Assessment and Error Analysis of Terra-MODIS and MISR Cloud-top Heights through Comparison with ISS-CATS lidar

Arka Mitra[1], Larry Di Girolamo[1], Yulan Hong[1], Yizhe Zhan[1,2] and Kevin J. Mueller[3]

[1].University of Illinois, Urbana-Champaign, USA

[2].Metservice Ltd., Wellington, New Zealand

[3] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

Corresponding author: Arka Mitra (mitraarka27@gmail.com)

Key Points:

- We present the first quasi-global (50°N to 50°S) comparison of Terra cloud-top heights with coincident samples from a space-based lidar.
- Using lidar as truth, Terra cloud-top height bias and precision are summarized as a function of cloud geometrical and optical properties.
- With the first measurement of stereo-opacity bias (-110 to -150 m, depending on cloud type), MISR cloud height error-budget is closed.

Abstract

Cloud-top heights (CTH) from the Multiangle Imaging Spectroradiometer (MISR) and the Moderate Resolution Imaging Spectroradiometer (MODIS) on Terra constitute our longest-running single-platform CTH record from a stable orbit. Here, we provide the first evaluation of the Terra Level 2 CTH record against collocated International Space Station Cloud-Aerosol Transport System (CATS) lidar observations between 50ºN - 50ºS. Bias and precision of Terra CTH relative to CATS is shown to be strongly tied to cloud horizontal and vertical heterogeneity and altitude. For single-layered, unbroken, optically thick clouds observed over all altitudes, the uncertainty in MODIS and MISR CTH are -540±690 m and -280±370 m, respectively. The uncertainties are generally smaller for lower altitude clouds and larger for optically thinner clouds. For multi-layered clouds, errors are summarized herein using both absolute CTH and CATS-layer-altitude proximity to Terra CTH. We show that MISR detects the lower cloud in a two-layered system, provided top-layer optical depth < ~0.3, but MISR low-cloud CTH errors are unaltered by the presence of thin cirrus. Systematic and random errors are propagated to explain inter-sensor disagreements, as well as to provide the first estimate of the MISR stereo-opacity bias. For MISR, altitude-dependent wind-retrieval bias (-90 to -110 m) and stereo-opacity bias (-110 to -150 m) and for MODIS, CO_2-slicing bias due to geometrically thick cirrus leads to overall negative CTH bias. MISR's precision is largely driven by precision in retrieved wind-speed (3.7 m s^{-1}), whereas MODIS precision is driven by forward-modeling uncertainty.

Plain Language Summary

Cloud-top height (CTH) is an essential climate variable that impacts the Earth's energy budget and hydrological cycle. We are greatly interested in CTHs for their possible application in detecting signatures of forced climate change in the nearly two-decade long CTH record from NASA's enduring mission, Terra. Since Terra has offered longevity and orbital stability, the remaining criterion for a successful climate dataset is an in-depth understanding and quantification of uncertainty in the data. To ascertain the accuracy of Terra sensors (a multi-view instrument, MISR & a multi-wavelength instrument, MODIS), we compare a subset of their observations against a lidar called CATS that operated from the International Space Station from 2015 to 2017. Through involved statistical analysis, we determined that both MISR and MODIS have provided us with robust CTHs, with MISR being about twice as accurate and precise as MODIS. We note that the MISR error budget is self-contained and that we were able to close the error budget. Each

instrument demonstrates strengths and weaknesses depending on the types of clouds being observed. This study has provided needed CTH error characteristics that can help inform future satellite architecture for observing CTH.

1 Introduction

Cloud altitude feedback is an important component of cloud feedbacks (Zelinka et al., 2017), with inter-model differences in cloud feedbacks being the largest source of uncertainty in climate predictions (Boucher, et al., 2013; Dufresne & Bony, 2008). One of the suggested techniques to lower these inter-model differences is to compare short-term model predictions with accurate global trends in cloud vertical distribution from stable satellite-based sensors. However, short-term trends in cloud-top height (CTH) are often quite small in magnitude and dominated by natural variability in the ocean-atmosphere system (Davies et al., 2017; Geiss & Marchand, 2019). Ohring et al., (2005) recommended a CTH accuracy of 150 m and a stability of 30 m/decade from a satellite sensor for monitoring decadal changes in CTH. Accurate CTH is also necessary in other meteorological research, such as in predicting vertical variations of freezing layers (Van Diedenhoven et al., 2016). As a result, it is imperative that the error characteristics of public standard CTH products be fully established and understood.

CTH retrievals are broadly classified as active or passive. Popular passive CTH retrieval techniques include CO_2-slicing (Menzel et al., 1983), 11-μm brightness temperature (Menzel et al., 2008) and simultaneous retrieval of CTH and winds through stereo photogrammetry (Muller et al., 2002). These passive techniques rely on a single-layered cloud assumption for a given field-of-view – an assumption that is valid only ~75-80% of the time over the globe (Stubenrauch et al., 2013). Active sensors (radars and lidars) can provide detailed hydrometeor vertical distributions, unlike passive sensors. As a result, many previous studies (Marchand et al., 2007; Naud et al., 2002; 2004; 2007) have employed active sensor CTH as the truth to quantify passive sensor CTH errors. Those studies arrived at the consensus that multi-layered clouds can lead to large differences in retrieved CTH amongst passive sensors. In those studies, and the present one, two imagers onboard the Terra satellite – the Multiangle Imaging Spectroradiometer (MISR) and the Moderate Resolution Imaging Spectroradiometer (MODIS) – were analyzed. Terra has provided us with a consistent equator crossing time or ECT (Stubenrauch et al., 2013; Zhao et al., 2016) for nearly two decades and is our longest-running stable climate record of CTH. Hence, understanding the CTH error characteristics from these two instruments is essential for interpreting CTH

variability within their records, to shed light on their strengths and weaknesses, to combine their strengths for improved CTH characterization, and to better inform future satellite missions on system design choices for reducing uncertainty in CTH retrievals.

Collection 5 1-km and 5-km resolution CTH from the MODIS instrument on-board Aqua were compared against near-coincident CALIPSO CTH globally for two months of 2006/2007 by Holz et al. (2008); made possible because both these instruments are part of NASA's A-Train constellation of satellites (1:30 pm ECT). Holz et al. (2008) reported globally averaged CTH differences between the 1-km MODIS and CALIPSO CTH to be -1.4 ± 2.9 km (the 5-km product exhibited worse accuracy and precision due to poorer resolution). Through a detailed analysis, the high negative bias of CTH was found to be largely due to the presence of optically thin high clouds (often, in multi-layered situations), and a failure of the-then CO_2-slicing algorithm to converge to a solution in many high cloud scenes. Random errors, meanwhile, were attributed to incorrect lapse-rates for marine low-level clouds and application of the brightness temperature technique for high clouds (Section 4.5.2 provides an in-depth discussion of these errors). The last two issues were specifically addressed in a series of improvements (Baum et al., 2012) that resulted in the latest Collection 6 MODIS 1-km CTH product. Comparisons of non-polar Aqua MODIS Collection 6 CTH against CALIPSO CTH showed higher deployment of CO_2-slicing than in Collection 5 for single-layered cirrus (less miscasting of high clouds as low clouds) and a reduction of the low-cloud positive bias from 424 m in Collection 5 to 197 m in Collection 6 (Baum et al., 2012). While ostensibly the same, Terra MODIS and Aqua MODIS are subject to key differences for CTH determination that stem from diurnal variability of lapse rates and cloud characteristics between morning and afternoon (Eastman & Warren, 2014), as well as the absence of Band 34 (13.6 μm) on Terra due to high noise. As such an independent study of Terra MODIS is also necessary, as well as for validating the Collection 6 CTH product.

MISR employs a stereoscopic technique for determining CTH and cloud-top advection or "winds", simultaneously (Mueller et al., 2013; Muller et al., 2002). The original CTH product, referenced to as TC_STEREO, often reported an uncertainty of 562 m, but this was specific to the error made if stereo correspondence was off by a single pixel (Moroney et al. 2002). Validation against ground-based radars and lidars showed that CTH uncertainties tended to be less than 1 km, irrespective of cloud height (Marchand et al., 2007; Naud et al., 2004). These studies showed that, when an optically thin upper cloud overlies an optically thick lower cloud (which is often the case

in multi-layered situations), MISR returns the CTH of the lower cloud, provided the upper cloud optical depth is less than 0.3-0.5, depending on surface type (Marchand et al., 2007). This is because lower cloud layers often provide the greatest observed spatial contrast in MISR's visible to near-IR images, even in the presence of thin upper clouds. TC_STEREO also produced many no-retrievals, in part due to overly strict quality control. More recently, the MISR algorithm underwent a series of improvements (Horváth, 2013) to generate the latest stereo product, called TC_CLOUD (Mueller et al., 2013). Although direct validation of TC_CLOUD CTH against active sensors has not yet been done, Horváth (2013) and Mueller et al. (2017) compared MISR winds with geostationary IR atmospheric motion-vectors (AMVs) from Meteosat-9 and GOES, respectively, revealing a pattern of mean and root mean-squared (RMS) differences between MISR and geostationary wind heights that vary with altitude and location. (Section 4.5.1 provides an in-depth discussion of these errors). Averaged globally, wind-related CTH bias relative to IR AMV heights were found to be ~ -200 m, with associated precision ranging from 0.5-1 km, depending on the dataset. The large deviation in the random error estimates can be attributed to the inherent uncertainties of the IR AMV retrievals; however, a better estimation require precise cloud height measurements, such as from a lidar.

The lack of a space-based active sensor with sufficient orbital overlap with Terra has so far impeded a global validation of MISR and Terra-MODIS CTH. To realize our goal of validating Level 2 Terra CTH, the database of 'true' active-sensor CTH came from the ISS-CATS (Yorks et al., 2016). ISS-CATS or simply, CATS (Cloud- Aerosol Transport System) was a space-based lidar that operated from the Japanese Experiment Module-Exposed Facility of the International Space Station (ISS) between 2015-2017. Although too short-lived to be a climate record, CATS was uniquely suited for a quasi-global validation of CTH from Terra-based sensors. Here we use the CATS dataset to examine the error characteristics of MISR and MODIS-Terra CTHs.

Section 2 briefly describes the instruments, their orbits, and the data sources. Section 3 elucidates the collocation among CATS, MISR and MODIS pixels and quantifies the random errors within our methods. Section 4 delves into CTH differences from the inter-comparison of the three instruments, the global distribution of these differences, and the chief reasons behind the disagreements. Concluding remarks follow in Section 5.

2 Data and Instruments

The flagship of NASA's Earth Observing System (EOS), Terra, is a near-polar, sun-synchronous satellite orbiting the Earth at a nominal altitude of 705 km above the surface, making its equator overpasses at 10:30 am local time. MISR and MODIS are two instruments on Terra that use two completely independent techniques for retrieving CTH. MISR employs a stereoscopic technique using 0.67 μm ("Red" channel), 275 m resolution radiance from the three least oblique angles (nadir and ±26.1°) to estimate CTH (Muller et al., 2002). One advantage of a stereoscopic technique over other passive CTH retrievals is that a stereo CTH is not sensitive to radiometric calibration (Naud et al., 2002). The operational MISR algorithm first estimates cloud-top winds, and then stereo heights for each 1.1 km pixel in a scene. The MISR data used here is the Level 2 TC_CLOUD Version F01_0001 orbit-level product, which provides a 1.1 km "wind-corrected" CTH over a swath of width 380 km.

MODIS is a broad-swath (swath width ~2330 km) imager with 36 spectral channels that has a nadir spatial resolution ranging from 250 to 1000 m, depending on the spectral channel. The MISR swath lies completely within the MODIS swath. MODIS employs a CO_2-slicing technique (Menzel et al., 2008) for CTH estimation, designed to calculate the cloud top pressure (CTP) and effective cloud amount for geometrically thin, single-layered mid-level and high clouds. These quantities are derived from ratios of differences between cloudy and clear-sky radiances from any of the following pairs: 14.2 μm/13.9 μm, 13.9 μm/13.6 μm, 13.9 μm/13.3 μm or 13.6 μm/13.3 μm, with MODIS CTP reporting the solution of the highest wavelength band-pair whose radiance difference exceeds instrument noise in the individual bands. It is assumed that cloud emissivity is equal for both wavelengths in the pair, an assumption better suited for ice clouds than water clouds. CTP retrieval occurs at 1-km resolution, provided that at least 4 out of the 25 pixels in a 5x5 pixel window surrounding it were flagged as either cloudy or probably cloudy by the MODIS cloud mask and an independent pixel-level phase detection flagged ice. CTP is converted to CTH using Global Data Assimilation System (GDAS) model output. For low-level (CTP > 650 hPa) or liquid-phase clouds or when none of the band pairs converge to a solution, the 11-μm brightness temperature (IR BT) technique estimates a cloud-top temperature (CTT) and from that, a CTP/CTH is calculated from gridded model output, with provisions to adjust lapse rate for marine

stratus (Baum et al., 2012). The Terra MODIS CTH product used here is the Collection 6.1 Level 2 MOD06, which is provided in granule form at a 5-minute temporal resolution.

