Effectiveness of Cirrus Detection with MODIS Cloud Mask data

Authors:

Affiliation:

Corresponding author:

Key points:

- The utility of MODIS Cloud Mask products for creating a cirrus mask by validating them against CALIOP data was assessed
- To assess data accuracy, we computed the probability of detection, false alarm rate, overall accuracy, and Cohen's kappa statistic to evaluate nine tests for cirrus detection
- The most effective method for detecting high-level clouds among the available options was self-developed ATC test

Abstract

All clouds influence the Earth's radiative budget, with their net radiative forcing being negative. However, high-level clouds warrant special attention due to their atmospheric warming effects. A comprehensive characterization of cirrus clouds requires information on their coverage, which can be obtained from various data types. ~~. A comprehensive characterization of cirrus requires information on cloud coverage, obtainable from various data types.~~ Active satellite sensors are presently the most accurate source for cirrus data, but their usefulness in climatological studies is limited (the narrow view and 16-day repeat cycle yield only ~20 observations per year per region, often insufficient for climatological studies). On the contrary, passive data, which has been available for the past 40 years with sufficient temporal resolution for climatological research, are less effective at detecting cirrus clouds compared to active vertical profiling sensors.~~were not specifically designed for cirrus detection.~~ In this study, we assessed the utility of MODIS standard products for creating a cirrus mask by validating them against CALIOP data. Our objective was to determine if a MODIS product exists that detects cirrus with the same accuracy as CALIOP.

Using CALIOP data as the reference, we evaluated six tests for cirrus detection considered in MODIS cloud masking algorithm and their combination (ALL TESTS CONSOLIDATION, ATC). Additionally we applied two ISCCP-originating tests: ISCCP3.6 and ISCCP23 tests. All tests have been applied to MODIS radiances.

Study revealed that the ATC test was the most effective resulting with the overall accuracy of 72.98% during daytime and 59.50% at night (probability of detection: 80.87% and 25.46%, false alarm rate of 34.86% and 6.90%, and Cohen's kappa~~kapppa~~ coefficient of 0.46 and 0.19 respectively). However, its effectiveness was notably reduced during nighttime compared to daytime. We conclude that the ATC test is suitable for creating a mask of high-level clouds.

~~**Plain Language Summary**~~

1. Introduction

Clouds are indispensable to Earth's environmental systems and human life, influencing weather, climate, water distribution, ecosystems, and various human activities. All of them affect the Earth's radiative budget, and their net radiative forcing for is negative and equal to -13 Wm^{-2} (Ramanathan et al., 1989). That means that clouds, in general, cools the atmosphere. Nevertheless a special attention should be paid to high-level clouds (according to WMO, high-level clouds include all types of Cirrus, Cirrocumulus, and Cirrostratus clouds. Additionally, clouds resulting from anthropogenic activities, such as aviation contrails, are classified within the high-level cloud category (WMO, 1977)) named with the customary term *cirrus*. Cirrus clouds have a complex role in climate regulation. The relation between cirrus particles (size, shape and albedo) and Earth's radiation budget has been examined (Kinne and Liou, 1989; Macke et al., 1998; Mishchenko et al., 1996; Stephens et al., 1990; Zhang et al., 1994, 1999), resulting in a general conclusion that cirrus play an important role and can warm the atmosphere. Cirrus typically have a base above about 8 000 metres and are composed of small particles – ice crystals. Because of cirrus specific properties (cloud height, temperature, effective particle size, surface thermal contrast, ice water path and cloud optical depth; Ackerman et al., 1988; Stephens et al., 1990; Stephens & Webster, 1981)), in contrast to low- and mid-level clouds, they heat the Earth (they allows shortwave radiation to reach Earth's surface and reduces outgoing longwave radiation). Recent research shows that cirrus radiative forcing varies from about 0.05 Wm^{-2} for contrails, to 35.5 Wm^{-2} for cirrus in general (Bock and Burkhardt, 2016; Campbell et al., 2016; Kärcher, 2018; Lolli et al., 2017; Oreopoulos et al., 2017). Additionally, their presence change the radiative forcing of other clouds for positive as well. For instance, when medium and low clouds co-occur, their radiative effect equals -18.8 Wm^{-2}. Additional presence of cirrus raises the radiative effect to 50,8 Wm^{-2} (Oreopoulos et al., 2017).

Cirrus properties description is incomplete without the information about cloud coverage. Most of the studies, have considered just a total cloud cover, but some of them also study high-level cloudiness. The global frequency of cirrus occurrence is between 28 and 42% . Research conducted using high resolution satellite data has shown that global cloud coverage is estimated at about 66% to 74% and 40% of all clouds are high-level clouds (Stubenrauch et al., 2010). According to Sassen et al. (2008) cirrus cover almost 17% of Earth's surface. The study of high-level cloud coverage and its trends has long intrigued scientists. In 1994, Wylie et al. presented global statistics on cirrus clouds over a four-year period, revealing an average cirrus coverage of 42% based on HIRS data. More recently, Li and

Groß (2022) analyzed a decade of CALIPSO lidar measurements, finding that cirrus clouds over Europe occur most frequently between 9 to 11 km altitude, with occurrence rates varying seasonally from about 5% in summer to 12% in winter. Another significant study used 16 years of ISCCP data to identify trends in cirrus clouds across Europe, noting an increase of 1-2% per decade in regions with high aircraft traffic, contrasting with a general decline elsewhere (Stordal et al., 2005). A 16-year ground-based lidar study in Gadanki, India, observed peak cirrus occurrence at 14.5 km with a 25% frequency (Pandit et al., 2015). The most extended study, spanning 20 years (1983-2004) with ISCCP data, documented high cirrus concentrations in regions such as the South Pacific Convergence Zone, the Amazon, and central Africa, while noting a global decrease in cirrus clouds except in the southern mid-latitudes, where no significant trend was observed (Eleftheratos et al., 2007). Numerous studies have explored changes in high-level cloud coverage. However, those relying on satellite data often lack a focus on cirrus clouds over sufficiently long periods—at least 30 years, as recommended by the WMO. Conducting such studies and identifying suitable data sources pose significant challenges.

Given the critical role of cloud cover, especially cirrus clouds, in atmospheric studies, the observation of clouds is considerably significant. Historically first method is visual observation from ground-based meteorological stations, which is simple and provides long time series data. However, this method has limitations, including difficulty in detecting high-level clouds due to cloud overlap at multiple altitudes, perspective issues near the horizon, and the optical thinness of cirrus clouds. Studies have shown that under optimal conditions, the probability of detecting cirrus clouds visually ranges from 44% to 83% during the day and 24% to 42% at night. With clouds at all levels, detection probabilities drop to 47%–71% during the day and 28%–43% at night (Kotarba & Nguyen Huu, 2022).

Present cloud climatologies benefit from satellite remote sensing. Initially, this information was obtained from various imagers, sounders, and radiometers, which utilize passive cloud detection methods (involving detecting natural radiation emitted or reflected by objects, such as clouds, without actively sending out signals). Researchers such as Ackerman et al. (2008); Amato et al. (2008); Chen et al. (2002); Frey et al. (2008, 2020); Gu et al. (2011); Kotarba (2016); Y. Liu et al. (2004); Minnis et al. (2008); Murino et al. (2014); Musial et al. (2014); Tang et al. (2013) have contributed to these studies. An example of passive utensil can be MODIS (Moderate Resolution Imaging Spectroradiometer), which is a key instrument aboard the Terra and Aqua satellites.

Active remote sensing technology relies on its own signal, directing it at an object and analyzing the response. This allows active sensors, in example CALIPSO's (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations) lidar, CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization), to operate day and night with similar efficiency. Active profiling instruments like CALIOP, which provide high-resolution vertical profiles of aerosols and clouds, have limitations such as a narrow field of view.

~~This narrow view, combined with a long 16-day repeat cycle, results in only about 20 observations per year of the same region, which is challenging and sometimes insufficient for climatological studies. Although active sensors, like CALIOP, are currently the best source of cirrus data (Heidinger and Pavolonis 2009), their potential for construing long-term climatologies is very limited. On the contrary, passive data are available for last more than 40 years with good enough for climatological research time resolution (i.e. for MODIS we have access to over 20 years of data), although they were not designed for cirrus detection. In this paper, we examined utility of MODIS products to create a cirrus mask by validating them with CALIOP data. Our objective is to determine whether any existing operational MODIS product detect cirrus clouds as accurately as the CALIPSO does. Specifically, we aim to assess whether MODIS Cloud Mask, when examining its individual tests could be easily adapted into an algorithm for masking cirrus clouds. We also seek to identify the conditions under which this approach would be effective and when it might not be suitable.~~

Clouds are indispensable to Earth's environmental systems and human life, influencing weather, climate, water distribution, ecosystems, and various human activities. They affect the Earth's radiation budget, with a net radiative forcing of approximately -20 Wm^{-2} (Boucher et al., 2013), which results in an overall cooling effect on the planet. Nevertheless, special attention should be paid to high-level clouds - according to the WMO, high-level clouds include Cirrus, Cirrocumulus, and Cirrostratus (WMO, 1977)(WMO, 1977) - commonly referred to as *cirrus*. Those clouds play a complex role in climate regulation. The relation between cirrus particles (size, shape and albedo) and Earth's radiation budget has been examined (Kinne and Liou, 1989; Macke et al., 1998; Mishchenko et al., 1996; Stephens et al., 1990; Zhang et al., 1994, 1999), resulting in a general conclusion that cirrus play an important role and can warm the atmosphere. They typically have a base above about 8,000 m and consist of small ice crystals.Those clouds play a complex role in climate regulation. The relation between cirrus particles (size, shape and albedo) and Earth's radiation budget has been examined (Kinne and Liou, 1989; Macke et al., 1998; Mishchenko et al., 1996; Stephens et al., 1990; Zhang et al., 1999, 1994), resulting in a general conclusion that cirrus play an important role and can warm the atmosphere. They typically have a base above about 8,000 meters and consist of small ice crystals. Due to their unique properties - such as altitude, temperature, effective particle size, surface thermal contrast, ice water path, and optical depth (Ackerman et al., 1988; Stephens et al., 1990; Stephens & Webster, 1981), they differ from low- and mid-level clouds in their effect on the Earth's radiation budget. Specifically, cirrus clouds allow shortwave radiation to reach the surface while reducing outgoing longwave radiation, thereby contributing to warming. Recent research estimates that cirrus radiative forcing of cirrus globally to approach 35.5 Wm^{-2} for cirrus globally (Campbell et al., 2016; Kärcher, 2018; Lolli et al., 2017; Oreopoulos et al., 2017). Furthermore, cirrus clouds can alter the radiative forcing of other cloud

types. For example, when medium and low clouds co-occur, their combined radiative effect is -18.8 Wm^{-2}, but the additional presence of cirrus raises this effect to 50.8 Wm^{-2} (Oreopoulos et al., 2017).

A description of cirrus cloud properties is incomplete without information about their coverage. Most studies have focused on total cloud cover, but some have also examined high-level cloudiness. The global frequency of cirrus occurrence is estimated to range between 17% and 42%. Research conducted using high-resolution satellite data indicates that global cloud coverage is approximately 66% to 74%, with 40% of all clouds classified as high-level clouds (Sassen et al., 2008; Stubenrauch et al., 2010). Numerous studies have explored changes in high-level cloud coverage. However, those relying on satellite data often do not address cirrus clouds over sufficiently long periods—at least 30 years, as recommended by the WMO. Conducting such long-term studies and identifying suitable data sources remain significant challenges.

Given the critical role of cloud cover, especially cirrus~~, observing~~ clouds, in atmospheric studies, observing clouds is of considerable importance. Historically first method is visual observation from ground-based meteorological stations, which is simple and provides long time series data. However, this method has limitations, including difficulty in detecting high-level clouds due overlapping clouds at multiple altitudes, perspective distortions near the horizon, and the optical thinness of cirrus clouds. Studies have shown that under optimal conditions, the probability of visually detecting cirrus clouds ranges from 44% to 83% during the day and from 24% to 42% at night. When clouds at all levels are present, detection probabilities drop to 47%–71% during the day and 28%–43% at night (Kotarba & Nguyen Huu, 2022).

Modern cloud climatologies benefit from satellite remote sensing. Initially, this information was obtained from various imagers, sounders, and radiometers, which utilize passive cloud detection methods (involving detecting natural radiation emitted or reflected by objects, such as clouds, without actively sending out signals). Researchers such as Ackerman et al. (2008); Amato et al. (2008); Chen et al. (2002); Frey et al. (2008, 2020); Gu et al. (2011); Kotarba (2016); Y. Liu et al. (2004); Minnis et al. (2008); Murino et al. (2014); Musial et al. (2014); Tang et al. (2013) have contributed to these studies. An example of passive sensor can be MODIS (Moderate Resolution Imaging Spectroradiometer), which is a key instrument aboard the Terra and Aqua satellites.Modern cloud climatologies benefit from satellite remote sensing. Initially, this information was obtained from various imagers, sounders, and radiometers, which utilize passive cloud detection methods (involving detecting natural radiation emitted or reflected by objects, such as clouds, without actively sending out signals). Researchers such as Ackerman et al. (2008); Amato et al. (2008); Chen et al. (2002); Frey et al. (2008, 2020); Gu et al. (2011); Kotarba (2016); Y. Liu et al. (2004); Minnis et al. (2008); Murino et al. (2014); Musial et al. (2014); Tang et al. (2013) have contributed to these studies. An example of passive utensil can be

MODIS (Moderate Resolution Imaging Spectroradiometer), which is a key instrument aboard the Terra and Aqua satellites.

Active remote sensing technology, in contrast, relies on its own signal, directing it at an object and analysing the response. This allows active sensors, for instancein example CALIPSO's (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations) lidar, CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization), to operate day and night with similar efficiency in cloud detection.. Active profiling instruments like CALIOP, which provide high-resolution vertical profiles of aerosols and clouds, have limitations, including a narrow field of view. This narrow view, combined with a long 16-day repeat cycle, results in only about 20 observations per year of the same region, which is challenging and sometimes insufficient for climatological studies (Kotarba, 2022)..

To standardize cloud classification and ensure consistency, the International Satellite Cloud Climatology Project (ISCCP) developed a system based on cloud height and optical thickness, providing a systematic framework for studying cloud types and their variability across regions and over time. This classification is crucial for advancing climate modelling, weather forecasting, and research on cloud-climate interactions. The ISCCP classification was applied to MODIS data, and its effectiveness in detecting cirrus clouds was also evaluated.

While active sensors like CALIOP remain the most reliable source of cirrus data (Heidinger and Pavolonis, 2009),(Heidinger and Pavolonis 2009), , their potential for building long-term climatologies is limited. In contrast, passive data have been available for over 40 years, offering temporal coverage suitable for climatological research. OneFor example of such sensors, MODIS has provided more than 20 years of data, although collecting data for over 20 years rather than 40, is MODIS, whoseits capabilities for detecting cirrus clouds are limited compared to those of active vertical profiling sensors.

