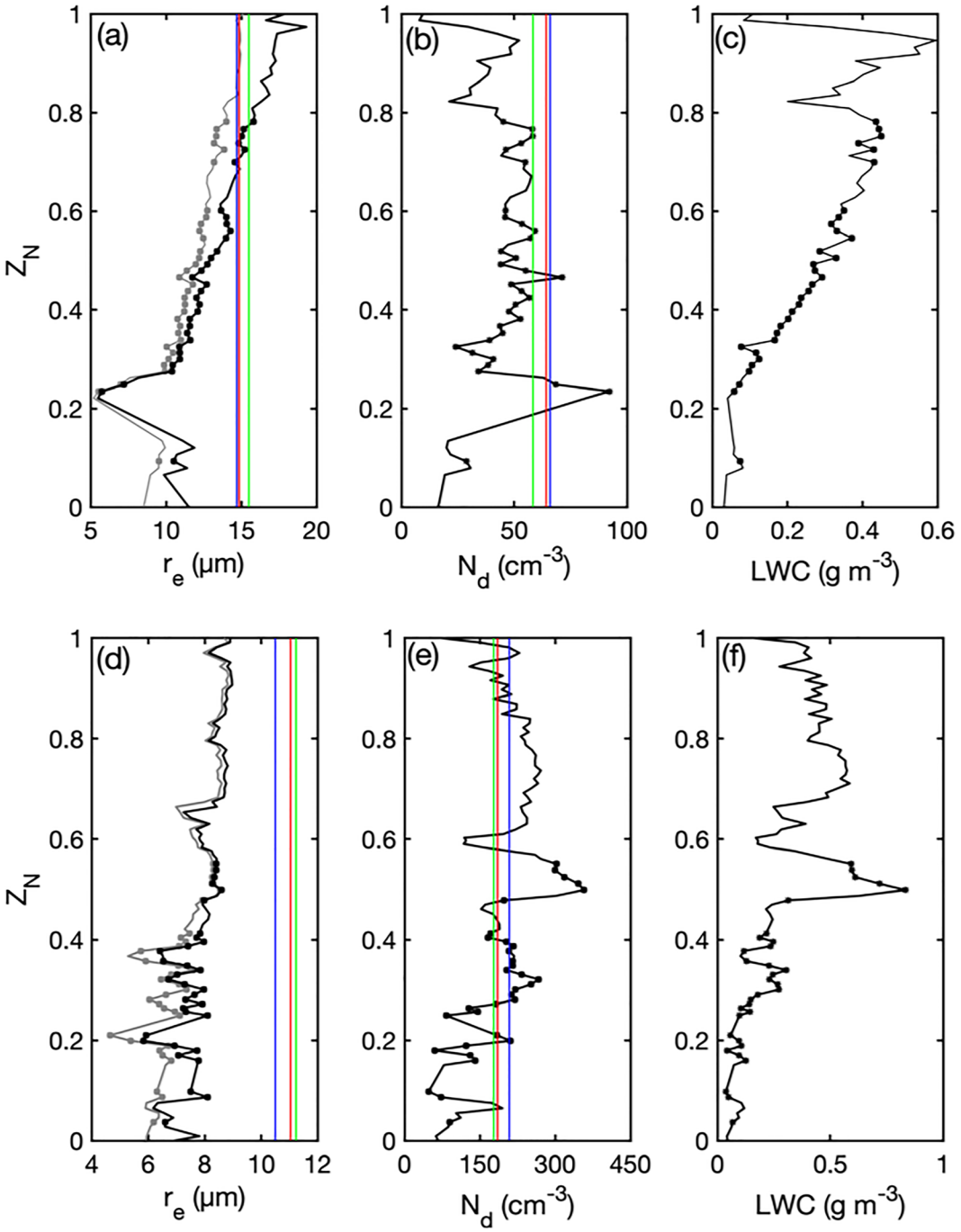

Figure 9. Vertical profiles of cloud droplet r_e, N_d, and LWC from the aircraft observations for two cases, along with the colocated MODIS retrieval results. Figures 9a–9c are for rf07 case (20180131 4:44:20–4:45:47 UTC) with relatively large cloud droplets in mixed-phase cloud, and Figures 9d–9f are for rf09 case (20180205 4:57:00–4:59:10 UTC) with small cloud droplets in pure liquid cloud. The green, red, and blue lines represent MODIS cloud properties of 1.6-, 2.1-, and 3.7-μm channels respectively. Gray and black lines in (a, d) represent the cloud droplet r_e obtained from CDP, and CDP + 2DS, respectively. In (b, e) and (c, f), the black lines represent the cloud N_d from aircraft observations, respectively. The dot symbol in each panel indicates the appearance of drizzle droplets.