The ISS is at a mean altitude of 409 km above the Earth, revolving in a nearly circular low-earth orbit with an inclination of 51.64° and completing about 16 revolutions/day. The Cloud-Aerosol Transport System (CATS) (McGill et al., 2015; Yorks et al., 2016) instrument onboard the ISS operated from 10 February 2015 to 30 October 2017 and consisted of two elastic backscatter lasers that used a combination of low energy, high repetition rate 532 nm and 1064 nm pulses to achieve greater output power than any previous space lidar (Pauly et al., 2019). Although instrument failure prevented its multiple intended operating modes, nadir-only information was retained. During its run, CATS data was continuously downlinked at 60 m vertical and 350 m horizontal resolution (except for loss-of-signal periods), and then pre-processed, geo-located and calibrated to produce CATS Level 1 attenuated total backscatter and depolarization ratio profiles (Yorks et al., 2016). Geophysical parameters derived from Level 1 information is compiled into 5-km resolution Level 2 data, including depolarization ratio and attenuated backscatter, along with their layer-integrated values. The CATS layer-detection algorithm follows the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) (Vaughan et al., 2005; Yorks et al., 2015), with the main difference being that CATS applied threshold-based feature-detection on 5-km backscatter profiles at 1064 nm, as opposed to 532 nm for CALIOP. CATS layer-detection operated only at a single 5 km horizontal resolution (60 m vertical), whereas the CALIOP algorithm successively runs at fine to coarse horizontal resolutions ranging from 5 to 80 km in order to detect progressively tenuous layers (Vaughan et al., 2009). Cloud-aerosol feature-mask discrimination and cloud phase detection are identical to CALIOP. Details of these techniques can be found in the CATS Algorithm Theoretical Basis Document (Yorks et al., 2015). CATS Version 2.01 Level 2 Product used in this study provided values at every lidar range-gate associated with successful layer-discrimination. For this study, only range-gates with cloudy feature-masks were considered.

3 Collocation Methodology

For an accurate inter-comparison between instruments, one needs to be able to compare spatially and temporally concurrent observations, due to the transient nature of atmospheric conditions. In our case, MODIS has the widest swath and CTH is stored in 5-minute granules at 1 km resolution; whereas, MISR, with a much narrower swath nestled within the MODIS swath,

provides CTH at 1.1 km resolution that are stored per orbit. This enables a one-to-one collocation between MODIS and MISR pixels. However, CATS has a narrow Ground Instantaneous Field-of-View (GIFOV) of 14.38 m diameter, which is equivalent to its swath-width since it does not scan cross-track. Each CATS Level 2 datum has an along-track resolution of 5 km. Thus, when overlap of the Terra and ISS orbits did happen, it was possible to have multiple MODIS and MISR pixels neighboring a single CATS Level 2 point. Here, we choose a one-to-one collocation between each CATS point and the nearest-neighbor MISR and MODIS points, since the spatial correlation length for cloud properties can be of the order of tens or even a few hundreds of kilometers (Marchand, 2012). This choice is further justified later in this section. The mean geolocation difference for collocated pixels was found to be ~0.4 km for both CATS-MISR and CATS-MODIS collocation. To find the collocated set of data, the following choices were made:

1) Only those MISR data points are selected that lie within a distance of 380 km (MISR swath width) and whose observation time is within 5 minutes (to later accommodate MODIS granule time) from a given CATS point. From within this chosen subset of MISR data, a nearest-neighbor search finds the nearest point lying within a 1 km distance from the CATS data point, if any. If collocated points are found, only then is a MODIS search conducted.
2) MODIS granules that lie within a 5-minute window of a given CATS-MISR datum are selected for a nearest-neighbor search. When the point is found, MISR and MODIS CTH, MODIS CTH detection technique, and CATS cloud layer-heights, associated 1064 nm backscatter, surface elevation, and geolocation are extracted and stored. The altitude of the center of the highest lidar range-gate having cloudy feature mask in a column is taken as the cloud-layer height, whereas the base of the cloud-layer is taken to be the height of the range-gate, which is followed by at least 5 successive clear-featured gates.

Figure 1 shows an example of successful collocation among all three instruments from 14th March 2016, over southeast Asia. Figure 1a shows the three different swaths along with CTH from MISR and MODIS, with the collocated points marked in black. Figure 1b-1d shows the same scene in MODIS RGB, 1.38μm reflectance and 11μm brightness temperature, respectively, whereas Figure 1e depicts the CATS-retrieved vertical profile of cloud-masked attenuated backscatter, along with collocated MISR and MODIS CTH. This particular scene was chosen as it has low-lying cumuli, both with and without cirrus cover. For single-layered clouds, as between 21°N-

22°N, there is greater agreement between MISR and MODIS CTH. However, the presence of cirrus between 22°N-24°N (see lidar in 1b, cooler cloud tops in 1c and 1.38 μm imagery in 1d) leads to severe disagreements between MISR and MODIS, with MISR CTH consistently picking up the lower cumuli and MODIS retrievals being highly variable. The range wherein collocations are feasibly within the MISR swath extends from 21°N and 24°N so MISR CTH are only shown therein. All heights are with respect to the World Geodetic System 1984 (WGS84) ellipsoid.

An intuitive sense for the collocation process can be formed from Figure 2. Figure 2a shows a highly zoomed-in view of a patch of MISR and MODIS geolocations from the same scene as in Figure 1, with a set of CATS pixels cutting through. The search for collocated data is conducted within the 1-km radii circular windows that are marked around each CATS geolocation in Figure 2a (the circles are merely representative and not to scale). With navigation errors (~100 m), collocation differences (~400 m), and mismatches in pixel-size amongst instruments (~1 km x 1 km vs ~14 m x 5 km, it is the local CTH variations below these scales that introduce uncertainty in comparing MISR or MODIS CTH with CATS. To quantify this random error, we found all the MISR and MODIS data that lay within circular regions for each of the 9538 CATS points that satisfied the co-location conditions for the year 2016 and examine CTH variations as a function of radius of the circular region. For example, the histograms of the standard deviations in CTH within each region of 1-km radius (number of neighbors at least 2), denoted as σ_{MISR} for MISR and σ_{MODIS} for MODIS, is shown in Figure 2b. Both σ_{MISR} and σ_{MODIS} peak at 0.1 km, with their mean values being 0.2 and 0.5 km, respectively. Thus, the CTH of each collocated point from MISR and MODIS can be taken to be generally representative of CTH of all other observations within a 1 km circle centered around the CATS data point, with an uncertainty of about 200 m for MISR and 500 m for MODIS. There is also a mismatch in resolution between MISR/MODIS (~1 km) and CATS (5 km), as well as wind displacement of clouds during the maximum allowed time-interval between observations of 5 minutes in our coincidence criteria (e.g., a high wind speed of ~30 m/s can displace clouds close to 10 km in five minutes). Thus, local CTH variations over scales up to ~ 10 km also introduce uncertainty in comparing the CTH between MISR or MODIS with CATS. Thus, σ_{MISR} and σ_{MODIS} are calculated for progressively increasing search radii up to 10 km and plotted in Figure 2c. It is observed that both σ_{MISR} and σ_{MODIS} exhibit asymptotic behavior with increasing distances, reaching 0.3 km and 0.8 km, respectively. These values can be interpreted as an upper limit of CTH error owing to our method of collocation. The error is larger for MODIS

because MISR is generally more sensitive to lower clouds (owing to the higher spatial contrast they offer relative to thin cirrus) than MODIS, where variability in CTH and emissivity are smaller compared to high and midlevel clouds – evident, for example, in Figure 1 (e).

In most of this study going forward, the topmost CATS cloud layer height is compared against MODIS and MISR CTH, since satellite derived CTH is often associated with the height of the topmost cloud layer. However, to investigate the sensitivity of sensors to individual layers, the closest CATS layer to MISR/MODIS CTH is studied in Section 4.4.

4 Results and Causes of CTH Differences

By applying the collocation method mentioned above, 36 months (February 2015-October 2017) of collocated MISR, MODIS and CATS CTH have been compared spanning a quasi-global domain. In total, 51622 collocated (clear + cloudy) points were collected, among which, 27% were rejected as flagged clear by MODIS; 12% are outside the region of MISR swath with valid retrievals; 22% reported MISR CTH "no-retrievals" – that is MISR stereo failed owing to a lack of contrast (e.g., over clear sky ocean); and 2% did not have valid CATS cloud-layer retrieval where MODIS and MISR retrieved a CTH. Over land, (provided enough surface texture), MISR stereo can retrieve surface elevation as stereo height. Such features have been dealt with in our study by subtracting surface elevation from MISR stereo heights for every collocated point and further, only retaining such points in our analysis whose surface-elevation-corrected stereo heights were at least greater than 562 m – the value used by MISR for cloud designation (Mueller et al. 2013). The CATS pixel-level surface elevation from the 1x1 km USGS GMTED2010 digital elevation map (DEM) is used for this purpose. Finally, our analysis on valid CTH retrievals was conducted on a dataset of 18986 cloudy points.

4.1 Global and Regional Biases from MISR, MODIS and CATS Inter-comparison

Figure 3 shows the global distribution of all 18986 collocated CATS, MISR and MODIS data. Unless otherwise noted, CATS CTH will refer to the topmost CATS cloud-layer altitude. Figure 3 shows that there is a much higher frequency of collocation near the 50° latitudes in both hemispheres, due to greater swath overlap of Terra with ISS. This study is restricted to an inter-comparison over the tropics and midlatitudes since the ISS orbit does not venture further poleward. Also, Figure 3 shows that CATS detects the presence of a lot more very high CTH (e.g., West

Pacific warm pool region) than MODIS or MISR, owing to the lidar's ability to detect optically thinner clouds. MISR detects a lower mean CTH than CATS or MODIS, because MISR stereo is sensitive to spatial texture in multi-angular views, which is greater for lower, textured clouds, even under cirrus. The textured nature of the radiance field in the Western Pacific warm pool was recently examined by Hong & Di Girolamo (2020), demonstrating that the texture of ice-above-liquid clouds was only slightly smoother than liquid-only clouds owing to the fact that cirrus in the region are generally optically thin. Hence, the spatial contrast observed by MISR has the largest contribution from liquid clouds under conditions of ice-over-liquid clouds in the region.

Figure 4 shows the latitudinal dependence of CTH differences between the three instruments, expressed as (a) CATS-MODIS, (b) CATS-MISR and (c) MODIS-MISR. In each individual panel, the median CTH differences for every 5 degrees latitude interval from 60°N – 60°S were plotted at the mid-point of each corresponding interval. Each figure shows the median CTH difference for the bin for all clouds in black, CATS single-layered clouds in red and multi-layered clouds in blue. The error bars for each point signify the median absolute deviation, a robust statistic that is directly proportional to statistical dispersion but is resilient to the presence of outliers in a non-normal distribution. For CATS-detected multi-layered clouds, there are at least 2 cloud layers present, with the layers being separated by a vertical distance of at least 600 m (10 range-gates). The last panel (d) depicts the latitudinal distribution of number of samples. As can be seen from Figure 4a-4c, the largest differences in median CTH for all clouds (in black) are observed about the equator in the tropical regions (between 20°N – 20°S), owing to the contribution from multi-layered clouds. Large differences near the tropics were also noticed in the CALIOP and Aqua MODIS CTH difference record by Holz et al. (2008) and is due to the frequent presence of high and optically thin cirrus, often overhanging low and optically thick cumuli (e.g., Li et al., 2015; Stubenrauch et al., 2013). Moreover, from Figure 4, the median deviations for both CATS-MODIS and CATS-MISR CTH for multi-layered scenes is much greater than single-layered clouds. This increase for multi-layered scenes is more pronounced for CATS-MISR than for CATS-MODIS, because MODIS and CATS are theoretically more sensitive to higher clouds under cloud overlap, whereas, MISR is more sensitive to textured low clouds, even in the presence of overlying optically thin cirrus (e.g., Naud et al., 2007). The comparatively modest increase in MODIS for multi-layered clouds can be attributed to MODIS underestimating the semi-transparent top layer height, when the lower layer is optically thick (Menzel et al., 2015). The jump for multi-

layered clouds for MODIS-MISR is striking in its absence, suggesting median MODIS and MISR CTH are closely similar; this will be explored in upcoming sections.

4.2 Height of the Top Cloud Layer

To further investigate CTH differences, histograms for the three instrument pairs have been plotted in Figure 5. The CTH differences here are (a and d) MODIS-CATS, (b and e) MISR-CATS and (c and f) MISR-MODIS, respectively. 100 equal-sized bins between -20 km and 20 km, and between -5 km and 5 km, have been used for the top- and bottom-panel, respectively, with all histograms centered at zero. CATS CTH is the topmost CATS layer height. While analyzing these results, one must be mindful that different instruments' CTH might be due to cloud occurrence at different altitudes; this issue of cloud overlap in the interpretation of CTH differences is examined in Section 4.4. In Figure 5 and in figures to follow, an inverted system of axes in red has been added showing mean CATS top-layer height in each histogram bin, each point further color-coded by mean CATS top-layer layer-integrated backscatter (γ), for all scenes in that bin. A lower γ denotes an optically thinner cloud. A CATS $\gamma = 0.02$ sr^{-1} approximately corresponds to mean layer-integrated optical depth (OD) of 0.8 (linear regression between CATS Level 2 OD with integrated backscatter). In each Figure 5 subplot, the purple line signifies CATS high clouds (CTH > 5 km), the blue line signifies CATS low clouds (CTH < 5 km), while the dashed black line signifies all collocated points. Of these 18986 points, 10315 were high clouds and the rest low clouds.

