Evaluation of the operational MODIS cloud mask product for detecting cirrus clouds

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 κ 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. Active satellite sensors (lidars) 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. In this study, we assessed the utility of Moderate Resolution Imaging Spectroradiometer (MODIS) standard products for creating a cirrus mask by validating them against Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) data. Our objective was to determine how well the operational cloud mask from the MODIS Science Team can be used to infer the presence of cirrus clouds relative to data products derived from the highly sensitive ~~Cloud-Aerosol Lidar with Orthogonal Polarization (~~CALIOP~~)~~ instrument by the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) Science Team.

Using CALIOP data as the reference, we evaluated six tests for cirrus and high clouds detection considered in MODIS cloud masking algorithm and their combination (all tests consolidation, ATC). Additionally, we applied classifications based on the cirrus definition from the International Satellite Cloud Climatology Project (ISCCP), which rely on retrieved MODIS cloud-top properties. These were used not as detection tests, but as a classification scheme for comparative purposes. All other tests were applied directly to MODIS radiances.~~Additionally we applied tests that originated in the International Satellite Cloud Climatology Project (ISCCP) definition of cirrus. 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% (probability of detection 80.9%, false alarm rate 34.9%, Cohen's ~~κ~~kappa 0.46) during daytime and 59.50% at night (probability of detection 25.5%, false alarm rate 6.9%, Cohen's ~~κ~~kappa 0.19). However, its effectiveness was notably reduced during nighttime compared to daytime. We conclude that the MODIS operational Cloud Mask after being modified into the ATC test is moderately suitable for creating a mask of high-level clouds, and only daytime. During the night-time MODIS ATC fails to reliably report the presence of cirrus.

63

64

65

66

67

68

69

70

71

1. Introduction

72

73 Clouds are indispensable to Earth's environmental systems and human life, influencing weather,

74 climate, water distribution, ecosystems, and various human activities. They affect the Earth's radiation

75 budget, with a net radiative forcing of approximately -20 Wm^{-2} (Boucher et al., 2013), which results in

76 an overall cooling effect on the planet. Nevertheless, special attention should be paid to high-level

77 clouds - according to the WMO, high-level clouds include Cirrus, Cirrocumulus, and Cirrostratus (WMO,

78 1977) - commonly referred to as *cirrus*. Those clouds play a complex role in climate regulation. The

79 relation between cirrus particles (size, shape and albedo) and Earth's radiation budget has been

80 examined (Kinne and Liou, 1989; Macke et al., 1998; Mishchenko et al., 1996; Stephens et al., 1990;

81 Zhang et al., 1994, 1999), resulting in a general conclusion that cirrus play an important role and can

82 warm the atmosphere. They typically have a base above about 8,000 m and consist of small ice crystals.

83 Due to their unique properties - such as altitude, temperature, effective particle size, surface thermal

84 contrast, ice water path, and optical depth (Ackerman et al., 1988; Stephens et al., 1990; Stephens &

85 Webster, 1981), they differ from low- and mid-level clouds in their effect on the Earth's radiation

86 budget. Specifically, cirrus clouds allow shortwave radiation to reach the surface while reducing

87 outgoing longwave radiation, thereby contributing to warming. Globally, it's been estimated that cirrus

88 clouds have a net warming effect of 35.5 Wm^{-2} (Campbell et al., 2016; Kärcher, 2018; Lolli et al., 2017;

89 Oreopoulos et al., 2017) in part because they trap and reduce outgoing longwave radiation more

90 efficiently than they reflect solar radiation back to space. Furthermore, cirrus clouds can alter the

91 radiative forcing of other cloud types. For example, when medium and low clouds co-occur, their

92 combined radiative effect is -18.8 Wm^{-2}, but the additional presence of cirrus raises this effect to 50.8

93 Wm^{-2} (Oreopoulos et al., 2017).

94 A description of cirrus cloud properties is incomplete without information about their coverage. Most

95 studies have focused on total cloud cover, but some have also examined high-level cloudiness. The

96 global frequency of cirrus occurrence is estimated to range between 17% and 42%. Research

97 conducted using high-resolution satellite data indicates that global cloud coverage is approximately

98 66% to 74%, with 40% of all clouds classified as high-level clouds (Sassen et al., 2008; Stubenrauch et

99 al., 2010). Numerous studies have explored changes in high-level cloud coverage. However, those

100 relying on satellite data often do not address cirrus clouds over sufficiently long periods—at least 30

101 years, as recommended by the WMO. Conducting such long-term studies and identifying suitable data

102 sources remain significant challenges.

103 Given the critical role of cloud cover, especially cirrus, observing clouds is of considerable importance.

104 Historically first method is visual observation from ground-based meteorological stations, which is

105 simple and provides long time series data. However, this method has limitations, including difficulty in

106 detecting high-level clouds due overlapping clouds at multiple altitudes, perspective distortions near

107 the horizon, and the optical thinness of cirrus clouds. Studies have shown that under optimal

108 conditions, the probability of visually detecting cirrus clouds ranges from 44% to 83% during the day

109 and from 24% to 42% at night. When clouds at all levels are present, detection probabilities drop to

110 47%–71% during the day and 28%–43% at night (Kotarba & Nguyen Huu, 2022).

111 Modern cloud climatologies benefit from satellite remote sensing. Initially, this information was

112 obtained from various imagers, sounders, and radiometers, which utilize passive cloud detection

113 methods (involving detecting natural radiation emitted or reflected by objects, such as clouds, without

114 actively sending out signals). Researchers such as Ackerman et al. (2008); Amato et al. (2008); Chen et

115 al. (2002); Frey et al. (2008, 2020); Gu et al. (2011); Kotarba (2016); Y. Liu et al. (2004); Minnis et al.

116 (2008); Murino et al. (2014); Musial et al. (2014); Tang et al. (2013) have contributed to these studies.

117 An example of passive sensor can be MODIS ~~(Moderate Resolution Imaging Spectroradiometer)~~, which

118 is a key instrument aboard the Terra and Aqua satellites.

119 Active remote sensing technology, in contrast, relies on its own signal, directing it at an object and

120 analysing the response. This allows active sensors, for instance CALIPSO's ~~(Cloud-Aerosol Lidar and~~

121 ~~Infrared Pathfinder Satellite Observations)~~ lidar, CALIOP ~~(Cloud-Aerosol Lidar with Orthogonal~~

122 ~~Polarization)~~, to operate day and night with similar efficiency in cloud detection. Active profiling

123 instruments like CALIOP, which provide high-resolution vertical profiles of aerosols and clouds, have

124 limitations, including a narrow field of view. This narrow view, combined with a long 16-day repeat

125 cycle, results in only about 20 observations per year of the same region, which is challenging and

126 sometimes insufficient for climatological studies (Kotarba, 2022).

127 To standardize cloud classification and ensure consistency, the International Satellite Cloud

128 Climatology Project (ISCCP) developed a system based on cloud height and optical thickness, providing

129 a systematic framework for studying cloud types and their variability across regions and over time. This

130 classification is crucial for advancing climate modelling, weather forecasting, and research on cloud-

131 climate interactions. The ISCCP classification was applied to MODIS data, and its effectiveness in

132 detecting cirrus clouds was also evaluated. In this study, we refer to the ISCCP classification not as a

133 detection method, but as a widely accepted climatological framework based on retrieved parameters.

134 While active sensors like CALIOP remain the most reliable source of cirrus data (e.g. Heidinger and

135 Pavolonis, 2009), their potential for building long-term climatologies is limited. In contrast, passive

136 data have been available for over 40 years, offering temporal coverage suitable for climatological

137 research. One example of such sensors, although collecting data for over 20 years rather than 40, is

138 MODIS, whose capabilities for detecting cirrus clouds are limited compared to those of active vertical

139 profiling sensors.

140 In this paper, we use cirrus characterizations from CALIOP data to explore the potential for creating a

141 cirrus mask from the operational MODIS cloud data products. Our objective is to determine how well

142 the MODIS products can be used to identify cirrus clouds compared to CALIPSO. Specifically, we aim

143 to assess whether the existing MODIS cloud detection tests used in the generation of the MYD35

144 operational data can be re-purposed for cirrus cloud masking without the need to develop a new cirrus

145 detection algorithm.

146

147 **2. Data and methods**

148 In this study, we use active sensor data for validating passive-based information for determining the

149 presence of cirrus (for the sake of clarity, throughout this manuscript, all high-level clouds will be called

150 as cirrus). The active sensor data was collected by the CALIOP lidar aboard the CALIPSO satellite, while

151 the passive data was obtained from the MODIS multi-band radiometer on the Aqua satellite. The

152 concept behind achieving the research objective was based on collocation of those two datasets in

153 time and space. In both instances, cirrus clouds are the same physical phenomenon; however, the

154 distinction arises from the varying sensitivities of the detection instruments employed, with optical

155 thickness serving as a crucial parameter. CALIPSO is capable of identifying cirrus clouds with an optical

156 thickness as low as approximately 0.01, while MODIS generally detects them when the optical

157 thickness is at least of 0.4 to 0.5 (e.g. Menzel et al., 2015). Data for the year 2015 were analyzed on a

158 global scale, comprising 136,272,209 collocated MODIS-CALIPSO observations. The primary

159 requirement was to obtain a sufficiently large sample of CALIPSO-MODIS match-ups across different

160 seasons and geographic regions, which necessitated one complete year of global observations.

