Earth and Space Science

RESEARCH ARTICLE
10.1029/2023EA003355

# Validation of Satellite-Based Cloud Phase Distributions Using Global-Scale In Situ Airborne Observations

**Dao Wang**[1,2], **Ching An Yang**[1], **and Minghui Diao**[1]

[1]Department of Meteorology and Climate Science, San Jose State University, San Jose, CA, USA, [2]Now at Department of Atmospheric Sciences, University of Utah, Salt Lake City, UT, USA

**Key Points:**

- Ice phase frequencies vary largely with altitudes, latitudes, and seasons but not much with longitudes or small spatiotemporal mismatches
- CALIPSO shows most similar ice phase frequency to in situ observations, CloudSat and DARDAR overestimate mixed and ice phase, respectively
- Liquid water transitions to ice gradually with increasing altitudes over Southern Ocean but more rapidly in tropics and NH extratropics

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

M. Diao,
minghui.diao@sjsu.edu

**Citation:**

Wang, D., Yang, C. A., & Diao, M. (2024). Validation of satellite-based cloud phase distributions using global-scale in situ airborne observations. *Earth and Space Science*, *11*, e2023EA003355. https://doi.org/10.1029/2023EA003355

Received 11 OCT 2023
Accepted 30 MAR 2024

**Author Contributions:**

**Conceptualization:** Minghui Diao
**Data curation:** Dao Wang, Ching An Yang, Minghui Diao
**Formal analysis:** Dao Wang, Minghui Diao
**Funding acquisition:** Minghui Diao
**Investigation:** Dao Wang, Minghui Diao
**Methodology:** Dao Wang, Minghui Diao
**Project administration:** Minghui Diao
**Resources:** Minghui Diao
**Software:** Dao Wang
**Supervision:** Minghui Diao

## Abstract

Understanding distributions of cloud thermodynamic phases is important for accurately representing cloud radiative effects and cloud feedback in a changing climate. Satellite-based cloud phase data have been frequently used to compare with climate models, yet few studies validated them against in situ observations at a near-global scale. This study aims to validate three satellite-based cloud phase products using a compositive in situ airborne data set developed from 11 flight campaigns. Latitudinal-altitudinal cross sections of cloud phase occurrence frequencies are examined. The Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) show the most similar vertical profiles of ice phase frequencies compared with in situ observations. The CloudSat data overestimate mixed-phase frequencies up to 15 km but provide better sampling through cloud layers than lidar data. The DARDAR (raDAR/liDAR) data show a sharp transition between ice and liquid phase and overestimate ice phase frequency at most altitudes and latitudes. The satellite data are further evaluated for various latitudes, longitudes, and seasons, which show higher ice phase frequency in the extratropics in their respective wintertime and smaller impacts from longitudinal variations. The Southern Ocean shows a thicker mixing region where liquid and ice phases have similar frequencies compared with tropics and Northern Hemisphere (NH) extratropics. Two comparison methods with different spatiotemporal windows show similar results, which demonstrates the statistical robustness of these comparisons. Overall, this study develops a near global-scale in situ observational data set to assess the accuracy of satellite-based cloud phase products and investigates the key factors affecting the distributions of cloud phases.

## Plain Language Summary

Accurate representations of cloud thermodynamic phase (i.e., ice, liquid, and mixed phase) play an important role in climate prediction. Even though satellite observations have been used to improve climate model simulations of cloud phase, few studies have validated satellite-based cloud phase distributions on a global scale. This work develops a large data set based on in situ aircraft-based observations from 11 flight campaigns in various regions. Three satellite-based cloud phase products are evaluated. Satellite observations that are either in proximity to the aircraft samples or in a similar domain are used for comparisons. CALISPO data show the best comparison results for representing the fraction of ice clouds among all types of cloud phases. CloudSat overestimates mixed phase frequency and DARDAR overestimates ice phase frequency, but they penetrate through cloud layers better when radar and lidar data are combined. The impacts of seasonal variability, spatial variability among various latitudes and longitudes, as well as temporal variability from a few hours to different seasons are examined. The results of this work help to identify the key factors affecting cloud phase distributions from a near-global perspective. The methodology developed can also guide future validations of satellite data using aircraft-based observations.

## 1. Introduction

Three cloud thermodynamic phases—liquid, mixed, and ice phase, have different radiative effects on the Earth's incoming and outgoing radiation. A previous study using satellite observations showed that the liquid, mixed and ice phases have negative, negative, and positive net cloud radiative effects on average at the top of the atmosphere, respectively (Matus & L'Ecuyer, 2017). The net cloud radiative effects are affected by the vertical level, microphysical and macrophysical properties of different types of clouds (Liou, 1992). For example, liquid clouds are generally located in the lower troposphere, which reflect more incoming shortwave radiation and have relatively longer lifetime. On the other hand, ice clouds tend to be located at higher altitudes and more transparent for shortwave radiation.

**Validation:** Dao Wang, Ching An Yang, Minghui Diao
**Visualization:** Dao Wang, Ching An Yang, Minghui Diao
**Writing – original draft:** Dao Wang, Minghui Diao
**Writing – review & editing:** Dao Wang, Minghui Diao

Simulations of Earth's climate are largely affected by the representation of cloud phases in global climate models (GCMs). Previously, the GCMs in the Coupled Model Intercomparison Projects 3 (CMIP3) and 5 (CMIP5) were found to have large radiative biases over the Southern Ocean (e.g., Bodas-Salcedo et al., 2012, 2014, 2016; Frey & Kay, 2018; Kay et al., 2016; McCoy et al., 2014a, 2014b) mainly due to the misrepresentation of liquid phase below 0°C (i.e., supercooled liquid water) in the model parameterizations. These studies show radiative biases in the GCMs can be up to 10–20 W $m^{-2}$ in the southern high latitudes due to underestimation of reflected shortwave radiation by the supercooled liquid water. In addition, other studies found that cloud feedback in a changing climate is also highly sensitive to the current state of cloud phase (e.g., Cesana & Storelvmo, 2017; Flynn & Mauritsen, 2020; Tan et al., 2016; Terai et al., 2016; Zelinka et al., 2020). This is because that the currently existing ice crystals could turn into liquid droplets as temperature increases and such melting process would absorb heat from the surrounding air, acting as a buffer to global warming. This process is also known as the cloud phase feedback in the climate system (Mitchell et al., 1989).

Satellite data have been frequently used to compare with GCM simulations and evaluate simulated global distributions of cloud phases and the associated radiative effects. NASA CALIPSO (Winker et al., 2003, 2010) and CloudSat (G. L. Stephens et al., 2002) are two of the satellites from the original Afternoon Constellation (A-Train), equipped with lidar and radar active sensors, respectively. These two sensors have the advantages of providing near global, daily coverage cloud properties for over a decade. Three cloud phase products derived from CALIPSO and CloudSat are frequently used for analyzing cloud phase distributions on a global scale, including lidar-only CALIPSO Vertical Feature Mask (VFM) (Hu et al., 2009; Vaughan et al., 2019), radar-lidar combined CloudSat (Wang, 2019) and DARDAR (Delanoë & Hogan, 2008a, 2008b) products. Several previous studies used different types of products derived from CALIPSO and CloudSat to validate GCMs as well as improving cloud microphysics parameterizations. For example, Bodas-Salcedo et al. (2014) evaluated climate models from the Cloud Feedback Model Intercomparison Project Phase 2 (CFMIP2), using the GCM-Oriented CALIPSO Cloud Product (GOCCP) data. They found underestimations of reflected shortwave radiation over the Southern Ocean especially associated with mid-level and low-level clouds. The study by Kay et al. (2016) used a community satellite simulator package to evaluate simulations of the National Center for Atmospheric Research (NCAR) Community Earth System Model (CESM). They found an insufficient amount of supercooled liquid water in the atmospheric component of CESM—Community Atmosphere Model version 5 (CAM5). Tan and Storelvmo (2016) used supercooled cloud fraction derived from the CALIPSO data to evaluate the parameters used to represent the Wegner-Bergeron-Findeisen process in the NCAR CAM5 model.

Data qualities of these satellite products are therefore important for increasing the fidelity of climate predictions. They are not only dictated by the nature of the sensors but also by the algorithms used to process the retrievals. Various types of airborne in situ or remote sensing observations have been used to compare with satellite-based cloud phase products. Cesana et al. (2016) compared several CALIPSO lidar-only cloud phase products with measurements by a Polar Nephelometer from two flight campaigns and identified three main reasons contributing to differences in cloud phase identification, including treatment of fully attenuated pixels, vertical resolution, and horizontal averaging. Ahn et al. (2018) compared in situ airborne observations with cloud phase optical property products derived from CALIPSO observations and found that satellite data underestimate the occurrence of mixed-phase clouds over the Southern Ocean. In situ observations of cloud properties collected by the Canadian Convair-580 aircraft were compared with the CloudSat data, which shows underestimated liquid water content (LWC) and overestimated liquid droplet effective radii by satellite data (Barker et al., 2008). Ceccaldi et al. (2013) compared DARDAR cloud phase classification against the airborne remote-sensing observations from the French airborne Radar-Lidar (RALI) platform, which carries a similar lidar (532 nm) and cloud radar (95 GHz) to the CALIPSO and CloudSat instrument configuration, respectively. They demonstrated an increase of supercooled liquid water fraction in a latitudinal cross section view in DARDAR version 2 compared with version 1 algorithm. While these previous studies provided valuable assessment of the potential biases in satellite-based cloud phase products, several main limitations still exist. For example, each previous study often focused on one satellite-based product instead of all three of them (i.e., CALIPSO, CloudSat, and DARDAR). In addition, each study often selected one spatiotemporal window for collocated comparisons, while lacking a quantification of the impacts of spatiotemporal misalignment on the comparisons results as documented in a similar satellite validation study by Diao et al. (2013). Lastly, the airborne observations used for satellite validation are generally limited to one specific region with smaller geographical coverage, while a global-scale comparison of cloud phase distribution is yet to be investigated.

**Table 1**
*Summary of Eleven NSF Flight Campaigns, Including Their Name, the Number of Research Flights, Time, Location, and Flight Hours*

| Campaign name and acronym | No. of flights | Time | Longitude and latitude | Flight hours | Clear-sky hours | In-cloud hours |
|---|---|---|---|---|---|---|
| HIAPER Pole-to-pole Observations - deployments 2–5 (HIPPO 2–5) | 46 | October–November 2009; March–April 2010; June–July 2011; August–September 2011 | 87°N–67°S, 128°E–105°W | 333 | 267 | 65 |
| Stratosphere-Troposphere Analyses of Regional Transport (START08) | 18 | April–June 2008 | 26°N–62°N, 86°W–118°W | 121 | 106 | 15 |
| Cloud Systems Evolution in the Trades (CSET) | 16 | July–August 2015 | 20°N–43°N, 118°W–157°W | 114 | 77 | 37 |
| Wintertime Investigation of Transport, Emissions, and Reactivity (WINTER) | 13 | February–March 2015 | 29°N–42°N, 67°W–86°W | 95 | 62 | 33 |
| Deep Convective Clouds and Chemistry Project (DC3) | 22 | May–June 2012 | 26°N–42°N, 79°W–107°W | 136 | 75 | 60 |
| CONvective TRansport of Active Species in the Tropics (CONTRAST) | 17 | January–February 2014 | 41°N–21°S, 133°E–105°W | 128 | 92 | 36 |
| PRE-Depression Investigation of Cloud-Systems in the Tropics (PREDICT) | 26 | August–September 2010 | 10°N–28°N, 37°W–87°W | 175 | 116 | 59 |
| Organization of Tropical East Pacific Convection (OTREC) | 22 | August–September 2019 | 3°N–13°N, 77°W–95°W | 123 | 86 | 34 |
| Tropical Ocean tRoposphere Exchange of Reactive halogen species and Oxygenated VOC (TORERO) | 17 | January–February 2012 | 13°N–42°S, 70°W–106°W | 134 | 105 | 28 |
| The $O_2/N_2$ Ratio and $CO_2$ Airborne Southern Ocean Study (ORCAS) | 18 | January–March 2016 | 19°S–75°S, 50°W–92°W | 95 | 75 | 20 |
| Southern Ocean Clouds, Radiation, Aerosol Transport Experimental Study (SOCRATES) | 15 | January–February 2018 | 42°S–62°S, 133°E–164°E | 112 | 77 | 33 |
| All campaigns | 230 | 2008–2019 | 87°N–67°S, 128°E–37°W | 1,566 | 1,138 | 420 |

*Note*. The campaigns are ordered by their northernmost latitude.

To address the three limitations, this study develops a near global-scale in situ observation data set and validates three frequently used satellite cloud phase products (i.e., CALIPSO, CloudSat, and DARDAR). A compositive in situ observation data set is developed using 11 US National Science Foundation (NSF) flight campaigns with a spatial coverage nearly from Pole to Pole. Several key factors affecting the cloud phase distributions and the satellite comparison results are investigated, including various latitudinal and longitudinal locations, seasonal variability, and various spatiotemporal windows used for the validation. This work is designed as follows. The detailed descriptions of in situ and satellite cloud phase data sets are given in Section 2. In Section 3, two different sampling methods to select satellite data for comparisons with in situ aircraft observations are described. The results are shown in Section 4, where we visualize, quantify, and investigate the individual factors affecting the occurrences of clouds and cloud phases. Lastly, Section 5 summarizes the key conclusions and provides implications for future satellite retrieval development.

## 2. Data and Methods

Several observational platforms and data sets are used in this study, including in situ observations from 11 flight campaigns funded by the US NSF and three satellite-based cloud phase products. Detailed information of in situ observations and remotely sensed observations are described in Tables 1 and 2, respectively.