Figure 5 shows that high absolute CTH differences in all cases arise from the presence of optically thin, high cloud layers. The peaks of the distributions for all clouds (dashed black) in the top panel are at (a) -800 m for MODIS-CATS, (b) -420 m for MISR-CATS and (c) -80 m for MISR-MODIS. These peaks exist where γ is largest. There exist prominent tails in all, extending up to about -15 km for MODIS-CATS and MISR-CATS and up to -12 km for MISR-MODIS. These long tails are due to optically thin, high clouds, with $\gamma < 0.02$ sr^{-1} and with mean CATS top-layer height > 10 km. Most cases in the MODIS-CATS (76%) and the MISR-CATS (89%) distributions involve negative CTH differences (i.e., MODIS and MISR CTH below CATS top-layer height). Most positive MODIS-CATS differences are for scenes with CATS top-layer height

below 7 km and $\gamma > 0.03$ sr^{-1} (OD ~ 1.2). Most positive MISR-CATS differences, however, arise from high clouds (CTH ~ 10 km), and moderate optical thickness ($\gamma \sim 0.02$ sr^{-1}, OD ~ 0.8).

As evident in Figure 5 and in other figures to follow, CTH differences follow approximately Gaussian distributions, exhibiting well-defined modes, with variable offsets from zero. So for a consistent inter-comparison between these datasets, we shall denote the distributions (restricted to absolute differences < 5 km) by their Mode and Mode standard deviation (σ), where $\sigma = \frac{FWHM}{2\sqrt{2\ln 2}}$ (assuming normality), and FWHM is the Full Width at Half-Maximum for our distribution. These statistics are chosen over a simple mean or variance of the data, because they are not skewed by outliers, that arise due to differing sensitivities of CATS, MISR and MODIS to different clouds in a scene. We note that such a mode and σ for the MODIS-CATS and MISR-CATS CTH difference distributions represent the bias and precision of MODIS and MISR CTH, assuming CATS CTH as the truth and the errors to be normally distributed about the mode. As we analyze the many parameters on which these sensitivities depend (e.g., top-layer properties and multi-layering), we shall gain more insight into these primary modes and outliers.

For high cloud scenes (purple lines, Figure 5), there is much disagreement between the three instruments. From Figures 5a and 5d, MODIS high-cloud CTH bias = -1200 m and precision = 1080 m, while from Figures 5b and 5e, MISR high cloud bias = -540 m and precision = 590 m. This difference in the MODIS and MISR distribution arises primarily from scenes where multiple cloud layers are present and the instruments are identifying different layers to report height, with MISR being more sensitive to lower clouds, while MODIS CTH is dependent on optical and geometrical properties of the multiple cloud layers in the scene (Holz et al., 2008; Naud et al., 2002; Naud et al., 2007; Stubenrauch et al., 2013). Further MODIS errors arise due to optically thin, geometrically thick cirrus, as the assumption of an infinitesimally thin cloud layer is central to the effectiveness of CO_2-slicing. The probabilities of MISR and MODIS detecting the true height of a CATS high cloud to within ±1 km is nearly equal at about 15%, in spite of MISR not being as sensitive as MODIS to optically thin cirrus. MODIS underestimation of high CTH for multi-layered scenes seems to be the primary reason behind this phenomenon.

There is much agreement between the instruments for low clouds (blue line, Figure 5). From Figures 5b and 5e, MISR-CATS CTH difference exhibits a sharp distribution, with MISR

low-cloud bias = -320 m and precision = 250 m. MISR-CATS low cloud CTH differences fall within 0 and -600 m 74% of the time. In comparison, MODIS low cloud CTH (Figure 5a and 5d) exhibits a bias = 40 m and precision = 730 m, with 14% of MODIS-CATS differences below -2 km and 29% of differences above 0. For low clouds, MODIS uses the IR BT technique with latitudinally varying climatological lapse rates (Baum et al., 2012). Significant deviations from these lapse rates is a source of uncertainty. Holz et al. (2008) and Harshvardhan et al. (2009) demonstrated that the Collection 5 MOD06 product was overestimating CTH by over 2 km in cases where a low-lying liquid phase cloud was present over the ocean, particularly in the presence of strong temperature inversions, due to poor representation in ancillary data. As rectification, the Collection 6 MOD06 started using zonally-averaged "apparent 11-μm brightness temperature (BT) lapse rates" from a combination of CALIOP CTH and modeled sea-surface temperatures to better capture boundary-layer lapse rates (Baum et al., 2012). This improvement manifests itself in the absence of the hump in positive MODIS-CATS differences that was observed in MODIS-CALIPSO differences reported in Figure 8 of Holz et al. (2008).

Despite MISR applying stereoscopy and MODIS a radiometric technique, the two passive sensors do produce reasonable agreement in CTH. The MISR-MODIS CTH difference distribution (Figure 5c and 5f) has Mode = -400 m and σ = 680 m. 62% of all MISR-MODIS CTH differences lie between ±2 km, with an optically thick top-layer at a mean altitude of about 5 km, and the spread of the distribution is attributable to the natural variability of clouds in a scene and the different sensitivities of MISR and MODIS to this variability (Section 3). About 36% of CTH differences is constituted by differences between 0 and -2 km, mostly for top layers of cloud with integrated backscatter larger than 0.02 sr^{-1} and with heights < 10 km, and is associated with MODIS IR BT CTH overestimation for stronger temperature inversions (note, IR BT technique is applied for all but high and mid-level ice clouds). A sizeable portion of MISR-MODIS differences in both high and low cloud scenes (25% and 36%, respectively) have positive values up to +2 km. Positive MISR-MODIS bias (mean difference = 0.6 km) for optically thin clouds is primarily due to

optically thin and geometrically thick cirrus (mean geometric depth of top layer in the 0 to +2 km interval from CATS ~ 1.2 km) and this role of OD on bias will be explored in the next section.

4.3 Optical Depth of the Top Cloud Layer

Figure 5 suggests that as one moves from large negative top-layer CTH differences to zero, there is a general tendency of the top-cloud layer to be lower and optically thicker for MISR and MODIS. As one moves from zero to positive CTH differences, the top layer starts to be slightly higher, with only a modest reduction in the mean backscatter. These tendencies are consistent with our knowledge of the three CTH retrieval techniques. This is especially true for CATS and MODIS, because their retrievals are highly dependent on cloud optical properties. An optically thicker top layer of cloud represents an opaque or a nearly opaque atmospheric column to the lidar, which leads to rapid attenuation of the lidar signal near the cloud top. This represents a strongly emissive cloud-top layer, hence the retrieved CTP (in case of the CO_2-slicing technique) or the CTT (in case of the 11µm-BT technique) is very close to actual values. However, for more transmissive cases, CTT and BT can diverge substantially, resulting in lower CTH under typical conditions; higher under atypical conditions (i.e., surface, or lower cloud layer being cooler than the cloud-top layer). The CO_2-slicing approach hinges on an assumption of a thin cloud layer, and any geometrical depth (especially accompanied by low optical depth) can lead to underestimation of CTH, through an overestimation of CTP. Smith & Platt (1978) estimated errors ~50 hPa in CTP for a cloud of ~100 hPa depth and CO_2-slicing is generally likened to a centre of mass problem (Menzel et al., 2008), with CTP errors co-varying with optical depth into the cloud (i.e., CTP close to true cloud-top for optically thick, and closer to the geometric centre for optically thin cases). As a result, the CTH difference in these cases, is a function of the vertical distribution of extinction in the cloud layer, as well as temperature throughout the column. On the other hand, although MISR makes use of a stereoscopy, MISR-CATS differences are also expected to depend on the vertical distribution of single scattering properties of the top cloud layer, as well as its horizontal distribution that gives rise to the spatial contrast for stereoscopy to work. For a single layer cloud, the contrast is expected to emerge over some depth of the cloud layer that is likely deeper than a lidar-derived height. For an optically thin upper cloud overlapping an optically thick lower cloud, the largest spatial contrast may well emerge from the lower cloud layer, allowing stereo to retrieve

the CTH of the lower cloud layer. The exact relationship of this '*stereo-opacity bias*' with the 3D distribution of cloud optical properties has yet to be quantified from theory or experiments.

To gauge the impact of the top-layer cloud optical properties on the retrieval of CTH for low and high clouds from MODIS and MISR, Figure 6 shows histograms of CTH differences for the three instrument pairs with 100 equal-width bins between -5 km and +5 km, for optically thick top cloud layers ($\gamma > 0.02$ sr^{-1}) in purple and optically thin top cloud layers ($\gamma < 0.02$ sr^{-1}) in blue. The top panel (Figure 6a-6c) is for CATS high clouds (CATS CTH > 5 km), while the lower panel (Figure 6d-6f) is for CATS low clouds (CATS CTH < 5 km). Based on the observed relationships between CTH differences and backscatter in the previous figures, $\gamma = 0.02$ sr^{-1} (OD~0.8) is simply chosen as a distinction between optically thick and optically thin cloud.

From Figure 6a, the MODIS-CATS CTH difference for optically thin, high topmost cloud layer shows much variation, especially for negative differences. The issues faced by the CO_2-slicing technique for semi-transparent clouds are many-fold, including errors due to cloud geometrical depth and the presence of lower cloud layers (Smith & Platt, 1978; Wielicki & Coakley, 1981; Wylie & Menzel, 1989), leading to MODIS optically thin high clouds bias = -1160 m and precision = 1020 m, with 84% of MODIS-CATS CTH differences being negative. MODIS optically thick high cloud CTH has bias = -280 m and precision = 730 m, and the MODIS-CATS distribution has a sharper peak. From Figure 6b, MISR optically thick high cloud bias = -440 m and precision = 470 m, with the MISR-CATS distribution being less noisy than the corresponding MODIS-CATS distribution, although the larger MISR bias suggests that MODIS heights are closer to the lidar cloud top than MISR for these clouds. The distributions for optically thick (mode = -20 m, σ = 610 m) and thin high clouds (mode = -320 m, σ = 640 m) for the MISR-MODIS difference (Figure 6c) are both symmetrical about their modes.

For CATS low clouds, from Figure 6d, the MODIS-CATS distribution shows two distinct peaks – the optically thin cloud distribution (bias = -440 m, precision = 600 m) and the optically thick cloud distribution (bias = 500 m, precision = 430 m) are both consistent with the limitations of the IR BT technique and with the Collection 6 improvements. Optically thin clouds being more transmissive, allow more IR radiation from closer to the warm surface to reach the satellite, leading to negative CTH bias, whereas the positive bias for optically thicker or more emissive clouds

presumably owes its origin to a larger deviation of true boundary-layer lapse rates from the Collection 6 climatological lapse rates. However, it needs to be noted that the bias for both optically thin and thick low clouds show a marked improvement from Collection 5 (Figure 11b of Baum et al. 2012). The MISR-CATS distributions (Figure 6e) for optically thick low clouds (bias = -280 m, σ = 260 m) and optically thin low clouds (bias = -320 m, precision = 310 m) exhibit a slight dependence of MISR low cloud retrieval on the optical depth (see discussion above). This will be explored in the next section after we have quantified the relationship of CTH differences with multi-layering. The distributions for optically thick and thin low clouds for the MISR-MODIS difference (Figure 6f) closely resemble that of the MODIS-CATS CTH distributions as the dependence of MISR CTH on OD is considerably lesser than that of MODIS CTH.

To recap, the previous two sections have quantified CTH differences between sensors, examined how these differences depend on the top layer properties, and provided evidence of significant contribution from cloud overlap in explaining these differences. The next section isolates those contributions and, in their absence, examines the depth within the cloud that these instruments are most sensitive to.

4.4 Multi-layered Clouds

Past research (Marchand et al., 2007; Naud et al., 2007; Naud et al., 2004; Holz et al., 2008) and the previous sections have flagged multi-layered clouds as leading to passive sensor CTH errors. To quantify this, histograms of CTH differences is shown in Figure 7, based on multi-layering for CATS high cloud (CTH > 5 km), with 100 equal-sized bins between -20 km and 20 km. The purple line indicates single-layered high cloud, and the blue line indicates at least more than one layer (with minimum vertical separation of 600 m) and the black line is a histogram for all high clouds. Moreover, since an optically thick high cloud can completely attenuate the lidar signal (preventing low-cloud detection), we further restrict single-layered clouds to those scenes with CATS Percentage Opacity lesser than or equal to 0.5. CATS reports a Percentage Opacity, defined as the fraction of “opaque” (no surface detection) 350 m-resolution Level 1 samples that constitute a Level 2 5-km datum. A value of 0 signifies ‘all profiles transparent’, while 1 signifies ‘all profiles opaque’. Note, the term ‘Percentage Opacity’ should not to be confused with a measure of cloud optical depth; rather it should be thought of as a measure of the sub-pixel transmittance

488 homogeneity of a CATS datum. This threshold is applied to reduce the occurrence of multi-layered
489 broken clouds, which can make comparisons between the different product resolutions tenuous.

490 From Figure 7, the greatest occurrence of negative MODIS-CATS (Figure 7a) and MISR-
491 CATS CTH differences (Figure 7b) are found in multi-layered cases where the top-layer has $\gamma <$
492 0.02 sr^{-1} and a mean CTH of more than 10 km. In multi-layered cases, CATS top-layer CTH and
493 MODIS CTH are within 1 km of each other for only about 10% of the time, while it is only 4%
494 for CATS top-layer CTH and MISR CTH. Compared to that, negative differences less than -2 km
495 are observed in 7% of all single-layered cases in the MODIS-CATS distribution (bias = -1160 m,
496 precision = 510 m) and a total of 5% in the MISR-CATS distribution (bias = -720 m, precision =
497 460 m), with these high negative values also due to semi-transparent high clouds (same as multi-
498 layered scenes). It is worth noting that even in single-layered cases, we cannot rule out the
499 presence of an optically thick lower layer below (Percentage Opacity = 0.5 can mean a maximum
500 of 7 out of the 14 350-m Level 1 profiles in the CATS datum were transparent).