$$H_\sigma = \frac{\text{stdev}\,(R_{0.86})}{\text{mean}\,(R_{0.86})} \tag{7}$$

where stdev($R_{0.86}$) and mean($R_{0.86}$) denote the standard deviation and mean of the reflectance products, respectively. MODIS 0.86-μm channel reflectance product which is aggregated to 1 km × 1 km

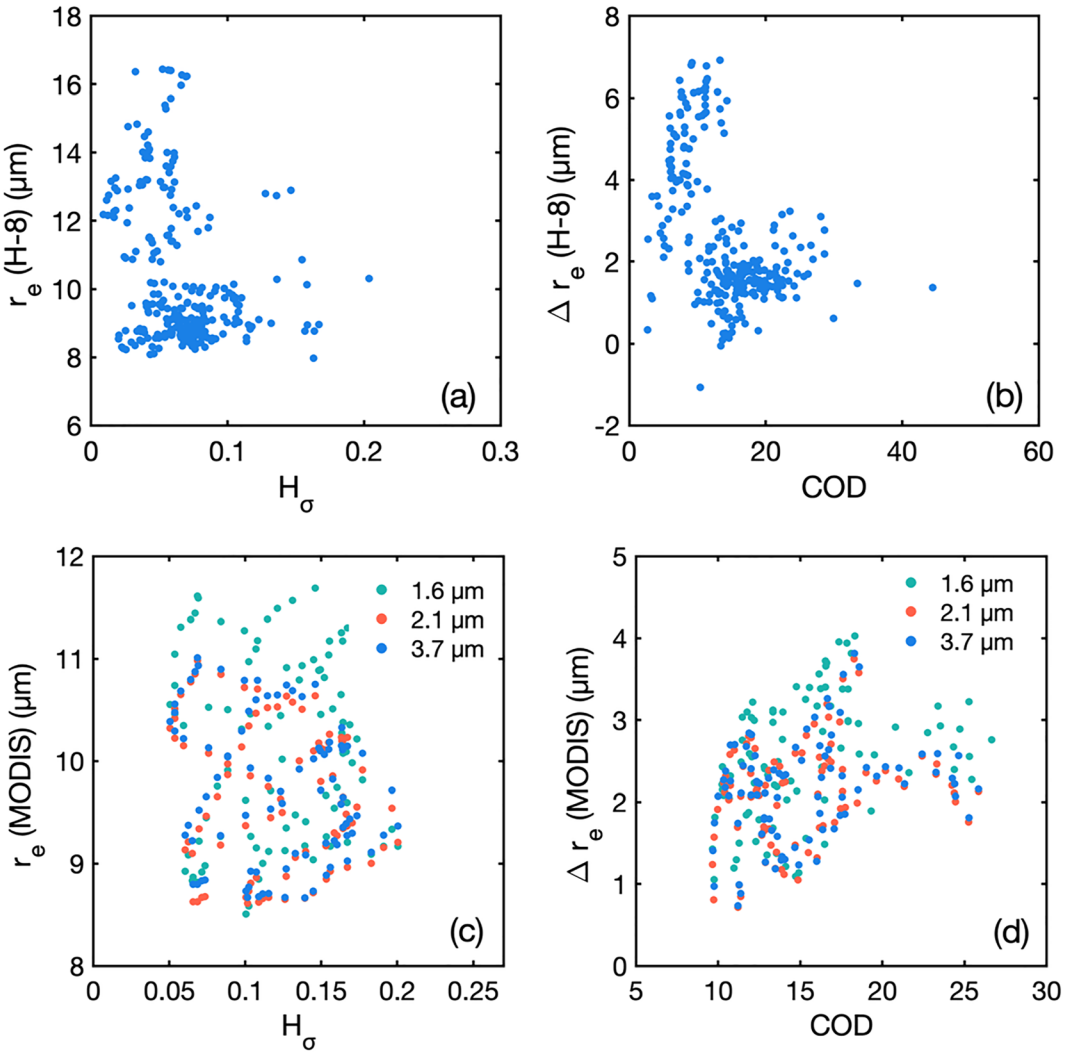

Figure 10. The scatter plots between the H_σ and cloud droplet r_e from SatCORPS H-8 (a) or from MODIS (c) and between cloud optical depth (COD) and the SatCORPS H-8 versus aircraft r_e difference (b) or the MODIS versus aircraft r_e difference (d). Three channels of MODIS are represented using three different colors.