### 2.1. In Situ Airborne Observations

A total of 11 NSF campaigns are used in this study, including the Stratosphere-Troposphere Analyses of Regional Transport (START08) (Pan et al., 2010), the HIAPER Pole-to-Pole Observations (HIPPO) (Wofsy, 2011), the PRE-Depression Investigation of Cloud-Systems in the Tropics (PREDICT) campaign (Montgomery

**Table 2**
*Descriptions of Three Satellite Products and Their Cloud Phase Variables*

| Satellite and instrument | Data availability | Data set | Variable | Phase categorization | Resolution |
|---|---|---|---|---|---|
| CloudSat—CPR (Cloud Profiling Radar) | 2006/06/12 d.163 ~ current<br>(Missing: 2017/12/06 ~ 2018/08/10) | 2B-CLDCLASS-LIDAR | CloudPhase | 1 = Ice Cloud<br>2 = Mix Cloud<br>3 = Liquid Cloud<br>−9 = Missing Value | Vertical: 240 m<br>Horizontal: 1.4 km × 1.8 km |
| CALIPSO—CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) | (V4.20) 2006/06/12 ~ 2020/06/30<br>(V4.21) 2020/07/01 ~ current | LID_L2_VFM-Standard-V4-20 | Feature_Classification_Flags | 0 = Unknown/Not Determined<br>1 = Ice<br>2 = Water<br>3 = Horizontal Oriented Ice | Vertical/Horizontal resolution:<br>30 m/333 m (−0.5–8.2 km)<br>60 m/1000 m (8.2–20.2 km)<br>180 m/1,667 m (20.3–30.1 km) |
| DARDAR (raDAR/liDAR) | (v2.1.1) 2006/06/15 ~ 2017/10/25 | DARDAR-CLOUD_v2.1.1 | DARMASK_Simplified_Categorization | −9 = ground<br>−1 = do not know<br>0 = clear<br>1 = ice<br>2 = ice + supercooled<br>3 = liquid warm<br>4 = supercooled<br>5 = rain<br>6 = aerosol<br>7 = maybe insects<br>8 = stratospheric feature | Vertical: 60 m<br>Horizontal: 1.1 km |

et al., 2012), the Deep Convective Clouds and Chemistry Project (DC3) (Barth et al., 2015), the Tropical Ocean tRoposphere Exchange of Reactive halogen species and Oxygenated VOC (TORERO) campaign (Volkamer et al., 2015), CONvective TRansport of Active Species in the Tropics (CONTRAST) campaign (Pan et al., 2017), the Wintertime Investigation of Transport, Emissions, and Reactivity (WINTER) campaign (Thornton et al., 2014), the Cloud Systems Evolution in the Trades (CSET) campaign (Albrecht et al., 2019), the $O_2$/$N_2$ Ratio and $CO_2$ Airborne Southern Ocean Study (ORCAS) (B. B. Stephens et al., 2018), the Southern Ocean Clouds, Radiation, Aerosol Transport Experimental Study (SOCRATES) (McFarquhar et al., 2021), and the Organization of Tropical East Pacific Convection (OTREC) campaign (Fuchs-Stone et al., 2020). The combined data set covers a latitude range of 67°S–87°N and a longitude range of 128°E–37°W, providing us with a unique analysis of near global-scale cloud phase distributions. The NSF Gulfstream-V research aircraft was the main platform for 10 out of 11 NSF campaigns, except for the WINTER campaign, which used the NSF C-130 research aircraft. Various sizes of cloud hydrometeors are measured by several in situ cloud probes, including the Cloud Droplet Probe (CDP), the Fast-2 Dimensional Cloud (Fast-2DC) probe, and the Two Dimension Stereo (2DS) cloud probe. The measurement ranges of these cloud probes are as follows: 2–50 μm for CDP, 62.5–3,200 μm for Fast-2DC, and 40–5,000 μm for 2DS. For 10 out of the 11 NSF campaigns, the CDP and Fast-2DC probes are used to derive cloud microphysical properties. Only for the NSF SOCRATES campaign, CDP and 2DS probes are used due to quality issues with Fast-2DC probe in several research flights in that campaign.

For each in situ cloud probe, the number concentrations of particles are reported at 1-Hz resolution. The 1-Hz measurements of each cloud probe are further categorized into large aerosols, liquid droplets (including supercooled liquid droplets), and ice crystals. This phase identification method was previous described in D'Alessandro et al. (2019) and Yang et al. (2021) with more details. Fundamentally, this method first identifies cloud phase for each cloud probe. Using the relationship between particle number concentration ($N_{CDP}$) and mass concentration ($M_{CDP}$), 1-Hz CDP observations are defined as (a) large aerosols (satisfying either $N_{CDP} \leq 10^{-1.5}$ cm$^{-3}$ or $M_{CDP} \leq 10^{-3.4}$ g m$^{-3}$); (b) liquid droplets (satisfying both $N_{CDP} \geq 10^{-0.5}$ cm$^{-3}$ and $M_{CDP} \geq 10^{-3.4}$ g m$^{-3}$); or (c) ice crystals (satisfying both $10^{-1.5} < N_{CDP} < 10^{-0.5}$ cm$^{-3}$ and $M_{CDP} \geq 10^{-3.4}$ g m$^{-3}$). Phase identifications for the Fast-2DC and 2DS probes are similar to each other, which use parameters such as ambient temperature, particle number concentrations, maximum particle diameter, and the standard deviation of particle size distribution to define either liquid or ice dominant 1-Hz measurements. Furthermore, visual inspections as well as intercomparisons with other in situ probes such as Rosemount ICEing (RICE) probe and King probe are also part of the quality control routine to verify the automatically designated cloud phase.

After categorizing the 1-Hz observations of individual cloud probes, LWC and ice water content (IWC) are calculated for each second of measurement by each probe. For the 1-Hz observations defined as liquid droplets, spherical shape and liquid water density are used for calculating LWC. To derive IWC for observations defined as ice crystals, IWC of small ice particles (i.e., diameter ≤75 μm) is derived using the mass-Dimension (m-D) relationship for small ice based on Brown and Francis (1995), while the m-D relationship of large ice in that paper was used to derive IWC for large ice crystals (diameter >75 μm). Lastly, the total IWC and total LWC are derived by combining the IWC or LWC from individual cloud probe, respectively. Ice mass fraction is calculated using the total IWC divided by the sum of total IWC and total LWC. Such ice mass fraction at three ranges of 0–0.1, 0.1–0.9, and 0.9–1 each represents liquid, mixed and ice phase based on integrated in situ observations at 1-Hz resolution, respectively. These three cloud phases are further used for validating three satellite products described below.

Even though in situ cloud probe measurements of hydrometeors are generally considered more accurate than remote sensing measurements, measurement uncertainties are unavoidable, especially for the small ice size range. For most of the flight campaigns using Fast-2DC probe, ice crystals with a diameter smaller than 62.5 μm have been removed to reduce the uncertainties. Anti-shattering tips are also equipped on the Fast-2DC and 2DS probes, even though an elimination of shattering is still not possible (e.g., Korolev et al., 2013). Nevertheless, the identification of cloud phases in this analysis relies on mass concentrations of hydrometeors (i.e., IWC and LWC). These values are dominated by larger size particles and therefore are less affected by smaller particles.

### 2.2. NASA CALIPSO (LID_L2_VFM_Standard-V4-20)

The NASA CALIPSO satellite carries the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP), which is a spaceborne dual-wavelength polarization lidar using an Nd:YAG laser that operates at 1,064 and 532 nm (Hu et al., 2009; Winker et al., 2007). CALIOP, which is designed with a broad dynamic range, has the unique advantage of measuring a variety of aerosols and cloud hydrometeors. High-resolution vertical profiles of clouds and aerosols, including their microphysical and optical properties, have been provided near globally since CALIPSO satellite was launched in April 2006.

The primary data product from CALIPSO used in this paper is the Vertical Feature Mask (VFM), which is a Level 2 product. The VFM data categorize the spatial distributions of various features, such as cloud type, aerosol type (e.g., dust, polluted smoke, etc.), and cloud thermodynamic phase (Vaughan et al., 2004). Quantities derived from CALIOP retrievals, such as backscatter and linear depolarization ratio, can be used to identify particle sizes and orientations, which can be further used to separate ice and liquid phases (Sassen, 1991). The CALIPSO ice-water algorithm classifies cloud phases using depolarization ratio, and cloud top and cloud base temperatures. Temperature profiles are obtained from synoptic analysis data. Probability functions for ice, liquid, and oriented plates are parameterized using these variables. A flowchart is used to eventually define three scenarios—ice, water, and unknown.

Cloud phase product is provided by "Feature_Classification_Flags" inside the files named "LID_L2_VFM-Standard-V4-20" (Table 2). The cloud phase product has different vertical and horizontal resolutions at three different altitudinal regions, as explained in Hunt et al. (2009) in their Table 4. The low-altitude region is located between 0.5 and 8.2 km, with a vertical (horizontal) resolution of 30 m (333 m). The mid-altitude region is located from 8.2 to 20.2 km, with a vertical (horizontal) resolution of 60 m (1 km). The high-altitude region is from 20.2 to 30.1 km, with a vertical (horizontal) resolution of 180 m (1.667 km). The cloud phase product includes three categories—liquid, ice, and horizontal-oriented ice (HOI). In this study, HOI is grouped into the ice phase.

### 2.3. NASA CloudSat (2B-CLDCLASS-LIDAR)

The CloudSat cloud phase product is derived from the combination of CALIPSO CALIOP lidar and CloudSat Cloud Profiling Radar (CPR) observations. The CloudSat CPR radar uses a 94 GHz sensor. The signals travel through the atmospheric medium with bidirectional attenuation due to absorption by liquid water droplets, precipitation-sized particles, and gasses (mainly water vapor). Since the lidar and radar have different sensitivities to different sizes and orientations of particles, the combined radar and lidar data could potentially complement each other when identifying cloud layers and cloud phases.

CloudSat utilizes radar and lidar signals, temperatures of cloud top and cloud base, and depolarization ratio for cloud phase identification. The process involves stepwise assignment, considering microphysical and optical properties. Despite higher droplet concentrations, mixed-phase clouds in CloudSat are dominated by ice particles due to Rayleigh scattering effects. Several fundamental differences in ice and liquid phase are used to develop the algorithm for identifying cloud phases (Wang, 2019; Wang & Sassen, 2001), such as variations of fallspeeds (i.e., larger for ice particles than liquid droplets), vertical levels (i.e., generally higher for ice phase), and number concentrations (i.e., higher for liquid droplets). The "Cloud Phase" variable from the 2B-CLDCLASS-LIDAR product is used in this work (Table 2). This cloud phase data set has a vertical resolution of 240 m and a horizontal resolution of 1.4 km by 1.8 km (Deng et al., 2010).

### 2.4. DARDAR Project (DARDAR_CLOUD.v2.1.1)

The DARDAR project uses the integrated CloudSat radar and CALIPSO lidar observations and retrieves cloud properties. This project is supported by Laboratoire Atmosphères, Milieux, Observations Spatiales and the University of Reading. The cloud phase product used in this study is part of the DARDAR-CLOUD data set. It is similar to the CloudSat 2B-CLDCLASS-LIDAR product as they both use CALIOP and CPR measurements, but DARDAR uses a different algorithm to define cloud phases. DARDAR integrates radar and lidar forward models, refining a state vector iteratively for cloud phase determination. The basic variables used include visible extinction coefficient, extinction-to-backscatter ratio, normalized number concentration parameter, etc. Among all three satellite algorithms, the DARDAR algorithm is the only one that has explicit temperature dependence for phase identification, partly because of the temperature dependence of the number concentration parameter of a size distribution. The DARDAR-MASK categorization excludes lidar signals below detected liquid pixels and uses radar, lidar, and infrared radiances for ice cloud properties.

The algorithm "Varcloud" is developed based on Delanoë and Hogan (2008a, 2008b), providing several cloud properties, including IWC, visible extinction coefficients, ice cloud effective radius, etc. The cloud phase variable is stored as "DARMASK_Simplified_Categorization," which has a vertical resolution of 60 m and a horizontal resolution of 1.4 km (Delanoë & Hogan, 2008a, 2008b). DARDAR cloud mask includes ice, ice + supercooled, liquid warm, supercooled, rain, and other aerosol indicators (Table 2). In this study, liquid warm, supercooled, and rain are classified as liquid phase. The category called "ice + supercooled" is categorized as the mixed phase.

## 3. Comparison Methods

This study developed two methods to compare satellite observations with in situ airborne observations (Figure 1). The main differences between the two methods are the spatial and temporal windows used to select satellite observations surrounding in situ observations. Ideally, selecting a satellite observation sample that is perfectly collocated at the same time and location as an in situ observation sample would minimize the impacts of spatial and temporal mismatches on the comparison results. Yet this type of perfect collocation is very rare in reality, since both satellites and aircraft are moving objects in space and time with different trajectories and velocities. For example, CALIPSO and CloudSat are both sun-synchronous and orbit the Earth at an altitude of 705 km with a speed of approximately 7 km/s (G. L. Stephens et al., 2002; Winker et al., 2003). Conversely, the NSF Gulfstream-V research aircraft usually operates between the surface and 14 km in altitude with a true airspeed around 200–250 m/s. Furthermore, aircraft campaigns often have limited geographical coverage and durations compared with the near global coverage of satellite observations. In addition, very few campaigns were specifically designed to follow the satellite trajectories in order to validate satellite observations, such as those comparisons conducted in Barker et al. (2008). To overcome these challenges, this study developed two methods for sampling satellite data at various ranges of spatiotemporal windows, described as follows.

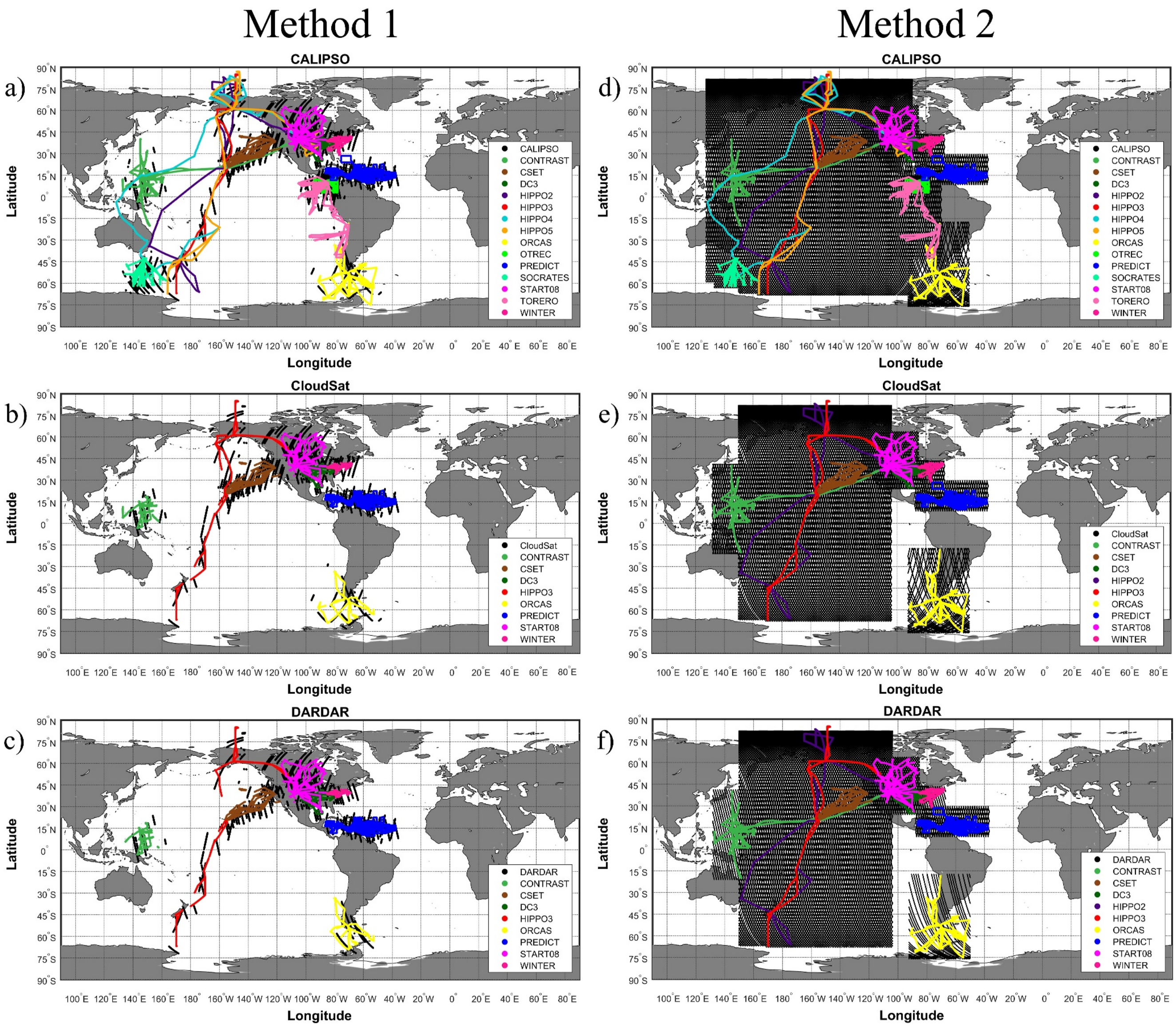

**Figure 1.** A global map for flight tracks of 11 NSF campaigns and three satellite cloud phase products used for comparisons. CALIPSO, CloudSat and DARDAR are selected via (a–c) Method 1 and (d–f) Method 2, respectively. Aircraft flight tracks in (d–f) represent the entire campaigns, while (a–c) only show those flight tracks in close proximity to satellite samples.