501 For MODIS-CATS and MISR-CATS differences, positive values are found as well,
502 primarily for CATS single-layered clouds. While these positive values comprise 11% of the
503 MODIS-CATS and 9% of the MISR-CATS distribution, it is worth noting that these values extend
504 up to +5 km for MODIS-CATS, and are due to optically thick top-layers with mean CTH of 7.8
505 km; while for MISR-CATS, these positive differences extend up to about +2.5 km, but are due to
506 optically thick top-layers with mean CTH of 11.6 km. Positive MISR height bias for high clouds
507 is due to wind-retrieval bias at those heights (Horváth, 2013); positive MODIS bias for high clouds
508 requires an independent discussion provided in Section 4.6.

509 MISR-MODIS CTH differences (Figure 7c) does not show striking differences between
510 single-layered (Mode = -80 m, σ = 670 m) and multi-layered scenes (Mode = -80 m, σ = 710 m),
511 except in the tail. Overall, MISR and MODIS sense the same cloud to within 1 km of each other
512 nearly 30% of the times – 25% for multi-layered and 32% of all single-layered scenes. These
513 scenes constitute the primary peak of the distributions and have a top-layer mean backscatter of
514 0.012 (OD~0.5) and mean altitude of 11.8 km.

515 For multi-layered cases, it is also necessary to quantify which cloud layer the passive sensor
516 is sensitive to; hence, separate histograms of differences between MODIS and MISR CTH and

CATS Layer 1 (top layer) and Layer 2 (bottom layer) heights are plotted in Figure 8, for scenes where CATS detected just two distinct layers of clouds at least 600 m apart. The difference between MISR or MODIS CTH and the closest CATS layer height is plotted in black. Figure 8 shows that MISR is highly sensitive to CATS Layer 2 (with bias = -400 m and precision = 350 m), seen in the tight overlap of the 'MISR - CATS Layer 2' and 'MISR - Closest CATS Layer' curves. For scenes where MISR detected CATS Layer 2, the mean and SD of the top-layer γ (OD) were 0.002 (~0.08) and 0.0052 (~0.18), respectively. MISR detection of CATS Layer 1 has bias = -820 m and precision = 850 m), with a mean and SD top-layer γ (OD) of 0.01 (~0.4) and 0.009 (~0.3), respectively. This suggests the possibility of a threshold OD necessary for MISR stereo to detect thin cirrus overhanging a textured low cloud, as was suggested in Marchand et al. (2007). One might expect this threshold to be a function of sun-satellite geometry, texture, and resolution, requiring future investigations using observations and radiative transfer modeling.

Figure 8b shows that MODIS CTH tends to lie between the two layers as indicated by large negative and positive tails for Layer 1 and Layer 2, respectively, and being strongly aligned with neither. For small negative MODIS-CATS differences, CATS Layer 1 is preferred (bias = -1200 m, precision = 1190 m) as the closest CATS layer (CO_2-slicing – negative bias); while for small positive MODIS-CATS CTH differences, CATS layer 2 (bias = 20 m, precision = 850 m) is preferred (BT technique – positive bias). This is consistent with Sections 4.2 and 4.3.

4.5 CTH Bias and Precision by Instrument

Sections 4.2-4.4 investigated the effects of cloud parameters (top-layer height and optical depth, and multi-layering) on the error characteristics of MISR and MODIS CTH retrievals, by assuming CATS CTH to be the truth; these results are summarized in Table 1. However, to constrain our error estimates further, we seek to remove the inherent uncertainty in the collocation process, as well as eliminate the possibility of having multiple layers in a scene. To this end, for the determination of instrument bias and precision, we now restrict ourselves to only single-layered CATS Level 2 profiles with Percentage Opacity = 1 (i.e., all constituent Level 1 profiles that went into the 5-km product being opaque); suggesting an absence of broken, multi-layered clouds and with minimum layer-integrated OD ~ 3 (the OD at which a CATS signal is completely attenuated) and where the absolute values of MISR-CATS and MODIS-CATS differences are less than 2.5

546 km (approximately, the largest FWHM from results above). This leaves us with ~6000 data points,
547 each for both MISR and MODIS investigation.

548 MISR and MODIS bias (offset of the distribution mode from 0) and precision (σ from the
549 FWHM approach) are calculated and summarized in Table 2 for all, high (CATS CTH > 5 km),
550 mid-level (10 km > CTH > 5 km) and low-level (CTH < 5 km) clouds. Moreover, as summarized
551 in Table 1, MISR and MODIS bias and precision exhibits variable dependence on the height and
552 OD of the cloud layer, and to investigate further, Figure 9 shows the distribution of MISR (Figure
553 9a) and MODIS (Figure 9b) bias and precision (1σ errors-bars) with altitude (for every 2 km
554 interval) for the same scenes. Each such interval contains a minimum of 150 collocated pixels (~7-
555 10 independent scenes). The mean CATS integrated backscatter from cloud top to 300 m below
556 cloud top, γ_{300}, is also shown for each bin. It is readily apparent from both Table 2 and Figure 9
557 that MISR exhibits lower bias and greater precision than MODIS CTH, and that both MISR and
558 MODIS precision deteriorates with increasing altitude of the cloud. A detailed discussion of the
559 error budget for MISR and MODIS CTH are discussed in the next two sub-sections.

560 **4.5.1 MISR CTH Errors**

561 The MISR bias reported in Table 2 arises from three principal sources – bias in co-
562 registration of oblique radiances with nadir, wind-retrieval bias, and a stereo-opacity bias (retrieval
563 of stereo height at a depth into the cloud due to low extinction near the top). We assume CATS
564 CTH to be unbiased. Sources of random error that determine the overall MISR CTH precision
565 include geo-registration errors of MISR imagery, correspondence errors of conjugate cloud
566 features in MISR imagery, random wind-retrieval errors, and random sub-pixel CTH variability
567 due to geo-collocations (~300 m from Section 3). We assume the random error in CATS CTH to
568 be the result of equal probability of successful and failed detection over the depth of one range-
569 gate: thus, contributing a random error of 30 m. Globally, MISR image geo-registration error is
570 estimated to be 0.05 ± 0.25 pixels, which translates to height errors of about 30 ± 140 m (Davies
571 et al., 2007; Jovanovic et al., 2007).

572 Wind-retrieval errors also propagate to height errors, although these contributions have
573 been reduced from the TC_STEREO to TC_CLOUD product (Horváth, 2013; Lonitz & Horváth,
574 2011; Mueller et al., 2017). Comparison of MISR near-surface heights to ground targets allows

for the evaluation of CTH errors due to the combined effects of registration, correspondence, and DEM errors, as done in Horváth (2013). We repeated their analysis using MISR data between 50°N and 50°S, finding the mode in height to be = -40 m and σ = 170 m, using the FWHM approach. These values are similar to the values (mean height error = -31 m, RMS error = 171 m) reported in Horváth (2013).

If the bias was not a function of wind speed, then we would conclude for the overall cloud samples used in Table 2 that the stereo-opacity bias is -280 m + 40 m = -240 m (-260 for high clouds; -200 m for low clouds). However, based on the analysis of Horváth (2013), wind errors do vary with altitude based on comparisons with geostationary wind data, with an along-track wind bias of ~1 m s^{-1} for low clouds and ~1.2 m s^{-1} for high clouds, respectively (cross-track winds are essentially unbiased). Using ~90 m height error for each 1 m s^{-1} error in along-track wind-speed (Mueller et al., 2017; Table 1), these wind-errors translate to CTH errors of ~90 m for low clouds and ~110 m for high clouds. Thus, we deduce that the stereo-opacity bias is ~ -110 m (= -200 m + 90 m) and ~ -150 m (= -260 + 110 m) for low and high clouds, respectively. The difference in stereo-opacity bias between high and low clouds is likely due to lower clouds having larger γ_{300} (i.e., greater extinction coefficients in the upper parts of the cloud; Figure 9a).

For the MISR CTH precision budget, we noted earlier that the MISR CTH co-registration/correspondence/DEM precision = 170 m and that the maximum CTH geolocation precision of our method = 300 m. But, since we are dealing with Percentage Opacity = 1 in this section, we expect the geolocation-related variations in heights to be much smaller here, and our overall observed precision of 370 m may be almost entirely dictated by the precision of MISR stereo. Here, it should be noted that our observed precision is about twice as good as was reported in both Horváth (2013) and Mueller et al. (2017), and is most likely due to the highly precise CTH that a lidar is able to offer (taken here as 30 m) compared to the IR AMV heights used in those studies. Assuming geolocation-related height error = 0 m, and the overall MISR precision from Table 2 to be 373 m, the MISR wind-height precision = $\sqrt{\mathbf{370^2 - 170^2 - 30^2}}$ = 330 m (360 m and 250 m for high and low clouds, respectively). Using the 90 m (m s)$^{-1}$ wind-height error proportionality again, we get an overall MISR wind speed precision of 3.7 m s^{-1} (4.0 m s^{-1} and 2.8 m s^{-1} for high and low clouds, respectively). Our MISR wind speed precision estimates backed out through MISR-CATS comparison are remarkably close to those determined by both Horváth

(2013) and Mueller et al. (2017), thus providing closure. This result also implies that the MISR operational quality assurance procedures, most notably the required agreement (and subsequent averaging) of forward and aft-derived height estimates, are filtering and improving the accuracy of raw stereo retrievals to an extent that mitigates the difficulty of obtaining heights from highly dynamic or poorly textured clouds.

4.5.2 MODIS CTH Errors

A similar accounting of MODIS CTH bias and precision is not strictly possible as MODIS uses *a priori* assumptions and ancillary data to determine CTH. Its errors covary with the departures from these assumptions and deviations from reality in the ancillary data. The errors in IR sensors have been historically quantified as CTP errors (Menzel et al., 2015; Wielicki & Coakley, 1981), although in recent literature (Baum et al., 2012; Holz et al., 2008), CTH errors have been quantified by comparing low-level and single-layered clouds against lidar. For example, from Figure 12 of Baum et al. (2012), we can estimate the bias and precision (FWHM method as above) to be -1100 m and 930 m, respectively, for single-layered cirrus, and a bias and precision of 200 m and 550 m for low clouds. The corresponding values of bias and precision for high and low clouds in Table 2 are quite similar, even though we define high and low clouds differently than that study. The MODIS bias (Table 2) seems to be largely due to high clouds, which goes back to systematic bias in the CO_2-slicing technique, which employs an infinitesimally thin cloud assumption. In these high cloud samples (optically thick cirrus), the negative bias presumably arises because optically thick cirrus also tends to be geometrically thick, leading to CO_2-slicing underestimating CTH. Again, owing to the choice of Percentage Opacity = 1, MODIS precision is assumed to be mostly unaffected by collocation errors, and originates from the forward modelling.

Systematic CTH overestimation by MODIS for low clouds and underestimation for semi-transparent high clouds is due to the retrieval techniques it employs and cannot be explained by just top-layer height, OD and overlap. For MODIS, the low and high cloud distinction used here nearly coincides with the 75^{th} percentile heights (green-dashed lines in Fig. 9b) where IR BT and CO_2-slicing techniques are applied, whereas mid-level clouds employ both. As a result, the bias

and precision of the two techniques can be roughly estimated by MODIS bias and precision for high and low clouds (Table 2).

To investigate the MODIS CTH bias and precision for the two CTH techniques, Figure 10 presents histograms for MODIS-CATS (10a and 10c) and MISR-MODIS (10b and 10d) CTH differences for all CATS high (CTH > 5 km) clouds (top panel) and single-layered high clouds (bottom panel). Only high cloud retrievals are chosen to focus on scenes where CO_2-slicing is preferred, but IR BT is still possible. A simple pressure-based distinction is not applicable as CO_2-slicing is only reserved for ice clouds. Aqua-MODIS phase flag is understood to accurately determine ice phase 65-80% of the time globally, through inter-comparisons with CLOUDSAT/CALIPSO data, with >90% agreement for multiple surface types for single-phase clouds (Marchant et al., 2016; Platnick et al., 2017). From Figure 10, we see that 57% of all collocated high clouds and 70% of single-layered high clouds were retrieved using CO_2-slicing, in keeping with the improvements of Collection 6 MOD06 updates aimed at increasing frequency of CO_2-slicing retrievals (Baum et al., 2012). In both high and single-layered high clouds, the smallest differences are associated with CO_2-slicing for MODIS-CATS and with IR BT technique for the MISR-MODIS. This discrepancy is because CO_2-slicing is more sensitive to optically thin high clouds than MISR and has a mean CTH closer to mean CATS CTH in most cases (especially, single-layered high clouds), as shown in Figure 5. However, due to reasons explained earlier, MODIS often detects mid-tropospheric CTH about 3-5 km above MISR CTH. Large MODIS-lidar differences occur for IR BT technique, as noted in previous studies (Holz et al., 2008; Naud et al., 2004), for semi-transparent high clouds (OD < 1), where MODIS opts for the IR BT technique over the more precise CO_2-slicing. The mean MODIS CTH error associated with the application of the IR BT technique is found to be -5.8 km overall for scenes with CATS CTH > 5 km, with CATS mean top-layer backscatter less than 0.02 sr^{-1} and mean top-layer height greater than 10 km. For CATS single-layered high clouds, mean CTH error from the application of IR BT is -2.3 km.

