~~AGCPP: ~~All-day ~~G~~global ~~C~~cloud ~~P~~physical ~~P~~properties products~~dataset~~ with 0.07° resolution retrieved from geostationary satellite imagers covering the period from 2000 to 2022

Lingxiao Zhao[1], Feng Zhang[1], Zhijun Zhao[1]~~Jingwei Li[1]~~, Feng Lu[2], Jingwei Li[1], Bin Guo[3], and Wenwen Li[4]~~and Zhijun Zhao[1]~~

[1]Key Laboratory of Polar Atmosphere-Ocean-Ice System for Weather and Climate of Ministry of Education / Shanghai Key Laboratory of Ocean-Land-Atmosphere Boundary Dynamics and Climate Change, Department of Atmospheric and Oceanic Sciences & Institutes of Atmospheric Sciences, Fudan University, Shanghai, China
[2]CMA Key Laboratory for Cloud Physics, Weather Modification Center, China Meteorological Administration (CMA), Beijing, China
[3]College of Physics and Electronic Information Engineering, Zhejiang Normal University, Jinhua, China
[4]Engineering Research Center of Optical Instrument and System, the Ministry of Education, Shanghai Key Laboratory of Modern Optical System, University of Shanghai for Science and Technology, Shanghai, China

Correspondence to: Feng Zhang (fengzhang@fudan.edu.cn)

Abstract. ~~The use of remote sensing to accurately measure cloud properties and their spatial and temporal variability has become an important area of atmospheric science research. However, the heterogeneity of data formats across national agencies and the calibrate and navigate associated with the use of data from different agencies have prevented the climate research community from using the full continuum of global cloud physical properties products. In this paper, All-day Global Cloud Physical Properties (AGCPP) is proposed, which provides cloud physical properties covering nearly the entire globe, from latitude -70° to 70° and longitude -180° to 180°. The main attributes of this dataset include cloud phase, cloud top height, cloud optical thickness, and cloud effective radius, with a time range from 1 January 2000 to 31 December 2022. AGCPP combines the observational advantages of geostationary satellites and polar-orbiting satellites. It uses the Moderate Resolution Imaging Spectroradiometer (MODIS) Level-2 cloud product (MOD06/MYD06) to train the cloud-based attention-UNet (CloudAtUNet) model, and then evaluates AGCPP using MOD06/MYD06 and the Cloud–Aerosol Lidar with Orthogonal Polarisation (CALIOP) 1 km cloud layer product. The evaluation results indicate that AGCPP demonstrates excellent continuity and consistency in both temporal and spatial accuracy, as well as high consistency in diurnal accuracy. Due to the long time series and all-day global nature of the dataset, it is expected that the dataset AGCPP will significantly increase the potential for climate change research, particularly with respect to potential feedback effects between clouds, surface albedo, and radiation. AGCPP is stored in the Network Common Data Format (netCDF), a standard that allows various tools and libraries to process the data quickly and easily. The AGCPP dataset is freely available on the Science Data Bank at https://doi.org/10.57760/sciencedb.26292 (Zhao et al., 2025), and the corresponding code can be found at https://github.com/lingxiao-zhao/AGCPP (last access: 25 June 2025).~~

Clouds play a crucial role in the Earth’s energy budget and the hydrological cycle. However, differences in the spatiotemporal resolution of satellite sensors and in retrieval algorithms lead to substantial heterogeneity among retrieved cloud products. Therefore, this study is based on global geostationary satellite thermal infrared brightness temperature data from the Gridded Satellite (GridSat-B1) project. Therefore, this study utilises the single-layer cloud inversion model small attention-UNet (Cloud-SmaAtUNet) within the DaYu Cloud Analysis System (DaYu-CLAS). Based on global geostationary satellite thermal infrared channel brightness temperature data from the Gridded Satellite (GridSat-B1) project, it has retrieved the first set of all-day Global Cloud physical properties Products (GCP), referred to as DaYu-GCP. This dataset achieves a temporal resolution of 3 h, spatial resolution of 0.07°, and a temporal span of 23 years (2000–2022). The DaYu-GCP includes cloud phase (CLP), cloud top height (CTH), cloud optical thickness (COT), and cloud effective radius (CER), covering all regions between 70°S–70°N in latitude and 180°W–180°E in longitude. Evaluation based on the Moderate-resolution Imaging Spectroradiometer (MODIS) official cloud products shows that the annual CLP identification accuracy of DaYu-GCP remains stable at 85% ± 0.7%, while the annual RMSE for CTH, COT, and CER stabilize at 1.50 ± 0.03 km, 10.71 ± 0.15, and 6.75 ± 0.10 μm, respectively. The multi-year variations in accuracy are within 2%, with no evident interannual differences, and the spatiotemporal distributions are continuous. In addition, evaluation based on observations from the Cloud Profiling Radar and the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) indicates that the DaYu-GCP products exhibit no significant day–night differences and maintain high accuracy under both conditions. Furthermore, the DaYu-GCP products are compared with other global cloud products. Taking the Northern Hemisphere as an example, the interannual variations of cloud cover frequency (CCF), CTH, COT, and CER retrieved from DaYu-GCP show correlation coefficients of 0.760, 0.486, 0.764, and 0.514 with the ISCCP product, respectively, and 0.444, 0.778, 0.171, and 0.412 with the CLARA-A3 product. The DaYu-GCP dataset, which is stored in the Network Common Data Format (NetCDF), is freely available on the Science Data Bank at https://doi.org/10.57760/sciencedb.26292 (Zhao et al., 2026). The corresponding code can be found at https://github.com/lingxiao-zhao/DaYu-GCP (last access: 25 June 2025).

1 Introduction

~~Cloud processes occupy a central role throughout the lifecycle of both severe convective storms and tropical cyclones. From their initial formation and growth to eventual dissipation, the evolution of cloud microphysics exerts a direct control on precipitation intensity and its spatial distribution (Zhuge and Zou, 2018; Yan et al., 2024). Likewise, the structural evolution and intensity fluctuations of tropical cyclones are intimately tied to ice–water phase transitions and the size distribution of cloud droplets within convective cores (Zhuge et al., 2015; Hsieh et al., 2024). Beyond modulating rainfall patterns, cloud layers also regulate Earth’s radiation budget by scattering incoming shortwave radiation and absorbing outgoing longwave radiation (Liu et al., 2024b; Viggiano et al., 2025), and they play a pivotal role in the global hydrological cycle (Liu et al.,~~

~~2024b; Viggiano et al., 2025). Consequently, acquiring high-quality, global-scale observations of cloud physical properties and resolving their spatiotemporal variability remains indispensable for advancing both weather forecasting and climate-system research.~~

Clouds within the Earth–atmosphere system play a vital role in regulating the planetary radiation balance (Wang et al., 2024; Zhang et al., 2025a) and in the global hydrological cycle (Liu et al., 2024; Viggiano et al., 2025). Therefore, obtaining a high-accuracy, long-term global dataset of cloud physical properties is essential for weather forecasting and climate-change research (Shi et al., 2025; Letu et al., 2023; Tang et al., 2025).

~~Currently, polar-orbiting satellite sensors, such as the Aqua and Terra (Platnick et al., 2015) satellites equipped with the Moderate Resolution Imaging Spectroradiometer (MODIS), have been continuously providing high-spatial-resolution (approximately 1 km) physical property datasets for global cloud cover since 2000. Compared to geostationary satellites operating in geosynchronous orbits, polar-orbiting satellites, due to their low Earth orbits of approximately 700 km, can obtain more accurate brightness temperature and cloud property retrieval data (Frey et al., 2008). These data are widely used in studying the interactions between clouds and climate change (Brennan et al., 2005; Kaps et al., 2023). For example, MODIS's cloud mask algorithm supports multiple spectral bands (day and night compatible) and has been verified through radar/lidar experiments to have higher accuracy than the Advanced Very High Resolution Radiometer (AVHRR) (Liu et al., 2004)). However, polar-orbiting satellites only scan along the Earth's poles in narrow bands (approximately 2,000 km wide), making it impossible to achieve continuous, comprehensive observations of the global cloud field (Menzel et al., 2008). In contrast, geostationary satellites (orbiting at an altitude of approximately 36,000 km) can continuously monitor approximately one-third of the Earth's surface day and night, providing high-frequency observations at minute intervals for long-term cloud changes.~~

Satellite remote sensing is the primary means of obtaining cloud physical properties. Among these platforms, geostationary satellites can continuously monitor approximately one-third of the Earth’s surface day and night, providing high-frequency observations at the minute scale for long-term cloud variability studies. For example, the Advanced Geostationary Radiation Imager (AGRI) onboard the FengYun (FY)-4A/B satellites operated by the National Satellite Meteorological Centre of the China Meteorological Administration (NSMC–CMA) (Min et al., 2017; Min et al., 2020), as well as the Advanced Himawari Imager (AHI) onboard the Himawari-8/9 (H8/9) satellites operated by the Japan Aerospace Exploration Agency (JAXA), can monitor East Asia and the Pacific region (Bessho et al., 2016a, b; Iwabuchi et al., 2018). The Spinning Enhanced Visible and Infra-Red Imager (SEVIRI) onboard the Meteosat Second Generation (MSG) satellites operated by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) provides observations over Africa and Europe (Donny Maladji et al., 1997; Coste et al., 2017; Kocaman et al., 2022). The Geostationary Operational Environmental Satellites (GOES)-R series operated by the National Oceanic and Atmospheric Administration (NOAA), including GOES-16 to GOES-19, are equipped with the Advanced Baseline Imager (ABI) to monitor the Americas (Bin et al., 2018, 2019; Bin et al., 2020; Heidinger et al., 2020). These geostationary satellite sensors provide observations every 10–15 min, with spatial resolutions of 0.5–1 km in the visible channels and 2–5 km in the infrared channels. In contrast to geostationary satellites, polar-orbiting satellites, such as the Moderate-resolution Imaging Spectroradiometer (MODIS) onboard Aqua and Terra (Platnick et al., 2015), cannot provide high-frequency continuous observations over a given region; however, since 2000 they have offered observations with higher spatial resolution (0.25–1 km).

~~However, due to the higher orbital altitude of geostationary satellites, the accuracy of the brightness temperature and cloud physical property retrievals obtained is slightly inferior to that of polar orbiting satellites (Zhang et al., 2021). Additionally, the sensors on geostationary satellites from various countries each have their own limitations: The Advanced Himawari Imager (AHI) on Japan's Himawari-8 (Wang et al., 2024) and the Meteosat Second Generation/Spinning Enhanced Visible and Infra-Red Imager (MSG/SEVIRI) on European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) (Poulsen et al., 2012; Watts et al., 2011) only release daytime cloud physical property products using the Optimal Cloud Analysis (OCA) algorithm, lacking nighttime products; China's Fengyun series (FY-4A/B) Advanced Geostationary Radiation Imager (AGRI) has also not yet released nighttime cloud microphysical parameter products including cloud optical thickness and cloud effective radius (Chen et al., 2020; Zhou et al., 2024); the U.S. GOES 16/17 series Advanced Baseline Imager (ABI) provides cloud microphysical parameters covering both daytime and nighttime, but due to the use of two separate Cloud Optical and Microphysical Properties (COMP) retrieval algorithms for daytime and nighttime (Walther and Heidinger, 2012), the algorithm switch causes discontinuous jumps or biases in the product during day-night transitions, introducing false signals in long-term climate trend analyses at timescales longer than a day (Smalley and Lebsock, 2023).~~

As shown in Table 1, based on satellite observation data, these sensors all provide official cloud physical characteristics product datasets, such as official products from the AGRI, AHI, ABI, and SEVIRI, as well as datasets from research initiatives like the Cloud Remote Sensing, Atmospheric Radiation and Renewable Energy (CARE). Most include physical characteristics such as cloud phase (CLP), cloud top height (CTH), cloud optical thickness (COT), and cloud effective radius (CER), with spatial resolutions of 2–5 km and temporal resolutions of 10–15 minutes. However, these are all regional cloud products, and most lack night-time cloud coverage. In addition, channel spectral responses differ among sensors, and the official cloud product algorithms also vary across platforms. For example, the official AHI cloud products are mainly retrieved using the Comprehensive Analysis Program for Cloud Optical Measurement (CAPCOM) multifunctional algorithm system, which integrates multi-channel threshold methods and a dual visible–near-infrared lookup table (LUT) approach. For liquid water clouds, the Mie–Lorenz scattering model is applied (Nakajima and Nakajma, 1995; Kawamoto et al., 2001), while for ice clouds, an extended Voronoi irregular ice crystal scattering model is used (Letu et al., 2020b), enabling daytime cloud detection and the retrieval of COT and CER (Imai and Yoshida, 2016; Mouri et al., 2016). The ABI official products are developed by the GOES-R Algorithm Working Group, which employs LUT-based retrievals constructed from visible and near-infrared radiances during daytime, while nighttime retrievals rely on thermal infrared channels (Walther and Heidinger, 2012; Walther et al., 2013; Minnis and Heck, 2012), ultimately achieving the retrieval of CLP, CTH, COT, and

CER (Pavolonis, 2010; Heidinger, 2012). These differences prevent the direct integration of official geostationary satellite cloud products into a spatiotemporally continuous global cloud product.

~~The limitations of using geostationary satellite data to derive cloud physical products are primarily due to the shortcomings of traditional retrieval algorithms. Firstly, daytime retrievals primarily rely on the visible and shortwave infrared dual-spectral bands (e.g., DCOMP and OCA algorithms). These methods can only accurately estimate cloud optical thickness and particle size under conditions with solar radiation, but they often fail or experience a significant drop in accuracy during twilight, dawn, and nighttime conditions (Wolters et al., 2008). Second, while traditional nighttime methods (such as Optimal Estimation (Iwabuchi et al., 2014) or infrared split-window (Heidinger and Pavolonis, 2009)) can be used at night, they are limited by the penetration capability of thermal infrared radiation, leading to systematic biases in estimating the microphysical properties of thick clouds or highly reflective cloud layers (Mayer et al., 2024).~~

Although several global cloud physical property datasets have been developed, such as the third edition of the Satellite Application Facility on Climate Monitoring's (CM SAF) cloud, albedo, and surface radiation dataset from Advanced Very High Resolution Radiometer (AVHRR) observations (CLARA-A3), which retrieves cloud amount, CTH, COT, and CER based on AVHRR measurements, its coverage is global but the temporal resolution is limited to 24 h and the spatial resolution to 0.25° (Karlsson et al., 2023b; Karlsson et al., 2023a). The International Satellite Cloud Climatology Project (ISCCP) uses AVHRR and approximately 10 km geostationary imagery to produce cloud amount, cloud types, cloud top temperature (CTT), and COT products. However, the D series (3 h; 2.5°) (Schiffer and Rossow, 1983; Rossow and Schiffer, 1991; Rossow et al., 1985) and H series (3 h; 1°) (Young et al., 2018; Rossow et al., 2022) were discontinued in 2009 and 2017, respectively. The National Oceanic and Atmospheric Administration (NOAA) Satellite ClOud and Radiation Property retrieval System (SatCORPS) employs multiple sensors, such as AHI and SEVIRI, to generate CLP, CTH, COT, and CER products, and although its temporal and spatial resolutions can reach 1 h and 3 km, respectively, the data are currently only available from 2023 onward (Trepte et al., 2019; Yost et al., 2021). These facts indicate that existing global cloud products are unable to simultaneously achieve both long temporal coverage and high spatiotemporal resolution.

~~In recent years, advances in high-performance computing and artificial intelligence have promoted the application of machine learning and deep learning methods in the field of cloud property retrieval. Pérez et al. (2009) first used neural networks to retrieve the MODIS infrared radiation model to support nighttime microphysical processes. Subsequent studies further constructed a unified day-night retrieval model based on cross-channel feature learning to improve retrieval accuracy and efficiency (Lee et al., 2021; Kurihana et al., 2022; Kotarba and Wojciechowska, 2025; Gao et al., 2024), particularly achieving significant improvements under thick cloud conditions (Zhao et al., 2023; Min et al., 2020). However, these methods are only applicable to sensors with similar orbits. Additionally, since each geostationary satellite only covers a regional area and stores data in fragmented, heterogeneous archives, creating a global continuous cloud layer attribute product poses significant challenges in data collection and pre-processing. For example, calibrating raw sensor signals to radiant brightness (Helder et al., 2020; Lee et al., 2024) and brightness temperature, as well as navigation to map each pixel to Earth surface coordinates (Knapp et al., 2011; Jiao et al., 2024). To address this issue, some studies have begun to~~

~~integrate polar orbiting satellite data with geostationary satellite observations (Tong et al., 2023; Li et al., 2023; Zhao et al., 2024; Liu et al., 2025). The Gridded Satellite (GridSat-B1) project pioneered a truly global, spatio-temporally continuous brightness temperature dataset by stitching together infrared channel data from multiple geostationary satellites (Knapp et al., 2011), laying the foundation for seamless, long-term climate analysis products (Shi et al., 2025; Letu et al., 2023; Tang et al., 2025).~~

~~The objective of this study was to combine the high spatio-temporal resolution brightness temperature observations of geostationary satellites with the high-precision advantages of polar-orbiting satellite cloud products to construct the All-day Global Cloud Physical Properties (AGCPP).First, this study constructed a large-scale training sample by matching the infrared brightness temperature of GridSat-B1 and ERA5 meteorological fields with the high-resolution cloud physical products of the MODIS satellite at the pixel level. Then, a deep neural network was trained to learn the mapping relationship. Finally, the model weights obtained from the training were applied to the entire GridSat-B1 brightness temperature time series to produce the AGCPP dataset. The innovation of this study lies in the fact that AGCPP is the first global dataset with a spatial resolution of 0.07° and a temporal resolution of 3 h. As shown in Table 1, this is the latest cloud product dataset currently available worldwide. Firstly FY-4B cloud products, CARE, GOES-R ABI cloud products, Himawari-8 cloud products are regional cloud products while AGCPP is a global cloud product. Secondly International Satellite Cloud Climatology Project (ISCCP) and CLARA-A3 have coarser spatial resolution of 0.3° and 0.25° and AGCPP has 0.07°. Finally, SatCORPS Global Cloud Product starts only from 2023 and is not capable to do long time climate analysis, while AGCPP covers 23 years from 2000-2022. Due to the long-term time series and all-day global characteristics of AGCPP dataset, it is anticipated that the dataset will significantly enhance the potential for climate change research, particularly studies on the potential feedback effects between clouds, surface albedo, and radiation.~~

Therefore, this study employs the single-layer cloud retrieval model small attention-UNet (Cloud-SmaAtUNet) within the DaYu Cloud Analysis System (DaYu-CLAS) to develop an all-day Global Cloud physical properties Products (GCP) retrieval algorithm. The DaYu-CLAS Cloud-SmaAtUNet model takes globally mosaicked thermal infrared channel brightness temperature data from the Gridded Satellite (GridSat-B1) project as input, and is trained using MODIS official cloud products as labels. This approach not only combines the global coverage of the geostationary satellite product GridSat-B1 with the high-precision observation advantages of the polar-orbiting satellite MODIS, but also maximises the extraction of spatial structural information within clouds, thereby enabling all-day, high-accuracy global retrievals of CLP, CTH, CER, and COT products, collectively referred to as DaYu-GCP. The DaYu-GCP products achieve a temporal resolution of 3 h, a spatial resolution of 0.07°, and a time span of 23 years (2000–2022), covering all regions between 70°S–70°N in latitude and 180°W–180°E in longitude, and can provide essential data support for studies of radiative balance, the hydrological cycle, and related processes.

Table 1. ~~Comparison results between our dataset~~ Specifications of our DaYu-GCP dataset and the latest cloud physical property product dataset.

~~Agencies~~	~~Names~~	~~Region~~	~~Latitude~~	~~Longitude~~	~~Products~~	~~Spatial resolution~~	~~Time resolution~~	~~Years~~
~~Ours~~	~~AGCPP~~	~~Global~~	~~70° S–70° N~~	~~180° W–180° E~~	~~cloud phase, cloud top height, cloud optical thickness, cloud effective radius~~	~~0.07°~~	~~3 h~~	~~2000–2022~~
~~National Satellite Meteorological Centre, China Meteorological Administration (NSMC-CMA) (Zhang et al., 2024)~~	~~FY-4B cloud products~~	~~China/Full Disc~~	~~80° S–80° N~~	~~23.8° E–173.8° W~~	~~cloud phase, cloud top height, cloud optical thickness (daytime), cloud effective radius (daytime)~~	~~0.036°(macro)/ 0.018°(micro)~~	~~0.25 h (macro)/1 h (micro)~~	~~2018-present~~
~~Aerospace Information Research Institute, Chinese Academy of Sciences (AIRI-CAS) (Letu et al., 2020)~~	~~CARE~~	~~East Asia–Pacific region~~	~~10° S–60° N~~	~~60° E–180° E~~	~~cloud phase, cloud top height, cloud optical thickness (daytime), cloud effective radius (daytime)~~	~~0.1°~~	~~0.5 h~~	~~2016–present~~
~~National Oceanic and Atmospheric Administration (NOAA) (Heidinger et al., 2020)~~	~~GOES-R ABI cloud products~~	~~Western Hemisphere (Full Disc, CONUS, mesoscale)~~	~~80° S–80° N~~	~~142° E–56° W/156° W–6° E~~	~~cloud phase, cloud top height, cloud optical thickness and cloud effective radius~~	~~0.018°~~	~~5–15 min (FD),~~ ~~5 min (CONUS),~~ ~~30–60 s (Mesoscale)~~	~~2018–present~~
~~Japan Aerospace Exploration Agency (JAXA) (Mouri, 2019)~~	~~Himawari-8 cloud products~~	~~East Asia–Pacific region~~	~~60° S–60° N~~	~~70° E–150° W~~	~~cloud phase, cloud top height, cloud optical thickness (daytime), cloud effective radius (daytime)~~	~~0.045°~~	~~0.25 h (10 min)~~	~~2015–present~~
~~National Aeronautics and Space Administration (NASA) (Young et al., 2018)~~	~~ISCCP (DX, D1, C1, H-series)~~	~~Global~~	~~90° S–90° N~~	~~180° W–180° E~~	~~cloud phase, cloud top height, cloud optical thickness~~	~~DX: 0.3°, C1/D1: 2.5°, H: 1°~~	~~3 h (DX, D1, C1), 1 month (H)~~	~~DX/D1/C1: 1983–2009, H: 1983–2017~~
~~EUMETSAT (Karlsson et al., 2023b; Karlsson et al., 2023a)~~	~~CLARA-A3~~	~~Global~~	~~90° S-90° N~~	~~180° W-180° E~~	~~cloud phase, cloud top height, cloud optical thickness, cloud effective radius (daytime)~~	~~0.25°~~	~~24 h/ 1 month~~	~~1979-2020~~
~~NASA (Minnis et al., 2008; Minnis et al., 2021)~~	~~SatCORPS Global Cloud~~	~~Global~~	~~90° S-90° N~~	~~180° W-180° E~~	~~cloud phase, cloud top height, cloud optical thickness (daytime), cloud effective radius~~	~~0.027°~~	~~1h~~	~~2023-present~~

~~Product~~	~~(daytime)~~

Agency	Source	Region	Latitude	Longitude	Products	Spatial resolution	Time resolution	Years (Reference)
Fudan University (DaYu-GCP)	GridSat-B1	Global	70° S–70° N	0° E-360° E	cloud phase, cloud top height, cloud optical thickness, cloud effective radius	0.07°	3 h	2000-2022
National Satellite Meteorological Centre, China Meteorological Administration	FY4A/B-AGRI	China/ Full Disc	81° S–81° N	23.8° E-186° E	cloud detection, cloud types, cloud top height, cloud top temperature	4 km	15 min	2018-present (Min et al., 2017; Min et al., 2020)
Aerospace Information Research Institute, Chinese Academy of Sciences (CARE)	FY4A/B-AGRI; H8-AHI	East Asia–Pacific region	60° S–60° N	60° E-180° E	cloud phase, cloud top height, cloud optical thickness, cloud effective radius	0.1°	Daytime: 0.5 h	2016 (Letu et al., 2020a)
National Oceanic and Atmospheric Administration	GOES16/17/18/19-ABI	Western Hemisphere/Full Disc	81° S–81° N	142° E-304° E	cloud phase, cloud top height, cloud optical thickness and cloud effective radius	2 km	1 h	2017–present (Bin et al., 2018, 2019; Bin et al., 2020; Heidinger et al., 2020)
Japan Aerospace Exploration Agency	H8/9-AHI	East Asia–Pacific region	60° S–60° N	80° E-200° E	cloud phase, cloud top height, cloud optical thickness, cloud effective radius	5 km	Daytime: 10 min	2015-present (Bessho et al., 2016b; Letu et al., 2016; Iwabuchi et al., 2018)
European Organisation for the Exploitation of Meteorological Satellites	MSG-SEVIRI	European region	81° S–81° N	81° W-81° E	cloud phase, cloud top height (9km), cloud optical thickness, cloud effective radius	3 km	15 min	2013-present (Donny Maladji et al., 1997; Coste et al., 2017; Kocaman et al., 2022)

National Aeronautics and Space Administration (ISCCP)	AVHRR and ~10 km geostationary imagery	Global	90° S–90° N	0° E-360° E	cloud amount, cloud types, cloud top temperature, cloud optical thickness	D: 2.5°	3 h	1983–2009 (Schiffer and Rossow, 1983; Rossow and Schiffer, 1991; Rossow et al., 1985)
						H: 1°	3 h	1983–2017 (Young et al., 2018; Rossow et al., 2022)
European Organisation for the Exploitation of Meteorological Satellites (CLARA-A3)	AVHRR	Global	90° S-90° N	0° E-360° E	cloud amount, cloud top height, cloud optical thickness, cloud effective radius	0.25°	24 h/ 1 month	1979-2020 (Karlsson et al., 2023b; Karlsson et al., 2023a)
National Aeronautics and Space Administration (SatCORPS)	H8-AHI; MSG-SEVIRI; GOES16/17/18/19-ABI; MODIS-Aqua/Terra	Global	90° S-90° N	0° E-360° E	cloud phase, cloud top height, cloud optical thickness, cloud effective radius	3 km	1 h	2023-present (Trepte et al., 2019; Yost et al., 2021)

~~This paper provides further details on the AGCPP dataset records, including input data, algorithm explanations, product examples, and validation results.~~ This paper provides further details on the DaYu-GCP dataset records, including input data, algorithm explanations, product examples, and validation results. Section 2 briefly introduces data preparation and methods, while Section 3 introduces, discusses and evaluates~~evaluation~~ four major products: CLP, CTH, COT and CER. ~~the four main product groups: cloud phase (CLP), cloud top height (CTH), cloud optical thickness (COT), and cloud effective radius (CER).~~ Additionally, DaYu-GCP was compared with existing global cloud datasets ISCCP and CLARA-A3. Section 4 is the conclusion. Section 5 describes the availability of the data.~~Section 4 presents the basic characteristics of AGCPP. Section 5 describes the availability of the data. Section 6 is the conclusion.~~

2 Data and methods

2.1 Data

~~2.1.1. Geostationary satellite data~~

This study employed two datasets: the geostationary satellite integrated data GridSat-B1 and the polar-orbiting satellite data MODIS. Further details are provided in .