### 3.1. Method 1: Selecting Satellite Samples in Proximity to 1-Hz In Situ Observations

Method 1 first selects satellite samples that are within a time window, which is from 12 hr before the first second of each research flight (RF) to 12 hr after the last second of that flight. These samples are further reduced to a spatial window bounded by the maximum and minimum values of the latitude and longitude of that specific flight. For each pair of satellite and 1-Hz in situ observation samples being compared, their differences in time (*dTime*) and in horizontal space (*dDist*) are calculated. Specifically, *dDist* represents the arc distance between two samples using their latitudinal and longitudinal positions. An additional restriction is applied to each satellite sample, that is, *dDist* has to be within ±200 km surrounding each second of in situ observations. This method of selecting a range of *dTime* and *dDist* for comparing satellite and aircraft observations is similar to that used in Diao et al. (2013) as illustrated in their Figure 4.

After applying these restrictions, vertical columns of satellite samples are selected surrounding each second of in situ observations. To further restrict the vertical levels for comparisons, each second of aircraft is matched with only one altitudinal bin inside a vertical column of certain satellite product. This altitudinal bin is the layer where the 1-D flight track transected through. The satellite samples selected via Method 1 for three data sets—CALIPSO, CloudSat and DARDAR are illustrated in Figures 1a–1c, respectively. The CALIPSO and CloudSat longitudinal and latitudinal positions were based on their respective orbit as part of the A-Train constellation, while the location of DARDAR product is based on the collocated footprints of CALIPSO-CloudSat at the ground (Delanoë & Hogan, 2008a, 2008b).

### 3.2. Method 2: Selecting Satellite Samples for Each Flight Campaign

Compared with Method 1, Method 2 uses a wider spatiotemporal window to select satellite data for individual campaigns, which allows more satellite samples to be evaluated in the comparisons. The temporal window of Method 2 starts from 1 day before the first day of each campaign and ends at 1 day after the last day of each campaign. A campaign domain is defined by the minimum and maximum values of latitude and longitude for each campaign. Method 2 extends 1° from each side of such campaign domain to select satellite observations surrounding each campaign. Another difference between the two methods is that Method 2 uses the entire vertical columns of satellite samples to conduct a statistical comparison with in situ observations. Spatial distributions of satellite data sampled via Method 2 are illustrated in the global map in Figures 1d–1f.

Overall, Method 1 serves the purpose of choosing satellite samples that are relatively closer to the in situ observations to reduce impacts of spatial and temporal mismatches on the comparison results. Method 2, on the other hand, serves the purpose of providing a larger sample size of satellite observations to be evaluated against in situ observations, which can be further used to examine variabilities in different seasons and latitudinal and longitudinal bands.

## 4. Results

### 4.1. A Case Study Comparison Using One Research Flight

A case study is conducted to illustrate the comparisons of cloud phase detected by in situ aircraft observations and three satellite products (Figure 2). This case is the RF08 from the NSF START08 flight campaign on 6 May 2008. The START08 campaign targeted the physical and chemical processes in the extratropical upper troposphere (UT) and lower stratosphere (LS). The outbound part of the flight sampled across a weakening squall line in eastern Oklahoma and northern Texas based on radar observations from the National Weather Service. The inbound part of the flight followed an air mass with low stability. The final descent back to Broomfield, Colorado sampled a weak convection with virga.

The satellite data shown in this case study are selected in a similar way as Method 1 but do not apply any restrictions on *dDist* or altitudinal levels for illustrative purposes. Figure 2a illustrates the relative positions between in situ and satellite observations in a 2-D map, while Figures 2b and 2c illustrate the constantly changing *dDist* and *dTime* between observation platforms. The three satellite products show almost identical locations and times relative to the aircraft observations, indicating minimal impacts from spatial and temporal variations among these three satellite products on their evaluation results.

The vertical structure of cloud layers are investigated through curtain plots (Figures 2d–2f). The satellite products show an extensive, thick layer of ice clouds in the middle to upper troposphere, which may originate from the anvil outflow cirrus from the squall line. All three satellite products detected cloud layers most of the time when cloud hydrometeors were detected in aircraft observations. For the flight track segments closer to Broomfield, CO (i.e., latitudes above 40°N), the satellites were further away from the aircraft track as illustrated by larger *dDist* values in Figure 2b. This may be the main reason that ice crystals sampled around this area in the flight were not detected by the satellites. Comparing the in-cloud occurrences among three satellite products, CALIPSO shows the lowest cloud coverage, with mostly clear-sky conditions between 2 and 8 km. This could be due to the lidar signal attenuation through thicker ice cloud layers at higher altitudes. On the contrary, both CloudSat and DARDAR show extensive cloud coverage from surface to 12 km in this case. The in-cloud occurrences of CloudSat data are also very similar to DARDAR, showing similar cloud fraction and location.

When further examining the separation between ice and liquid phase, the ice and liquid hydrometeors observed by the research aircraft cloud probes are mostly captured correctly by three satellite products, with ice hydrometeors being detected as either ice or mixed phase, and liquid hydrometeors being detected as liquid or mixed phase by

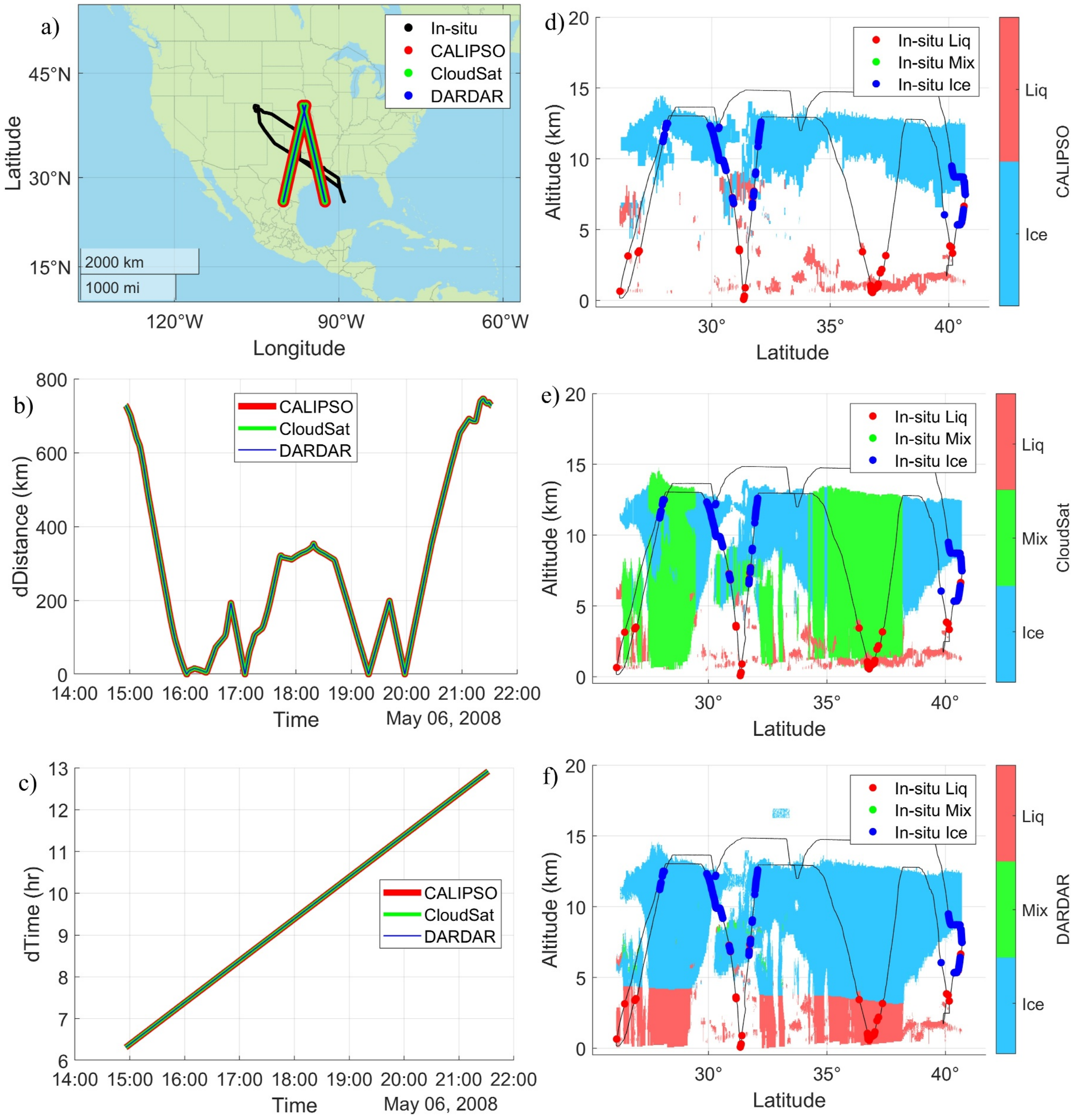

**Figure 2.** A case study of NSF START08 RF08 comparing in situ observations with three satellite products. (a) A map of the locations of in situ and satellite observations. (b) *dDist*, arc distance between research aircraft and satellites. (c) Time differences, *dTime*. (d–f) Curtain plots of satellite-based cloud phases (in shades) and in situ observations of liquid, mixed and ice phase (in markers).

satellites. To further contrast the three satellites on their phase partitioning, DARDAR shows a distinct separation between ice and liquid phases around 3–4 km, which indicates a temperature-dependent threshold used in their cloud phase partitioning algorithm. CALIPSO and CloudSat have almost identical occurrences of liquid phase, except for a few occasions when CALIPSO allows liquid phase to occur up to 8 km (e.g., near 32° latitude), while CloudSat categorizes these points as mixed phase. A unique feature of CloudSat product is the thick layers of mixed phase stretching from 2 to 12 km. In comparison, the CALIPSO data do not include mixed-phase clouds. The DARDAR cloud phase includes mixed phase but has very few occurrences. Both CloudSat and DARDAR detected deep convection and anvil structures in both outbound and inbound parts of the flight, but their cloud phases differ from each other. CloudSat identified two convective pillar clouds with anvil clouds on the side, likely altostratus and/or cirrus clouds, as well as rainbands on the south of deep convection categorized as liquid and mixed-phase clouds in an alternating pattern. The two convective pillars are mostly categorized as mixed

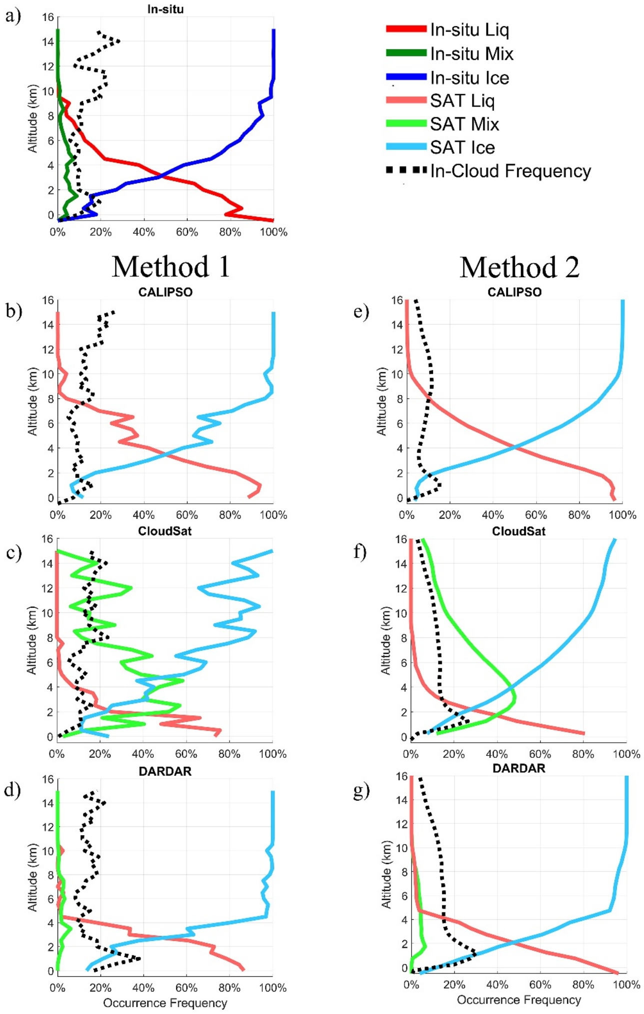

**Figure 3.** Vertical profiles of occurrence frequencies of three cloud phases and in-cloud conditions. (a) In situ observations. (b–d) Satellite data using Method 1. (e–g) Satellite data using Method 2.

phase in CloudSat but are categorized as partly ice and partly liquid by DARDAR. For the location around 36.5°N–37.5°N, DARDAR shows better agreement with the in situ observations as they both capture liquid phase, while CloudSat shows mixed phase.

Overall, this case study illustrates that all three satellite products can capture cloud occurrences at similar locations as the in situ observations, even though that observations have spatiotemporal mismatches from satellite data. This case study also illustrates the key differences among three satellite products in terms of in-cloud occurrences and cloud phase partition. The distinct feature of lower in-cloud occurrences seen in CALIPSO is consistent with the physical limitation of CALIOP lidar as previously mentioned for penetrating thicker clouds (Winker et al., 2007; Zhang et al., 2010). On the other hand, CloudSat and DARDAR cloud phase products show higher in-cloud occurrences than CALIPSO, since CPR radar can penetrate optically thick clouds (Im et al., 2005).

## 4.2. Vertical Distributions of Cloud Phase Frequencies Using Eleven NSF Campaigns

The cloud phase occurrence frequencies at various altitudes are examined using all 11 NSF flight campaigns (Figure 3). In-cloud frequency is calculated in each 500-m bin as the number of in-cloud samples (i.e., the sum of all cloud phase samples) divided by the total number of samples (i.e., clear-sky plus in-cloud samples) in that bin. In addition, cloud phase frequency is calculated for each vertical bin as the number of samples of a specific cloud phase divided by the number of in-cloud samples. That is, cloud phase frequency represents the frequency of a cloud phase among all cloud phases, while the in-cloud frequency represents the frequency of in-cloud conditions among all-sky conditions. The number of samples of each cloud phase and total in-cloud conditions are shown in Figure S1 in Supporting Information S1.

The in-cloud frequency based on in situ observations (Figure 3a) ranges from 10% to 30%, with two peak frequencies—20% located at 1 km and 30% at 14 km in altitude. We first examine samples selected by Method 1 to evaluate satellite samples in closer proximity to aircraft observations (Figures 3b–3d). All three satellite products selected by Method 1 show in-cloud frequencies around 10%–20% from 2 to 15 km, similar to the magnitudes seen in in situ observations. CALIPSO also shows lower in-cloud frequency than CloudSat by ∼10% at 2–12 km, likely due to its lidar signal attenuation, consistent with the finding in Figure 2d. For low-cloud distributions, DARDAR shows a higher peak position of low-cloud occurrence frequency (40%) at 1 km (Figure 3d), while CALIPSO and CloudSat show peak frequency of 18% at 1 and 2.5 km, respectively. Comparing the high-cloud frequencies, the slightly higher in-cloud frequency above 13 km in the in situ observations may be the optically thin cirrus clouds that are below the detection limits of remoting sensing techniques. In fact, CALIPSO shows a slightly higher in-cloud frequency in the 14.5–15 km bin compared with CloudSat and DARDAR, possibly originating from the higher sensitivity of lidar signals to small ice particles compared with radar signals.