161 Therefore, 2015 was chosen arbitrarily.

162 **2.1. MODIS data**

163 MODIS, an advanced instrument aboard NASA's Terra and Aqua satellites, acquires data across 36

164 spectral bands, spanning wavelengths from visible to thermal infrared (0.4 to 14.4 µm). Its passive

165 sensors rely primarily on naturally available energy: solar energy reflected from objects or absorbed

166   and re-emitted (e.g. Ackerman et al., 1998). MODIS provides data at various spatial resolutions - 250

167   m, 500 m, and 1 km - with a swath width of 2,330 km, enabling it to observe the entire Earth twice

168   daily, one observation during the day and one at night. Cloud detection results are stored in the 48-

169   bit "Cloud Mask" product, known as MYD35 for Aqua, while corresponding cloud properties can be

170   found in MYD06 dataset. As an imager, MODIS provides column-integrated radiances, which limits its

171   ability to retrieve cirrus-specific information. For this research, we used Collection 061 data, which is

172   available in 5-minute granules at a spatial resolution of 1 km per pixel (at nadir). Each MYD35 and

173   MYD06 file is paired with a MYD03 "Geolocation file" product that contains longitude and latitude

174   information for each individual cloud mask IFOV (instantaneous field of view; Guenther et al., 2002).

175

176   **2.1.1. The MODIS Cloud Mask product**

177   The MODIS Cloud Mask product is a Level 2 dataset produced at spatial resolutions of 1 km and 250 m

178   (at nadir). The cloud masking procedure was described in details by Ackerman et al. (1998), Frey et al.

179   (2008), and Baum et al. (2012). The algorithm utilizes a sequence of visible and infrared threshold and

180   consistency tests to determine the confidence level that an unobstructed view of the Earth's surface

181   is achieved.

182   The primary MODIS routine for identifying clouds is the MODIS Cloud Mask (product MOD35), which

183   applies a series of spectral threshold tests to each pixel. The cloud mask algorithm does not explicitly

184   label cloud type (no specific "cirrus" output flag); instead, it provides a confidence level that the pixel

185   is cloudy or clear.

186   However, certain tests within the algorithm are specifically designed to detect high thin clouds like

187   cirrus. It particularly applies to tests using the spectral band centred at 1.38μm, a MODIS-introduced

188   wavelength for cirrus detection (Bo-Cai Gao and Kaufman, 1995). In this research, we considered 6

189   individual cloud detection tests of MODIOS cloud mask, which – according to the cloud mask detection

190   algorithm (Ackerman et al., 1998) - have a potential for cirrus or high cloud detection:

191   −   Thin Cirrus test (SOLAR) – the solar channels in MODIS cover a range of wavelengths primarily

192   in the visible and near-infrared spectrum (0.4 to 2.5 μm). Test is only applied during daytime.

193   −   Thin Cirrus test (IR) – the purpose of this test is detecting thin cirrus clouds. Channels used for

194   this test are 11 μm an 12 μm (infrared (IR) range), incorporated to the split window technique.

195   This test leverages the fact that cirrus clouds, composed of ice, are more transparent at 11 μm

196   than at 12 μm, resulting in a positive BTD signature.

197 − High Cloud Test (BT13.9) – applying CO_2 absorption channels (around 14 µm) is a simple
198 technique got from the CO_2 slicing method (suitable for determining middle and upper
199 troposphere ice clouds heights and effective amounts). This test is useful for high-level cloud
200 detection, while it can reveal clouds above 500 hPa – it helps identify high clouds by detecting
201 colder cloud tops using the CO2 absorption band.
202 − High Cloud Test (BT6.7) – test designed for detecting thick high clouds. Starting from the
203 ground level, the 6.7 µm radiation emitted by the surface or low clouds is absorbed in the
204 atmosphere, therefore the signal is not received by an instrument. Water vapor in the
205 atmospheric layer between 200 hPa and 500 hPa is the only source of the 6.7 µm radiation in
206 clear-sky observation. Thick clouds placed above or near the 200 hPa level can be distinguish
207 from clear sky or lower clouds.
208 − High Cloud Test (BT1.38) – the 1.38 µm channel lies in the strong water vapor absorption
209 region. That results in obscuration of the most of Earth's surfaces, as well as attenuation of
210 reflectance from low- and mid-level clouds. Pixels subjected to this test reveal high-level thin
211 clouds as brighter. Unfortunately, the test has certain limitations, including its applicability to
212 nighttime conditions, polar regions, midlatitude winters, and high elevations..
213 − High Cloud Test (BT3.9-12.0) – the 3.9-12.0 µm BTD (Brightness Temperatures Difference) test
214 is specifically designed for nighttime observations over land and polar snow/ice surfaces. It
215 effectively distinguishes thin cirrus clouds from cloud-free conditions and cloud-free
216 conditions and exhibits relative insensitivity to the atmospheric water vapor content (e.g.
217 Hutchinson and Hardy, 1995).

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

221 $$\text{If } \exists_{i \in \{1,2,\dots,9\}} (t_i=1), \text{ then ATC}_{OF}=1$$

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

224 $$\text{If } \forall_{i \in \{1,2,\dots,9\}} (t_i=0), \text{ then ATC}_{OF}=0$$

225 ATC is essentially an adaptation of the MYD35 approach, but it is limited to tests that specifically
226 provide insights 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. For Aqua satellite, dataset is called MYD06.

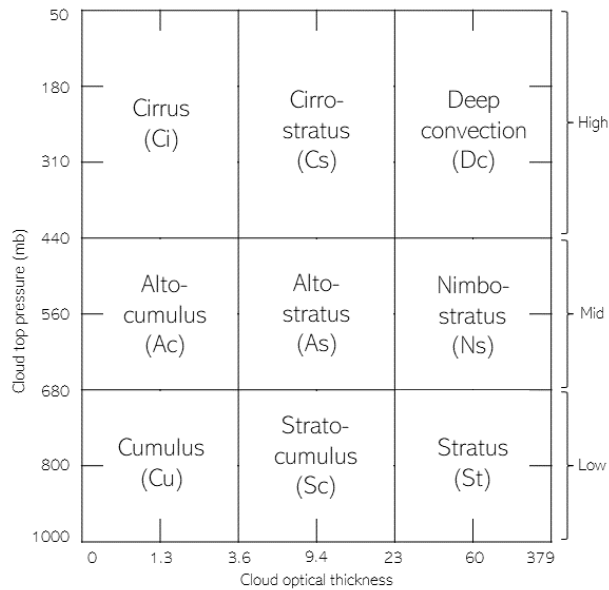
In addition to the ready-to-use MODIS tests (Section 2.2.1), other criteria can be applied using data available from MODIS and CALIOP. For instance, the ISCCP's definition of cloud types. By examining visible and infrared radiances from geostationary and polar-orbiting meteorological satellites and making 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).

COT and CTP are also available for MODIS, within MYD06 standard product, and we used them to generate cirrus masks 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).

It is important to clarify that the ISCCP-based ~~tests~~classes used here are not interpreted as detection algorithms in the same sense as the MODIS cloud mask tests. Instead, they serve as a classification of cloud type based on retrieved cloud optical thickness and top pressure, following established ISCCP thresholds. These classifications are included for reference as they are widely used in long-term cloud climatology studies.

253

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

255

2.2.   CALIOP data

256

257 CALIOP provides atmospheric profiles with vertical resolutions ranging from 30 m below 8.2 km to 180

258 m above 20.1 km, and 60 m between these altitudes (Winker et al., 2006). This capability allows for

259 clear distinction between cirrus and lower cloud layers, making CALIOP excellent for cirrus detection.

260 Furthermore, lidar can detect cirrus clouds with an optical depth as low as 0.01 (Vaughan et al., 2009),

261 a capability beyond the reach of passive imagers (Ackerman et al., 2008). Being an active sensor, lidar

262 offers similar effectiveness in cloud detection both daytime and nighttime, or even higher during night,

263 when backscattered light does not interfere with diffused solar radiation (McGill et al., 2007).

