Technical note: Applicability of physics-based and machine-learning-based algorithms of geostationary satellite in retrieving the diurnal cycle of cloud base height

Mengyuan Wang[1], Min Min[1]*, Jun Li[2], Han Lin[3], Yongen Liang[1], Binlong Chen[2], Zhigang Yao[4], Na Xu[2], Miao Zhang[2]

[1]School of Atmospheric Sciences, Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), and Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies, Zhuhai 519082, China

[2]Key Laboratory of Radiometric Calibration and Validation for Environmental Satellites and Innovation Center for FengYun Meteorological Satellite (FYSIC), National Satellite Meteorological Center (National Center for Space Weather), China Meteorological Administration, Beijing 100081, China

[3]Key Laboratory of Spatial Data Mining and Information Sharing of Ministry of Education, National and Local Joint Engineering Research Center of Satellite Geospatial Information Technology, Fuzhou University, Fuzhou 350108, China

[4]Beijing Institute of Applied Meteorology, Beijing 100029, China

Correspondence to: Min Min (minm5@mail.sysu.edu.cn)

Abstract. Two groups of retrieval algorithms, physics-based and the other machine-learning (ML) based, each consisting of two independent approaches, have been developed to retrieve cloud base height (CBH) and its diurnal cycle from Himawari-8 geostationary satellite observations. Validations have been conducted using the joint CloudSat/CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) CBH products in 2017, ensuring independent assessments. Results show that the two ML-based algorithms exhibit markedly superior performance (the optimal method is with a correlation coefficient of $R > 0.91$ and an absolute bias of approximately 0.8 km) compared to the two physics-based algorithms. However, validations based on CBH data from the ground-based lidar at the Lijiang station in Yunnan province and the cloud radar at the Nanjiao station in Beijing, China, explicitly present contradictory outcomes ($R < 0.60$). An identifiable issue arises with significant underestimations in the retrieved CBH by both ML-based algorithms, leading to an inability to capture the diurnal cycle characteristics of CBH. The strong consistence observed between CBH derived from ML-based algorithms and the spaceborne active sensors of CloudSat/CALIOP may be attributed to utilizing the same dataset for training and validation, sourced from the CloudSat/CALIOP products. In contrast, the CBH derived from the optimal physics-based algorithm demonstrates the good agreement in diurnal variations of CBH with ground-based lidar/cloud radar observations during the daytime (with an R value of approximately 0.7). Therefore, the findings in this investigation from ground-based observations advocate for the more reliable and adaptable nature of physics-based algorithms in retrieving CBH from geostationary satellite measurements. Nevertheless, under ideal conditions, with an ample dataset of spaceborne cloud profiling radar observations encompassing the entire day for training purposes, the ML-based algorithms may hold promise in still delivering accurate CBH outputs.

Key words: Geostationary meteorological satellite; cloud base height; physics-based algorithm; machine learning.

1 Introduction

Clouds, comprising visible aggregates like atmospheric water droplets, supercooled water droplets, ice crystals, etc., cover roughly 70% of the Earth's surface (Stubenrauch et al., 2013). They play a pivotal role in global climate change, the hydrometeor cycle, aviation safety, and serve as a primary focus in weather forecasting and climate research, particularly storm clouds (Hansen, 2007; Hartmann and Larson, 2002). From advanced geostationary (GEO) and polar-orbiting (LEO, low earth orbit) satellite imagers, various measurable cloud properties, such as cloud fraction, cloud phase, cloud top height (CTH), and cloud optical thickness (D_{COT}), are routinely retrieved. However, the high-quality cloud geometric height (CGH) and cloud base height (CBH), a fundamental macro physical parameter delineating the vertical distribution of clouds, remains relatively understudied and underreported. Nonetheless, for boundary-layer clouds, the cloud base height stands as a critical parameter depending on other cloud-controlling variables. These variables encompass the cloud base temperature (Zhu et al., 2014), cloud base vertical velocity (Zheng et al., 2020), activation of CCN (Cloud Condensation Nuclei) at the cloud base (Rosenfeld et al., 2016; Miller et al., 2023), and the cloud-surface decoupling state (Su et al., 2022). These factors significantly impact convective cloud development and ultimately the climate.

There are distinct diurnal cycle characteristics of clouds in different regions across the globe (Li et al., 2022). These diurnal cycle characteristics primarily stem from the daily solar energy cycle absorbed by both the atmosphere and Earth's surface. Besides, vertical atmospheric motions are shaped by imbalances in atmospheric heating and surface configurations, also leading to a range of cloud movements and structures (Miller et al., 2018). Cloud base plays a pivotal role in weather and climate processes. It is critical for predicting fog and cloud-related visibility issues important in aviation and weather forecasting. For instance, lower cloud bases often lead to more intense rainfall. In climate modeling, CBH is integral for accurate long-term weather predictions and understanding the radiative balance of the Earth, which influences global temperatures (Zheng and Rosenfeld, 2015). Hence, the accurate determination of CBH and its diurnal cycle with high spatiotemporal resolution becomes very important, necessitating comprehensive investigations (Viúdez-Mora et al., 2015; Wang et al., 2020). Such efforts can provide deeper insights into potential ramifications of clouds on radiation equilibrium and global climate systems.

However, as one of the most crucial cloud physical parameters in atmospheric physics, the CBH poses challenges in terms of measurement or estimation from space. Presently, the primary methods for measuring CBH rely on ground-based observations, utilizing tools such as sounding balloons, Mie-scattering lidars, stereo-imaging cloud-height detection technologies, and cloud probe sensors (Forsythe et al., 2000; Hirsch et al., 2011; Seaman et al., 2017; Zhang et al., 2018; Zhou et al., 2019; Zhou et al., 2024). While *in-situ* ground-based observation methods offer highly accurate, reliable, and timely continuous CBH results, they are constrained by localized observation coverage and the sparse distribution of observation sites (Aydin and Singh, 2004). In recent decades, with the rapid advancement of meteorological satellite observation technology, spaceborne observing methods have emerged that provide global cloud observations with high spatiotemporal resolution compared to conventional ground-based remote sensing methods. In this realm, satellite remote sensing techniques for measuring CBH fall primarily into two categories: active and passive methods. Advanced active remote sensing technologies like CloudSat (Stephens et al., 2002) and Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) (Winker et al., 2009) in the National Aeronautics and Space Administration (NASA) A-Train (Afternoon-Train) series (Stephens et al., 2002) can capture global cloud profiles, including CBH, with high quality by detecting unique return signals from cloud layers using onboard active millimeter wave radar or lidar. However, their viewing footprints are limited along the nadir of the orbit, implying that observation coverage remains confined primarily to a horizontal scale (Min et al., 2022; Lu et al., 2021).

In addition to active remote sensing methods, satellite-based passive remote sensing technologies can also play an important role in estimating CBH (Meerkötter and Bugliaro, 2009; Lu et al., 2021). The physics-based principles and retrieval methods for CTH have reached maturity and are now widely employed in satellite passive remote sensing field (Heidinger and Pavolonis, 2009; Wang et al., 2022). However, the corresponding physical principles or methods for measuring CBH using satellite passive imager measurements are still not entirely clear and unified (Heidinger et al., 2019; Min et al., 2020). A recent study by (Yang et al., 2021) utilized oxygen A-band data observed by the Orbiting Carbon Observatory 2 (OCO-2) to retrieve single-layer marine liquid CBH. These passive space-based remote sensing methods aforementioned, such as satellite imagery, play a key role in retrieving CBH. In terms of detection principles, the first method involves the extrapolation technique for retrieving CBH for clouds of the same type. For instance, (Wang et al., 2012) proposed a method to extrapolate CBH from CloudSat using spatiotemporally matched MODIS (Moderate Resolution Imaging Spectroradiometer) cloud classification data (Baum et al., 2012; Platnick et al., 2017). The second physics-based retrieval method first approximates the cloud geometric thickness using its optical thickness. It then employs the previously derived CTH product to compute the corresponding CBH using the respective NOAA (National Oceanic and Atmospheric Administration) SNPP/VIIRS (Suomi National Polar-orbiting Partnership/Visible Infrared Imaging Radiometer Suite) products (Noh et al., 2017). Hutchison et al. (2002 and 2006) also formulated an empirical algorithm that estimates both cloud geometric thickness (CGT) and CBH. This algorithm relies on statistical analyses derived from MODIS D_{COT} and cloud liquid water path products (Hutchison et al., 2006; Hutchison, 2002).

Machine learning (ML) has proven to be highly effective in addressing nonlinear problems within remote sensing and meteorology fields, such as precipitation estimation and CTH retrieval (Min et al., 2020; HåKansson et al., 2018; Kühnlein et al., 2014). In recent years, several studies have leveraged ML-based algorithms to retrieve CBH, establishing nonlinear connections between CBH and GEO satellite observations. For instance, Tan et al. (2020) integrated CTH and cloud optical properties products from Fengyun-4A (FY-4A) GEO satellite with spatiotemporally matched CBH data from CALIPSO/CloudSat. They developed a random forest (RF) model for CBH retrieval. Similarly, Lin et al. (2022) constructed a gradient boosted regression tree (GBRT) model using U.S. new-generation Geostationary Operational Environmental Satellites-R Series (GOES-R) Advanced Baseline Imager (ABI) level 1B radiance data and the ERA5 (the fifth generation ECMWF) reanalysis dataset (Lin et al., 2022; Hersbach et al., 2020) (https://cds.climate.copernicus.eu/cdsapp#!/search?type=dataset). They employed CALIPSO CBH data as labels to achieve single-layer CBH retrievals. Notably, the CBH quality of ML-based algorithms was found to surpass that of physics-based algorithms (Lin et al., 2022). Moreover, Tana et al. (2023) utilized Himawari-8 data and the RF algorithm to develop a novel CBH algorithm, achieving a similar high correlation coefficient (R) of 0.92 and a low root mean square error (RMSE) of 1.17 km compared with CloudSat/CALISPO data.

However, these former studies did not discuss whether both physics-based and ML-based algorithms of GEO satellite could retrieve the diurnal cycle of CBH well. This gap in research could be mainly attributed to potential influences from the fixed LEO satellite (with active radar or lidar) passing time in the previous CBH retrieval model (Lin et al., 2022). The diurnal cycles of CBH have not been well investigated in both GEO and LEO remote sensing research. Hence, it is crucial to thoroughly investigate the diurnal cycle features of CBH derived from GEO satellite measurements by comparing them with ground-based radar and lidar observations (Min and Zhang, 2014; Warren and Eastman, 2014). In this study, we aim to assess the applicability and feasibility of both physics-based and ML-based algorithms of GEO satellites in capturing the diurnal cycle characteristics of CBH.

The subsequent sections of this paper are structured as follows. Section 2 provides a concise overview of the data employed in this study. Following that, section 3 introduces the four distinct physics/ML-based CBH retrieval algorithms. In section 4, the CBH results obtained from these four algorithms are analyzed, and comparisons are drawn with spatiotemporally matched CBHs from ground-based cloud radar and lidar. Finally, section 5 encapsulates the primary conclusions and new findings derived from this study.