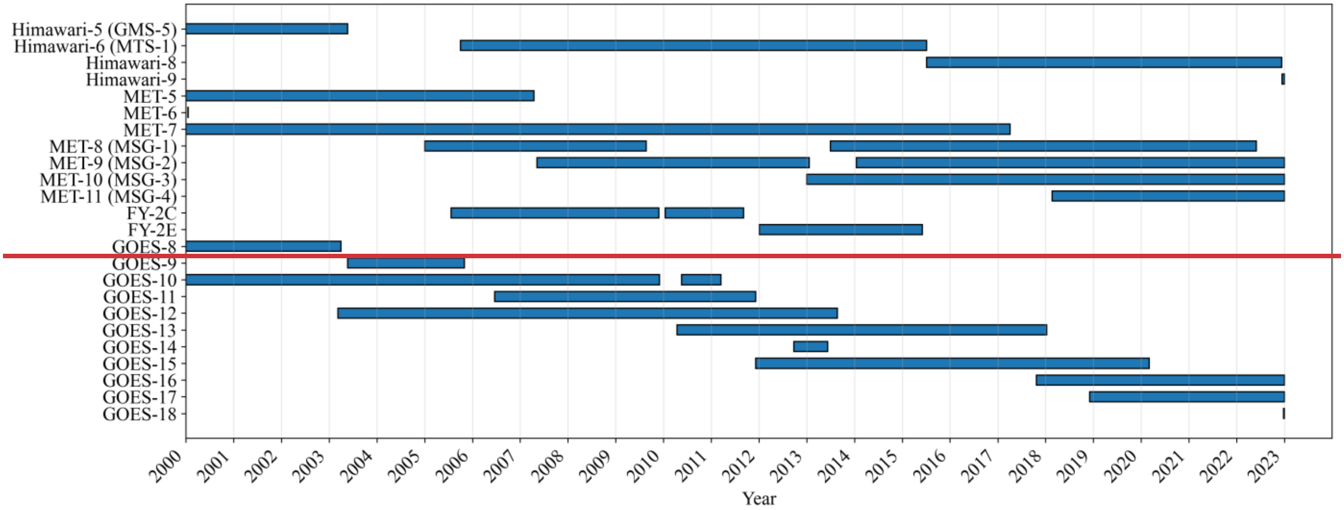
2.1.1. GridSat-B1 data

The primary source of geostationary infrared-channel brightness temperature data used in this study is the GridSat-B1 dataset. Over the 23-year study period (2000–2022), a total of 24 satellites from four countries or regions contributed to the construction of the GridSat-B1 dataset, as shown in Fig. 1. This global dataset integrates observations from multiple geostationary satellites to provide top-of-atmosphere (TOA) infrared brightness temperatures from two bands: the infrared water-vapor band (IRWVP, 6.7 μm) and the infrared window band (IRWIN, 11 μm). In addition to IRWVP and IRWIN, and in consideration of the physical generation mechanisms and developmental processes of clouds, the satellite identifiers (Satid) associated with the infrared brightness temperature data, together with the corresponding satellite zenith (SAZ) angles, have been included in the input data. The parsing procedure for Satid and SAZ in the auxiliary data is described in the GridSat-B1 documentation (Knapp et al., 2011). Additionally, the official documentation indicates that zenith-angle correction for infrared brightness temperature images may be referenced in Joyce et al. (2001). Because the IRWVP and IRWIN grid integrated products may originate from different satellites, Satid is divided into IRWVP satellite identifiers (Satid_VP) and IRWIN satellite identifiers (Satid_IN), and SAZ is similarly separated into IRWVP SAZ (SAZ_VP) and IRWIN SAZ (SAZ_IN). GridSat-B1 has a standard spatial grid resolution of 0.07° and a temporal resolution of 3 h, corresponding to standard synoptic times of 0000, 0300, ..., 2100 UTC. For each grid point, the measurement closest to the satellite nadir is selected. Missing data at these specified times are supplemented from ISCCP (Young et al., 2018) using the temporally nearest available image.

~~Since the 1970s, geostationary satellites have been regularly providing high-temporal-resolution Earth observation data. However, conducting climate research using their extensive historical data has typically faced significant obstacles. Key challenges include the absence of a global central repository integrating all international satellite data, the difficulty of processing massive amounts of spatiotemporal resolution data, and the heterogeneity of calibration and navigation formats across different satellite systems. These factors have added complexity to the unified processing required for multi-satellite climate research. To overcome these limitations, institutions such as the National Oceanic and Atmospheric Administration (NOAA) National Climatic Data Center (now NCEI) have remapped data to standard projections, recalibrated them to improve temporal consistency, and ultimately created datasets such as GridSat-B1 (Knapp et al., 2011).~~

~~The primary source of the geostationary infrared channel brightness temperature data used in this study is the GridSat-B1 dataset. Over the 23-year study period from 2000 to 2022, a total of 24 satellites from four countries or regions~~

~~participated in the construction of the GridSat-B1 dataset, as shown in Fig. 1. This global dataset integrates observational data from multiple geostationary satellites to provide top-of-atmosphere (TOA) infrared brightness temperature (from two bands at infrared water vapor (IRWVP) 6.7 μm and infrared window (IRWIN) 11 μm). Specifically, for each grid point, the measurement closest to the satellite's nadir is selected. GridSat-B1 features a standard grid resolution of 0.07° (8 km) and a temporal resolution of 3 hours, corresponding to standard weather observation times of 0000, 0300, ..., 2100 UTC. Image acquisition is conducted within 15 minutes after the start of these weather forecast periods. Missing data at these specified time points are supplemented by the ISCCP (Young et al., 2018) using the best available image temporally closest to the target time period.~~



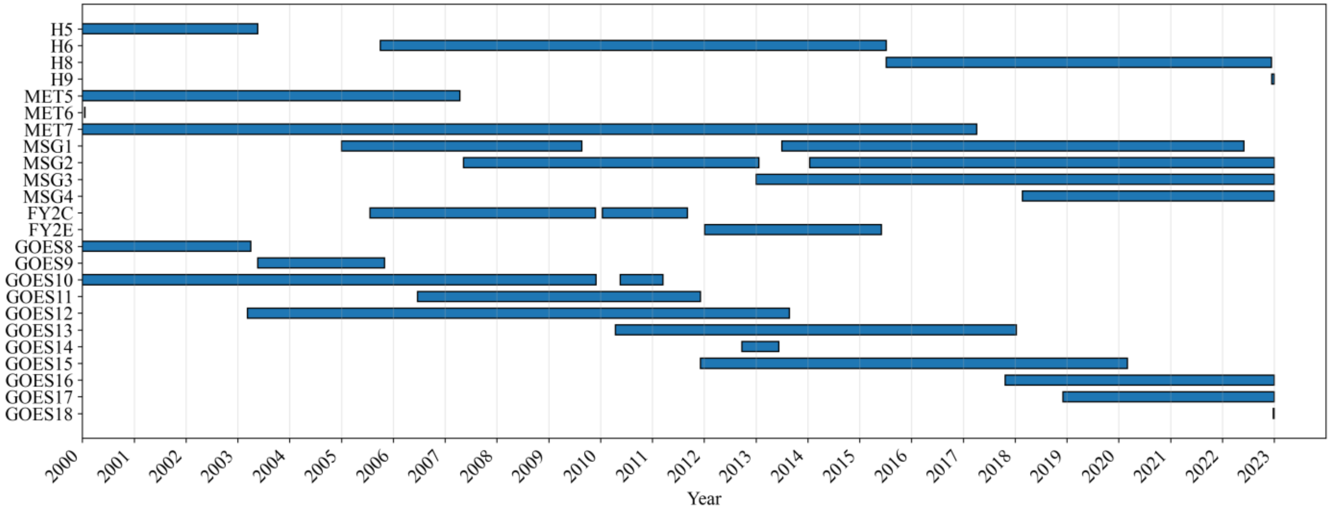


Figure 1.

Gantt chart of the 24 satellites contributing to the GridSat-B1 project during the study period.

~~Gantt chart of 24 satellites within the research time interval and corresponding participation times: Himawari-5 (GMS-5), Himawari-6 (MTS-1), Himawari-8, Himawari-9, MET-5, MET-6, MET-7, MET-8 (MSG-1), MET-9 (MSG-2), MET-10 (MSG-3), MET-11 (MSG-4), FY-2C, FY-2E, GOES-8, GOES-9, GOES-10, GOES-11, GOES-12, GOES-13, GOES-14, GOES-15, GOES-16, GOES-17, and GOES-18.~~

~~Due to the long time span and the large number of satellites and observation instruments involved, the primary challenge in constructing global TOA infrared brightness temperatures lies in ensuring consistency. GridSat-B1 effectively addresses radiation calibration and temporal consistency issues through the following methods: (1) This dataset first uses ISCCP calibration methods to preliminarily integrate data from different satellites. Specifically for infrared (IR) channels, GridSat-B1 also implements a secondary calibration process. Secondary calibration utilises high-resolution infrared radiation detectors (HIRS) as a reference standard, focusing on correcting systematic biases under low-temperature conditions. (2) To ensure the consistency and uniformity of long-term time series data, GridSat-B1 undergoes time normalisation processing. This process also utilises HIRS data as a calibration anchor point, effectively eliminating time offsets between different satellite observation systems, thereby significantly enhancing the temporal consistency of the entire IR brightness temperature historical dataset.~~

2.1.2. Polar orbit satellite data

The MODIS instrument operates aboard two polar-orbiting satellites: Terra (launched in December 1999) and Aqua (launched in April 2002). With 36 spectral channels and a global revisit frequency of 1–2 days, MODIS's broad spectral coverage supports a wide range of applications, including vegetation-health monitoring, land-cover classification, sea-surface temperature retrieval, and cloud analysis (Hosen et al., 2023; Cai et al., 2011; Menzel et al., 2008). In this study, cloud physical properties from the MODIS Level-2 cloud product (Collection 6.1) were used as training labels, specifically CLP, CTH, CER, and COT. This product, identified by the code "06", provides data from the Terra platform (MOD06) and the Aqua platform (MYD06). Owing to its well-characterized accuracy and high data quality, the MODIS Collection 6.1 product is widely utilized as a benchmark in remote-sensing studies (Zhang et al., 2017). The Cloud Profiling Radar and Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument, mounted on the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite, was launched in April 2006 and ceased operations in June 2023 (Winker et al., 2010). CALIPSO provides global observations of the vertical structure and characteristics of aerosols and thin clouds (Zhang et al., 2017; Hagihara et al., 2010). The assessment of model accuracy primarily utilized official cloud products from MODIS and CALIOP.

The CLP, CTH, COT, and CER selected from the MODIS official cloud products in this study originate from the Cloud_Phase_Infrared_1km, cloud_top_height_1km, Cloud_Optical_Thickness, and Cloud_Effective_Radius variables respectively. The CLP, CTH, and COT from the CALIOP official cloud product are derived from the Feature_Classification_Flags, Layer_Top_Altitude, and Column_Optical_Depth_Cloud_532 variables respectively. ~~MODIS comprises the Terra satellite, launched into a polar orbit in December 1999, and the Aqua satellite, launched into a polar orbit in April 2002. These satellites continuously collect data every 1–2 days across 36 spectral channels that cover the entire globe. Its exceptionally wide spectral range enables MODIS data to be used in a wide range of studies, including vegetation health, land cover, sea surface temperature, and cloud analysis (Hosen et al., 2023; Cai et al., 2011; Menzel et al., 2008). MOD represents Terra products, and MYD represents Aqua products. In this study, CLP, CTH, CER, and COT of MOD/MYD were used as training labels. Due to MODIS's lower orbital altitude and higher data quality, it is widely used as a ground truth label (Zhang et al., 2017).~~

~~The Cloud Profiling Radar and Cloud-Aerosol Lidar with Orthogonal Polarisation (CALIOP) lidar instrument, mounted on the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite, was launched in April 2006 and ceased operations in June 2023. CALIPSO provides global vertical structure and characteristic observation data for aerosols and thin clouds (Zhang et al., 2017; Hagihara et al., 2010). The assessment of model accuracy primarily utilised cloud products from MODIS and CALIOP.~~

~~2.1.3. Meteorological field and auxiliary data~~

~~Considering the physical generation mechanism and development process of clouds, some meteorological fields and satellite IDs (Satid) constituting infrared brightness temperature data, as well as the corresponding satellite zenith angle (SZA), have been added to the input data. The parsing process for Satid and SZA in the auxiliary data is mentioned in the documentation for GridSat-B1 (Knapp et al., 2011). Additionally, the official documentation notes that the zenith angle correction for infrared brightness temperature images can be referenced in the work of Joyce et al. (2001).~~

~~Since the data from IRWVP and IRWIN grid fusion may originate from different satellites, Satid is divided into IRWVP satellite IDs (Satid_VP) and IRWIN satellite IDs (Satid_IN), and SZA is divided into IRWVP SZA (SZA_VP) and IRWIN SZA (SZA_IN). Meteorological field data are sourced from the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5 (ERA5) (Hersbach et al., 2020). This study selected ERA5 hourly air temperature profiles (ATP), relative humidity profiles (RHP), surface skin temperature (SKT), total column water vapour (TCWV), and soil type with a spatial resolution of 0.25°. To prevent unnecessary model redundancy caused by overly dense pressure levels in the input data, which could affect the efficiency of model training and data production. ATP and RHP are each selected at four identical pressure levels: 1000, 850, 500, and 300 hPa.~~

2.1.3. ERA5 data

Meteorological field data are obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5 (ERA5) dataset (Hersbach et al., 2020). This study selected ERA5 hourly air temperature profiles (ATP), relative humidity profiles (RHP), skin temperature (SKT), total column water vapor (TCWV) and land cover with a spatial resolution of 0.25°. Excessively dense pressure levels in the input data may introduce unnecessary model redundancy and adversely affect training and operational efficiency. Therefore, ATP and RHP were each extracted at four identical pressure levels: 1000, 850, 500, and 300 hPa (Zhao et al., 2024a; Liu et al., 2025).

Table 2. Input, target and evaluation data preparation for building the dataset.

	Variable	Source	Spatial resolution	Temporal resolution
Input	TOA Brightness Temperature (2 bands: IRWVP 6.7 and IRWIN 11 µm) Satellite Zenith Angle (SAZ_VP and SAZ_IN) Satellite Index (Satid_VP and Satid_IN)	GridSat-B1	8 km	3 h
	Skin Temperature Total Column Water Vapor Air Temperature Profile (4 pressure levels: 1000, 850, 500, and 300 hPa) Relative Humidity Profile (4 pressure levels: 1000, 850, 500, and 300 hPa) Land Cover	ERA5	0.25°	1 h
Target	Cloud Phase Cloud Top Height Cloud Optical Thickness Cloud Effective Radius	Aqua and Terra/MODIS	1 km	5 min
Evaluation	Cloud Phase Cloud Top Height Cloud Optical Thickness	CALIPSO/CALIOP	1 km	/

~~2.1.4. Detailed information~~

~~Table 2 provides all the training and evaluation datasets used in this study. Due to the different institutional sources of the datasets, they may have different projection methods and spatio-temporal resolutions. To ensure the correct correspondence of pixels and data consistency, the data were first aligned to a unified 0.07° latitude and longitude grid~~

~~before model construction. The nearest neighbour interpolation method (Huang et al., 2012) was used for resampling the MODIS Level-2 cloud product (MOD06/MYD06), while the bilinear interpolation method (Kim et al., 2019) was used for resampling the ERA5 meteorological field. Due to differences in satellites and onboard sensors, there are differences in the spatial observation range and temporal observation frequency of the data. In this study, the input and target/evaluation data were matched in space and time to construct the infrared brightness temperature to cloud physical properties (IRBT2CPP) required for training. The following is a more detailed introduction to the data from different sources.~~

~~**Table 2.** Input, target and evaluation data preparation for building the dataset.~~

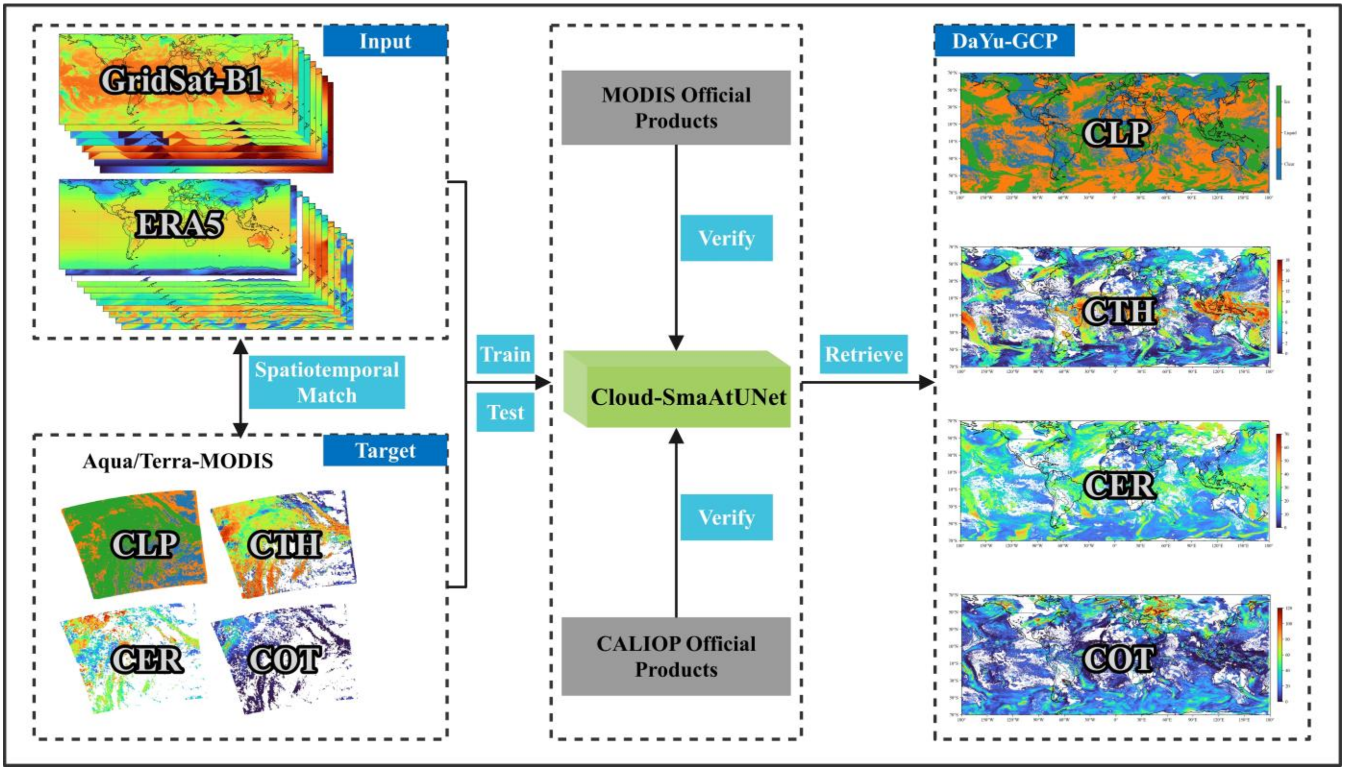
	~~Variable~~	~~Source~~	~~Spatial resolution~~	~~Temporal resolution~~
~~Input~~	~~TOA Brightness Temperature (2 bands: IRWVP 6.7 and IRWIN 11 μm)~~	~~Gridsat~~	~~8 km~~	~~3 h~~
	~~Satellite zenith angle (SZA_VP and SZA_IN)~~			
	~~Satellite Index (Satid_VP and Satid_IN)~~			
	~~Surface skin temperature~~	~~ERA5~~	~~0.25°~~	~~1 h~~
	~~Total column water vapor~~			
	~~Soil type~~			
	~~Air temperature profile (4 pressure levels: 1000, 850, 500, and 300 hPa)~~			
	~~Relative humidity profile (4 pressure levels: 1000, 850, 500, and 300 hPa)~~			
~~Target~~	~~Cloud phase~~	~~Aqua and Terra/MODIS~~	~~1 km~~	~~5 min~~
	~~Cloud top height~~			
	~~Cloud optical thickness~~			
	~~Cloud effective radius~~			
~~Evaluation~~	~~Cloud phase~~	~~Aqua and Terra/MODIS~~	~~1 km~~	~~5 min~~
	~~Cloud top height~~			
	~~Cloud optical thickness~~			
	~~Cloud effective radius~~			
	~~Cloud phase~~	~~CALIPSO/CALIOP~~	~~1 km~~	~~-~~
	~~Cloud top height~~			
	~~Cloud optical thickness~~			

2.2 Method

~~2.2.1. The main framework~~

~~The research technical route for AGCPP production are shown in Fig. 2. First, we match the infrared brightness temperature of the GridSat-B1 and ERA5 meteorological field with the high-resolution cloud physical product of the MODIS satellite at the pixel level to construct a large-scale training sample IRBT2CPP. Then, we train the deep neural network cloud-based Attention-enhanced (At) UNet model (AtUNet) (Trebing et al., 2021) (CloudAtUNet) to learn the mapping relationship between brightness temperature and cloud physical parameters. Finally, we directly apply the model weights obtained from training to the entire GridSat-B1 brightness temperature time series to produce the AGCPP dataset, thereby achieving continuous cloud physical parameter generation on a global scale and over a long time series. This method combines the advantages of polar-orbiting and geostationary satellites in terms of spatial resolution and observational continuity, while effectively overcoming inconsistencies caused by calibration, navigation, and sensor differences in multi-source data fusion.~~ In our previous studies, the single-layer and double-layer cloud retrieval models within the DaYu CLoud Analysis System (DaYu-CLAS) have been demonstrated to perform well in all-day cloud physical properties retrievals. DaYu-CLAS includes single-layer cloud retrieval models such as Cloud-ResUNet (Zhao et al., 2023; Zhao et al., 2024b; Tong et al., 2023), Cloud-SmaAtUNet (Li et al., 2023), and the CloudDiff model (Xiao et al., 2025), as well as the double-layer cloud retrieval model OverlapCloudDiff (Li et al., 2025). Cloud-SmaAtUNet is an improved version of UNet, in which depthwise separable convolutions and convolutional block attention modules (CBAM) are integrated into both the encoder and decoder paths. Li et al. (2023) applied Cloud-SmaAtUNet to H8-AHI data and demonstrated that Cloud-SmaAtUNet achieves high accuracy and efficiency in cloud physical properties retrieval tasks. However, that study focused on a single sensor and produced cloud products for only one year (2017), which is insufficient to support studies of global cloud physical properties at high spatiotemporal resolution.

Therefore, in this study, the Cloud-SmaAtUNet model is applied to achieve global all-day cloud physical properties retrievals, as shown in Fig. 2. Due to the different institutional sources of the datasets, they may have different projection methods and spatiotemporal resolutions. To ensure the correct correspondence of pixels and data consistency, the data were first aligned to a unified 0.07° latitude and longitude grid before model construction. Temporal matching was conducted based on the respective temporal characteristics of the datasets. MODIS and ERA5 data, available at synoptic hours (0000, 0300, ..., 2100 UTC), were directly matched. In contrast, CALIOP data, were included by selecting profiles within a ±2 min window around each corresponding synoptic hour. Because the GridSat-B1 images are too large to be directly used as model inputs, each image is divided into multiple 64 × 64 pixel sub-images. After data preprocessing, Cloud-SmaAtUNet is trained using brightness temperature (BT) from the 6.7 and 11 μm channels and the SAZ as the primary inputs, with CLP, CTH, COT, and CER from the MODIS official cloud products as labels. Considering the influence of meteorological conditions on cloud formation and development, additional meteorological variables, such as temperature and humidity profiles, are incorporated as auxiliary inputs. In this way, a DaYu-GCP dataset with a temporal resolution of 3 h and a spatial resolution of 0.07° is retrieved. Finally, the products are validated using the MODIS and CALIOP official cloud products to evaluate their spatiotemporal continuity and day–night consistency.



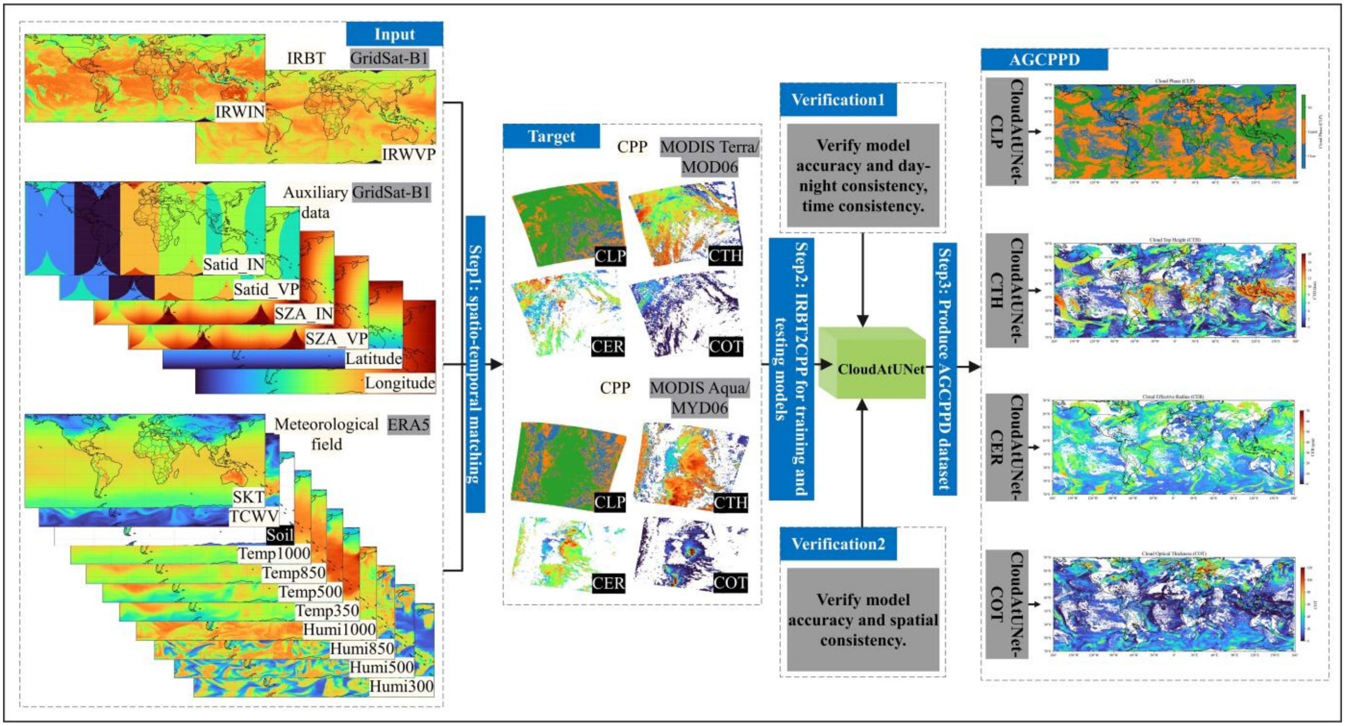


Figure 2.

Flowchart of DaYu-GCP production. It is worth noting that the input data consists of 19 channels, including GridSat-B1 and ERA5, while the target is CLP, CTH, COT, and CER images from MODIS official product processed to the same resolution. The training input and output image sizes are both 64×64. For each target, this study conducted model training and validation and evaluation separately. Finally, a sliding window fusion strategy was used to produce global-scale cloud product data.

~~Flowchart of AGCPP production. It is worth noting that the input data consists of 19 channels, including infrared brightness temperature (IRBT), auxiliary data, and meteorological fields, while the target is CLP, CTH, COT, and CER images processed to the same resolution. The training input and output image sizes are both 64×64. For each target, this study conducted model training and validation and evaluation separately. Finally, a sliding window fusion strategy was used to produce global-scale cloud product data.~~

During model parameter training, the spatiotemporally matched dataset comprised approximately 700,000 sample pairs. Data spanning 2000–2021 were used as the training set, while the entire dataset from 2022 constituted the testing set. These samples were evenly distributed across the spatial domain, and the strategy for selecting the training and testing sets satisfied the basic requirements of Cloud-SmaAtUNet. This selection strategy, together with the large overall sample size, jointly reduces the risk of model overfitting. To further expand the training data, data augmentation operations—including horizontal flipping, vertical flipping, and rotations of 90°, 180°, and 270°—were applied to the training set, increasing its size to six times the original. Meanwhile, the testing set remained unchanged, without any augmentation, in order to objectively evaluate the effectiveness of data augmentation in improving model performance. Indeed, the application of data augmentation led to improved accuracy. Detailed information on the dataset is provided in Table 3. Sample size and division of the matched dataset into training-set and testing-set.