264 In this research, the lidar level-2 cloud layer at 5-km horizontal resolution, version 4.20

265 (CAL_LID_L2_05kmCLay-Standard-V4–20) product was used. As described by Liu et al. (2009) and

266 Vaughan et al. (2009) this product reports cloud layers and cloud type information, with cirrus as a

267 separate class (categorized as type 6). In the CALIPSO system, cirrus clouds are detected using the SIBYL

268 and CAD algorithms. The SIBYL algorithm identifies cloud layers based on enhanced backscatter signals

269 in CALIOP lidar profiles. Subsequently, using a probabilistic approach, the CAD (Cloud-Aerosol

270 Discrimination) algorithm differentiates between clouds (including cirrus) and aerosols.. The

271 depolarization ratios for Cirrus clouds are higher than those for water-based clouds, enabling their

272 identification. Additionally, CALIOP provides an information about the cloud base and top altitudes,

273 allowing for determination of their position in the atmosphere. In CALIOP data, cirrus clouds are

274 identified as high-altitude (cloud-top pressure < 440 hPa) and transparent layers—meaning the lidar

275 can detect the surface or a lower atmospheric layer beneath. If a layer is opaque, even at high altitude,

276 it is not classified as cirrus. This ensures that only optically thin, high-level ice clouds are labeled as

277 cirrus. Clouds above the tropopause, such as polar stratospheric clouds, were excluded as they

278 constitute a separate feature type in CALIPSO data. The quality of CALIOP's detection is reflected in

279 CAD score, which ranges from -100 to 100. Value -100 indicates high confidence of aerosol detection,

280 while a value of 100 indicates high confidence in cloud detection. A medium value (0) signifies equal

281 probability that the feature is a cloud or aerosol. To mitigate misclassification, particularly over dust

282 regions (e.g., the Sahara), the CAD algorithm dynamically adjusts depolarization thresholds (Liu et al.,

283 2009; Vaughan et al., 2009). In this study, we used only observations with a CAD score higher than 80.

284 The optical depth is also provided in this (CAL_LID_L2_05kmCLay-Standard-V4–20) CALIOP product.

285 For the purpose of this research, we consider CALIPSO as the reference for cirrus clouds detection.

286

287   **2.3. Matching datasets**

288 In order to achieve the goal of this study, MODIS and CALIOP data were collocated in space and time.

289 It was possible because Aqua and CALIPSO operated for 12 years (2006-2018) as a part of satellite

290 constellation commonly known as the Afternoon Constellation. Members of the constellation used

291 sun-synchronous polar orbits with 16-day revisit cycle, and with equatorial crossing time at 13:30 local

292 solar time (ascending node). CALIPSO followed the Aqua spacecraft by approximately one minute (e.g.

293 Stephens et al., 2018), enabling quasi-simultaneous observation of the same part of the atmosphere,

294 as 1 km ground track of CALIOP always overlapped with 2330 km wide imagery of MODIS.

295 Collocating MODIS with CALIOP has been frequently used to validate reliability of MODIS datasets, or

296 to developed a new, joint imager-lidar atmospheric products (e.g. Baum et al., 2012; Holz et al., 2009;

297 Kotarba, 2020; Sun-Mack et al., 2014; Wang et al., 2016; Xie et al., 2010). Either 333 m, 1 km, or 5 km

298 lidar data may be considered, however only 1 km and 5 km products offers cloud type classification.

299 Additionally, only the 5 km product informs about cloud optical thickness per cloud layer, and provides

300 superior cirrus detection due to higher sensitivity (noise level decreases as more profiles is integrated

301 into retrieval).

302 From the geometry point of view, a 5 km profile is an aggregation of five consecutive 1 km profiles,

303 and the geo-coordinates of the central one are saved as representative for 5 km profile. It possess a

304 challenge when MODIS and CALIOP are to be matched: one 5 km profile of CALIOP only can be

305 accurately matched to one 1 km MODIS pixel, while 5 km data actually covers five MODIS pixels. To

306 overcome this problem we matched CALIOP with MODIS using non-aggregated, 1 km data, and only

307 then assigned 5 km data to already collocated MODIS-CALIOP pairs. As a result, one 5 km profile of

308 CALIOP was used to characterize five MODIS pixels.

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

314 This study relied on MODIS-CALIOP observations for 2015, and the year was selected arbitrary, as the
315 only requirements was to consider a relatively large (year-long) sample of global observations of
316 clouds. Eventually, our database consisted of 136,272,209 paired MODIS-CALIOP observations; the
317 average spatial distance between geometrical centers of matched MODIS pixel and CALIOP profile was
318 444 m (std. dev. = 231 m), while the average temporal separation reached 84 seconds (std. dev. = 12
319 seconds).

320 The final, aggregated MODIS–CALIOP statistics were compiled into global maps, each with a spatial
321 resolution of 5° in both longitude and latitude.

322 ### 2.4. Evaluation of MODIS data

323 We consider a test as a useful for cirrus detection whenever it results with a high agreement with the
324 reference data (CALIOP). The degree of agreement was calculated using a confusion matrix approach
325 for a binary classifier ('cirrus' and 'no cirrus' classes) . The approach compares the model's predictions
326 (MODIS performance) against the actual results (CALIPSO detections of cirrus). Table 1 provides list of
327 statistical indices we used as a measure for MODIS-CALIPSO agreement in cirrus detection.

328 *Table 1. Confusion matrix-based measures used for assessing MODIS agreement with CALIPSO in this study*

Abbreviation	Full name
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
~~K~~kappa	Cohen's kappa ~~K~~k coefficient
PE	Expected agreement
n	Number of elements in the set

329

330 The structure of confusion matrix is presented in Table 2. and includes the following elements:

331   −   True Positives (TP): The number of cases where MODIS accurately identified the existing

332   (according to CALIOP) cirrus.

333   −   False Positives (FP): The number of cases where MODIS incorrectly identified the high-level

334   cloud, meaning it detected cirrus presence when it was actually absent.

335   −   True Negatives (TN): The number of cases where MODIS correctly did not detect the presence

336   of the cloud.

337   −   False Negatives (FN): The number of cases where MODIS overlooked the cirrus occurrence.

338 *Table 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)

339 Every result undergoes thorough validation through different parameter estimation using feature-

340 based statistics (Stanski et al., 1989). To describe the data accuracy, the probability of detection (POD)

341 [1] and false alarm rate (FAR) [2] metrics were calculated:

342 Probability of detection (POD) – is a metric used to assess the effectiveness of a detection system. In

343 the context of cloud detection, POD indicates how well the detection algorithm correctly identifies the

344 presence of clouds when they are actually present. A higher POD value signifies better performance of

345 the detection system.

346 $POD = TP/(TP+FN)$ [1]

347 False alarm rate (FAR) – is a metric that measures the frequency of incorrect positive detections by a

348 system. In the context of cloud detection, a lower FAR indicates a more accurate system, with fewer

349 instances of falsely identifying clouds when they are not present.

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

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

352 measures:

353 Overall accuracy (OA) – is a metric that measures the proportion of correct predictions made by a

354 detection system out of all predictions. In cloud detection, higher overall accuracy indicates that the

355 system effectively identifies both the presence and absence of clouds correctly.

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

357 Cohen's kappa κk – is a statistical metric used to assess the degree of agreement between two raters

358 or classification methods (Cohen, 1960). Its scale ranges from -1 to 1, where a value of 1 represents

359 perfect agreement, 0 indicates agreement no better than chance, and negative values indicate

360 agreement worse than chance. In cloud detection, a higher kappa value indicates stronger agreement

361   between the detected presence of clouds and their actual presence, while considering the possibility

362   of random agreement.

363   κk = (OA - PE)/(1-PE) [4]

364   where

365   PE – expected agreement

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

367   n = TP + FP + FN + TN [6]

368

369   The accuracy of high-level cloud detection was evaluated using the aforementioned metrics,

370   differentiated by day and night, latitude, cloud optical depth, the number of detected cloud layers, and

371   land classification. This assessment was conducted for the entire year 2015, as well as specifically for

372   January and July (those two months are presented to exemplify the characteristics of two distinct

373   seasons).

374

375   **2.5. Bootstrap sampling**

376   Due to the nature of cirrus cloud occurrences (18.7% in 2015, see Section 3), we can assume that the

377   data sample will be imbalanced and one class (without cirrus) significantly outnumbers the other.

378   Therefore, for such data, the appropriate statistical method to apply is bootstrap sampling (Efron,

379   1980).

380   The balancing of the sample stems from the issue of class imbalance, potentially skewing the statistical

381   analysis and leading to biased results. To mitigate this, the bootstrap method is employed to artificially

382   balance the dataset. This involves resampling the data with replacement, to ensure that each class has

383   a comparable number of instances. By doing so, the analysis can yield more accurate results, rather

384   than being dominated by the majority class. When a sample is drawn from a population, the statistical

385   measures derived from that exhibit sampling variability. The fundamental concept of bootstrap

386   revolves around resampling the original dataset with replacement to generate multiple bootstrap

387   samples. In our study, for 1000 iterations, we selected a sample with replacement that included all

388   observations indicating the presence of cirrus clouds (according to CALIPSO), as well as an equal

389   number randomly drawn from the remaining observations. Each time, the previously described

390   measures were calculated. After performing these calculations 1000 times, the average of these

391   measures was computed.