2 Data

In this study, observations from the Himawari-8 (H8) Advanced Himawari Imager (AHI) are utilized for the retrieval of high spatiotemporal resolution CBH. Launched successfully by the Japan Meteorological Administration on October 7, 2014, the H8 geostationary satellite is positioned at 140.7°E. The AHI onboard H8 encompasses 16 spectral bands ranging from 0.47 μm to 13.3 μm, featuring spatial resolutions of 0.5–2 km. This includes 3 visible (VIS) bands at 0.5–1 km, 3 near-infrared (NIR) bands at 1–2 km, and 10 infrared (IR) bands at 2 km. The H8/AHI can scan a full disk area within 10 minutes, two specific areas within 2.5 minutes, a designated area within 2.5 minutes, and two landmark areas within 0.5 minutes (Iwabuchi et al., 2018). Its enhanced temporal resolution and observation frequency facilitate the tracking of rapidly changing weather systems, enabling the accurate determination of quantitative atmospheric parameters (Bessho et al., 2016).

Operational H8/AHI Level-1B data, accessible from July 7, 2015, are freely available on the satellite product homepage of the Japan Aerospace Exploration Agency (Letu et al., 2019). The Level-2 cloud products utilized in this study, including cloud mask (CLM), CTH, cloud effective particle radius (CER or R_{eff}), and D_{COT}, are generated by the Fengyun satellite science product algorithm testbed (FYGAT) (Wang et al., 2019; Min et al., 2017) of the China Meteorological Administration (CMA) for various applications. According to previous CALIPSO validations (Min et al., 2020), the absolute bias of cloud top height retrieved by the H8 satellite is approximately 3 km, with an absolute bias of 1 to 2 km for samples below 5 km. The accuracy of CTH is crucial for estimating CBH in the subsequent algorithm. It is important to note that certain crucial preliminary cloud products, such as CLM, have been validated in prior studies (Wang et al., 2019; Liang et al., 2023). Nevertheless, before initiating CBH retrieval, it is imperative to validate the H8/AHI cloud optical and microphysical products from the FYGAT retrieval system. This validation has been carried out by using analogous MODIS Level-2 cloud products as a reference. Additional details regarding the validation of cloud products are provided in the Appendix A section.

In addition to the H8/AHI Level-1/2 data, the Global Forecast System (GFS) numerical weather prediction (NWP) data are employed for CBH retrieval in this study. The variables include land/sea surface temperature and the vertical profiles of temperature, humidity, and pressure. Operated by the U.S. NOAA (Kalnay et al., 1996), the GFS serves as a global and advanced NWP system. The operational GFS system routinely delivers global high-quality and gridded NWP data at 3-hour intervals, with four different initial forecast times per day (00:00, 06:00, 12:00, and 18:00 UTC). The three-dimensional NWP data cover the Earth in a 0.5°×0.5° grid interval and resolve the atmosphere with 26 vertical levels from the surface (1000 hPa) up to the top of the atmosphere (10 hPa).

As previously mentioned, the official MODIS Collection-6.1 Level-2 cloud product Climate Data Records (Platnick et al., 2017) are utilized in this study to validate the H8/AHI cloud products (CTH, CER, and D_{COT}) generated by the FYGAT system. High-quality, long-term series MODIS data is often used as a validation reference to evaluate the products of new satellites. MODIS sensors are onboard

NASA Terra and Aqua polar-orbiting satellites. Terra functions as the morning satellite, passing through the equator from north to south at approximately 10:30 local time, while Aqua serves as the afternoon satellite, traversing the equator from south to north at around 13:30 local time. As a successor to the NOAA Advanced Very High Resolution Radiometer (AVHRR), MODIS features 36 independent spectral bands and a broad spectral range from 0.4 μm (VIS) to 14.4 μm (IR), with a scanning width of 2330 km and spatial resolutions ranging from 0.25 to 1.0 km. Recent studies (Baum et al., 2012; Platnick et al., 2017) have highlighted significant improvements and collective changes in cloud top, optical, and microphysical properties from Collection-5 to Collection-6.

In addition to the passive spaceborne imaging sensors mentioned above, the CloudSat satellite , equipped with a 94-GHz active cloud profiling radar (CPR), holds the distinction of being the first sun-synchronous orbit satellite specifically designed to observe global cloud vertical structures and properties. It is part of the A-Train series of satellites, akin to the Aqua satellite, launched and operated by NASA (Heymsfield et al., 2008). CALIPSO is another polar-orbiting satellite within the A-Train constellation, sharing an orbit with CloudSat and trailing it by a mere 10–15 seconds. CALIPSO is the first satellite equipped with an active dual-channel CALIOP at 532 and 1064 nm bands (Hunt et al., 2009). Both CloudSat and CALIPSO possess notable advantages over passive spaceborne sensors due to the 94-GHz radar of CloudSat and the joint return signals of lidar and radar on CALIPSO. These features enhance their sensitivity to optically thin cloud layers and ensure strong penetration capability, resulting in more accurate CTH and CBH detections compared to passive spaceborne sensors (CAL_LID_L2_05kmCLay-Standard-V4-10). The joint cloud type products of 2B-CLDCLASS-LIDAR, derived from both CloudSat and CALIPSO measurements, offer a comprehensive description of cloud vertical structure characteristics, cloud type, CTH, CBH, etc. The time interval between each profile in this product is approximately 3.1 seconds, and the horizontal resolution is 2.5 km (along track)×1.4 km (cross-track). Each profile is divided into 125 layers with a 240-m vertical interval. For more details on 2B-CLDCLASS-LIDAR products, please refer to the CloudSat official product manual (Sassen and Wang, 2008). In this study, we consider the lowest effective cloud base height from the joint CloudSat/CALIOP data as the true values for training and validation. Please note that for this study, we utilized one-year H8/AHI data and matched it with the joint CloudSat/CALIOP data from January 1 to December 31 of 2017.

3 Physics and machine-learning based cloud base height algorithms

3.1 GEO cloud base height retrieval algorithm from the interface data processing segment of the Visible Infrared Imaging Radiometer Suite

The Joint Polar Satellite System (JPSS) program is a collaborative effort between NASA and NOAA. The operational CBH retrieval algorithm, part of the 30 Environmental Data Records (EDR) of JPSS, can be implemented operationally through the Interface Data Processing Segment (IDPS) (Baker, 2011). In this study, our geostationary satellite CBH retrieval algorithm aligns with the IDPS CBH algorithm developed by (Baker, 2011). Utilizing the geostationary H8/AHI cloud products discussed earlier, this new GEO CBH retrieval algorithm is succinctly outlined below. It is important to note that multilayer cloud scenes remain a challenge for retrieving both CTH and CBH, especially when considering the column-integrated cloud water path (CWP) used in physics-based algorithms (Noh et al., 2017). In this study, we will simplify the scenario by assuming a single-layer cloud for all algorithms.

The new GEO IDPS CBH algorithm initiates the process by first retrieving the CGT from bottom to top. Subsequently, CGT is subtracted from the corresponding CTH to calculate CBH (CBH = CTH − CGT). The algorithm is divided into two independent executable modules based on cloud phase, distinguishing between liquid water and ice clouds. CBH of water cloud retrieval requires D_{COT} and CER as inputs. For ice clouds, an empirical equation is employed for CBH retrieval. However, the standard deviations of error in IDPS CBH for individual granules often exceed the JPSS VIIRS minimum uncertainty requirement of ±2km (Noh et al., 2017). For a more comprehensive understanding of this CBH algorithm, please refer to the IDPS algorithm documentation (Baker, 2011). Note that, similar to previous studies on cloud retrieval (Noh et al., 2017; Platnick et al., 2017), this investigation also assumes a single-layer cloud for all CBH algorithms, due to the challenges associated with determining multilayer cloud structures.

3.2 GEO cloud base height retrieval algorithm implemented in the Clouds from Advanced Very High Resolution Radiometer Extended system

As mentioned above, the accuracy of the GEO IDPS algorithm is highly dependent on the initial input parameters such as cloud phase, D_{COT} and R_{eff}, which may introduce some uncertainties in the final retrieval results. In contrast, another statistically-based algorithm is proposed and implemented here, which is named the GEO CLAVR-x (Clouds from AVHRR Extended, NOAA's operational cloud processing system for the AVHRR) CBH algorithm (Noh et al., 2017), and it mainly refers to NOAA AWG CBH algorithm (ACBA) (Noh et al., 2022). Previous studies have also demonstrated a R of 0.569 and a RMSE of 2.3 km for the JPSS VIIRS CLAVR-x CBH algorithm. It is anticipated that this algorithm will also be employed for the NOAA GOES-R geostationary satellite imager (Noh et al., 2017; Seaman et al., 2017).

Similar to the GEO IDPS CBH retrieval algorithm mentioned earlier, the GEO CLAVR-x CBH retrieval algorithm also initially obtains CGT and CTH, subsequently calculating CBH by subtracting CGT from CTH (CTH−CGT). However, the specific calculation method for the CGT value differs. This algorithm is suitable for single-layer and the topmost layer of multi-layer clouds, computing CBH using the CTH at the top layer of the cloud. In comparison with the former GEO IDPS CBH algorithm, the GEO CLAVR-x CBH algorithm considers two additional cloud types: deep convection clouds and thin cirrus clouds (Baker, 2011). For more details on this CLAVR-x CBH algorithm, please refer to the original algorithm documentation (Noh et al., 2017).

3.3 Random-forest-based cloud base height estimation algorithm

RF, one of the most significant ML algorithms, was initially proposed and developed by (Breiman, 2001). It is widely employed to address classification and regression problems based on the law of large numbers. The RF method is well-suited for capturing complex or nonlinear relationships between predictors and predictands.

In this study, two distinct ML-based GEO CBH algorithms, namely VIS+IR and IR-single (only uses observations of H8/AHI IR channels), are devised to retrieve or predict the CBH using different sets of predictors. The RF training of the chosen predictors is formulated as follows:

$$\text{CBH}=RF_{\text{reg}}[x_1, x_2, \ldots, x_n], \quad (1)$$

where RF_{reg} denotes the regression RF model, and x_i represents the ith predictor. The selected predictors from H8/AHI for both the VIS+IR and IR RF model training and prediction are detailed in Table 1, mainly referencing Min et al. (2020) and Tan et al. (2020). The VIS+IR algorithm retrieves CBH using NWP data (atmospheric temperature and altitude profiles, total precipitable water (TPW), surface temperature), surface elevation, air mass 1 (air mass 1=1/cos(view zenith angle)), and air mass 2 (air mass 2=1/cos(solar zenith angle)). The rationale for choosing air mass and TPW is their ability to account for the potential absorption effect of water vapor along the satellite viewing angle. The predictors in CBH retrieval also include the IR band Brightness Temperature (BT) and VIS band reflectance. The IR-single algorithm selects the same GFS NWP data as the VIS+IR algorithm but employs only view zenith angles and azimuth angles.

To optimize the RF prediction model, the hyperparameters of the RF model are tuned individually. The parameters and their dynamic ranges involved in tuning the RF prediction models include the number of trees [100, 200, 300, 400, 500], the maximum depth of trees [10, 20, 30, 40, 50], the minimum number of samples required to split an internal node [2, 4, 6, 8, 10], and the minimum number of samples required to be at a leaf node [1, 3, 5, 7, 9]. In this study, we set the smallest number of trees in the forest to 100 and the maximum depth of the tree to 40.