(EV_250_Aggr1km_RefSB) is used to calculate H_σ. For SatCORPS H-8 cloud product, the H_σ is derived from 0.8-μm channel reflectance product (reflec08), since it does not include 0.86-μm channel reflectance. $H_\sigma < 0.3$ represents relatively homogeneous clouds, whereas $H_\sigma > 0.3$ could be used to define heterogeneous clouds (Liang et al., 2009). Figure 10 shows changes of SatCORPS H-8 and MODIS cloud droplet r_e with H_σ, along with the changes of satellite (SatCORPS H-8 and MODIS) versus aircraft difference of cloud droplet r_e with COD. Figure 10a shows that the SatCORPS H-8 H_σ is generally smaller than 0.2 with most values below 0.1, implying that the clouds are more homogeneous. However, there is relatively high discrepancy between the in situ r_e and SatCORPS H-8 r_e, as illustrated in Figure 10b. Figure 10b also shows that the difference of r_e between SatCORPS H-8 and aircraft could be even larger when COD is relatively low (COD < 10). This finding is similar as that found by Zhang and Platnick (2011). Figures 10c and 10d show similar findings for MODIS retrieval results, while the exact values are slightly different. The MODIS H_σ is also smaller than 0.2 for most values, and for COD between 10 and 30, the MODIS versus aircraft difference of cloud droplet r_e is mainly between 0.7 and 4 μm.

4.2 Droplet Size Spectra

Standard deviation of lognormal droplet size distribution (σ) is used in this study to analyze the cloud droplet spectra effect on satellite retrievals. Noted that only the cloud droplet lognormal distribution measured from

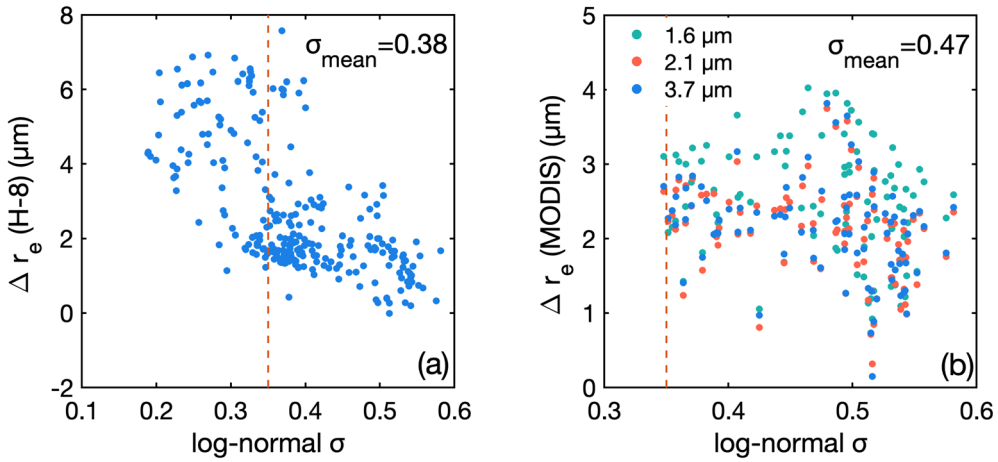

Figure 11. The change of positive biases in cloud droplet r_e from (a) SatCORPS H-8 cloud product and (b) MODIS with log-normal standard deviations (σ) compared with in situ aircraft observations. The orange dashed line represents $\sigma = 0.35$ which is used in the calculation of lookup tables for both SatCORPS H-8 and MODIS retrievals.

CDP is discussed here without including the drizzle mode. The MODIS and SatCORPS H-8 retrieval algorithm generally assumed σ as 0.35 in their calculation of lookup table. As Painemal and Zuidema (2011) indicated, retrieved r_e would be increasingly overestimated if true σ is smaller than the assumed value. Consistently with this, Figure 11 shows generally decreasing biases in the retrieved cloud droplet r_e with σ for both MODIS and SatCORPS H-8 cloud product compared with in situ aircraft observations. However, there are a few cases with even larger positive biases in cloud r_e when σ is larger than 0.35, which could be caused by other influential factors that are beyond the understanding of current study. Figure 11a shows that the σ is mainly between 0.1 and 0.6 with a mean value of 0.38 for aircraft observations collocated with SatCORPS H-8 observations, and Figure 11b shows that the σ is mainly between 0.3 and 0.6 with a mean value of 0.47 for aircraft observed cases collocated with MODIS. Associated with the existence of much smaller σ values, the positive bias in SatCORPS H-8 r_e could reach as large as 6–7 μm with $\sigma < 0.35$ compared with in situ aircraft observations, as shown in Figure 11a. In contrast, the positive biases in MODIS r_e are generally less than 4 μm.