392   To demonstrate the concept of bootstrap sampling, we conducted a simple experiment using a dataset

393   consisting of 100 observations. Of these, 15 correspond to cirrus clouds (positive class), and 85

394   correspond to non-cirrus clouds (negative class). Given the significant class imbalance, many models

395 tend to favor the majority class, leading to overly optimistic accuracy metrics. For example, a naive

396 model that predicts "non-cirrus" for all observations achieves an overall accuracy (OA) of 85%,

397 correctly classifying all negative instances while entirely disregarding the minority class:

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

399 To mitigate this imbalance, we applied bootstrap sampling to generate a balanced dataset through

400 resampling with replacement, ensuring an equal number of positive and negative instances (e.g., 15

401 cirrus and 15 non-cirrus cases). When the same naive model was applied to the balanced dataset, the

402 overall accuracy dropped to 50%, highlighting the model's inability to correctly classify the minority

403 class:

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

405 This experiment illustrates how bootstrap sampling can reveal the shortcomings of models trained on

406 imbalanced datasets, offering a more accurate and realistic assessment of model performance.

407 The bootstrap has been already widely used among climatological studies. It has been employed to,

408 among others, estimate confidence interval (Jolliffe, 2007), forecast storm track (Wilks et al., 2009),

409 project future climate (Orlowsky et al., 2010), verify potential predictability of seasonal mean

410 temperature and precipitation (Feng et al., 2011), study seasonal prediction of drought (Behrangi et

411 al., 2015), inspect macrophysical properties of tropical cirrus clouds (Thorsen et al., 2013), evaluate

412 sampling error in TRMM/PR rainfall products (Iida et al., 2010).

413 **3.   Cirrus distribution in 2015 according to CALIOP**

414 Before conducting an analysis to assess the agreement in high-level cloud detection between CALIOP

415 and MODIS data, we examined the cirrus coverage in 2015 according to reference data (CALIOP). The

416 Cirrus cloud mask was generated by applying a condition that classified each 5-degree latitude-

417 longitude grid cell based on the proportion of observations identified as Cirrus. Specifically, the number

418 of Cirrus observations and non-Cirrus observations within each 5-degree grid cell were counted. The

419 percentage of Cirrus observations for a given 5-degree grid cell was a fraction of observations with

420 cirrus detected to all observations.

421 This approach ensures that the mask reflects the relative frequency of Cirrus clouds within each 5-

422 degree grid cell, providing a spatially resolved representation of their distribution.

423 The distribution of cirrus clouds (Fig. Fig. 2.) varies globally and is affected by factors such as latitude

424 and atmospheric dynamics. According to Sassen et al. (2008), the total frequency of cirrus clouds from

425 15 June 2006 to 15 June 2007 was reported as 16.7%, compared to 18.7% observed in our study for

426   2015. However, according to the research by Kotarba (2022), annual mean values of cloud amount,

427   derived from CALIPSO, can vary significantly (over 10 p.p. (percentage points)) between years due to

428   sampling frequency.

429   Cirrus clouds are more frequently observed at night, particularly in tropical and mid-latitude regions,

430   with their occurrence peaking around midnight and reducing during the day also according to Noel et

431   al. (2018). Moreover, frequencies of stratospheric cirrus clouds measured by CALIPSO from 2006 to

432   2012 detected at nighttime are 2-3 times higher those detected during daytime (Zou et al., 2020).

433   Nevertheless, the day-night difference observed in Sassens's (2008) study was smaller than in ours,

434   with values of 15.2% during the day and 18.3% at night, compared to 13.2% (Fig. 2a.) and 23.3% (Fig.

435   2b.), respectively, in our analysis. These differences may stem from the use of different (earlier)

436   versions of source datasets and the application of different criteria for selecting and screening cloud

437   layer data. The higher detectability of nighttime cirrus clouds may also be attributed to reduced noise

438   in lidar signals under nighttime conditions (the lack of solar background). Additionally, the differences

439   might also reflect more intense convective activity and increased formation of cirrus clouds during the

440   night. However, diurnal differences in cirrus occurrence are complex, the artificial diurnal difference,

441   driven by the varying levels of background noise during the day and night, likely outweighs the actual

442   diurnal variations.In our study, near the equator, especially within the tropical belt, cirrus cloud cover

443   exhibits peak values throughout the year, reaching approximately 35% during nighttime and 20%

444   during daytime. In certain locations, particularly during nighttime, the high-level cloudiness has been

445   observed to exceed 50%. In the mid-latitudes of both hemispheres, the distribution of clouds varies,

446   generally showing lower coverage compared to low latitudes, with approximately 10% during daytime

447   and 20% at night. In polar regions, particularly above approximately 60° latitude, cirrus cloud cover

448   tends to be higher than in mid-latitudes, with nighttime coverage generally higher than daytime (Fig.

449   3.).

450

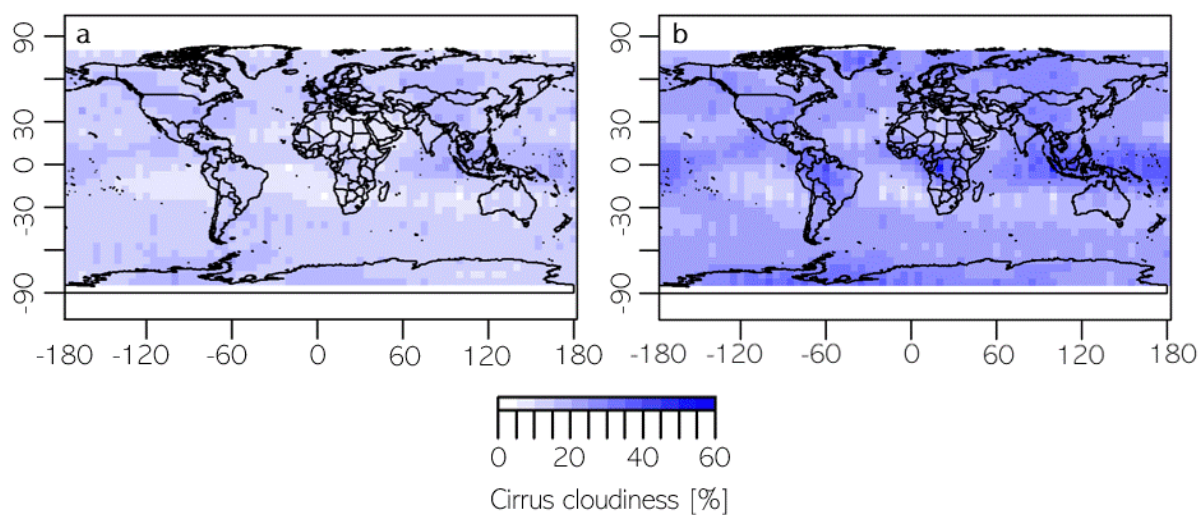



Fig. 2. CALIOP-based cirrus cloud coverage in 2015- daytime (a) and nighttime (b)

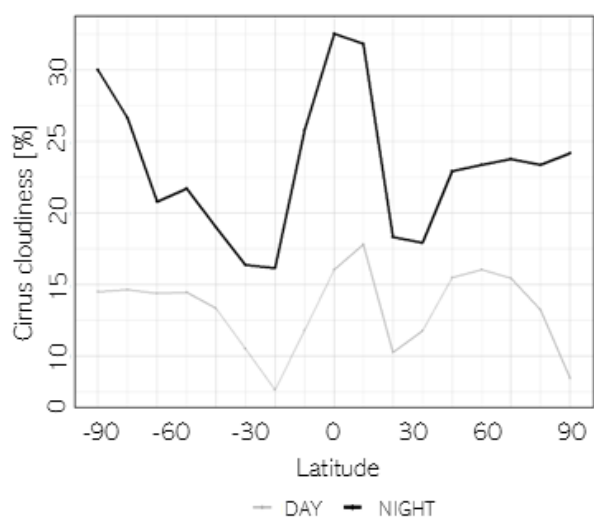



Fig. 3. Cirrus coverage as a function of latitude
(CALOP data)

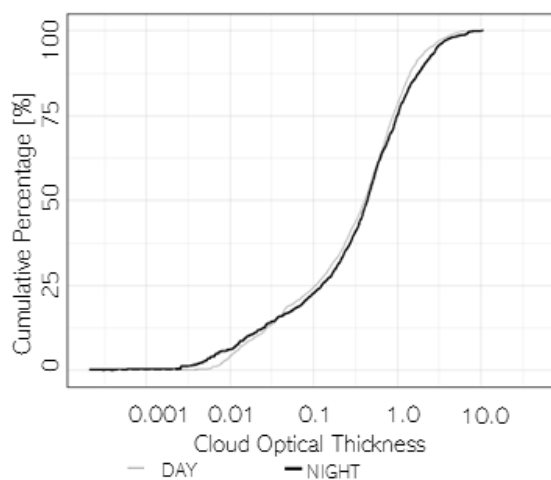



Fig. 4. Cumulative ratio of cirrus clouds with respect to COT
(CALIOP data)

451

452

453 Additionally, CALIOP observations allows for retrieving the cloud optical thickness for individual layers
454 as well as for the entire atmospheric column (Fig. 4.). When CALIOP detects multiple cirrus cloud layers,
455 the COT values for all layers flagged as Cirrus are summed. The mean cirrus COT was observed to be
456 0.72 during daytime and 0.84 at nighttime, indicating a notable increase in optical thickness at night.
457 This can raise important question about the underlying cause of this difference. One possible
458 explanation is that the increased nighttime COT enhances the likelihood of cirrus cloud detection, as
459 lidar systems like CALIPSO have greater sensitivity to optically thicker clouds, and even more so at night
460 due to the absence of solar background and a higher signal-to-noise ratio. Consequently, this could
461 lead to a higher observed cloud cover at night simply due to improved detectability rather than actual
462 physical differences in cloud properties. However, this argument may not fully align with the observed
463 data. Given that CALIOP is more sensitive at night, it could be expected to detect more thin clouds,
464 potentially lowering the average COT compared to daytime. The observed increase in nighttime cirrus
465 COT could stem from multiple sources, including genuine diurnal variability, retrieval algorithm
466 behavior, or screening-induced bias. Further investigation would be needed to isolate their individual
467 contributions. Alternatively, data filtering processes might contribute to the observed disparity.