3.4 Evaluation method

The performance of RF models and physics-based methods will be assessed using mean absolute error (MAE), mean bias error (MBE), RMSE, R, and standard deviation (STD) scores using the testing dataset. These scores mentioned above are used to understand different aspects of the predictive performance of model: MAE and RMSE provide insights into the average error magnitude, MBE indicates bias in the predictions, R evaluates the linear association between observed and predicted values, and STD assesses the variability of the predictions. In the RF IR-single algorithm, 581,783 matching points are selected from H8/AHI and CloudSat data for 2017. Seventy percent of these points are randomly assigned to the training dataset, and the remainder serves as the testing dataset. For the RF VIS+IR algorithm, a total of 418,241 matching points are chosen, with 70% randomly allocated to the training set. Note that the reduced data amount is because only daytime data can be used for the VIS+IR method training. It's important to note that the two training datasets in

CloudSat will also be used to verify the CBHs obtained by cloud radar and lidar. The statistical formulas for evaluation are as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - x_i|, \quad (2)$$

$$\mathrm{MBE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i), \quad (3)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2}, \quad (4)$$

$$R = \frac{\sum_{i=1}^{n}(y_i - y)(x_i - x)}{\sqrt{\sum_{i=1}^{n}(y_i - y)^2}\sqrt{\sum_{i=1}^{n}(x_i - x)^2}}, \quad (5)$$

$$\mathrm{STD} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - x)^2}, \quad (6)$$

where n is the sample number, y_i is the ith CBH retrieval result, and x_i is the ith joint CloudSat/CALIOP CBH product.

Since the two RF models (VIS+IR and IR-single) select 230 typical variables to fit CBHs, the importance scores of these predictors in the two ML-based algorithms are ranked for better optimization. In a RF model, feature importance indicates how much each input variable contributes to the model's predictive accuracy by measuring the decrease in impurity or error when the feature is used to split data (Gregorutti et al., 2017). In the VIS+IR model, the top-ranked predictors are CTH and cloud top temperature (CTT) from the H8/AHI Level-2 product (see Fig. B1 in Appendix B). It is important to note that D_{COT} is a crucial and sensitive factor for these ML-based algorithms. Retrieving CBH samples with relatively low D_{COT} remains challenging due to the low signal-to-noise ratio when D_{COT} is low (Lin et al., 2022). To address this issue, samples with D_{COT} less than 1.6 are filtered in the VIS+IR model, and samples with relatively large BTs at Channel-14 are filtered in the IR-single model. This filtering process significantly improves the R value from 0.869 to 0.922 in the VIS+IR model and from 0.868 to 0.911 in the IR-single model. For more details on the algorithm optimization, please refer to Appendix B.

In this study, the H8/AHI satellite CBH data retrieved by the four algorithms mentioned before are matched spatiotemporally with the 2B-CLDCLASS-LIDAR cloud product from joint CloudSat/CALIPSO observations in 2017. In this process, the nearest distance matching method is employed, ensuring that collocating the closest points and the observation time difference between the CloudSat/CALIPSO observation point and the matched Himwari-8 data is less than 5 minutes (Noh et al., 2017). As in earlier study (Min et al., 2020), we also used 70% of the matched data for training and 30% of an independent sample for validation. Figure 1 displays a comparison of CBH results over the full disk at 02:00 UTC on January 1, 2017, retrieved by the GEO IDPS algorithm, the GEO CLAVR-x algorithm, the RF VIS+IR algorithm, and the RF IR-single algorithm for all cloud conditions including single and multilayer cloud scenes. A similar distribution pattern and magnitude of CBHs retrieved by these four independent algorithms can be observed in Figure 1. However, notable differences exist between physics-based and ML-based algorithms. Further comparisons are conducted and analyzed with spaceborne and ground-based lidar and radar observations in the subsequent sections of this study.

4 Results and Discussions

4.1 Comparisons with the joint CloudSat/CALIPSO cloud base height product

4.1.1 Joint scatter plots

Figure 2 presents the density scatter plot of the CBHs retrieved from the GEO IDPS and GEO CLAVR-x algorithms compared with the CBHs from the joint CloudSat/CALIPSO product, along with the related scores of MAE, MBE, RMSE, and R calculated and labeled in each panel. The calculated R exceeds the 95% significance level ($p < 0.05$). For the GEO IDPS algorithm, the R is 0.62, the MAE is 1.83 km, and the MBE and RMSE are -0.23 and 2.64 km (Fig. 2a). In comparison, (Seaman et al., 2017) compared the operational VIIRS CBH product retrieved by the similar SNPP/VIIRS IDPS algorithm with the CloudSat CBH results. In their results, the R is 0.57, and the RMSE is 2.3 km. For the new GEO CLAVR-x algorithm (Fig. 2b), the R is 0.645, and the RMSE is 2.91 km. The larger RMSE from two independent physics-based CBH algorithms demonstrate a slightly poorer performance and precision of these retrieval algorithms for GEO satellites. Particularly, the larger RMSEs (2.64 and 2.91 km) indicate weaker stabilities of the GEO IDPS and CLAVR-x CBH algorithms, compared with VIIRS CBH product (Seaman et al., 2017). In this figure, more samples can be found near the 1:1 line, implying the good quality of retrieved CBHs. However, in stark contrast, quite a number of CBH samples retrieved by both GEO IDPS and GEO CLAVR-x algorithms (compared with the official VIIRS CBH product) fall below 1.0 km, indicating relatively large errors when compared with the joint CloudSat/CALIPSO CBH product. Moreover, Figure 2 reveals that relatively large errors are also found in the CBHs lower than 2 km for the four independent algorithms, primarily caused by the weak penetration ability of VIS or IR bands on thick and low clouds.

Referring to the joint CloudSat/CALIPSO CBH product, Figures 2c and 2d present the validations of the CBH results retrieved from two ML-based algorithms using the VIS+IR (only retrieving the CBH during the daytime) and IR-single models. Figure 2c demonstrates better consistency of CBH between the VIS+IR model and the joint CloudSat/CALIPSO product with R = 0.91, MAE = 0.82 km, MBE = 0.43 km, and RMSE = 1.71 km. Figure 2d also displays a relatively high R of 0.876 when validating the IR-single model, with MAE = 0.88, MBE = −0.45, and RMSE = 2.00. Therefore, both VIS+IR and IR-single models can obtain high-quality CBH retrieval results from geostationary imager measurements. In comparison, previous studies also proposed similar ML-based algorithms for estimating CBH using FY-4A satellite imager data. For example, (Tan et al., 2020) used the variables of CTH, D_{COT}, R_{eff}, cloud water path, longitude/latitude from FY-4A imager data to build the training and prediction model and obtained CBH with MAE=1.29 km and R=0.80. In this study, except CTH, the other Level-2 products and geolocation data (longitude/latitude) used in (Tan et al., 2020) are abandoned, while the matched atmospheric profile products (such as temperature and relative humidity) from NWP data are added. These changes in ML-based model training and prediction lead to more accurate CBH retrieval results. Note that, in accordance with the previous study conducted by (Noh et al., 2017), we excluded CBH samples obtained from CloudSat/CALIPSO that were smaller than 1 km in our comparisons. This exclusion was primarily due to the presence of ground clutter contamination in the CloudSat CPR data (Noh et al., 2017).

4.1.2 Test case

Figure 3 displays two cross-sections of CBH from various sources overlaid with CloudSat radar reflectivity (unit: dBZ) for spatiotemporally matched cases. The periods covered are from 03:16 to 04:55 UTC on January 13, 2017 (154.0°E–160.0°E; 40.56°S–53.39°S) and from 05:38 to 07:17 UTC on January 14, 2017 (107.1°E–107.8°E; 8.35°N–11.57°N). The CloudSat radar reflectivity and joint CloudSat/CALIPSO product provide insights into the vertical structure or distribution of clouds and their corresponding CBHs. The results from the four GEO CBH retrieval algorithms (GEO IDPS, GEO CLAVR-x, RF VIS+IR model, and RF

IR-single model) mentioned earlier are individually marked with different markers in each panel. According to Figure 3a, the GEO IDPS algorithm faces challenges in accurately retrieving CBHs for geometrically thicker cloud samples near 157°E. Optically thick mid- and upper-level cloud layers may obscure lower-level cloud layers. However, the CBH results retrieved by the GEO IDPS algorithm near 155°E (in Fig. 3a) and 107.4°E (in Fig. 3b) align with the joint CloudSat/CALIPSO CBH product. It is worth noting that the inconsistency observed between 107.2°E and 107.3°E in Figure 3b, specifically regarding the CBHs around 1 km obtained from CloudSat/CALIPSO, can likely be attributed to ground clutter contamination in the CloudSat CPR data (Noh et al., 2017). The GEO CLAVR-x algorithm achieves improved CBH results compared to the GEO IDPS algorithm. It can even retrieve CBHs for some thick cloud samples that are invalid when using the GEO IDPS algorithm. However, the CBHs from the GEO CLAVR-x algorithm are noticeably higher than those from the joint CloudSat/CALIPSO product. In contrast, the CBHs from the two ML-based algorithms show substantially better results than those from the other two physics-based algorithms. Particularly, the ML-based VIS+IR model algorithm yields the best CBH results. However, compared with those from the two physics-based algorithms, the CBHs from the two ML-based algorithms still exhibit a significant error around 5 km.

4.2 Comparisons with the ground-based lidar and cloud radar measurements

Lidar actively emits laser pulses in different spectral bands into the air. When the laser signal encounters cloud particles during transmission, a highly noticeable backscattered signal is generated and received (Omar et al., 2009). The lidar return signal of cloud droplets is markedly distinct from atmospheric aerosol scattering signals and noise, making CBH easily obtainable from the signal difference or mutation (Sharma et al., 2016). In this study, continuous ground-based lidar data from the Twin Astronomy Manor in Lijiang City, Yunnan Province, China (26.454°N, 100.0233°E, altitude = 3175 m) are used to evaluate the diurnal cycle characteristics of CBHs retrieved using GEO satellite algorithms (Young and Vaughan, 2009). The geographical location and photo of this station are shown in Figure 4.

4.2.1 Comparison of CBH retrievals from ground and satellite data

The ground-based lidar data at Lijiang station on December 6, 2018, and January 8, 2019, are selected for validation. In fact, this lidar was primarily used for the calibration of ground-based lunar radiation instruments. During the two-month observation period (from December of 2018 to January of 2019), it was always operated only under clear sky conditions, resulting in the capture of cloud data on just two days. These two days have been cloudy, with stratiform clouds at an altitude of around 5 km and no precipitation occurring. The number of available and spatiotemporally matched CBH sample points from ground-based lidar is 78 and 64 on December 6, 2018, and January 8, 2019, respectively. Figure 5a and 5b show the point-to-point CBH comparisons between ground-based lidar and four GEO satellite CBH algorithms on December 6, 2018, and January 8, 2019. It is worth noting that the retrieved CBHs of the two physics-based algorithms on December 6, 2018, are in good agreement with the reference values from the lidar measurements, and, in particular, the GEO CLAVR-x algorithm can obtain better results. From the results on January 8, 2019, more accurate diurnal cycle characteristics of CBHs are revealed by the GEO CLAVR-x algorithm than by the GEO IDPS algorithm.