4.3 Drizzle

The existence of drizzle within clouds could influence the evaluation of cloud droplet r_e retrieval. Nakajima et al. (2010b) found positive r_e bias of nearly 2.5 μm for clouds with drizzles based on aircraft observations. For the clouds investigated in this study, the maximum drizzle occurrence frequency is 39%. As shown in Figure 12, the biases in retrieved cloud r_e for H-8 drizzle, H-8 nondrizzle, MODIS drizzle, and MODIS nondrizzle cases are 2.95, 2.55, 2.18, and 2.79 μm compared to in situ aircraft observations, respectively. The biases in MODIS retrieved cloud r_e for cases with drizzle are slightly smaller than that without drizzle, while the biases in SatCORPS H-8 retrieved cloud r_e for cases with drizzle are slightly larger than that without drizzle. Roughly, the existence of drizzle has weak influence on the performance of both MODIS and SatCORPS H-8 retrievals, both of which are based on spectral radiation measurements. In other words, for clouds studied here, the existence of drizzle seems to have weak influence on the spectral radiation measured by the satellites.

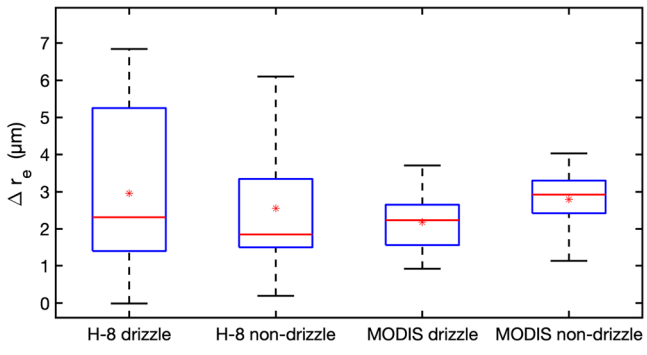

Figure 12. Boxplot of the difference in cloud droplet r_e between MODIS/SatCORPS H-8 retrievals and in situ aircraft observations for clouds with drizzle and without drizzle. The asterisk symbols are the mean bias, and $\triangle r_e$ is the difference of cloud droplet r_e between retrieval and in situ measurements.

5. Conclusions

Using the in situ measurements, this study evaluated the retrieval performance of Himawari-8 and MODIS cloud products over the Southern Ocean. Seven flights during the SOCRATES are used to match with satellite observations and then evaluate the performance of cloud

microphysical properties of effective radius (r_e) and droplet number concentration (N_d) from satellite retrievals.

SatCORPS H-8 cloud r_e and N_d are overestimated by about 33% and 3%, respectively, near the cloud top compared with in situ measurements over the Southern Ocean when aircraft flew horizontally. The performance of satellite retrievals varies a little with cloud phase. When we compare the SatCORPS H-8 retrievals with the vertically averaged in situ aircraft observations, SatCORPS H-8 r_e and N_d could be overestimated about 18% and 39%, respectively. However, the retrieval biases in r_e would decrease a lot if comparing the cloud properties for the top 20% of clouds, which is only 3% overestimation, indicating that the satellite retrievals are heavily weighted to cloud tops.

MODIS retrievals systematically overestimated cloud droplet r_e by 26–31% compared to the in situ observations over the Southern Ocean. The performance of MODIS retrievals has a strong sensitivity to the wavelength channel that the retrieval method used. In general, the positive retrieval bias is smaller for $r_{e2.1}$ and $r_{e3.7}$ compared to $r_{e1.6}$. In contrast, the MODIS retrieved N_d has a good consistency with the in situ aircraft observations regardless of the channel. Due to the different absorption strength of cloud liquid droplets, the photon penetration depth in clouds vary with wavelength channels used. In other words, the cloud retrievals based on different channels have different representation of cloud layer depth. Considering the vertical distribution of cloud properties, the performance of cloud retrievals could be very different when retrieved using different wavelength channels and evaluated by the in situ aircraft observations at different heights. In this study, the vertical distribution of cloud properties along with the entrainment effect could be the reason that makes $r_{e3.7}$ and $r_{e2.1}$ smaller than $r_{e1.6}$.

This study has suggested that both SatCORPS H-8 cloud product and MODIS cloud product overestimate the cloud droplet r_e and N_d compared to the in situ observations. The influential factors to the evaluation of satellite retrievals have been discussed, including the cloud heterogeneity, the droplet size spectra, and the existence of drizzle. Particularly, it shows that the standard deviation of cloud droplet size spectra could lead large biases to cloud droplet r_e which could be as large as 7 μm. To improve the quality of satellite retrievals, we should improve the identification of cloud phases and seriously consider the impacts of cloud inhomogeneity and droplet size spectra over different regions in future.