In this paper, we use cirrus characterizations from CALIOP data to exploreevaluate the potential for creating a cirrus mask from theusing operational MODIS cloud data products. Our objective is to determine how well the MODIS products can be used to identify cirrus clouds compared to CALIPSO. Specifically, we aim to assess whether MODIS cloud detectionCloud Mask, by analysing its individual tests used to generate MYD35 operational data can, could be re-used for a time-adapted into an algorithm for cirrus masking. We also seek to identify the conditions under which this approach would be effective masking of cirrusand when it might not be suitable.

2. Data and methods

In this study, we use active sensor data for validating passive-based information for determining the presence of cirrus (for the sake of clarity, throughout this manuscript, all high-level clouds will be called as cirrus). The active ~~dataset~~ sensor data was collected by the CALIOP lidar aboard the CALIPSO

satellite, while the passive data was obtained from the MODIS ~~spectroradiometer~~ multi-band radiometer on the Aqua satellite. The concept behind achieving the research objective was based on collocation of those two datasets in time and space. In both instances, cirrus clouds are the same physical phenomenon; however, the distinction arises from the varying sensitivities of the detection instruments employed, with optical thickness serving as a crucial parameter. CALIPSO is capable of identifying cirrus clouds with an optical thickness as low as approximately 0.01, while MODIS generally detects them when the optical thickness is in the range of 0.4 to 0.5 (Menzel et al., 2015). Data for the year 2015 were analyzed on a global scale, comprising 136,272,209 combined observations from the aforementioned satellites. The primary requirement was to obtain a sufficiently large sample of CALIPSO-MODIS match-ups across different seasons and geographic regions, which necessitated one complete year of global observations. Therefore, 2015 was chosen arbitrarily.~~were analyzed for the whole globe. These include 136,272,209 combined observations from the aforementioned satellites.~~

2.1. MODIS data

MODIS, an advanced instrument aboard NASA's Terra and Aqua satellites, acquires data across 36 spectral bands, spanning wavelengths from visible to thermal infrared (0.4 to 14.4 µm). Its passive sensors rely primarily on naturally available energy: solar energy reflected from objects or absorbed and re-emitted ~~The MODIS, an advanced instrument aboard NASA's Terra and Aqua satellites, acquires data across 36 spectral bands, encompassing wavelengths from visible to thermal infrared (0.4 to 14.4 µm). Its passive sensors relies mostly on naturally available energy: Sun's energy reflected from the object or absorbed and reemitted~~ (Ackerman et al., 1998). MODIS provides data at various spatial resolutions - —250 mmeters, 500 mmeters, and 1 km - kilometer—with a swath width of 2,330 kmkilometers, enabling it to observe the entire Earth twice daily, every one observation during the day andto two days. ~~It provides data at various spatial resolutions—250 meters, 500 meters, and 1 kilometer with swath width of 2,330 kilometers which observes the entire Earth every one~~ at night. ~~to two days.~~ Cloud detection results are stored in the 48-bit "Cloud Mask" product, known as MYD35 for Aqua, while corresponding cloud properties can be found in MYD06 dataset. As an imager, MODIS provides column-integrated radiances, ~~what~~ which limits its ability to retrieve cirrus-specific information. ~~the possibilities for cirrus retrieval.~~

For this research, we used Collection 061 data, which is available in 5-minute granules at a spatial resolution of 1 km per pixel (at nadir). ~~For this research, we assessed the version of Collection 061 data, which is available in 5-minute granules at a spatial resolution of 1 km per pixel (at nadir).~~ Each MYD35 and MYD06 file is paired with a MYD03 "Geolocation file" product that contains longitude and latitude information for each individual cloud mask IFOV (instantaneous field~~Instantaneous Field~~ of view;~~View,~~ Guenther et al., 2002).

2.1.1. The MODIS Cloud Mask product

The MODIS Cloud Mask product is a Level 2 dataset produced at spatial resolutions of 1 km and 250 m (at nadir). The cloud masking procedure was described in details by Ackerman et al. (1998), Frey et al. (2008), and Baum et al. (2012). The algorithm utilizes a sequence of visible and infrared threshold and consistency tests to determine the confidence level that an unobstructed view of the Earth's surface is achieved. ~~The MODIS Cloud Mask product is a Level 2 dataset produced at spatial resolutions of 1 km and 250 m (at nadir). The cloud masking procedure is detailed in the works of Ackerman et al. (1998), Frey et al. (2008), and Baum et al. (2012). The algorithm utilizes a sequence of visible and infrared threshold and consistency tests to determine the confidence level that an unobstructed view of the Earth's surface is achieved. The MYD35 dataset includes data from the Aqua satellite.~~

In this research, we ~~considered~~ analyzed 6 ready-to-use MODIS tests. Individual tests were described by Ackerman et al. (1998):

- Thin Cirrus test (SOLAR) – the solar channels in MODIS cover a range of wavelengths primarily in the visible and near-infrared spectrum (0.4 to 2.5 µm). ~~This test uses the solar range to set the confident clear and middle thresholds to define the range of expected reflectances from thin cirrus. It indicates that a thin cirrus cloud is likely to be present.~~ Test is only useful during daytime.
- Thin Cirrus test (IR) – the purpose of this test is detecting thin cirrus clouds. Channels used for this test are 11 µm an 12 µm (infrared (IR) range), incorporated to the split window technique.
- High Cloud Test (BT13.9) – applying CO_2 absorption channels (around 14 µm) is a simple technique got from the CO_2 slicing method (suitable for determining middle and upper troposphere ice clouds heights and effective amounts). This test is useful for high-level cloud detection, while it can reveal clouds above 500 hPa.
- High Cloud Test (BT6.7) – test designed for detecting thick high clouds. Starting from the ground level, the 6.7 µm radiation emitted by the surface or low clouds is absorbed in the atmosphere, therefore the signal is not received by an instrument. Water vapor in the atmospheric layer between 200 hPa and 500 hPa ~~The water vapor in layer in the atmosphere between 200 hPa and 500 hPa~~ is the only source of the 6.7 µm radiation in clear-sky observation. Thick clouds placed above or near the 200 hPa level can be distinguish from clear sky or lower clouds.

- High Cloud Test (BT1.38) – the 1.38 µm channel lies in the strong water vapor absorption region. That results in obscuration of the most of Earth's surfaces, as well as attenuation of reflectance from low- and mid-level clouds. Pixels subjected to this test reveal high-level thin clouds as brighter.~~Pixels with this test applied, reveals high level thin clouds as brighter.~~ Unfortunately, the test has certain limitations, including its applicability to nighttime conditions, polar regions, midlatitude winters, and high elevations..
- High Cloud Test (BT3.9-12.0) – the 3.9-12.0 µm BTD (Brightness Temperatures Difference) test is specifically designed for nighttime observations over land and polar snow/ice surfaces. It effectively distinguishes thin cirrus clouds from cloud-free conditions ~~It is effective in distinguishing between thin cirrus clouds~~ and cloud-free conditions and exhibits relative insensitivity to the atmospheric water vapor content (Hutchinson and Hardy, 1995)~~(Hutchinson and Hardy, 1995)~~.

Additionally, we independently developed a unified approach to combine all tests, which we termed **All Tests Consolidation (ATC)**. If any)~~unification of all tests, which we called **All Tests Consolidation (ATC)**. If any one~~ (∃ - there is at least one) of the nine tests (t) detected cirrus clouds, the output flag (OF) was set to indicate the presence of cirrus.~~we set the output flag (OF) to indicate cirrus.~~

$$\text{If } \exists_{i\in\{1,2,\ldots,9\}}\,(t_i=1), \text{ then } ATC_{OF}=1$$

Conversely, if no cirrus clouds were detected by ~~all~~ any of the tests (∀ - for every), provided they were all conducted, no cirrus flag was set.

$$\text{If } \forall_{i\in\{1,2,\ldots,9\}}\,(t_i=0), \text{ then } ATC_{OF}=0$$

ATC is essentially an adaptation of the MYD35~~MOD35~~ approach, but it is limited to tests that specifically provide insights ~~specifically~~ about cirrus clouds.

2.1.2. The MODIS Cloud Product

As described by Menzel et al. (2015) the MODIS Cloud Product uses a combination of infrared and visible techniques to determine cloud physical and radiative properties. It derives cloud-particle phase, effective particle radius, and optical thickness from visible and near-infrared radiances, and indicates cloud shadows. Infrared methods provide cloud-top temperature, height, effective emissivity, phase, and cloud fraction, both day and night, at 1-km-pixel resolution. Additionally, the product includes

cirrus reflectance at 1-km resolution, used to correct for cirrus scattering in land-surface reflectance. ~~to correct for cirrus scattering in land-surface reflectance.~~ For Aqua satellite, dataset is called MYD06.

In addition to the ready-to-use MODIS tests (Section 2.2.1), other criteria can be applied using data available from MODIS and CALIOP. For instance, the ISCCP's definition of cloud types. By examining visible and infrared radiances from geostationary and polar-orbiting meteorological satellites and making assumptions about cloud layering, thermodynamic phases, and properties, ISCCP characterizes a cloudy satellite pixel using the column visible optical depth (COT) and the cloud-top pressure (CTP) of the highest cloud layer. This information is used to classify different cloud types as shown in the figure 1 (Rossow and Schiffer, 1999). In addition to the ready-to-use MODIS tests (Section 2.2.1), other criteria can be applied using data available from MODIS and CALIOP. For instance, the ISCCP's ~~Beside ready-to-use MODIS tests (Section 2.2.1), other criteria may be applied using available for MODIS and CALIOP data, e.g. the ISCCP's (The International Satellite Cloud Climatology Project; https://isccp.giss.nasa.gov), which was established in 1982 as part of the World Climate Research Programme (WCRP; https://www.wcrp-climate.org/) to gather the global distribution of clouds, their properties, and their diurnal, seasonal, and interannual variations) definition of cloud types. The developers of ISCCP deserve significant recognition for their foresight, as more than fourty years later, ISCCP~~ remains a leading reference~~continues to be a leading reference for describing the cloudy atmosphere. By examining visible and infrared radiances from geostationary and polar-orbiting meteorological satellites and making assumptions about cloud layering, thermodynamic phases, and properties, ISCCP characterizes a cloudy satellite pixel using the column visible optical depth (COT) and the cloud-top pressure (CTP) of the highest cloud layer. This information~~ is used to classify ~~can be used to classify different cloud types as shown in the figure 1 (Rossow and Schiffer, 1999).~~

COT and CTP areare~~is~~ also available for MODIS, within MYD06 standard product, and we used ~~it~~ them to generate cirrus masks based~~according to~~based on the ISCCP definition. We considered two variants of the mask, defining cirrus as:

- a cloud with an optical thickness less than 3.6 and a top pressure below 440 hPa (hereinafter ISCCP3.6 test),

- a cloud with an optical thickness less than 23 and a top pressure below 440 hPa (hereinafter ISCCP23 test).

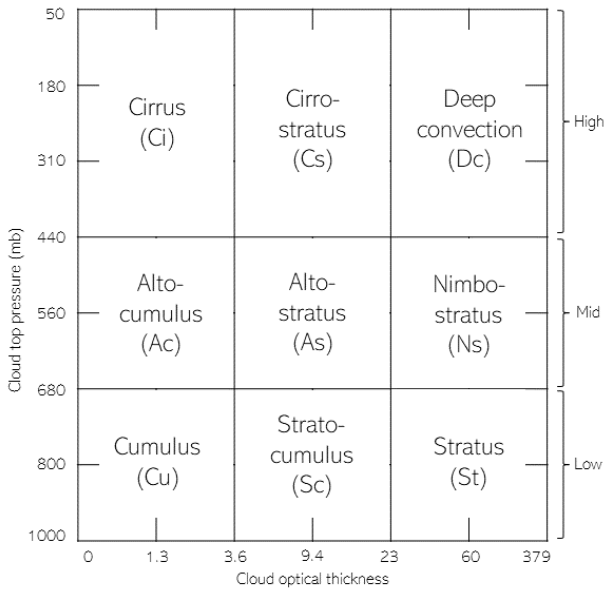


Fig. 1. Cloud-type definitions used in the ISCCP D-series

2.2. CALIOP data

CALIOP provides atmospheric profiles with vertical resolutions ranging from 30 m below 8.2 km to 180 m above 20.1 km, and 60 m between these altitudes (Winker et al., 2006). This capability allows for clear distinction between cirrus and lower cloud layers, making CALIOP excellent for cirrus detection. Furthermore, lidar can detect cirrus clouds with an optical depth as low as 0.01 (Vaughan et al., 2009), a capability beyond the reach of other imagers (Ackerman et al., 2008). Being an active sensor, lidar offers similar effectiveness in cloud detection both daytime and nighttime, or even higher during night, when backscattered light does not interfere with diffused solar radiation (McGill et al., 2007).~~Active sensors, for in example CALIOP, operate~~ both ~~day and night. Unlike passive methods, CALIOP's cloud detection accuracy is even higher at night than during the day (McGill et al., 2007). CALIOP provides high-resolution atmospheric profiles, with vertical resolutions ranging from 30 m below 8.2 km to 180 m above 20.1 km, and 60 m between these altitudes (M. Winker et al., 2006). This capability allows for clear distinction between cirrus and lower cloud layers, making CALIOP excellent for cirrus detection. Furthermore, lidar can detect cirrus clouds with an optical depth as low as 0.01 Additionally, lidar can detect cirrus clouds with an optical depth of 0.01 or less (Vaughan et al. 2009), a capability beyond the reach of other imagers (Ackerman et al. 2008; Hutchison et al. 2014).~~

In this research, *the lidar level-2 cloud layer at 5-km horizontal resolution, version 4.20* (CAL_LID_L2_05kmCLay-Standard-V4–20) product was used. As described by Liu et al. (2009)~~(2009)~~ and Vaughan et al. (2009) this product reports cloud layers and cloud type information, with cirrus as

a separate class (. There are seven primary categories, including clouds and aerosols, and within the cloud category, eight subtypes are defined (e.g., cirrus is categorized as type 6). There are seven categories including clouds and aerosols. Inside the cloud class, 8 subtypes can be found (i.e. cirrus).The detection of cirrus clouds is based on the analysis of the backscatter coefficient and the lidar signal's depolarization ratio, which differentiates ice particles, characteristic of cirrus clouds, from water droplets. The depolarization ratios for Cirrus clouds are higher than those for water-based clouds, enabling their identification. Additionally, CALIOP provides anprecise information about the cloud base and top altitudes, allowing for accurate determination of their position in the atmosphere. The quality of CALIOP's detection is reflected in described by the CAD (cloud-aerosol discrimination (CAD) score, which ranges from -100 to 100. Quality of CALIOP's detection is described by CAD (cloud-aerosol discrimination) score, which ranges from -100 to 100. Value -100 indicates high confidence of aerosol detection, while a value of 100 indicates high confidence in cloud detection. A medium value (0) signifies equal probability that the feature is a cloud or aerosol ; value 100 shows that cloud was detected with high confidence; medium value (0) means that there is the same probability that the feature is cloud or aerosol (Liu et al., 2009; Vaughan et al., 2009)(Liu et al., 2009; Vaughan et al., 2009). In this study, we usedonly useused only observations with a CAD score higher than 80. The optical depth is also provided in this (CAL_LID_L2_05kmCLay-Standard-V4–20) CALIOP product.