468 **4. Evaluation of MODIS data**

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

472 The measures described in section 2 are presented in Table 3. The parameters that, in our opinion,
473 precluded the use of the test are highlighted in bold. Additionally, they are preceded by the rate of
474 observations performed (ROP) parameter, which is fraction of all observations under test.

475 During daytime, the first four methods (SOLAR, IR, BT13.9, BT6.7) exhibited notably low detection
476 effectiveness (with POD ranging between 0.33 and 15.79%), as well as low κ kappa coefficients (0.01-
477 0.48). Although the test was performed on a relatively high proportion of observations (78.37% -
478 97.59%), with a low number of false alarms (FAR between 1.23% and 13.16%) and good overall
479 accuracy (OA ranging between 48.61% and 53.80%), the poor detection capabilities (indicated by POD)
480 rendered these data inadequate as reliable sources of information on the occurrence of Ci clouds. The
481 differing parameters excluded tests BT3.9-12.0 and those with ISCCP criteria from consideration. The
482 limited number of observations with available results from these tests rendered them impractical for
483 use.

484 The two tests most effective globally were BT1.38 and ATC. With very similar parameters (POD, FAR,

485 OA and ~~κκappa~~) the ATC test demonstrated superiority due to a significantly higher number of available

486 observations (78.37% vs 98.67%, respectively).

487 Among the night tests, IR, BT13.9, and BT6.7 exhibited low detection capabilities (POD 0.60% - 10.59%),

488 whereas the BT3.9-12.0 test was performed only on 38.09% of the observations. As with the daytime

489 tests, the ATC test proved to be the most suitable for detection.

490 Considering that global statistics for January and July were not markedly different from the yearly

491 averages (Table 3.), subsequent analyses were conducted using data from the entire year.

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

	Daytime					Nighttime				
	Accuracy measures									
	ROP [%]	POD	FAR	OA	Kκ	ROP [%]	POD	FAR	OA	Kκ
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

494   All statistical measures also were calculated for different latitudes (Fig. 5.). The observed latitudinal

495   variability can be attributed to the physical properties of the different radiation wavelengths used by

496   each channel, as well as their specific functions. Additionally, this variability is influenced by factors

497   such as the spatial distribution of Cirrus clouds and the varying illumination conditions across latitudes.

498   For almost all of the tests we observe the ROP (Fig. 5a. & Fig. 5b.) decrease with the latitude increase.

499   This is related to presence of solar illumination. The exception is ROP according to BT3.9-12.0 (which

500   increase from 0% in tropics to almost 30% in polar region) and was specifically designed for nighttime

501   observations over land and polar snow/ice surfaces. ROP for both tests using ISCCP criteria is equal.

502   The latitudinal distribution of POD during the day (Fig. 5c.) shows that ISCCP criteria most accurately

503   detected cirrus clouds in the tropical regions (up to 75% for ISCCP23 and almost 100% for ISCCP3.6),

504   with POD reduction with latitude decrease (to about 10% and 40% respectively). A similar pattern was

505   observed i.e. for BT13.9 method, but with cirrus detection capabilities about 3 times inferior.

506   Depending on the test, latitudinal variability of POD could be also higher for mid-latitudes (ATC), low

507   latitudes (test utilizing the solar radiation range), or remained relatively unchanged. There is no clear

508   trend of increasing/decreasing POD with latitude during the night (Fig. 5d.; slightly more cirrus

509   correctly detected for polar regions by IR, BT13.9 and BT3.9-12.0 tests). The mid-latitudes exhibit POD

510   drop for BT6.7 test, and consequently ATC test.

511   Figure 5 (Fig. 5e. & Fig. 5f.) shows also the latitudinal variability of FAR. In the tropical regions most of

512   the tests show peak of falsely reported cirrus clouds during daytime in equatorial region (with

513   maximum exceeding 90% for ISCCP23 and 50% for ISCCP3.6). Additionally, BT1.38 test falsely detects

514   cirrus more often with increasing latitude, which results in 'bimodal' FAR distribution with peaks in

515   tropics (about 35%) and midlatitudes (75% for northern hemisphere and 30% for southern). The ATC

516   test exhibited a latitudinal distribution of false alarm rate (FAR) that closely resembled the pattern

517   observed for the BT1.38 test, but with values shifted upward by approximately 10 percentage pointsA

518 ~~distribution resembling BT1.38 exhibited test ATC, but with an upward shift of about 10 percentage~~

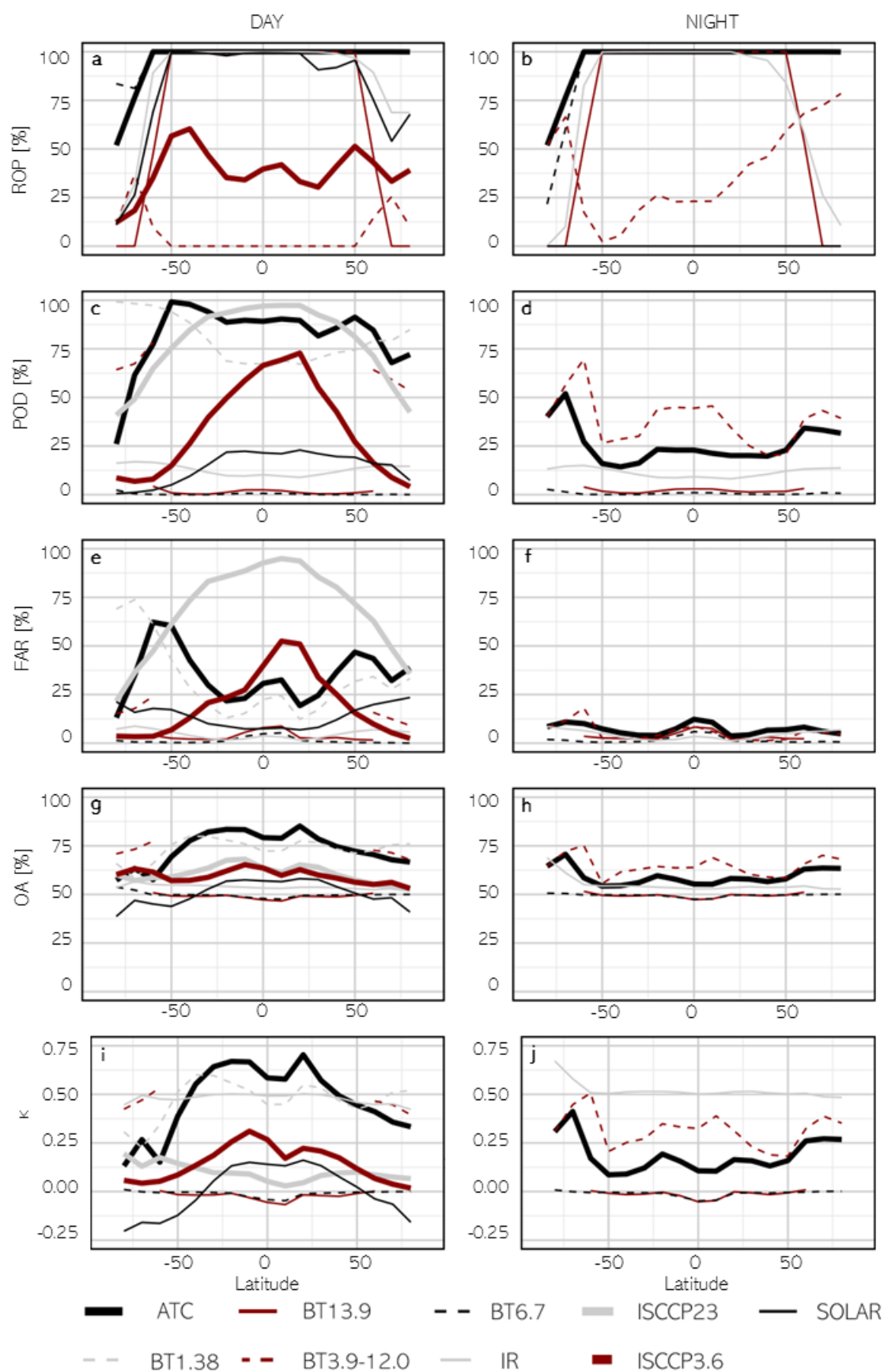
519 ~~points~~. Relatively few falsely observed cloud cases, with similar to the daytime distribution, were

520 detected at night.