Acknowledgments
This work was supported by the National Natural Science Foundation of China (NSFC) (41925022, 91837204, and 41575143), the Ministry of Science and Technology of China National Key R&D Program on Monitoring, Early Warning and Prevention of Major Natural Disasters (2017YFC1501403), the State Key Laboratory of Earth Surface Processes and Resource Ecology (2017-ZY-02), and the Fundamental Research Funds for the Central Universities (2017EYT18). We appreciate the efforts of the entire SOCRATES team in collecting a high-quality data set. The data that support the findings of this study are available online (https://data.eol.ucar.edu/master_list/?project=SOCRATES). We also highly appreciate the SatCORPS H-8 cloud product provided by the NASA SatCORPS group through website (https://data.eol.ucar.edu/dataset/552.027).

References

Abdelmonem, A., Jarvinen, E., Duft, D., Hirst, E., Vogt, S., Leisner, T., & Schnaiter, M. (2016). PHIPS-HALO: The airborne particle habit imaging and polar scattering probe—Part 1: Design and operation. *Atmospheric Measurement Techniques*, *9*(7), 3131–3144. https://doi.org/10.5194/amt-9-3131-2016

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*, 11,120–11,140. https://doi.org/10.1029/2018JD028535

Albrecht, B. A. (1989). Aerosols, cloud microphysics, and fractional cloudiness. *Science*, *245*(4923), 1227–1230. https://doi.org/10.1126/science.245.4923.1227

Albrecht, B. A., Fairall, C. W., Thomson, D. W., White, A. B., Snider, J. B., & Shubert, W. H. (1990). Surface-based remote sensing of the observed and the adiabatic liquid water content of stratocumulus clouds. *Geophysical Research Letters*, *17*, 89–92.

Bao, S., Letu, H., Zhao, C., Tana, G., Shang, H., Wang, T., et al. (2018). Spatiotemporal distributions of cloud parameters and the temperature response over the Mongolian Plateau during 2006–2015 based on MODIS data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *99*, 1–10. https://doi.org/10.1109/JSTARS.2018.2857827

Baumgardner, D., & Korolev, A. (1997). Airspeed corrections for optical array probe sample volumes. *Journal of Atmospheric and Oceanic Technology*, *14*, 1224–1229. https://doi.org/10.1175/1520-0426(1997)014<1224:ACFOAP>2.0.CO;2

Baumgardner, D., & Rodi, A. (1989). Laboratory and wind-tunnel evaluations of the Rosemount icing detector. *Journal of Atmospheric and Oceanic Technology*, *6*(6), 971–979. https://doi.org/10.1175/1520-0426(1989)006<0971:lawteo>2.0.co;2

Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., et al. (2016). An introduction to Himawari-8/9-Japan's new-generation geostationary meteorological satellites. *Journal of the Meteorological Society of Japan*, *94*(2), 151–183. https://doi.org/10.2151/jmsj.2016-009

Biter, C. J., Dye, J. E., Huffman, D., & King, W. D. (1987). The drop-size response of the CEIRO liquid water probe. *Journal of Atmospheric and Oceanic Technology*, *4*, 359–367. https://doi.org/10.1175/1520-0426(1987)004<0359:TDSROT>2.0.CO;2

Cess, R. D. (1976). Stratospheric aerosols—Effect upon atmospheric-temperature and global climate. *Tellus*, *28*(1), 1–10.

Chang, F. L., & Li, Z. Q. (2002). Estimating the vertical variation of cloud droplet effective radius using multispectral near-infrared satellite measurements. *Journal of Geophysical Research*, *107*(D15), 4257. https://doi.org/10.1029/2001JD000766

Davis, S. M., Hallar, A. G., Avallone, L. M., & Engblom, W. (2007). Measurement of total water with a tunable diode laser hygrometer: Inlet analysis, calibration procedure, and ice water content determination. *Journal of Atmospheric and Oceanic Technology*, *24*(3), 463–475. https://doi.org/10.1175/jtech1975.1

Dorsi, S. W., Shupe, M. D., Persson, P. O. G., Kingsmill, D. E., & Avallone, L. M. (2015). Phase-specific characteristics of wintertime clouds across a midlatitude mountain range. *Monthly Weather Review*, *143*(10), 4181–4197. https://doi.org/10.1175/mwr-d-15-0135.1

Faber, S., French, J. R., & Jackson, R. (2018). Laboratory and in-flight evaluation of measurement uncertainties from a commercial Cloud Droplet Probe (CDP). *Atmospheric Measurement Techniques*, *11*(6), 3645–3659. https://doi.org/10.5194/amt-11-3645-2018