For the purpose of this research, we regard consider CALIPSO as the reference for cirrus clouds detection. This choice is based on the lidar's high sensitivity to optically thin clouds and its reliable performance in both daytime and nighttime conditions.This choice is driven by the lidar's high sensitivity to optically thin clouds and its reliable performance in both daytime and nighttime conditions.

2.3. Matching datasets

In order to achieve the goal of this study, MODISNASA and CALIOP data were collocatedits partners operate a group of Earth-observing satellites in space and time. It was possible because Aqua and CALIPSO operated for 12 years (2006-2018) as a part of satellite constellation commonlysun-synchronous polar orbits, known as the Afternoon Constellation. MembersThis constellation has changed over time as satellites have moved out of the constellation used sun-synchronous polar orbits of 16-day revisit cycle, and with equatorial crossing time at 13or have deorbited, but Aqua remained a key member while CALIPSO began to move out of it in 2018. Afternoon Constellation crosses the equator in a northbound direction around 1:30 PM local solar time (ascending node). CALIPSO followed Aqua spacecraft, providing near-simultaneous observations from multiple instruments. Aqua and CALIPSO, with nearly identical orbital configurations, operated in close proximity from 2006 to 2018, trailing by approximately one minute (Stephens et al., 2018), enabling quasi-simultaneous

~~synchronized~~ observation of the same part of the atmosphere, as 1 km~~times and a shared 16-day revisit cycle despite slightly differing~~ ground track of CALIOP always overlapped with 2330 km wide imagery of MODIS~~tracks~~.

Collocating MODIS with CALIOP has been frequently used to validate reliability of MODIS datasets, or to developed a new, joint imager-lidar atmospheric products (e.g. Baum et al., 2012; Holz et al., 2009; Kotarba, 2020; Sun-Mack et al., 2014; Wang et al., 2016; Xie et al., 2010). Either 333 m, 1 km, or 5 km lidar data may be considered, however only 1 km and 5 km products offers cloud type classification. Additionally, only 5 km product informs about cloud optical thickness per cloud layer, and provides superior cirrus detection due to higher sensitivity (noise level decreases as more profiles is integrated into retrieval).

From the geometry point of view, a 5 km profile is an aggregation of five consecutive 1 km profiles, and the geo-coordinates of the central one are saved as representative for 5 km profile. It possess a challenge when MODIS and~~Using CALIOP data for the calibration and validation of atmospheric products from various space missions is a well-established practice. This method has been extensively applied to Aqua MODIS (Baum et al., 2012; Holz et al., 2009; Kotarba, 2020; Sun-Mack et al., 2014; Wang et al., 2016; Xie et al., 2010).~~

~~For this study, Aqua MODIS data and corresponding~~ CALIOP are to be matched: one 5 km profile of CALIOP only can be accurately matched to one 1 km MODIS pixel, while 5 km data actually covers five MODIS pixels. To overcome this problem we matched CALIOP with MODIS using non-aggregated, 1 km data, and only then assigned 5 km data to already collocated MODIS-CALIOP pairs. As a result, one 5 km profile of CALIOP was used to characterize five MODIS pixels.

Aqua and CALIPSO ground tracks are offset by 100-120 km at the equator (decreasing towards the poles). It means, that they observe the atmosphere from slightly different angles, causing a parallax shift. We did not correct the data for parallax, as its impact only would be observed close to the edges of clouds, which are small fraction of all observations, or for investigating dynamically-changing cloud top properties (Wang et al., 2011) which was not the case of our investigation.

This study relied on MODIS-CALIOP observations for 2015, and the year was ~~were matched. The matching process involved selecting a MODIS IFOV and comparing it with the corresponding CALIOP profile, ensuring the geometric center fell within the~~ selected arbitrary, as the only requirements was to consider a relatively large (year-long) ~~MODIS IFOV. Due to the orbital configuration of the two missions, CALIOP could only~~ sample of global observations of clouds. Eventually, our database consisted of ~~MODIS IFOVs near the MODIS nadir because of nadir-pointing instrument, preventing matching observations across the entire MODIS swath. Despite the length of the period (1 year), the~~

~~procedure resulted in a sufficient number of observations (~~136,272,209 paired MODIS-CALIOP observations; the average spatial distance between geometrical centers of matched MODIS pixel and CALIOP profile was 444 m (std. dev. = 231 m), while the average temporal separation reached 84 seconds (std. dev. = 12 seconds).

The final,~~) as each MODIS granule contains approximately 2,030 IFOVs, and a full day of Aqua observations produces 288 granules. The~~ aggregated MODIS–CALIOP statistics were compiled into global maps, each with a spatial resolution of 5° in both longitude and latitude.

2.4. Evaluation of MODIS data

The comparison was conducted at the pixel level, using a confusion matrix as the basis for calculations. This~~It gives~~This approach provides a detailed comparison of the model's predictions against the actual results. For clarity, Table 1 provides an explanation of abbreviations related to statistical measures.

Table 1. Abbreviations

Abbreviation	Definition
TP	True Positives
FP	False Positives
TN	True Negatives
FN	False Negatives
ROP	Rate of Observations Performed
POD	Probability of Detection
FAR	False Alarm Rate
OA	Overall Accuracy
Kappakappa	Cohen's kappa k coefficient
PE	Expected agreement
n	Number of elements in the set

The structures~~Structure~~ of confusion matrix is presented in Table 2.~~–21~~. and includes the following elements:

- True Positives (TP): The ~~count~~ number of cases where MODIS accurately identified the existing (according to CALIOP) cirrus.
- False Positives (FP): The number ~~count~~ of cases where MODIS incorrectly identified the high-level cloud, meaning it detected cirrus presence when it was actually absent.
- True Negatives (TN): The number ~~count~~ of cases where MODIS correctly did not detect the presence of the cloud.

– False Negatives (FN): The number ~~count~~ of cases where MODIS overlooked the cirrus occurrence.

Table ~~1~~2. Confusion matrix

		CALIPSO (REFERENCE DATA)	
		Cirrus	No Cirrus
MODIS	Cirrus	True positive (TP)	False positive (FP)
	No Cirrus	False negative (FN)	True negative (TN)

Every result undergoes ~~a~~ thorough validation through different parameter~~parameters~~ estimation using feature-based statistics (Stanski et al., 1989). To describe the data accuracy, the probability of detection (POD) ~~characteristics~~ [1] and false alarm rate (FAR) ~~statistic~~ [2] metrics were calculated:

Probability of detection (POD) – is a metric used to assess the effectiveness of a detection system. In the context of cloud detection, POD indicates how well the detection algorithm correctly identifies the presence of clouds when they are actually present. A higher POD value signifies better performance of the detection system.

$$POD = TP/(TP+FN) \quad [1]$$

False alarm rate (FAR) – is a metric that measures the frequency of incorrect positive detections by a system. In the context of cloud detection, a lower FAR indicates a more accurate system, with fewer instances of falsely identifying clouds when they are not present.

$$FAR = FP/(FP+TN) \quad [2]$$

The incident frequencies within the matrix enabled the identification of two more diagnostic measures:

Overall accuracy (OA) – is a metric that measures the proportion of correct predictions made by a detection system out of all predictions. In cloud detection, higher overall accuracy indicates that the system effectively identifies both the presence and absence of clouds correctly.

$$OA = (TP+TN)/n \quad [3]$$

Cohen's kappa k – ~~Cohen's kappa~~ is a statistical metric used to assess the degree of agreement between two raters or classification methods. Its scale ranges from -1 to 1, where a value of 1 represents perfect agreement, 0 indicates agreement no better than chance, and negative values indicate agreement worse than chance. In cloud detection, a higher kappa value indicates stronger agreement between the detected presence of clouds and their actual presence, while considering the possibility of random agreement.

$$k = (OA - PE)/(1-PE) \quad [4]$$

where

PE – expected agreement

$$PE = [(TP + FP)(TP+FN) + (TN+FP)(TN+FN)]/n^2 \quad [5]$$

$$n = TP + FP + FN + TN \quad [6]$$

The accuracy of high-level cloud detection was evaluated using the aforementioned metrics, differentiated by day and night, latitude, cloud optical depth, the number of detected cloud layers, and land classification. This assessment was conducted for the entire year 2015, as well as specifically for January and July (those two months are presented to exemplify the characteristics of two distinct seasons).

2.5. Bootstrap sampling

Due to the nature of cirrus cloud occurrences (18.7% in 2015, see Section 3), we can assume that the data sample will be imbalanced and one class (without cirrus) significantly outnumbers the other. Therefore, for such data, the appropriate statistical method to apply is bootstrap sampling (Efron, 1980).

The balancing the sample stems from the issue of class imbalance, potentially skewing the statistical analysis and leading to biased results. To mitigate this, the bootstrap method is employed to artificially balance the dataset. This involves resampling the data with replacement, to ensure that each class has a comparable number of instances. By doing so, the analysis can yield more reliable, rather than being dominated by the majority class. When a sample is drawn from a population, the statistical measures derived from that exhibit sampling variability. The fundamental concept of bootstrap revolves around resampling the original dataset with replacement to generate multiple bootstrap samples. In our study, for 1000 iterations, we selected a sample with replacement that included all observations indicating the presence of cirrus clouds (according to CALIPSO), as well as an equal number randomly drawn from the remaining observations. Each time, the previously described measures were calculated. After performing these calculations 1000 times, the average of these measures was computed.

To demonstrate the concept of bootstrap sampling, we conducted a simple experiment using a dataset consisting of 100 observations. Of these, 15 correspond to cirrus clouds (positive class), and 85 correspond to non-cirrus clouds (negative class). Given the significant class imbalance, many models tend to favor the majority class, leading to overly optimistic accuracy metrics. For example, a naive model that predicts "non-cirrus" for all observations achieves an overall accuracy (OA) of 85%, correctly classifying all negative instances while entirely disregarding the minority class:

OA = (TP+TN)/n =(0+85)/100=0.85 (85%)

To mitigate this imbalance, we applied bootstrap sampling to generate a balanced dataset through resampling with replacement, ensuring an equal number of positive and negative instances (e.g., 15 cirrus and 15 non-cirrus cases). When the same naive model was applied to the balanced dataset, the overall accuracy dropped to 50%, highlighting the model's inability to correctly classify the minority class:

OA = (TP+TN)/n =(0+15)/30=0.50 (50%)

This experiment illustrates how bootstrap sampling can reveal the shortcomings of models trained on imbalanced datasets, offering a more accurate and realistic assessment of model performance.

The bootstrap has been already widely used among climatological studies. It has been employed to, among others, estimate confidence interval (Jolliffe, 2007), forecast storm track (Wilks et al., 2009), project future climate (Orlowsky et al., 2010), verify potential predictability of seasonal mean temperature and precipitation (Feng et al., 2011), study seasonal prediction of drought (Behrangi et al., 2015), inspect macrophysical properties of tropical cirrus clouds (Thorsen et al., 2013), evaluate sampling error in TRMM/PR rainfall products (Iida et al., 2010).

3. Cirrus clouds in 2015

Before conducting an analysis to assess the agreement in high-level cloud detection between CALIOP and MODIS data, we examined the cirrus coverage in 2015 according to reference data (CALIOP). The Cirrus cloud mask (Ci) was generated by applying a condition that classified each 54-degree pixel based on the proportion of observations identified as Cirrus. Specifically, the number of Cirrus observations (nCi) and non-Cirrus observations (nNONCi) within each pixel were counted. The percentage of Cirrus observations (CiCoverage) for a given pixel was a fraction of observations with cirrus detected to all observations.calculated using the formula:

$$CiCoverage = \frac{\mathrm{nCi}}{\mathrm{nCi} + \mathrm{nNONCi}} * 100$$

This approach ensures that the mask reflects the relative frequency of Cirrus clouds within each 54-degree pixel, providing a spatially resolved representation of their distribution.

The distribution of cirrus clouds (~~Fig.~~ 2.) varies globally and is affected by factors such as latitude and atmospheric dynamics. According to (Sassen et al. (2008),., 2008), the total frequency of cirrus clouds from 15 June 2006 to 15 June 2007 was reported as 16.7%, compared to 18.7% observed in our study for 2015. However, according to the research by Kotarba (2022), annual mean values of cloud amount, derived from CALIPSO, can vary significantly (over 10 p.p.) between years due to sampling frequency.

Cirrus clouds are more frequently observed at night, particularly in tropical and mid-latitude regions, with their occurrence peaking around midnight and reducing during the day also according to Noel et al. (2018). Moreover, frequencies of stratospheric cirrus clouds measured by CALIPSO from 2006 to 2012 are 2-3 times higher are detected at night time rather than at daytime (Zou et al., 2020). Nevertheless, the day-night difference observed in Sassens's (2008)their study was smaller than in ours, with values of 15.2% during the day and 18.3% at night, compared to 13.2% (Fig. 2a.) and 23.3% (Fig. 2b.), respectively, in our analysis. These differences may stem from the use of different versions of source datasets or the application of varying data quality filtering criteria. The higher detectability of nighttime cirrus clouds may also be attributed to reduced noise in lidar signals under nighttime conditions. Additionally, the differences might also reflect more intense convective activity and increased formation of cirrus clouds during the night.

In our study, near~~Based on the CALIOP dataset, cirrus cloud coverage reached 18.7% in 2015, daytime coverage of high-level clouds in 2015 was recorded at 13.2% (Fig. 2a.), whereas nighttime coverage was higher, measured at 23.3% (Fig. 2b.). Near~~ the equator, especially within the tropical belt, cirrus cloud cover exhibits peak values throughout the year, reaching approximately 35% during nighttime and 20% during daytime. In certain locations, particularly during nighttime, the high-level cloudiness has been observed to exceed 50%. In the mid-latitudes of both hemispheres, the distribution of clouds varies~~ with the seasons~~, generally showing lower coverage compared to low latitudes, with approximately 10% during daytime and 20% at night. In polar regions, particularly above approximately 60° latitude, cirrus cloud cover tends to be higher than in mid-latitudes, with nighttime coverage generally higher than daytime (Fig. 3.).