Method 2 (Figures 2e–2g), which allows more satellite samples in wider spatiotemporal windows for comparisons, are contrasted with Method 1. The altitudinal fluctuations of in-cloud frequencies shown in Method 1 are more smoothed out when using larger sample sizes in Method 2. In addition, Method 2 shows similar in-cloud frequencies as Method 1 between 2 and 12 km for three satellite data, with ±5% differences between the two methods. All satellite data show low-cloud peak frequencies at 1 km, consistent with in situ observations. However, Method 2 shows lower high-cloud frequencies at 12–15 km for all three satellite products, which may be due to the flight designs of certain campaigns, such as the DC3 campaign targeting anvil outflow cirrus, and the PREDICT campaign targeting the upper levels of tropical cyclones. These specific flight designs may lead to a slightly higher in-cloud frequency in the UT/LS region around the flight track compared with other regions in the entire flight domain. This result indicates that in-cloud frequency comparisons are highly sensitive to the spatial and temporal proximity between different observation platforms. As a result, smaller spatiotemporal comparison windows are recommended for in-cloud frequency comparison.

Cloud phase frequency distributions of all observations show increasing frequencies of ice phase and decreasing frequencies of liquid phase with increasing altitudes, consistent with the fundamental thermodynamic process of liquid water converting to ice as temperatures decrease. Unlike the comparisons of in-cloud frequencies that show different results using Methods 1 and 2, both comparison methods show similar features of cloud phase frequency distributions, including the trend relative to altitude and the frequency range of each cloud phase.

The locations where ice and liquid-containing phase (i.e., liquid plus mixed phase) have the same occurrence frequencies are located at 3.75, 3.75, 4.5, and 3 km for in situ observations, CALIPSO, CloudSat, and DARDAR, respectively (Figures 3a–3d). The reason that DARDAR shows slightly lower altitudes of this level than in situ observations is because it has higher frequencies of ice phase from surface to 10 km than in situ observations. The locations above which ice phase frequencies exceed 95% are 9, 9, 15, and 4.5 km for the in situ and three satellite data sets, respectively. In addition, the locations below which liquid phase frequencies exceed 90% are 0.5, 1, 0.5, and 0.5 km for these data sets, respectively. Comparing the slopes of ice frequency distributions where ice phase transitions across 10% and 90%, CALIPSO has the most similar slopes at both locations compared with in situ observations. Both CloudSat and DARDAR do not capture the slope at 10% ice frequency. CloudSat shows the other slope at 90% ice frequency being too gradual, while the DARDAR shows it being too steep.

The frequencies of mixed phase from in situ observations are relatively low (i.e., less than 5%). DARDAR shows similar low frequencies of mixed phase less than 5%. CloudSat significantly overestimates mixed phase frequencies in all altitudes with frequency values ranging from 10% to 60%, consistent with the feature shown in the case study (Figure 2e). These spurious mixed phase in CloudSat likely should have been categorized as all ice phase above 8 km and as all liquid phase below 1 km, since the in situ observations show >90% of ice and liquid phase at these two altitude ranges, respectively. The peak position of in situ observed mixed phase is around 1.5–5 km, which is close to where ice and liquid phases have similar frequencies at 3 km. Interestingly, for CloudSat and DARDAR, the peak positions of mixed phase frequencies are also concurrent with their respective levels with equal frequencies of ice and liquid phases. This feature indicates that the locations where both ice and liquid phases occur frequently as separate cloud segments (near 50% frequency) also have the highest probability of generating and maintaining mixed-phase cloud segments.

The unknown categories in satellite products are excluded from the cloud phase analysis in this work. To test the impacts of the unknown category (only provided in CALIPSO and DARDAR data) on the cloud phase frequencies, Figure S2 in Supporting Information S1 illustrates the contrast between including or excluding unknown as a cloud phase type. Satellite data in the entire year of 2010 are shown without the collocation restriction to aircraft observations. Even though including the unknown category decreases the frequencies of other cloud phases, the altitude that marks the transition of the dominant cloud phase (ice vs. liquid) remains the same. The ice and liquid phase frequency lines intersect at 3 and 2 km for CALIPSO and DARDAR, respectively.

### 4.3. Latitudinal Distributions of Ice Phase Frequency in a Near-Global View

The zonal average of ice phase occurrence frequency is compared between in situ and satellite observations using comparison Method 1 (Figure 4) and Method 2 (Figure 5). The ice phase frequency is first calculated as the number of ice phase samples divided by the total number of in-cloud samples in a given bin, which has a vertical resolution of 500 m and a horizontal resolution of 2.5° latitude. In addition, a $3 \times 3$ moving average is applied to each latitudinal by altitudinal bin to reduce the impacts of missing data in certain bins. Ice frequency lines (IceF) are defined as the lowest altitudes where ice phase frequency reaches a certain percentage within each 2.5° latitudinal bin, including frequencies of 10% ($\mathrm{IceF}_{10}$), 50% ($\mathrm{IceF}_{50}$), and 90% ($\mathrm{IceF}_{90}$). A moving average is also applied to three IceF lines after they are calculated from high-frequency data. IceF lines illustrate the transition between liquid-containing phase and ice phase at various altitudes and latitudes. We define the altitudinal range between $\mathrm{IceF}_{10}$ and $\mathrm{IceF}_{90}$ as the mixing region. The number of samples used for the analysis of Figures 4 and 5 are shown in Figure S3 in Supporting Information S1.

Three different sub-sets of in situ observations are shown in Figures 4a–4c, which represent the in situ observations that satisfy the spatiotemporal collocation criteria of Method 1 and can be further compared with

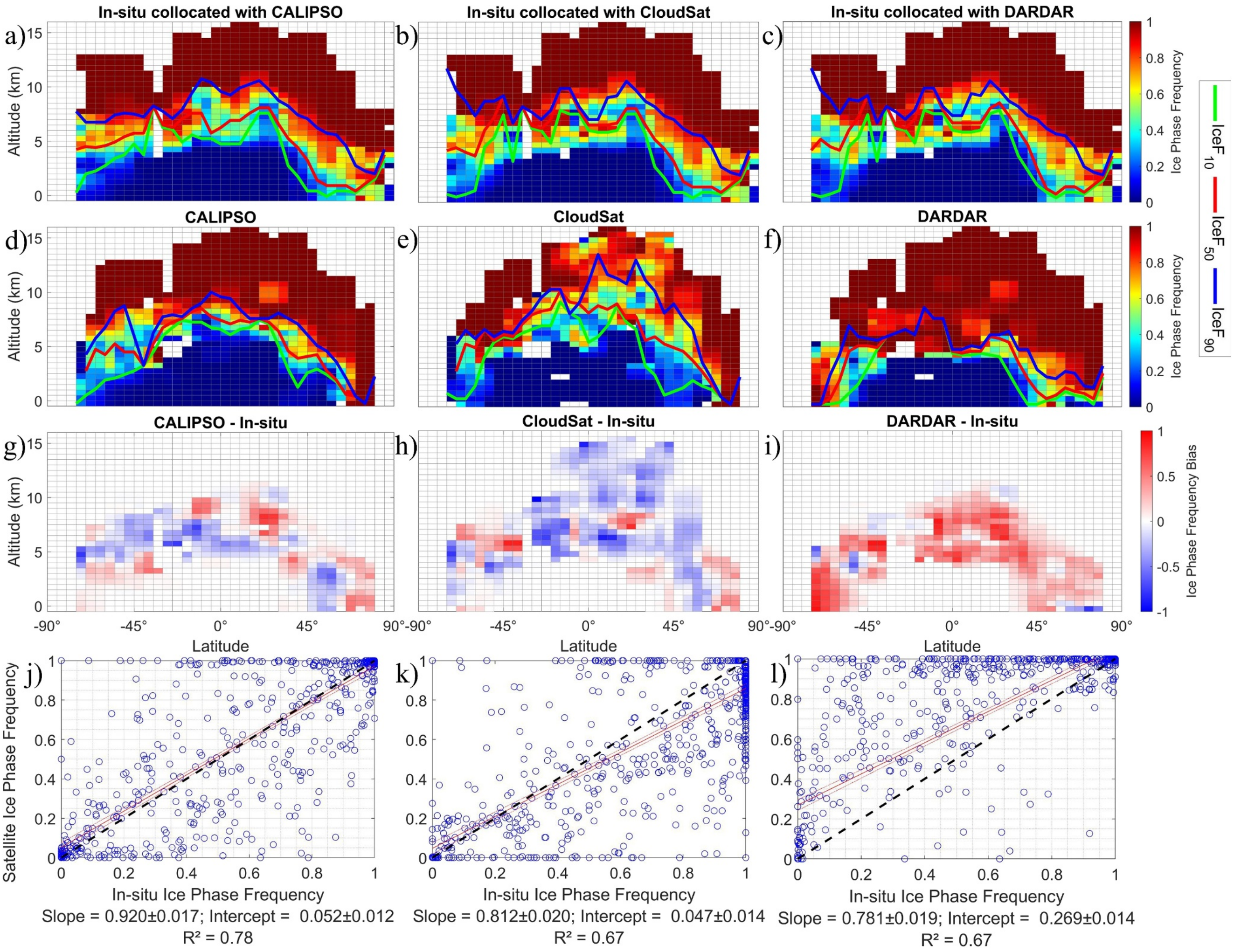

**Figure 4.** Latitudinal—altitudinal distributions of ice phase occurrence frequency using Method 1. Columns 1–3 represent comparisons for three satellite products—CALIPSO, CloudSat, and DARDAR, respectively. Rows 1–3 represent in situ data, satellite data, and their differences (satellite minus in situ), respectively. Note that in situ data in row 1 are not identical to each other because they are selected via Method 1 to be collocated with different satellite products. Row 4 represents the correlation analysis between in situ and satellite data from the top two rows. Each blue marker represents a pair of ice phase frequencies from the same latitudinal and altitudinal bin. Linear regression fit is shown as the middle red solid line, while the upper and lower red dotted lines represent the 95% simultaneous confidence bounds.

CALIPSO, CloudSat, and DARDAR, respectively. Similarly, the top row of Figure 5 represents the sub-sets of in situ observations that satisfy the spatiotemporal selection criteria of Method 2 for the comparisons with respective satellite data set. Note that even though Method 2 allows satellite products within the entire domain of the flight campaigns to be used for comparison, data availability still varies among different satellite data and therefore not all aircraft observations are matched with all three satellite products, leading to the variations in the sub-sets of in situ observations in the top row.

Latitudinal dependence of ice phase frequency distribution is shown in all data sets, with decreasing altitudes of three IceF lines from the lower to the higher latitudes. Comparing with the in situ observations, CALIPSO shows lower $\text{IceF}_{90}$ by 1–2 km, CloudSat shows higher $\text{IceF}_{90}$ by 2–5 km, and DARDAR shows lower $\text{IceF}_{90}$ by 2–5 km. On the other hand, CALIPSO and CloudSat show more similar altitudes of $\text{IceF}_{10}$ lines compared with in situ observations, while DARDAR still shows lower $\text{IceF}_{10}$ line by 1–3 km. This result indicates that all satellite data capture the vertical levels of where ice crystals start to contribute to a non-negligible fraction of clouds more accurately, but they have more difficulty of capturing where supercooled liquid water disappears at lower temperatures.

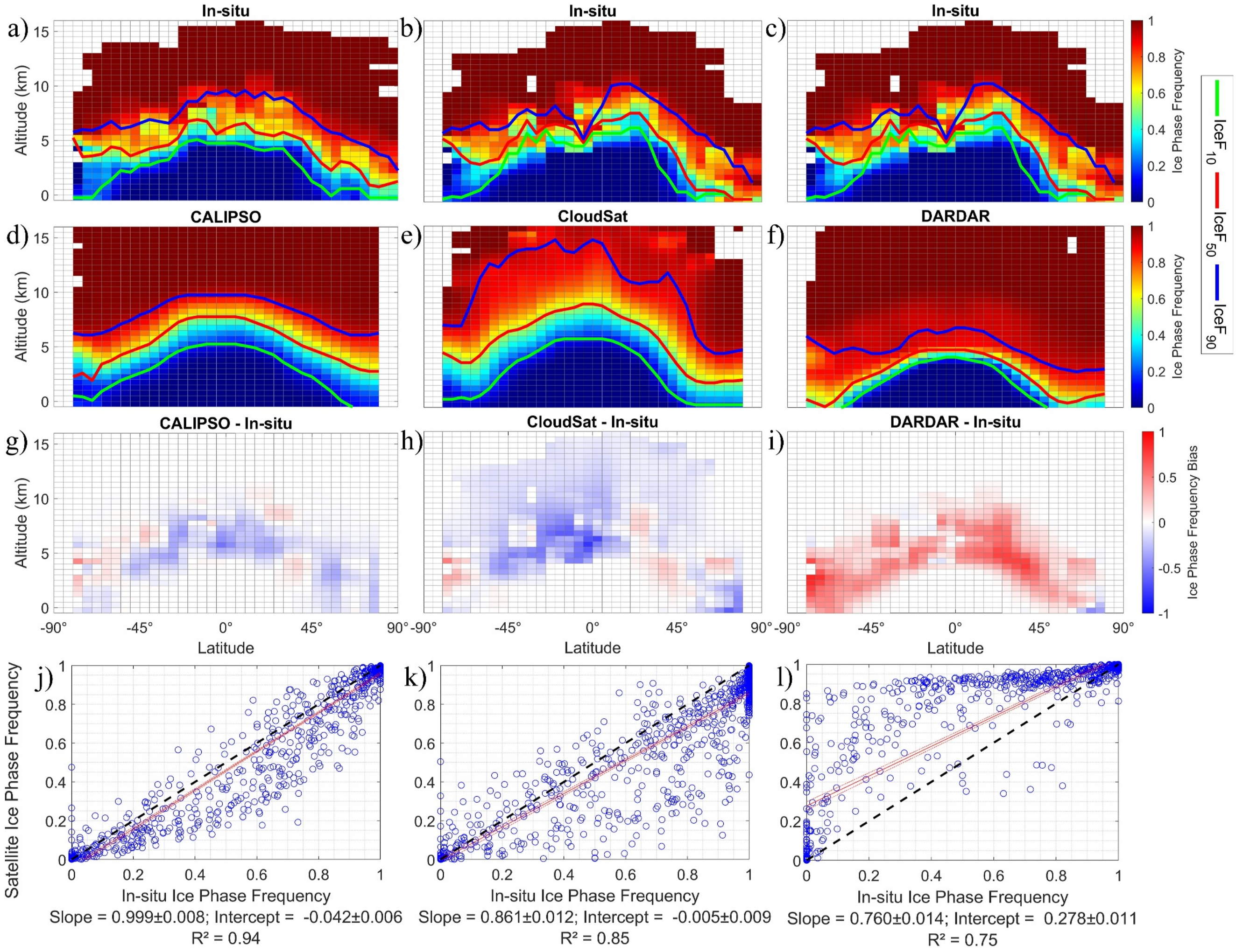

**Figure 5.** Similar to Figure 4, except for using Method 2. Unlike Figure 4, in situ observations in this figure represent the entire flight campaigns.