Field, P. R., Hogan, R. J., Brown, P. R. A., Illingworth, A. J., Choularton, T. W., Kaye, P. H., et al. (2004). Simultaneous radar and aircraft observations of mixed-phase cloud at the 100m scale. *Quarterly Journal of the Royal Meteorological Society*, *130*(600), 1877–1904. https://doi.org/10.1256/qj.03.102

Garrett, T. J., & Zhao, C. (2006). Increased Arctic cloud longwave emissivity associated with pollution from mid-latitudes. *Nature*, *440*(636), 787–789. https://doi.org/10.1038/nature04636

Garrett, T. J., & Zhao, C. (2013). Ground-based remote sensing of thin clouds in the Arctic. *Atmospheric Measurement Techniques*, *6*(5), 1227–1243. https://doi.org/10.5194/amt-6-1227-2013

Haney, C. O. (2013). Cloud drop effective radius for trade wind cumuli observed during RICO by aircraft and MODIS. M.S. Thesis, Univ. of Illinois at Urbana-Champaign, 89 pp.

Haynes, J. M., Jakob, C., Rossow, W. B., Tselioudis, G., & Brown, J. (2011). Major characteristics of Southern Ocean cloud regimes and their effects on the energy budget. *Journal of Climate*, *24*(19), 5061–5080. https://doi.org/10.1175/2011jcli4052.1

Hu, Y., Rodier, S., Xu, K. M., Sun, W., Huang, J., Lin, B., et al. (2010). Occurrence, liquid water content, and fraction of supercooled water clouds from combined CALIOP/IIR/MODIS measurements. *Journal of Geophysical Research*, *115*, D00H34. https://doi.org/10.1029/2009JD012384

Iwabuchi, H., Putri, N. S., Saito, M., Tokoro, Y., Sekiguchi, M., Yang, P., & Baum, B. A. (2018). Cloud property retrieval from multiband infrared measurements by Himawari-8. *Journal of the Meteorological Society of Japan*, *96B*, 27–42. https://doi.org/10.2151/jmsj.2018-001

Jiang, J. H., Su, H., Zhai, C., Perun, V. S., del Genio, A., Nazarenko, L. S., et al. (2012). Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA A-train satellite observations. *Journal of Geophysical Research*, *117*, D1410. https://doi.org/10.1029/2011JD017237

King, W. D., Parkin, D. A., & Handsworth, R. J. (1978). Hot-wire liquid water device having fully calculable response characteristics. *Journal of Applied Meteorology*, *17*(12), 1809–1813. https://doi.org/10.1175/1520-0450(1978)017<1809:ahwlwd>2.0.co;2

Klein, S. A., & Hartmann, D. L. (1993). The seasonal cycle of low stratiform clouds. *Journal of Climate*, *6*(8), 1587–1606. https://doi.org/10.1175/1520-0442(1993)006<1587:tscols>2.0.co;2

Korolev, A. V., Strapp, J. W., Isaac, G. A., & Nevzorov, A. N. (1998). The Nevzorov airborne hot-wire LWC-TWC probe: Principle of operation and performance characteristics. *Journal of Atmospheric and Oceanic Technology*, *15*(6), 1495–1510. https://doi.org/10.1175/1520-0426(1998)015<1495:tnahwl>2.0.co;2

Lance, S., Brock, C. A., Rogers, D., & Gordon, J. A. (2010). Water droplet calibration of the Cloud Droplet Probe (CDP) and in-flight performance in liquid, ice and mixed-phase clouds during ARCPAC. *Atmospheric Measurement Techniques*, *3*(6), 1683–1706. https://doi.org/10.5194/amt-3-1683-2010

Lawson, R. P. (2011). Effects of ice particles shattering on the 2D-S probe. *Atmospheric Measurement Techniques*, *4*(7), 1361–1381. https://doi.org/10.5194/amt-4-1361-2011

Lawson, R. P., Baker, B., Pilson, B., & Mo, Q. X. (2006). In situ observations of the microphysical properties of wave, cirrus, and anvil clouds. Part II: Cirrus clouds. *Journal of the Atmospheric Sciences*, *63*(12), 3186–3203. https://doi.org/10.1175/jas3803.1

Letu, H., Yang, K., Nakajima, T. Y., Ishimoto, H., Nagao, T. M., Riedi, J., et al. (2020). High-resolution retrieval of cloud microphysical properties and surface solar radiation using Himawari-8/AHI next-generation geostationary satellite. *Remote Sensing of Environment*, *239*, 111583. https://doi.org/10.1016/j.rse.2019.111583

Letu, H. S., Nagao, T. M., Nakajima, T. Y., Riedi, J., Ishimoto, H., Baran, A. J., et al. (2018). Ice cloud properties from Himawari-8/AHI next-generation geostationary satellite: Capability of the AHI to monitor the DC cloud generation process. *IEEE Transactions on Geoscience and Remote Sensing*, 1–11. https://doi.org/10.1109/TGRS.2018.2882803

Li, J., Jian, B., Huang, J., Hu, Y., Zhao, C., & Kawamotod, K. (2018). Long term variation of cloud droplet number concentrations from space-based Lidar. *Remote Sensing of Environment*, *213*, 144–161.