For MODIS-CATS difference (Figure 9a and 9c), there exists a noticeable hump in the distribution for positive values and is associated with optically thick, single-layered high clouds (as in Figure 6a) with top-layer backscatter greater than 0.01 sr^{-1} and CTH between 5-10 km. From Figure 9, this hump is clearly associated with CO_2-slicing and not IR BT and is difficult to explain based on the information at hand. CO_2-slicing is subject to many sources of errors – instrument

noise, uncertainties in calculating clear-sky radiances, assumption of constant emissivity in band-pairs used to calculate CTP and deviations from constant lapse rates. Apart from these, there may be two more sources of error for the present data. Firstly, CATS, unlike the CALIOP lidar used in Holz et al. (2008), employs a single horizontal resolution (5 km) for layer-detection and is known to miss extremely tenuous cirrus layers during daytime (Rajapakshe et al., 2017). As a result, it might be possible that MODIS 1-km CTH can detect small, thin higher cirrus that CATS might miss. Secondly, a problem endemic to Terra MODIS, but not with Aqua-MODIS used in Holz et al. (2008), is that one of the bands used in the CO_2-slicing – Band 34 (13.6 μm) – remains unused due to severe noise, effectively reducing the algorithm to just 14.2/13.9 and 13.9/13.3 μm ratios (most sensitive to pressure regimes of 100-450 hPa and 550-650 hPa, respectively), instead of the full suite of options (Menzel et al., 2008). Hence, in this analysis alone, 73.1% of all CO_2-slicing retrievals for high clouds and 78% for single-layered high clouds were from the 14.2/13.9 μm band-pair, while the remainder came from 13.9/13.3 μm band-pair. The important 35/33 (13.9/13.6 μm) band pair, most sensitive to mid-level clouds and cloud edges (Menzel et al., 2015), is missing and is a possible reason for overestimation of mid-level CTH.

5 Conclusions

Terra is our longest running single-platform mission with a stable ECT for cloud top heights (CTH), now spanning more than 2 decades. Its long record from a stable orbit makes it valuable in climate research and in data assimilation in reanalysis products. Of course, its scientific application requires well characterized errors in the public geophysical products produced by the Terra mission. Here we have used the ISS CATS lidar to quantify the error characteristics of MODIS and MISR CTHs from Terra, producing the first quasi-global evaluation of these errors from space-based lidar. Ample collocated (< 1 km) and concurrent (< 5 minutes) MODIS, MISR and CATS samples were retrieved during the CATS 2015-2017 period for robust statistics. While CATS top-layer CTH is taken as truth in our analysis, the CATS-detected lower-level cloud tops underlying thin upper-level clouds were also used to examine MODIS and MISR CTH error characteristics – an approach that proved to be central in our understanding of MISR and MODIS CTH. Generally, we find that MISR and MODIS CTH errors are larger in the tropical regions and

690 smaller in the midlatitudes, and are strong functions of cloud type, defined by cloud height, optical
691 depth and multi-layering, as summarized in Tables 1 and 2.

692 For CATS CTH < 5 km (single or multi-layered), MISR and MODIS CTH biases and
693 precisions (bias ± precision) are -320 ± 250 m and 40 ± 720 m, respectively. MISR CTH bias
694 changes little with optical depth (Figure 6), but a reduction of MISR CTH bias to -240 m for
695 unbroken, single-layered and opaque low clouds is observed (Table 2). In contrast, MODIS CTH
696 bias for low clouds (hence, IR BT technique) is highly dependent on optical depth with average
697 bias of -440 m for thin clouds ($\gamma < 0.02$ sr^{-1} or OD < 0.8) and of +500 m for thick clouds ($\gamma > 0.02$
698 sr^{-1}) (Table 1). This dichotomy occurs because for optically thinner (more transmissive) clouds,
699 the IR BT technique senses a thermal signature of the warmer surface, whereas, for high OD (more
700 emissive) clouds, there is presumably greater lapse rate deviation from the climatology used in
701 Collection 6 MOD06 product. When considering the subset of unbroken, single-layered, and
702 opaque low clouds, MODIS CTH bias is +60 m (Table 2), with the positive bias for more emissive
703 clouds dominating, as low clouds tend to be thicker on average in our dataset.

704 For CATS CTH > 5 km (single or multi-layered), MISR and MODIS CTH biases are -540
705 ± 590 m and -1200 ± 1080 m, respectively. For both MISR and MODIS, high cloud biases do tend
706 to vary with optical depth. MODIS CTH bias is -1160 m for thin high clouds ($\gamma < 0.02$ sr^{-1}) and -
707 280 m for thick clouds ($\gamma > 0.02$ sr^{-1}). Low opacity near cloud-top in geometrically thick clouds
708 leads to underestimation of MODIS CTH as CO_2-slicing technique assumes an infinitesimally thin
709 single-layered cloud solution. Similarly, the MISR CTH bias is -680 m for high clouds with $\gamma <$
710 0.02 sr^{-1} and -440 m for those with $\gamma > 0.02$ sr^{-1}, suggesting the presence of a stereo-opacity bias –
711 the depth into the cloud in which spatial contrast is established in the emerging radiation field.
712 This study provides the first assessment of the MISR stereo-opacity bias, estimated here to range
713 between -110 and -150 m for clouds sampled in this study. It is larger for higher altitude clouds
714 owing to the optically thinner nature of cloud tops for higher altitude clouds.

715 For CATS-retrieved multi-layered clouds, which are often thin cirrus overlying thicker
716 clouds, CTH comparisons are more complicated. Both passive sensors severely underestimate top-
717 layer CTH, MISR by -820 ± 850 m and MODIS by -1200 ± 1190 m. These large biases necessitate
718 us to adopt a "closest layer" approach (i.e., comparing passive-sensor CTH to closest CATS layer

height). For two-layered cases, MISR is found to be sensitive to the lower cloud layer, with MISR CTH errors for this lower layer being -400 ± 350 m. This is almost identical to MISR single-layered low cloud bias and precision, suggesting that MISR low CTH accuracy is independent of the presence of a high, thin cirrus. The mean top-layer OD when MISR detects the higher layer is found to be 0.4 ± 0.3, agreeing with the result from Marchand et al. (2007). This is indicative of an opacity threshold for stereo detection, a parameter which would presumably be a function of sun-satellite geometry and spatial contrast. MODIS underestimates top-layer CTH by greater than 1 km due to the CO_2-slicing technique converging at a higher-pressure solution, when an optically thin (OD < 0.8) cloud is present. As a result, MODIS produces more midlevel CTH than MISR and MISR-MODIS CTH differences have generally low absolute values.

Optically thick, single-layered, unbroken clouds allow us to neglect random collocation errors (~300 m) for a complete error budget analysis for MISR stereo. Unlike MODIS, the MISR CTH error budget is self-contained since it does not rely on external ancillary products. MISR underestimates CTH for these clouds by -280 ± 370 m. Contributors to the bias are estimated as: (a) bias in imagery co-registration and feature correspondence (~ -40 m), (b) MISR stereo-opacity bias (-110 to -150 m, dependent on cloud altitude) and (c) MISR wind-correction bias (-90 to -110 m, also dependent on altitude). Random errors in this dataset are largely due to wind-driven errors (330 m for all samples, 250 m for low and 360 m for high clouds). Based on our estimated wind-height precision, we were able to provide an independent estimate of MISR wind-speed precision of 3.7 m s^{-1} (2.8 m s^{-1} and 4.0 m s^{-1} for low and high clouds, respectively). These values are quite similar to the findings of Horváth (2013) and Mueller et al. (2017). Thus, we conclude that we have essentially achieved closure on the MISR CTH error budget.

Similarly, MODIS underestimates CTH by -540 ± 690 m for these optically thick, single-layered, and unbroken clouds in our dataset. While it is difficult to exactly quantify, the largest contributor to MODIS CTH bias is the CO_2-slicing underestimation for geometrically thick cirrus. MODIS CTH random errors are due to inherent uncertainties in the forward model and reliance on external ancillary datasets. Since CO_2-slicing is best suited for thin cirrus, application of IR BT for high clouds (when CO_2-slicing does not converge to solution), can still lead to erroneous results, as discussed in Holz et al., (2008). However, compared to Collection 5, improvements of Collection 6 low-cloud CTH from the marine boundary-layer correction, as well in high-cloud

retrievals from adopting CO_2-slicing technique more frequently (Baum et al., 2012), have indeed led to a substantial reduction in errors in MODIS CTHs.

Based on the findings presented here, it is clear that MISR and MODIS CTHs can be combined for improved interpretation of CTH variability, particularly in multi-layered conditions. Our findings also point to recommendations for future satellite architecture designs that have CTH as a target product, such as the Aerosol and Cloud, Convection and Precipitation (ACCP) mission called out in NASEM (2018). As each of these sensors (lidar, IR, multi-view) occupies a niche that cannot be replaced by the others alone, these sensors on a single-orbit taking observations of the same physical reality can improve the short-comings of each by creating fused datasets that complement each other and provide greater insight to CTH variability than any of these sensors operating alone. Also, our analysis and closure of the MISR CTH error budget has several implications for future stereo-enabled technological designs. Since the largest contributor to the error budget are wind-driven errors, removing this error can be achieved by flying two (or more) multi-view imaging systems in close proximity and in close formation. This would allow for the same scene to be viewed at the same time; hence removing wind-driven errors. Improving resolution would also improve the precision in the stereo CTH (an instrument resolution of ~100 m would contribute ~60 m to the precision budget, assuming MISR viewing geometry). We recommend that detailed 3D radiative transfer modeling be undertaken to fully understand the nature of the remaining stereo-opacity bias – how it varies with sun-satellite geometry and cloud micro- and macro-physical properties.

Acknowledgments, Samples, and Data

This research was supported under MISR project contract 147871 with the Jet Propulsion Laboratory, California Institute of Technology. Partial support from the NASA ACCESS program under contract NNX16AMO7A is also acknowledged. The MISR data were obtained from NASA Langley Research Center Atmospheric Sciences Data Center (https://l0dup05.larc.nasa.gov/opendap/misrl2l3/MISR/contents.html). The MODIS data were obtained through the Level 1 and Atmosphere Archive and Distribution System of NASA Goddard Space Flight Center (https://ladsweb.modaps.eosdis.nasa.gov/archive/allData/61/). The CATS data was downloaded from the NASA Langley Research Center's ASDC DAAC (https://opendap.larc.nasa.gov/opendap/CATS/). We are thankful to the NASA MODIS, MISR and CATS teams for supplying the documentation and tools, especially the MISR toolkit (https://nasa.github.io/MISR-Toolkit/html/index.html). Data was stored and computation was conducted on the University of Illinois, Urbana-Champaign School of Earth, Society, and Environment (SESE) computer cluster, Keeling. We are thankful to Catherine Moroney for providing the data used to calculate MISR stereo errors for near-surface retrievals. We also thank Puja Roy and Jesse Loveridge for the many fruitful discussions and inputs.

References

Baum, B. A., W. P. Menzel, R. A. Frey, D. C. Tobin, R. E. Holz, S. A. Ackerman, A. K. Heidinger, and P. Yang (2012), MODIS Cloud-Top Property Refinements for Collection 6, *Journal of Applied Meteorology and Climatology*, *51*(6), 1145–1163, doi:10.1175/jamc-d-11-0203.1.

Boucher, O., D. Randall, P. Artaxo, C. Bretherton, G. Feingold, P. Forster, V.-M. Kerminen, Y. Kondo, H. Liao, U. Lohmann, P. Rasch, S.K. Satheesh, S. Sherwood, B. Stevens, and X.Y. Zhang, 2013: Clouds and aerosols. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Doschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley, Eds. Cambridge University Press, pp. 571-657, doi:10.1017/CBO9781107415324.016.

Davies, R., Á. Horváth, C. Moroney, B. Zhang, and Y. Zhu (2007), Cloud motion vectors from MISR using sub-pixel enhancements, *Remote Sensing of Environment*, *107*(1-2), 194–199, doi:10.1016/j.rse.2006.09.023.

Davies, R., V. M. Jovanovic, and C. M. Moroney (2017), Cloud heights measured by MISR from 2000 to 2015, *Journal of Geophysical Research: Atmospheres*, *122*(7), 3975–3986, doi:10.1002/2017jd026456.

Dufresne, J.-L., and S. Bony (2008), An Assessment of the Primary Sources of Spread of Global Warming Estimates from Coupled Atmosphere–Ocean Models, *Journal of Climate*, *21*(19), 5135–5144, doi:10.1175/2008jcli2239.1.

Eastman, R., and S. G. Warren (2014), Diurnal Cycles of Cumulus, Cumulonimbus, Stratus, Stratocumulus, and Fog from Surface Observations over Land and Ocean, *Journal of Climate*, *27*(6), 2386–2404, doi:10.1175/jcli-d-13-00352.1.

Geiss, A., and R. Marchand (2019), Cloud responses to climate variability over the extratropical oceans as observed by MISR and MODIS, *Atmospheric Chemistry and Physics*, *19*(11), 7547–7565, doi:10.5194/acp-19-7547-2019.

Harshvardhan, G. Zhao, L. D. Girolamo, and R. Green (2009), Satellite-Observed Location of Stratocumulus Cloud-Top Heights in the Presence of Strong Inversions, *IEEE Transactions on Geoscience and Remote Sensing*, *47*(5), 1421–1428, doi:10.1109/tgrs.2008.2005406.

Holz, R. E., S. A. Ackerman, F. W. Nagle, R. Frey, S. Dutcher, R. E. Kuehn, M. A. Vaughan, and B. Baum (2008), Global Moderate Resolution Imaging Spectroradiometer (MODIS) cloud detection and height evaluation using CALIOP, *Journal of Geophysical Research*, *113*, doi:10.1029/2008jd009837.