Among all three satellite cloud phase products, CALIPSO (Figure 4d) shows the most similar location and thickness of mixing region at various latitudes compared with in situ observations, even though CALIPSO still shows a slightly thinner of mixing region (2–3 km) than the in situ observations (3–5 km). Comparatively, the mixing region thickness is overestimated and underestimated by CloudSat and DARDAR, respectively. These biases of mixing region thickness are consistent with the transitions from liquid to ice being too gradual for CloudSat and too sharp for DARDAR (as seen in Figure 3). Comparing the two hemispheres, the mixing regions are located at 5–10 km in the tropics (30°N–30°S), 0–6 km in the Northern Hemisphere (NH) midlatitudes, and 2–7.5 km in the Southern Hemisphere (SH) midlatitudes based on in situ observations. One possible explanation for the higher ice phase frequencies in the NH shown in in situ observations is that the majority of flight campaigns in the SH took place in the austral summer (e.g., SOCRATES and ORCAS took place in January–February, and only a few flights from HIPPO were in austral fall), while the flight campaigns in the NH occur in all seasons. The impacts of seasonal variability will be further discussed in Section 4.5. Interestingly, the in situ observations show a small increase of $\mathrm{IceF}_{90}$ height at the NH and SH polar regions with increasing latitudes. Such an increase is captured by satellite data for the NH polar region, which indicates a possible role of insufficient amount of ice nucleating particles (INPs) that inhibits ice formation.

Differences in ice phase frequencies (i.e., calculated as satellite minus in situ observations) are shown in the third row of Figure 4. CALIPSO shows a mixture of positive and negative biases in ice phase frequency, which underestimates ice phase frequency in the mid-troposphere around the mixing region and overestimates it above and below the mixing region. CloudSat shows most negative biases except for the polar regions, while DARDAR shows almost exclusively positive biases.

Figure 5 shows similar main features compared with Figure 4, including similar location and thickness of mixing region for in situ observations, a mixture of positive and negative biases of ice phase frequencies in CALIPSO, significant negative biases in CloudSat and positive biases in DARDAR. Similar to the smoothing effect of Method 2 shown in Figure 3, Figure 5 also shows reduction of fluctuations in three IceF lines as well as more samples in higher altitudes in all satellite data when larger sample sizes are analyzed. The separation between three IceF lines is also more distinguishable. In situ observations show that the thickness from $IceF_{10}$ to $IceF_{50}$ is about the same as that from $IceF_{50}$ to $IceF_{90}$. This feature is also well represented by CALIPSO. CloudSat shows a similar depth from $IceF_{10}$ to $IceF_{50}$, but a much thicker level from $IceF_{50}$ to $IceF_{90}$, consistent with its overestimation of mixed phases at higher altitudes. DARDAR shows a similar depth from $IceF_{50}$ to $IceF_{90}$, but a narrower range from $IceF_{10}$ to $IceF_{50}$, consistent the sharp transition from liquid to ice seen in Figure 3. Overall, the similarities between Methods 1 and 2 for the main features seen in the in situ observations and each satellite product demonstrate the statistical robustness of these comparison results.

Linear regressions between satellite and aircraft data are applied at the bottom rows in Figures 4 and 5. The main features are consistently shown in the correlation analysis, including (a) CALIPSO producing the best comparisons with in situ observations with slopes closest to one (i.e., 0.920 and 0.999), (b) CloudSat underestimates ice phase frequency with slopes of 0.812 and 0.861, and (c) DARDAR overestimates ice phase frequency with slopes of 0.781 and 0.760 and large positive intercepts of 0.269 and 0.278 for Methods 1 and 2, respectively. Even though DARDAR has slopes even at lower values than CloudSat, the large intercepts of DARDAR dominate the overestimation of ice phase frequencies through a wide range of altitudes.

Figures 6a–6d shows the cloud phase frequency distributions of in situ and satellite data at various temperatures, using the aircraft temperature measurements as a proxy for the closest satellite point selected via Method 1. In addition, we applied temperature-based modifications on satellite cloud phase data, resetting all cloud phases to all liquid at temperatures $\geq 0°C$, and to all ice at temperatures $\leq -40°C$ (Figures 6e–6h). CALIPSO consistently shows the most similar cloud phase frequency to in situ observations before modification. DARDAR originally shows higher ice phase frequencies at temperature ranges of −25°C to −20°C and −5°C to 10°C, and lower ice phase frequencies at −20°C to −10°C. It becomes more similar to the observations after the phase modification. The largest discrepancies for CloudSat are seen in the overestimation of mixed-phase frequency, even at temperatures below −40°C or above 0°C, while the ice phase frequencies of CloudSat compare reasonably well with in situ data. This indicates that the main issue from CloudSat data originates from the algorithms calculating mixed-phase frequency distributions that control the partitioning between liquid and mixed phases. Even though the lidar-only product from CALIPSO exhibits a strong correlation with in situ observations, due to the physical limitations of lidar, its observations are confined to the top layer of the cloud as depicted in Figure 2. This limitation is also evident in the vertical profile of in-cloud frequencies of CALIPSO (Figure 3), which detected fewer low-level clouds compared with CloudSat and DARDAR. In contrast, radar-lidar combination products underscore the importance of full vertical detection. The cloud phase differences presented here emphasize the need for improved algorithms in upcoming generations of atmospheric observation satellites. Furthermore, it is worth noting that CloudSat's radar instrument is single-band. The proposed inclusion of a dual-band radar system in NASA's Atmosphere Observing System (AOS) and INvestigation of Convective UpdraftS (INCUS) missions could provide the satellite community with enhanced capabilities for analyzing and developing improved cloud phase detection algorithms.

### 4.4. Regional Variations of Ice Phase and In-Cloud Frequencies

Average vertical profiles of ice phase frequencies (Figure 7) and in-cloud frequencies (Figure 8) are examined for three regions: Southern Ocean (35°S–70°S), tropics (5°S–30°N), and NH extratropics (30°N–90°N). In addition, effects of temporal variability are examined using *dDist* within 200 km and three temporal windows, that is, *dTime* within 3, 6, and 12 hr. Such analysis of the temporal window is similar to the method used in Diao et al. (2013). Ice phase frequency and in-cloud frequency follow the same definition as those used in Figure 3. Only the in situ observations that have collocated satellite samples are shown in each panel, which leads to slight variation in the observed profiles depending on which satellite product is being evaluated.

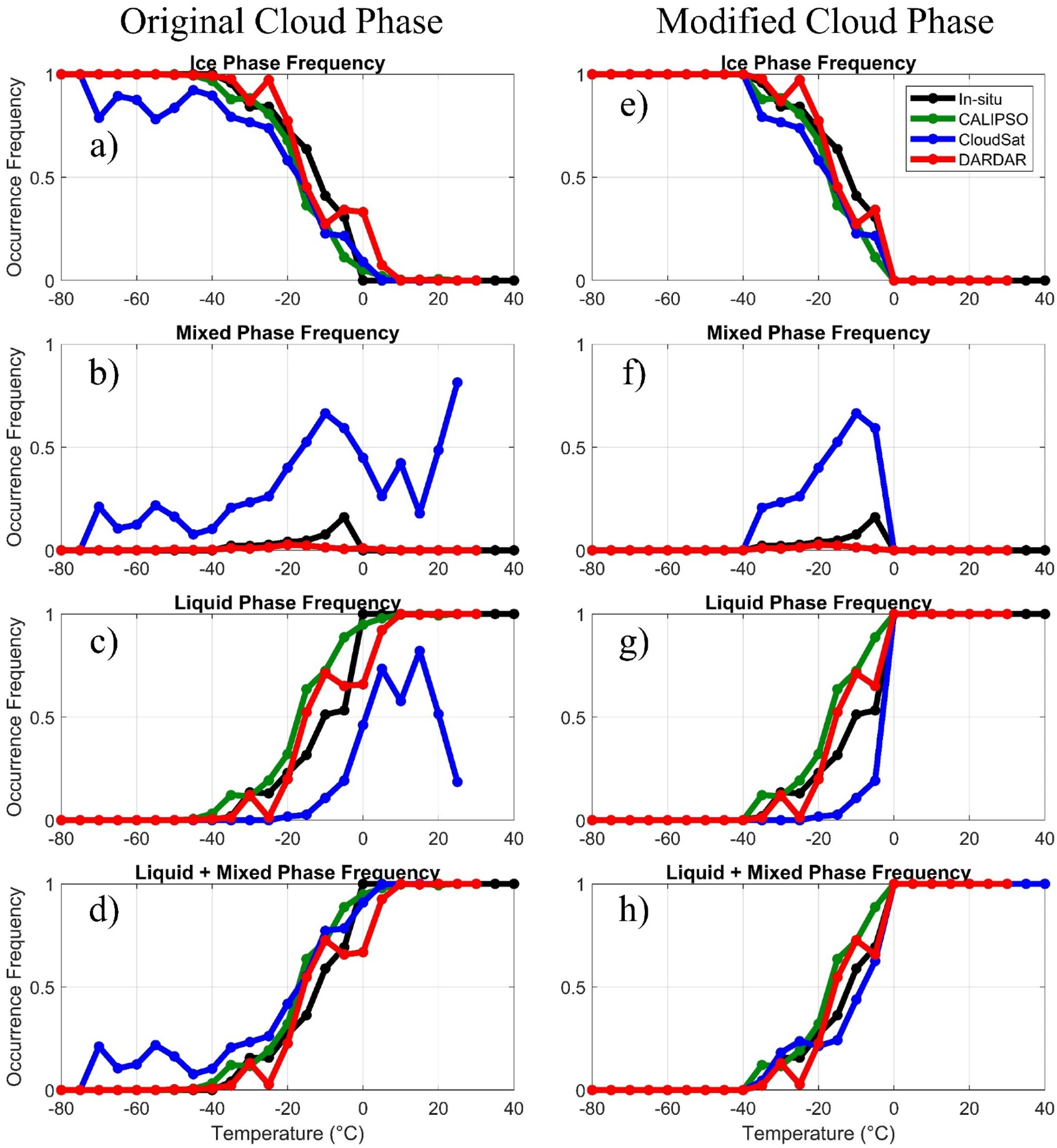

**Figure 6.** Cloud phase frequency distributions at various temperatures, using (a–d) the original satellite cloud phase data and (e–h) the modified satellite cloud phase data. That is, satellite cloud phases are modified to all liquid at temperatures ≥0°C, and all ice at temperatures ≤−40°C.

Regional variations are first examined based on in situ observations. The tropics have lower ice phase frequency at each altitudinal bin compared with the Southern Ocean and NH extratropics due to the higher elevations of isotherms in the tropics. Over the Southern Ocean, the transition from liquid to ice follows almost a linear line based on in situ observations (Figure 7a). That is, ice phase frequency increases from 0 to 0.5 when altitudes increase from surface to 5 km, then from 0.5 to 1 when altitudes increase from 5 to 8 km. In the tropics (Figure 7b), such phase transition occurs more rapidly when ice phase frequency increases from 0 to 0.5 within the range of 4–5.5 km, while the latter part of the increasing frequency from 0.5 to 1 occurs more gradually (i.e., 5–11 km). In the NH extratropics (Figure 7c), the ice phase frequency only changes by 0.1 below 2.5 km or above 6 km, but changes drastically from 0.2 to 0.9 from 2.5 to 6 km. Overall, the Southern Ocean has the thickest mixing region (i.e., $IceF_{10}$ to $IceF_{90}$ is between 1 and 7 km), indicating a longevity of mixed-phase clouds where supercooled liquid water may coexist with ice particles. This result is consistent with numerous observational studies that reported extensive distributions of supercooled liquid water over the Southern Ocean (e.g., Desai et al., 2023; Yang et al., 2021). The NH extratropics has the thinnest mixing region between 2.5 and 6 km, indicating that once

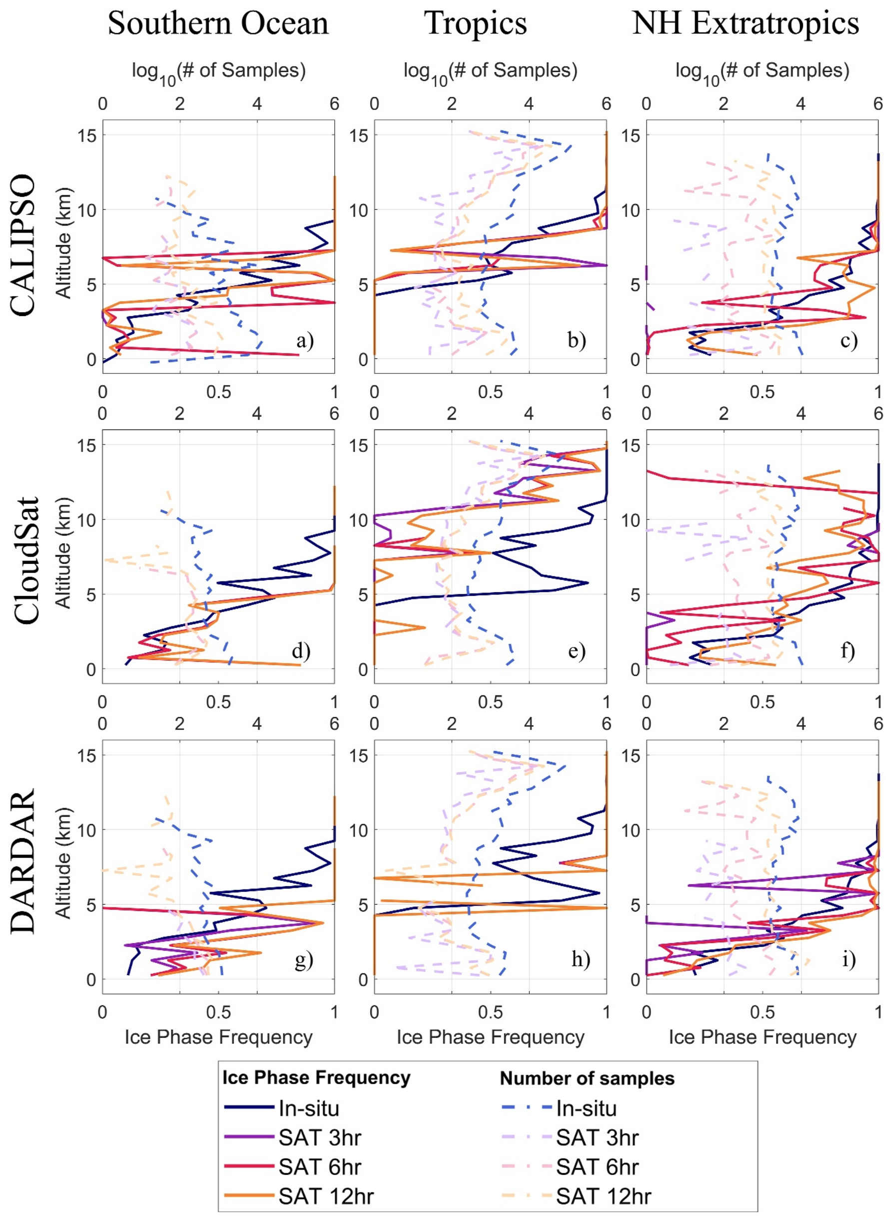

**Figure 7.** Regional variations of the vertical profiles of ice phase frequency. Columns 1–3 represent three latitudinal bands—the Southern Ocean, tropics, and NH extratropics. Solid lines denote ice phase frequency with different time window restrictions. Dotted lines denote the $\log_{10}$ number of samples of in-cloud conditions within different time windows restrictions.

ice crystal forms in the NH extratropics, supercooled liquid water converts to ice phase more rapidly, such as via Wegener-Bergeron-Findeisen process, glaciation, and riming.