Data matching pair	Total number of samples	Training-set	Testing-set	Training-set (data augmentation)	Testing-set (data augmentation)
GridSat-B1 & MOD06 Labels	373269	357744	15525	2146464	15525
GridSat-B1 & MYD06 Labels	324950	312090	12860	1872540	12860
Combined MOD06 & MYD06 Labels	698219	669834	28385	4019004	28385

.

~~IRBT2CPP is a corresponding matching dataset of infrared brightness temperature and cloud products containing information from 2000 to 2022, spanning 23 years, with a total of approximately 700,000 samples. This is a fairly large dataset. Considering the lengthy and extensive data processing involved in constructing IRBT2CPP, we have also chosen to make this dataset publicly available(https://doi.org/10.57760/sciencedb.27171) (Zhao, 2025).~~

~~When training model parameters, data from 2000 to 2021 was used as the training-set, and data from the entire year of 2022 was used as the testing-set. These samples were evenly distributed in spatial dimensions, and the strategy for selecting the training and testing sets met the basic requirements of AtUNet (Trebing et al., 2021). This strategy and the large total sample size reduce the risk of overfitting together. In order to expand the training data, we applied data augmentation operations such as horizontal flipping, vertical flipping, and 90°, 180°, and 270° rotation to the training set, expanding the training set to six times its original size. At the same time, the testing set remained unchanged without any augmentation in order to objectively verify the effectiveness of data augmentation in improving model performance. In fact, data augmentation operations did improve accuracy. Specific information about the data is shown in Table 3.~~

Table 3. Sample size and division of the matched dataset into training-set and testing-set.

Data matching pair	Total number of samples	Training-set	Testing-set	Training-set (data augmentation)	Testing-set (data augmentation)
GridSat-B1 & MOD06 Labels	373269	357744	15525	2146464	15525
GridSat-B1 & MYD06 Labels	324950	312090	12860	1872540	12860
Combined MOD06 & MYD06 Labels	698219	669834	28385	4019004	28385

~~**Table 3.** IRBT2CPP data amount statistics and training-set testing-set data amount division.~~

	~~Total number of~~	~~Training-~~	~~Testing-~~	~~Training-set (data~~	~~Testing-set (data~~

	~~samples~~	~~set~~	~~set~~	~~augmentation)~~	~~augmentation)~~
~~GridSat2MOD~~	~~373269~~	~~357744~~	~~15525~~	~~2146464~~	~~15525~~
~~GridSat2MYD~~	~~324950~~	~~312090~~	~~12860~~	~~1872540~~	~~12860~~
~~MOD+MYD~~	~~698219~~	~~669834~~	~~28385~~	~~4019004~~	~~28385~~

Key parameters for model training include: batch size = 512, maximum epochs = 300, and learning rate = 0.001. An early-stopping strategy was adopted, whereby training was terminated if the loss on the testing set did not decrease by more than 0.1 for 15 consecutive epochs. All models converged and stopped before reaching the maximum of 300 epochs. The loss functions used for model training varied by task. CrossEntropyLoss was employed for the CLP classification task, whereas MSELoss was applied to the CTH, COT, and CER regression tasks. To achieve seamless global coverage of the cloud products, an image sliding-window fusion strategy was implemented to eliminate gaps between adjacent small samples; details of this method are provided in Supplementary Material Text S1. For model evaluation, statistical metrics for classification performance included Accuracy, Recall, Precision, and F1-score, while regression performance was assessed using root mean squared error (RMSE), mean absolute error (MAE), mean bias error (MBE), and Pearson correlation coefficients (PearsonR).

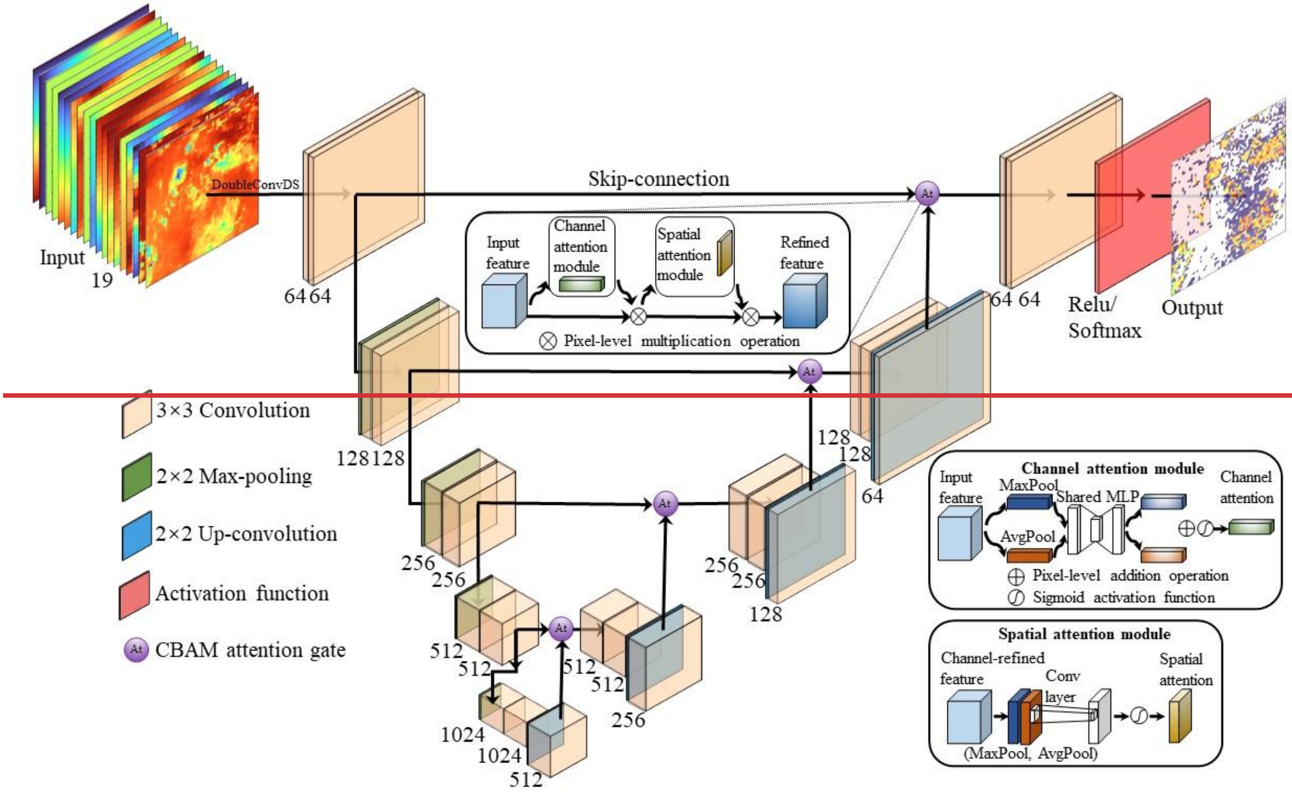
~~2.2.2. The machine learning model~~

~~Deep learning has achieved significant breakthroughs in the field of satellite remote sensing, particularly the UNet model, which is widely used in remote sensing image processing due to its exceptional spatial feature extraction capabilities (Liu et al., 2024a; Zhong et al., 2024). This study introduces an improved version of the UNet model called AtUNet (Trebing et al., 2021). This model incorporates the Convolutional Block Attention Module (CBAM) in the encoder section, reducing the number of parameters by 25% while maintaining the original UNet's accuracy, thereby significantly improving computational efficiency. Given that this model is specifically optimised for cloud physical retrieval, its input end integrates key physical prior parameters related to cloud formation (such as ERA5 temperature/humidity fields, etc.), we call it CloudAtUNet in this paper to highlight its embedded learning capability for cloud physics processes.~~

~~Figure 3 shows the complete structure of the CloudAtUNet model. Its encoder-decoder architecture synchronously captures the spectral response and spatial structural features of cloud systems through skip connections, making it particularly suitable for inverting cloud physical parameters with strong spatio-temporal inhomogeneity. The encoder learns the optimal nonlinear combination of multi-spectral brightness temperature and meteorological fields through convolutional layers, analysing sub-pixel-scale cloud physical attribute features while extracting the spatial distribution patterns of cloud systems. The attention mechanism CBAM module dynamically focuses on the core regions of cloud clusters while~~

~~suppressing irrelevant meteorological background noise. The upsampling convolutional layers in the decoder preserve the fine-grained structural features of cloud boundaries, preventing spatial information loss during decoding and enabling precise spatial reconstruction of cloud parameters.~~

~~The model was trained using a local high-performance computing cluster and NVIDIA GeForce RTX 3090 graphics cards, applying CloudAtUNet to terabyte-scale GridSat-B1 brightness temperature data and Aqua/Terra MODIS cloud physical products. Key parameters for model training include: batch size = 512, maximum epochs = 300, and learning rate = 0.001. An early stopping strategy was used, stopping when the loss on the testing set did not decrease by more than 0.1 for 15 consecutive epochs. All models stopped before reaching the maximum 300 epochs. The loss functions for model training differed. CrossEntropyLoss was selected for the CLP classification task, while MSELoss was selected for the CTH, COT, and CER regression tasks.~~



~~**Figure 3.** Image of the CloudAtUNet model structure, including the CBAM attention enhancement mechanism.~~

~~2.2.3. Image sliding window fusion strategy~~

~~Due to the large size of the images in the overlapping region between Gridsat-B1 and Aqua/Terra, they cannot be directly used as model inputs. Therefore, we divided each image into multiple small sample images with a matrix size of 64 × 64. In the actual production of AGCPP, each image takes approximately 2 minutes to process, so the 23-year product with 4 variables requires nearly 9,000 CPU hours (4 × 8 × 2 × 365 × 23 ÷ 60 = 8,955 h). Six RTX 3090 and four RTX A6000 GPUs were used to produce the dataset, taking nearly two months to complete.~~

~~The CloudAtUNet model was trained using a large number of image samples with a matrix size of 64×64 pixels. Therefore, when reconstructing the local prediction results of 64×64 pixels into a complete global cloud physical properties product image based on implicit prior knowledge (training weights), a certain stitching strategy needs to be adopted. To avoid gaps between adjacent prediction blocks, a sliding window fusion strategy based on linear weights is used.~~

~~First, the full image size: $H \times W$. PATCH: $P = 64$. STRIDE: $S = 10$. Since the image size may not be a multiple of PATCH, it is necessary to use rounding up to fill in $H_2 \times W_2$:~~

~~$$\boldsymbol{H_2 = \left\lceil \frac{H}{P} \right\rceil P, W_2 = \left\lceil \frac{W}{P} \right\rceil P} \qquad (1)$$~~

~~The starting coordinates and relative coordinates of the window are (i, j) and (x', y'), respectively.~~

~~$$(i, j), \quad 0 \leq i \leq H_2 - P, ; 0 \leq j \leq W_2 - P \qquad (2)$$~~

~~$$(x', y'), \quad 0 \leq x', y' \leq P - 1 \qquad (3)$$~~

~~The weight matrix is composed of piecewise linear functions in the row and column directions. When adjacent blocks exist at the window boundaries, the weights in the boundary regions transition linearly from 0 to 1 (left/top boundary) or from 1 to 0 (right/bottom boundary). By summing the weighted prediction values of the overlapping regions and normalizing them, a seamless stitched image is ultimately formed. The vertical weights $w_y^{(i)}(x')$ and horizontal weights $w_x^{(j)}(y')$ are respectively:~~

~~$$w_y^{(i)}(x') = \begin{cases} 0, & i = 0, \\ \frac{x'}{S-1}, & i > 0, 0 \leq x' < S, \\ 1, & S \leq x' < P - S, \\ \frac{P-1-x'}{S-1}, & i + P < H_2, P - S \leq x' < P, \\ 0, & \text{Otherwise.} \end{cases} \qquad (4)$$~~

~~$$w_x^{(j)}(y') = \begin{cases} 0, & j = 0, \\ \frac{y'}{S-1}, & j > 0, 0 \leq y' < S, \\ 1, & S \leq y' < P - S, \\ \frac{P-1-y'}{S-1}, & j + P < W_2, P - S \leq y' < P, \\ 0, & \text{Otherwise.} \end{cases} \qquad (5)$$~~

~~The final weight matrix W is constructed by the outer product of vertical weights and horizontal weights:~~

~~$$\mathrm{W}_{i,j}(x', y') = w_y^{(i)}(x') \cdot w_x^{(j)}(y') \qquad (6)$$~~

~~Where i and j are the starting position indices of the window in the filled large image, $x' = x - i$, $y' = y - j$, $x', y' \in [0, P - 1]$. The contributions of each window, $P_{i,j}$ (i.e., the predicted value at the relative coordinates (x', y') within window (i, j)), are weighted and normalized to obtain the final image:~~

~~$$I(x, y) = \frac{\sum_{i,j} W_{i,j}(x-i, y-j) \cdot P_{i,j}(x-i, y-j)}{\sum_{i,j} W_{i,j}(x-i, y-j)} \qquad (7)$$~~

3 Results and Discussion

3.1. Evaluation of model performance in testing-set

To enhance the physical interpretability of the model, the pixel-based (point-to-point) cloud retrieval machine learning model Random Forest (Cloud-RF) (Breiman, 2001) was employed as a comparative benchmark. Here, a grid-search approach was used to determine the Cloud-RF model parameters, with n_estimators = 200, max_depth = 50, min_samples_split = 3, and min_samples_leaf = 1. Using the 2022 test dataset, we compared the performance of the Cloud-SmaAtUNet and Cloud-RF models. Compared with the MODIS official products, Cloud-SmaAtUNet achieved an overall accuracy of 82.3% in the CLP classification task. The recognition accuracies of Cloud-SmaAtUNet for clear sky, water clouds, and ice clouds reached 81.52%, 80.21%, and 86.38%, respectively (Fig. 3(a)). In addition, Cloud-SmaAtUNet also exhibited relatively dense joint probability density distributions aligned along the diagonal in the regression tasks for CTH, COT, and CER. Although Cloud-SmaAtUNet showed a systematic underestimation (MBE < 0), the estimation errors for CTH, COT, and CER remained within acceptable ranges (Fig. 3 (c)(e)(g)). The RMSE (MAE) values were 1.617 km (0.954 km), 11.314 (6.871), and 7.181 μm (5.133 μm), respectively, with PearsonR of 0.926, 0.617, and 0.799. These results indicate that Cloud-SmaAtUNet exhibits excellent performance in the retrieval of cloud physical properties.

In contrast, the performance of the Cloud-RF model in cloud retrieval declined across all products. The overall accuracy of Cloud-RF in the CLP classification task decreased to 78.1% (Fig. 3(b)). Among the classes, the recognition accuracy for water clouds decreased most markedly (73.44%), while those for clear sky and ice clouds also showed slight reductions (78.93% and 85.19%, respectively). This indicates that Cloud-RF performs worse than Cloud-SmaAtUNet in CLP classification. Moreover, in the regression tasks for CTH, COT, and CER, Cloud-RF exhibited more pronounced underestimation than Cloud-SmaAtUNet (MBE < 0 with larger magnitude), and the joint probability density distributions aligned along the diagonal were less concentrated (Fig. 3 (d)(f)(h)). The RMSE (MAE) values were 2.369 km (1.483 km), 13.370 (8.935), and 8.860 μm (6.535 μm), respectively, with PearsonR of 0.843, 0.565, and 0.719. These results demonstrate that, compared with Cloud-SmaAtUNet, the Cloud-RF model shows inferior performance in retrieving cloud physical properties.

Compared with the conventional Cloud-RF, Cloud-SmaAtUNet can leverage the spatial structure information of clouds to improve cloud physical properties retrievals. The accuracy of CLP is increased by 5.4%, while the RMSE of CTH, COT, and CER are reduced by 31.7%, 18.2%, and 23.4%, respectively. Beyond the performance differences, Cloud-SmaAtUNet requires only about 100 s to retrieve a single global image, whereas Cloud-RF takes nearly six times longer, approximately 630 s. These findings are also reflected in other studies (Zhao et al., 2024a).

~~3 Evaluation of the AGCPP~~

~~We conducted a systematic evaluation of the generated AGCPP, which was divided into temporal consistency, spatial consistency evaluations and diurnal consistency. In the temporal consistency assessment, the model was first evaluated using the official MODIS products for the entire year of 2022 as a benchmark to assess CLP, CTH, COT, and CER. Additionally, we used Aqua/Terra MODIS data from 2000 to 2022 to assess the annual accuracy of CLP, CTH, COT, and CER. Spatial consistency assessment is also based on the official MODIS products for all years from 2000 to 2022, with errors statistically calculated by longitude and latitude.~~

~~Then, we performed a diurnal consistency analysis using the official CALIOP product for the year 2022, focusing on evaluating the accuracy of the CLP, COT, and CER during the day and night. Lastly, we again performed a yearly temporal consistency evaluation using CALIOP to evaluate the annual accuracy performance of CLP, COT, and CER in the AGCPP with the official CALIOP products from 2006 to 2022. Since AGCPP uses Aqua/Terra MODIS as the training target, and the official cloud product retrieval algorithms for MODIS and CALIOP are not identical, this may result in some systematic biases. Therefore, we also evaluated the annual accuracy performance of MODIS and CALIOP between the official cloud products CLP, COT, and CER from 2006 to 2022. To ensure the accuracy of the evaluation, the time difference between CALIOP and AGCPP was limited to ±2 minutes. The time difference between CALIOP and MODIS was also limited to ±2 minutes.~~

~~3.1. Time consistency evaluation with MODIS~~

~~In the evaluation, statistical error metrics for classification evaluation include Accuracy, Recall, Precision, and F1-score. Accuracy measures the proportion of correctly predicted samples out of the total number of samples, making it the most intuitive overall performance metric. Recall measures the proportion of actual positive examples correctly identified by the model out of all true positive examples. Precision measures the proportion of samples predicted as positive that are actually positive. F1-score is the harmonic mean of Precision and Recall, serving as a balanced comprehensive evaluation metric between Precision and Recall, better reflecting the model's robustness. For regression error metrics, RMSE, MAE, MBE, R^2, and PearsonR are used. RMSE imposes heavier penalties on larger errors, reflecting the overall dispersion of prediction errors. MAE assigns equal weight to all errors, providing an intuitive measure of error magnitude. MBE is the difference~~

~~between the model's calculated result and the true value, helping to diagnose whether the model has systematic bias. R^2 measures the proportion of variance in the target variable that the model can explain, reflecting the model's goodness of fit. PearsonR measures the strength and direction of the linear relationship between predicted and true values.~~

~~Based on thresholds established in earlier publications for comparing the FY4A (AGRI) official cloud product, the Himawari-8 (AHI) official cloud product, and the TL-ResUnet retrieved cloud product against MODIS, we applied the thresholds to evaluate our model results (Zhao et al., 2024). In detail: For the evaluation of CLP for the cloud classification task, the Accuracy is 71.77% and 79.82% for FY4A and TL-ResUnet, respectively. For the CTH evaluation of the cloud regression task, the RMSEs of FY4A and TL-ResUnet are 3.58 and 1.99 km. For the COT evaluation, the RMSEs of Himawari-8 and TL-ResUnet are 14.62 and 12.87. For the CER evaluation, the RMSEs of Himawari-8 and TL-ResUnet are 10.14 and 10.14 μm respectively. Moreover, although PearsonR ranges from −1 to 1 with higher values indicating better agreement, the quality limitations of the observational data mean that the threshold values for this metric differ among the three retrieved cloud physical properties: for TL-ResUnet they are 0.884, 0.596, and 0.765, respectively (Tong et al., 2023; Li et al., 2023; Zhao et al., 2024).~~

~~3.1.1. CloudAtUNet model performance testing~~

~~CloudAtUNet-CLP uses 0, 1, and 2 to represent clear skies, water clouds, and ice clouds, respectively. Table 4 shows the statistical results of the error evaluation indicators on the testing set (the whole year of 2022), and Fig. 4 shows the detailed data distribution under the testing set evaluation. The evaluation results show that the Accuracy, Recall, Precision, and F1-score of the CLP classification are 0.823, 0.827, 0.827, and 0.827, respectively. Compared with the threshold indicators, it is higher than FY4A and TL-ResUnet's 71.77% and 79.82%. The closeness of the four metrics indicates that the model achieves a good balance across the four metrics, demonstrating robust classification performance for both positive and negative category samples. Fig. 4 (a) shows the detailed data distribution of the CLP evaluation. It can be seen that the accuracy of clear sky forecasts is 81.52, water clouds are slightly lower at 80.21, and ice clouds are the highest at 86.38, which demonstrates that CloudAtUNet has good ability to analyse cloud physical properties in cloud classification tasks.~~

~~Additionally, the RMSE, MAE, MBE, R^2, and PearsonR values for the CTH regression task are 1.617, 0.954, −0.039, 0.857, and 0.926, respectively. For the threshold metrics, RMSE is lower than FY4A and TL-ResUnet's 3.58 and 1.99, and PearsonR is higher than 0.884. MBE = −0.039 indicates that although CloudAtUNet-CTH slightly underestimates CTH (<0), the underestimation is minimal, and the model does not exhibit significant systematic bias. Considering the 0–18 numerical range of CTH in the official MODIS cloud product, this performance is highly commendable. Fig. 4(b) shows the scatter density distribution diagram for CTH assessment. It can be seen that the majority of data are concentrated within 0–4 km, with a significant portion also clustered around the 1:1 line.~~

~~Secondly, the RMSE, MAE, MBE, R^2, and PearsonR values for the COT regression task are 11.314, 6.871, −0.649, 0.381, and 0.617, respectively. For threshold indicators, RMSE is lower than Himawari-8 and TL-ResUnet's 14.62 and 12.87,~~

~~and the PearsonR is well above 0.596. MBE = -0.649 indicates that while CloudAtUNet-COT underestimates COT, considering the 0–150 numerical range in the official MODIS cloud product, the degree of underestimation is acceptable and does not indicate severe systematic bias. Fig. 4(c) shows the scatter density distribution diagram for COT evaluation, with the majority of data concentrated between 0 and 20. For data above 20, the scatter density distribution diagram for CloudAtUNet-COT becomes more dispersed, which is the primary source of error.~~

~~Finally, the RMSE, MAE, MBE, R^2, and PearsonR values for the CER task are 7.181, 5.133, -0.132, 0.638, and 0.799, respectively. For threshold indicators, RMSE is lower than Himawari-8 and TL-ResUnet's 10.14 and 10.14 μm, and PearsonR is higher than 0.765. The MBE of -0.132 still indicates that CloudAtUNet-CER slightly underestimates CER, with the underestimation being greater than CTH but less than COT. This is primarily due to the 0–60 numerical range of CER in the official MODIS cloud product, so there is no significant systematic bias. Fig. 4(d) shows the scatter density distribution diagram for CER evaluation, indicating that CER values are primarily concentrated in the 7–20 μm and 30–45 μm ranges. The former primarily targets the retrieval of water clouds, while the latter primarily targets the retrieval of ice clouds (Liu et al., 2023). CloudAtUNet-CER demonstrates slightly stronger retrieval capability for water cloud CER than for ice cloud CER.~~

~~**Table 4.** Error indicators for the evaluation results of the testing set.~~

		~~Indicators~~			
~~Productions~~	~~Accuracy~~	~~Recall~~	~~Precision~~	~~F1-score~~	
~~CLP~~	~~0.823~~	~~0.827~~	~~0.827~~	~~0.827~~	
		~~Indicators~~			
	~~RMSE~~	~~MAE~~	~~MBE~~	~~R^2~~	~~PearsonR~~
~~CTH~~	~~1.617~~	~~0.954~~	~~-0.039~~	~~0.857~~	~~0.926~~
~~COT~~	~~11.314~~	~~6.871~~	~~-0.649~~	~~0.381~~	~~0.617~~
~~CER~~	~~7.181~~	~~5.133~~	~~-0.132~~	~~0.638~~	~~0.799~~

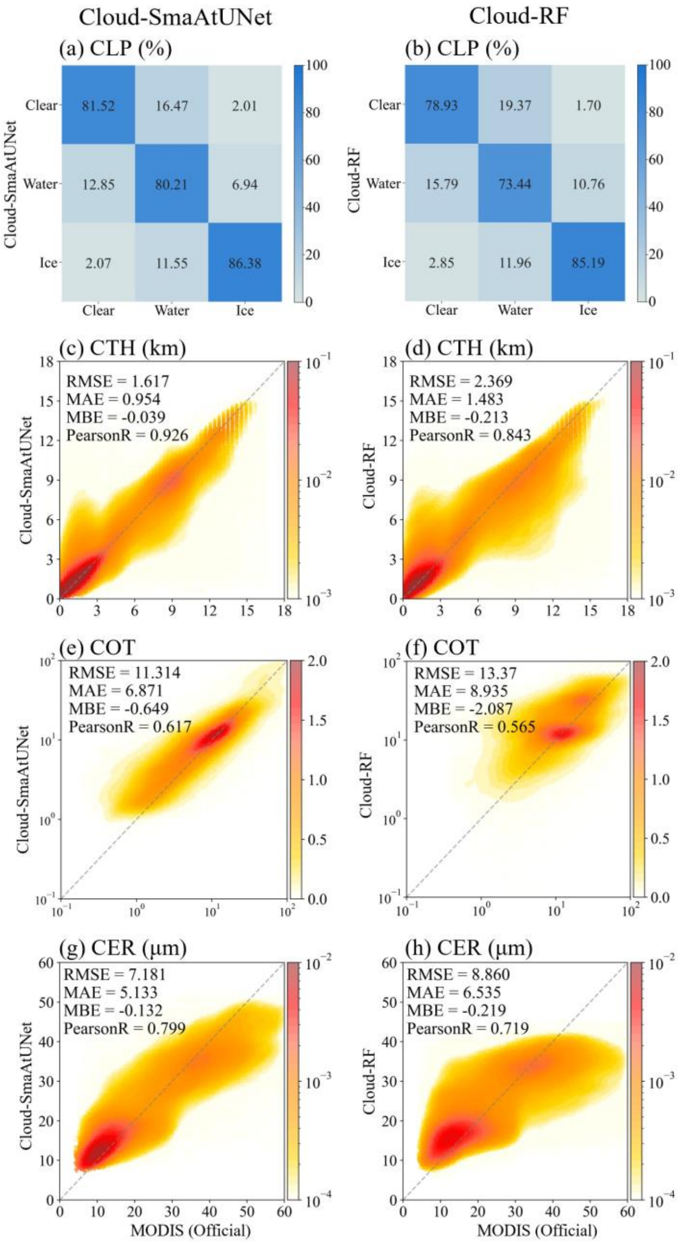
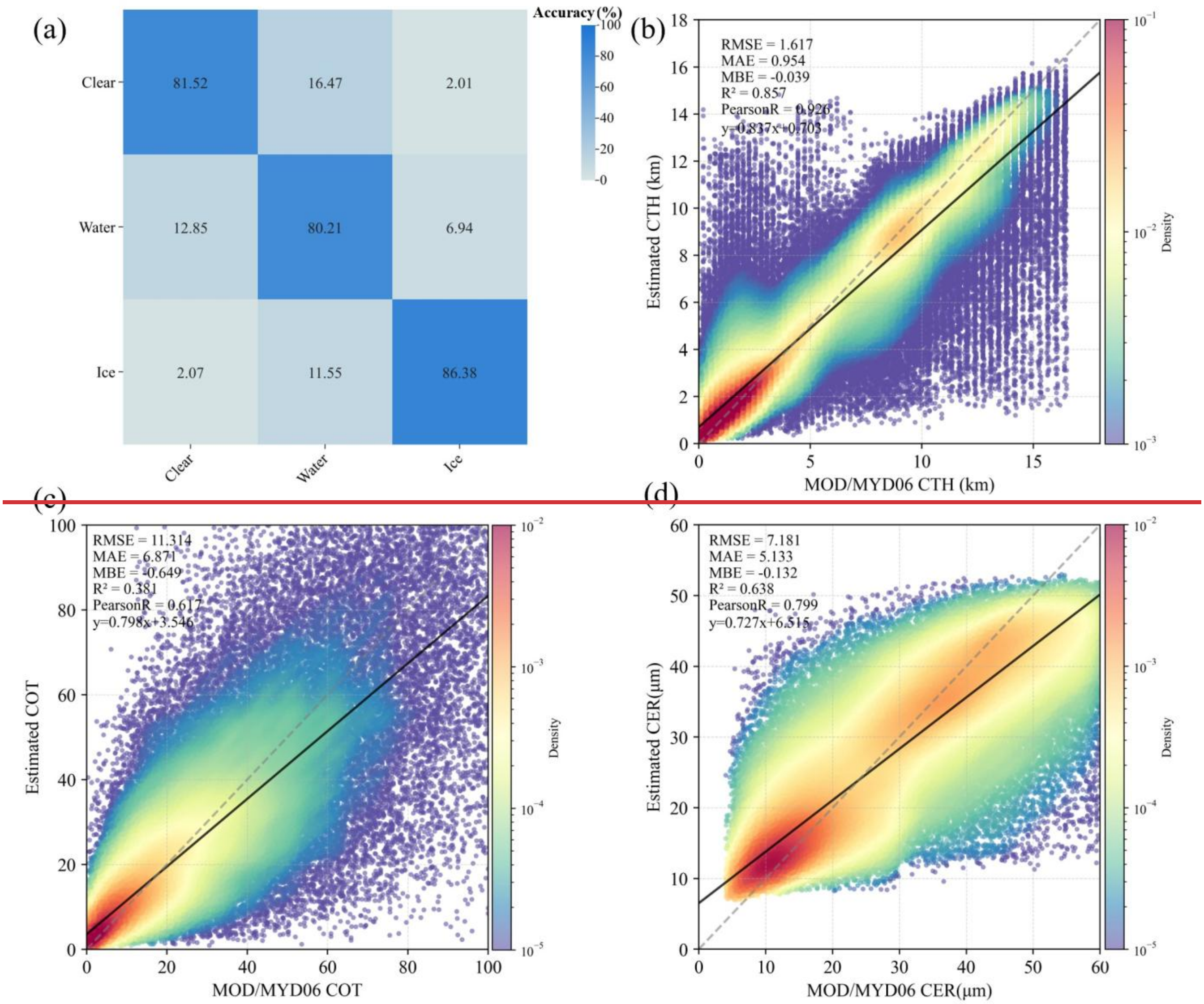


Figure 3. Detailed distribution of testing-set evaluations on Cloud-SmaAtUNet (left column) and Cloud-RF (right column). (a-b) Confusion matrix for CLP. Scatter density distribution diagram under the kernel density estimation of (c-d) CTH. (e-f) COT. (g-h) CER between MODIS official and DaYu-CLAS Cloud-SmaAtUNet or DaYu-CLAS Cloud-RF products.