Compared with the CBHs measured by ground-based lidar, the statistics of the results retrieved from the GEO IDPS algorithm are R = 0.67, MAE = 3.09 km, MBE = 0.86 km, and RMSE = 3.61 km (Fig. 5c). However, for cloud samples with CBH below 7.5 km, the GEO IDPS algorithm shows an obvious underestimation of CBH in Figure 5c. For the GEO CLAVR-x algorithm, it can also be seen that the matched samples mostly lie near the 1:1 line with R = 0.77 (the optimal CBH algorithm), MAE = 1.32 km, MBE = 0.22 km, and RMSE = 1.60 km. In addition, this figure also shows the CBH comparisons between the ML-based VIS+IR model/IR-single model algorithms and the lidar measurements, revealing that the retrieved CBH results from the ML-based VIS+IR model are better than those from the ML-based IR-single model algorithm. The comparison results between the CBHs of the ML-based VIS+IR model algorithm and the lidar measurements are around the 1:1 line, with smaller errors and R = 0.60. In contrast, the R between the CBHs of the ML-based IR-single model algorithm and the lidar measurements is only 0.50, with a relatively large error. By comparing the retrieved CBHs with the lidar measurements at Lijiang station, it indicates that CBH results from two physics-based algorithms are remarkably more accurate, particularly that the GEO CLAVR-x algorithm can well capture diurnal variation of CBH.

To further assess the accuracy and quality of the diurnal cycle of CBHs retrieved with these algorithms, CBHs from another ground-based cloud radar dataset covering the entire year of 2017 are also collected and used in this study. The observational instrument is a Ka-band (35 GHz) Doppler millimeter-wave cloud radar (MMCR) located at the Beijing Nanjiao Weather Observatory (a typical urban observation site) (39.81°N, 116.47°E, altitude = 32 m; see Fig. 4), performing continuous and routine observations. The MMCR provides a specific vertical resolution of 30 m and a temporal resolution of 1 minute for single profile detection, based on the radar reflectivity factor. In a previous study (Zhou et al., 2019), products retrieved by this MMCR were utilized to investigate the diurnal variations of CTH and CBH, and comparisons were made between MMCR-derived CBHs and those derived from a Vaisala CL51 ceilometer. The former study also found that the average R of CBHs from different instruments reached up to 0.65. It is worth noting that the basic physics principle for detecting cloud base height from both spaceborne cloud profiling radar and ground-based cloud radar and lidar measurements is the same. All these algorithms of detecting CBH are based on the manifest change of return signals between CBH and the clear sky atmosphere in the vertical direction (Huo et al., 2019; Ceccaldi et al., 2013). The diurnal variation of cloud base height over land is primarily influenced by solar heating, causing the cloud base to rise in the morning and reach its peak by midday. As the surface cools in the afternoon and evening, the cloud base lowers, playing a crucial role in weather patterns and forecasting (Zheng et al., 2020). Due to the density of points in the one-year time series, the point-to-point CBH comparison results for the entire year are not displayed here (monthly results are shown in the supplementary document), we only show 4 days results in the following Figure 6. Therefore, it is essential to rigorously compare the ML-based algorithm with ground-based observations to determine its ability to adapt to the daily variations in cloud base height caused by natural factors. The joint spaceborne CloudSat/CALIPSO detection might face limitations in penetrating extremely dense, optically thick, or areas with heavy precipitation clouds. Hence, in comparison, the CBH values gathered from ground-based lidar and cloud radar measurements are expected to be more accurate than the data derived from spaceborne CloudSat/CALIPSO detection.

Similar to Figure 5, Figure 6 presents two sample groups of CBH results from the cloud radar at Beijing Nanjiao station relative to the matched CBHs from the four retrieval algorithms (GEO IDPS, GEO CLAVR-x, ML-based IR-single, ML-based VIS+IR) on April 9–10 and July 26–28, 2017. Similar to the results at Lijiang station discussed in Figure 5, we observe better and more robust performances in retrieving diurnal cycle characteristics of CBH from the two physics-based CBH retrieval algorithms. In contrast, more underestimated CBH samples are retrieved by the two ML-based algorithms.

4.2.2 Diurnal cycle analysis of CBH retrieval accuracy

To further investigate the diurnal cycle characteristics of retrieved CBH from GEO satellite imager measurements, Figure 7 presents box plots of the hourly CBH errors (relative to the results of cloud radar at Beijing Nanjiao station) in 2017 from the four different CBH retrieval algorithms. Remarkably, there are significant underestimations of the CBHs retrieved from the two ML-based algorithms. The ML-based VIS+IR method achieves relatively better results than the ML-based IR-single method during the daytime. Comparing the two ML-based algorithms, the errors of the IR-single model algorithm have a similar standard deviation (2.80 km) to those of the VIS+IR model algorithm (2.69 km) during the daytime. For the IR-single model algorithm, it can be applied during both daytime and nighttime, its nighttime performance degrades slightly, with an averaged RMSE (3.88 km) higher than that of daytime (3.56 km). The nighttime CBH of the IR-single model algorithm is the only choice that should be used with discretion.

Figure 8 shows the comparisons of hourly MAE, MBE, RMSE, and R relative to the CBHs from the cloud radar at Beijing Nanjiao station during daytime between four retrieval algorithms in 2017. The RMSE of the two ML-based algorithms shows stable diurnal variation. It is noted that all algorithms have lower R at sunrise, around 07:00 local time, which improve as the day progresses. However, the GEO CLAVR-x algorithm stands out for its relatively higher and more stable in R and RMSE during daytime.

Figure 9a displays scatter plots and relevant statistics of the CBHs retrieved from the GEO IDPS algorithm against the CBHs from cloud radar. The CBHs from the GEO IDPS algorithm align well with the matched CBHs from cloud radar at Beijing Nanjiao station, with R = 0.52, MAE = 2.08 km, MBE = 1.17 km, and RMSE = 2.67 km. In Figure 9b, the GEO CLAVR-x algorithm shows better results with R = 0.57, MAE = 2.06 km, MBE = −0.20 km, and RMSE = 2.60 km. It is not surprising that Figs. 8c and 8d reveal obvious underestimated CBH results from the two ML-based CBH algorithms. Particularly, the CBH results from the ML-based VIS+IR model algorithm concentrate in the range of 2.5 km to 5 km. Therefore, Figure 5 to Figure 9 further substantiates the weak diurnal variations captured by ML-based techniques, primarily attributed to the scarcity of comprehensive CBH training samples throughout the entire day. Besides, although the two robust physics-based algorithms of GEO IDPS and GEO CLAVR-x (the optimal one) can retrieve high-quality CBHs from H8/AHI data, especially the diurnal cycle of CBH during the daytime, they still struggle to retrieve CBHs below 1 km.

5. Conclusions and discussion

To explore and argue the optimal and most robust CBH retrieval algorithm from geostationary satellite imager measurements, particularly focusing on capturing the typical diurnal cycle characteristics of CBH over land, this study employs four different retrieval algorithms (two physics-based and two ML-based algorithms). High spatiotemporal resolution CBHs are retrieved using the H8/AHI data from 2017 and 2019. To assess the accuracies of the retrieved CBHs, point-to-point validations are conducted using spatiotemporally matched CBHs from the joint CloudSat/CALIOP product, ground-based lidar and cloud radar observations in China. The main findings and conclusions are outlined below.

Four independent CBH retrieval algorithms, namely physics-based GEO IDPS, GEO CLAVR-x, ML-based VIS+IR, and ML-based IR-single, have been developed and utilized to retrieve CBHs from GEO H8/AHI data under the assumption of single layer cloud. The two physics-based algorithms utilize cloud top and optical property products from AHI as input parameters to retrieve high spatiotemporal resolution CBHs, with operations limited to daytime. In contrast, the ML-based VIS+IR model and IR-single model algorithms use the matched joint CloudSat/CALIOP CBH product as true values for building RF prediction models. Notably, the ML-based IR-single algorithm, which relies solely on infrared band measurements, can retrieve CBH during both day and night.

The accuracy of CBHs retrieved from the four independent algorithms is verified using the joint CloudSat/CALIOP CBH products for the year 2017. The GEO IDPS algorithm shows an R of 0.62 and an RMSE of 2.64 km. The GEO CLAVR-x algorithm provides more accurate CBHs with an R of 0.65 and RMSE of 2.91 km. After filtering samples with optical thickness less than 1.6 and brightness temperature (at 11 μm band) greater than 281 K, the ML-based VIS+IR and ML-based IR-single algorithms achieve higher accuracy with an R(RMSE) of 0.92(1.21 km) and 0.91(1.42

km), respectively. This indicates strong agreement between the two ML-based CBH algorithms and the CloudSat/CALIOP CBH product.

However, in stark contrast, the results from the physics-based algorithms (with R and RMSE of 0.59/2.86 km) are superior to those from the ML-based algorithms (with R and RMSE of 0.39/3.88 km) when compared with ground-based CBH observations such as lidar and cloud radar. In the comparison with the cloud radar at Beijing Nanjiao station in 2017, the R of the GEO CLAVR-x algorithm is 0.57, while the R of the GEO IDPS algorithm is 0.52. Meanwhile, notable differences are observed in the CBHs between both ML-based algorithms. Similar conclusions are also evident in the 2-day comparisons at Yunnan Lijiang station.

The CBH results from the two ML-based algorithms (R > 0.91) can likely be attributed to the use of the same training and validation dataset source as the joint CloudSat/CALIOP product. However, this dataset has limited spatial coverage and small temporal variation, potentially limiting the representativeness of the training data. In contrast, the GEO CLAVR-x algorithm demonstrates the best performance and highest accuracy in retrieving CBH from geostationary satellite data. Notably, its results align well with those from ground-based lidar and cloud radar during the daytime. However, both physics-based methods, utilizing CloudSat CPR data for regression, struggle to accurately retrieve CBHs below 1 km, as the lowest 1 km above ground level of this data is affected by ground clutter. In general, the physics-based algorithms, such as GEO CLAVR-x and GEO IDPS, demonstrate notable advantages in capturing the diurnal cycle of CBH. Unlike ML-based methods, they offer more stable error metrics, especially with higher correlation and lower RMSE during the daytime. Additionally, they are more effective at capturing significant and natural variations in CBH, providing generally higher quality retrievals from H8/AHI data, even though challenges remain in accurately retrieving CBHs below 1 km.

设置了格式: 字体: (默认) Times New Roman, (中文) DengXian, 小四, 字体颜色: 文字 1

设置了格式: 字体: (默认) Times New Roman, (中文) DengXian, 小四, 字体颜色: 文字 1

设置了格式: 字体: (默认) Times New Roman, (中文) DengXian, 小四, 字体颜色: 文字 1

设置了格式: 字体: (默认) Times New Roman, (中文) DengXian, 小四, 字体颜色: 文字 1

Additionally, despite utilizing the same physics principles in spaceborne and ground-based lidar/radar CBH algorithms, the previous study (Thorsen et al., 2011) has highlighted differences in profiles between them. Therefore, this factor induced by detection principle could contribute to the relatively poorer results in CBH retrieval by ML-based algorithms compared to ground-based lidar and radar. The analysis and discussion above suggest that ML-based algorithms are constrained by the size and representativeness of their datasets.

Ideally, we guess that including more spaceborne cloud profiling radars with varying passing times (covering the entire day) in the training dataset could improve the machine learning technique, potentially leading to a higher-quality CBH product with more comprehensive observations. The CBH product using ML-based algorithms should continue to be improved in future work. Particularly, exploring the joint ML-physics-based method presents a promising direction, which can address the complexities and challenges in retrieving cloud properties. By integrating established physical relationships into ML models, we can potentially enhance the accuracy and reliability of predictions. This approach not only leverages the strengths of both physics-based models and data-driven techniques but also offers a pathway to more robust and interpretable solutions in atmospheric sciences. At present, we will focus on developing physics-based algorithms for cloud base height for the next generation of geostationary meteorological satellites, to support the application of these products in weather and climate domains.