521 No significant differences were found between the equatorial and polar regions for all the tests for OA.

522 For daytime the latitudinal variation was more readily observable and varied (Fig. 5g. & 5i. vs Fig. 5h.

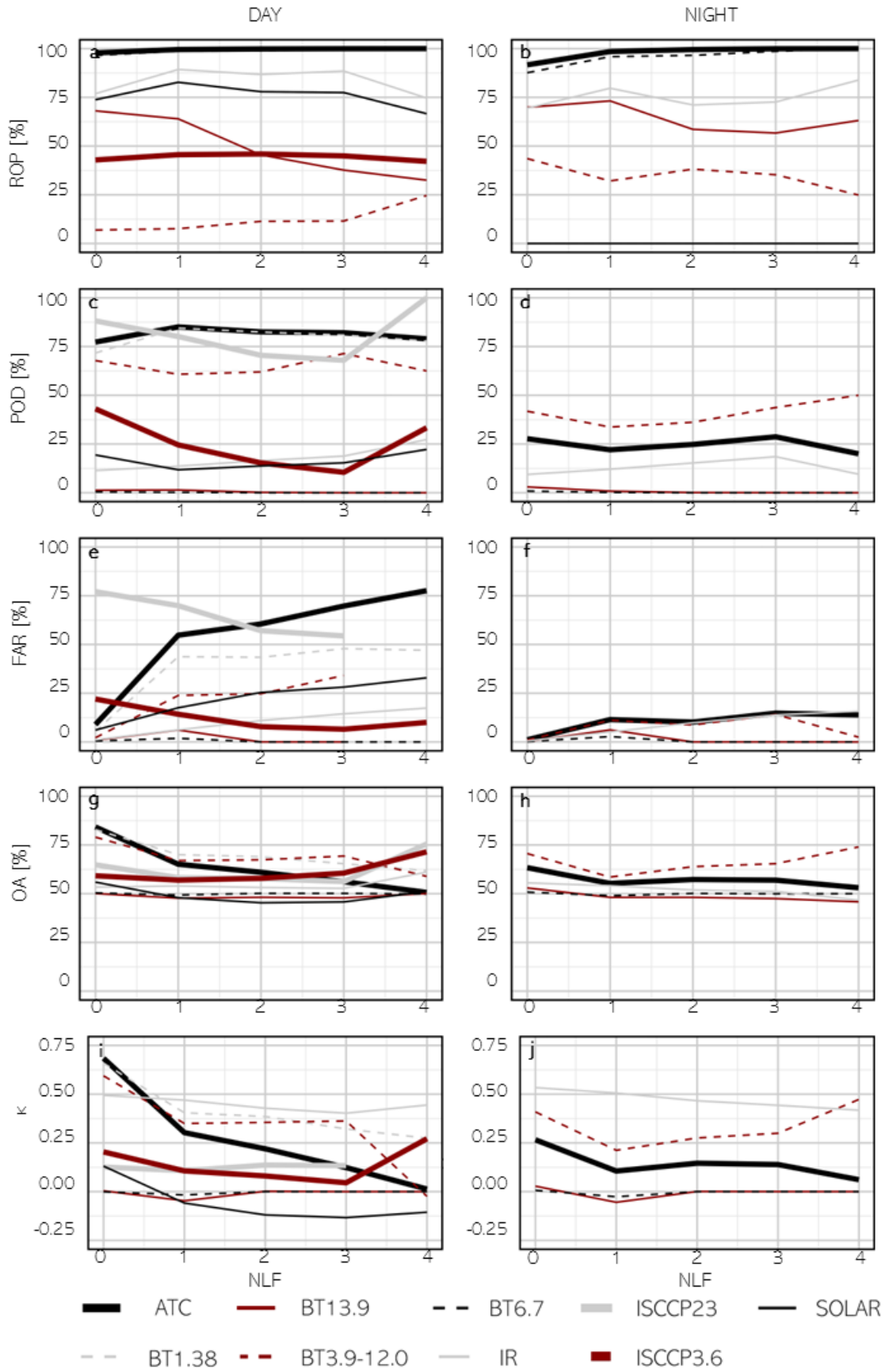
523 & 5j.).

524

525 *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 κkappa 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 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, Table 4).

CALIOP data products allow to report up to 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 κkappa. Specifically, as the number of non-cirrus layers increases, both OA and κkappa values decrease, for both day and night observations (Fig. 6g., Fig. 6h., Fig. 6i. & Fig. 6j.).

The International Geosphere–Biosphere Programme defines ecosystem surface classifications. For purpose of this study, 17 IGBP groups were aggregated to 3 classes: water, land and snow (goodness-of-fit with respect to land classification is presented in 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.

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

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

	Daytime					Nighttime				
	Accuracy measures									
	ROP [%]	POD	FAR	OA	Kk	ROP [%]	POD	FAR	OA	Kk
	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

563

564

565 The analysis with respect to NLF and land cover types confirmed that ATC test is best suited for

566 achieving the objective of this study. Therefore, the spatial distribution of the individual fit measures

567 for this test was examined (Fig. 7).

568 The spatial distribution reveals a very high ROP for both: daytime (Fig. 7a.) and nighttime (Fig. 7b.) for

569 the entire Earth. The southernmost regions of the Southern Hemisphere are an exception, exhibiting

570 lower values.

571 Spatial variations in correctly detected cirrus highlight differences between daytime and nighttime

572 POD distribution (Fig. 7c. & Fig. 7d.). During daytime, high values are observed over nearly the entire

573 Earth's surface, with exceptions in Antarctica, Greenland and the Himalayas (≥ 80% vs ≤ 20%

574 respectively), which are regions covered by snow and ice. However, at night, the highest difference is

575 between land and water (≥ 50% vs approximately 20%). Similar patterns to the POD distribution for

576 day and night can be observed in the OA results (Fig. 7g. - Fig. 7h.). On both sides of the equator, FAR

577 reaches the lowest values, being slightly higher during the day than at night (around 20% and ≤ 5%)

578 and increasing with latitude. However, there is a decrease in FAR observed in regions covered by snow

579 and ice (Fig. 7e. & Fig. 7f.). In regions with the highest rate of correctly detected and the lowest rate of

580 falsely reported cirrus, the general accuracy of classification (OA) exceeded 80% during daytime and

581 50% at night. Similar to OA, κkappa was higher during the day. During the day, κkappa values ranged

582 from 0.5 to 1.0 in regions at low latitudes. In contrast, at mid and high latitudes, κkappa values were

583 between 0.0 and 0.5, remaining positive (Fig. 7i.). At night (Fig. 7j.), nearly the entire Earth's surface

584 exhibited κkappa values between 0.0 and 0.5, with negative κkappa values observed near Micronesia.