~~**Figure 4.** Detailed distribution of testing set evaluations. (a) Confusion matrix for CLP. Scatter density distribution diagram under the kernel density estimation of (b) CTH. (c) COT. (d) CER.~~

3.2. Physical interpretability of the Cloud-SmaAtUNet model

To elucidate the physical interpretability of the model, we investigated the roles of BT and meteorological fields in cloud retrieval. This experiment involved feature ablation testing, comparing model performance across different input combinations (BT only, ERA5 only, BT and ERA5) to analyze each data source's contribution to retrieval accuracy. Specifically, for both the Cloud-SmaAtUNet and Cloud-RF models, three distinct input configurations were employed:

Group 1 incorporated 11 μm and 6.7 μm BT along with the full set of ERA5 meteorological fields; Group 2 comprised solely 11 μm and 6.7 μm BT data; Group 3 comprised solely ERA5 meteorological field data. These experiments utilized both Cloud-SmaAtUNet and Cloud-RF models for training and evaluation, with results presented in Table S1 (see Supplementary Materials) and Fig. 4.

The CLP classification results in Table S2 indicate that when using only ERA5, the Cloud-SmaAtUNet classification Accuracy for clear sky, water cloud, and ice cloud were 66.06%, 63.64%, and 68.16%, respectively; when using only BT, the Accuracy was 78.86%, 79.43%, and 83.96%. Using BT and ERA5 together, the classification Accuracy for clear sky, water cloud, and ice cloud improved to 81.52%, 80.21%, and 86.38%, representing increases of 23.4%, 26.0%, and 26.7% compared to using ERA5 alone, and increases of 3.4%, 1.0%, and 2.9% compared to using BT alone.

The results of the regression tasks are shown in Fig. 4. When using ERA5 alone, the RMSE of CTH, COT, and CER retrieved by Cloud-SmaAtUNet across different ranges were 2.1–7.2 km, 12.9–82.9, and 7.0–25.3 μm, respectively. When using only BT, the RMSE was 1.5–6.0 km, 9.2–63.7, and 4.8–24.1 μm. When using BT and ERA5 together, the RMSE of CTH, COT, and CER decreased to 1.1–5.1 km, 6.0–45.9, and 2.8–21.4 μm. These results indicate that using BT and ERA5 together, compared to using only BT (or only ERA5), reduces the RMSE of CTH, COT, and CER on average by 21.0% (63.5%), 39.5% (90.3%), and 42.6% (85.6%), respectively.

The above experimental results indicate that satellite-observed BT represents an integrated signature of clouds, the atmosphere, and surface types. In the cloud retrieval process based on BT, incorporating atmospheric background fields and surface information contributes to the accurate retrieval of cloud physical properties.

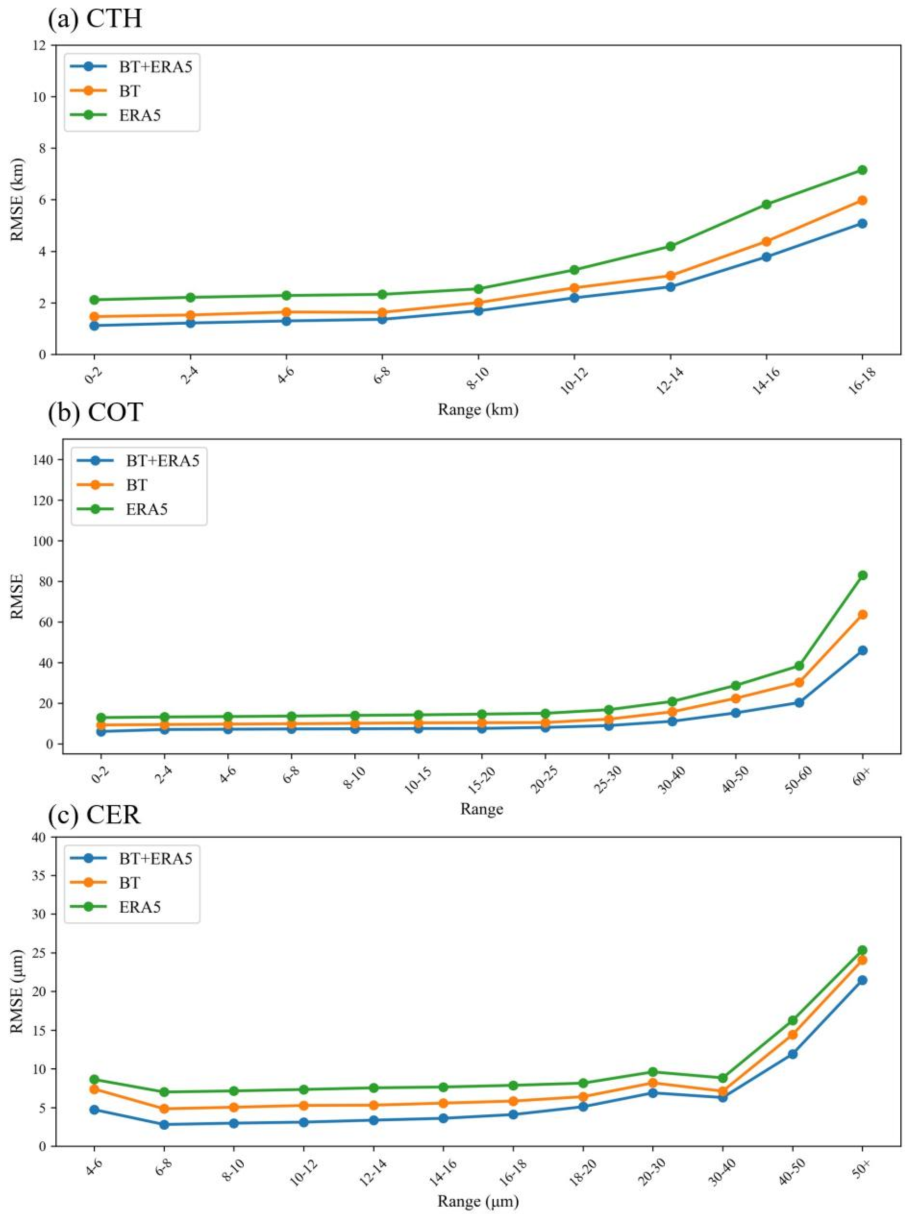


Figure 4. Cloud-SmaAtUNet (left column) and Cloud-RF (right column) were evaluated for RMSE across three distinct input datasets, with results presented for different value ranges. Group 1 inputs comprise 11 μm and 6.7 μm BT alongside ERA5. Group 2 inputs comprise 11 μm and 6.7 μm BT. Group 3 inputs comprise ERA5. (a) CTH. (b) COT. (c) CER.

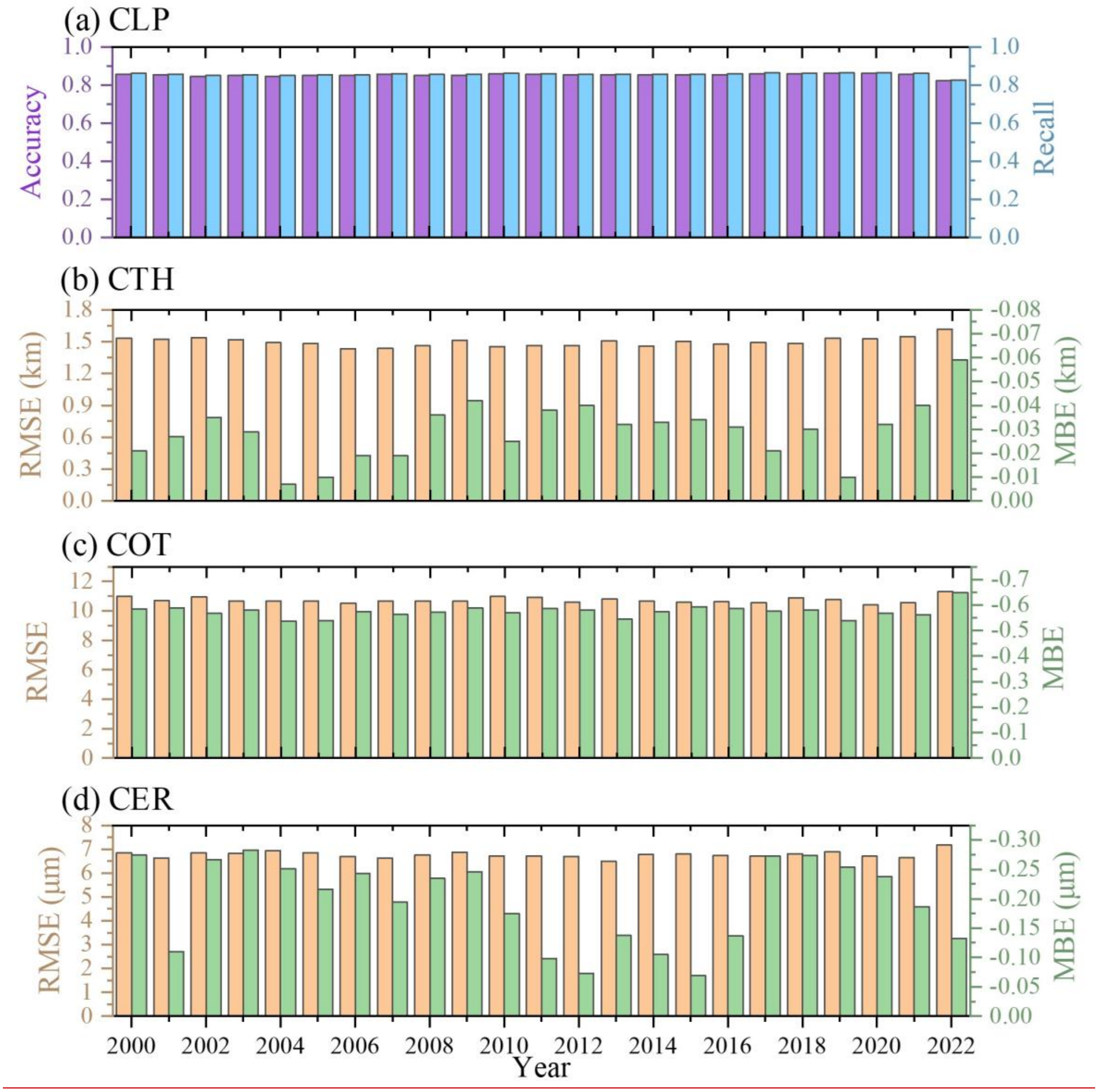
3.1.2. Evaluation of DaYu-GCP with MODIS official products

~~CloudAtUNet annual evaluation~~

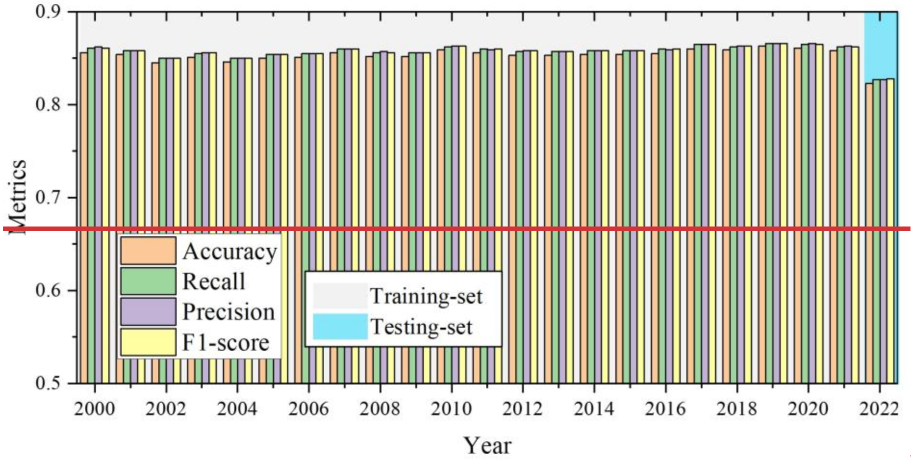
To assess the interannual stability of the DaYu-GCP products, we validated them using the Aqua/Terra MODIS official cloud products over the period 2000–2022, as shown in Fig. 5. In terms of CLP identification, the Accuracy and Recall for all years are approximately 0.85 ± 0.007. Although there are slight fluctuations in accuracy across years—for example, slightly lower in 2004 and slightly higher in 2010—the differences are minor. This demonstrates that the Cloud-SmaAtUNet model possesses robust cloud detection capability. For the retrievals of CTH, COT, and CER, the annual RMSE values remain stable at 1.50 ± 0.03 km, 10.71 ± 0.15, and 6.75 ± 0.10 μm, respectively, with coefficients of variation (standard deviation/mean) of 2.0%, 1.4%, and 1.5%. These results indicate that DaYu-GCP exhibits stable accuracy across years. The MBE for CTH, COT, and CER across all years range from -0.007 to -0.059 km, -0.536 to -0.649, and -0.069 to -0.283, respectively. Although the MBE deviations are very small, the dataset still exhibits a slight systematic negative bias.

~~This section uses Aqua/Terra MODIS data from 2000 to 2022 to evaluate the annual accuracy of CLP, CTH, COT, and CER. We expect the results of the CloudAtUNet model to remain consistent in annual assessments, but in reality, the accuracy metrics for each year cannot be completely consistent. This section evaluates accuracy using annual MODIS official cloud product data, providing users with some reference for using AGCPP, especially when using data from a specific year. Fig. 5 shows the evaluation results for CLP, and Fig. 6 shows the evaluation results for CTH, COT, and CER.~~

~~It can be seen that the accuracy index of all years in the training-set in Fig. 5 is approximately 0.85. Although there are slight fluctuations in accuracy from year to year, with slight differences, such as 2004 being slightly lower and 2010 being slightly higher. When users need to use AGCPP, they can use this chart to roughly determine whether the accuracy of the data usage year is slightly higher or slightly lower than the surrounding years. Additionally, the fluctuations in all data in the training-set are almost within three decimal places, so it can be said that the CloudAtUNet-CLP model has good learning and cloud classification capabilities, and the results of the annual tests are relatively continuous and consistent, with no extreme anomalies. The results on the testing-set are slightly lower than those on the training set, which is normal because the model has not seen the samples in the testing-set. A detailed analysis has been presented in 3.1.1 and will not be described here.~~

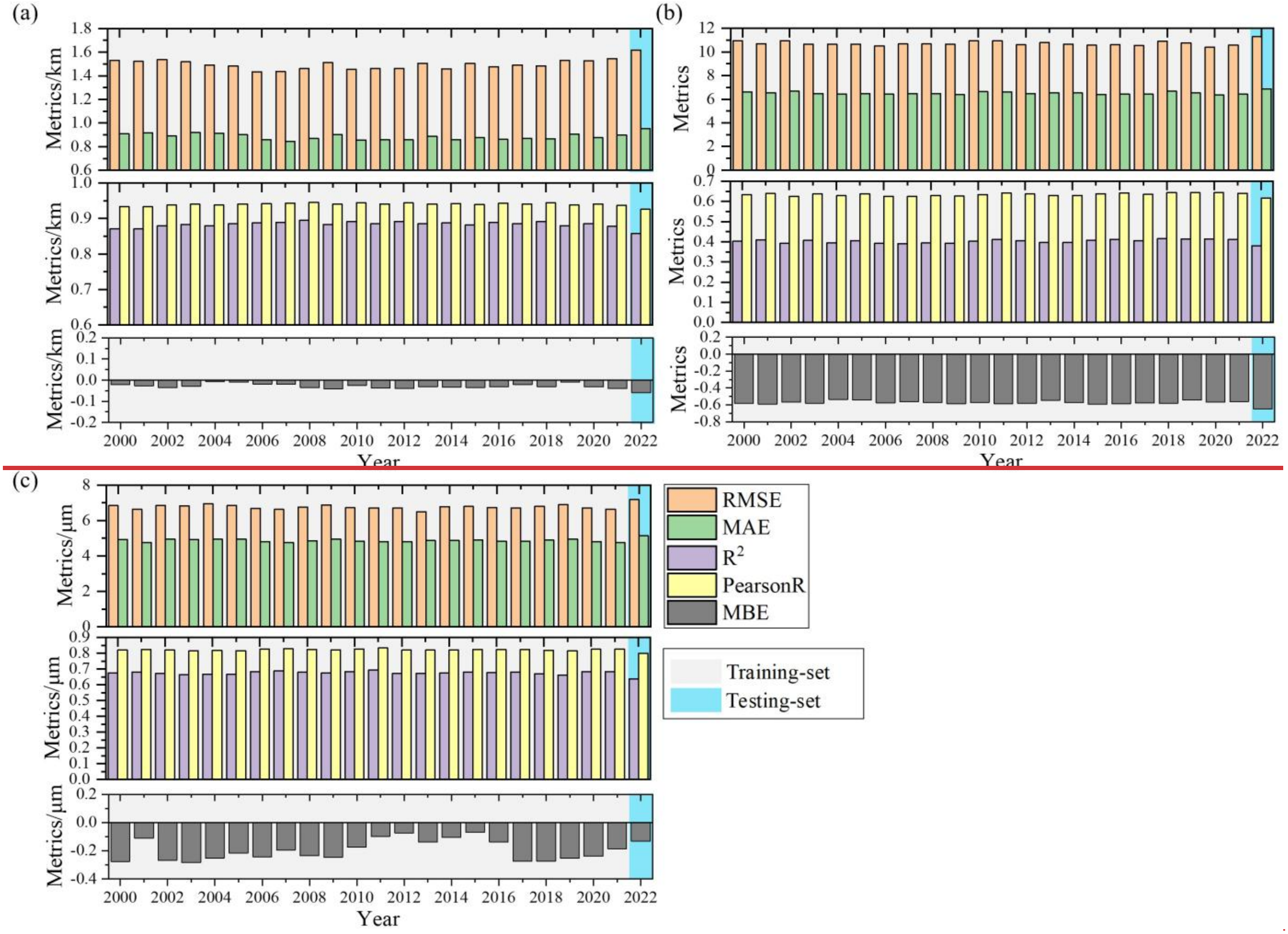


655 **Figure 5.** Evaluation of (a) CLP classification accuracy and Recall (b) CTH (b) COT (c) CER regression RMSE and MBE results based on training-set and testing-set of Aqua and Terra MODIS official cloud products.



~~**Figure 5.** Evaluation of CLP classification accuracy on training-set and testing-set based on annual Aqua and Terra MODIS official cloud products.~~

~~Figure 6 shows the statistical results of the regression metrics for CTH, COT, and CER. Since RMSE and MAE indicate higher model accuracy when their values are lower, they are plotted on the same graph. R^2 and PearsonR also indicate higher model accuracy when their values are higher, so they are also plotted on the same graph. Finally, MBE is plotted on a separate graph. In Fig. 6 (a), the RMSE of the CTH training-set is around 1.5, while the testing-set is slightly higher at around 1.6. The same is true for the RMSE in Fig. 6 (b) and Fig. 6 (c). Through the MBE indicator (model result minus true value), we also found that whether it is the training-set or the testing-set, the model tends to give smaller values in regression prediction, slightly underestimating. Finally, although MSE was used as the loss function during training, R^2 and PearsonR were more stable than RMSE and MAE in annual statistics. For example, the R^2 of all years in the CTH training-set was almost within the range of 0.88±0.01, while the annual RMSE and MAE varied to a certain extent. However, all statistical results and data prove that the CloudAtUNet annual detection results are stable.~~



~~**Figure 6.** Evaluation of (a) CTH. (b) COT. (c) CER. regression results based on training-set and testing-set of official Aqua and Terra MODIS cloud products.~~

~~3.2. Spatial consistency evaluation with MODIS~~

~~This section presents the spatial distribution of spatial differences between AGCPP and MODIS for all years from 2000 to 2022. Fig. 7 shows the spatial distribution of spatial differences between AGCPP and MODIS along latitude, with statistical results averaged every 3°. Fig. 8 shows the spatial distribution of differences between AGCPP and MODIS along longitude, with statistical results averaged every 5°. Fig. 7 and Fig. 8 present the mean values (blue dots), one standard deviation (STD) above and below the mean (blue shading), and box plots (upper bound, lower bound, upper quartile, lower quartile, and median). Firstly, the latitudinal distribution of CLP is less different, but the classification effect is slightly lower at low latitudes than at high latitudes in Fig. 7(a). This is mainly due to the complex cloud structure in the tropics (e.g.,~~

~~cumulonimbus, high convective clouds), which makes it easier to confuse water clouds with ice clouds. Although deep learning models well trained, may still misclassify when encountering inhomogeneous cloud phases in the tropics. This phenomenon is mentioned in the study of Meyer et al. (2016), which points out that especially in tropical strong convective clouds with strong ice-water mixing, leading to more significant classification errors.~~

~~Secondly, the spatial error distributions of CTH and CER in Fig. 7(b) and Fig. 7(d) are relatively consistent, with higher values near the equator than in other regions. This is primarily because the CTH and CER values near the equator are inherently larger, resulting in correspondingly larger error margins during model evaluation. Mitra et al. (2021) compared MODIS CTH with Lidar observations and found that for CTH, errors are smaller for low clouds and larger for high clouds. Since cloud heights near the equator are inherently higher, the higher cloud heights result in larger errors in CTH retrieval. Zhang et al. (2025) also found that in the equatorial and low-latitude regions, due to high ice/liquid cloud mixing and complex infrared absorption patterns, CER values are higher, resulting in larger errors during model evaluation.~~

~~Finally, as shown in Fig. 7(c), the spatial error distribution of COT is smaller near the equator and increases gradually with increasing latitude. This is primarily because ice clouds are more abundant in mid-to-high latitude regions, and multiple studies have shown that ice clouds typically have higher optical thickness than liquid water clouds (Takahashi et al., 2016). Additionally, Alexandrov et al. (2025) also pointed out that if the 3D radiative effects are ignored in cloud retrieval calculations, the COT of ice clouds at high latitudes would be significantly underestimated. Therefore, in these regions, both the MODIS official algorithm and our deep learning model face greater challenges in estimating COT.~~

We also analyzed the spatiotemporal distribution of spatial discrepancies between DaYu-GCP and MODIS official cloud products from 2000 to 2022. Fig. 6 shows the latitudinal distribution of differences between DaYu-GCP and MODIS official cloud products, with statistical results averaged every 3°. Fig. 7 shows the longitudinal distribution of differences between DaYu-GCP and MODIS official products, with statistical results averaged every 5°. In both Fig. 6 and Fig. 7, the mean values (blue dots), one standard deviation (STD) above and below the mean (blue shading), and box plots (upper bound, lower bound, upper quartile, lower quartile, and median) are presented.

Firstly, the latitudinal distribution of CLP exhibits minimal variation, with overall accuracy reaching approximately 0.85. Accuracy increases slightly with increasing latitude. Additionally, the spatial error distributions of CTH and CER in Fig. 6(b) and Fig. 6(d) are relatively consistent, showing higher values near the equator than in other regions. This is primarily because CTH and CER values near the equator are inherently larger, leading to correspondingly larger error margins during model evaluation (Fig. 11(l) and Fig. 11(o)). Mitra et al. (2021) compared MODIS CTH with Lidar observations and found that CTH errors are smaller for low clouds and larger for high clouds. Since cloud heights near the equator are inherently higher, this results in larger errors in CTH retrieval. Zhang et al. (2025b) also found that in equatorial and low-latitude regions, the high mixing of ice and water clouds and complex infrared absorption patterns lead to higher CER values, which in turn result in larger errors during model evaluation. Finally, as shown in Fig. 6(c), COT errors are smaller near the equator and gradually increase with latitude.