Besides, at night, current GEO satellite imaging instruments encounter challenges in accurately determining CBH due to limited or absent solar illumination. Because it is unable to retrieve cloud optical depth in the visible band, the current method faces limitations. However, there is potential for enhanced accuracy in deriving cloud optical and microphysical properties, as well as CBH, by incorporating the Day/Night Band (DNB) observations during nighttime in the future (Walther et al., 2013).

Data availability. The authors would like to acknowledge NASA, JMA, Colorado State University, and NOAA for freely providing the MODIS (https://ladsweb.modaps.eosdis.nasa.gov/search), CloudSat/CALIOP (https://www.cloudsat.cira.colostate.edu/), Himawari-8 (ftp.ptree.jaxa.jp), and GFS NWP (ftp://nomads.ncdc.noaa.gov/GFS/Grid4) data online, respectively.

Author contributions. MM proposed the essential research idea. MW, MM, JL, HL, BC, and YL performed the analysis and drafted the manuscript. ZY and NX provided useful comments. All the authors contributed to the interpretation and discussion of results and the revision of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors would like to acknowledge NASA, JMA, University of Colorado, and NOAA for freely providing satellite data online, respectively. The authors thank NOAA, NASA, and their VIIRS algorithm working groups (AWG) for freely providing the VIIRS cloud base height algorithm theoretical basic documentations (ATBD). In addition, the authors appreciate the power computer tools developed by the Python and scikit-learn groups (http://scikit-learn.org). Besides the authors also thank Rundong Zhou and Pan Xia for drawing some pictures of this manuscript. Last but not the least, the authors sincerely thank Prof. Yong Zhang and Prof. Jianping Guo for freely providing cloud base height results retrieved by ground-based cloud radar at Beijing Nanjiao station. This work was supported partly by the Guangdong Major Project of Basic and Applied Basic Research (Grant 2020B0301030004), National Natural Science Foundation of China under Grants 42175086 and U2142201, FengYun Meteorological Satellite Innovation Foundation under Grant FY-APP-ZX-2022.0207, the Innovation Group Project of Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai) (No. SML2023SP208), and the Science and Technology Planning Project of Guangdong Province (2023B1212060019). We would like to thank the editor and anonymous reviewers for their thoughtful suggestions and comments.

Appendix A

Based on the previously discussed description of two physics-based cloud base height (CBH) retrieval algorithms (GEO IDPS and GEO CLAVR-x retrieval algorithms), cloud products such as cloud top height (CTH), effective particle radius

(R_{eff}), and cloud optical thickness (D_{COT}) will be utilized in both algorithms. To validate the reliability of these cloud products derived from the Advanced Himawari Imager (AHI) aboard the Himawari-8 (H8), a pixel-by-pixel comparison is conducted with analogous MODIS Collection-6.1 Level-2 cloud products. Both Aqua and Terra MODIS Level-2 cloud products (MOD06 and MYD06) are accessible for free download from the MODIS official website. For verification purposes, the corresponding Level-2 cloud products from January, April, July, and October of 2018 are chosen to assess CTH, D_{COT}, and R_{eff} retrieved by H8/AHI.

Figure S2 (in the supplementary document) shows the spatiotemporally matched case comparisons of CTH, D_{COT} and R_{eff} from H8/AHI and Terra/MODIS (MYD06) at 03:30 UTC on January 15, 2018. It can be seen that the CTH, D_{COT} and R_{eff} from H8/AHI are in good agreement with the matched MODIS cloud products. However, there are still some differences in R_{eff} at the regions near 35°N, 110°E in Figures S2d and S2c. The underestimated R_{eff} values from H8/AHI relative to MODIS have been reported in previous studies. (Letu et al., 2019) compared the ice cloud products retrieved from AHI and MODIS, and concluded that the R_{eff} from both products differ remarkably in the ice cloud region and the D_{COT} from them are roughly similar. However, the D_{COT} from AHI data is higher in some areas. Looking again at the cloud optical thickness that at the same time, the slight underestimation of H8/AHI D_{COT} can be found in Figures S2e and S2f. Figure S3 (in the supplementary document) shows another case at 02:10 UTC on January 15, 2018. Despite of the good consistence between H8/AHI and MODIS cloud products, there are slight differences in CTH in the area around 40°S–40.5°S, 100°E–110°E in Figs. S3a and S3b. Besides, as shown in Figure S2, there are still underestimations in the R_{eff} of H8/AHI.

To further compare and validate these three H8/AHI cloud products, the spatiotemporally matched samples from H8/AHI and Aqua/Terra MODIS in four months of 2018 are counted within the three intervals of 0.1 km (CTH), 1.0 μm (R_{eff}), and 1 (D_{COT}) in Figure S4 (in the supplementary document). The corresponding mean absolute error, mean bias error, RMSE and R values are also calculated and marked in each subfigure. As can be seen, the R of CTH is around 0.75 in all four months and is close to 0.8 in August. The results of D_{COT} show the highest R, reaching above 0.8. In contrast, the underestimation trend in R_{eff} is also shown in this figure. These different consistencies between two satellite-retrieved cloud products may be attributed to: (1) different spatiotemporal resolutions between H8/AHI and MODIS; (2) different wavelength bands, bulk scattering model, and specific algorithm used for retrieving cloud products; (3) different view zenith angle between GEO and low-earth-orbit satellite platforms (Letu et al., 2019). In addition, other external factors such as surface type also can affect the retrieval of cloud product. However, according to Figure S4, the bulk of the analyzed samples are still around the 1:1 line, indicating the good quality of H8/AHI cloud products.

Appendix B

The ML-based visible (VIS)+infrared (IR) model algorithm mentioned above uses 230 typical variables (see Table 1) as model predictors, and the importance scores of top-30 predictors are ranked in Figure S5 (in the supplementary document). It can be seen that the most important variables are CTH and CTT, and D_{COT} is an important or sensitive factor affecting these two quantities. A sensitivity test is also performed to further investigate the potential influence of D_{COT} on the CBH retrieval by the VIS+IR model (see Table S1 in the supplementary document). From Figure S7a, we find that the samples with D_{COT} lower than 5 cause the relatively large CBH errors compared with the matched CBHs from the joint CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation)/CloudSat product.

According to the results in this Figure S7b, we may filter the samples with relatively small D_{COT} to further improve the accuracy of CBH retrieval by the VIS+IR model (see Table S1). Figure S7b shows that after filtering the samples with the D_{COT} less than 1.6, the R increases from 0.895 to 0.922, implying a better performance of CBH retrieval. According to the ranking of predictor importance (see Fig. S6 in the supplementary document), we also conduct another sensitivity test on the BT observed by H8/AHI IR Channel-14 (Cha14) at 11 µm, which plays an important role in the IR-single model. Figure S7c shows that the BT values of H8/AHI Channel-14 ranges from 160 K to 316 K, and the samples with BT higher than 300 K show large CBH errors. Similarly, by filtering the samples with BT higher than 281 K, we can get a better IR-single model algorithm for retrieving high-quality CBH (see Table S2 in the supplementary document). Figure S7d also proves that the R value increases from 0.868 to 0.911.

删除了:

Reference

Aydin, K. and Singh, J.: Cloud Ice Crystal Classification Using a 95-GHz Polarimetric Radar, Journal of Atmospheric and Oceanic Technology, 21, 1679–1688, https://doi.org/10.1175/JTECH1671.1, 2004.

Baker, N.: Joint Polar Satellite System (JPSS) VIIRS Cloud Base Height Algorithm Theoretical Basis Document (ATBD), 2011.

Baum, B., Menzel, W. P., Frey, R., Tobin, D., Holz, R., and Ackerman, S.: MODIS cloud top property refinements for Collection 6, Journal of Applied Meteorology and Climatology, 51, 1145-1163, 10.1175/JAMC-D-11-0203.1, 2012.

Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., Kumagai, Y., Miyakawa, T., Murata, H., Ohno, T., Okuyama, A., Oyama, R., Sasaki, Y., Shimazu, Y., Shimoji, K., Sumida, Y., Suzuki, M., Taniguchi, H., Tsuchiyama, H., Uesawa, D., Yokota, H., and Yoshida, R.: An introduction to Himawari-8/9—Japan's new-generation geostationary meteorological satellites, Journal of the Meteorological Society of Japan, 94, 151-183, 10.2151/jmsj.2016-009, 2016.

Breiman, L.: Random forests, Machine Learning, 45, 5-32, 2001.

Ceccaldi, M., Delanoë, J., Hogan, R. J., Pounder, N. L., Protat, A., and Pelon, J.: 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, 7962-7981, 10.1002/jgrd.50579, 2013.

Forsythe, J. M., Haar, T. H. V., and Reinke, D. L.: Cloud-Base height estimates using a combination of Meteorological Satellite Imagery and Surface Reports, Journal of Applied Meteorology and Climatology, 39, 2336–2347, https://doi.org/10.1175/1520-0450(2000)039<2336:CBHEUA>2.0.CO;2, 2000.

Gregorutti, B., Michel, B., and Saint-Pierre, P.: Correlation and variable importance in random forests, Statistics and Computing, 27, 659-678, 10.1007/s11222-016-9646-1, 2017.

Håkansson, N., Adok, C., Thoss, A., Scheirer, R., and Hörnquist, S.: Neural network cloud top pressure and height for MODIS, Atmospheric Measurement Techniques, 11, 3177–3196, 10.5194/amt-11-3177-2018, 2018.

Hansen, B.: A Fuzzy Logic–Based Analog Forecasting System for Ceiling and Visibility, Weather and Forecasting, 22, 1319-1330, 10.1175/2007waf2006017.1, 2007.

Hartmann, D. L. and Larson, K.: An important constraint on tropical cloud - climate feedback, Geophys Res Lett, 29, 12-11-12-14, 10.1029/2002gl015835, 2002.

Heidinger, A. and Pavolonis, M.: Gazing at cirrus clouds for 25 years through a split window, part 1: Methodology, Journal of Applied Meteorology and Climatology, 48, 1110-1116, 10.1175/2008JAMC1882.1, 2009.

Heidinger, A. K., Bearson, N., Foster, M. J., Li, Y., Wanzong, S., Ackerman, S., Holz, R. E., Platnick, S., and Meyer, K.: Using sounder data to improve cirrus cloud height estimation from satellite imagers, Journal of Atmospheric and Oceanic Technology, 36, 1331-1342, 10.1175/jtech-d-18-0079.1, 2019.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999-2049, https://doi.org/10.1002/qj.3803, 2020.

Heymsfield, A. J., Bansemer, A., Matrosov, S., and Tian, L.: The 94-GHz radar dim band: Relevance to ice cloud properties and CloudSat, Geophys. Res. Lett., 35, 10.1029/2007GL031361, 2008.

Hirsch, E., Agassi, E., and Koren, I.: A novel technique for extracting clouds base height using ground based imaging, Atmospheric Measurement Techniques, 4, 117-130, 10.5194/amt-4-117-2011, 2011.

Hunt, W. H., Winker, D. M., Vaughan, M. A., Powell, K. A., Lucker, P. L., and Weimer, C.: CALIPSO lidar description and performance assessment, J. Atmos. Oceanic. Technol., 26, 2009.