~~nighttime coverage generally higher than daytime (Fig.~~

589 ~~3.).~~

Fig. 2. CALIOP-based cirrus cloud coverage in 2015- daytime (a) and nighttime (b)

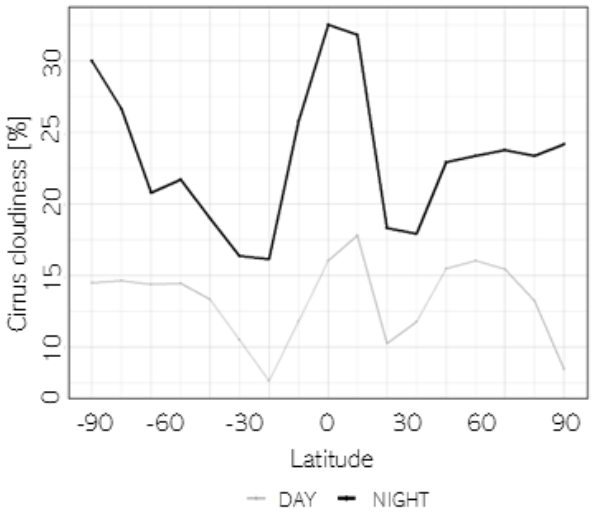


Fig. 3. Cirrus coverage as a function of latitude (CALOP data)

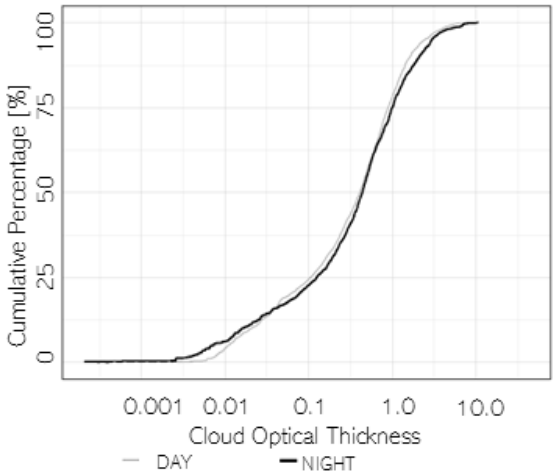
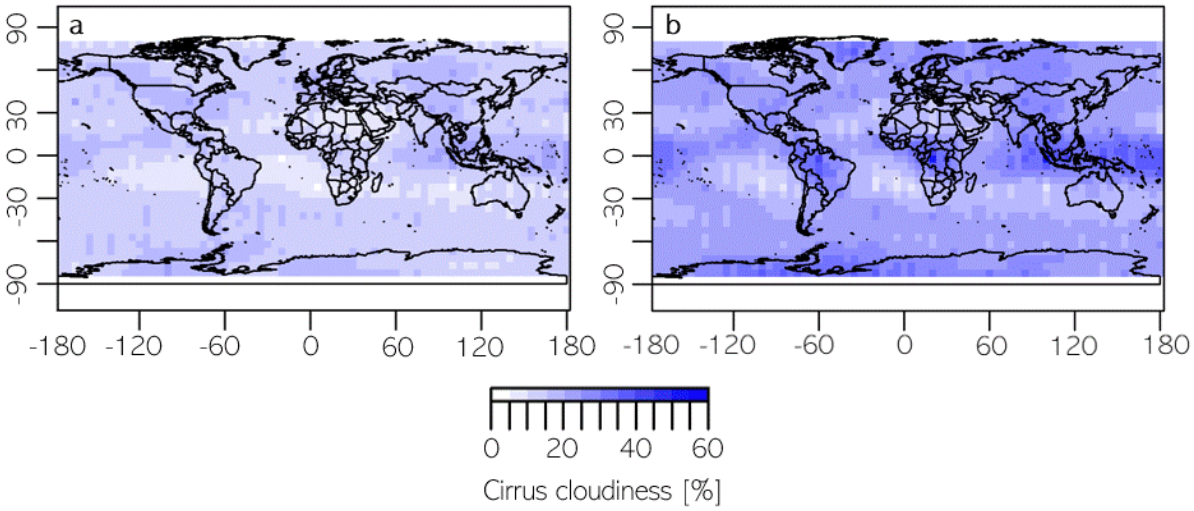
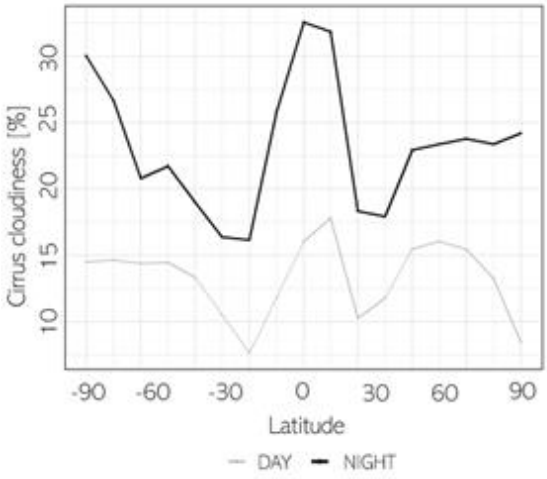


Fig. 4. Cumulative ratio of cirrus clouds with respect to COT (CALIOP data)

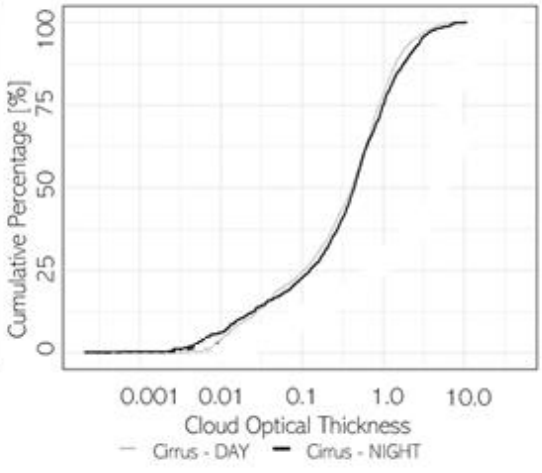


~~Fig. 2. CALIOP based cirrus cloud coverage in 2015 daytime (a) and nighttime (b)~~

590



~~Fig. 3. Cirrus coverage as a function of latitude (CALOP data)~~



~~Fig. 4. Cumulative ratio of cirrus clouds with respect to COT (CALIOP data)~~

591

592 Additionally, CALIOP~~,~~ measures the cloud optical thickness ~~(COT)~~ for individual layers as well as for the
593 entire atmospheric column (Fig. 4.). When CALIOP detects multiple cirrus cloud layers, the COT values
594 for all layers flagged as Cirrus are summed. The mean ~~COT for~~ cirrus COT~~clouds~~ was observed to be
595 0.72 during daytime and 0.84 at nighttime, indicating a notable increase~~. For the entire column (all~~
596 ~~cloud layers in~~ optical thickness~~column), the average COT measured by CALIOP was 4.26 during the~~
597 ~~day and 4.20 at night.~~ This can raise important question about the underlying cause of this difference.
598 One possible explanation is that the increased nighttime COT enhances the likelihood of cirrus cloud
599 detection, as lidar systems like CALIPSO have greater sensitivity to optically thicker clouds.
600 Consequently, this could lead to a higher observed cloud cover at night simply due to improved

detectability rather than actual physical differences in cloud properties. Alternatively, data filtering processes might contribute to the observed disparity.

4. Evaluation of MODIS data

Using CALIPSO data as the reference, nine methods for detecting cirrus clouds with MODIS data were evaluated. All tests were applicable during daytime, whereas only five could be utilized at nighttime due to the requirement of solar illumination.

The measures described~~Ddescribed~~ in section 2 ~~measures~~ are presented in ~~tab.~~Table 3~~2~~3. The parameters that, in our opinion, precluded the use of the test are~~have been~~are highlighted in bold. Additionally, they are preceded ~~with~~ by the rate of observations performed (ROP) parameter, which is the fraction of total observations for which the specific test could be conducted.

During daytime, the first four methods (SOLAR, IR, BT13.9, BT6.7) exhibited notably low detection effectiveness (with POD ranging between 0.33 and 15.79%), as well as low kappa coefficients (0.01-0.48). Although the test was performed on a relatively high proportion of observations (78.37% - 97.59%), with a low number of false alarms (FAR between 1.23% and 13.16%) and good overall accuracy (OA ranging between 48.61% and 53.80%), the poor detection capabilities (indicated by POD) rendered these data inadequate as reliable sources of information on the occurrence of Ci clouds. The differing parameters excluded tests BT3.9-12.0 and those with ISCCP criteria from consideration. The limited number of observations with available results from these tests rendered them impractical for use.

The two tests most effective globally were BT1.38 and ATC. With very similar parameters (POD, FAR, OA and kappa~~Kappa~~) the ATC test demonstrated superiority due to a significantly higher number of available observations (78.37% vs 98.67%, respectively).

Among the night tests, IR, BT13.9, and BT6~~B6~~.7 exhibited low detection capabilities (POD 0.60% - 10.59%), whereas the BT3.9-12.0 test was performed only on 38.09% ~~-~~of the observations. As with the daytime tests, the ATC test proved to be the most suitable for detection.

Considering that global statistics for January and July were not markedly different from the yearly averages (Table 3~~. 2~~3.), subsequent analyses were conducted using data from the entire year.

Table 3~~. 2~~3~~.~~ Goodness-of-fit of cloud detection between MODIS and CALIOP. Bold - parameters that precluded the use of the test

Daytime					Nighttime				
Accuracy measures									

	ROP [%]	POD	FAR	OA	k	ROP [%]	POD	FAR	OA	k
					2015					
SOLAR	78.37	**15.79**	13.16	51.66	**0.03**	**0.00**	NA	NA	NA	NA
IR	83.32	**12.56**	4.37	53.80	0.48	73.98	**10.59**	3.27	54.94	0.52
BT13.9	65.52	**1.35**	3.59	48.61	**-0.02**	71.02	**2.13**	3.42	50.67	**-0.01**
BT6.7	97.59	**0.33**	1.23	49.92	**-0.01**	91.44	**0.60**	1.58	50.23	**-0.01**
BT1.38	78.37	77.76	28.28	74.71	0.49	**0.00**	NA	NA	NA	NA
BT3.9-12.0	**7.39**	64.48	15.36	72.41	0.46	**38.09**	39.09	5.46	65.26	0.33
ATC	98.67	80.87	34.86	72.98	0.46	94.84	25.46	6.90	59.50	0.19
ISCCP23	**37.97**	84.16	72.00	61.26	**0.13**	**0.00**	NA	NA	NA	NA
ISCCP3.6	**37.97**	33.30	16.54	58.96	**0.17**	**0.00**	NA	NA	NA	NA
					January 2015					
SOLAR	74.84	**15.08**	13.50	49.28	**0.02**	**0.00**	NA	NA	NA	NA
IR	78.95	**12.47**	4.54	51.81	0.46	72.30	**10.53**	3.46	54.07	0.51
BT13.9	67.59	**1.66**	3.66	46.28	**-0.02**	72.26	**2.36**	3.32	49.65	**-0.01**
BT6.7	97.95	**0.23**	1.09	49.68	**-0.01**	99.97	**0.59**	1.43	49.59	**-0.01**
BT1.38	74.84	79.65	31.69	74.22	0.48	**0.00**	NA	NA	NA	NA
BT3.9-12.0	**7.02**	56.89	13.50	69.48	0.41	**41.19**	35.00	3.80	64.37	0.30
ATC	98.98	80.23	34.17	73.03	0.46	99.98	23.38	6.12	58.63	99.98
ISCCP23	**38.55**	84.27	68.88	64.10	**0.17**	**0.00**	NA	NA	NA	NA
ISCCP3.6	**38.55**	33.38	14.58	59.27	**0.19**	**0.00**	NA	NA	NA	NA
					July 2015					
SOLAR	84.32	**16.57**	11.58	53.99	**0.05**	**0.00**	NA	NA	NA	NA
IR	92.26	**11.99**	3.76	54.17	0.49	68.77	**10.02**	2.61	57.81	0.56
BT13.9	65.65	**1.89**	3.72	49.61	**-0.02**	67.48	**2.62**	3.93	53.88	**-0.01**
BT6.7	99.69	**0.15**	1.06	49.63	**-0.01**	81.30	**0.84**	1.96	52.06	**-0.01**
BT1.38	84.32	74.97	22.06	76.52	0.53	**0.00**	NA	NA	NA	NA
BT3.9-12.0	**7.67**	72.20	21.54	74.30	0.47	**37.58**	47.02	7.95	67.82	0.38
ATC	99.96	83.14	31.76	75.69	0.51	88.61	30.47	7.99	62.05	0.23
ISCCP23	**36.57**	85.54	74.77	61.16	**0.12**	**0.00**	NA	NA	NA	NA
ISCCP3.6	**36.57**	32.84	16.26	58.67	**0.17**	**0.00**	NA	NA	NA	NA

As previously mentioned, all statistical measures were also calculated for different latitudes (Fig. 5.).

The observed latitudinal variability can be attributed to the physical properties of the different radiation wavelengths used by each channel, as well as their specific functions. Additionally, this variability is influenced by factors such as the spatial distribution of Cirrus clouds and the varying illumination conditions across latitudes. ~~The observed latitudinal variability can be attributed to the~~

physical properties of the respective radiation range and the intended function of the specific channel, as well as to the spatial distribution of cirrus clouds occurrence. For almost all of the tests we observe the ROP (Fig. 5a. & Fig. 5b.) decrease with the latitude increase. This is related to presence of solar illumination. The exception is ROP according to BT3.9-12.0 (which increase from 0% in tropics to almost 30% in polar region) which and was specifically designed for nighttime observations over land and polar snow/ice surfaces. ROP for both tests using ISCCP criteria is equal.

The latitudinal distribution of POD during the day (Fig. 5c.) showsshowsn that ISCCP criteria the most accurately detected cirrus clouds in the tropical regions (up to 75% for ISCCP23 and almost 100% for ISCCP3.6), with POD reduction with latitude decrease (to about 10% and 40% respectively). A similarsSimilar pattern was observed i.e. for BT13.9 method, but with cirrus detection capabilities about 3 times inferior. Depending on the test, latitudinal variability of POD could be also higher for mid-latitudes (ATC), low latitudes (test utilizing the solar radiation range), or remained relatively unchanged. There is no clear trend of increasing/decreasing POD with latitude during the night (Fig. 5d.; slightly more cirrus correctly detected for polar regions by IR, BT13.9 and BT3.9-12.0 tests). The mid-latitudes exhibit POD drop for BT6.7 test, and consequently ATC test.

Figure 5 (Fig. 5e. & Fig. 5f.) shows also the latitudinal variability of FAR. In the tropical regions most of the tests show peak of falsely reported cirrus clouds during daytime in equatorial region (with maximum exceeding 90% for ISCCP23 and 50% for ISCCP3.6). Additionally, BT1.38 test falsely detects cirrus more often with increasing latitude, which increase, whichat results with in 'bimodal' FAR distribution with peaks in tropics (about 35%) and midlatitudes (75% for northern hemisphere and 30% for southern). A distribution resembling BT1.38 exhibited test ATC, but with an upward shift of about 10 percentage points. Relatively few falsely observed cloud cases, with similar to the daytime distribution, were detected at night.

No significant differences were found between the equatorial and polar regions for all the tests for OA. For the daytime the latitudinal variation was more readily observable and varied (Fig. 5g. & 5i. vs Fig. 5h. & 5j.).