Hong, Y., and L. Di Girolamo (2020), Cloud phase characteristics over Southeast Asia from A-Train satellite observations, *Atmospheric Chemistry and Physics*, *20*(13), 8267–8291, doi:10.5194/acp-20-8267-2020.

Horváth, Á. (2013), Improvements to MISR stereo motion vectors, *Journal of Geophysical*

Research: Atmospheres, *118*(11), 5600–5620, doi:10.1002/jgrd.50466.

Jovanovic, V., C. Moroney, and D. Nelson (2007), Multi-angle geometric processing for globally geo-located and co-registered MISR image data, *Remote Sensing of Environment*, *107*(1-2), 22–32, doi:10.1016/j.rse.2006.08.013.

Li, J., J. Huang, K. Stamnes, T. Wang, Q. Lv, and H. Jin (2015), A global survey of cloud overlap based on CALIPSO and CloudSat measurements, *Atmospheric Chemistry and Physics*, *15*(1), 519–536, doi:10.5194/acp-15-519-2015.

Lonitz, K., and Á. Horváth (2011), Comparison of MISR and Meteosat-9 cloud-motion vectors, *Journal of Geophysical Research: Atmospheres*, *116*(D24), doi:10.1029/2011jd016047.

Marchand, R. T., T. P. Ackerman, and C. Moroney (2007), An assessment of Multiangle Imaging Spectroradiometer (MISR) stereo-derived cloud top heights and cloud top winds using ground-based radar, lidar, and microwave radiometers, *Journal of Geophysical Research*, *112*(D6), doi:10.1029/2006jd007091.

Marchand, R. (2012), Spatial correlation of hydrometeor occurrence, reflectivity, and rain rate from CloudSat, *Journal of Geophysical Research: Atmospheres*, *117*(D6), doi:10.1029/2011jd016678.

Marchant, B., S. Platnick, K. Meyer, G. T. Arnold, and J. Riedi (2016), MODIS Collection 6 shortwave-derived cloud phase classification algorithm and comparisons with CALIOP, *Atmospheric Measurement Techniques*, *9*(4), 1587–1599, doi:10.5194/amt-9-1587-2016.

McGill, M. J., J. E. Yorks, V. S. Scott, A. W. Kupchock, and P. A. Selmer (2015), The Cloud-Aerosol Transport System (CATS): a technology demonstration on the International Space Station, *Lidar Remote Sensing for Environmental Monitoring XV*, doi:10.1117/12.2190841.

Menzel, W. P., R. A. Frey, and B. A. Baum (2015), Cloud Top Properties And Cloud Phase Algorithm Theoretical Basis Document Collection 006 Update. [Available at https://atmosphere-imager.gsfc.nasa.gov/sites/default/files/ModAtmo/MOD06-ATBD_2015_05_01_1.pdf],

Menzel, W. P., W. L. Smith, and T. R. Stewart (1983), Improved Cloud Motion Wind Vector and Altitude Assignment Using VAS, *Journal of Climate and Applied Meteorology*, *22*(3), 377–384, doi:10.1175/1520-0450(1983)022<0377:icmwva>2.0.co;2.

Menzel, W. P., R. A. Frey, H. Zhang, D. P. Wylie, C. C. Moeller, R. E. Holz, B. Maddux, B. A. Baum, K. I. Strabala, and L. E. Gumley (2008), MODIS Global Cloud-Top Pressure and Amount Estimation: Algorithm Description and Results, *Journal of Applied Meteorology and Climatology*, *47*(4), 1175–1198, doi:10.1175/2007jamc1705.1.

Moroney, C., R. Davies, and J.-P. Muller (2002), Operational retrieval of cloud-top heights using MISR data, *IEEE Transactions on Geoscience and Remote Sensing*, *40*(7), 1532–1540,

doi:10.1109/tgrs.2002.801150.

Mueller, K. J., C. P. Moroney, V. Jovanovic, J.-P. Muller, L. Di Girolamo, and R. Davies (2013), MISR Level 2 Cloud Product Algorithm Theoretical Basis. [Available at https://eospso.gsfc.nasa.gov/sites/default/files/atbd/MISR_L2_CLOUD_ATBD-1.pdf].,

Mueller, K. J., D. L. Wu, Á. M. Horváth, V. M. Jovanovic, J.-P. J. Muller, L. Di Girolamo, M. J. Garay, D. J. Diner, C. Moroney, and S. Wanzong (2017), Assessment of MISR Cloud Motion Vectors (CMVs) Relative to GOES and MODIS Atmospheric Motion Vectors (AMVs), *Journal of Applied Meteorology and Climatology*, *56*(3), 555–572, doi:10.1175/jamc-d-16-0112.1.

Muller, J.-P., A. Mandanayake, C. Moroney, R. Davies, D. Diner, and S. Paradise (2002), MISR stereoscopic image matchers: techniques and results, *IEEE Transactions on Geoscience and Remote Sensing*, *40*(7), 1547–1559, doi:10.1109/tgrs.2002.801160.

National Academies of Sciences, Engineering, and Medicine. 2018. *Thriving on Our Changing Planet: A Decadal Strategy for Earth Observation from Space.* Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/24938.

Naud, C., B. Baum, M. Pavolonis, A. Heidinger, R. Frey, and H. Zhang (2007), Comparison of MISR and MODIS cloud-top heights in the presence of cloud overlap, *Remote Sensing of Environment*, *107*(1-2), 200–210, doi:10.1016/j.rse.2006.09.030.

Naud, C., J.-P. Muller, and E. E. Clothiaux (2002), Comparison of cloud top heights derived from MISR stereo and MODIS CO2-slicing, *Geophysical Research Letters*, *29*(16), doi:10.1029/2002gl015460.

Naud, C., J.-P. Muller,M. Haeffelin, Y. Morille, and A. Delaval (2004), Assessment of MISR and MODIS cloud top heights through inter-comparison with a back-scattering lidar at SIRTA, *Geophys. Res. Lett.*, *31*, L04114, doi:10.1029/2003GL018976.

Ohring, G., B. Wielicki, R. Spencer, B. Emery, and R. Datla (2004), Satellite instrument calibration for measuring global climate change, *Bulletin of the American Meteorological Society*, *86*(9), 1303–1313, doi:10.6028/nist.ir.7047.

Pauly, R. M. et al. (2019), Cloud-Aerosol Transport System (CATS) 1064 nm calibration and validation, *Atmospheric Measurement Techniques*, *12*(11), 6241–6258, doi:10.5194/amt-12-6241-2019.

Platnick, S. et al. (2017), The MODIS Cloud Optical and Microphysical Products: Collection 6 Updates and Examples From Terra and Aqua, *IEEE Transactions on Geoscience and Remote Sensing*, *55*(1), 502–525, doi:10.1109/tgrs.2016.2610522.

Rajapakshe, C., Z. Zhang, J. E. Yorks, H. Yu, Q. Tan, K. Meyer, S. Platnick, and D. M. Winker (2017), Seasonally transported aerosol layers over southeast Atlantic are closer to underlying clouds than previously reported, *Geophysical Research Letters*, *44*(11), 5818–5825,

doi:10.1002/2017gl073559.

Smith, W. L., and C. M. R. Platt (1978), Comparison of Satellite-Deduced Cloud Heights with Indications from Radiosonde and Ground-Based Laser Measurements, *Journal of Applied Meteorology*, *17*(12), 1796–1802, doi:10.1175/1520-0450(1978)017<1796:cosdch>2.0.co;2.

Stubenrauch, C. J., and Coauthors (2013), Assessment of Global Cloud Datasets from Satellites: Project and Database Initiated by the GEWEX Radiation Panel, *Bulletin of the American Meteorological Society*, *94*(7), 1031–1049, doi:10.1175/bams-d-12-00117.1.

Van Diedenhoven, B., A. M. Fridlind, B. Cairns, A. S. Ackerman, and J. E. Yorks (2016), Vertical variation of ice particle size in convective cloud tops, *Geophysical Research Letters*, *43*(9), 4586–4593, doi:10.1002/2016gl068548.

Vaughan, M. A., D. M. Winker, and K. A. Powell (2005), CALIOP Algorithm Theoretical Basis Document Part 2: Feature Detection and Layer Properties Algorithms. PC-SCI-202 Part 2 Release 1.01. [Available at https://www-calipso.larc.nasa.gov/resources/pdfs/PC-SCI-202_Part2_rev1x01.pdf],

Vaughan, M. A., K. A. Powell, D. M. Winker, C. A. Hostetler, R. E. Kuehn, W. H. Hunt, B. J. Getzewich, S. A. Young, Z. Liu, and M. J. Mcgill (2009), Fully Automated Detection of Cloud and Aerosol Layers in the CALIPSO Lidar Measurements, *Journal of Atmospheric and Oceanic Technology*, *26*(10), 2034–2050, doi:10.1175/2009jtecha1228.1.

Wielicki, B. A., and J. A. Coakley (1981), Cloud Retrieval Using Infrared Sounder Data: Error Analysis., *Journal of Applied Meterology*, *20*(2), 157–169, doi:10.1175/1520-0450(1981)020<0157:CRUISD>2.0.CO;2.

Wylie, D. P., and W. P. Menzel (1989), Two Years of Cloud Cover Statistics Using VAS, *Journal of Climate*, *2*(4), 380–392, doi:10.1175/1520-0442(1989)002<0380:tyoccs>2.0.co;2.

Yorks, J. E., S. P. Palm, D. L. Hlavka, M. J. McGill, E. D. Nowottnick, P. D. Selmer, and W. undefined Hart (2016), The Cloud-Aerosol Transport System (CATS) algorithm theoretical basis document. [Available at https://cats.gsfc.nasa.gov/media/docs/CATS_ATBD.pdf],

Yorks, J. E., M. J. Mcgill, S. P. Palm, D. L. Hlavka, P. A. Selmer, E. P. Nowottnick, M. A. Vaughan, S. D. Rodier, and W. D. Hart (2016), An overview of the CATS level 1 processing algorithms and data products, *Geophysical Research Letters*, *43*(9), 4632–4639, doi:10.1002/2016gl068006.

Zelinka, M. D., D. A. Randall, M. J. Webb, and S. A. Klein (2017), Clearing clouds of uncertainty, *Nature Climate Change*, *7*(10), 674–678, doi:10.1038/nclimate3402.

Zhao, G., L. Di Girolamo, D. J. Diner, C. J. Bruegge, K. L. Mueller, and D. Wu (2016), Regional Changes in Earths Color and Texture as Observed From Space Over a 15-Year Period, *IEEE Transactions on Geoscience and Remote Sensing*, *54*(7), 4240–4249, doi:10.1109/tgrs.2016.2538723.

Tables and Figures

Figure 1. A collocation case from 14 March 2016 over South-East Asia, between MODIS, MISR and CATS. (a) MODIS and MISR CTH. The part of the CATS orbit that was nearly coincident with MISR and MODIS is in black. (b) MODIS RGB. (c) MODIS 1.38μm Reflectance (d) MODIS 11μm Brightness Temperature (e) The vertical cross-section of attenuated backscatter of cloud layers as detected from CATS. The approximate MISR and MODIS CTH are also plotted. All CTH are with respect to the WGS84 ellipsoid.