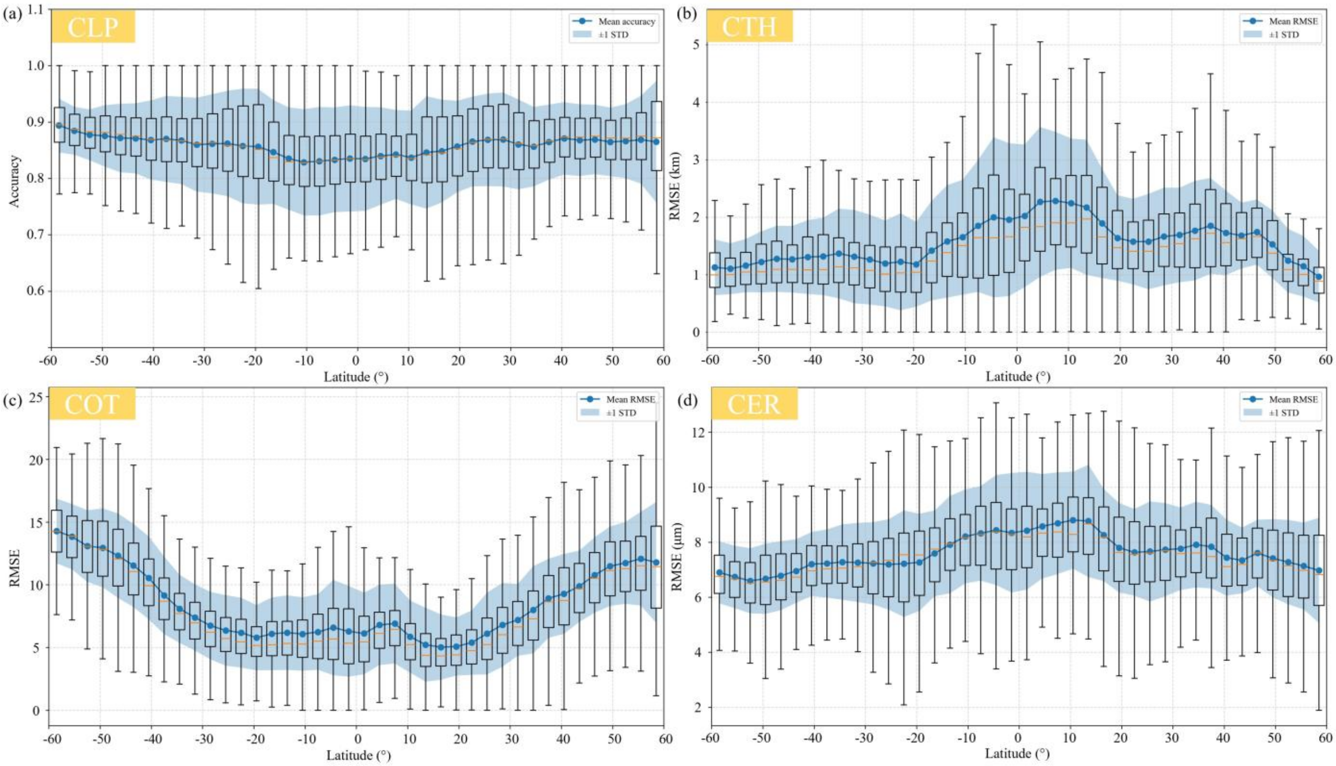


Figure 6. Variation in 23-year average Accuracy and RMSE across different latitudes. (a) CLP. (b) CTH. (c) COT. (d) CER.~~Figure 7. Distribution of spatial differences in latitude between AGCPP and MODIS. (a) CLP accuracy. (b) CTH RMSE. (c) COT RMSE. (d) CER RMSE.~~

Figure 7. Variation in 23-year average Accuracy and RMSE across different longitudes. (a) CLP. (b) CTH. (c) COT. (d) CER.

shows that for all products, the indicators exhibit clear peak–trough cycles along the longitudinal direction, with the mean Accuracy of CLP fluctuating between 0.82 and 0.88. The mean RMSE of CTH, COT, and CER vary between 1.3–2.1 km, 7–13, and 6.7–8.2 μm, respectively. This may be because regions with larger fluctuations correspond to fewer samples in the MODIS products, resulting in less stable evaluation results.

~~Figure 8 shows that the STD of the longitude-direction metrics for all products exhibits a distinct peak-trough cycle, with eight peaks occurring at −135°, −90°, −45°, 0°, 45°, 90°, 135°, and 180°, and the midpoints between the peaks representing the trough values. This pattern is due to the fact that the AGCPP is derived from the GridSat-B1 product, which is constructed by merging the observed brightness temperatures from multiple geostationary satellites. At the meridian point (subsatellite point) of each satellite, the observational geometry and calibration are most consistent, resulting in the lowest errors and STDs. As the zenith angle of the same satellite increases, the STD also gradually increases. However, once the observation area crosses into the coverage zone of an adjacent satellite, discontinuities in radiometric calibration and~~

~~geometric alignment occur, introducing additional errors and resulting in STD peaks at these longitude positions. Gunshor et al. (2009) demonstrated that inter-calibration errors between satellites primarily stem from differences in temporal alignment, spatial resolution, and geometric alignment, with these factors being most pronounced at the mosaic boundaries. Although GridSat-B1 has undergone calibration and navigation operations, it can only minimise differences in radiometric calibration and geometric alignment rather than completely eliminate them, resulting in this phenomenon.~~

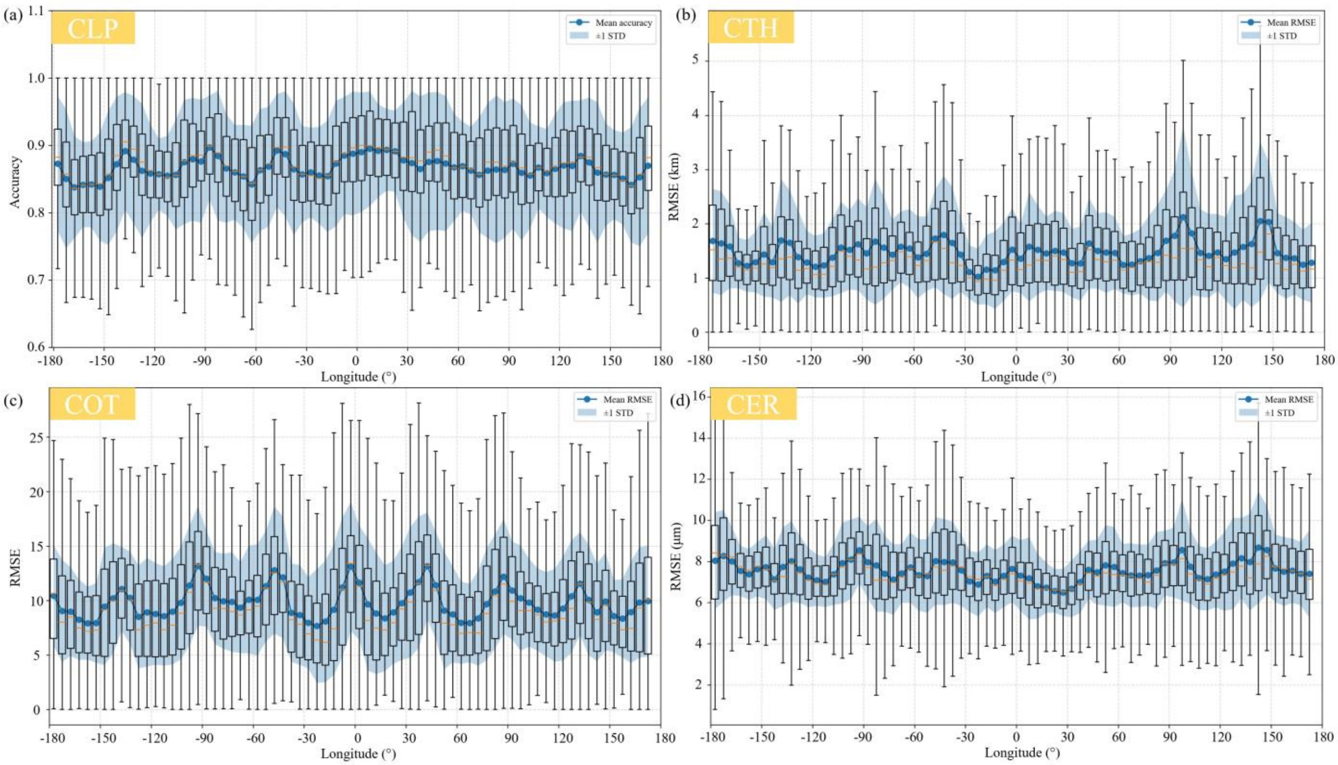


Figure 7. Variation in 23-year average Accuracy and RMSE across different longitudes. (a) CLP. (b) CTH. (c) COT. (d) CER.

~~**Figure 8.** Distribution of spatial differences in longitude between AGCPP and MODIS. (a) CLP accuracy. (b) CTH RMSE. (c) COT RMSE. (d) CER RMSE.~~

3.4. Evaluation of DaYu-GCP with CALIOP official products

In this study, we evaluated the consistency of daytime and nighttime performance of the DaYu-GCP results using CALIPSO/CALIOP data in 2022. Compared with the active sensors CALIPSO/CALIOP, the Accuracy of CLP during the day and night are 0.787 and 0.775, respectively (Fig. 8(a)). The daytime and nighttime RMSE of CTH are 3.384 km and 3.568 km, respectively (Fig. 8(b)). For the evaluation of COT, given that CALIOP cannot penetrate thick cloud layers, this assessment considered only thin cloud layers labeled as transparent in the CALIOP data as ground-truth references.

Correspondingly, within DaYu-GCP, only thin cloud regions with a thickness less than 5 were compared against CALIOP measurements. Finally, the daytime and nighttime RMSE of the COT product are 2.836 and 3.002, respectively (Fig. 8(c)). Overall, the model results demonstrate good consistency between day and night.

~~3.3. Day-night and time consistency evaluation with CALIPSO/CALIOP~~

~~3.3.1. Evaluation of day-night consistency~~

~~In this study, we evaluated the consistency of daytime and nighttime accuracy of the CloudAtUNet model results using CALIPSO/CALIOP. Additionally, the comparison results between the FY4A (AGRI) official cloud product, the Himawari-8 (AHI) official cloud product, and the CALIOP official cloud product are provided in the literature (Li et al., 2023; Zhao et al., 2024). Specifically, on CLP, the accuracy value of AHI compared to CALIOP is 0.736 (Zhao et al., 2024); On CTH, the RMSE value of AGRI and CALIOP results is 4 (Zhao et al., 2024); On COT, the RMSE value of AHI and CALIOP results is 23.71 (Li et al., 2023).~~

~~Figure 9(a) shows that compared with the active sensors CALIPSO/CALIOP, the accuracy of the CloudAtUNet model CLP during the day and at night is 0.787 and 0.775, respectively. Compared with the official cloud product 0.736 of Himawari-8 (AHI), the accuracy of the CLP product output by the CloudAtUNet model is higher, and there is little difference in accuracy between day and night, indicating good consistency between day and night. Secondly, Fig. 9(b) shows that the daytime and nighttime accuracies of the CloudAtUNet model CTH are 3.384 and 3.568, respectively. Compared to the official cloud product FY4A (AGRI) with an accuracy of 4, the CloudAtUNet model's CTH product has higher accuracy, and there is little difference in accuracy between daytime and nighttime, demonstrating good consistency between day and night. Finally, Fig. 9(c) shows that the daytime and nighttime accuracies of the CloudAtUNet model's COT product are 16.87 and 17.936, respectively. Similarly, using the Himawari-8 (AHI) official cloud product's accuracy of 23.71 as the standard, the CloudAtUNet model's COT product accuracy is higher. Overall, due to the use of infrared brightness temperature and the day-night unified CloudAtUNet model, the model results exhibit good day-night consistency.~~

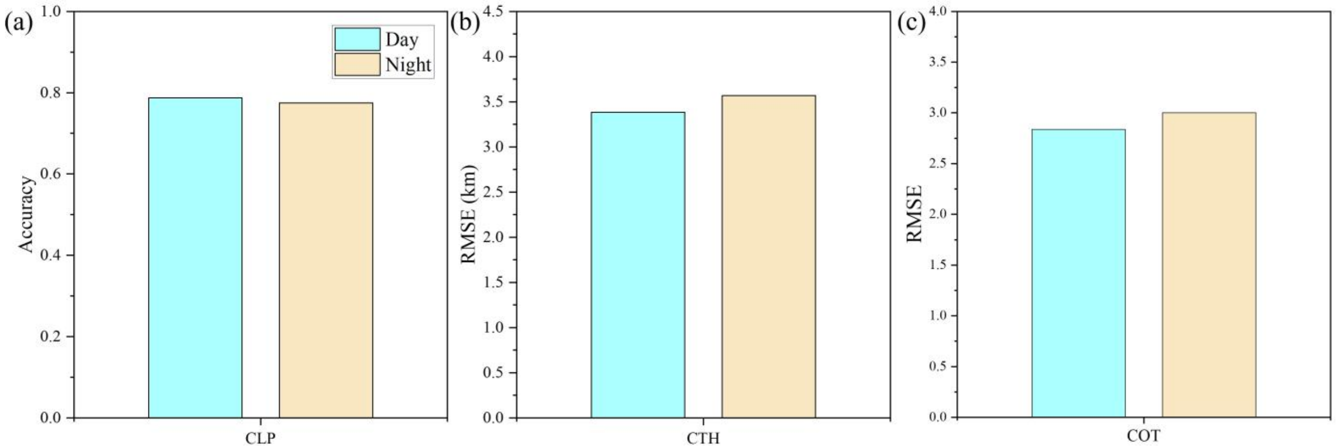
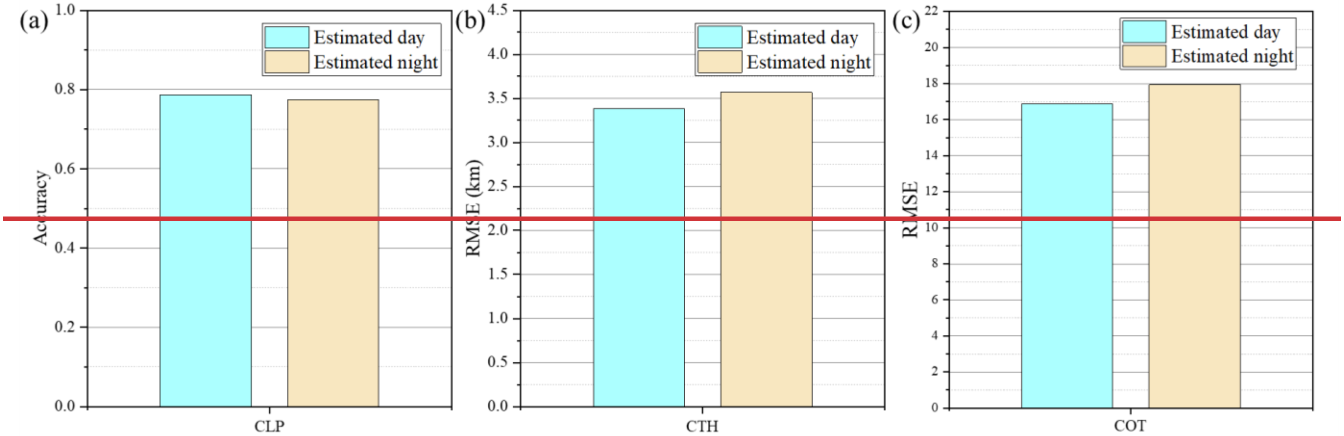


Figure 8. The accuracy of the DaYu-GCP (daytime and nighttime) was evaluated using data from the active sensors CALIPSO/CALIOP. (a) CLP accuracy. (b) The RMSE of CTH. (c) The RMSE of COT. The blue bar chart shows the model's daytime accuracy, and the yellow bar chart shows the model's nighttime accuracy.



~~**Figure 9.** The accuracy of the CloudAtUNet model (daytime and nighttime) was evaluated using data from the active sensors CALIPSO/CALIOP. (a) CLP accuracy. (b) The RMSE of CTH. (c) The RMSE of COT. The blue bar chart shows the model's daytime accuracy, and the yellow bar chart shows the model's nighttime accuracy.~~

~~3.3.2. Annual accuracy evaluation between AGCPP and CALIPSO/CALIOP~~

~~After producing the AGCPP dataset for 2000-2022, we matched the AGCPP with CALIPSO/CALIOP data for the overlapping time period and conducted an evaluation. To ensure the accuracy of the evaluation, the time difference between CALIOP and AGCPP was limited to ±2 minutes. Additionally, since AGCPP is labelled using Aqua/Terra MODIS as the training target, and the official cloud product retrieval algorithms for MODIS and CALIOP are not identical, this may~~

~~introduce some systematic biases. Therefore, we also conducted an evaluation for comparison, with the time difference between CALIOP and MODIS restricted to ±2 minutes.~~

~~Figure 10 shows the comparison results of AGCPP-CALIOP (G-C) and MODIS-CALIOP (M-C) for the three variables CLP, CTH, and COT. It can be seen that the accuracy of G-C is generally slightly lower than that of M-C in Fig. 10(a). The classification accuracy of M-C is approximately 0.8, while that of G-C is around 0.75. This is similar to the evaluation results in Wang et al. (2016) (the consistency between MODIS and CALIOP was 77.8%). In years where M-C has higher accuracy, G-C also has higher accuracy, as the labels used for AGCPP training are from MODIS rather than CALIOP. Therefore, some of the errors in G-C's evaluation are due to errors between M-C. Fig. 11 shows the confusion matrix information of the detection results. It can be seen that the M0-C0 values for 2013 and 2015 are 68.20 and 65.64, respectively, which are slightly lower than those of other years. Similarly, the G0-C0 values for 2013 and 2015 are 64.79 and 64.65, respectively, which are also slightly lower than those of other years. This further indicates that the results of the G-C evaluation are to some extent dependent on the gap between M and C. After being trained using MODIS as the label, AGCPP effectively captures and fits the MODIS information, with consistent and continuous accuracy year by year.~~

~~Additionally, the accuracy of the regression tasks CTH and COT is compared using RMSE, as shown in Fig. 10(b) and Fig. 10(c), respectively. Similarly, the regression accuracy of G-C is slightly lower than that of M-C. Since a smaller RMSE value indicates higher model accuracy, the value for G-C is slightly larger. However, the annual accuracy is relatively consistent and continuous for both CTH and COT. References also note that the errors in CTH for MODIS and CALIOP are significant, for example, the bias for pixels larger than 2 km exceeds 3 km (Weisz et al., 2007), which aligns with our statistical results.~~

presents the validation of DaYu-GCP and MODIS official products against the CALIOP official products for three variables: CLP, CTH, and COT. For CLP, the overall Accuracy of DaYu-GCP shows relatively small variation, with a mean value of approximately 0.76. The Accuracy for clear sky, water clouds, and ice clouds are about 0.74, 0.76, and 0.77, respectively, which are slightly lower than those of MODIS, whose overall Accuracy is around 0.80, with clear sky, water cloud, and ice cloud Accuracy of approximately 0.76, 0.79, and 0.81, respectively. For CTH, the multi-year RMSE of DaYu-GCP ranges from 3.35 to 3.53 km, slightly higher than that of MODIS (3.09–3.32 km). For COT, the multi-year RMSE of DaYu-GCP ranges from 3.59 to 4.08, also slightly higher than MODIS (2.57–3.18). Using the CALIOP official products as the reference, the annual RMSE for CTH and COT remain stable at 3.42 ± 0.15 km and 3.84 ± 0.16, with coefficients of variation of 4.4% and 4.1%, respectively. Overall, the DaYu-GCP products demonstrate stability in both day–night accuracy and interannual variations.

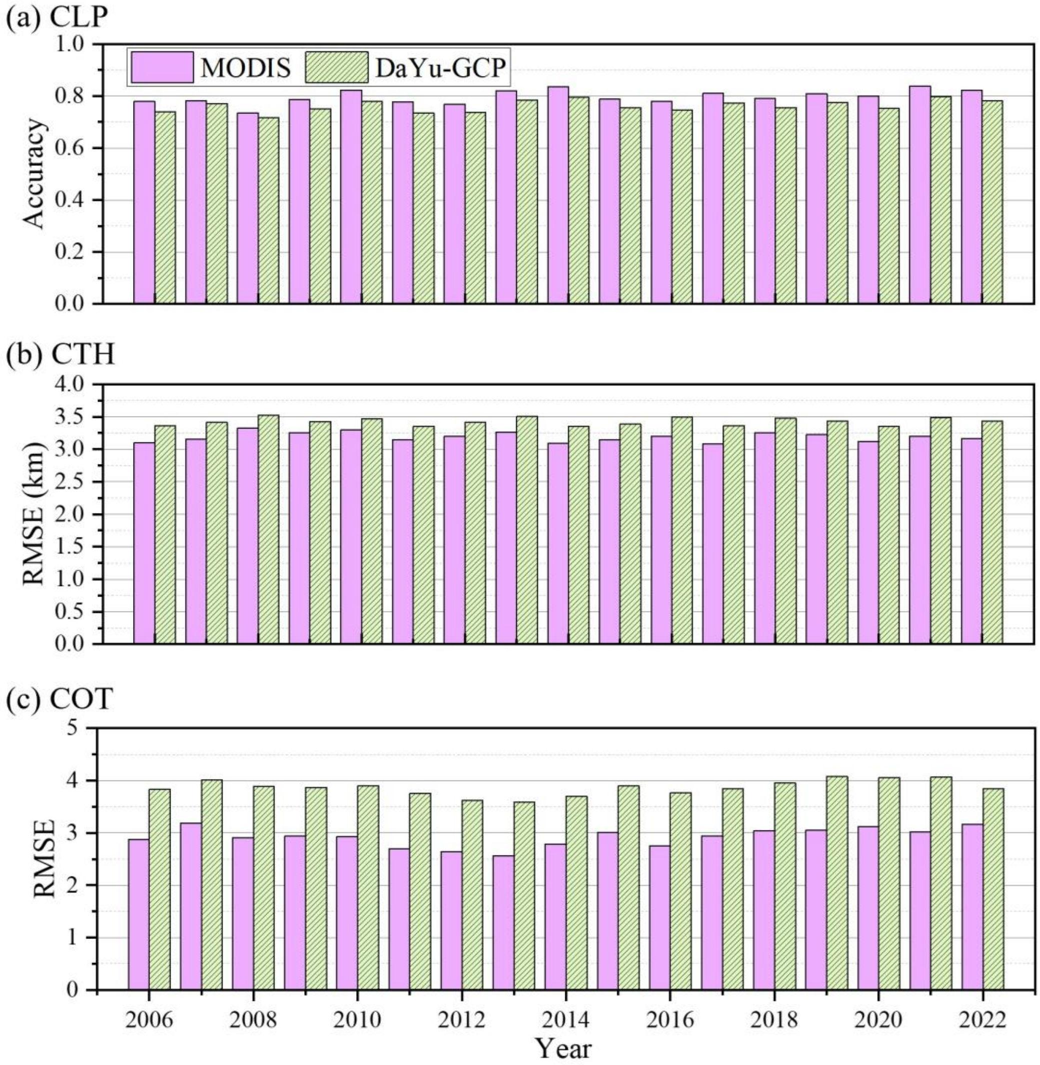
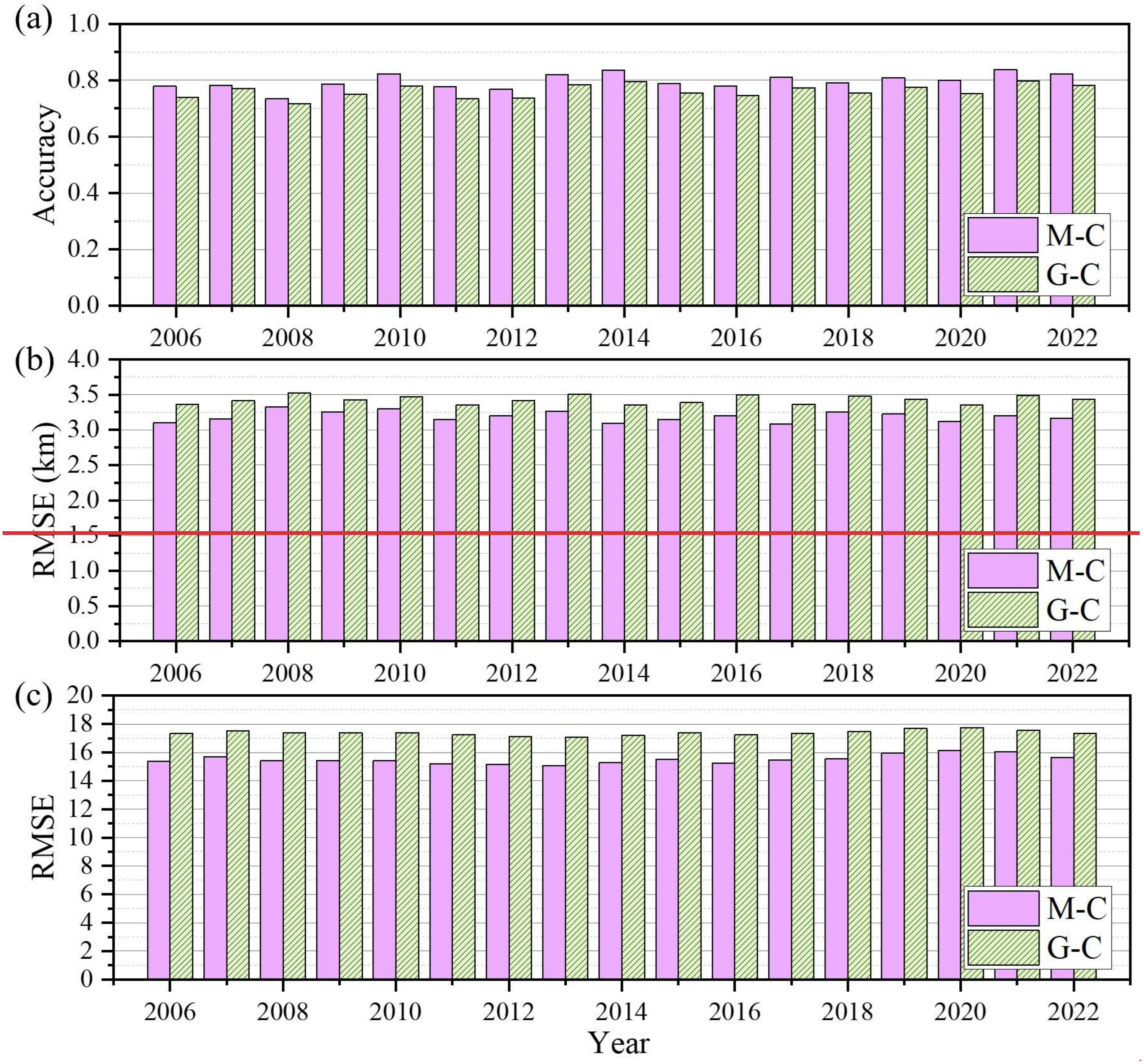


Figure 9. Annual evaluation results for MODIS and DaYu-GCP based on CALIOP standards. (a) Comparison of CLP accuracy. (b) Comparison of CTH RMSE. (c) Comparison of COT RMSE.

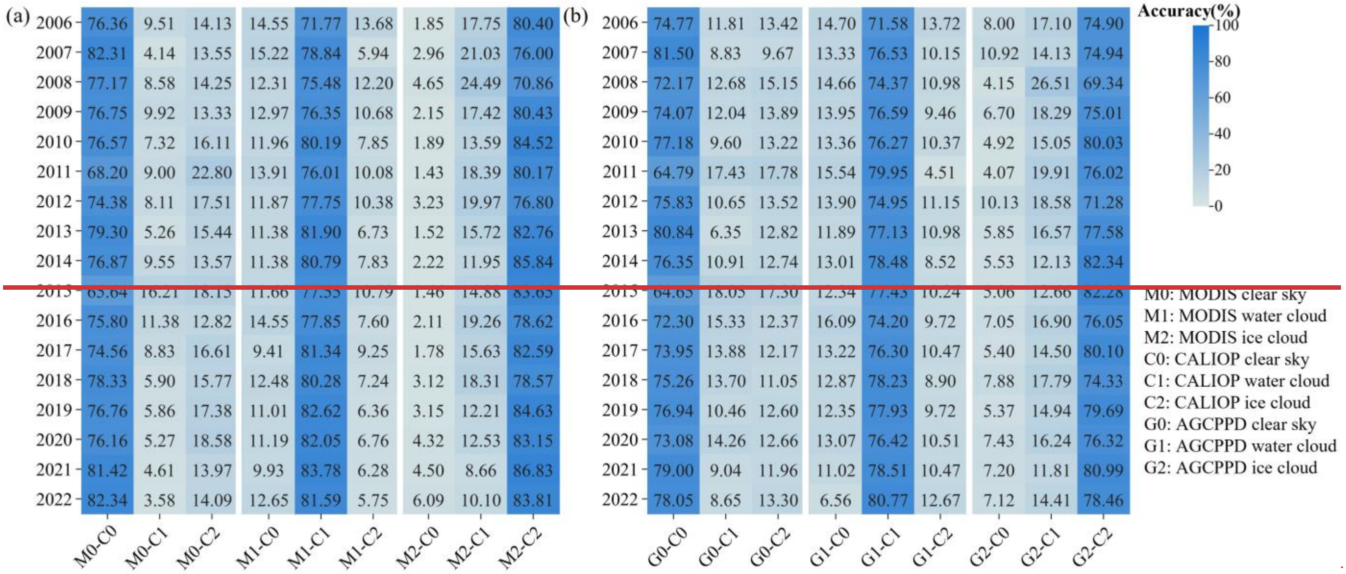
Table 4. The classification accuracy of clear sky, water cloud, and ice cloud for MODIS and DaYu-GCP on an annual basis, using CALIOP as the standard.