Three satellite data are able to capture the increasing ice phase frequencies with altitudes, as well as the lower ice phase frequencies in the tropics for the same altitudes compared with the other two regions. Main biases in three satellite data are summarized as follows. CALIPSO and CloudSat overestimate ice phase frequency in the Southern Ocean mid-troposphere (4–6 km) and underestimate it in the NH extratropical lower troposphere (0–4 km). CloudSat also underestimates ice phase frequency at 5–15 km in the tropics. DARDAR consistently overestimates ice phase frequencies in most locations.

Broadening the temporal window from 3 to 6 and 12 hr shows relatively small impacts on ice phase frequencies over the Southern Ocean and the tropics, but larger impacts over the NH extratropical lower troposphere, where

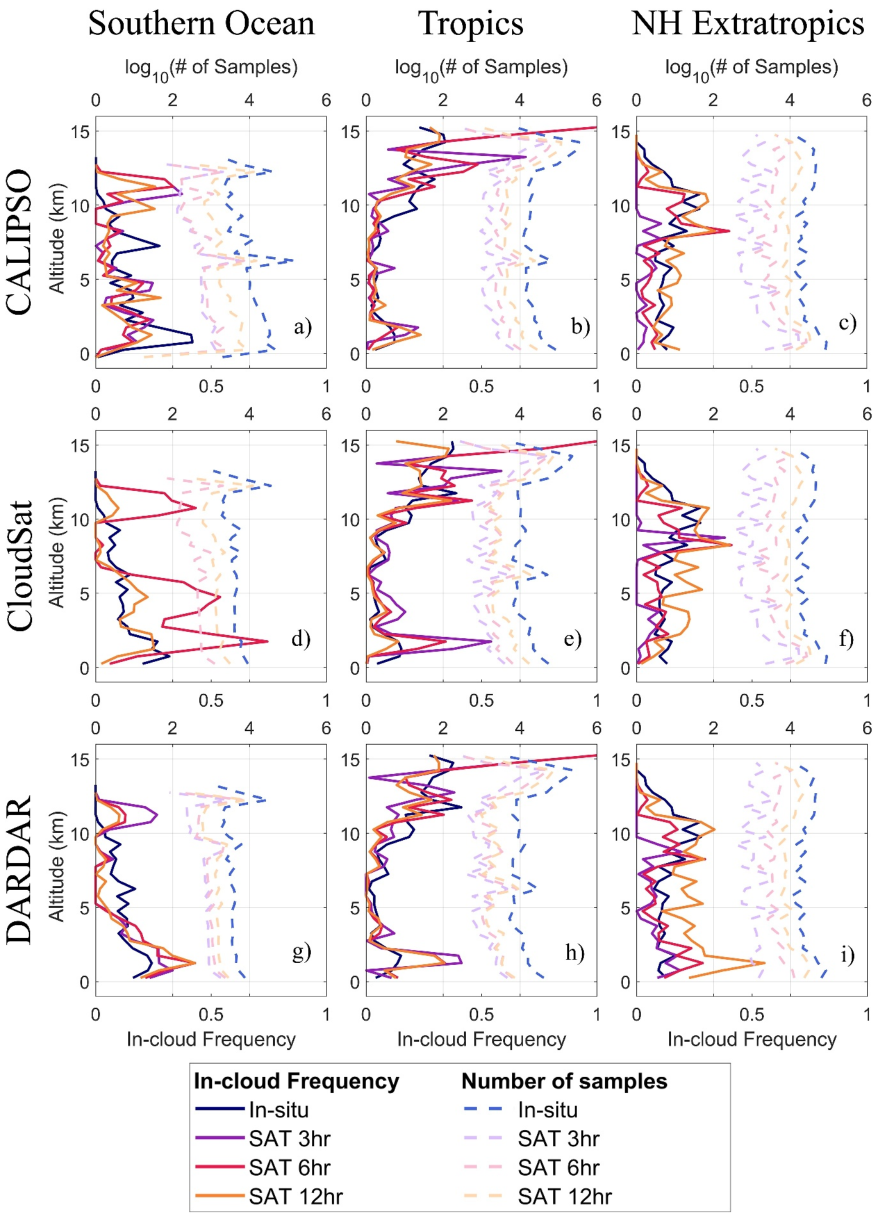

**Figure 8.** Similar to Figure 7, but for in-cloud frequency. Dotted lines denote the $\log_{10}$-scale number of samples of all-sky conditions within different time windows restrictions.

ice phase frequency increases by 0.2–0.5 when including satellite data 12 hr away from the flight data. This result is likely caused by a larger diurnal variation in ice phase frequency in the NH extratropical lower troposphere that is mostly sampled over land, while the other flight tracks in the Southern Ocean and tropics are mostly over open water. Since most of the research flights took place during daytime, within 6 hr of the flights are mostly still in daytime while within 12 hr of the flights could be both daytime and nighttime. Therefore, the night cooling in the lower troposphere over land in the NH extratropics likely leads to the increasing frequency of ice phase inside clouds. DARDAR shows a smaller sensitivity to various temporal windows compared with CALIPSO and CloudSat, possibly because DARDAR already shows positive biases of ice phase frequency and thus the diurnal variation is not as significant.

In-cloud frequencies are compared among three regions in Figure 8. A high-altitudinal cirrus cloud layer is seen in the in situ observations, which is located at 12–15 and 8–11 km in the tropics and NH extratropics, respectively. The higher altitudes of cirrus layer in the tropics are likely due to the higher tropopause height. This is consistent with previous studies that reported high occurrence frequencies of cirrus clouds slightly below the thermal tropopause (e.g., Diao et al., 2015). All satellite data also capture the existence of such cirrus layer. In situ observations from the SOCRATES and ORCAS campaigns mostly targeted low- and mid-level clouds, which causes the underrepresentation of the high-level clouds over the Southern Ocean. In-cloud frequencies in the mid and lower troposphere are around 25%–40% in both in situ and satellite observations. Using different temporal windows shows smaller impacts on in-cloud frequencies in the Southern Ocean and tropics, but higher impacts for NH extratropics, where higher in-cloud frequencies are seen in the 12-hr time window, likely associated with nighttime. The higher in-cloud frequencies seen in Figure 8d for CloudSat data within the 3-hr window are unique features for the ORCAS campaign, since SOCRATES campaign did not overlap with CloudSat data availability and HIPPO campaign only had comparable samples within the 12-hr window, not within the 6-hr window. When a wider spatial domain is investigated (not shown), we found in-cloud frequencies over the Southern Ocean similar to what is shown by the 12-hr window profile.

Correlation analyses between satellite and aircraft data are conducted in Figures S4 and S5 in Supporting Information S1 using the vertical profiles shown in Figures 7 and 8, respectively. Linear regressions are applied, with their information shown in Tables S1 and S2 in Supporting Information S1, respectively. The linear regressions show consistent results as the other analyses, including CALIPSO produces generally the best results compared with in situ observations, and CloudSat and DARDAR overestimates and underestimates ice phase frequencies, respectively. Additional details are also revealed on the specific region that each satellite data set produces the worst linear regression fit, which are NH extratropics, tropics, and Southern Ocean for CALIPSO, CloudSat, and DARDAR, respectively. As for the differences between three temporal windows, no clear trend of an improvement or worsening of comparison results is seen with the increasing window sizes, possibly due to the combined effects of number of samples available and proximity of the collocation restriction.

### 4.5. Latitudinal, Longitudinal, and Seasonal Variability of Ice Phase Frequency

Three latitudinal ranges (i.e., below 30°S, 30°S–30°N, and above 30°N) are investigated for vertical profiles of cloud phase frequencies (Figure 9). The overestimations of mixed phase in CloudSat and the overestimations of ice phase in DARDAR are consistent with those seen in Figure 3. The comparison results on the thickness of mixing region are also consistent with those seen in Figures 4 and 5, with CALIPSO showing the most similar transition from liquid to ice to in situ observations. CloudSat shows a too gradual transition as clouds shift from mixing region (i.e., ice frequency between 0.1 and 0.9) to ice dominated region (i.e., ice frequency above 0.9), while DARDAR shows a too rapid transition. Regional variations in the three latitudinal bands are consistent with those seen among three regions described in Figure 7, including variations in height and thickness of the mixing region. The SH extratropics have higher mixed-phase frequencies compared with the NH extratropics, as illustrated by in situ observations, CloudSat and DARDAR.

We further compare satellite observations in various longitudinal ranges (Figure 10) and four seasons (Figure 11). Since larger sample sizes are needed for these variability analysis, 1-year samples from three satellite products with near-global coverage are examined without applying collocation criteria to in situ observations. In situ observations are not involved in this analysis since the average of all in situ observations are shown in Figure 5. Three longitudinal bands are examined: West Pacific (120°E−180°E), East Pacific (120°W–180°W), and Americas (20°W–120°W). The global distributions of ice phase frequencies of all three satellites (Figures 10a–10c) show similar features to those shown in Figures 4 and 5. Among three satellite data, CALIPSO (Figure 10a) still shows the most similar mixing region height and thickness compared with in situ observations (Figure 5a). In addition, the main biases in the satellite observations are similar to those shown in Figures 4 and 5, including the overestimation (underestimation) of ice phase frequencies in DARDAR (CloudSat), as well as the overestimation (underestimation) of the mixing region thickness in CloudSat (DARDAR). This further corroborates the statistical robustness of the validation results based on consistent findings using three sample sizes shown in Figures 4, 5, and 10.

In terms of seasonal variability (Figure 11), CALIPSO and CloudSat show similar variabilities in ice phase frequencies in four seasons: December, January, and February (DJF); March, April, and May (MAM); June, July, and August (JJA); and September, October, and November (SON). That is, when NH is in boreal winter (DJF) or when SH is in austral winter (JJA), the winter hemisphere shows higher ice phase frequencies and lower altitudes of IceF lines in the extratropics compared with other seasons for the same latitudinal/altitudinal bin. Such

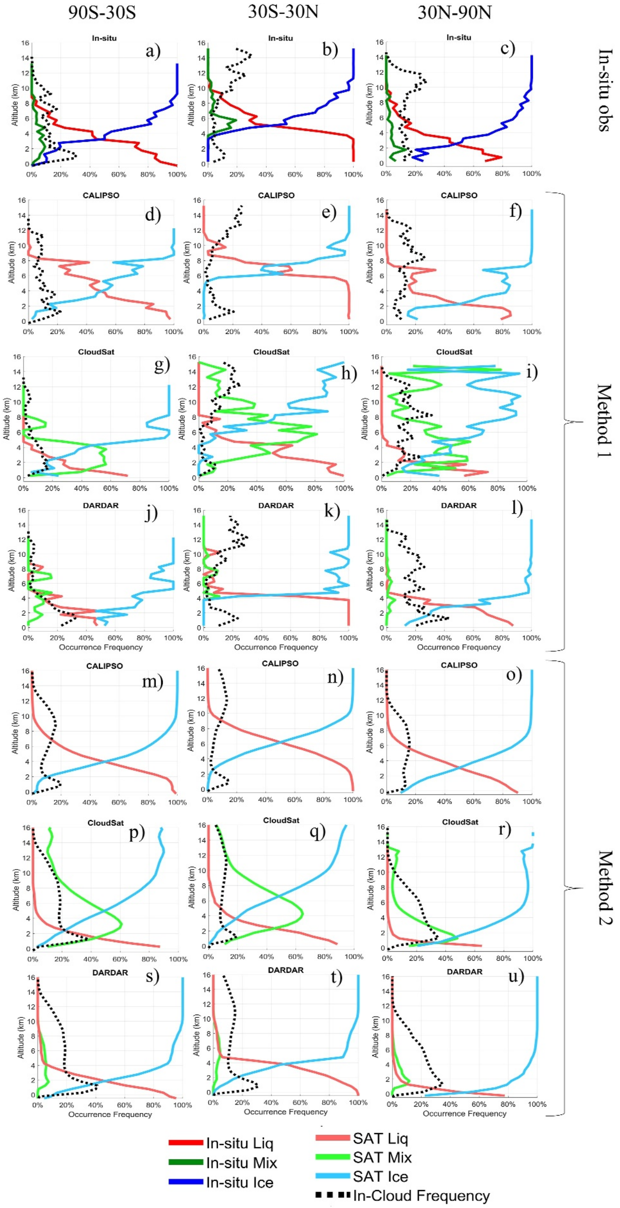

**Figure 9.** Latitudinal variations in cloud phase frequency distributions. Similar to Figure 3 but separated into three latitudinal bands. (a–c) In situ observations. Satellite data selected by (d–l) Method 1 and (m–u) Method 2.

seasonal variability is shown most clearly in the extratropics by CALIPSO and CloudSat, while the tropics have relatively constant ice phase frequencies among four seasons. DARDAR shows much smaller variations among the four seasons, with only slightly higher altitudes of IceF$_{90}$ line in the winter than summer for both hemispheres. This lack of seasonal variability in DARDAR is likely caused by its overestimation of ice phase in general, which is consistent with its smaller sensitivity to variations of temporal windows shown in Figure 7i. Overall, the results above indicate that longitudinal variations of ice phase frequencies are smaller than the latitudinal variations.

We further examined the general variabilities of the vertical profiles of cloud phase occurrence frequencies and in-cloud frequencies in different latitudinal, longitudinal bands and four seasons (Figures S6 and S7 in Supporting Information S1). A large data set with a near-global coverage that is solely based on satellite data for the entire year of 2010 is used. This data set is identical to that used in Figures 10 and 11 (i.e., satellite data only).

For both cloud phase and in-cloud frequencies, the latitudinal variability is larger than the longitudinal or seasonal variability. The differences between the two hemispheres are smaller than the differences between the lower and higher latitudes. The tropics have the lowest ice phase frequency at the same altitude compared with mid- and high-latitudes, since temperature is higher at the same altitude in the tropics. The longitudinal and seasonal analyses show less than 10% of differences in cloud phase frequencies at the same altitude.

In-cloud frequencies in the tropics have lower values than the mid- and high-latitudes except for the upper troposphere (10–15 km). The in-cloud frequency shows similar results in various longitudes and seasons as well, but with larger variations near surface (<1 km) and in the upper troposphere (12–15 km). Overall, these locations where variations of cloud phase and in-cloud frequencies are large are not the transitional altitudes where liquid and ice phases switch their dominance. Therefore, the main conclusions of this work, which focuses on mixed-cloud regimes in the mid-troposphere, are not largely affected by these variations.

## 5. Discussion and Conclusion

This study aims to validate cloud phase identification of three satellite products against in situ aircraft-based observations in a near-global perspective. To achieve this, a large in situ data set was developed using 11 NSF flight campaigns. Several factors influencing cloud phase frequency distributions are examined. Two methods for selecting collocated samples are contrasted.

Among all three satellite-based cloud phase products, CALIPSO data show the most similar distributions of cloud phase frequencies illustrated by the locations of ice phase frequency lines and the mixing region in latitudinal cross sections compared with in situ observations (Figures 4 and 5). The main biases in CloudSat and DARDAR are the overestimations of mixed phase and ice phase frequencies, respectively. When examining the thickness of the mixing region, CALIPSO slightly underestimates it, CloudSat significantly overestimates it, and DARDAR underestimates it. These results indicate that even though CALIPSO data only use lidar data to derive cloud phase and mixed phase is not provided in CALIPSO cloud phase data, these limitations have a smaller impact on CALIPSO cloud phase frequency distribution, but have a larger impact on its in-cloud frequency especially at the lower altitudes (Figure 8a). Larger discrepancies for representing the IceF$_{50}$ to IceF$_{90}$ transition are seen in all three satellites compared with the Ice$_{10}$ to IceF$_{50}$ transition, indicating the algorithms responsible for differentiating supercooled liquid water from ice particles at lower temperatures still have large uncertainties and should be investigated further.