585

586
587
588
589

DAY   NIGHT

591

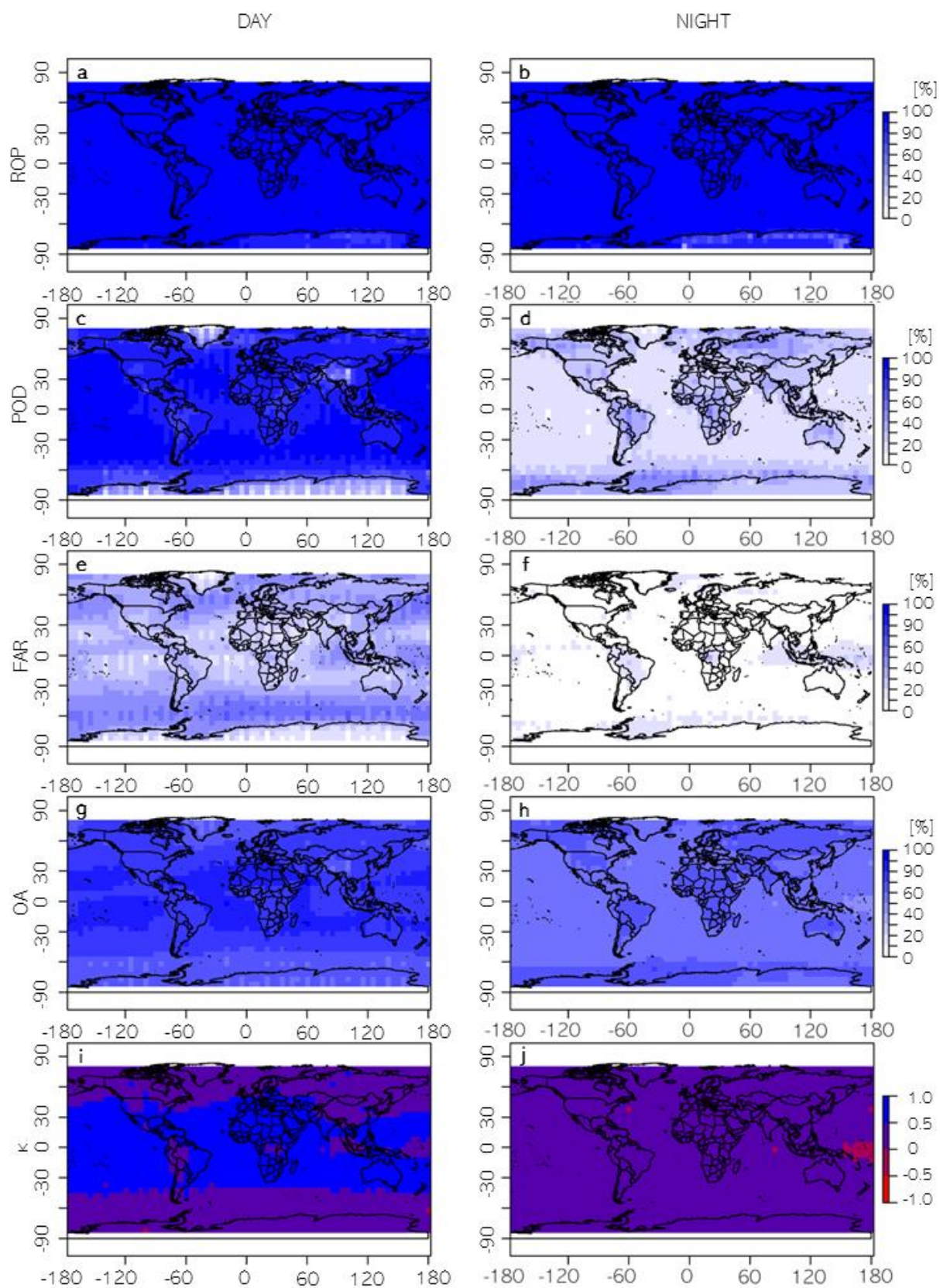



Fig. 7. Spatial distribution of the accuracy detection of cirrus using ATC test (letters (a, …, j) used to facilitate reference in the
text)

5. **Discussion**

This study demonstrated that MODIS operational cloud mask product can be used for producing a relatively accurate cirrus mask daytime (73% of overall accuracy, ~~K~~kappa = 0.46), and a questionable quality cirrus mask night-time (60% with ~~K~~kappa of only 0.2). We suggested the best approach to achieve such a goal, and reported related limitations, specifically for nighttime conditions.

During daytime, the two most effective tests were BT1.38 and ATC. With very similar parameters (POD, FAR, OA and ~~K~~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 a 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 of POD 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

625 ~~observations from MODIS.~~

626 The ISCCP provides a widely used framework for classifying clouds based on retrieved properties such

627 as optical thickness and cloud-top pressure. In this study, two ISCCP-based classifications were applied

628 using MODIS data: ISCCP3.6, which defines cirrus as clouds with optical thickness < 3.6 and top

629 pressure < 440 hPa, and ISCCP23, which extends the optical thickness threshold to < 23.

630 ISCCP3.6 showed moderate daytime performance but was limited by a relatively low ROP of 37.97%.

631 ISCCP23 achieved a high POD of 84.16%, but with a correspondingly high FAR of 72.00%, and a slightly

632 better OA of 61.26%. Both tests are only applicable to daytime data and share the same ROP.

633 These ISCCP-based classifications are included here as a reference framework, not as detection

634 algorithms. Furthermore, the ISCCP classifications used in this study are based on MODIS retrievals,

635 and therefore differ methodologically from early ISCCP climatologies, which relied on AVHRR-based

636 satellite observations.

637

638 Considering all mentioned above, the ATC test is proved to be the best among the available methods

639 for detecting high-level clouds. However, it is evident that its utility during nighttime is limited

640 compared to daytime. A notable factor contributing to this is the sensitivity of CALIOP. Lidar is known

641 to have significantly greater sensitivity at night, which explains its ability to detect nearly twice as many

642 cirrus clouds globally at night compared to daytime. This diurnal pattern in CALIOP data, while

643 highlighting the sensor's advantages in nighttime detection, should not be misinterpreted as a

644 definitive indicator of diurnal differences in cirrus cloud occurrence. Instead, it reflects the increased

645 detection capabilities of CALIOP at night.

646 Additionally, MODIS faces further limitations at night due to the unavailability of the 1.38 μm band (a

647 'cirrus band', introduced specifically to detect high, ice clouds; Bo-Cai Gao and Kaufman, 1995), which

648 is highly effective for detecting cirrus clouds during the day. As shown in the statistical analysis,

649 alternative tests exhibit significantly lower performance compared to the 1.38 μm band, emphasizing

650 its critical role in daytime cirrus cloud detection. This limitation further impacts the low effectiveness

651 of MODIS-based cirrus detection during nighttime observations.

652 Consequently, we have determined that the ATC test may be suitable for creating a cirrus cloud mask

653 and conducting a long-term climatological analysis of cirrus cloud coverage. This approach

654 simultaneously allows us to address the second research gap mentioned in this paragraph, which

655 concerns our lack of knowledge regarding the long-term variability of high-level cloud coverage.

656 Obtained from the CALIOP data, the cirrus mask mentioned in Section 3 allows us to investigate the

657 distribution of cirrus clouds (Fig. 2.) in 2015. Based on the CALIOP dataset, cirrus cloud coverage

658 reached 18.7% in 2015, daytime coverage of high-level clouds in 2015 was recorded at 13.2%, whereas

659 nighttime coverage was higher, measured at 23.3%. The day-night differences result from CALIOP's

660   higher nighttime sensitivity, reduced lidar signal noise, and increased nocturnal convective activity

661   leading to more cirrus formation. Additionally, annual variations in cloud amount (over 10 percentage

662   points) may occur due to CALIPSO's sampling frequency, as noted by Kotarba (2022).

663   Similarly, a cirrus mask was generated based on the MODIS data using the ATC test. Derived from this

664   data, cirrus cloud coverage (Fig. 8a.) showed daytime coverage of high-level clouds at 41.0%, while

665   nighttime coverage was lower, measured at 10.9% (Fig. 8b.).

666   We also compared cirrus cloud coverage in 2015 obtained from CALIOP and MODIS data (Fig. 9.). Each

667   point in Figure 9 represents mean annual cirrus cloud amount within 5-deg grid box, calculated from

668   all available observations within that grid cell. The mean difference between cirrus coverage derived

669   from CALIOP and MODIS was -27.71 p.p. for daytime observations (Fig. 9a.), with MODIS generally

670   indicated higher cloud cover compared to CALIOP. On the contrary, the mean difference between

671   cirrus coverage derived from CALIOP and MODIS was -12.31 p.p. for the nighttime observations (Fig.

672   9b.). A linear regression was performed to evaluate how well MODIS-derived cirrus coverage

673   corresponds to CALIOP values. While the relationship between MODIS and CALIOP is statistically

674   significant ($p < 0.001$), the R^2 value of 0.165 indicates that MODIS captures only 16.5% of the variability.

675   In the nighttime dataset, the R^2 improves to 0.422, meaning MODIS cloud coverage aligns better with

676   CALIOP at night. Additionally, we highlighted in blue the grid cells where the agreement between

677   MODIS and CALIOP cloud classification was moderate or higher, i.e., Cohen's ~~K~~kappa ≥ 0.5. During the

678   day, high-~~K~~kappa points tend to cluster in regions with low to moderate cirrus amounts. At night, the

679   distribution of high-~~K~~kappa points is more dispersed, indicating that even without the 1.38 µm band,

680   MODIS can achieve strong agreement with CALIOP in selected regions.

681   The divergence between pixel-level classification metrics and the aggregated cirrus cloud cover

682   comparisons arises from differences in MODIS detection performance between daytime and nighttime

683   conditions. During daytime, MODIS exhibits higher pixel-level agreement with CALIOP, as indicated by

684   elevated ~~K~~kappa and POD values. However, this is accompanied by a substantial false alarm rate,

685   largely attributable to inherited detections from the 1.38 µm cirrus test within the MODIS ATC. This