Year	MODIS (%)			DaYu-GCP (%)		
	Clear sky	Water cloud	Ice cloud	Clear sky	Water cloud	Ice cloud
2006	76.36	71.77	80.40	74.77	71.58	74.90
2007	82.31	78.84	76.00	81.50	76.53	74.94
2008	77.17	75.48	70.86	72.17	74.37	69.34

2009	76.75	76.35	80.43	74.07	76.59	75.01
2010	76.57	80.19	84.52	77.18	76.27	80.03
2011	68.20	76.01	80.17	64.79	79.95	76.02
2012	74.38	77.75	76.80	75.83	74.95	71.28
2013	79.30	81.90	82.76	80.84	77.13	77.58
2014	76.87	80.79	85.84	76.35	78.48	82.34
2015	65.64	77.55	83.65	64.65	77.43	82.28
2016	75.80	77.85	78.62	72.30	74.20	76.05
2017	74.56	81.34	82.59	73.95	76.30	80.10
2018	78.33	80.28	78.57	75.26	78.23	74.33
2019	76.76	82.62	84.63	76.94	77.93	79.69
2020	76.16	82.05	83.15	73.08	76.42	76.32
2021	81.42	83.78	86.83	79.00	78.51	80.99
2022	82.34	81.59	83.81	78.05	80.77	78.46



~~**Figure 10.** Annual evaluation results of AGCPP and CALIOP (G-C) and annual evaluation results of MODIS and CALIOP (M-C). M stands for MODIS, C stands for CALIOP, and G stands for AGCPP. (a) Comparison of CLP accuracy. (b) Comparison of CTH RMSE. (c) Comparison of COT RMSE.~~



(a)

	M0-C0	M0-C1	M0-C2	M1-C0	M1-C1	M1-C2	M2-C0	M2-C1	M2-C2
2006	76.36	9.51	14.13	14.55	71.77	13.68	1.85	17.75	80.40
2007	82.31	4.14	13.55	15.22	78.84	5.94	2.96	21.03	76.00
2008	77.17	8.58	14.25	12.31	75.48	12.20	4.65	24.49	70.86
2009	76.75	9.92	13.33	12.97	76.35	10.68	2.15	17.42	80.43
2010	76.57	7.32	16.11	11.96	80.19	7.85	1.89	13.59	84.52
2011	68.20	9.00	22.80	13.91	76.01	10.08	1.43	18.39	80.17
2012	74.38	8.11	17.51	11.87	77.75	10.38	3.23	19.97	76.80
2013	79.30	5.26	15.44	11.38	81.90	6.73	1.52	15.72	82.76
2014	76.87	9.55	13.57	11.38	80.79	7.83	2.22	11.95	85.84
2015	65.64	16.21	18.15	11.66	77.55	10.79	1.46	14.88	83.65
2016	75.80	11.38	12.82	14.55	77.85	7.60	2.11	19.26	78.62
2017	74.56	8.83	16.61	9.41	81.34	9.25	1.78	15.63	82.59
2018	78.33	5.90	15.77	12.48	80.28	7.24	3.12	18.31	78.57
2019	76.76	5.86	17.38	11.01	82.62	6.36	3.15	12.21	84.63
2020	76.16	5.27	18.58	11.19	82.05	6.76	4.32	12.53	83.15
2021	81.42	4.61	13.97	9.93	83.78	6.28	4.50	8.66	86.83
2022	82.34	3.58	14.09	12.65	81.59	5.75	6.09	10.10	83.81

(b)

	G0-C0	G0-C1	G0-C2	G1-C0	G1-C1	G1-C2	G2-C0	G2-C1	G2-C2
2006	74.77	11.81	13.42	14.70	71.58	13.72	8.00	17.10	74.90
2007	81.50	8.83	9.67	13.33	76.53	10.15	10.92	14.13	74.94
2008	72.17	12.68	15.15	14.66	74.37	10.98	4.15	26.51	69.34
2009	74.07	12.04	13.89	13.95	76.59	9.46	6.70	18.29	75.01
2010	77.18	9.60	13.22	13.36	76.27	10.37	4.92	15.05	80.03
2011	64.79	17.43	17.78	15.54	79.95	4.51	4.07	19.91	76.02
2012	75.83	10.65	13.52	13.90	74.95	11.15	10.13	18.58	71.28
2013	80.84	6.35	12.82	11.89	77.13	10.98	5.85	16.57	77.58
2014	76.35	10.91	12.74	13.01	78.48	8.52	5.53	12.13	82.34
2015	64.65	18.05	17.30	12.34	77.43	10.24	5.06	12.66	82.28
2016	72.30	15.33	12.37	16.09	74.20	9.72	7.05	16.90	76.05
2017	73.95	13.88	12.17	13.22	76.30	10.47	5.40	14.50	80.10
2018	75.26	13.70	11.05	12.87	78.23	8.90	7.88	17.79	74.33
2019	76.94	10.46	12.60	12.35	77.93	9.72	5.37	14.94	79.69
2020	73.08	14.26	12.66	13.07	76.42	10.51	7.43	16.24	76.32
2021	79.00	9.04	11.96	11.02	78.51	10.47	7.20	11.81	80.99
2022	78.05	8.65	13.30	6.56	80.77	12.67	7.12	14.41	78.46

M0: MODIS clear sky
M1: MODIS water cloud
M2: MODIS ice cloud
C0: CALIOP clear sky
C1: CALIOP water cloud
C2: CALIOP ice cloud
G0: AGCPPD clear sky
G1: AGCPPD water cloud
G2: AGCPPD ice cloud

~~**Figure 11.** Confusion matrix for CLP model evaluation. M stands for MODIS, C stands for CALIOP, and G stands for AGCPP. 0, 1, and 2 represent the pixel being classified as clear sky, water cloud, and ice cloud in the model, respectively. (a) Evaluation results of MODIS official cloud products and CALIOP official cloud products. (b) Evaluation results of AGCPP and CALIOP official cloud products.~~

3.5. Spatiotemporal distribution characteristics of clouds

Since MODIS is carried on polar-orbiting satellites, the spatial coverage of cloud products at any given time is limited, making large-scale, synchronous global observations difficult. Consequently, the spatiotemporal distribution characteristics of the retrieved cloud physical properties lack representativeness.

Here, we initially defined and calculated three physical quantities characterizing cloud cover based on the CLP product: ice cloud fraction (ICF), water cloud fraction (WCF), and total cloud fraction (TCF) (Zhao et al., 2024a). To ensure that spatiotemporally matched samples achieve global coverage, this study employed a 1.5 h temporal resolution and used MODIS transit times as the reference for spatiotemporal matching between DaYu-GCP and MODIS data. Monthly and annual average distributions were subsequently calculated.

Annual mean values for each year from 2020 to 2022 are presented in Fig. 10, while the corresponding seasonal means (DJF, MAM, JJA, and SON) are shown in Supplementary Figures S1–S4. Comparison of the results indicates that the MODIS official product and the DaYu-GCP product are highly consistent with respect to ICF, WCF, TCF, CTH, COT, and CER at MODIS transit times (Fig. 10 (a,d,g,j,m,p) and Fig. 10 (b,e,h,k,n,q)), particularly in equatorial regions. Since MODIS transits only at 10:30 AM and 1:30 PM local time, these annual averages do not fully capture the spatiotemporal characteristics of cloud physical properties throughout the day. In contrast, the all-day global DaYu-GCP, which provides complete spatial coverage, compensates for this limitation, representing a significant improvement over the official MODIS products (Fig. 10 (a,d,g,j,m,p) and Fig. 10 (c,f,i,l,o,r)).

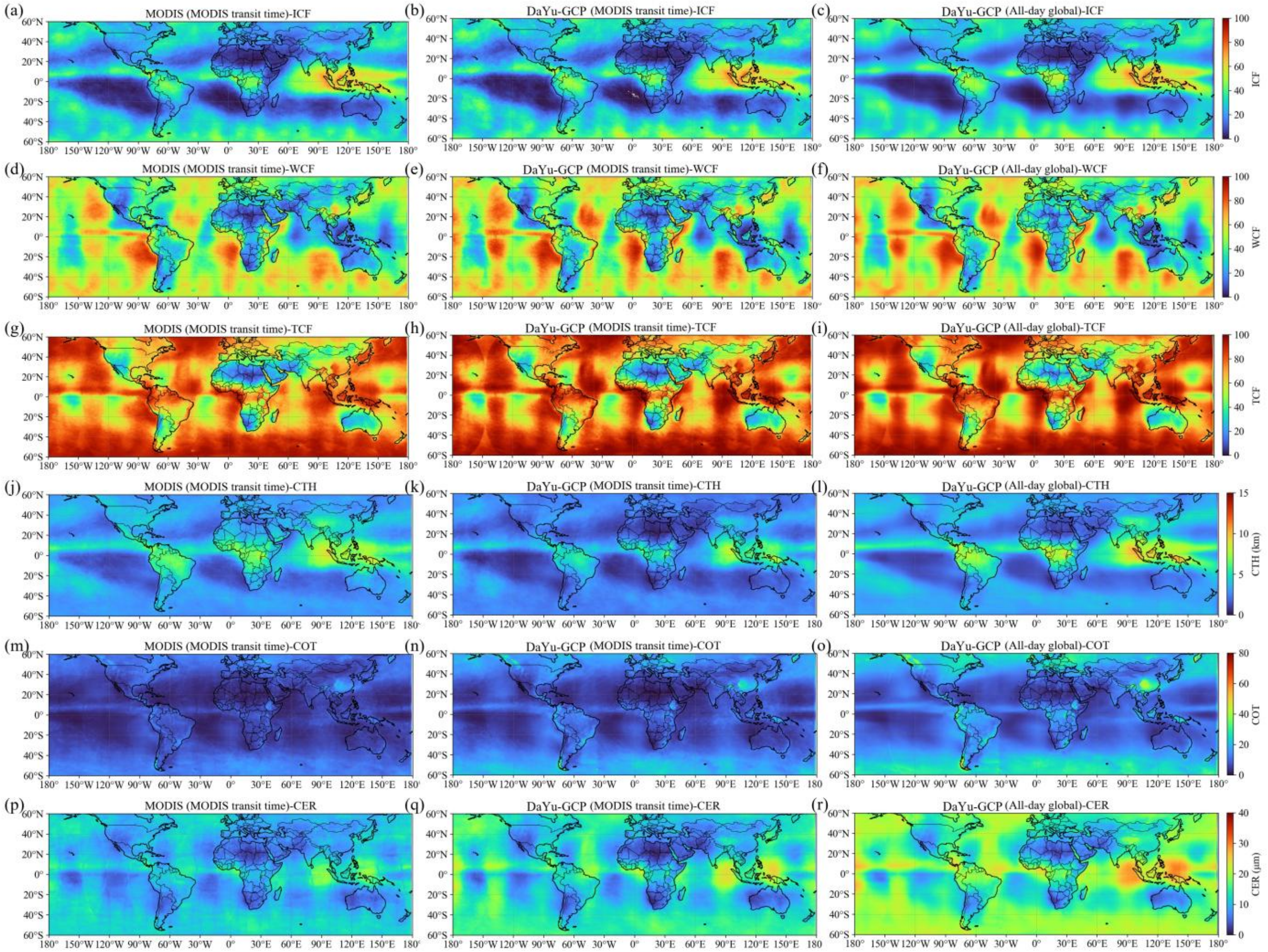


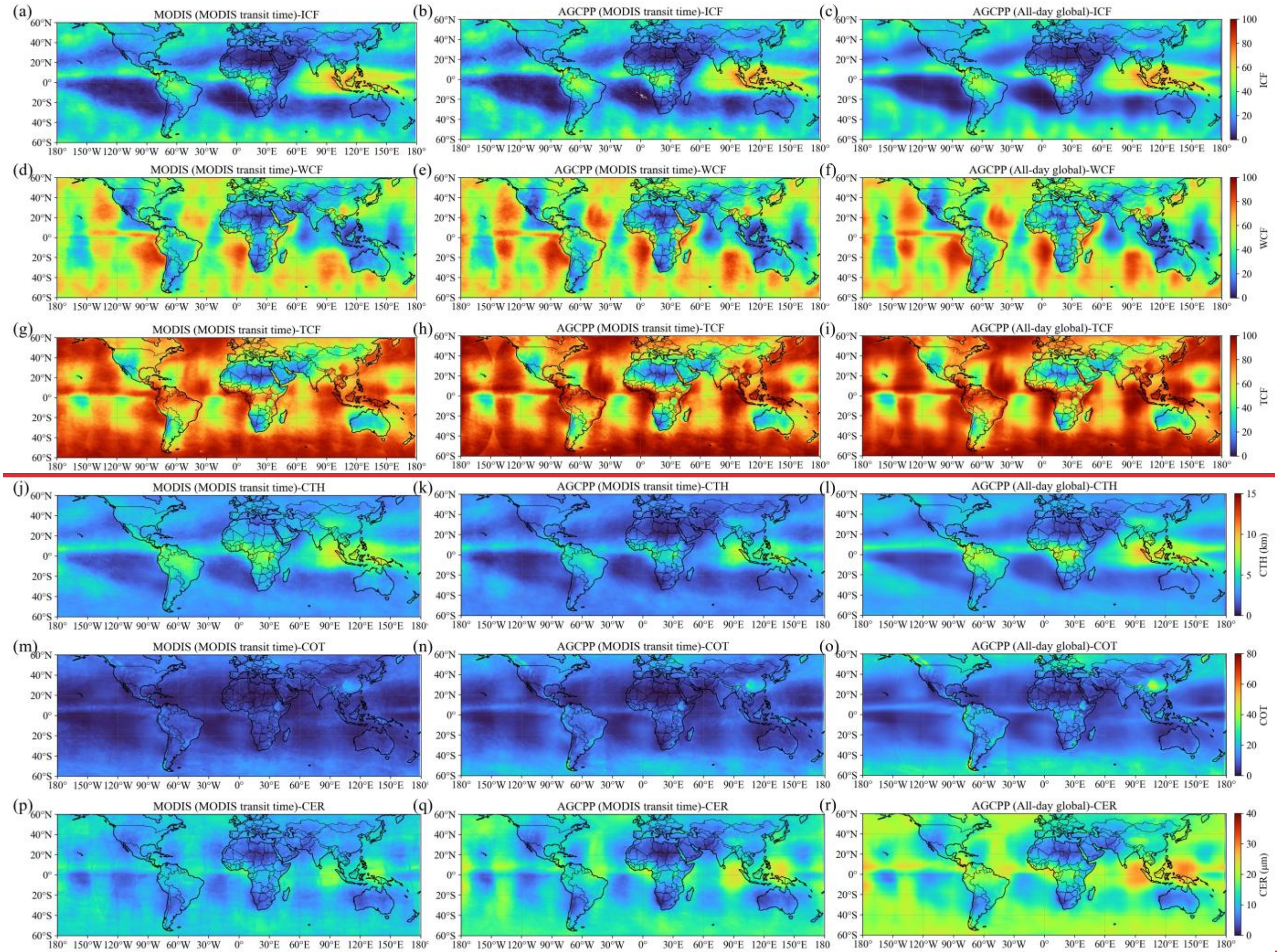
Figure 10. Annual average spatial distribution of global cloud physical products from 2020 to 2022, respectively: ICF (a–c), WCF (d–f), TCF (g–i), CTH (j–l), COT (m–o), and CER (p–r). Vertically, these are the averages of the MODIS official cloud products with a 1.5 h transit time interval (a, d, g, j, m, p), the averages of DaYu-GCP at the same MODIS transit time and coverage location (b, e, h, k, n, q), and the averages of DaYu-GCP's all-day global products at all times (c, f, i, l, o, r).

~~4 Basic characterization of AGCPP~~

~~AGCPP is also currently the world's first dataset capable of providing high spatio-temporal resolution cloud physical property products for the period 2000-2022. Since MODIS is located on polar-orbiting satellites, the spatial coverage of cloud products at the same time is insufficient, making it difficult to conduct large-scale synchronous observations of the entire globe. Therefore, the spatio-temporal distribution characteristics of the obtained cloud physical properties lack representativeness.~~

~~Here, we initially defined and calculated three physical quantities characterising cloud cover based on the CLP product: ice cloud fraction (ICF), water cloud fraction (WCF), and total cloud fraction (TCF) (Zhao et al., 2024). Since MODIS is insufficient to cover the entire globe at a 3 h time resolution, especially in the equatorial regions. Due to its 3 h temporal resolution, MODIS is insufficient to cover the entire globe, particularly in equatorial regions. Therefore, we employed a 1.5 h temporal resolution, averaging MODIS transit time images to obtain monthly average MODIS (MODIS transit time) images, as shown in the figure. Similarly, AGCPP also selected images with the same temporal and spatial coverage to obtain monthly average AGCPP (MODIS transit time) images. Missing time intervals were interpolated using adjacent time points. Finally, we also calculated the monthly average images of the AGCPP (All-day global) product.~~

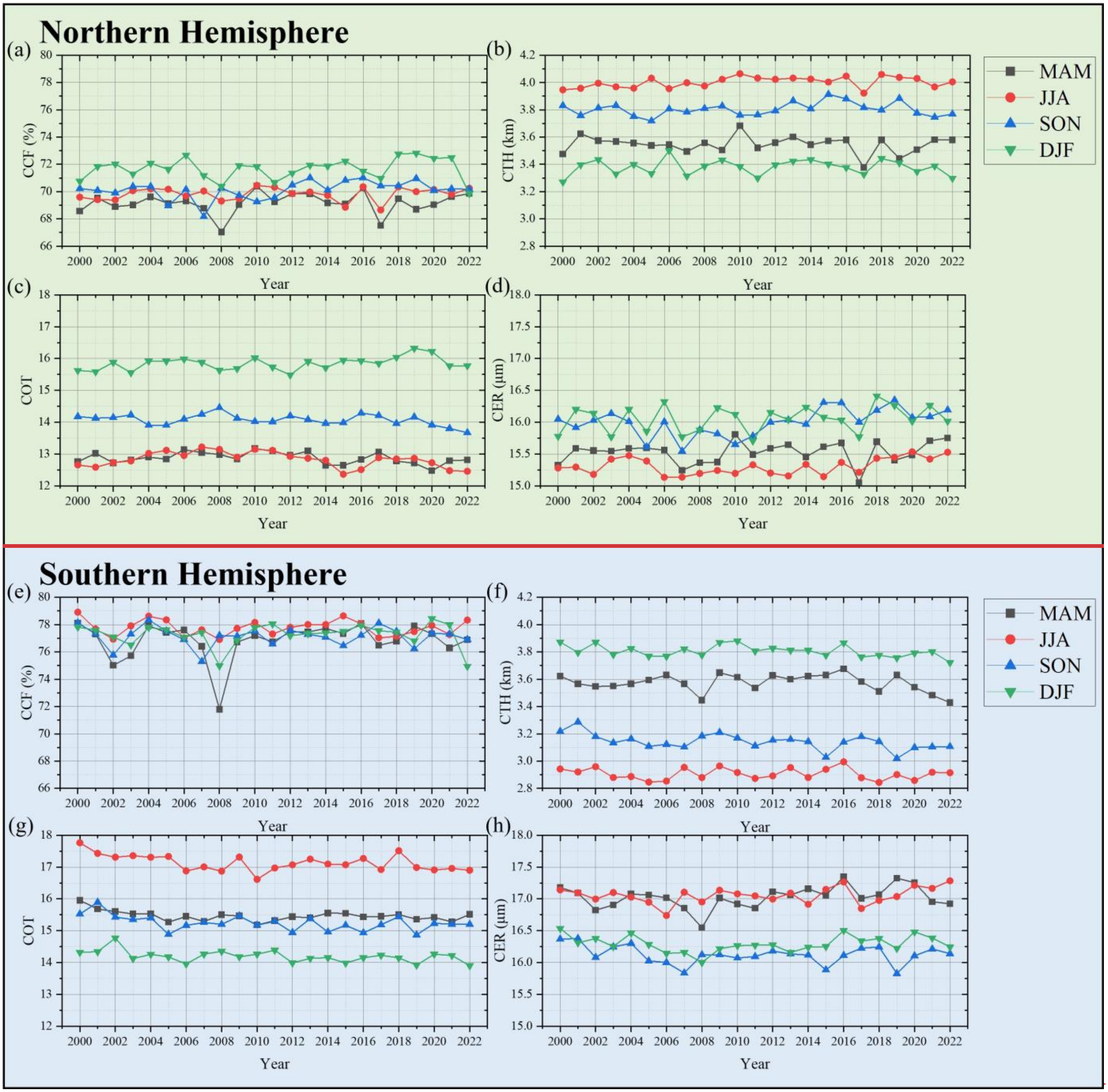
~~By calculating the seasonal averages (DJF, MAM, JJA, SON) of the monthly average results for each year from 2020 to 2022, we plotted Fig. 12, Fig. S 1, Fig. S 2, Fig. S 3, and Fig. S 4 (see Supplementary Materials Figure S1–S4). Based on the seasonal average results, we further calculated the annual average results for 2020–2022, as shown in Fig. 12. By comparing the results, it can be seen that the MODIS official product and the AGCPP product are highly consistent in terms of ICF, WCF, TCF, CTH, COT, and CER at the MODIS transit time, as shown in Fig. 12(a,d,g,j,m,p) and Fig. 12(b,e,h,k,n,q), especially in the equatorial regions. However, since the Terra/MODIS overpass time is approximately 10:30 AM local time, while the Aqua/MODIS overpass time is approximately 1:30 PM local time, the spatiotemporal distribution characteristics of the obtained cloud physical properties lack representativeness. Nevertheless, the all-day global product of AGCPP, which has global spatial coverage, can compensate for this deficiency. This is significantly different from the official MODIS products, as shown in Fig. 12(a,d,g,j,m,p) and Fig. 12(c,f,i,l,o,r).~~



~~**Figure 12.** Annual average spatial distribution of global cloud physical products from 2020 to 2022, respectively: ICF (a–c), WCF (d–f), TCF (g–i), CTH (j–l), COT (m–o), and CER (p–r). Vertically, these are the seasonal averages of the MODIS official cloud products with a 1.5 h transit time interval (a, d, g, j, m, p), the seasonal averages of AGCPP at the same MODIS transit time and coverage location (b, e, h, k, n, q), and the seasonal averages of AGCPP's all–day global products at all times (c, f, i, l, o, r).~~

~~Therefore, based on the spatio-temporal continuity and high accuracy of AGCPP products, this paper can comprehensively and accurately analyse the physical characteristics of clouds on a global scale. Here, we have compiled statistics on the changes in cloud cover frequency (CCF), CTH, COT, and CER over a period of 23 years between latitudes 60°S and 60°N. The basic cloud feature analysis based on AGCPP is shown in Fig. 13, with statistics separated for the Northern Hemisphere and Southern Hemisphere. For the Northern Hemisphere, the seasons are spring (MAM), summer (JJA), autumn (SON), and winter (DJF). For the Southern Hemisphere, the seasons are reversed: spring (SON), summer (DJF), autumn (MAM), and winter (JJA). As shown in Fig. 13(a)(e), CCF indicates that cloud cover in all seasons is higher~~

~~in the Southern Hemisphere than in the Northern Hemisphere. Additionally, the CTH values for the seasons in the Northern Hemisphere, ranked from highest to lowest, are summer, autumn, spring, and winter, which is completely consistent with the seasonal distribution of CTH in the Southern Hemisphere, as shown in Fig. 13(b)(f). The COT values in the Northern Hemisphere are ranked from highest to lowest as winter, autumn, spring/summer, while in the Southern Hemisphere they are winter, autumn/spring, summer, as shown in Fig. 13(c)(g). Finally, the CER values in the Northern Hemisphere are higher in autumn/winter than in spring/summer, while in the Southern Hemisphere they are consistent with the Northern Hemisphere, also higher in autumn/winter than in spring/summer, as shown in Fig. 13(d)(h).~~



~~**Figure 13.** Seasonal average change curves for CCF, CTH, COT, and CER for different years between latitudes 60°S-60°N from 2000 to 2023 based on AGCPP products. Statistics are divided into the northern and southern hemispheres. The black, red, blue, and green lines represent MAM, JJA, SON, and DJF, respectively. (a) and (e) are CCF variables. (b) and (f) are CTH variables. (c) and (g) are CTH variables. (d) and (h) are CTH variables.~~

Additionally, DaYu-GCP was compared with the global cloud datasets ISCCP and CLARA-A3. The four variables involved in the comparison were CCF, CTH, CER, and COT. For ISCCP, the base HGG data itself provides global 3 h CCF and COT. CTH can be derived from the provided CTT, and CER can be calculated from the available COT and cloud water path (CWP) (Schiffer and Rossow, 1983; Rossow et al., 1996; Young et al., 2018). For CLARA-A3, the dataset provides global daily mean CCF and CTH directly, whereas COT and CER can be derived from ice water path (IWP) and liquid water path (LWP) (Nakajima and King, 1990).

To clearly compare the annual variability of cloud physical properties among different products, monthly anomaly time series of each parameter were constructed for the Northern and Southern Hemispheres based on the monthly mean data from each product, and Pearson correlation coefficients (R) between DaYu-GCP and the other two datasets were calculated (Fig. 11). DaYu-GCP exhibits higher correlations with ISCCP in CCF, COT, and CER; for the Northern Hemisphere, the R reach 0.760, 0.764, and 0.514, respectively. DaYu-GCP shows higher correlation with CLARA-A3 in CTH, with R = 0.778 for the Northern Hemisphere. The Southern Hemisphere exhibits similar results. Among all R values, CER shows the lowest correlation compared to CCF and CTH; for the Northern Hemisphere, R values between DaYu-GCP and ISCCP, DaYu-GCP and CLARA-A3, and ISCCP and CLARA-A3 are 0.514, 0.412, and 0.219, respectively. This indicates that CER remains a challenging parameter in future cloud physical properties datasets. These comparison results demonstrate that, as a new cloud product, DaYu-GCP exhibits reasonable consistency in the spatiotemporal variability of major cloud parameters with the two long-standing global datasets, ISCCP and CLARA-A3. This underscores the reliable data quality of DaYu-GCP and its potential as a consistent data source for global cloud-climate studies.