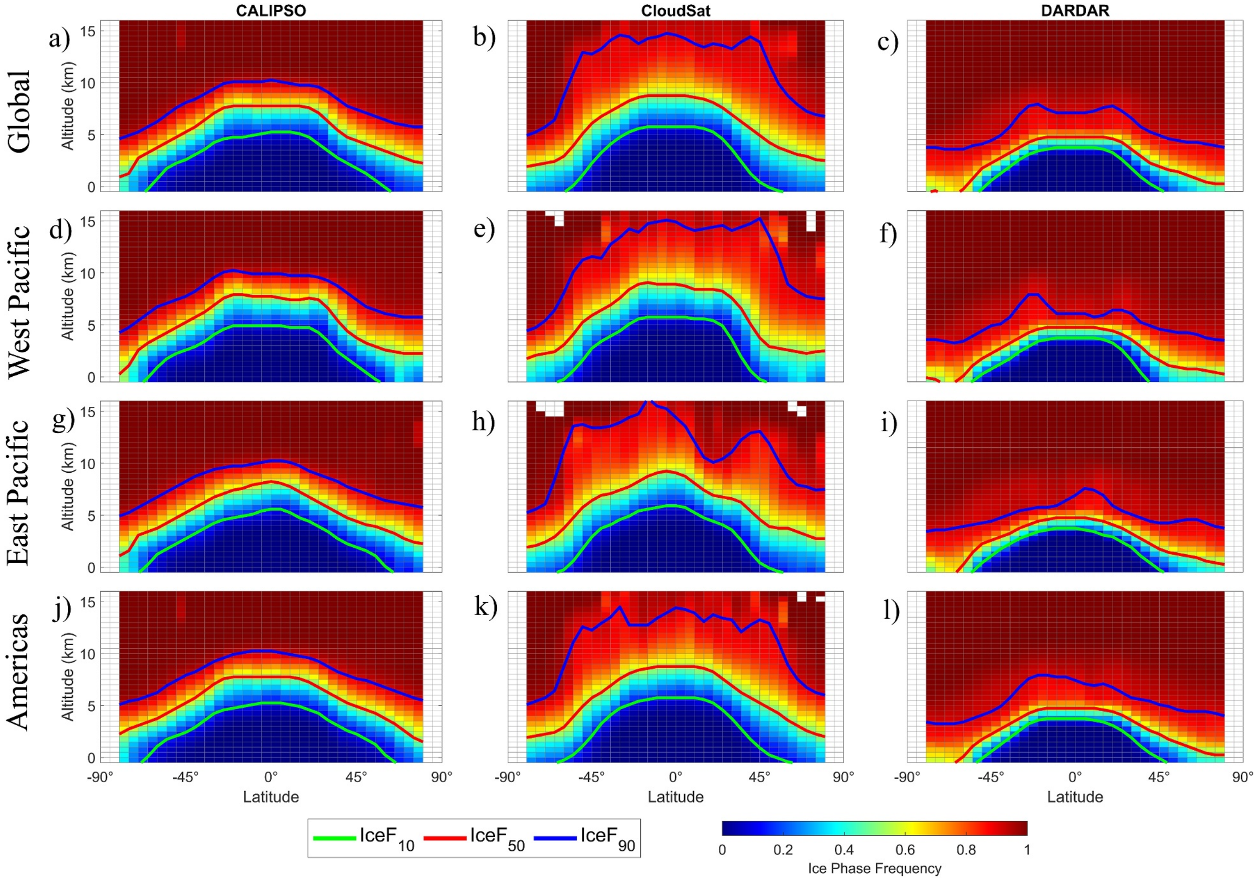

**Figure 10.** Latitudinal-altitudinal view of ice phase frequency for three longitudinal bands: West Pacific, East Pacific, and Americas. Satellite data represent the entire year of 2010. Colored solid lines denote 10% (green), 50% (red), and 90% (blue) of ice phase frequency.

Spatial variabilities of cloud phase distributions are examined across a wide range of latitudes, as well as for specific longitudinal bands and geographical regions. Compared with the latitudinal variations, longitudinal variations are much smaller, indicating a possibility of relaxing the collocation requirements on different longitudes when future comparisons between aircraft and satellite observations are conducted. The regional variations among the Southern Ocean, tropics, and NH extratropics have indicated a more rapid transition from liquid to ice phase in the NH extratropics compared with the Southern Ocean. The longevity of mixed-phase clouds over the Southern Ocean may be attributed to the smaller amount of INPs in that region compared with the NH extratropics (e.g., DeMott et al., 2010; McCluskey et al., 2018), as well as other dynamical influences that likely sustain higher relative humidity in that region (e.g., Maciel & Diao, 2022). The polar regions in the NH and SH also show a slightly higher elevation of the $\text{IceF}_{90}$ line based on in situ observations, which may also indicate insufficient amount of INPs in these high latitudinal regions contributing to delayed formation of ice particles (e.g., Shi et al., 2022). Note that this work does not further examine the relationships of cloud phase distributions with respect to meteorological conditions in specific regions. Previous studies (such as Bodas-Salcedo et al., 2012, 2014, 2016) have demonstrated the large impacts of extratropical cyclones on the existence of supercooled liquid water over the Southern Ocean. Other studies such as G. L. Stephens et al. (2010) also discussed the difficulties of directly comparing gridded climate model output with CloudSat data due to differences in temporal and spatial geometry. Those challenges apply to satellite-aircraft comparisons as well. Another example is Ahn et al. (2018), where the authors limited their time window to 30 min to minimize any cloud evolution arising from the high wind speeds and strong vertical wind shear over the Southern Ocean. But restricting the comparisons to 30 min in this work significantly reduces the number of samples by orders of magnitude to only 1–2 min in only a few vertical bins. Therefore, future studies are recommended to examine the impacts of various ambient conditions (e.g., synoptic conditions, mesoscale thermodynamic and dynamic conditions) on the comparisons between different observing platforms.

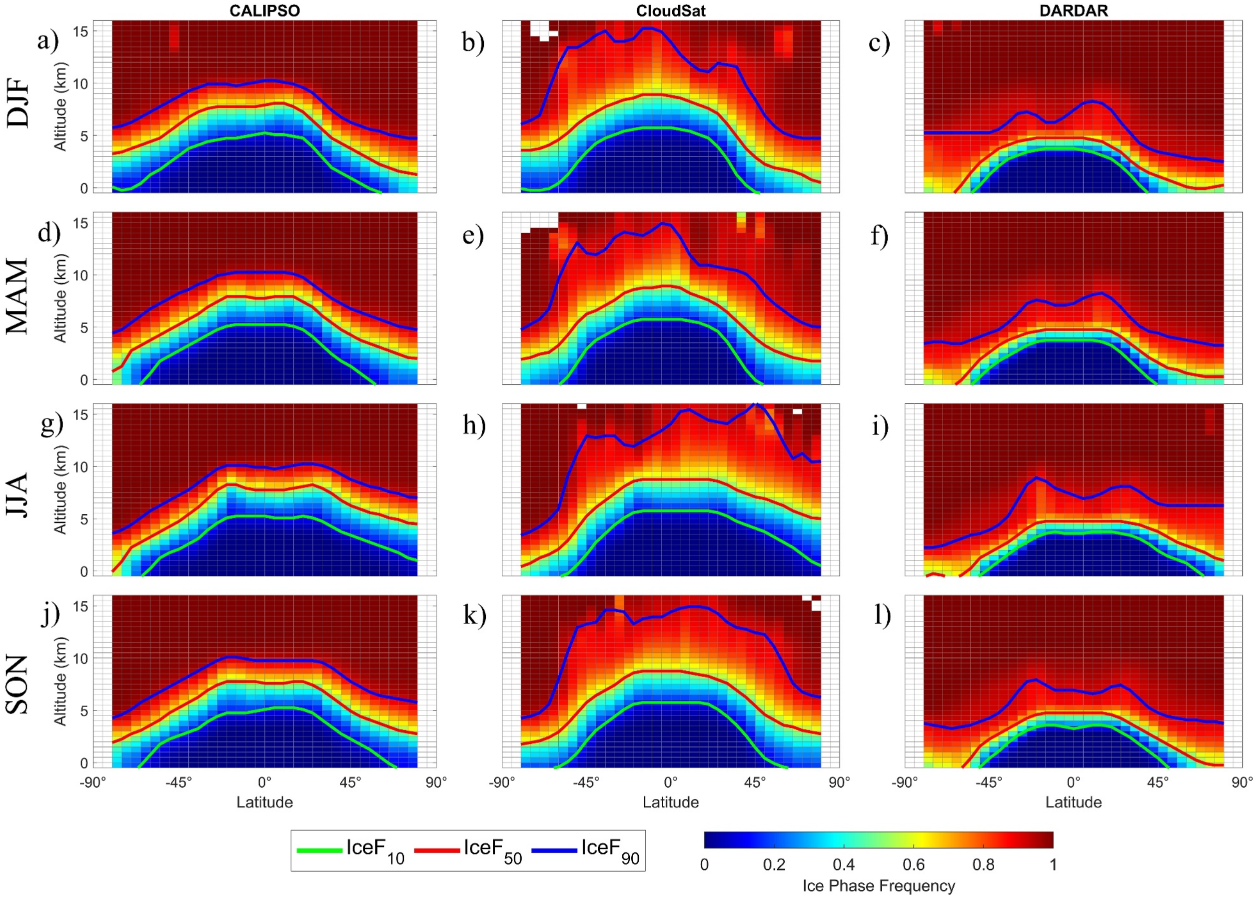

**Figure 11.** Similar to Figure 10 but separated by four seasons in each row using three satellite data sets in 2010.

Three sample sizes of satellite data are contrasted with each other (Figures 4, 5, and 10) and two methods are used to compare satellite data with in situ observations with different spatiotemporal windows of mismatches. All key biases seen in the cloud phase distributions by satellite data are consistently shown using two methods of comparisons, demonstrating the statistical robustness of these validation results. The mismatches of spatiotemporal windows have a larger impact on in-cloud frequencies than cloud phase frequencies, since the former has a stronger dependence on whether a cloud is being sampled or not, while the latter relies more on physical parameters such as thermodynamic, dynamical and aerosol conditions. Recommendations for future campaigns can also be inferred based on the lack of samples in certain locations or time periods. For example, very few cirrus clouds have been sampled in these flight campaigns over the Southern Ocean and Arctic, and few flights took place at nighttime. Locations closer to the North and South Poles are also lacking representations by flight samples. For future flight campaigns, providing more samples in the SH higher altitudes as well as high latitudes in both NH and SH can help examine the hemispheric differences in each cloud type and cloud phase. In addition, providing more nighttime flights can help investigate diurnal cycles of these cloud properties.

The definitions of cloud phase slightly vary between satellite and aircraft-based in situ observations. For the satellite data, cloud phase identification is a pixel-by-pixel identification, basically based on the spatial fraction of ice pixels among all in-cloud pixels. Whereas the in situ observations are based on mass concentrations, basically the mass fraction of ice phase among all cloud phases. An aircraft-based observational study by Maciel and Diao (2022) examined the variability of cloud phase frequencies based on different approaches of defining cloud phases (i.e., using spatial fraction, number fraction, or mass fraction of ice hydrometeors). They found small differences (i.e., less than 0.1 within a 5-degree temperature bin) in the cloud phase frequencies at −40 to 0°C between definitions based on spatial fraction and mass fraction. Thus, the discrepancies of cloud phase frequencies larger than 0.1 at various temperature bins between in situ aircraft and satellite data are more likely caused by biases in the satellite phase identification algorithms rather than different definitions of cloud phases.

The in situ data set developed in this work has other applications besides satellite validation, such as comparisons with ground-based, ship-based observations or global climate model simulations. Sensitivity tests to different factors affecting the in-cloud and cloud phase frequency distributions can be used to assess potential differences between studies using different spatiotemporal windows for comparisons or studies conducted for different geographical locations. With the increasing data volume and spatial coverage of flight campaigns, it is expected that airborne observations will have an increasing capacity for representing cloud properties on a global scale. Future investigations are still warranted to connect remote-sensing and in situ observations to represent cloud properties from microscale to synoptic scale.

## Data Availability Statement

The observational data sets used in this work can be downloaded from the CALIPSO, CloudSat, and DARDAR websites using the following links: https://www-calipso.larc.nasa.gov/, https://www.cloudsat.cira.colostate.edu/data-products/2b-cldclass-lidar, https://www.icare.univ-lille.fr/dardar/documentation-dardar-cloud/, respectively. In situ airborne observations can be accessed at the website hosted by the University Center of Atmospheric Research (UCAR) Earth Observation Laboratory: https://www.eol.ucar.edu/.

**Acknowledgments**

D.W., C.Y. and M.D. would like to acknowledge the funding from the NASA ACCDAM Grant ROSES-2020 80NSSC21K1457, the National Science Foundation Office of Polar Programs (OPP) Grant 1744965 and SJSU Division of Research and Innovation award number 22-LUG-08-006. D.W. and C.Y. received the Walker Fellowship from San Jose State University (SJSU). D.W. also received the Outstanding Thesis Award from SJSU in 2024. We would like to thank the computing resources and support from the National Center for Atmospheric Resesarch University Small Request allocation award USJS0009. We would like to thank the science team that support each of the eleven NSF field campaigns, the scientists, and staff from the UCAR Earth Observing Laboratory/Research Aviation Facility, and the science team of CALIPSO, CloudSat, and DARDAR.