Huo, J., Bi, Y., Lü, D., and Duan, S.: Cloud Classification and Distribution of Cloud Types in Beijing Using Ka-Band Radar Data, Advances in Atmospheric Sciences, 36, 793-803, 10.1007/s00376-019-8272-1, 2019.

Hutchison, K., Wong, E., and Ou, S. C.: Cloud base heights retrieved during night‐time conditions with MODIS data, Int J Remote Sens, 27, 2847-2862, 10.1080/01431160500296800, 2006.

Hutchison, K. D.: The retrieval of cloud base heights from MODIS and three-dimensional cloud fields from NASA's EOS Aqua mission, Int J Remote Sens, 23, 5249-5265, 10.1080/01431160110117391, 2002.

Iwabuchi, H., Putri, N. S., Saito, M., Tokoro, Y., Sekiguchi, M., Yang, P., and Baum, B. A.: Cloud Property Retrieval from Multiband Infrared Measurements by Himawari-8, Journal of the Meteorological Society of Japan. Ser. II, 96B, 27-42, 10.2151/jmsj.2018-001, 2018.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., W.Higgins, Janowiak, J., Mo, K. C., Ropelewski, C., and Wang, J.: The NCEP NCAR 40-Year Reanalysis Project, 1996.

Kühnlein, M., Appelhans, T., Thies, B., and Nauß, T.: Precipitation Estimates from MSG SEVIRI Daytime, Nighttime, and Twilight Data with Random Forests, Journal of Applied Meteorology and Climatology, 53, 2457-2480, 10.1175/jamc-d-14-0082.1, 2014.

Letu, H., Nagao, T. M., Nakajima, T. Y., Riedi, J., Ishimoto, H., Baran, A. J., Shang, H., Sekiguchi, M., and Kikuchi, M.: Ice cloud properties from Himawari-8/AHI next-generation geostationary satellite: Capability of the AHI to monitor the DC cloud generation process, IEEE Transactions on Geoscience and Remote Sensing, 57, 3229-3239, 10.1109/tgrs.2018.2882803, 2019.

Li, Y., Yi, B., and Min, M.: Diurnal variations of cloud optical properties during day-time over China based on Himawari-8 satellite retrievals, Atmospheric Environment, 277, 119065, 10.1016/j.atmosenv.2022.119065, 2022.

Liang, Y., Min, M., Yu, Y., Wang, X., and Xia, P.: Assessing diurnal cycle of cloud covers of Fengyun-4A geostationary satellite based on the manual observation data in China, IEEE Transactions on Geoscience and Remote Sensing, 61, 10.1109/TGRS.2023.3256365, 2023.

Lin, H., Li, Z., Li, J., Zhang, F., Min, M., and Menzel, W. P.: Estimate of daytime single-layer cloud base height from Advanced Baseline Imager measurements, Remote Sensing of Environment, 274, 112970, 10.1016/j.rse.2022.112970, 2022.

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

Meerkötter, R. and Bugliaro, L.: Diurnal evolution of cloud base heights in convective cloud fields from MSG/SEVIRI data Atmospheric Chemistry and Physics, 9, 1767–1778, 10.5194/acp-9-1767-2009, 2009.

Miller, R. M., Rauber, R. M., Girolamo, L. D., Rilloraza, M., Fu, D., McFarquhar, G. M., Nesbitt, S. W., Ziemba, L. D., Woods, S., and Thornhill, K. L.: Influence of natural and anthropogenic aerosols on cloud base droplet size distributions in clouds over the South China Sea and West Pacific, Atmospheric

Chemistry and Physics, 23, 8959–8977, 10.5194/acp-23-8959-2023, 2023.

Miller, S. D., Rogers, M. A., Haynes, J. M., Sengupta, M., and Heidinger, A. K.: Short-term solar irradiance forecasting via satellite/model coupling, Solar Energy, 168, 102-117, 10.1016/j.solener.2017.11.049, 2018.

Min, M. and Zhang, Z.: On the influence of cloud fraction diurnal cycle and sub-grid cloud optical thickness variability on all-sky direct aerosol radiative forcing, Journal of Quantitative Spectroscopy and Radiative Transfer, 142, 25-36, 10.1016/j.jqsrt.2014.03.014., 2014.

Min, M., Li, J., Wang, F., Liu, Z., and Menzel, W. P.: Retrieval of cloud top properties from advanced geostationary satellite imager measurements based on machine learning algorithms, Remote Sensing of Environment, 239, 111616, 10.1016/j.rse.2019.111616 2020.

Min, M., Chen, B., Xu, N., He, X., Wei, X., and Wang, M.: Nonnegligible diurnal and long-term variation characteristics of the calibration biases in Fengyun-4A/AGRI infrared channels based on the oceanic drifter data, IEEE Transactions on Geoscience and Remote Sensing, 60, 1-15, 10.1109/TGRS.2022.3160450, 2022.

Min, M., Wu, C., Li, C., Liu, H., Xu, N., Wu, X., Chen, L., Wang, F., Sun, F., Qin, D., Wang, X., Li, B., Zheng, Z., Cao, G., and Dong, L.: Developing the science product algorithm testbed for Chinese next-generation geostationary meteorological satellites: FengYun-4 series, Journal of Meteorological Research, 31, 708-719, 10.1007/s13351-017-6161-z, 2017.

Noh, Y.-J., Miller, S. D., Seaman, C. J., Haynes, J. M., Li, Y., Heidinger, A. K., and Kulie, M. S.: Enterprise AWG Cloud Base Algorithm (ACBA), 2022.

Noh, Y.-J., Forsythe, J. M., Miller, S. D., Seaman, C. J., Li, Y., Heidinger, A. K., Lindsey, D. T., Rogers, M. A., and Partain, P. T.: Cloud-base height estimation from VIIRS. Part II: A statistical algorithm based on A-Train satellite data, Journal of Atmospheric and Oceanic Technology, 34, 585–598, 10.1175/JTECH-D-16-0110.1, 2017.

Omar, A., Winker, D., Kittaka, C., Vaughan, M., Liu, Z., Hu, Y., Trepte, C., Rogers, R., Ferrare, R., Kuehn, R., and Hostetler, C.: The CALIPSO automated aerosol classification and lidar ratio selection algorithm, J. Atmos. Oceanic. Technol., 26, 1994-2014, 10.1175/2009JTECHA1231, 2009.

Platnick, S., Meyer, K. G., King, M. D., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Holz, R. E., Yang, P., Ridgway, W. L., and Riedi, J.: The MODIS cloud optical and microphysical products: Collection 6 updates and examples from Terra and Aqua, IEEE Trans Geosci Remote Sens, 55, 502-525, 10.1109/TGRS.2016.2610522, 2017.

Rosenfeld, D., Zheng, Y., Hashimshoni, E., Pohlker, M. L., Jefferson, A., Pohlker, C., Yu, X., Zhu, Y., Liu, G., Yue, Z., Fischman, B., Li, Z., Giguzin, D., Goren, T., Artaxo, P., Barbosa, H. M., Poschl, U., and Andreae, M. O.: Satellite retrieval of cloud condensation nuclei concentrations by using clouds as CCN chambers, Proc. Natl. Acad. Sci., 113, 5828-5834, 10.1073/pnas.1514044113, 2016.

Sassen, K. and Wang, Z.: Classifying clouds around the globe with the CloudSat radar: 1-year of results, Geophys. Res. Lett., 35, 1-5, doi:10.1029/2007GL032591, 2008.

Seaman, C. J., Noh, Y.-J., Miller, S. D., Heidinger, A. K., and Lindsey, D. T.: Cloud-base height estimation from VIIRS. Part I: Operational algorithm validation against CloudSat, Journal of Atmospheric and Oceanic Technology, 34, 567-583, 10.1175/jtech-d-16-0109.1, 2017.

Sharma, S., Vaishnav, R., Shukla, M. V., Kumar, P., Kumar, P., Thapliyal, P. K., Lal, S., and Acharya, Y. B.: Evaluation of cloud base height measurements from Ceilometer CL31 and MODIS satellite over Ahmedabad, India, Atmospheric Measurement Techniques, 9, 711-719, 10.5194/amt-9-711-2016, 2016.

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., and Sassen, K.: The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation, Bull. Amer. Meteor. Soc., 83, 1771-1790, 2002.

Stubenrauch, C. J., Rossow, W. B., Kinne, S., Ackerman, S., Cesana, G., Chepfer, H., Di Girolamo, L., Getzewich, B., Guignard, A., Heidinger, A., Maddux, B. C., Menzel, W. P., Minnis, P., Pearl, C., Platnick, S., Poulsen, C., Riedi, J., Sun-Mack, S., Walther, A., Winker, D., Zeng, S., and Zhao, G.: Assessment of global cloud datasets from satellites: project and database initiated by the GEWEX radiation panel, Bulletin of the American Meteorological Society, 94, 1031-1049, 10.1175/bams-d-12-00117.1, 2013.

Su, T., Zheng, Y., and Li, Z.: Methodology to determine the coupling of continental clouds with surface and boundary layer height under cloudy conditions from lidar and meteorological data, Atmospheric Chemistry and Physics, 22, 1453-1466, 10.5194/acp-22-1453-2022, 2022.

Tan, Z., Huo, J., Ma, S., Han, D., Wang, X., Hu, S., and Yan, W.: Estimating cloud base height from Himawari-8 based on a random forest algorithm, Int J Remote Sens, 42, 2485-2501, 10.1080/01431161.2020.1854891, 2020.

Thorsen, T. J., Fu, Q., and Comstock, J.: Comparison of the CALIPSO satellite and ground-based observations of cirrus clouds at the ARM TWP sites, Journal of Geophysical Research: Atmospheres, 116, 10.1029/2011jd015970, 2011.

Viúdez-Mora, A., Costa-Surós, M., Calbó, J., and González, J. A.: Modeling atmospheric longwave radiation at the surface during overcast skies: The role of cloud base height, Journal of Geophysical Research: Atmospheres, 120, 199-214, 10.1002/2014jd022310, 2015.

Wang, F., Min, M., Xu, N., Liu, C., Wang, Z., and Zhu, L.: Effects of linear calibration errors at low temperature end of thermal infrared band: Lesson from failures in cloud top property retrieval of FengYun-4A geostationary satellite, IEEE Transactions on Geoscience and Remote Sensing, 60, 5001511, 10.1109/TGRS.2022.3140348, 2022.

Wang, T., Shi, J., Ma, Y., Letu, H., and Li, X.: All-sky longwave downward radiation from satellite measurements: General parameterizations based on LST, column water vapor and cloud top temperature, ISPRS Journal of Photogrammetry and Remote Sensing, 161, 52-60, 10.1016/j.isprsjprs.2020.01.011, 2020.

Wang, X., Min, M., Wang, F., Guo, J., Li, B., and Tang, S.: Intercomparisons of cloud mask product among Fengyun-4A, Himawari-8 and MODIS, IEEE Transactions on Geoscience and Remote Sensing, 57, 8827-8839, 10.1109/TGRS.2019.2923247 2019.

Wang, Z., Vane, D., Stephens, G., Reinke, D., and TBD: Level 2 combined radar and lidar cloud scenario classification product process description and interface control document, 2012.

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

Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the CALIPSO mission and CALIOP data processing algorithms, J. Atmos. Oceanic. Technol., 26, 2310-2323, 10.1175/2009JTECHA1281.1, 2009.