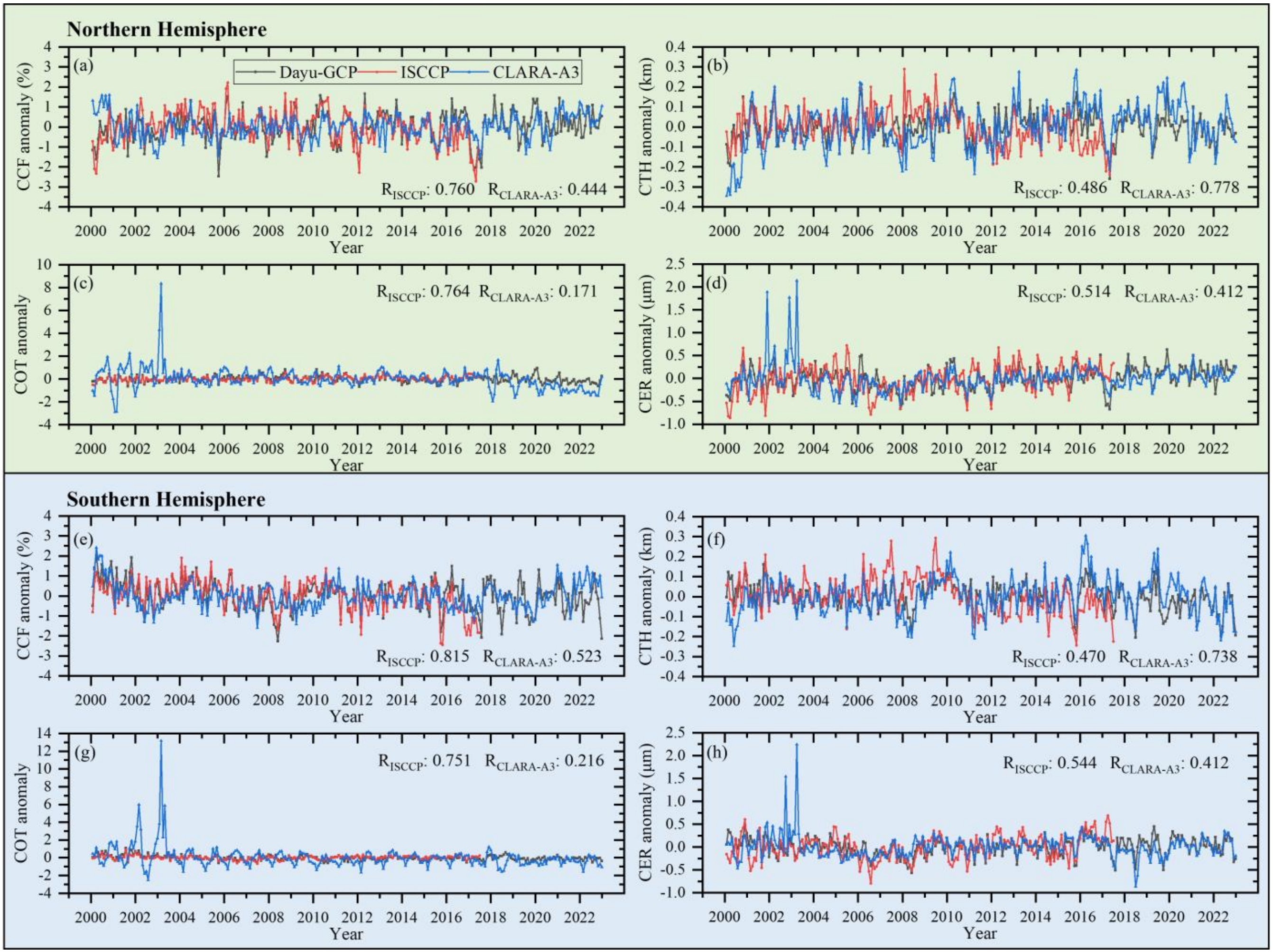


Figure 11. CCF, CTH, COT and CER anomaly change curves for the period 2000 to 2022 across the 60°N to 60°S latitude zone, derived from DaYu-GCP, ISCCP and CLARA-A3 products. Statistical data are presented separately for the Northern and Southern Hemispheres. The black, red and blue curves represent DaYu-GCP, ISCCP and CLARA-A3 respectively. In the corner, we have labelled the Pearson correlation coefficients R_{ISCCP} and $R_{CLARA-A3}$. R_{ISCCP} denotes the correlation coefficient between DaYu-GCP and ISCCP. $R_{CLARA-A3}$ denotes the correlation coefficient between DaYu-GCP and CLARA-A3. (a) and (e) are CCF variables. (b) and (f) are CTH variables. (c) and (g) are CTH variables. (d) and (h) are CTH variables.

4 Conclusion

This study developed and released the first all-day, high spatiotemporal resolution global cloud physical properties product, DaYu-GCP. The product spans the period from 2000 to 2022, with a temporal resolution of 3 h and a spatial resolution of 0.07°, covering all regions between 70°S and 70°N. It includes four key cloud physical properties: CLP, CTH, COT, and CER. Systematic quality assessments indicate that the DaYu-GCP product demonstrates reliable accuracy, strong stability, and good consistency. The specific conclusions are as follows:

(1) Compared with the MODIS official products, the classification Accuracy of CLP reaches 82.3%, and the RMSE of CTH, COT, and CER are 1.617 km, 11.314, and 7.181 μm, respectively. Interannual evaluations show that the CLP Accuracy remains stable at approximately 0.85 ± 0.007, while the RMSE of CTH, COT, and CER remain stable at 1.50 ± 0.03 km, 10.71 ± 0.15, and 6.75 ± 0.10 μm, with coefficients of variation of 2.0%, 1.4%, and 1.5%, respectively. This indicates that the accuracy of DaYu-GCP exhibits no significant interannual differences.

(2) Compared with the CALIOP official products, the day–night Accuracy of CLP are 0.787 and 0.775, respectively; the day–night RMSE of CTH are 3.384 km and 3.568 km, and the day–night RMSE of COT are 2.836 and 3.002. These results demonstrate good day–night consistency of the product.

(3) The product shows reasonable correlations with ISCCP and CLARA-A3 datasets in both hemispheres. For example, in the Northern Hemisphere, the correlations of CCF and COT with ISCCP are high (R = 0.760 and 0.764), while CTH shows a stronger correlation with CLARA-A3 (R = 0.778). The correlations of CER between DaYu-GCP and ISCCP, DaYu-GCP are slightly lower (R = 0.514 and 0.412), whereas the R between ISCCP and CLARA-A3 themselves is only 0.219, indicating that DaYu-GCP maintains higher spatiotemporal consistency with global cloud products ISCCP and CLARA-A3, although the CER retrieval algorithm still requires further development in future studies.

In summary, DaYu-GCP fills the gap in existing global cloud property products that simultaneously offer high spatiotemporal resolution and long-term coverage. The dataset combines high accuracy, strong stability, and robust consistency, and is publicly accessible, making it a valuable resource for studies of cloud–climate feedbacks and the global energy and water cycles. Furthermore, additional research is required in the future to explore how to enhance the data for cloud physical property products in the polar regions, thereby obtaining more reliable CER products.

5 Data availability

The DaYu-GCP data described in this paper have been made publicly available. We would like to express our special thanks to Science Data Bank and its staff for their help and support in the data publication process. DaYu-GCP can be accessed at https://doi.org/10.57760/sciencedb.26292 (Zhao et al., 2026). Moreover, the matching dataset employed to train the models in this study is also publicly available and can be accessed at https://doi.org/10.57760/sciencedb.27171 (Zhao, 2025).

~~5 Data availability~~

~~The AGCPP data described in this paper have been made publicly available. We would like to express our special thanks to Science Data Bank and its staff for their help and support in the data publication process. All data can be accessed at https://doi.org/10.57760/sciencedb.26292 (Zhao et al., 2025).~~

~~6 Conclusion~~

~~In order to provide a full-time, long time series, and complete continuum of global cloud physical property products, a dataset called AGCPP (containing CLP, CTH, COT, and CER products) was produced, covering latitudes from -70° to 70°, and longitudes from -180° to 180°. AGCPP provides data for every 3 h throughout the day from 2000 to 2022, with a spatial resolution of 0.07° for cloud physical properties. In this study, we evaluated the AGCPP data quality with the official Aqua and Terra/MODIS (MOD06/MYD06) cloud product and the official CALIPSO/CALIOP 1km cloud product. The results showed that AGCPP passed the temporal consistency assessment, the diurnal consistency assessment and the spatial consistency assessment. In addition, MODIS official cloud products and AGCPP were compared in terms of seasonal and annual averages using the same MODIS transit time and coverage area to ensure consistency. Furthermore, AGCPP all-day global spatio-temporal coverage was used to compensate for the lack of representativeness of MODIS spatial distribution. Finally, a simple statistical analysis of the physical characteristics of clouds in the northern and southern hemispheres was performed based on the AGCPP all-day global product. Due to the long time-series and all-day global nature of the dataset, it is expected that the dataset AGCPP will significantly increase the potential for climate change studies, especially on the potential feedback effects between clouds, surface albedo, and radiation. Considering the lengthy and extensive data processing involved in constructing IRBT2CPP, we have also chosen to make this dataset publicly available(https://doi.org/10.57760/sciencedb.27171) (Zhao, 2025). We look forward to researchers worldwide building upon our work to iterate and develop cloud physical property products with higher accuracy, thereby collectively advancing research in the field of cloud physical properties. In addition, more research needs to be conducted in the future regarding how to complement the data on cloud physical properties products at the North and South poles.~~

Author contribution

LZ contributed to conceptualization, methodology, writing–original draft, investigation, validation, data curation, drawing. FZ contributed to supervision, writing-review & editing, funding acquisition. ZZ contributed to investigation, discussion, writing-review & editing. FL contributed to supervision, discussion. JL contributed to investigation, validation. BG contributed to discussion, writing-review & editing. WL contributed to discussion, investigation. All authors reviewed the manuscript.

~~LZ contributed to conceptualization, methodology, writing–original draft, investigation, validation, data curation, drawing. FZ contributed to supervision, writing-review & editing, funding acquisition. JL contributed to investigation, validation. FL contributed to supervision, discussion. ZZ contributed to investigation, discussion. All authors reviewed the manuscript.~~

Competing interests

The authors declare that they have no conflict of interest.

References

~~Alexandrov, M. D., Cairns, B., Van Diedenhoven, B., and Emde, C.: Correction of Cloud Optical Thickness Retrievals from Nadir Reflectances in the Presence of 3D Radiative Effects. Part II: Theory, J. Atmos. Sci., 82, 933 − 941, https://doi.org/10.1175/JAS-D-24-0209.1, 2025.~~

~~Brennan, J. I., Kaufman, Y. J., Koren, I., and Rongrong, L.: Aerosol − cloud interaction—Misclassification of MODIS clouds in heavy aerosol, IEEE Trans. Geosci. Remote Sens., 43, 911, https://doi.org/10.1109/TGRS.2005.844662, 2005.~~

~~Cai, H., Zhang, S., Bu, K., Yang, J., and Chang, L.: Integrating geographical data and phenological characteristics derived from MODIS data for improving land cover mapping, J. Geogr. Sci., 21, 705 − 718, https://doi.org/10.1007/s11442-011-0874-1, 2011.~~

~~Chen, Y., Chen, G., Cui, C., Zhang, A., Wan, R., Zhou, S., Wang, D., and Fu, Y.: Retrieval of the vertical evolution of the cloud effective radius from the Chinese FY-4 (Feng Yun 4) next-generation geostationary satellites, Atmos. Chem. Phys., 20, 1131 − 1145, https://doi.org/10.5194/acp-20-1131-2020, 2020.~~

~~Frey, R. A., Ackerman, S. A., Liu, Y., Strabala, K. I., Zhang, H., Key, J. R., and Wang, X.: Cloud Detection with MODIS. Part I: Improvements in the MODIS Cloud Mask for Collection 5, J. Atmos. Ocean. Technol., 25, 1057 − 1072, https://doi.org/10.1175/2008JTECHA1052.1, 2008.~~

~~Gao, S., Lu, C., Zhu, J., Li, Y., Liu, Y., Zhao, B., Hu, S., Liu, X., and Lv, J.: Using Machine Learning to Predict Cloud Turbulent Entrainment-Mixing Processes, J. Adv. Model. Earth Syst., 16, e2024MS004225, https://doi.org/10.1029/2024MS004225, 2024.~~

~~Gunshor, M. M., Schmit, T. J., Menzel, W. P., and Tobin, D. C.: Intercalibration of Broadband Geostationary Imagers Using AIRS, J. Atmos. Ocean. Technol., 26, 746 − 758, https://doi.org/10.1175/2008JTECHA1155.1, 2009.~~

~~Hagihara, Y., Okamoto, H., and Yoshida, R.: Development of a combined CloudSat-CALIPSO cloud mask to show global cloud distribution, J. Geophys. Res. Atmos., 115, https://doi.org/10.1029/2009JD012344, 2010.~~

~~Heidinger, A. K. and Pavolonis, M. J.: Gazing at Cirrus Clouds for 25 Years through a Split Window. Part I: Methodology, J. Appl. Meteorol. Climatol., 48, 1100 – 1116, https://doi.org/10.1175/2008JAMC1882.1, 2009.~~

~~Heidinger, A. K., Pavolonis, M. J., Calvert, C., Hoffman, J., Nebuda, S., Straka, W., Walther, A., and Wanzong, S.: Chapter 6 – ABI Cloud Products from the GOES-R Series, in: The GOES-R Series, edited by Goodman, S. J., Schmit, T. J., Daniels, J., and Redmon, R. J., Elsevier, 43 – 62, https://doi.org/10.1016/B978-0-12-814327-8.00006-8, 2020.~~

~~Helder, D., Doelling, D., Bhatt, R., Choi, T., and Barsi, J.: Calibrating Geosynchronous and Polar Orbiting Satellites: Sharing Best Practices, https://doi.org/10.3390/rs12172786, 2020.~~

~~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, Q. J. R. Meteorol. Soc., 146, 1999 – 2049, https://doi.org/10.1002/qj.3803, 2020.~~

~~Hosen, M. K., Alam, M. S., Chakraborty, T., and Golder, M. R.: Monitoring spatiotemporal and seasonal variation of agricultural drought in Bangladesh using MODIS-derived vegetation health index, J. Earth Syst. Sci., 132, 188, https://doi.org/10.1007/s12040-023-02200-3, 2023.~~

~~Hsieh, T.-L., Vecchi, G. A., Wang, C., Yang, W., Zhang, B., and Soden, B. J.: Dependence of tropical cyclone seeds and climate sensitivity on tropical cloud response, Sci. Adv., 10, eadi2779, https://doi.org/10.1126/sciadv.adi2779, 2024.~~

~~Huang, H., Cui, C., Cheng, L., Liu, Q., and Wang, J.: Grid interpolation algorithm based on nearest neighbor fast search, Earth Sci. Inform., 5, 181 – 187, https://doi.org/10.1007/s12145-012-0106-y, 2012.~~

~~Iwabuchi, H., Yamada, S., Katagiri, S., Yang, P., and Okamoto, H.: Radiative and Microphysical Properties of Cirrus Cloud Inferred from Infrared Measurements Made by the Moderate Resolution Imaging Spectroradiometer (MODIS). Part I: Retrieval Method, J. Appl. Meteorol. Climatol., 53, 1297 – 1316, https://doi.org/10.1175/JAMC-D-13-0215.1, 2014.~~

~~Jiao, N., Xiang, Y., Wang, F., Zhou, G., and You, H.: Investigation of Global International GNSS Service Control Information Extraction for Geometric Calibration of Remote Sensing Images, Remote Sens., 16, 3860, 2024.~~

~~Joyce, R., Janowiak, J., and Huffman, G.: Latitudinally and Seasonally Dependent Zenith-Angle Corrections for Geostationary Satellite IR Brightness Temperatures, J. Appl. Meteorol., 40, 689 – 703, https://doi.org/10.1175/1520-0450(2001)040<0689:LASDZA>2.0.CO;2, 2001.~~

~~Kaps, A., Lauer, A., Camps-Valls, G., Gentine, P., Gómez-Chova, L., and Eyring, V.: Machine-Learned Cloud Classes From Satellite Data for Process-Oriented Climate Model Evaluation, IEEE Trans. Geosci. Remote Sens., 61, 1 – 15, https://doi.org/10.1109/TGRS.2023.3237008, 2023.~~

~~Karlsson, K.-G., Devasthale, A., and Eliasson, S.: Global Cloudiness and Cloud Top Information from AVHRR in the 42-Year CLARA-A3 Climate Data Record Covering the Period 1979 – 2020, Remote Sens., 15, 3044, 2023a.~~

~~Karlsson, K. G., Stengel, M., Meirink, J. F., Riihelä, A., Trentmann, J., Akkermans, T., Stein, D., Devasthale, A., Eliasson, S., Johansson, E., Håkansson, N., Solodovnik, I., Benas, N., Clerbaux, N., Selbach, N., Schröder, M., and Hollmann, R.: CLARA-A3: The third edition of the AVHRR-based CM SAF climate data record on clouds, radiation and surface albedo covering the period 1979 to 2023, Earth Syst. Sci. Data, 15, 4901–4926, https://doi.org/10.5194/essd-15-4901-2023, 2023b.~~

~~Kim, K.-H., Shim, P.-S., and Shin, S.: An Alternative Bilinear Interpolation Method Between Spherical Grids, Atmosphere, 10, 123, 2019.~~

~~Knapp, K. R., Ansari, S., Bain, C. L., Bourassa, M. A., Dickinson, M. J., Funk, C., Helms, C. N., Hennon, C. C., Holmes, C. D., Huffman, G. J., Kossin, J. P., Lee, H.-T., Loew, A., and Magnusdottir, G.: Globally Gridded Satellite Observations for Climate Studies, Bull. Am. Meteorol. Soc., 92, 893 – 907, https://doi.org/10.1175/2011BAMS3039.1, 2011.~~

~~Kotarba, A. Z. and Wojciechowska, I.: Satellite-based detection of deep-convective clouds: the sensitivity of infrared methods and implications for cloud climatology, Atmos. Meas. Tech., 18, 2721 – 2738, https://doi.org/10.5194/amt-18-2721-2025, 2025.~~

~~Kurihana, T., Moyer, E. J., and Foster, I. T.: AICCA: AI-Driven Cloud Classification Atlas, Remote Sens., 14, 5690, 2022.~~

~~Lee, Y., Kummerow, C. D., and Ebert-Uphoff, I.: Applying machine learning methods to detect convection using Geostationary Operational Environmental Satellite-16 (GOES-16) advanced baseline imager (ABI) data, Atmos. Meas. Tech., 14, 2699 – 2716, https://doi.org/10.5194/amt-14-2699-2021, 2021.~~

~~Lee, Y., Ahn, M.-H., Kang, M., Eo, M., Kim, D., and Moon, K.-J.: Advantages of Inter-Calibration for Geostationary Satellite Sensors Onboard Twin Satellites, Geophys. Res. Lett., 51, e2024GL109364, https://doi.org/10.1029/2024GL109364, 2024.~~

~~Letu, H., Yang, K., Nakajima, T., Ishimoto, H., Nagao, T., Riedi, J., Baran, A., Ma, R., Wang, T., Khatri, P., Chen, L., Shi, C., and Shi, J.: High-resolution retrieval of cloud microphysical properties and surface solar radiation using Himawari-8/AHI next-generation geostationary satellite, Remote Sens. Environ., 239, 1 – 16, https://doi.org/10.1016/j.rse.2019.111583, 2020.~~

~~Letu, H., Ma, R., Nakajima, T. Y., Shi, C., Hashimoto, M., Nagao, T. M., Baran, A. J., Nakajima, T., Xu, J., Wang, T., Tana, G., Bilige, S., Shang, H., Chen, L., Ji, D., Lei, Y., Wei, L., Zhang, P., Li, J., Li, L., Zheng, Y., Khatri, P., and Shi, J.: Surface Solar Radiation Compositions Observed from Himawari-8/9 and Fengyun-4 Series, Bull. Am. Meteorol. Soc., 104, E1772 – E1789, https://doi.org/10.1175/BAMS-D-22-0154.1, 2023.~~

~~Li, J., Zhang, F., Li, W., Tong, X., Pan, B., Li, J., Lin, H., Letu, H., and Mustafa, F.: Transfer-Learning-Based Approach to Retrieve the Cloud Properties Using Diverse Remote Sensing Datasets, IEEE Trans. Geosci. Remote Sens., 61, 1–10, https://doi.org/10.1109/TGRS.2023.3318374, 2023.~~

~~Liu, C., Zhang, F., Ouyang, H., Li, W., and Zhao, Z.: Diurnal Variation of Cloud Physical Properties for Tropical Cyclones Over North Atlantic in 2019 – 2023, Geophys. Res. Lett., 52, e2025GL115566, https://doi.org/10.1029/2025GL115566, 2025.~~

~~Liu, C., Song, Y., Zhou, G., Teng, S., Li, B., Xu, N., Lu, F., and Zhang, P.: A cloud optical and microphysical property product for the advanced geosynchronous radiation imager onboard China's Fengyun-4 satellites: The first version, Atmos. Oceanic Sci. Lett., 16, 100337, https://doi.org/10.1016/j.aosl.2023.100337, 2023.~~

~~Liu, J., Shi, C., Ge, L., Tie, R., Chen, X., Zhou, T., Gu, X., and Shen, Z.: Enhanced Wind Field Spatial Downscaling Method Using UNET Architecture and Dual Cross-Attention Mechanism, Remote Sens., 16, 1867, 2024a.~~

~~Liu, W. H., Yue, P., Wu, X. H., Li, J. J., Shao, N. F., Zhu, B., and Lu, C. S.: Why does a decrease in cloud amount increase terrestrial evapotranspiration in a monsoon transition zone?, Environ. Res. Lett., 19, https://doi.org/10.1088/1748-9326/ad3569, 2024b.~~

~~Liu, Y., Key, J. R., Frey, R. A., Ackerman, S. A., and Menzel, W. P.: Nighttime polar cloud detection with MODIS, Remote Sens. Environ., 92, 181 – 194, https://doi.org/10.1016/j.rse.2004.06.004, 2004.~~

~~Mayer, J., Bugliaro, L., Mayer, B., Piontek, D., and Voigt, C.: Bayesian cloud-top phase determination for Meteosat Second Generation, Atmos. Meas. Tech., 17, 4015 – 4039, https://doi.org/10.5194/amt-17-4015-2024, 2024.~~

~~Menzel, W. P., Frey, R. A., Zhang, H., Wylie, D. P., Moeller, C. C., Holz, R. E., Maddux, B., Baum, B. A., Strabala, K. I., and Gumley, L. E.: MODIS Global Cloud-Top Pressure and Amount Estimation: Algorithm Description and Results, J. Appl. Meteorol. Climatol., 47, 1175 – 1198, https://doi.org/10.1175/2007JAMC1705.1, 2008.~~

~~Meyer, K., Yang, Y., and Platnick, S.: Uncertainties in cloud phase and optical thickness retrievals from the Earth Polychromatic Imaging Camera (EPIC), Atmos. Meas. Tech., 9, 1785 – 1797, https://doi.org/10.5194/amt-9-1785-2016, 2016.~~

~~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 Sens. Environ., 239, 111616, https://doi.org/10.1016/j.rse.2019.111616, 2020.~~

~~Minnis, P., Nguyen, L., Palikonda, R., Heck, P., Spangenberg, D., Doelling, D., Ayers, J., Smith Sr, W., Khaiyer, M., Trepte, Q ., Avey, L., Chang, F.-L., Yost, C., Chee, T., and Sun-Mack, S.: Near-real time cloud retrievals from operational and research meteorological satellites, Proc. SPIE, 7107, https://doi.org/10.1117/12.800344, 2008.~~

~~Minnis, P., Sun-Mack, S., Chen, Y., Chang, F. L., Yost, C. R., Smith, W. L., Heck, P. W., Arduini, R. F., Bedka, S. T., Yi, Y., Hong, G., Jin, Z., Painemal, D., Palikonda, R., Scarino, B. R., Spangenberg, D. A., Smith, R. A., Trepte, Q. Z., Yang, P.,~~

~~and Xie, Y.: CERES MODIS Cloud Product Retrievals for Edition 4 — Part I: Algorithm Changes, IEEE Trans. Geosci. Remote Sens., 59, 2744 – 2780, https://doi.org/10.1109/TGRS.2020.3008866, 2021.~~

~~Mitra, A., Di Girolamo, L., Hong, Y., Zhan, Y., and Mueller, K. J.: Assessment and Error Analysis of Terra-MODIS and MISR Cloud-Top Heights Through Comparison With ISS-CATS Lidar, J. Geophys. Res. Atmos., 126, e2020JD034281, https://doi.org/10.1029/2020JD034281, 2021.~~

~~Mouri, K.: Improvement of the cloud top height algorithm for the fundamental cloud product and related evaluation, Japan Meteorological Agency, 2019.~~

~~Pérez, J. C., Cerdeña, A., González, A., and Armas, M.: Nighttime cloud properties retrieval using MODIS and artificial neural networks, Adv. Space Res., 43, 852 858, https://doi.org/10.1016/j.asr.2008.06.013, 2009.~~

~~Platnick, S., Ackerman, S. A., King, M. D., Meyer, K., Menzel, W. P., Holz, R. E., Baum, B. A., and Yang, P.: MODIS atmosphere L2 cloud product (06_L2). NASA MODIS Adaptive Processing System, Goddard Space Flight Center [data set], Terra (MOD06_L2): 10.5067/MODIS/MOD06_L2.006, 2015.~~

~~Poulsen, C. A., Siddans, R., Thomas, G. E., Sayer, A. M., Grainger, R. G., Campmany, E., Dean, S. M., Arnold, C., and Watts, P. D.: Cloud retrievals from satellite data using optimal estimation: evaluation and application to ATSR, Atmos. Meas. Tech., 5, 1889 – 1910, https://doi.org/10.5194/amt-5-1889-2012, 2012.~~

~~Shi, C., Letu, H., Nakajima, T. Y., Nakajima, T., Wei, L., Xu, R., Lu, F., Riedi, J., Ichii, K., Zeng, J., Shang, H., Ma, R., Yin, S., Shi, J., Baran, A. J., Xu, J., Li, A., Tana, G., Wang, W., Na, Q., Sun, Q., Yang, W., Chen, L., and Shi, G.: Near-global monitoring of surface solar radiation through the construction of a geostationary satellite network observation system, The Innovation, 6, 100876, https://doi.org/10.1016/j.xinn.2025.100876, 2025.~~

~~Smalley, K. M. and Lebsock, M. D.: Corrections for Geostationary Cloud Liquid Water Path Using Microwave Imagery, J. Atmos. Ocean. Technol., 40, 1049 – 1061, https://doi.org/10.1175/JTECH-D-23-0030.1, 2023.~~

~~Takahashi, N., Hayasaka, T., and Okamoto, H.: Differences in Ice Cloud Microphysical Properties between Western and Eastern Tropical Pacific Regions Derived from CloudSat and CALIPSO Measurements, Sola, 12, 91 – 95, 2016.~~

~~Tang, C., Shi, C., Letu, H., Yin, S., Nakajima, T., Sekiguchi, M., Xu, J., Zhao, M., Ma, R., and Wang, W.: Development of a hybrid algorithm for the simultaneous retrieval of aerosol optical thickness and fine-mode fraction from multispectral satellite observation combining radiative transfer and transfer learning approaches, Remote Sens. Environ., 319, 114619, https://doi.org/10.1016/j.rse.2025.114619, 2025.~~

~~Tong, X., Li, J., Zhang, F., Li, W., Pan, B., Li, J., and Letu, H.: The Deep-Learning-Based Fast Efficient Nighttime Retrieval of Thermodynamic Phase From Himawari-8 AHI Measurements, Geophys. Res. Lett., 50, e2022GL100901, https://doi.org/10.1029/2022GL100901, 2023.~~

~~Trebing, K., Sta ǹ czyk, T., and Mehrkanoon, S.: SmaAt-UNet: Precipitation nowcasting using a small attention-UNet architecture, Pattern Recognit. Lett., 145, 178 – 186, https://doi.org/10.1016/j.patrec.2021.01.036, 2021.~~

~~Viggiano, M., Cimini, D., De Natale, M. P., Di Paola, F., Gallucci, D., Larosa, S., Marro, D., Nilo, S. T., and Romano, F.: Combining Passive Infrared and Microwave Satellite Observations to Investigate Cloud Microphysical Properties: A Review, Remote Sens., 17, 337, 2025.~~