## References

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), 11120–11140. https://doi.org/10.1029/2018JD028535

Albrecht, B., Ghate, V., Mohrmann, J., Wood, R., Zuidema, P., Bretherton, C., et al. (2019). Cloud system evolution in the trades (CSET): Following the evolution of boundary layer cloud systems with the NSF–NCAR GV. *Bulletin of the American Meteorological Society*, *100*(1), 93–121. https://doi.org/10.1175/BAMS-D-17-0180.1

Barker, H. W., Korolev, A. V., Hudak, D. R., Strapp, J. W., Strawbridge, K. B., & Wolde, M. (2008). A comparison between CloudSat and aircraft data for a multilayer, mixed phase cloud system during the Canadian CloudSat-CALIPSO Validation Project. *Journal of Geophysical Research*, *113*(D8). https://doi.org/10.1029/2008JD009971

Barth, M. C., Cantrell, C. A., Brune, W. H., Rutledge, S. A., Crawford, J. H., Huntrieser, H., et al. (2015). The deep convective clouds and Chemistry (DC3) field campaign. *Bulletin of the American Meteorological Society*, *96*(8), 1281–1309. https://doi.org/10.1175/BAMS-D-13-00290.1

Bodas-Salcedo, A., Andrews, T., Karmalkar, A. V., & Ringer, M. A. (2016). Cloud liquid water path and radiative feedbacks over the Southern Ocean. *Geophysical Research Letters*, *43*(20), 10938–10946. https://doi.org/10.1002/2016GL070770

Bodas-Salcedo, A., Williams, K. D., Field, P. R., & Lock, A. P. (2012). The surface downwelling solar radiation surplus over the Southern Ocean in the met office model: The role of midlatitude cyclone clouds. *Journal of Climate*, *25*(21), 7467–7486. https://doi.org/10.1175/JCLI-D-11-00702.1

Bodas-Salcedo, A., Williams, K. D., Ringer, M. A., Beau, I., Cole, J. N. S., Dufresne, J. L., et al. (2014). Origins of the solar radiation biases over the Southern Ocean in CFMIP2 models. *Journal of Climate*, *27*(1), 41–56. https://doi.org/10.1175/JCLI-D-13-00169.1

Brown, P. R. A., & Francis, P. N. (1995). Improved measurements of the ice water content in cirrus using a total-water probe. *Journal of Atmospheric and Oceanic Technology*, *12*(2), 410–414. https://doi.org/10.1175/1520-0426(1995)012<0410:imotiw>2.0.co;2

Ceccaldi, M., Delanoë, J., Hogan, R. J., Pounder, N. L., Protat, A., & Pelon, J. (2013). From CloudSat-CALIPSO to EarthCare: Evolution of the DARDAR cloud classification and its comparison to airborne radar-lidar observations. *Journal of Geophysical Research: Atmospheres*, *118*(14), 7962–7981. https://doi.org/10.1002/JGRD.50579

Cesana, G., Chepfer, H., Winker, D., Getzewich, B., Cai, X., Jourdan, O., et al. (2016). Using in situ airborne measurements to evaluate three cloud phase products derived from CALIPSO. *Journal of Geophysical Research: Atmospheres*, *121*(10), 5788–5808. https://doi.org/10.1002/2015JD024334

Cesana, G., & Storelvmo, T. (2017). Improving climate projections by understanding how cloud phase affects radiation. *Journal of Geophysical Research: Atmospheres*, *122*(8), 4594–4599. https://doi.org/10.1002/2017JD026927

D'Alessandro, J. J., Diao, M., Wu, C., Liu, X., Jensen, J. B., & Stephens, B. B. (2019). Cloud phase and relative humidity distributions over the Southern Ocean in austral summer based on in situ observations and CAM5 simulations. *Journal of Climate*, *32*(10), 2781–2805. https://doi.org/10.1175/JCLI-D-18-0232.1

Delanoë, J., & Hogan, R. J. (2008a). A variational scheme for retrieving ice cloud properties from combined radar, lidar, and infrared radiometer. *Journal of Geophysical Research*, *113*(D7), D07204. https://doi.org/10.1029/2007JD009000

Delanoë, J., & Hogan, R. J. (2008b). DARDAR-CLOUD Document. Retrieved from https://www.icare.univ-lille.fr/asd-content/projects_tools/dardar/docs/varcloud-algorithm_description-v1.0.pdf

DeMott, P. J., Prenni, A. J., Liu, X., Kreidenweis, S. M., Petters, M. D., Twohy, C. H., et al. (2010). Predicting global atmospheric ice nuclei distributions and their impacts on climate. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(25), 11217–11222. https://doi.org/10.1073/pnas.0910818107

Deng, M., Mace, G. G., Wang, Z., & Okamoto, H. (2010). Tropical composition, cloud and climate coupling experiment validation for cirrus cloud profiling retrieval using CloudSat radar and CALIPSO lidar. *Journal of Geophysical Research*, *115*(17). https://doi.org/10.1029/2009JD013104

Desai, N., Diao, M., Shi, Y., Liu, X., & Silber, I. (2023). Ship-based observations and climate model simulations of cloud phase over the Southern Ocean. *Journal of Geophysical Research: Atmospheres*, *128*(11), e2023JD038581. https://doi.org/10.1029/2023JD038581

Diao, M., Jensen, J. B., Pan, L. L., Homeyer, C. R., Honomichl, S., Bresch, J. F., & Bansemer, A. (2015). Distributions of ice supersaturation and ice crystals from airborne observations in relation to upper tropospheric dynamical boundaries. *Journal of Geophysical Research: Atmospheres*, *120*(10), 5101–5121. https://doi.org/10.1002/2015JD023139

Diao, M., Jumbam, L., Sheffield, J., Wood, E. F., & Zondlo, M. A. (2013). Validation of AIRS/AMSU-A water vapor and temperature data with in situ aircraft observations from the surface to UT/LS from 87°N–67°S. *Journal of Geophysical Research: Atmospheres*, *118*(12), 6816–6836. https://doi.org/10.1002/jgrd.50483

Flynn, C. M., & Mauritsen, T. (2020). On the climate sensitivity and historical warming evolution in recent coupled model ensembles. *Atmospheric Chemistry and Physics*, *20*(13), 7829–7842. https://doi.org/10.5194/ACP-20-7829-2020

Frey, W. R., & Kay, J. E. (2018). The influence of extratropical cloud phase and amount feedbacks on climate sensitivity. *Climate Dynamics*, *50*(7–8), 3097–3116. https://doi.org/10.1007/s00382-017-3796-5

Fuchs-Stone, Ž., Raymond, D. J., & Sentić, S. (2020). OTREC2019: Convection over the East Pacific and southwest Caribbean. *Geophysical Research Letters*, *47*(11), e2020GL087564. https://doi.org/10.1029/2020GL087564

Hu, Y., Winker, D., Vaughan, M., Lin, B., Omar, A., Trepte, C., et al. (2009). CALIPSO/CALIOP cloud phase discrimination algorithm. *Journal of Atmospheric and Oceanic Technology*, *26*(11), 2293–2309. https://doi.org/10.1175/2009JTECHA1280.1

Hunt, W. H., Vaughan, M. A., Powell, K. A., Weimer, C., & Lucker, P. L. (2009). CALIPSO lidar description and performance assessment. *Journal of Atmospheric and Oceanic Technology*, *26*(7), 1214–1228. https://doi.org/10.1175/2009JTECHA1223.1

Im, E., Wu, C., & Durden, S. L. (2005). Cloud profiling radar for the CloudSat mission. *IEEE Aerospace and Electronic Systems Magazine*, *20*(10), 15–18. https://doi.org/10.1109/MAES.2005.1581095

Kay, J. E., Bourdages, L., Miller, N. B., Morrison, A., Yettella, V., Chepfer, H., & Eaton, B. (2016). Evaluating and improving cloud phase in the Community Atmosphere Model version 5 using spaceborne lidar observations. *Journal of Geophysical Research: Atmospheres*, *121*(8), 4162–4176. https://doi.org/10.1002/2015JD024699

Korolev, A., Emery, E., & Creelman, K. (2013). Modification and tests of particle probe tips to mitigate effects of ice shattering. *Journal of Atmospheric and Oceanic Technology*, *30*(4), 690–708. https://doi.org/10.1175/JTECH-D-12-00142.1

Liou, K.-N. (1992). *Radiation and cloud processes in the atmosphere. Theory, observation, and modeling*. Oxford University Press.

Maciel, F. V., & Diao, M. (2022). The transition from supercooled liquid water to ice crystals in mixed-phase clouds based on airborne in-situ observations. *Atmospheric Measurement Techniques Discussions*. [preprint]. in review. https://doi.org/10.5194/amt-2022-256

Matus, A. V., & L'Ecuyer, T. S. (2017). The role of cloud phase in Earth's radiation budget. *Journal of Geophysical Research: Atmospheres*, *122*(5), 2559–2578. https://doi.org/10.1002/2016JD025951

McCluskey, C. S., Ovadnevaite, J., Rinaldi, M., Atkinson, J., Belosi, F., Ceburnis, D., et al. (2018). Marine and terrestrial organic ice-nucleating particles in pristine marine to continentally influenced Northeast Atlantic air masses. *Journal of Geophysical Research: Atmospheres*, *123*(11), 6196–6212. https://doi.org/10.1029/2017JD028033

McCoy, D. T., Hartmann, D. L., & Grosvenor, D. P. (2014a). Observed Southern Ocean cloud properties and shortwave reflection. Part I: Calculation of SW flux from observed cloud properties. *Journal of Climate*, *27*(23), 8836–8857. https://doi.org/10.1175/JCLI-D-14-00287.1

McCoy, D. T., Hartmann, D. L., & Grosvenor, D. P. (2014b). Observed Southern Ocean cloud properties and shortwave reflection. Part II: Phase changes and low cloud feedback. *Journal of Climate*, *27*(23), 8858–8868. https://doi.org/10.1175/JCLI-D-14-00288.1

McFarquhar, G. M., Bretherton, C. S., Marchand, R., Protat, A., DeMott, P. J., Alexander, S. P., et al. (2021). Observations of clouds, aerosols, precipitation, and surface radiation over the southern ocean. *Bulletin of the American Meteorological Society*, *102*(4), E894–E928. https://doi.org/10.1175/BAMS-D-20-0132.1

Mitchell, J. F. B., Senior, C. A., & Ingram, W. J. (1989). $CO_2$ and climate: A missing feedback? *Nature*, *341*(6238), 132–134. https://doi.org/10.1038/341132a0

Montgomery, M. T., Davis, C., Dunkerton, T., Wang, Z., Velden, C., Torn, R., et al. (2012). The pre-depression investigation of cloud-systems in the tropics (PREDICT) experiment: Scientific basis, new analysis tools, and some first results. *Bulletin of the American Meteorological Society*, *93*(2), 153–172. https://doi.org/10.1175/BAMS-D-11-00046.1

Pan, L. L., Atlas, E. L., Salawitch, R. J., Honomichl, S. B., Bresch, J. F., Randel, W. J., et al. (2017). The convective Transport of active species in the tropics (CONTRAST) experiment. *Bulletin of the American Meteorological Society*, *98*(1), 106–128. https://doi.org/10.1175/BAMS-D-14-00272.1

Pan, L. L., Bowman, K. P., Atlas, E. L., Wofsy, S. C., Zhang, F., Bresch, J. F., et al. (2010). The stratosphere–troposphere analyses of regional transport 2008 experiment. *Bulletin of the American Meteorological Society*, *91*(3), 327–342. https://doi.org/10.1175/2009BAMS2865.1

Sassen, K. (1991). The polarization lidar technique for cloud research: A review and current assessment. *Bulletin of the American Meteorological Society*, *72*(12), 1848–1866. https://doi.org/10.1175/1520-0477(1991)072<1848:TPLTFC>2.0.CO;2

Shi, Y., Liu, X., Wu, M., Zhao, X., Ke, Z., & Brown, H. (2022). Relative importance of high-latitude local and long-range-transported dust for Arctic ice-nucleating particles and impacts on Arctic mixed-phase clouds. *Atmospheric Chemistry and Physics*, *22*(4), 2909–2935. https://doi.org/10.5194/acp-22-2909-2022

Stephens, B. B., Long, M. C., Keeling, R. F., Kort, E. A., Sweeney, C., Apel, E. C., et al. (2018). The $O_2/N_2$ ratio and $CO_2$ airborne Southern Ocean study. *Bulletin of the American Meteorological Society*, *99*(2), 381–402. https://doi.org/10.1175/BAMS-D-16-0206.1

Stephens, G. L., L'Ecuyer, T., Forbes, R., Gettelman, A., Golaz, J.-C., Bodas-Salcedo, A., et al. (2010). Dreary state of precipitation in global models. *Journal of Geophysical Research*, *115*(D24), D24211. https://doi.org/10.1029/2010JD014532

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., et al. (2002). The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation. *Bulletin of the American Meteorological Society*, *83*(12), 1771–1790. https://doi.org/10.1175/bams-83-12-1771

Tan, I., & Storelvmo, T. (2016). Sensitivity study on the influence of cloud microphysical parameters on mixed-phase cloud thermodynamic phase partitioning in CAM5. *Journal of the Atmospheric Sciences*, *73*(2), 709–728. https://doi.org/10.1175/JAS-D-15-0152.1

Tan, I., Storelvmo, T., & Zelinka, M. D. (2016). Observational constraints on mixed-phase clouds imply higher climate sensitivity. *Science*, *352*(6282), 224–227. https://doi.org/10.1126/science.aad5300

Terai, C. R., Klein, S. A., & Zelinka, M. D. (2016). Constraining the low-cloud optical depth feedback at middle and high latitudes using satellite observations. *Journal of Geophysical Research: Atmospheres*, *121*(16), 9696–9716. https://doi.org/10.1002/2016JD025233

Thornton, J. A., Brown, S. S., Jaeglè, L., Cohen, R., Jimenez, J., Weber, R., & Roberts, J. M. (2014). Winter investigation of transport, emissions, and reactivity: An NSF funded aircraft and ground-based investigation. Retrieved from https://csl.noaa.gov/groups/csl7/measurements/2015winter/whitepaper.pdf

Vaughan, M., Pitts, M., Trepte, C., Winker, D., Detweiler, P., Garnier, A., et al. (2019). Cloud-aerosol LIDAR infrared pathfinder satellite observations data management system data products catalog release 4.50 document no: PC-SCI-503. Retrieved from https://www-calipso.larc.nasa.gov/products/CALIPSO_DPC_Rev4x50.pdf

Vaughan, M. A., Young, S. A., Winker, D. M., Powell, K. A., Omar, A. H., Liu, Z., et al. (2004). Fully automated analysis of space-based lidar data: An overview of the CALIPSO retrieval algorithms and data products. In *Laser radar techniques for atmospheric sensing* (Vol. 5575). https://doi.org/10.1117/12.572024

Volkamer, R., Baidar, S., Campos, T. L., Coburn, S., DiGangi, J. P., Dix, B., et al. (2015). Aircraft measurements of BrO, IO, glyoxal, $NO_2$, $H_2O$, $O_2$-$O_2$ and aerosol extinction profiles in the tropics: Comparison with aircraft-/ship-based in situ and lidar measurements. *Atmospheric Measurement Techniques*, *8*(5), 2121–2148. https://doi.org/10.5194/AMT-8-2121-2015

Wang, Z. (2019). CloudSat 2B-CLDCLASS-LIDAR product process description and interface control document. Retrieved from https://www.cloudsat.cira.colostate.edu/cloudsat-static/info/dl/2b-cldclass-lidar/2B-CLDCLASS-LIDAR_PDICD.P1_R05.rev0_.pdf

Wang, Z., & Sassen, K. (2001). Cloud type and macrophysical property retrieval using multiple remote sensors. *Journal of Applied Meteorology*, *40*(10), 1665–1682. https://doi.org/10.1175/1520-0450(2001)040<1665:CTAMPR>2.0.CO;2

Winker, D. M., Hunt, W. H., & McGill, M. J. (2007). Initial performance assessment of CALIOP. *Geophysical Research Letters*, *34*(19). https://doi.org/10.1029/2007GL030135

Winker, D. M., Pelon, J., Coakley, J. A., Ackerman, S. A., Charlson, R. J., Colarco, P. R., et al. (2010). The CALIPSO mission: A global 3D view of aerosols and clouds. *Bulletin of the American Meteorological Society*, *91*(9), 1211–1230. https://doi.org/10.1175/2010BAMS3009.1

Winker, D. M., Pelon, J. R., & McCormick, M. P. (2003). The CALIPSO mission: Spaceborne lidar for observation of aerosols and clouds. In *Lidar remote sensing for industry and environment monitoring III* (Vol. 4893). https://doi.org/10.1117/12.466539

Wofsy, S. C. (2011). HIAPER Pole-to-Pole Observations (HIPPO): Fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, *369*(1943), 2073–2086. https://doi.org/10.1098/rsta.2010.0313

Yang, C. A., Diao, M., Gettelman, A., Zhang, K., Sun, J., McFarquhar, G., & Wu, W. (2021). Ice and supercooled liquid water distributions over the Southern Ocean based on in situ observations and climate model simulations. *Journal of Geophysical Research: Atmospheres*, *126*(24). https://doi.org/10.1029/2021JD036045

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., et al. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, *47*(1), e2019GL085782. https://doi.org/10.1029/2019GL085782

Zhang, Y., Klein, S. A., Boyle, J., & Mace, G. G. (2010). Evaluation of tropical cloud and precipitation statistics of Community Atmosphere Model version 3 using CloudSat and CALIPSO data. *Journal of Geophysical Research*, *115*(12). https://doi.org/10.1029/2009JD012006