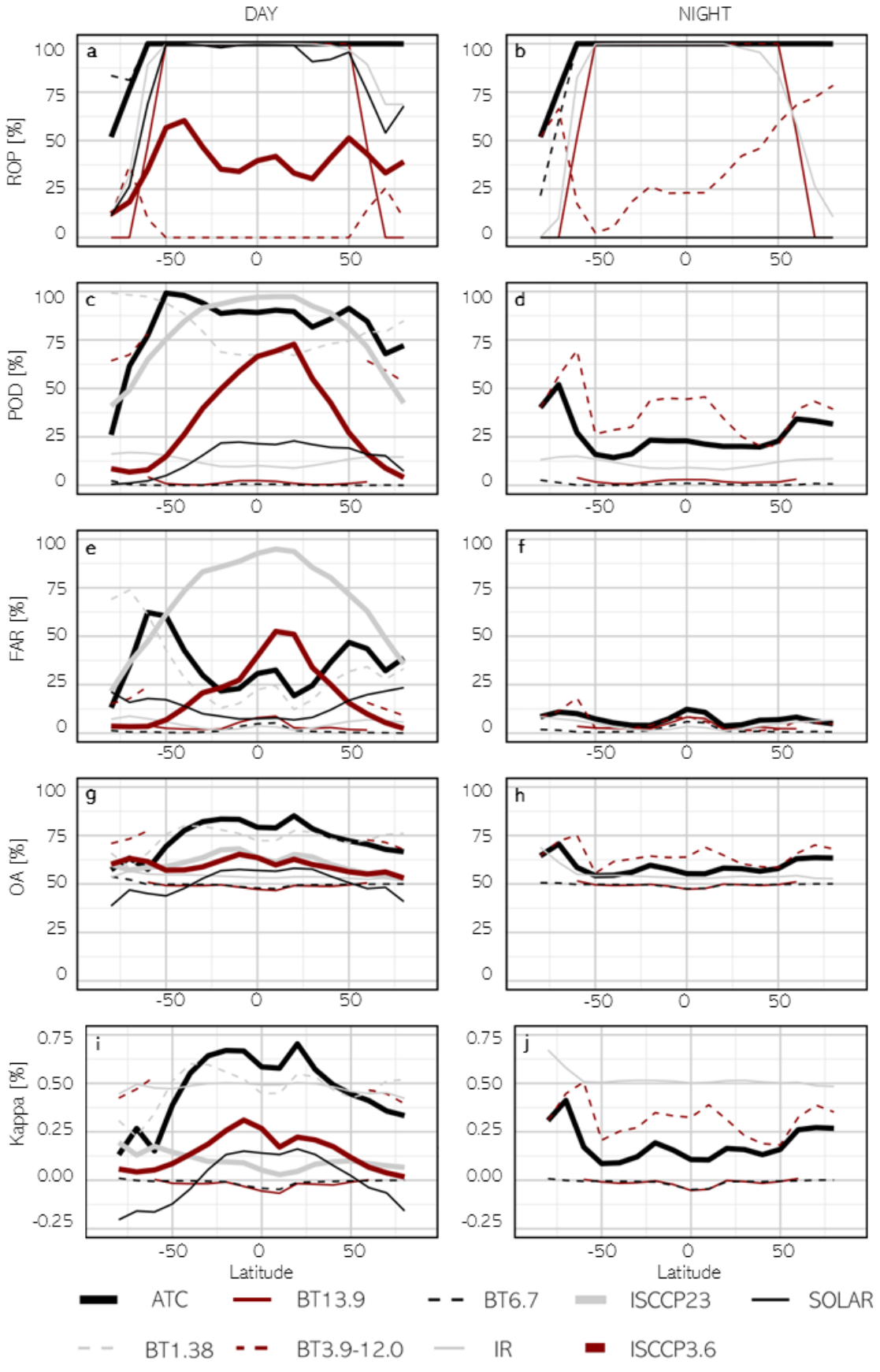


Fig. 5. Cirrus detection accuracy with respect to the latitude (letters (a, …, j) used to facilitate reference in the text)

Considering the very high proportion of correctly detected cirrus clouds, the high overall accuracy and kappa coefficient (degree of agreement between two classification methods), ATC test showed the highest agreement with CALIOP data. Additionally, it covers nearly all observations in the test (96.7%) and shows relatively low variability of statistical measures across different latitudes. Therefore, it can be used as a basis for studies~~studiesy~~ evaluating cirrus cloud coverage in long term perspective.

To ensure the ATC test performs optimally under various conditions and to provide a comprehensive analysis, fit measures were additionally evaluated for "number of layers found" (NLF, Fig. 6.) and IGBP (The International Geosphere–Biosphere Programme, ~~tab.~~Table 4).

~~Since~~ CALIOP data products allow~~is a lidar providing high-resolution atmospheric profiles, with vertical resolutions ranging from 30 m~~ to report up to ~~180 m, it's output could be divided into maximum~~ 10 cloud layers within a profile.~~.~~ When multiple cloud layers overlap, the lidar signal may be attenuated, potentially leading to underestimation of cloud detection. Our research evaluated the collocation of MODIS data to the reference CALIOP data, segmented by the number of detected cloud layers excluding cirrus clouds. A zero indicated that no other cloud layers were detected besides possible cirrus in a given profile. Both day and night observations revealed a maximum of four additional cloud layers. Based on the test conducted, ROP either decreased (i.e. BT13.9 70% to 30% at daytime or BT3.9-12.0 at nighttime), increased (7% to 25% at daytime for BT3.9-12.0), or remained stable with an increasing number of cloud layers (Fig. 6a. & Fig. 6b.). For ATC test, no discernible trend was identified. No clear trend could be observed for POD, both day and night (Fig. 6c. & Fig. 6d.). However, the distribution of the FAR parameter exhibited a different pattern. In several tests, particularly the ATC test, the FAR value (Fig. 6e & Fig. 6f) significantly increased with the number of detected cloud layers (from 9% to 78% during the day and from 1% to 15% at night for the ATC test). This pattern suggests that for clouds with significant vertical development (i.e., those containing multiple layers), MODIS tended to identify only the uppermost layer, mistakenly classifying it as the entire cloud profile. As a result, the increasing number of falsely detected cirrus clouds, particularly in cases of non-cirrus layers (NLF), is reflected in the distributions of OA and kappaKappa. Specifically, as the number of non-cirrus layers increases, both OA and kappaKappa values decrease, for both day and night observations ~~However, the distribution of FAR parameter exhibited a different pattern. In multiple tests, notably ATC test, the value of FAR (Fig. 6e. & Fig. 6f.) significantly elevated with an increasing number of cloud layers (9% to 78% during day and 1% to 15% at night for ATC). Presumably, for clouds with significant vertical development (with more detected layers), MODIS identified only the uppermost layer, incorrectly categorizing it as the complete cloud profile. Increasing number of falsely reported cirrus with NLF manifests itself in OA and Kappa distribution. With the increase in non-cirrus layers found, there is a corresponding decrease in OA and Kappa, both for day and night (Fig.~~ 6g., Fig. 6h., Fig. 6i. & Fig. 6j.).

The International Geosphere–Biosphere Programme ~~(IGBP)~~ defines ecosystem~~ecosystems~~ surface classifications. For purpose of this study, 17 IGBP groups ~~was~~ were aggregated to 3 classes: water, land and snow (goodness-of-fit with respect to land classification is presented in ~~tab.~~Table 4.). Bright surfaces like snow, ice deserts, or complex terrain with varying surface types can make it challenging to distinguish clouds from the ground. The first noticeable aspect is the significantly lower ROP for snow compared to other classes. Generally, the fit measures are similar to those in previous analyses. During the day, ATC test performs better over water, whereas SOLAR test performs better over land. On the contrary, during nighttime, BT3.9-12.0 test performs better over water, whereas ATC test performs better over land.

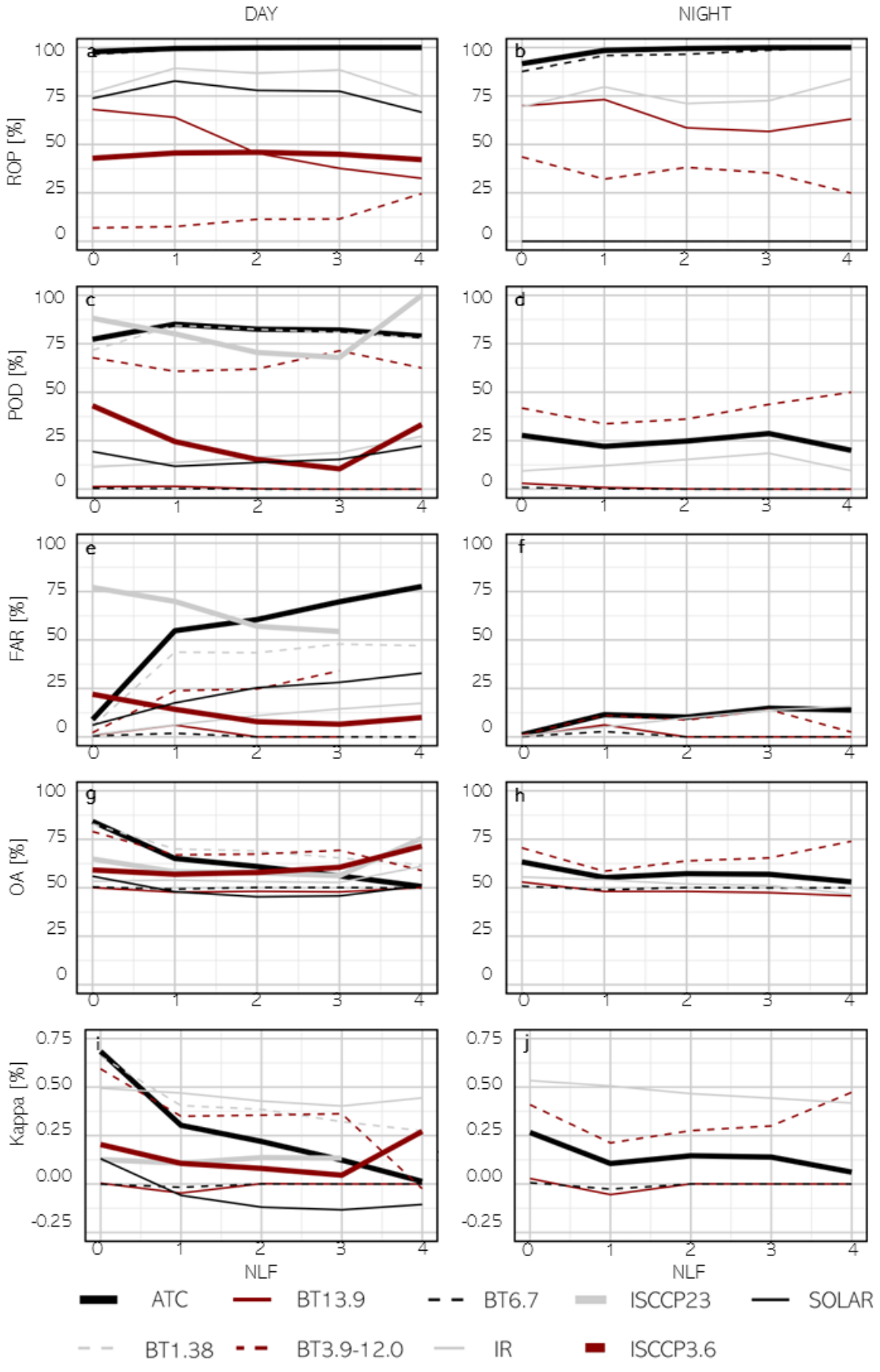


Fig. 6. Cirrus detection accuracy with respect to the NLF (letters (a, …, j) used to facilitate reference in the text)

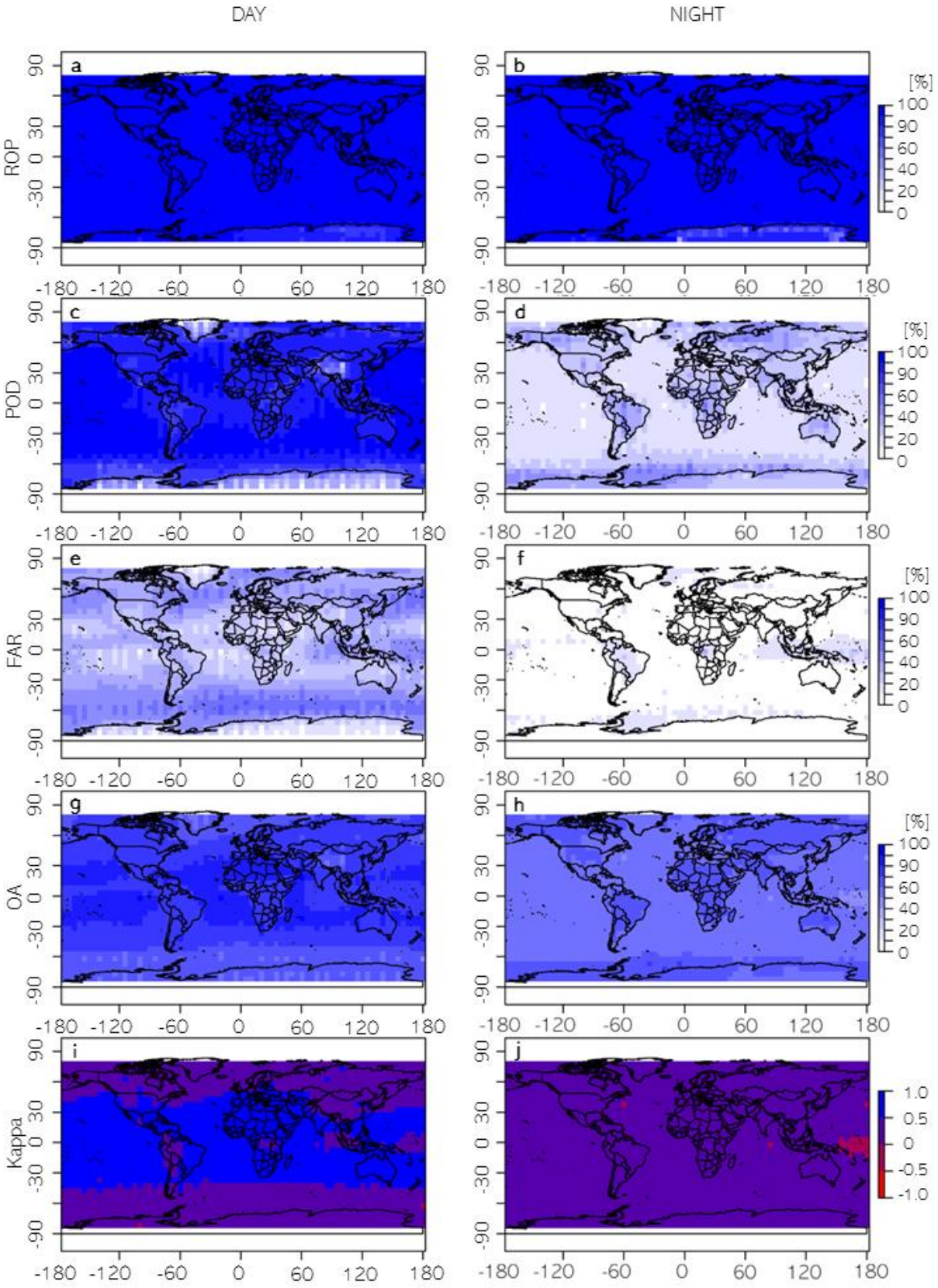
Table 4. ~~34.~~ Goodness-of-fit of cloud detection between MODIS and CALIOP with respect to land classification