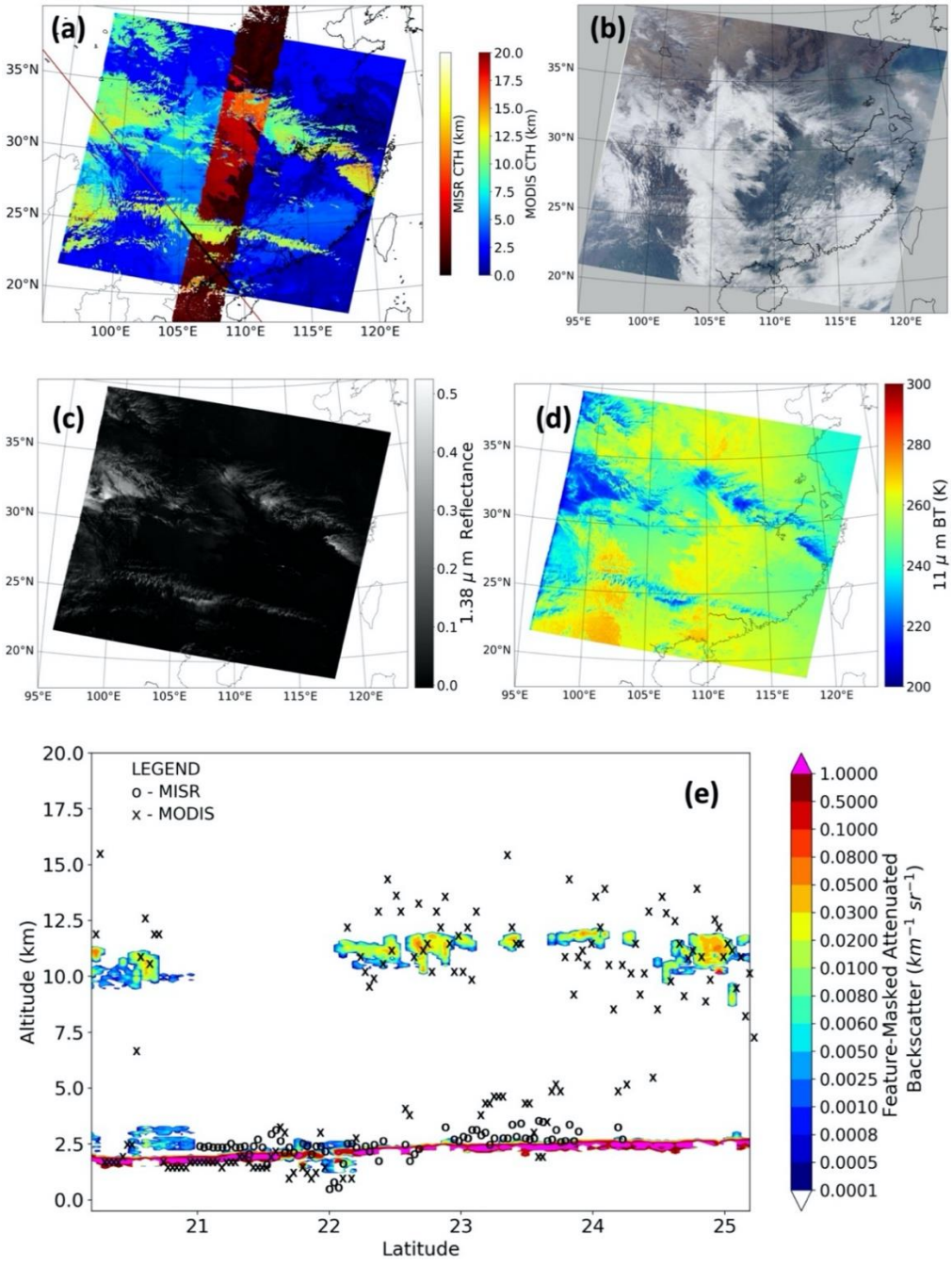

Figure 2. (a) A highly zoomed view of the collocation case from 14 March 2016, between MODIS, MISR and CATS, from Figure 1. MODIS geolocations are in blue, MISR in red. The black arrows signify the general direction and 5-km along-track extent of the CATS pixels, with the dark circles approximately signifying 1-km radii (not precisely to scale) from the CATS geolocation, within which the nearest neighboring MISR and MODIS pixels were searched; (b) Histograms of the standard deviation of MISR (purple) and MODIS (blue) CTH within each 1-km radii search windows for all successful collocation incidents for the year 2016; (c) Mean standard deviation of MISR (purple) and MODIS (blue) CTH for progressively bigger search radii, for all successful collocation incidents from 2016.

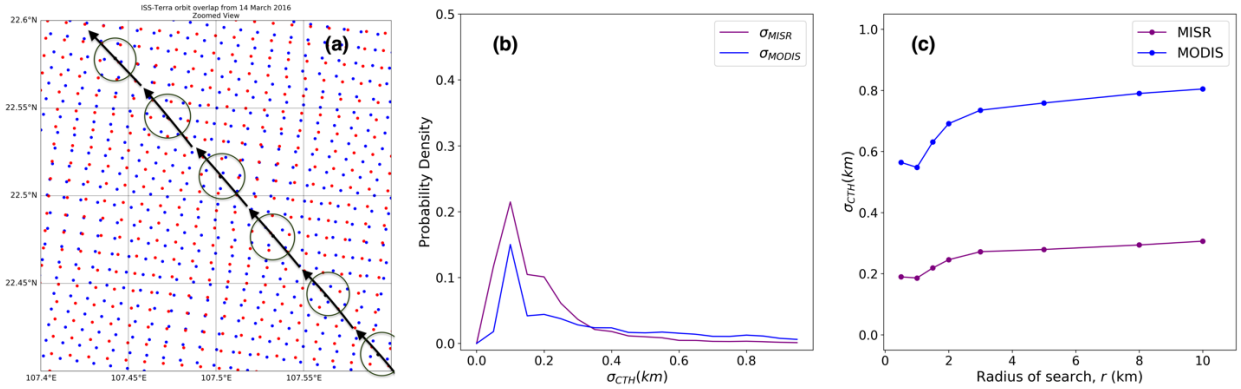

Figure 3. The global distribution of collocations of CATS, MISR and MODIS, where all three instruments recorded valid cloud top height retrievals for all of CATS operation. 18,986 individual collocated points have been plotted in the figure.

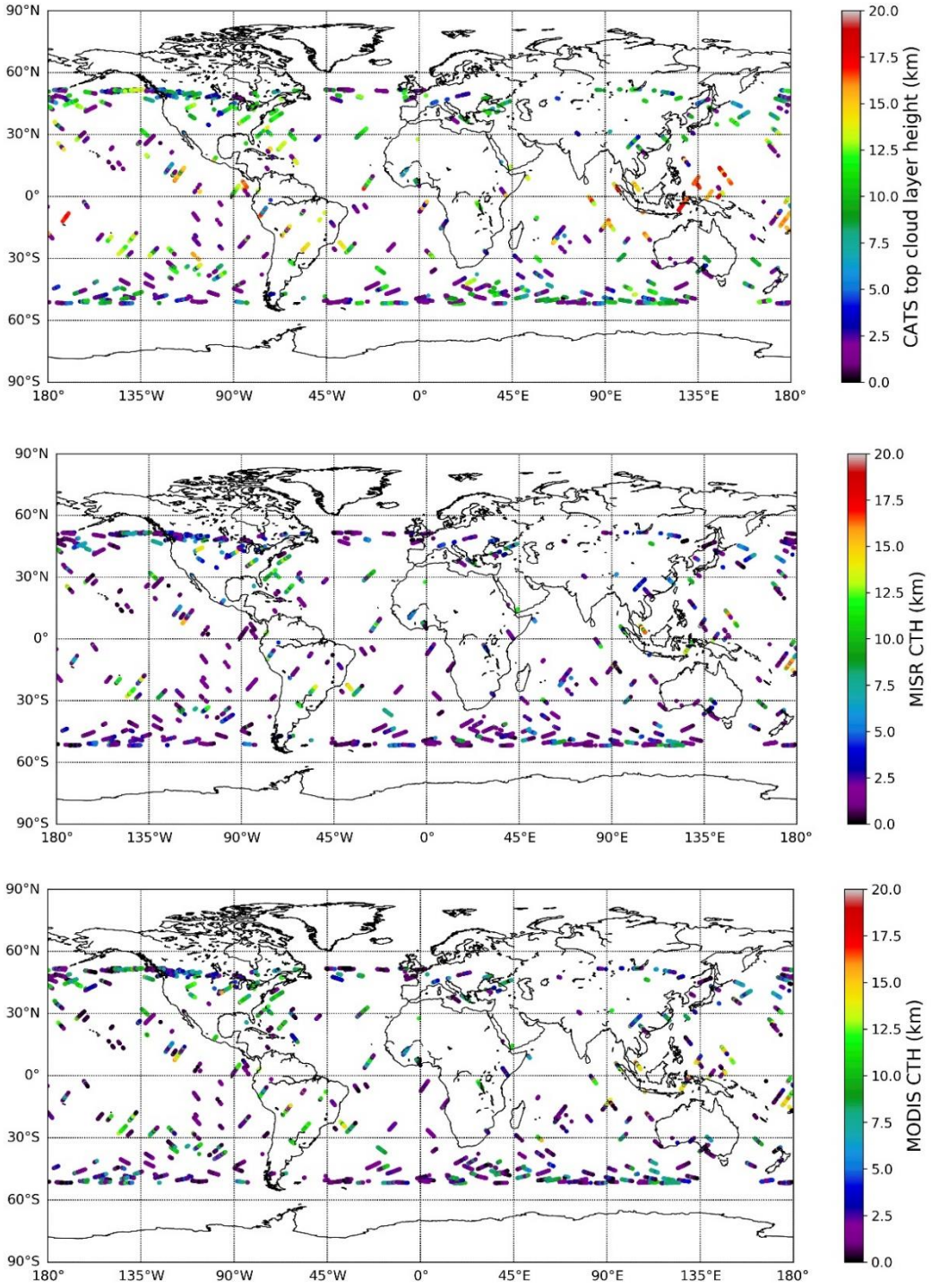

Figure 4. Latitudinal distribution of median CTH differences for CATS-MODIS (top left), CATS-MISR (top right), and MODIS-MISR (bottom left) for all clouds (black), CATS single-layered clouds (red) and CATS multi-layered clouds (blue). The number of samples in each latitudinal bin is shown in the bottom right panel. Error bars in the first three subplots represent the median absolute deviation statistic.

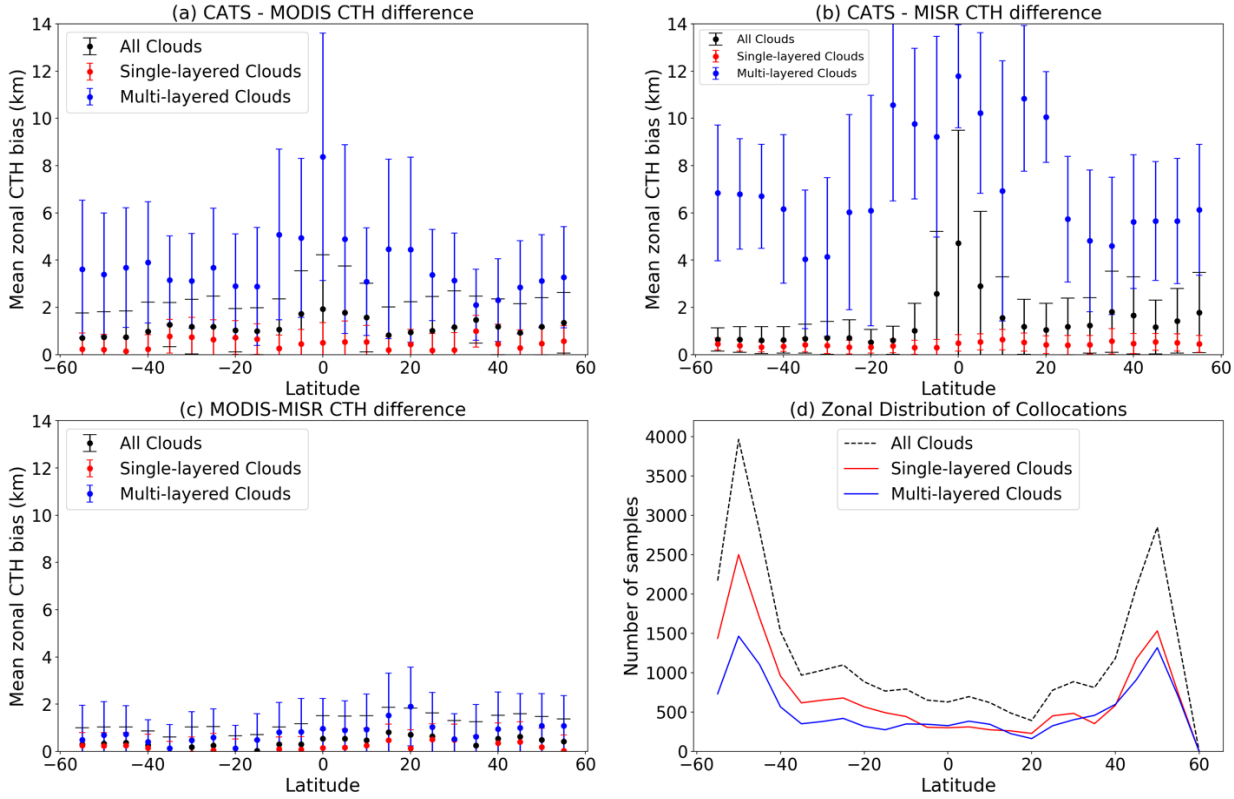

Figure 5. Histograms of global CTH differences for high clouds (CATS CTH > 5 km, purple) and low clouds (CATS CTH < 5 km, blue), with 100 bins between +20 km and -20 km (top panels) and +5 km and -5 km (bottom panels) The overall distribution is marked by a black dashed line and contains 18986 collocated data points, out of which 10315 were high clouds. For each histogram bin, the average CATS top layer height has been marked in an inverted system of axes in red, with each point denoting the CATS layer-integrated backscatter (γ) for the topmost cloud layer.

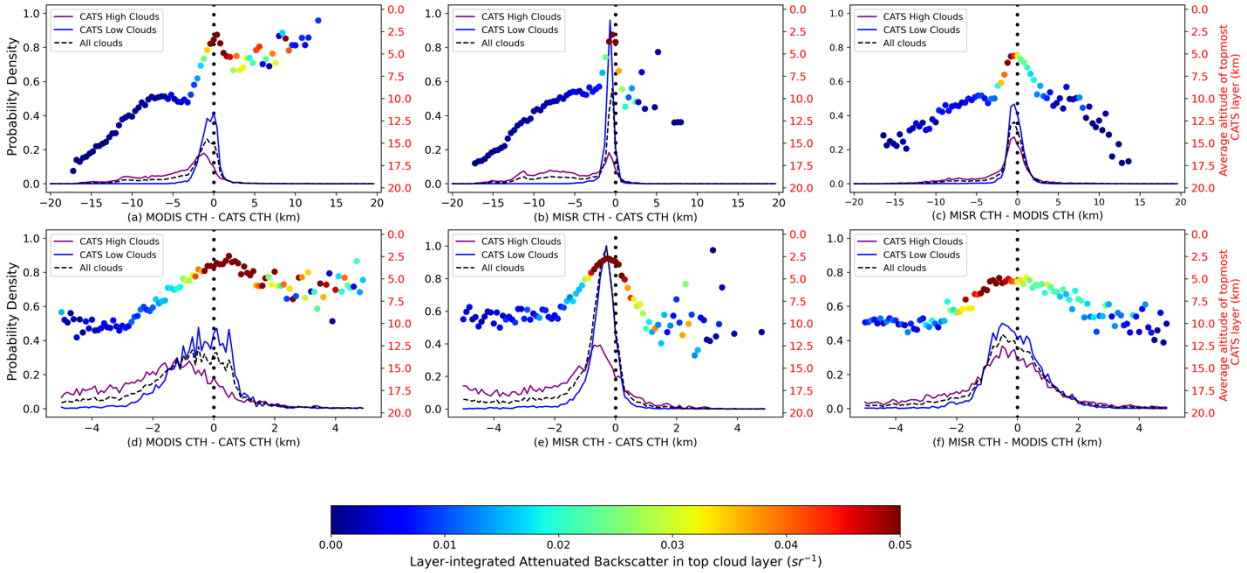

Figure 6. Histograms of global CTH differences for (a-c) high clouds and (d-f) low clouds from CATS, for optically thick top layer of cloud ($\gamma > 0.02$ sr^{-1}, purple) and optically thin top layer of cloud ($\gamma < 0.02$ sr^{-1}, blue), with the black dashed line denoting overall distributions.