Li, J., Lv, Q., Zhang, M., Wang, T., Kawamoto, K., Chen, S., & Zhang, B. (2017). Effects of atmospheric dynamics and aerosols on the fraction of supercooled water clouds. *Atmospheric Chemistry and Physics*, *17*, 1847–1863.

Liang, L. S., Di Girolamo, L., & Platnick, S. (2009). View-angle consistency in reflectance, optical thickness and spherical albedo of marine water-clouds over the northeastern Pacific through MISR-MODIS fusion. *Geophysical Research Letters*, *36*, L09811. https://doi.org/10.1029/2008GL037124

Liang, L. S., Di Girolamo, L., & Sun, W. B. (2015). Bias in MODIS cloud drop effective radius for oceanic water clouds as deduced from optical thickness variability across scattering angles. *Journal of Geophysical Research-Atmospheres*, *120*, 7661–7681. https://doi.org/10.1002/2015JD023256

Lohmann, U., & Feichter, J. (2005). Global indirect aerosol effects: A review. *Atmospheric Chemistry and Physics*, *5*, 715–737. https://doi.org/10.5194/acp-5-715-2005

Mace, G. G. (2010). Cloud properties and radiative forcing over the maritime storm tracks of the Southern Ocean and North Atlantic derived from A-Train. *Journal of Geophysical Research*, *115*, D10201. https://doi.org/10.1029/2009JD012517

Mace, G. G., Zhang, Y., Platnick, S., King, M. D., Minnis, P., & Yang, P. (2004). Evaluation of cirrus cloud properties derived from MODIS data using cloud properties derived from ground-based observations collected at the ARM SGP site. *Journal of Applied Meteorology*, *44*, 221–240.

McFarquhar, G. M., Zhang, G., Poellot, M. R., Kok, G. L., McCoy, R., Tooman, T., et al. (2007). Ice properties of single-layer stratocumulus during the mixed-phase Arctic cloud experiment: 1. Observations. *Journal of Geophysical Research*, *112*, D24201. https://doi.org/10.1029/2007JD008633

Miles, N. L., Verlinde, J., & Clothiaux, E. E. (2000). Cloud droplet size distributions in low-level stratiform clouds. *Journal of the Atmospheric Sciences*, *57*(2), 295–311. https://doi.org/10.1175/1520-0469(2000)057<0295:cdsdil>2.0.co;2

Morrison, A. E., Siems, S. T., & Manton, M. J. (2011). A three-year climatology of cloud-top phase over the Southern Ocean and North Pacific. *Journal of Climate*, *24*, 2405–2418. https://doi.org/10.1175/2010JCLI3842.1

Nakajima, T., King, M. D., Spinhirne, J. D., & Radke, L. F. (1991). Determination of the optical-thickness and effective particle radius of clouds from reflected solar-radiation measurements. 2. Marine stratocumulus observations. *Journal of the Atmospheric Sciences*, *48*(5), 728–750. https://doi.org/10.1175/1520-0469(1991)048<0728:dotota>2.0.co;2

Nakajima, T. Y., & Nakajima, T. (1995). Wide-area determination of cloud microphysical properties from NOAA AVHRR measurements for FIRE and ASTEX regions. *Journal of the Atmospheric Sciences*, *52*(23), 4043–4059. https://doi.org/10.1175/1520-0469(1995)052<4043:wadocm>2.0.co;2

Nakajima, T. Y., Suzuki, K., & 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. *Journal of the Atmospheric Sciences*, *67*(6), 1884–1896. https://doi.org/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 multi-sensor view. *Journal of the Atmospheric Sciences*, *67*(6), 1897–1907. https://doi.org/10.1175/2010jas3276.1

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. *Journal of Geophysical Research*, *116*, D24206. https://doi.org/10.1029/2011jd016155

Platnick, S., King, M. D., Ackerman, S. A., Menzel, W. P., Baum, B. A., Riedi, J. C., & Frey, R. A. (2003). The MODIS cloud products: Algorithms and examples from Terra. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(2), 459–473. https://doi.org/10.1109/tgrs.2002.808301

Protat, A., Bouniol, D., Delanoe, J., May, P. T., Plana-Fattori, A., Hasson, A., et al. (2009). Assessment of cloudsat reflectivity measurements and ice cloud properties using ground-based and airborne cloud radar observations. *Journal of Atmospheric and Oceanic Technology*, *26*, 1717–1741.