Yang, J., Li, S., Gong, W., Min, Q., Mao, F., and Pan, Z.: A fast cloud geometrical thickness retrieval algorithm for single-layer marine liquid clouds using OCO-2 oxygen A-band measurements, Remote Sensing of Environment, 256, 10.1016/j.rse.2021.112305, 2021.

Young, S. A. and Vaughan, M. A.: The retrieval of profiles of particulate extinction from Cloud Aerosol Lidar Infrared Pathfinder Satellite Observations (CALIPSO) data: Algorithm description, J. Atmos.

Oceanic. Technol., 26, 1105-1119, 10.1175/2008JTECHA1221.1, 2009.

Zhang, Y., Zhang, L., Guo, J., Feng, J., Cao, L., Wang, Y., Zhou, Q., Li, L., Li, B., Xu, H., Liu, L., An, N., and Liu, H.: Climatology of cloud-base height from long-term radiosonde measurements in China, Advances in Atmospheric Sciences, 35, 158-168, 10.1007/s00376-017-7096-0, 2018.

Zheng, Y. and Rosenfeld, D.: Linear relation between convective cloud base height and updrafts and application to satellite retrievals, Geophys Res Lett, 42, 6485-6491, 10.1002/2015gl064809, 2015.

Zheng, Y., Sakradzija, M., Lee, S.-S., and Li, Z.: Theoretical Understanding of the Linear Relationship between Convective Updrafts and Cloud-Base Height for Shallow Cumulus Clouds. Part II: Continental Conditions, J Atmos Sci, 77, 1313-1328, 10.1175/jas-d-19-0301.1, 2020.

Zhou, Q., Zhang, Y., Li, B., Li, L., Feng, J., Jia, S., Lv, S., Tao, F., and Guo, J.: Cloud-base and cloud-top heights determined from a ground-based cloud radar in Beijing, China, Atmospheric Environment, 201, 381-390, 10.1016/j.atmosenv.2019.01.012, 2019.

Zhou, R., Pan, X., Xiaohu, Z., Na, X., and Min, M.: Research progress and prospects of atmospheric motion vector based on meteorological satelliteimages, Reviews of Geophysics and Planetary Physics (In Chinese), 55, 184-194, 10.19975/j.dqyxx.2022-077, 2024.

Zhu, Y., Rosenfeld, D., Yu, X., Liu, G., Dai, J., and Xu, X.: Satellite retrieval of convective cloud base temperature based on the NPP/VIIRS Imager, Geophys Res Lett, 41, 1308-1313, 10.1002/2013gl058970, 2014.

Tables and Figures

Table 1. Predictand and predictor variables for both visible (VIS)+infrared (IR) model and IR-single regression model training, which are divided according to the different predictor variables from satellite and NWP data

Predictand	**IR-single model input**	**VIS+IR model input**
Predictor [satellite measurements]	BT(3.9μm), BT(6.2μm), BT(6.9μm), BT(7.3μm), BT(8.6μm), BT(9.6μm), BT(10.4μm), BT(11.2μm), BT(12.4μm), BT (13.3μm), BTD(11.2–12.4μm), BTD(11.2–13.3μm) [Unit = K], Air Mass (1/cos(VZA)), View azimuth angles [Unit = degree], Cloud top height from H8/AHI [unit: m], Cloud top temperature from H8/AHI [unit: K]	BT(3.9μm), BT(6.2μm), BT(6.9μm), BT(7.3μm), BT(8.6μm), BT(9.6μm), BT(10.4μm), BT(11.2μm), BT(12.4μm), BT(13.3μm), BTD(11.2–12.4μm), BTD(11.2–13.3μm) [Unit = K], Air Mass(1/cos(VZA)), Air Mass(1/cos(SZA)), View/Solar Azimuth angles [Unit = degree], Cloud top height from H8/AHI [unit: m], Cloud top temperature from H8/AHI [unit: K] Ref(0.47μm), Ref(0.51μm), Ref(0.64μm), Ref(0.86μm), Ref(1.64μm), Ref(2.25μm)
Predictor [GFS NWP]	Altitude profile (from surface to about 21 km, 67 layers) [unit: m], Temperature profile (from surface to about 21 km, 67 layers) [unit: K], Relative humidity profile (from surface to about 21 km, 67 layers) [unit: %], Total precipitable water, Surface temperature [unit: K]	Altitude profile (from surface to about 21 km, 67 layers) [unit: m], Temperature profile (from surface to about 21 km, 67 layers) [unit: K], Relative humidity profile (from surface to about 21 km, 67 layers) [unit: %], Total precipitable water, Surface temperature [unit: K]
Predictor	Surface elevation [unit: m]	Surface elevation [unit: m]

[other]		

Notes: VZA = view zenith angle [unit: degree]; SZA = solar zenith angle [unit: degree]

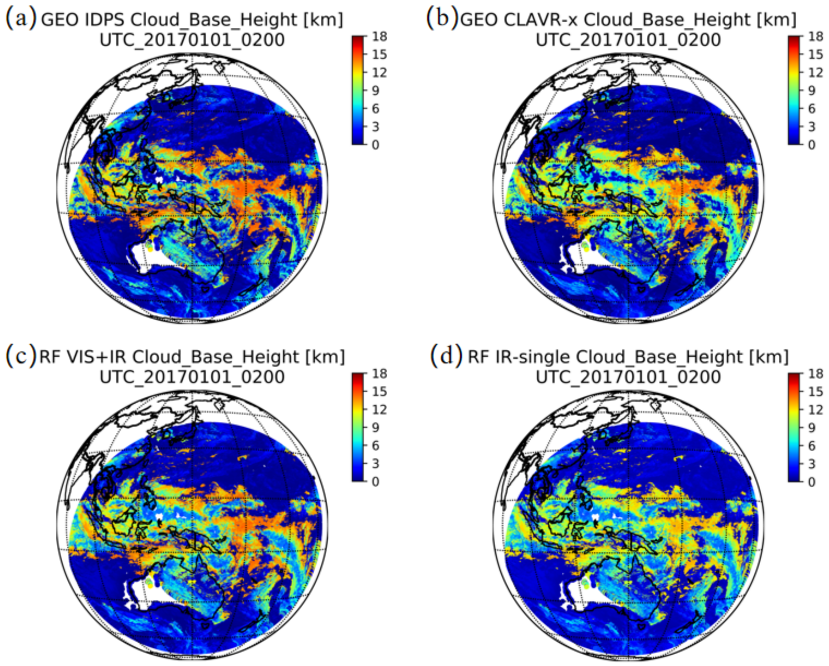

Figure 1. Comparison of full disk CBH results retrieved by the four independent algorithms at 02:00 UTC on January 1, 2017. (a) GEO IDPS algorithm, (b) GEO Clouds from AVHRR Extended (CLAVR-x) algorithm, (c) ML-based (RF, random forest) VIS+IR algorithm and (d) ML-based (RF) IR-single algorithm.

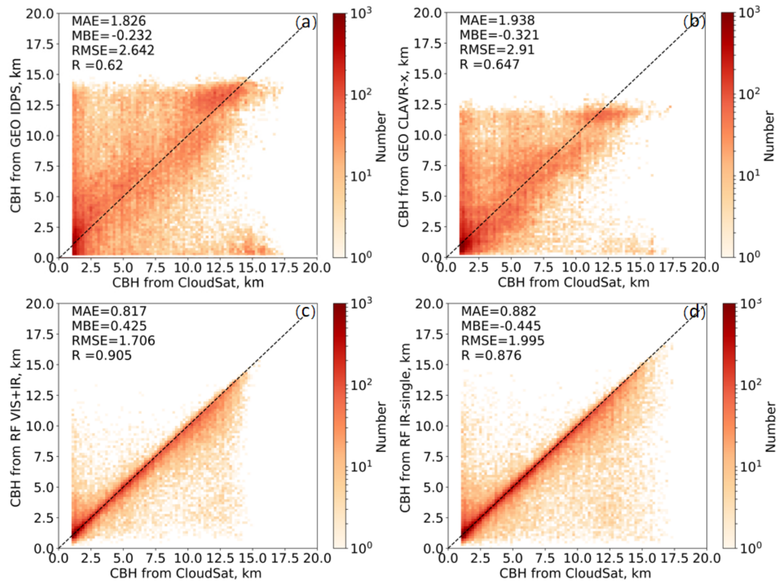

Figure 2. Density distributions of CBHs retrieved from (a) GEO IDPS, (b) GEO CLAVR-x, (c) VIS+IR and (d) IR-single algorithms compared with the CBHs from the joint CloudSat/CALIPSO product (taken as true values) in 2017 for both single and multilayer clouds. The mean absolute error (MAE), mean bias error (MBE), root mean square error (RMSE) and R are listed in each subfigure where the difference exceeds the 95% significance level ($p < 0.05$) according to the Pearson's $\chi 2$ test.

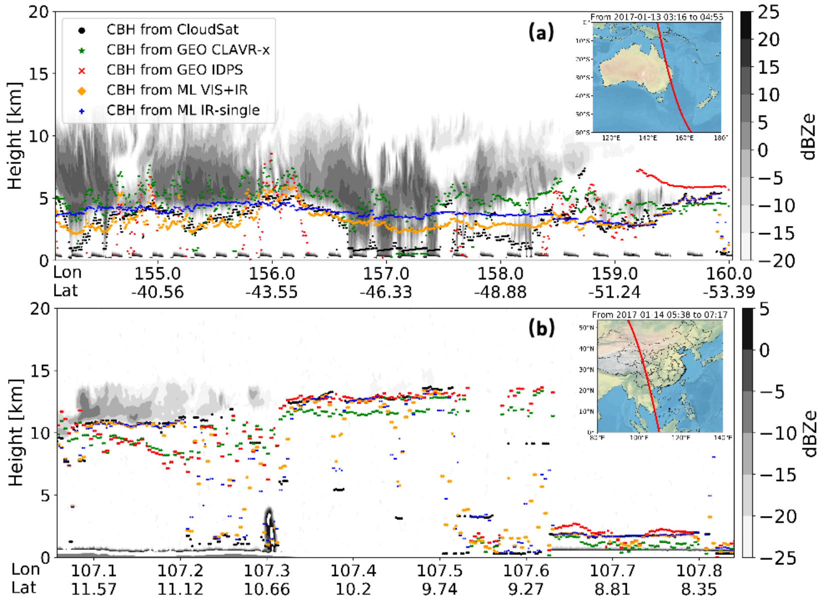

Figure 3. Inter-comparisons of CBH products retrieved by CloudSat (red solid circle), the GEO IDPS algorithm (blue solid circle), the GEO CLAVR-x (green solid circle), the ML-based VIS+IR model algorithm (orange solid circle), and the ML-based IR-single model algorithm (pink solid circle) at (a) 03:16–04:55 UTC on January 13, 2017 (a) and (b) 05:38–07:17 UTC on January 14, 2017. The black and gray colormap represents the matched CloudSat radar reflectivity.