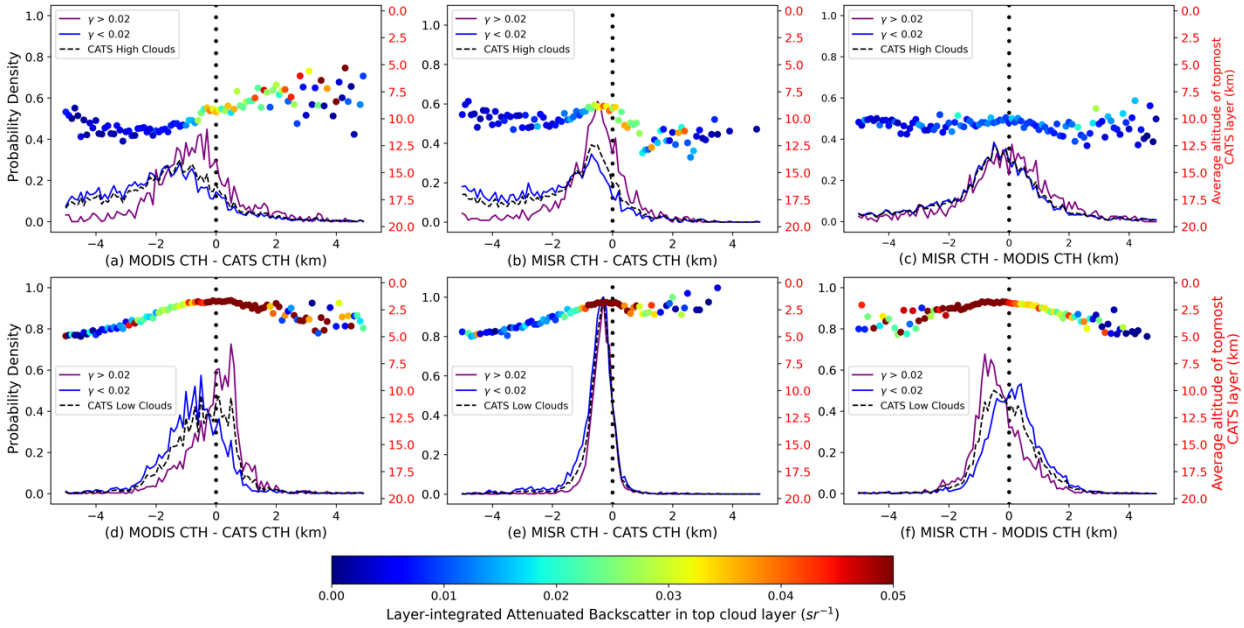

Figure 7. Histograms of global CTH differences for CATS single-layered clouds (CATS Percentage Opacity < 50%, purple) and multi-layered clouds (CATS detected at least two layers, blue), with the CATS top layer height being greater than 5 km.

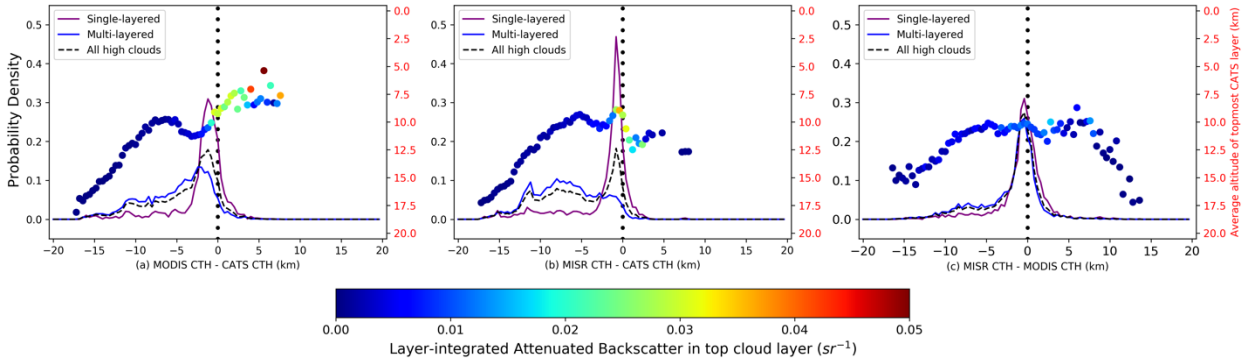

Figure 8. Histograms of global CTH differences for double-layered clouds from CATS, for each passive sensor – (a) MISR and (b) MODIS - and the first layer (blue) and second layer (red) of CATS clouds, respectively, with 100 bins between +20 km and -20 km. The distribution of the difference in cloud top height from each passive sensor and the closest CATS layer is given by a black dashed line and contains 7454 collocated data points.

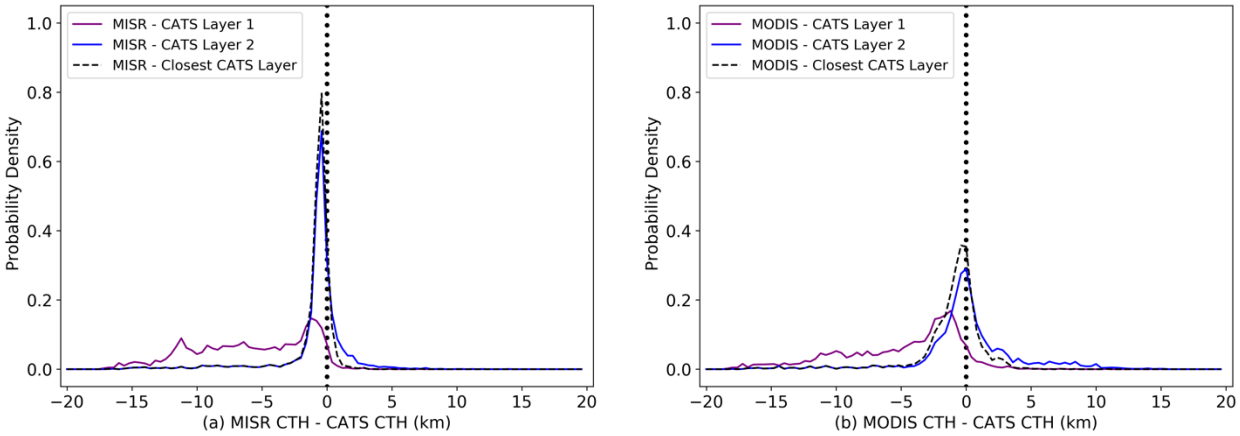

Table 1. MODIS and MISR CTH bias and precision (rounded to the nearest multiple of 10) with respect to CATS, summarizing Sections 4.2-4.4. In each row, MISR and MODIS errors are probed by imposing conditions on a cloud "parameter of interest" (e.g., top-layer height), thus extracting from our dataset a subset of scenes that is representative of a "type of cloud" (e.g., high/low).

Parameter of Interest	Type of Cloud	MODIS		MISR	
		Bias (m)	Precision (m)	Bias (m)	Precision (m)
Topmost Cloud-Layer Height	High (CATS CTH > 5km)	-1200	1080	-540	590
	Low (CATS CTH < 5km)	40	730	-320	250
Topmost Cloud OD (High clouds)	Optically thick ($\gamma > 0.02$ sr^{-1})	-280	730	-440	470
	Optically thin ($\gamma < 0.02$ sr^{-1})	-1160	1020	-680	550
Topmost Cloud OD (Low clouds)	Optically thick ($\gamma > 0.02$ sr^{-1})	500	430	-280	260
	Optically thin ($\gamma < 0.02$ sr^{-1})	-440	600	-320	310
Cloud Overlap	Single-layered (High)	-1160	510	-720	460
	Multi-layered (Highest layer)	-2380	1030	N/A*	N/A*
	Multi-layered (Top Layer Closest)	-1200	1190	-820	850
	Multi-layered (Bottom Layer Closest)	20	850	-400	350

*Distribution does not resemble Gaussian.

Table 2. MISR and MODIS bias and precision (rounded to nearest multiple of 10) for all, high, mid-level and low clouds, as seen by CATS, with absolute CTH difference with respect to CATS ≤ 2.5 km and CATS Percentage Opacity = 1.

Instrument	Overall		High (CATS CTH > 10 km)		Mid-level (10 km > CTH > 5 km)		Low (CATS CTH < 5 km)	
	Bias (m)	**Precision (m)**	**Bias (m)**	**Precision (m)**	**Bias (m)**	**Precision (m)**	**Bias (m)**	**Precision (m)**
MISR	-280	370	-300	400	-370	400	-240	300
MODIS	-540	690	-950	740	-350	690	60	660

Figure 9. Distribution with altitude of (a) MISR and (b) MODIS CTH bias and precision (1σ error-bars) for CATS single-layered clouds with Percentage Opacity = 1 and an absolute CTH difference <= 2.5 km. The results are binned every 2 km (bin centers are odd integers), with mean CATS integrated backscatter for the top 300 m into the cloud (γ_{300} in sr^{-1}) shown in red. Green dotted lines in (b) denote the 75th-percentile CATS CTH for scenes employing IR BT (left line) and CO_2-slicing, respectively. Each bin has a minimum of 150 samples.

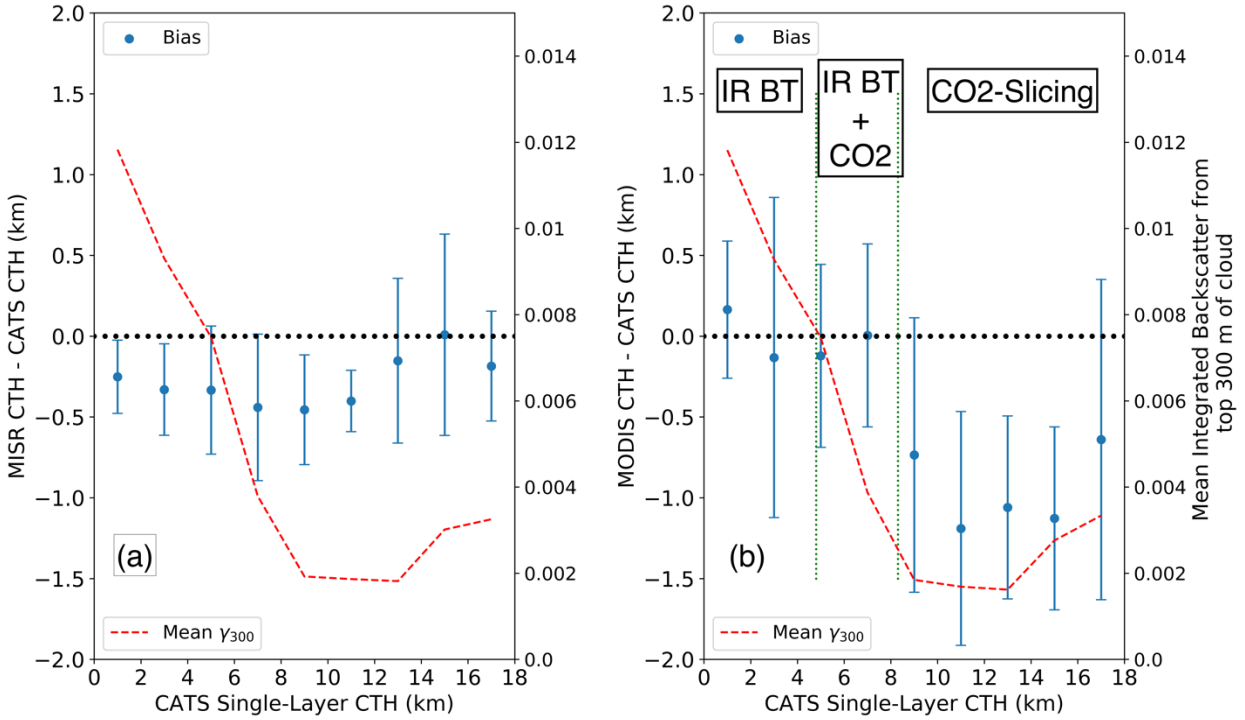

Figure 10. Histograms of global CTH differences for CATS high (CTH > 5 km) clouds (upper panel) and CATS single-layered high clouds (lower panel). CO_2 retrievals are in purple and 11-μm brightness temperature retrievals are in blue. For each histogram bin of CTH difference, the average CATS top layer height has been marked in an inverted system of axes in red, with each point denoting the layer-integrated CATS backscatter for the topmost cloud layer.