Qiu, Y. M., Zhao, C. F., Guo, J. P., & Li, J. M. (2017). 8-year ground-based observational analysis about the seasonal variation of the aerosol-cloud droplet effective radius relationship at SGP site. *Atmospheric Environment*, *164*, 139–146. https://doi.org/10.1016/j.atmosenv.2017.06.002

Ramanathan, V., Cess, R. D., Harrison, E. F., Minnis, P., Barkstrom, B. R., Ahmad, E., & Hartmann, D. (1989). Cloud-radiative forcing and climate-results from the earth radiation budget experiment. *Science*, *243*(4887), 57–63. https://doi.org/10.1126/science.243.4887.57

Rossow, W. B., & Schiffer, R. A. (1991). ISCCP cloud data products. *Bulletin of the American Meteorological Society*, *72*(1), 2–20.

Schnaiter, M., Järvinen, E., Abdelmonem, A., & Leisner, T. (2018). PHIPS-HALO: The airborne particle habit imaging and polar scattering probe—Part 2: Characterization and first results. *Atmospheric Measurement Techniques*, *11*, 341–357. https://doi.org/10.5194/amt-11-341-2018

Schneider, S. H. (1972). Cloudiness as a global climatic feedback mechanism—Effects on radiation balance and surface-temperature of variations in cloudiness. *Journal of the Atmospheric Sciences*, *29*(8), 1413. https://doi.org/10.1175/1520-0469(1972)029<1413:caagcf>2.0.co;2

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., et al., & The CloudSat Science Team (2002). The CloudSat mission and the A-train. *Bulletin of Atmospheric Meteorological Society*, *83*, 1771–1790.

Trenberth, K. E., & Fasullo, J. T. (2010). Simulation of present-day and twenty-first-century energy budgets of the Southern Oceans. *Journal of Climate*, *23*(2), 440–454. https://doi.org/10.1175/2009jcli3152.1

Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., & Powell, K. A. (2009). Overview of the CALIPSO mission and CALIOP data processing algorithms. *Journal of Atmospheric and Oceanic Technology*, *26*, 2310–2323.

Yang, Y., Zhao, C. F., Dong, X. B., Fan, G. C., Zhou, Y. Q., Wang, Y., et al. (2019). Toward understanding the process-level impacts of aerosols on microphysical properties of shallow cumulus cloud using aircraft observations. *Atmospheric Research*, *221*, 27–33. https://doi.org/10.1016/j.atmosres.2019.01.027

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

Zhao, C. F., Chen, Y., Li, J., Letu, H., Su, Y., Chen, T., & Wu, X. (2019). 15-year statistical analysis of cloud characteristics over China using Terra and Aqua MODIS observations. *International Journal of Climatology*, *38*(5), 2612–2629.

Zhao, C. F., & Garrett, T. J. (2015). Effects of Arctic haze on surface cloud radiative forcing. *Geophysical Research Letters*, *42*, 557–564. https://doi.org/10.1002/2014GL062015

Zhao, C., Qiu, Y., Dong, X., Wang, Z., Peng, Y., Li, B., et al. (2018). Negative Aerosol-Cloud r e Relationship From Aircraft Observations Over Hebei, *China. Earth and Space Science*, *5*(1), 19–29. https://doi.org/10.1002/2017ea000346

Zhao, C. F., Xie, S. C., Chen, X., Jensen, M. P., & Dunn, M. (2014). Quantifying uncertainties of cloud microphysical property retrievals with a perturbation method. *Journal of Geophysical Research – Atmospheres*, *119*, 5375–5385. https://doi.org/10.1002/2013JD021112

Zhao, C. F., Xie, S., Klein, S. A., Protat, A., Shupe, M. D., McFarlane, S. A., et al. (2012). Toward understanding of differences in current cloud retrievals of ARM ground-based measurements. *Journal of Geophysical Research*, *117*, D10206. https://doi.org/10.1029/2011JD016792

Zhao, C. F., Zhao, L. J., & Dong, X. B. (2019). A case study of stratus cloud properties using in situ aircraft observations over Huanghua, China. *Atmosphere*, *10*(1), 11. https://doi.org/10.3390/atmos10010019