	Daytime					Nighttime				
	Accuracy measures									
	ROP [%]	POD	FAR	OA	k	ROP [%]	POD	FAR	OA	k
	WATER									
SOLAR	88.95	11.40	13.05	49.55	-0.02	0.00	NA	NA	NA	NA
IR	92.44	12.93	4.24	54.23	0.48	85.56	11.10	3.36	54.35	0.51
BT13.9	74.18	1.32	3.66	48.21	-0.02	79.41	1.98	3.37	49.90	-0.01
BT6.7	99.99	0.20	1.25	49.48	-0.01	99.98	0.52	1.57	49.48	-0.01
BT1.38	88.95	84.91	30.78	76.99	0.54	0.00	NA	NA	NA	NA
BT3.9-12.0	5.45	67.67	16.17	74.23	0.49	14.64	51.57	8.69	70.13	0.42
ATC	100.00	90.10	40.63	74.73	0.49	99.99	18.94	6.62	56.16	0.12
ISCCP23	29.32	86.27	73.48	62.45	0.14	0.00	NA	NA	NA	NA
ISCCP3.6	29.32	34.69	16.22	59.89	0.19	0.00	NA	NA	NA	NA
	LAND									
SOLAR	84.11	27.39	12.70	57.53	0.15	0.00	NA	NA	NA	NA
IR	93.02	11.42	4.47	52.87	0.48	80.87	9.16	2.92	53.14	0.51
BT13.9	77.48	1.41	3.40	49.65	-0.02	86.30	2.49	3.58	49.46	-0.01
BT6.7	100.00	0.22	1.32	49.45	-0.01	100.00	0.49	1.64	49.42	-0.01
BT1.38	88.95	84.91	30.78	76.99	0.54	0.00	NA	NA	NA	NA
BT3.9-12.0	8.09	62.80	14.99	71.91	0.45	97.78	33.85	3.61	65.15	0.30
ATC	100.00	79.62	29.87	74.88	0.50	100.00	39.34	7.80	65.77	0.32
ISCCP23	45.98	83.88	76.09	58.95	0.08	0.00	NA	NA	NA	NA
ISCCP3.6	45.98	35.73	22.99	55.46	0.12	0.00	NA	NA	NA	NA
	SNOW									
SOLAR	10.35	6.01	20.12	41.56	-0.14	0.00	NA	NA	NA	NA
IR	13.98	15.12	7.27	50.73	0.43	1.12	13.76	5.52	56.13	0.52
BT13.9	0.16	0.72	5.12	47.19	-0.04	0.16	2.59	5.06	49.70	-0.03
BT6.7	78.83	1.70	1.04	54.07	0.01	27.05	2.48	1.86	49.83	0.01
BT1.38	10.35	90.90	53.45	69.55	0.38	0.00	NA	NA	NA	NA
BT3.9-12.0	13.95	61.99	15.30	69.73	0.41	47.02	39.67	7.85	65.31	0.31
ATC	88.29	27.48	10.83	59.27	0.17	55.73	33.67	7.17	62.25	0.26
ISCCP23	11.34	46.54	31.07	57.85	0.16	0.00	NA	NA	NA	NA
ISCCP3.6	11.34	8.00	3.64	59.62	0.05	0.00	NA	NA	NA	NA

The analysis with respect to NLF and land cover types confirmed that ATC test is bestthe most suitablebest suited for achieving the objective of this study. Therefore, the spatial distribution of the individual fit measures for this test was examined (Fig. 7).

The spatialsSpatial distribution indicates reveals a very high level of ROP for both: daytime (Fig. 7a.) and nighttime (Fig. 7b.) for the entire Earth. The southernmost regions of the Southern Hemisphere are an exception, exhibiting lower values.

Spatial variations observed in correctly detected cirrus highlight differences between daytime and nighttime POD distribution (Fig. 7c. & Fig. 7d.). During the daytime, high values are observed over nearly the entire Earth's surface, with exceptions in the regions of Antarctica, Greenland and the Himalayas (≥ 80% vs ≤ 20% respectively), which are regions covered with by snow and ice. However, at night, the highest difference is between land and water (≥ 50% vs c.a.approximately 20%). Similar patterns to the POD distribution for day and night can be observed inwhen considering OAin the OA results (Fig. 7g. - Fig. 7h.). On both sides of the equator, FAR reaches the lowest values, being slightly higher during the day than at night (aroundaboutround 20% and ≤ 5%) and increasing with latitude. However., Hhowever, there is a reduction decrease in FAR observed in regions covered by snow and ice (Fig. 7e. & Fig. 7f.). In regions with the highest rate of correctly detected and the lowest raterateio of falsely reported cirrus, the general accuracy of classification (OA) exceeded 80% duringatduring daytime and an 50% at night. SimilarAs well asSimilar to OA, kappaKappa was higher during the day. During the day, kappaKappa values ranged from 0.5 to 1.0 for in regions at low latitudes. In contrast, at mid and high latitudes, kappaKappa values were between 0.0 and 0.5, remaining positive (Fig. 7i.). At night (Fig. 7j.), nearly the entire Earth's surface exhibited kappaKappa values between 0.0 and 0.5, with a negative kappaKappa values observed in the vicinity of near Micronesia.


DAY
NIGHT
a
b
c
d
e
f
g
h
i
j
ROP
POD
FAR
OA
Kappa
[%]
100
90
60
40
20
0
1.0
0.5
0
-0.5
-1.0
90
30
0
-30
-90
-180
-120
-60
0
60
120
180

739

Fig. 7. Spatial distribution of the accuracy detection of cirrus using ATC test (letters (a, …, j) used to facilitate reference in the text)

5. Discussion ~~and summary~~

This study proved that~~aims to address the research gap by evaluating whether~~ MODIS ready-to-use cloud mask product can be used for producing a reliable cirrus mask. We suggested the best ~~also seek to identify the conditions under which this~~ approach to achieve such a goal, and reported related ~~would be effective and when it might not be suitable. The study found that it was possible however, certain~~ limitations, specifically for ~~particularly those related to~~ nighttime conditions~~, must be consistently considered~~.

During daytime, the two most effective tests were BT1.38 and ATC. With very similar parameters (POD, FAR, OA and kappa~~Kappa~~) the ATC test demonstrated superiority due to a significantly higher number of available observations. Among the nighttime tests the ATC test proved to be the most suitable for cirrus detection.

Additionally, the ATC test covers nearly all observations in the test (96.7%) and shows relatively low variability of statistical measures across different latitudes. Spatial analysis indicates very high level of ROP for both: day and night for the entire Earth. Spatial variations observed in correctly detected cirrus highlight differences between daytime and nighttime POD distribution. During the daytime, high values are observed over nearly the entire Earth's surface, with exceptions in the polar regions and Himalayas. However, at night, land regions display higher POD values compared to the surrounding areas.

The International Satellite Cloud Climatology Project (ISCCP) has long provided a framework for cloud classification and detection, offering standardized methods to analyse cloud properties on a global scale. Within this framework, the ISCCP3.6 and ISCCP23 tests were applied in this study to evaluate their performance in detecting cirrus clouds using MODIS data.

The results of the ISCCP3.6 and ISCCP23 tests highlight their respective strengths and limitations. The ISCCP3.6 test defines cirrus clouds as having an optical thickness below 3.6 and cloud-top pressure below 440 hPa. It demonstrated moderate detection performance during daytime. However, its use is limited to daytime observations, and it achieves a relatively low Rate of Observations Performed (ROP) at 37.97%. Conversely, the ISCCP23 test, which broadens the definition of cirrus to include clouds with optical thicknesses below 23, achieved a significantly higher POD of 84.16% but at the expense of a much higher FAR of 72.00%, resulting in a slightly better OA of 61.26%. Like ISCCP3.6, the ISCCP23 test was also restricted to daytime observations and exhibited the same ROP of 37.97%. When compared to the ATC test, both ISCCP-based tests exhibit notable limitations. Moreover, the ISCCP statistics

presented in this study do not reflect the characteristics of the early years of the ISCCP climatology, which primarily utilized data from satellites equipped with AVHRR sensors, rather than the more recent observations from MODIS.

Considering all mentioned above, the ATC test is proved to be the best among the available methods for detecting high-level clouds. However, it is evident that its utility during nighttime is limited compared to daytime. A notable factor contributing to this is the sensitivity of CALIOP. Lidar is known to have significantly greater sensitivity at night, which explains its ability to detect nearly twice as many cirrus clouds globally at night compared to daytime. This diurnal pattern in CALIOP data, while highlighting the sensor's advantages in nighttime detection, should not be misinterpreted as a definitive indicator of diurnal differences in cirrus cloud occurrence. Instead, it reflects the increased detection capabilities of CALIOP at night.

Additionally, MODIS faces further limitations at night due to the unavailability of the 1.38 μm band, which is highly effective for detecting cirrus clouds during the day. As shown in the statistical analysis, alternative tests exhibit significantly lower performance compared to the 1.38 μm band, emphasizing its critical role in daytime cirrus cloud detection. This limitation further impacts the effectiveness of MODIS-based cirrus detection during nighttime observations.

Consequently, we have determined that the ATC test may be suitable for creating a high-level cloud mask and conducting a long-term climatological analysis of cirrus cloud coverage. This approach simultaneously allows us to address the second research gap mentioned in this paragraph, which concerns our lack of knowledge regarding the long-term variability of high-level cloud coverage. ~~Considering all mentioned above ATC test is proved to be the best among the available methods for detecting high-level clouds. However, it is evident that its utility during nighttime is significantly limited compared to daytime. Consequently, we have determined that it may be suitable for creating a high-level clouds mask and conducting a long-term climatological analysis of cirrus cloud coverage. This approach simultaneously allows us to address the second research gap mentioned in this paragraph, which concerns our lack of knowledge regarding the long-term variability of high-level cloud coverage.~~

Obtained from the CALIOP data, the cirrus mask mentioned in Section 3 allows us to investigate the distribution of cirrus clouds (Fig. 2.) in 2015. Based on the CALIOP dataset, cirrus cloud coverage reached 18.7% in 2015, daytime coverage of high-level clouds in 2015 was recorded at 13.2%, whereas nighttime coverage was higher, measured at 23.3%. TheThis day-night differences result from CALIOP'sdifference can be attributed to the higher nighttime sensitivity, reduced lidar signal noise, and increased nocturnal convective activity leading to more of CALIOP at night, which enhances its ability to detect optically thin cirrus formation. Additionally, annual variations in cloud amount (over 10

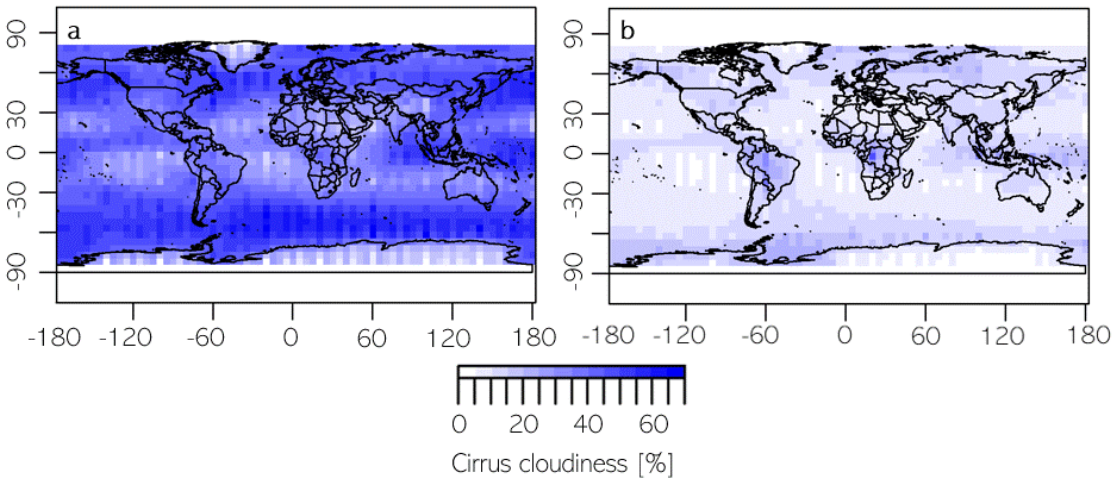
percentage points) may occur due to CALIPSO's sampling frequency, as noted by Kotarba (2022).clouds during nighttime observations.

Similarly, a cirrus mask was generated based on the MODIS data using the ATC test. Derived from this data, cirrus cloud coverage (Fig. 8a.) showed daytime coverage of high-level clouds at 41.0%, while nighttime coverage was lower, measured at 10.9% (Fig. 8b.).

We also compared cirrus cloud coverage in 2015 obtained from CALIOP and MODIS data (Fig. 9.). The mean~~In a similar manner~~Similarly~~, a cirrus mask was generated based on the MODIS data using~~ the ~~ATC test. Derived from this data,~~ ~~cirrus~~ cloud ~~coverage (Fig. 8a.)~~ showed ~~daytime coverage of high-level clouds was recorded at 41.0%, whereas~~ while ~~nighttime coverage was lower, measured at 10.9% (Fig. 8b.). Regrettably, this indicates that although achieving comparable outcomes as reported in existing literature (Sassen et al., 2008), the data concerning the diurnal pattern of cloud coverage is entirely lost.~~

~~We also compared cirrus cloud coverage in 2015 obtained from CALIOP and MODIS data (Fig. 9.).~~ The m~~Mean~~ difference between cirrusc~~Cirrus~~ coverage derived from CALIOP and MODIS was -27.71 p.p. for ~~the~~ daytime observations (Fig. 9a.), with MODIS generally indicated higher cloud cover compared to CALIOP. On the contrary, the mean difference between cirrusc~~Cirrus~~ coverage derived from CALIOP and MODIS was -12.31 p.p. for the nighttime observations (Fig. 9b.). While the relationship between MODIS and CALIOP is statistically significant ($p < 0.001$), the R^2 value of 0.165 indicates that MODIS captures only 16.5% of the variability. In the nighttime dataset, the R^2 improves to 0.422, meaning MODIS cloud coverage aligns better with CALIOP at night. Although the majority of fit metrics show improved performance during the day, the high number of false alarms ultimately results in the nighttime fit being more accurate when cirrus coverage is examined in the subsequent analysis.