~~Walther, A. and Heidinger, A. K.: Implementation of the Daytime Cloud Optical and Microphysical Properties Algorithm (DCOMP) in PATMOS-x, J. Appl. Meteorol. Climatol., 51, 1371 – 1390, https://doi.org/10.1175/JAMC-D-11-0108.1, 2012.~~

~~Wang, M., Min, M., Li, J., Lin, H., Liang, Y., Chen, B., Yao, Z., Xu, N., and Zhang, M.: Technical note: Applicability of physics-based and machine-learning-based algorithms of a geostationary satellite in retrieving the diurnal cycle of cloud base height, Atmos. Chem. Phys., 24, 14239 – 14256, https://doi.org/10.5194/acp-24-14239-2024, 2024.~~

~~Wang, T., Fetzer, E. J., Wong, S., Kahn, B. H., and Yue, Q.: Validation of MODIS cloud mask and multilayer flag using CloudSat-CALIPSO cloud profiles and a cross-reference of their cloud classifications, J. Geophys. Res. Atmos., 121, 11,620 – 11,635, https://doi.org/10.1002/2016JD025239, 2016.~~

~~Watts, P. D., Bennartz, R., and Fell, F.: Retrieval of two-layer cloud properties from multispectral observations using optimal estimation, J. Geophys. Res. Atmos., 116, https://doi.org/10.1029/2011JD015883, 2011.~~

~~Weisz, E., Li, J., Menzel, W. P., Heidinger, A. K., Kahn, B. H., and Liu, C.-Y.: Comparison of AIRS, MODIS, CloudSat and CALIPSO cloud top height retrievals, Geophys. Res. Lett., 34, https://doi.org/10.1029/2007GL030676, 2007.~~

~~Wolters, E. L. A., Roebeling, R. A., and Feijt, A. J.: Evaluation of Cloud-Phase Retrieval Methods for SEVIRI on Meteosat-8 Using Ground-Based Lidar and Cloud Radar Data, J. Appl. Meteorol. Climatol., 47, 1723 – 1738, https://doi.org/10.1175/2007JAMC1591.1, 2008.~~

~~Yan, Y., Tan, J., Ying, M., Liu, Y., and Tang, Y.: Vertical structures of typhoon cloud microphysical and radiative features associated with the precipitation type over the western North Pacific, Atmos. Res., 311, 107693, https://doi.org/10.1016/j.atmosres.2024.107693, 2024.~~

~~Young, A. H., Knapp, K. R., Inamdar, A., Hankins, W., and Rossow, W. B.: The International Satellite Cloud Climatology Project H-Series climate data record product, Earth Syst. Sci. Data, 10, 583–593, https://doi.org/10.5194/essd-10-583-2018, 2018.~~

~~Zhang, H., Dong, X., Xi, B., Xin, X., Liu, Q., He, H., Xie, X., Li, L., and Yu, S.: Retrieving high-resolution surface photosynthetically active radiation from the MODIS and GOES-16 ABI data, Remote Sens. Environ., 260, 112436, https://doi.org/10.1016/j.rse.2021.112436, 2021.~~

~~Zhang, M., Xu, N., Chen, L., Liu, J., and Yang, C.: Consistency Assessment of Cloud Top Property Products between FY-4B and FY-4A, J. Trop. Meteorol., 40, 389 – 396, https://doi.org/10.16032/j.issn.1004-4965.2024.035, 2024.~~

~~Zhang, Q., Liu, C., Zhang, R., Yang, J., Li, J., Li, B., Lu, F., Yin, Y., and Jin, H.: Assessment of Geostationary Satellite Cloud Effective Radius over East China with In Situ Aircraft Measurements, J. Meteorol. Res., 39, 240 – 251, https://doi.org/10.1007/s13351-025-4164-8, 2025.~~

~~Zhang, Z., Dong, X., Xi, B., Song, H., Ma, P.-L., Ghan, S. J., Platnick, S., and Minnis, P.: Intercomparisons of marine boundary layer cloud properties from the ARM CAP-MBL campaign and two MODIS cloud products, J. Geophys. Res. Atmos., 122, 2351 – 2365, https://doi.org/10.1002/2016JD025763, 2017.~~

~~Zhao, L.: IRBT2CPP: A training and testing dataset for machine learning training, ranging from infrared brightness temperature to cloud physical properties, Science Data Bank [data set], https://doi.org/10.57760/sciencedb.27171, 2025.~~

~~Zhao, L., Zhang, F., Li, J., Lu, F., and Zhao, Z.: AGCPP: All-day Global Cloud Physical Properties Dataset with 0.07° Resolution Retrieved from Geostationary Satellite Imagers (2000 – 2022), Science Data Bank [data set], https://doi.org/10.57760/sciencedb.26292, 2025.~~

~~Zhao, Z., Zhang, F., Li, W., and Li, J.: Image-Based Retrieval of All-Day Cloud Physical Parameters for FY4A/AGRI and Its Application Over the Tibetan Plateau, J. Geophys. Res. Atmos., 129, e2024JD041032, https://doi.org/10.1029/2024JD041032, 2024.~~

~~Zhao, Z., Zhang, F., Wu, Q., Li, Z., Tong, X., Li, J., and Han, W.: Cloud Identification and Properties Retrieval of the Fengyun-4A Satellite Using a ResUnet Model, IEEE Trans. Geosci. Remote Sens., 61, 1 – 18, https://doi.org/10.1109/TGRS.2023.3252023, 2023.~~

~~Zhong, Y., Shi, Z., Zhang, Y., Zhang, Y., and Li, H.: CSAN-UNet: Channel Spatial Attention Nested UNet for Infrared Small Target Detection, Remote Sens., 16, 1894, 2024.~~

~~Zhou, Y., Liu, Y., Han, W., Zeng, Y., Sun, H., Yu, P., and Zhu, L.: Exploring the characteristics of Fengyun-4A Advanced Geostationary Radiation Imager (AGRI) visible reflectance using the China Meteorological Administration Mesoscale (CMA-MESO) forecasts and its implications for data assimilation, Atmos. Meas. Tech., 17, 6659 – 6675, https://doi.org/10.5194/amt-17-6659-2024, 2024.~~

~~Zhuge, X. and Zou, X.: Summertime Convective Initiation Nowcasting over Southeastern China Based on Advanced Himawari Imager Observations, J. Meteorol. Soc. Jpn. Ser. II, 96, 337 – 353, https://doi.org/10.2151/jmsj.2018-041, 2018.~~

~~Zhuge, X. Y., Guan, J., Yu, F., and Wang, Y.: A New Satellite-Based Indicator for Estimation of the Western North Pacific Tropical Cyclone Current Intensity, IEEE Trans. Geosci. Remote Sens., 53, 5661 – 5676, https://doi.org/10.1109/TGRS.2015.2427035, 2015.~~

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, J. Meteorol. Soc. Jpn. Ser. II, 94, 151–183, https://doi.org/10.2151/jmsj.2016-009, 2016a.

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, J. Meteorol. Soc. Jpn., 94, 151–183, 2016b.

Bin, T., John, D., Robert, W., and Alan, R.: GOES-16 ABI navigation assessment, Proc. SPIE, 107640G, https://doi.org/10.1117/12.2321170, 2018.

Bin, T., John, D., Robert, W., and Alan, R.: GOES-16 and GOES-17 ABI INR assessment, Proc. SPIE, 111271D, https://doi.org/10.1117/12.2529336, 2019.

Bin, T., John, J. D., Christopher, N. F., Thomas, J. G., Scott, H., Peter, J. I., Patrick, D. J., Brian, C. P., Alan, D. R., Pradeep, T., and Robert, E. W.: GOES-R series image navigation and registration performance assessment tool set, J. Appl. Remote Sens., 14, 032405, https://doi.org/10.1117/1.JRS.14.032405, 2020.

Breiman, L.: Random Forests, Mach. Learn., 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Cai, H., Zhang, S., Bu, K., Yang, J., and Chang, L.: Integrating geographical data and phenological characteristics derived from MODIS data for improving land cover mapping, J. Geogr. Sci., 21, 705–718, https://doi.org/10.1007/s11442-011-0874-1, 2011.

Coste, P., Pasternak, F., Faure, F., Jacquet, B., Bianchi, S., Donny, M. A. A., Luhmann, H. J., Hanson, C., Pili, P., and Fowler, G.: SEVIRI, the imaging radiometer on Meteosat second generation: in-orbit results and first assessment, Proc. SPIE, 105680L, https://doi.org/10.1117/12.2308023, (year not given).

Donny Maladji, A. A., Bernard, J., and Frederick, P.: Characteristics of the Meteosat Second Generation (MSG) radiometer/imager: SEVIRI, Proc. SPIE, 19–31, https://doi.org/10.1117/12.298084, 1997.

Hagihara, Y., Okamoto, H., and Yoshida, R.: Development of a combined CloudSat-CALIPSO cloud mask to show global cloud distribution, J. Geophys. Res. Atmos., 115, https://doi.org/10.1029/2009JD012344, 2010.

Heidinger, A. K.: GOES-R Advanced Baseline Imager (ABI) Algorithm Theoretical Basis Document for Cloud Height (Version 3.0). 79 pp., https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Baseline/ATBD_GOES-R_Cloud_Height_v3.0_Jul2012.pdf, 2012.

Heidinger, A. K., Pavolonis, M. J., Calvert, C., Hoffman, J., Nebuda, S., Straka, W., Walther, A., and Wanzong, S.: Chapter 6 - ABI Cloud Products from the GOES-R Series, in: The GOES-R Series, edited by: Goodman, S. J., Schmit, T. J., Daniels, J., and Redmon, R. J., Elsevier, 43–62, https://doi.org/10.1016/B978-0-12-814327-8.00006-8, 2020.

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, Q. J. R. Meteorol. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hosen, M. K., Alam, M. S., Chakraborty, T., and Golder, M. R.: Monitoring spatiotemporal and seasonal variation of agricultural drought in Bangladesh using MODIS-derived vegetation health index, J. Earth Syst. Sci., 132, 188, https://doi.org/10.1007/s12040-023-02200-3, 2023.

Imai, T. and Yoshida, R.: Algorithm Theoretical Basis for Himawari-8 Cloud Mask Product, Meteorological Satellite Center Technical Note, 61, 1–17, https://www.data.jma.go.jp/mscweb/technotes/msctechrep61-1.pdf, 2016.

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, J. Meteorol. Soc. Jpn., 96B, 27–42, https://doi.org/10.2151/jmsj.2018-001, 2018.

Joyce, R., Janowiak, J., and Huffman, G.: Latitudinally and Seasonally Dependent Zenith-Angle Corrections for Geostationary Satellite IR Brightness Temperatures, J. Appl. Meteorol., 40, 689, https://doi.org/10.1175/1520-0450(2001)040<0689:LASDZA>2.0.CO;2, 2001.

Karlsson, K.-G., Devasthale, A., and Eliasson, S.: Global Cloudiness and Cloud Top Information from AVHRR in the 42-Year CLARA-A3 Climate Data Record Covering the Period 1979–2020, Remote Sens., 15, 3044, 2023a.

Karlsson, K. G., Stengel, M., Meirink, J. F., Riihelä, A., Trentmann, J., Akkermans, T., Stein, D., Devasthale, A., Eliasson, S., Johansson, E., Håkansson, N., Solodovnik, I., Benas, N., Clerbaux, N., Selbach, N., Schröder, M., and Hollmann, R.: CLARA-A3: The third edition of the AVHRR-based CM SAF climate data record on clouds, radiation and surface albedo covering the period 1979 to 2023, Earth Syst. Sci. Data, 15, 4901–4926, https://doi.org/10.5194/essd-15-4901-2023, 2023b.

Kawamoto, K., Nakajima, T., and Nakajima, T. Y.: A Global Determination of Cloud Microphysics with AVHRR Remote Sensing, J. Climate, 14, 2054–2068, https://doi.org/10.1175/1520-0442(2001)014<2054:AGDOCM>2.0.CO;2, 2001.

Knapp, K. R., Ansari, S., Bain, C. L., Bourassa, M. A., Dickinson, M. J., Funk, C., Helms, C. N., Hennon, C. C., Holmes, C. D., Huffman, G. J., Kossin, J. P., Lee, H.-T., Loew, A., and Magnusdottir, G.: Globally Gridded Satellite Observations for Climate Studies, Bull. Am. Meteorol. Soc., 92, 893–907, https://doi.org/10.1175/2011BAMS3039.1, 2011.

Kocaman, S., Debaecker, V., Bas, S., Saunier, S., Garcia, K., and Just, D.: A comprehensive geometric quality assessment approach for MSG SEVIRI imagery, Adv. Space Res., 69, 1462–1480, https://doi.org/10.1016/j.asr.2021.11.018, 2022.

Letu, H., Ishimoto, H., Riedi, J., Nakajima, T. Y., C.-Labonnote, L., Baran, A. J., Nagao, T. M., and Sekiguchi, M.: Investigation of ice particle habits to be used for ice cloud remote sensing for the GCOM-C satellite mission, Atmos. Chem. Phys., 16, 12287–12303, https://doi.org/10.5194/acp-16-12287-2016, 2016.

Letu, H., Yang, K., Nakajima, T., Ishimoto, H., Nagao, T., Riedi, J., Baran, A., Ma, R., Wang, T., Khatri, P., Chen, L., Shi, C., and Shi, J.: High-resolution retrieval of cloud microphysical properties and surface solar radiation using Himawari-8/AHI next-generation geostationary satellite, Remote Sens. Environ., 239, 1–16, https://doi.org/10.1016/j.rse.2019.111583, 2020a.

Letu, H., Yang, K., Nakajima, T. Y., Ishimoto, H., Nagao, T. M., Riedi, J., Baran, A. J., Ma, R., Wang, T., Shang, H., Khatri, P., Chen, L., Shi, C., and Shi, J.: High-resolution retrieval of cloud microphysical properties and surface solar radiation using Himawari-8/AHI next-generation geostationary satellite, Remote Sens. Environ., 239, 111583, https://doi.org/10.1016/j.rse.2019.111583, 2020b.

Letu, H., Ma, R., Nakajima, T. Y., Shi, C., Hashimoto, M., Nagao, T. M., Baran, A. J., Nakajima, T., Xu, J., Wang, T., Tana, G., Bilige, S., Shang, H., Chen, L., Ji, D., Lei, Y., Wei, L., Zhang, P., Li, J., Li, L., Zheng, Y., Khatri, P., and Shi, J.: Surface Solar Radiation Compositions Observed from Himawari-8/9 and Fengyun-4 Series, Bull. Am. Meteorol. Soc., 104, E1772–E1789, https://doi.org/10.1175/BAMS-D-22-0154.1, 2023.

Li, J., Pan, B., Zhang, F., Guo, B., Li, W., Jiang, G.-M., Wu, X., and Wang, Q.: Probabilistic Retrieval of All-Day Overlapping Cloud Microphysical Properties, Adv. Atmos. Sci., https://doi.org/10.1007/s00376-025-5234-7, 2025.

Li, J., Zhang, F., Li, W., Tong, X., Pan, B., Li, J., Lin, H., Letu, H., and Mustafa, F.: Transfer-Learning-Based Approach to Retrieve the Cloud Properties Using Diverse Remote Sensing Datasets, IEEE Trans. Geosci. Remote Sens., 61, 1–10, https://doi.org/10.1109/TGRS.2023.3318374, 2023.

Liu, C., Zhang, F., Ouyang, H., Li, W., and Zhao, Z.: Diurnal Variation of Cloud Physical Properties for Tropical Cyclones Over North Atlantic in 2019–2023, Geophys. Res. Lett., 52, e2025GL115566, https://doi.org/10.1029/2025GL115566, 2025.

Liu, W. H., Yue, P., Wu, X. H., Li, J. J., Shao, N. F., Zhu, B., and Lu, C. S.: Why does a decrease in cloud amount increase terrestrial evapotranspiration in a monsoon transition zone?, Environ. Res. Lett., 19, https://doi.org/10.1088/1748-9326/ad3569, 2024.

Menzel, W. P., Frey, R. A., Zhang, H., Wylie, D. P., Moeller, C. C., Holz, R. E., Maddux, B., Baum, B. A., Strabala, K. I., and Gumley, L. E.: MODIS Global Cloud-Top Pressure and Amount Estimation: Algorithm Description and Results, J. Appl. Meteorol. Climatol., 47, 1175–1198, https://doi.org/10.1175/2007JAMC1705.1, 2008.

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 Sens. Environ., 239, 111616, https://doi.org/10.1016/j.rse.2019.111616, 2020.

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, J. Meteorol. Res., 31, 708–719, https://doi.org/10.1007/s13351-017-6161-z, 2017.

Minnis, P. and Heck, P.: GOES-R Advanced Baseline Imager (ABI) Algorithm Theoretical Basis Document for Nighttime Cloud Optical Depth, Cloud Particle Size, Cloud Ice Water Path, and Cloud Liquid Water Path (Version 3.0), 85 pp., https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Baseline/ATBD_GOES-R_Cloud_NCOMP_v3.0_Jul2012.pdf, 2012.

Mitra, A., Di Girolamo, L., Hong, Y., Zhan, Y., and Mueller, K. J.: Assessment and Error Analysis of Terra-MODIS and MISR Cloud-Top Heights Through Comparison With ISS-CATS Lidar, J. Geophys. Res. Atmos., 126, e2020JD034281, https://doi.org/10.1029/2020JD034281, 2021.

Mouri, K., Izumi, T., Suzue, H., and Yoshida, R.: Algorithm Theoretical Basis Document for Cloud Type/Phase Product, Meteorological Satellite Center Technical Note, 19–31, https://www.data.jma.go.jp/mscweb/technotes/msctechrep61-2.pdf, 2016.

Nakajima, T. and King, M. D.: Determination of the Optical Thickness and Effective Particle Radius of Clouds from Reflected Solar Radiation Measurements. Part I: Theory, J. Atmos. Sci., 47, 1878–1893, https://doi.org/10.1175/1520-0469(1990)047<1878:DOTOTA>2.0.CO;2, 1990.

Nakajima, T. Y. and Nakajma, T.: Wide-Area Determination of Cloud Microphysical Properties from NOAA AVHRR Measurements for FIRE and ASTEX Regions, J. Atmos. Sci., 52, 4043–4059, https://doi.org/10.1175/1520-0469(1995)052<4043:WADOCM>2.0.CO;2, 1995.

Pavolonis, M. J.: GOES-R Advanced Baseline Imager (ABI) Algorithm Theoretical Basis Document for Cloud Type and Cloud Phase (Version 2.0). 85 pp., https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Baseline/ATBD_GOES-R_Cloud_Phase_Type_v2.0_Sep2010.pdf, 2010.

Platnick, S., Ackerman, S. A., King, M. D., Meyer, K., Menzel, W. P., Holz, R. E., Baum, B. A., and Yang, P.: MODIS atmosphere L2 cloud product (06_L2), NASA MODIS Adaptive Processing System, Goddard Space Flight Center [data set], Terra (MOD06_L2): 10.5067/MODIS/MOD06_L2.006, 2015.

Rossow, W. B. and Schiffer, R. A.: ISCCP Cloud Data Products, Bull. Am. Meteorol. Soc., 72, 2–20, https://doi.org/10.1175/1520-0477(1991)072<0002:ICDP>2.0.CO;2, 1991.

Rossow, W. B., Knapp, K. R., and Young, A. H.: International Satellite Cloud Climatology Project: Extending the Record, J. Climate, 35, 141–158, https://doi.org/10.1175/JCLI-D-21-0157.1, 2022.

Rossow, W. B., Walker, A. W., Beuschel, D. E., and Roiter, M. D.: International Satellite Cloud Climatology Project (ISCCP) Documentation of New Cloud Datasets, World Meteorol. Organ., 1996.

Rossow, W. B., Mosher, F., Kinsella, E., Arking, A., Desbois, M., Harrison, E., Minnis, P., Ruprecht, E., Seze, G., Simmer, C., and Smith, E.: ISCCP Cloud Algorithm Intercomparison, J. Appl. Meteorol. Climatol., 24, 877–903, https://doi.org/10.1175/1520-0450(1985)024<0887:ICAI>2.0.CO;2, 1985.

Schiffer, R. A. and Rossow, W. B.: The International Satellite Cloud Climatology Project (ISCCP): The First Project of the World Climate Research Programme, Bull. Am. Meteorol. Soc., 64, 779–784, https://doi.org/10.1175/1520-0477-64.7.779, 1983.

Shi, C., Letu, H., Nakajima, T. Y., Nakajima, T., Wei, L., Xu, R., Lu, F., Riedi, J., Ichii, K., Zeng, J., Shang, H., Ma, R., Yin, S., Shi, J., Baran, A. J., Xu, J., Li, A., Tana, G., Wang, W., Na, Q., Sun, Q., Yang, W., Chen, L., and Shi, G.: Near-global monitoring of surface solar radiation through the construction of a geostationary satellite network observation system, The Innovation, 6, 100876, https://doi.org/10.1016/j.xinn.2025.100876, 2025.

Tang, C., Shi, C., Letu, H., Yin, S., Nakajima, T., Sekiguchi, M., Xu, J., Zhao, M., Ma, R., and Wang, W.: Development of a hybrid algorithm for the simultaneous retrieval of aerosol optical thickness and fine-mode fraction from multispectral satellite observation combining radiative transfer and transfer learning approaches, Remote Sens. Environ., 319, 114619, https://doi.org/10.1016/j.rse.2025.114619, 2025.

Tong, X., Li, J., Zhang, F., Li, W., Pan, B., Li, J., and Letu, H.: The Deep-Learning-Based Fast Efficient Nighttime Retrieval of Thermodynamic Phase From Himawari-8 AHI Measurements, Geophys. Res. Lett., 50, e2022GL100901, https://doi.org/10.1029/2022GL100901, 2023.

Trepte, Q. Z., Minnis, P., Sun-Mack, S., Yost, C. R., Chen, Y., Jin, Z., Hong, G., Chang, F. L., Smith, W. L., Bedka, K. M., and Chee, T. L.: Global Cloud Detection for CERES Edition 4 Using Terra and Aqua MODIS Data, IEEE Trans. Geosci. Remote Sens., 57, 9410–9449, https://doi.org/10.1109/TGRS.2019.2926620, 2019.

Viggiano, M., Cimini, D., De Natale, M. P., Di Paola, F., Gallucci, D., Larosa, S., Marro, D., Nilo, S. T., and Romano, F.: Combining Passive Infrared and Microwave Satellite Observations to Investigate Cloud Microphysical Properties: A Review, Remote Sens., 17, 337, 2025.

Walther, A. and Heidinger, A. K.: Implementation of the Daytime Cloud Optical and Microphysical Properties Algorithm (DCOMP) in PATMOS-x, J. Appl. Meteorol. Climatol., 51, 1371–1390, https://doi.org/10.1175/JAMC-D-11-0108.1, 2012.

Walther, A., Straka, W., and Heidinger, A. K.: GOES-R Advanced Baseline Imager (ABI) Algorithm Theoretical Basis Document for Daytime Cloud Optical and Microphysical Properties (DCOMP) (Version 3.0), 66 pp., https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Baseline/ATBD_GOES-R_Cloud_DCOMP_v3.0_Jun2013.pdf, 2013.

Wang, M., Su, J., Han, X., Deng, X., Peng, N., and Liu, L.: Changes in Daytime Cirrus Properties From the ISCCP-H Data Set and Their Impacts on the Radiation Energy Budget, Earth Space Sci., 11, e2023EA003352, https://doi.org/10.1029/2023EA003352, 2024.

Winker, D., Pelon, J., Coakley, J., Ackerman, S., Charlson, R., Colarco, P., Flamant, P., Fu, Q., Hoff, R., Kittaka, C., Kubar, T., Treut, H., McCormick, M., Mégie, G., Poole, L., Trepte, C., Vaughan, M., and Wielicki, B.: The CALIPSO Mission: A

Global 3D View of Aerosols and Clouds, Bull. Am. Meteorol. Soc., 91, 1211–1230, https://doi.org/10.1175/2010BAMS3009.1, 2010.

Xiao, H., Zhang, F., Wang, L., Pan, B., Zhu, Y., Wang, M., Li, W., Guo, B., and Li, J.: High-resolution ensemble retrieval of cloud properties for all-day based on geostationary satellite, npj Clim. Atmos. Sci., 8, 386, https://doi.org/10.1038/s41612-025-01263-x, 2025.

Yost, C. R., Minnis, P., Sun-Mack, S., Chen, Y., and Smith, W. L.: CERES MODIS Cloud Product Retrievals for Edition 4—Part II: Comparisons to CloudSat and CALIPSO, IEEE Trans. Geosci. Remote Sens., 59, 3695–3724, https://doi.org/10.1109/TGRS.2020.3015155, 2021.

Young, A. H., Knapp, K. R., Inamdar, A., Hankins, W., and Rossow, W. B.: The International Satellite Cloud Climatology Project H-Series climate data record product, Earth Syst. Sci. Data, 10, 583–593, https://doi.org/10.5194/essd-10-583-2018, 2018.Zhang, H., Zhao, C., Li, J., Yu, Y., Chen, A., Yang, Y., Xia, Y., Yang, J., Sun, Y., Chi, Y., Zhao, X., Wang, Y., and Huang, J.: Shortwave cloud warming effect observed over highly reflective Greenland, Sci. Bull., 70, 951–959, https://doi.org/10.1016/j.scib.2025.01.027, 2025a.

Zhang, Q., Liu, C., Zhang, R., Yang, J., Li, J., Li, B., Lu, F., Yin, Y., and Jin, H.: Assessment of Geostationary Satellite Cloud Effective Radius over East China with In Situ Aircraft Measurements, J. Meteorol. Res., 39, 240–251, https://doi.org/10.1007/s13351-025-4164-8, 2025b.

Zhang, Z., Dong, X., Xi, B., Song, H., Ma, P.-L., Ghan, S. J., Platnick, S., and Minnis, P.: Intercomparisons of marine boundary layer cloud properties from the ARM CAP-MBL campaign and two MODIS cloud products, J. Geophys. Res. Atmos., 122, 2351–2365, https://doi.org/10.1002/2016JD025763, 2017.

Zhao, L.: IRBT2CPP: A training and testing dataset for machine learning training, ranging from infrared brightness temperature to cloud physical properties, Science Data Bank [data set], doi:10.57760/sciencedb.27171, 2025.

Zhao, L., Zhang, F., Li, J., Lu, F., and Zhao, Z.: DaYu-GCP: DaYu all-day global cloud physical properties products with 0.07° resolution retrieved from geostationary satellite imagers covering the period from 2000 to 2022, Science Data Bank [data set], doi:10.57760/sciencedb.26292, 2026.

Zhao, Z., Zhang, F., Li, W., and Li, J.: Image-Based Retrieval of All-Day Cloud Physical Parameters for FY4A/AGRI and Its Application Over the Tibetan Plateau, J. Geophys. Res. Atmos., 129, e2024JD041032, https://doi.org/10.1029/2024JD041032, 2024a.

Zhao, Z., Zhang, F., Wu, Q., Li, Z., Tong, X., and Li, J.: Cloud Identification and Properties Retrieval of the Fengyun-4A Satellite Using a ResUnet Model, IEEE Trans. Geosci. Remote Sens., 61, 1–18, https://doi.org/10.1109/TGRS.2023.3252023, 2023.

Zhao, Z., Zhang, F., Li, W., Meteorology, J. L. C.-F. J. L. o. M., Atmospheric, D. o., Sciences, O., Sciences, I. o. S., University, F., China, Waves, K. L. f. I. S. o. E., Education, M. o., Science, S. o. M., and Technology: Image-Based Retrieval of All-Day Cloud Physical Parameters for FY4A/AGRI and Its Application Over the Tibetan Plateau, J. Geophys. Res. Atmos., 129, 2024b.