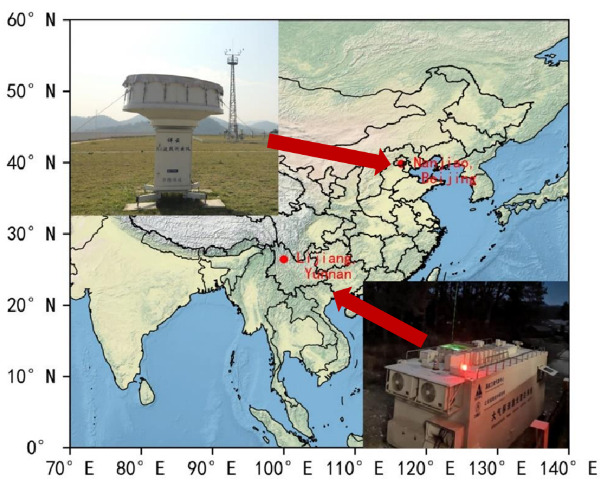

Figure 4. Geographical locations and photos of lidar and cloud radar at Yunnan Lijiang and Beijing Nanjiao stations.

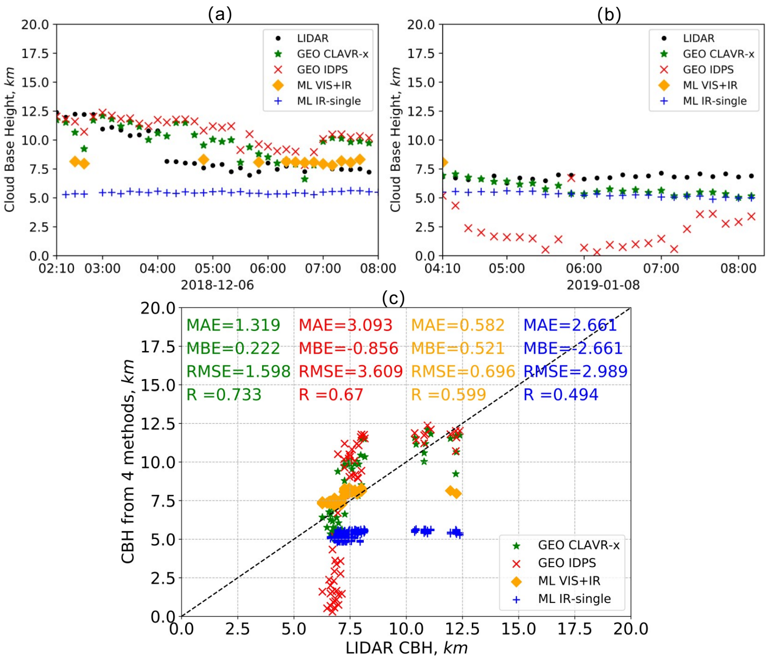

Figure 5. Comparisons of the CBHs from the ground-based lidar measurements (black solid circle) at Yunnan Lijiang station and the four GEO satellite retrieval algorithms, namely the GEO IDPS (red cross symbol), the GEO CLAVR-x (green solid asterisk), the ML-based VIS+IR model (orange solid diamond) and the ML-based IR-single model (blue plus sign) algorithms. Figure 5a and 5b show the time series of CBHs from lidar and the four GEO satellite retrieval algorithms on December 6, 2018 and January 8, 2019, respectively. Fig 5c shows the scatterplots of

CBH samples from the lidar measurements and the four retrieval algorithms.

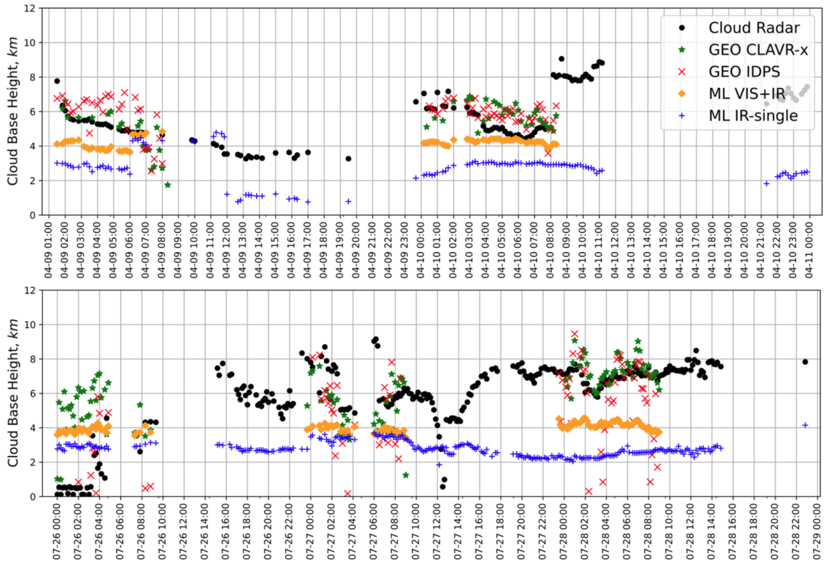

Figure 6. Same as Figure 5, but for the CBH sample results from the cloud radar at Beijing Nanjiao station (black solid circle) on April 9–10, 2017 (top panel) and July 26–28, 2017 (bottom panel).

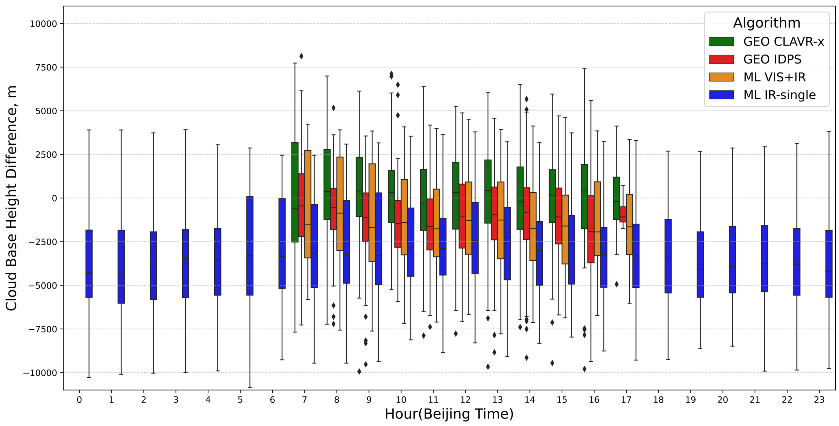

Figure 7. Box plots of the hourly CBH errors of four GEO satellite retrieval algorithms (GEO IDPS, GEO CLAVR-x, ML-based VIS+IR and ML-based IR-single) relative to the CBHs from the cloud radar at Beijing Nanjiao station in 2017. The box symbols signify the 25th, 50th and 75th percentiles of errors. The most extreme sample points between the 75th and outlier, and the 25th percentiles and outliers are marked as whiskers and diamonds, respectively. Except for the period between 7 and 17 (local time), the three algorithms of GEO CLAVR-x, GEO IDPS, and ML VIS+IR are unavailable due to the lack of reflected solar radiance measurements.

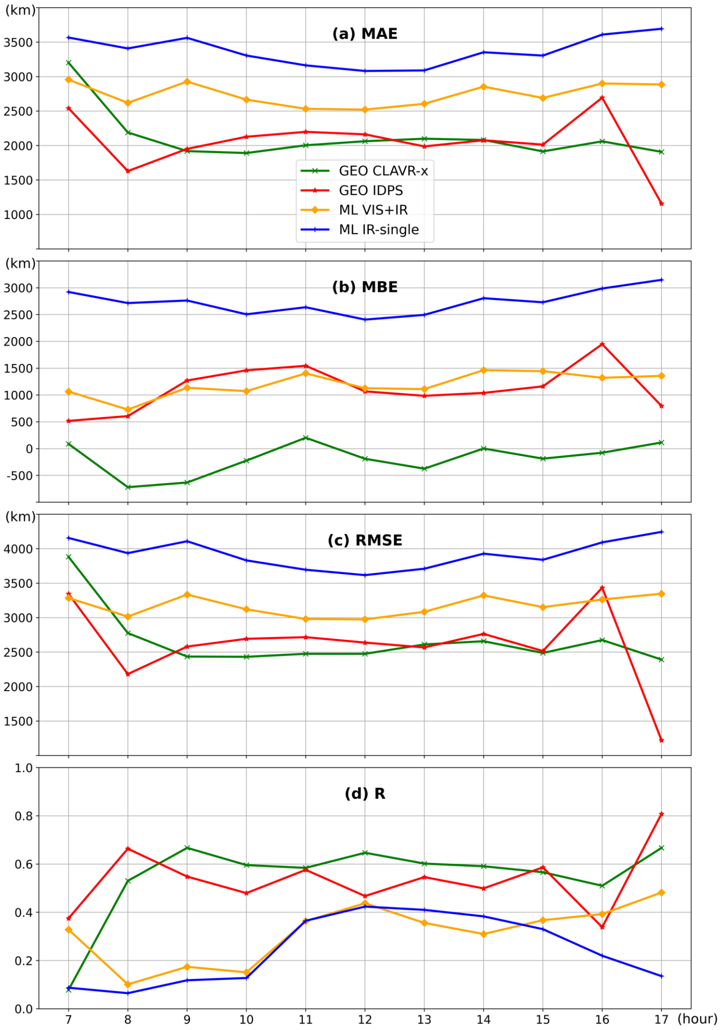

Figure 8. Comparisons of hourly (a) MAE, (b) MBE, (c) RMSE, and (d) R of CBH (relative to the CBHs from the cloud radar at Beijing Nanjiao station) from 07 to 17 (local time) between four retrieval algorithms (GEO IDPS, GEO CLAVR-x,

ML-based VIS+IR and ML-based IR-single) in 2017.

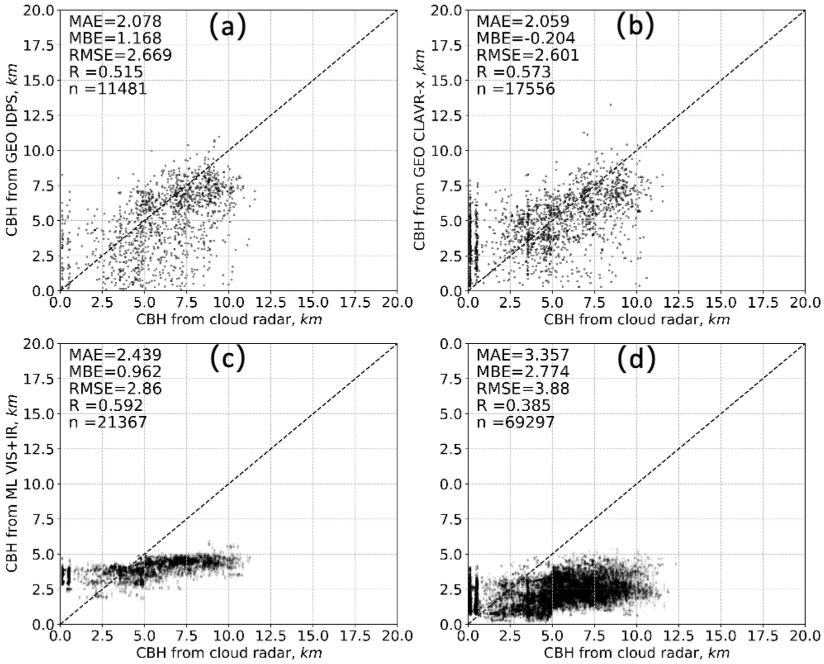

Figure 9. Comparisons between the CBHs from the cloud radar at Beijing Nanjiao station and the matched CBHs from the four retrieval algorithms (GEO IDPS, GEO CLAVR-x, ML-based VIS+IR and ML-based IR-single) in 2017.