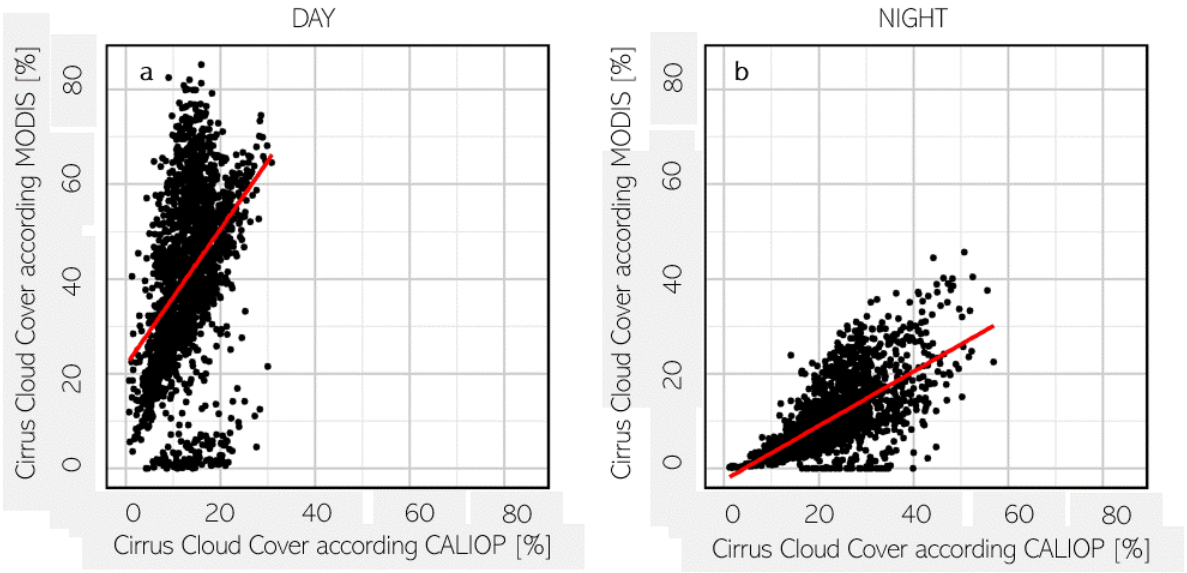
~~In conclusion, our study has shown that ATC test, developed based on MODIS Cloud Product data, demonstrated the highest agreement with reference data~~, achieving ~~(the~~ an ~~overall accuracy of 72.98% during daytime and 59.50% at night,~~ a probability of detection: ~~80.87% and 25.46%,~~ a ~~false alarm rate of 34.86% and 6.90%, and~~ a ~~Cohen's kappa coefficient of 0.46 and 0.19, respectively). Although the study had certain limitations, such as nighttime cirrus detections, its results may have important implications for understanding the nature of high-level cloudiness~~. Future research should focus on investigating long-term trends in cirrus cloudiness, which may help refine our understanding of their variability and improve detection methodologies. ~~. Future research may focus on long-term trends in cirrus cloudiness. These conclusions represent an important step toward a better understanding of the impact of clouds on the climate.~~



841

842 *Fig. 8. MODIS-based cirrus cloud coverage in 2015- daytime (a) and nighttime (b)*

843



844

845 *Fig. 9. CALIOP-MODIS cirrus cloud coverage comparison in 2015- daytime (a) and nighttime (b)*

846 Our goal was to assess the extent to which MODIS detects cirrus clouds in comparison to CALIPSO,
847 while acknowledging that MODIS will inevitably miss a significant portion of cirrus clouds due to its
848 lower sensitivity. This comparison offers valuable insights into the practical efficiency of the MODIS
849 instrument. We accepted MODIS data as it is; however, we examined the fit measures as a function of
850 COT (Fig. 10.), as this primarily explains the differences between MODIS and CALIOP measurements.

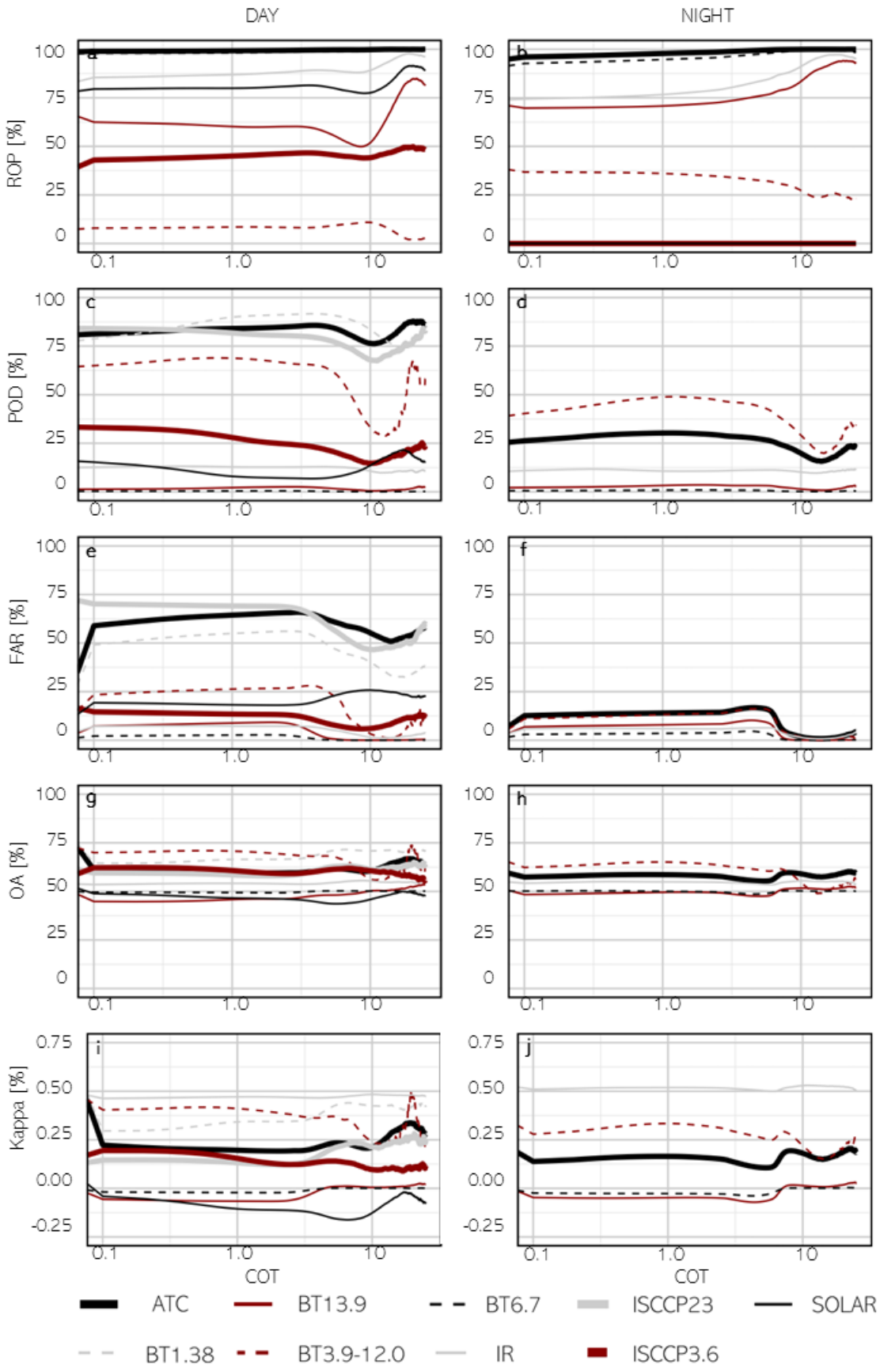


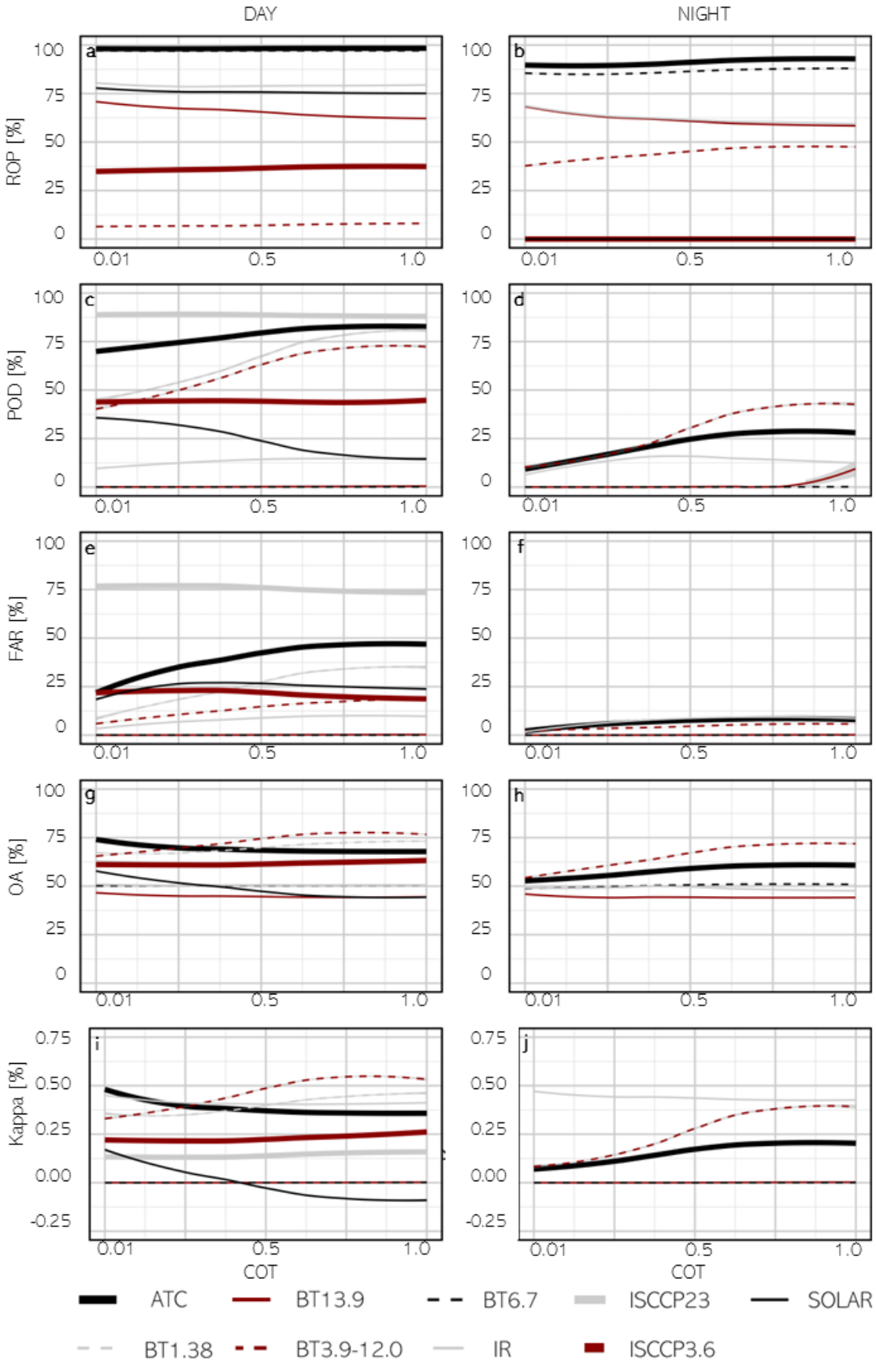
Figure 10. Cirrus detection accuracy with respect to the COT (0-25) (letters (a, …, j) used to facilitate reference in the text)

As observed in the graph (Fig. 10.), there are no significant changes within the range of 0.1 to 1.0, and even up to 10.0. The most noticeable changes occur at COT values close to 10, though these may be influenced by the sample size, as the occurrence of cirrus clouds with a COT near 10 is limited or may represent a misclassification by CALIOP. Notably, differences in parameter values are apparent between a COT of 0 (indicating no cirrus according to CALIOP, at the start of the graph) and 0.1. Upon examining the ATC test results, FAR increases from approximately 30 to 60 during the day, with a similar rise observed at night. The reduced sensitivity of MODIS is reflected in a small but observable increase in POD values as COT increases. Additionally, as thin cirrus clouds become more prevalent, both OA and kappa values decrease.

As mentioned earlier, CALIPSO can detect cirrus clouds with an optical thickness as low as 0.01, whereas MODIS typically detects them when COT ranges between 0.4 and 0.5. Therefore, we analysed the changes in fit measure as a function of COT within the range of 0 to 1, using a finer step size of 0.01 instead of 0.1 as in previous analyses (Fig. 11.).

During the daytime, most methods show a steady increase in POD as COT rises, while at night, POD also improves significantly with increasing COT, with ATC outperforming other tests. When solar radiation is present, FAR increases with higher COT for most methods, indicating more false positives as clouds become optically thicker. At night, FAR remains relatively low but shows a slight upward trend with increasing COT. OA remains stable during both day and night. Kappa improves at night for all methods as COT increases but remains lower than daytime values. For daytime, Kappa is highest for ATC and gradually decreases as COT rises.

Given that MODIS inevitably misses a significant portion of cirrus clouds due to its lower sensitivity, we conducted a detailed analysis for COT values between 0 and 10 and between 0 and 1 with a finer step. The results reveal that fit measures change noticeably with increasing COT for small COT values (<1), a trend that stabilizes for higher COT values. Although it is certain that the issue of thin cirrus clouds generally lowers the fit measures of MODIS to CALIOP, it cannot be said that this is the sole reason for the imperfect fit, as at higher COT values (>1) it also deviates from the full fit. Despite MODIS's limited ability to detect thin cirrus clouds, we do not dismiss its utility. Notably, the ATC method consistently outperforms other approaches across all evaluated metrics, making it a reliable choice for cirrus detection.



883

884 *Figure 11. Cirrus detection accuracy with respect to the COT (0-1) (letters (a, …, j) used to facilitate reference in the text)*

6. Summary

This study evaluates the utility of MODIS data (Aqua satellite) for detecting cirrus clouds by comparing it to CALIOP, a lidar instrument (CALIPSO satellite). Cirrus clouds, located above 6,000-8,000 meters and composed of ice crystals, play a significant role in Earth's radiative budget due to their warming effects. In both sensor's data, cirrus clouds are the same physical phenomenon; however, the distinction arises from the varying sensitivities of the detection instruments employed, with optical thickness serving as a crucial parameter. The research aims to determine how well the MODIS products can be used to identify cirrus clouds compared to CALIPSO.

The study assessed six MODIS tests, the ATC test and two methods originating from ISCCP, using 136 million observations from 2015.

Key findings reveal that the ATC test was the most effective for detecting cirrus clouds:

- During daytime, it achieved an overall accuracy of 72.98%, with a probability of detection (POD) of 80.87%, a false alarm rate (FAR) of 34.86%, and a Cohen's kappa coefficient of 0.46.
- At nighttime, its overall accuracy dropped to 59.50%, with a POD of 25.46%, FAR of 6.90%, and kappa coefficient of 0.19.

The CALIOP-based cirrus mask revealed a global cirrus cloud coverage of 18.7% in 2015, with higher nighttime coverage (23.3%) compared to daytime (13.2%) due to CALIOP's enhanced nighttime sensitivity. In contrast, the MODIS-based ATC test estimated daytime cirrus coverage at 41.0%, but significantly lower nighttime coverage at 10.9%. Equatorial regions exhibited the highest cirrus frequencies, particularly at night.

Despite its limitations, the ATC test shows promise for creating a high-level cloud mask and conducting long-term climatological studies.

This study represents a critical step toward leveraging MODIS data for understanding high-level cloud coverage and its climatic impacts.

Acknowledgements

This work was supported by the National Science Center of Poland [grant number 2021/41/N/ST10/02274].

We gratefully acknowledge Poland's high-performance Infrastructure PLGrid ACK Cyfronet AGH for providing computer facilities and support within computational grant no PLG/2024/016949.

Ackerman, S. A., Liou, K.-N., Valero, F. P. J., and Pfister, L.: Heating Rates in Tropical Anvils, J. Atmos. Sci., 45, 1606–1623, 1988.

Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., and Gumley, L. E.: Discriminating clear sky from clouds with MODIS, J. Geophys. Res. Atmos., 103, 32141–32157, https://doi.org/10.1029/1998JD200032, 1998.

Ackerman, S. A., Holz, R. E., Frey, R., Eloranta, E. W., Maddux, B. C., and McGill, M.: Cloud detection with MODIS. Part II: Validation, J. Atmos. Ocean. Technol., 25, 1073–1086, https://doi.org/10.1175/2007JTECHA1053.1, 2008.

Amato, U., Antoniadis, A., Cuomo, V., Cutillo, L., Franzese, M., Murino, L., and Serio, C.: Statistical cloud detection from SEVIRI multispectral images, Remote Sens. Environ., 112, 750–766, https://doi.org/10.1016/j.rse.2007.06.004, 2008.

Baum, B. A., Menzel, W. P., Frey, R. A., Tobin, D. C., Holz, R. E., Ackerman, S. A., Heidinger, A. K., and Yang, P.: MODIS cloud-top property refinements for collection 6, J. Appl. Meteorol. Climatol., 51, 1145–1163, https://doi.org/10.1175/JAMC-D-11-0203.1, 2012.

Behrangi, A., Nguyen, H., and Granger, S.: Probabilistic seasonal prediction of meteorological drought using the bootstrap and multivariate information, J. Appl. Meteorol. Climatol., 54, 1510–1522, https://doi.org/10.1175/JAMC-D-14-0162.1, 2015.

Bock, L. and Burkhardt, U.: Reassessing properties and radiative forcing of contrail cirrus using a climate model, J. Geophys. Res., 121, 9717–9736, https://doi.org/10.1002/2016JD025112, 2016.

Campbell, J. R., Lolli, S., Lewis, J. R., Gu, Y., and Welton, E. J.: Daytime cirrus cloud top-of-the-atmosphere radiative forcing properties at a midlatitude site and their global consequences, J. Appl. Meteorol. Climatol., 55, 1667–1679, https://doi.org/10.1175/JAMC-D-15-0217.1, 2016.

Chen, P. Y., Srinivasan, R., Fedosejevs, G., and Narasimhan, B.: An automated cloud detection method for daily NOAA-14 AVHRR data for Texas, USA, Int. J. Remote Sens., 23, 2939–2950, https://doi.org/10.1080/01431160110075631, 2002.

Efron, B.: The Jackknife, the bootstrap, and other resampling plans, 1980.

Eleftheratos, K., Zerefos, C. S., Zanis, P., Balis, D. S., Tselioudis, G., Gierens, K., and Sausen, R.: A study on natural and manmade global interannual fluctuations of cirrus cloud cover for the period 1984-2004, Atmos. Chem. Phys., 7, 2631–2642, https://doi.org/10.5194/acp-7-2631-2007, 2007.

Feng, X., Delsole, T., and Houser, P.: Bootstrap estimated seasonal potential predictability of global temperature and precipitation, Geophys. Res. Lett., 38, 1–6, https://doi.org/10.1029/2010GL046511, 2011.

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.

Frey, R. A., Ackerman, S. A., Holz, R. E., Dutcher, S., and Griffith, Z.: The continuity MODIS-VIIRS cloud mask, Remote Sens., 12, 1–18, https://doi.org/10.3390/rs12203334, 2020.

Gu, L., Ren, R., and Zhang, S.: Automatic cloud detection and removal algorithm for MODIS remote sensing imagery, J. Softw., 6, 1289–1296, https://doi.org/10.4304/jsw.6.7.1289-1296, 2011.

Guenther, B., Xiong, X., Salomonson, V. V., Barnes, W. ., and Young, J.: On-orbit performance of the Earth Observing System Moderate Resolution Imaging Spectroradiometer; first year of data, Remote Sens. Environ., 83, 16–30, 2002.

Holz, R. E., Ackerman, S. A., Nagle, F. W., Frey, R., Dutcher, S., Kuehn, R. E., Vaughan, M. A., and Baum, B.: Global Moderate Resolution Imaging Spectroradiometer (MODIS) cloud detection and height evaluation using CALIOP, J. Geophys. Res. Atmos., 114, 1–17, https://doi.org/10.1029/2008JD009837, 2009.

Hutchinson, K. D. and Hardy, K. R.: Threshold functions for automated cloud analyses of global meteorological satellite imagery, Int. J. Remote Sens., 16, 3665–3680, https://doi.org/10.1080/01431169508954653, 1995.

Iida, Y., Kubota, T., Iguchi, T., and Oki, R.: Evaluating sampling error in TRMM/PR rainfall products by the bootstrap method: Estimation of the sampling error and its application to a trend analysis, J. Geophys. Res. Atmos., 115, 1–14, https://doi.org/10.1029/2010JD014257, 2010.

Jolliffe, I. T.: Uncertainty and inference for verification measures, Weather Forecast., 22, 637–650, https://doi.org/10.1175/WAF989.1, 2007.

Kärcher, B.: Formation and radiative forcing of contrail cirrus, Nat. Commun., 9, 1–17, https://doi.org/10.1038/s41467-018-04068-0, 2018.

Kinne, S. and Liou, K.-N.: The Effects of the Nonsphericity and Size Distribution of Ice Crystals on the Radiative Properties of Cirrus Clouds, Atmos. Res., 24, 273–284, 1989.

Kotarba, A. Z.: Regional high-resolution cloud climatology based on MODIS cloud detection data, Int. J. Climatol., 36, 3105–3115, https://doi.org/10.1002/joc.4539, 2016.

Kotarba, A. Z.: Calibration of global MODIS cloud amount using CALIOP cloud profiles, Atmos. Meas. Tech., 13, 4995–5012, https://doi.org/10.5194/amt-13-4995-2020, 2020.

Kotarba, A. Z. and Nguyen Huu, Ż.: Accuracy of Cirrus Detection by Surface-Based Human Observers, J. Clim., 35, 3227–3241, https://doi.org/10.1175/JCLI-D-21-0430.1, 2022.

Li, Q. and Groß, S.: Satellite observations of seasonality and long-term trends in cirrus cloud properties over Europe : investigation of possible aviation impacts, Atmos. Chem. Phys., 22, 15963–15980, 2022.

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.

Liu, Z., Vaughan, M., Winker, D., Kittaka, C., Getzewich, B., Kuehn, R., Omar, A., Powell, K., Trepte, C.,

988 and Hostetler, C.: The CALIPSO lidar cloud and aerosol discrimination: Version 2 algorithm and initial
989 assessment of performance, J. Atmos. Ocean. Technol., 26, 1198–1213,
990 https://doi.org/10.1175/2009JTECHA1229.1, 2009.
991 Lolli, S., Campbell, J. R., Lewis, J. R., Gu, Y., Marquis, J. W., Chew, B. N., Liew, S. C., Salinas, S. V., and
992 Welton, E. J.: Daytime top-of-the-atmosphere cirrus cloud radiative forcing properties at Singapore, J.
993 Appl. Meteorol. Climatol., 56, 1249–1257, https://doi.org/10.1175/JAMC-D-16-0262.1, 2017.
994 Macke, A., Francis, P. N., Mcfarquhar, G. M., and Kinne, S.: The role of ice particle shapes and size
995 distributions in the single scattering properties of cirrus clouds, J. Atmos. Sci., 55, 2874–2883,
996 https://doi.org/10.1175/1520-0469(1998)055<2874:TROIPS>2.0.CO;2, 1998.
997 Menzel, W. P., Frey, R. A., and Baum, B. A.: Cloud Top Properties and Cloud Phase Algorithm
998 Theoretical Basis Document Collection 006 Update, 73, 2015.
999 Minnis, P., Trepte, Q. Z., Sun-Mack, S., Chen, Y., Doelling, D. R., Young, D. F., Spangenberg, D. A.,
1000 Miller, W. F., Wielicki, B. A., Brown, R. R., Gibson, S. C., and Geier, E. B.: Cloud detection in nonpolar
1001 regions for CERES using TRMM VIRS and Terra and Aqua MODIS data, IEEE Trans. Geosci. Remote
1002 Sens., 46, 3857–3884, https://doi.org/10.1109/TGRS.2008.2001351, 2008.
1003 Mishchenko, M. I., Rossow, W. B., Macke, A., and Lacis, A. A.: Sensitivity of cirrus cloud albedo,
1004 bidirectional reflectance and optical thickness retrieval accuracy to ice particle shape, J. Geophys.
1005 Res., 101, 16973–16985, 1996.
1006 Murino, L., Amato, U., Carfora, M. F., Antoniadis, A., Huang, B., Menzel, W. P., and Serio, C.: Cloud
1007 detection of modis multispectral images, J. Atmos. Ocean. Technol., 31, 347–365,
1008 https://doi.org/10.1175/JTECH-D-13-00088.1, 2014.
1009 Musial, J. P., Hüsler, F., Sütterlin, M., Neuhaus, C., and Wunderle, S.: Daytime low stratiform cloud
1010 detection on AVHRR imagery, Remote Sens., 6, 5124–5150, https://doi.org/10.3390/rs6065124,
1011 2014.
1012 Nguyen Huu, Ż. and Kotarba, A. Z.: Reliability of visual detections of cirrus over Poland, Theor. Appl.
1013 Climatol., 14, 1–11, 2021.
1014 Oreopoulos, L., Cho, N., and Lee, D.: New Insights about Cloud Vertical Structure from CloudSat and
1015 CALIPSO observations, J. Geophys. Res. Atmos., 122, 9280–9300,
1016 https://doi.org/10.1002/2017JD026629.New, 2017.
1017 Orlowsky, B., Bothe, O., Fraedrich, K., Gerstengarbe, F. W., and Zhu, X.: Future climates from bias-
1018 bootstrapped weather analogs: An application to the Yangtze River basin, J. Clim., 23, 3509–3524,
1019 https://doi.org/10.1175/2010JCLI3271.1, 2010.
1020 Pandit, A. K., Gadhavi, H. S., Venkat Ratnam, M., Raghunath, K., Rao, S. V. B., and Jayaraman, A.: 16
1021 year climatology of cirrus clouds over a tropical station in southern India using ground and space-
1022 based lidar observations, Atmos. Chem. Phys. Discuss., 15, 15791–15830,

https://doi.org/10.5194/acpd-15-15791-2015, 2015.
Ramanathan, V., Cess, R. ., Harrison, E. ., Minnis, P., Barkstrom, B. ., Ahmad, E., and Hartmann, D.: Cloud-Radiative Forcing and Climate : Results, Science (80-.)., 243, 57–63, 1989.
Rossow, W. B. and Schiffer, R. A.: Advances in Understanding Clouds from ISCCP, Bull. Am. Meteorol. Soc., 80, 2261–2287, https://doi.org/10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2, 1999.
Sassen, K., Wang, Z., and Liu, D.: Global distribution of cirrus clouds from CloudSat/cloud-aerosol lidar and infrared pathfinder satellite observations (CALIPSO) measurements, J. Geophys. Res. Atmos., 114, 1–12, https://doi.org/10.1029/2008JD009972, 2008.
Stanski, H. R., Wilson, L. J., and Burrows, W. R.: Survey of Common Verification Methods in Meteorology, World weather watch technical report no 8, WMO/TD No 358, 1989.
Stephens, G. L. and Webster, P. J.: Clouds and Climate: Sensitivity of Simple Systems, J. Atmos. Sci., 38, 235–247, 1981.
Stephens, G. L., Tsay, S. C., Stackhouse, P. W., and Flatau, P. J.: The relevance of the microphysical and radiative properties of cirrus clouds to climate and climatic feedback, https://doi.org/10.1175/1520-0469(1990)047<1742:trotma>2.0.co;2, 1990.
Stephens, G. L., Winker, D., Pelon, J., Trepte, C., Vane, D., Yuhas, C., L'Ecuyer, T., and Lebsock, M.: Cloudsat and calipso within the a-train: Ten years of actively observing the earth system, Bull. Am. Meteorol. Soc., 99, 569–581, https://doi.org/10.1175/BAMS-D-16-0324.1, 2018.
Stordal, F., Myhre, G., Stordal, E. J. G., Rossow, W. B., Lee, D. S., Arlander, D. W., and Svendby, T.: Is there a trend in cirrus cloud cover due to aircraft traffic?, Atmos. Chem. Phys., 5, 2155–2162, https://doi.org/10.5194/acp-5-2155-2005, 2005.
Stubenrauch, C. J., Cros, S., Guignard, A., and Lamquin, N.: A 6-year global cloud climatology from the Atmospheric InfraRed Sounder AIRS and a statistical analysis in synergy with, Atmos. Chem. Phys. Discuss., 15, https://doi.org/10.5194/acp-10-7197-2010, 2010.
Sun-Mack, S., Minnis, P., Chen, Y., Kato, S., Yi, Y., Gibson, S. C., Heck, P. W., and Winker, A. D. M.: Regional apparent boundary layer lapse rates determined from CALIPSO and MODIS data for cloud-height determination, J. Appl. Meteorol. Climatol., 53, 990–1011, https://doi.org/10.1175/JAMC-D-13-081.1, 2014.
Tang, H., Yu, K., Hagolle, O., Jiang, K., Geng, X., and Zhao, Y.: A cloud detection method based on a time series of MODIS surface reflectance images, Int. J. Digit. Earth, 6, 157–171, https://doi.org/10.1080/17538947.2013.833313, 2013.
Thorsen, T. J., Fu, Q., Comstock, J. M., Sivaraman, C., Vaughan, M. A., Winker, D. M., and Turner, D. D.: Macrophysical properties of tropical cirrus clouds from the CALIPSO satellite and from ground-based micropulse and Raman lidars, J. Geophys. Res. Atmos., 118, 9209–9220, https://doi.org/10.1002/jgrd.50691, 2013.

Vaughan, M. A., Powell, K. A., Kuehn, R. E., Young, S. A., Winker, D. M., Hostetler, C. A., Hunt, W. H., Liu, Z., Mcgill, M. J., and Getzewich, B. J.: Fully automated detection of cloud and aerosol layers in the CALIPSO lidar measurements, J. Atmos. Ocean. Technol., 26, 2034–2050, https://doi.org/10.1175/2009JTECHA1228.1, 2009.

Wang, T., Fetzer, E. J., Wong, S., Kahn, B. H., and Yue, Q.: Validation of MODIS cloud mask and multilayer fl ag using CloudSat-CALIPSO cloud profiles and a cross-reference of their cloud classifications, J. Geophys. Res., 121, https://doi.org/10.1038/175238c0, 2016.

Wilks, D. S., Neumann, C. J., and Lawrence, M. B.: Statistical extension of the National Hurricane Center 5-day forecasts, Weather Forecast., 24, 1052–1063, https://doi.org/10.1175/2009WAF2222189.1, 2009.

WMO: International Cloud Atlas, Volume I: Manual on the Observation of Clouds and Other Meteors, 180 pp., https://doi.org/10.2307/1550553, 1977.

Wylie, D. P., Menzel, W. P., Woolf, H. M., and Strabala, K. I.: Four Years of Global Cirrus Cloud Statistics Using HIRS, J. Clim., 7, 1972–1986, 1994.

Xie, Y., Qu, J. J., and Xiong, X.: Improving the CALIPSO VFM product with Aqua MODIS measurements, Remote Sens. Lett., 1, 195–203, https://doi.org/10.1080/01431161003720387, 2010.

Zhang, Y., Laube, M., and Raschke, E.: Numerical simulations of cirrus properties, Contrib. to Atmos. Phys., 67, 109–120, 1994.

Zhang, Y., MacKe, A., and Albers, F.: Effect of crystal size spectrum and crystal shape on stratiform cirrus radiative forcing, Atmos. Res., 52, 59–75, https://doi.org/10.1016/S0169-8095(99)00026-5, 1999.