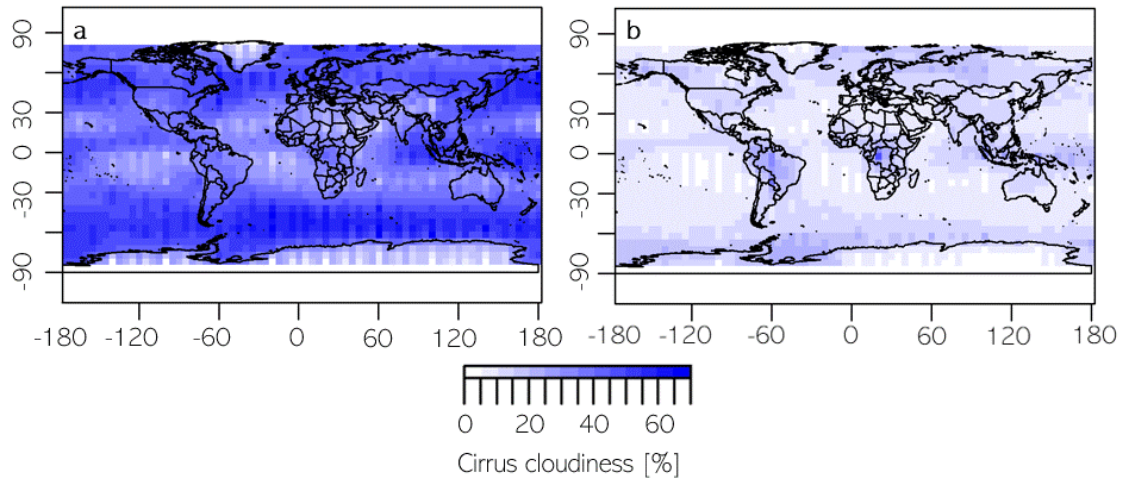
686   spectral test, while highly sensitive to high-level clouds, is prone to overestimations during daytime

687   due to factors such as surface reflection and sun-glint. Consequently, MODIS systematically

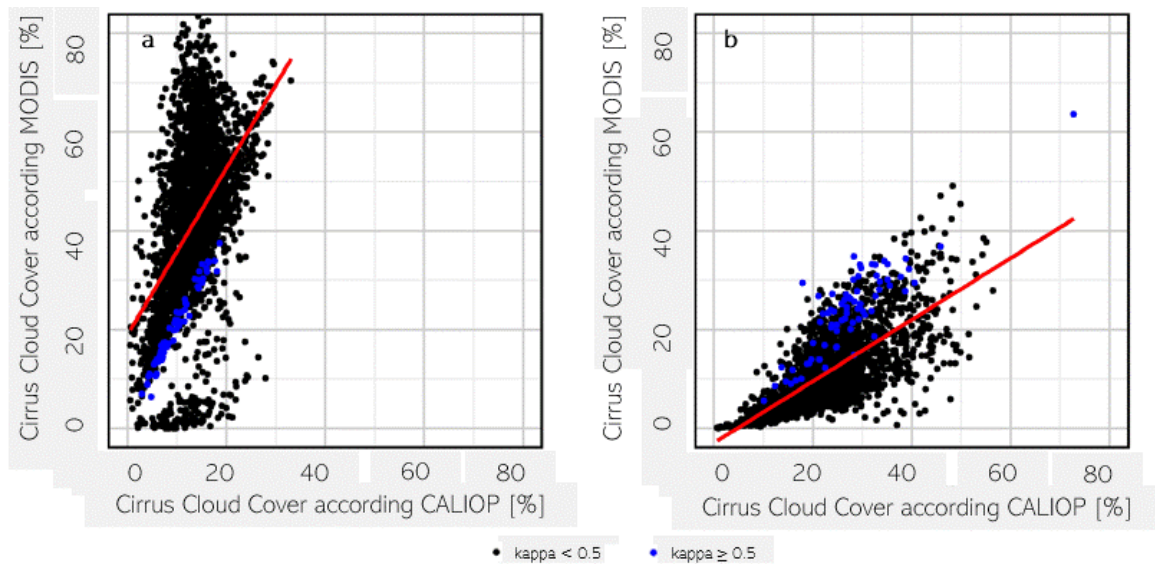
688   overestimates cirrus cover relative to CALIOP in daytime aggregated statistics. At night, the absence of

689   the 1.38 µm channel reduces the occurrence of false alarms, leading to a closer agreement in total

690   cirrus cover between MODIS and CALIOP, despite the weaker agreement observed at the pixel scale.

691

692   *Fig. 8. MODIS-based cirrus cloud coverage in 2015- daytime (a) and nighttime (b) derived with ATC approach of this study*

693

694

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

696   Our goal was to assess the extent to which selected tests of MODIS cloud mask can be used for

697   producing a cirrus mask, acknowledging that MODIS will miss a significant portion of cirrus due to the

698   sensor's lower sensitivity to optically thin clouds. ~~This comparison offers valuable insights into the~~

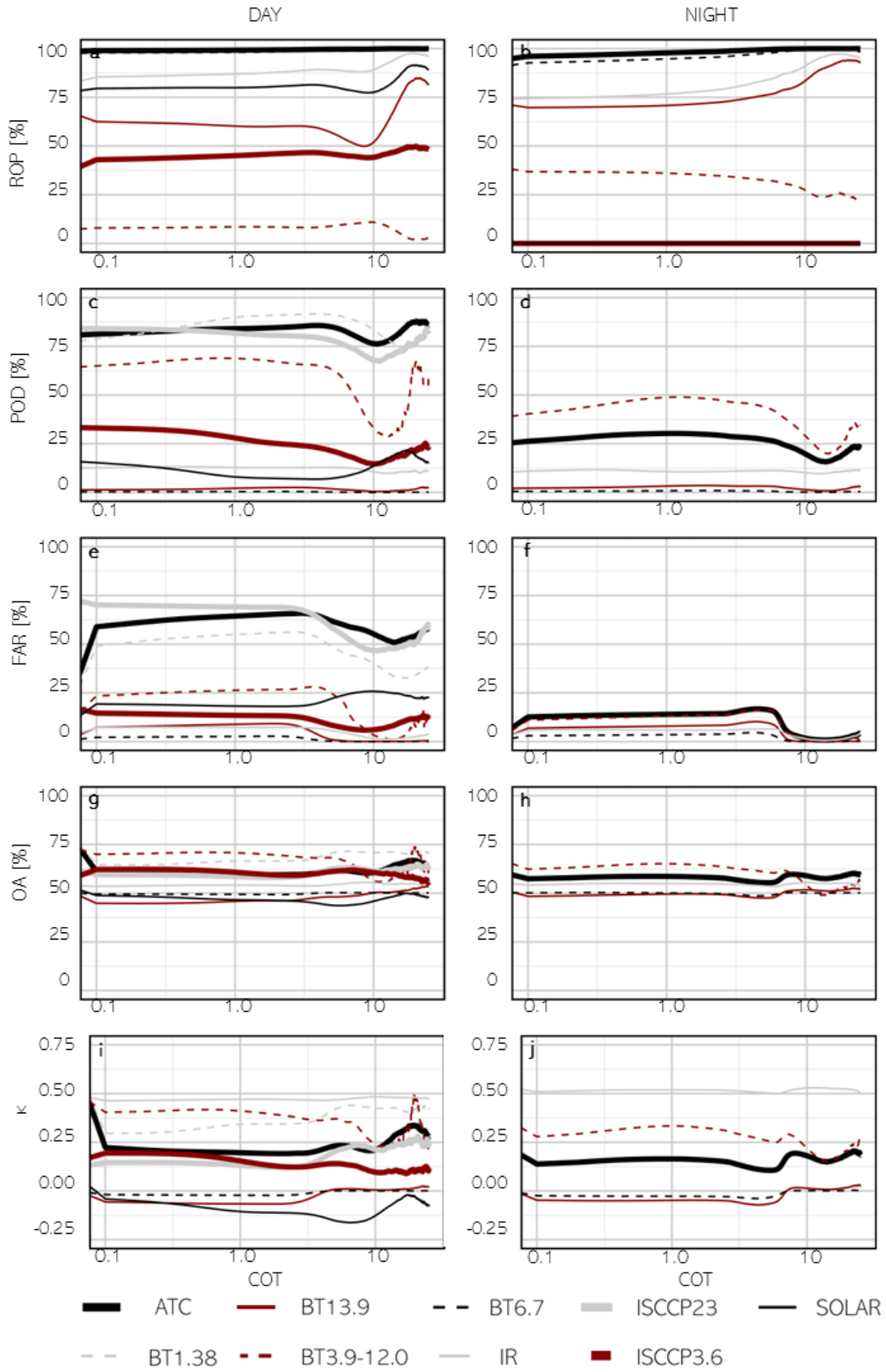
699   ~~practical efficiency of the MODIS instrument. We accepted MODIS data as it is; however,~~ In order to

700   discuss how the sensitivity possibly impacted our results we examined the MODIS-CALIOP agreement

701   as a function of COT. Information on the cirrus optical thickness was obtained from CALIOP data, and

702   the results are shown in Fig. 10.~~, as this primarily explains the differences between MODIS and CALIOP~~

703   ~~measurements.~~

704

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

706

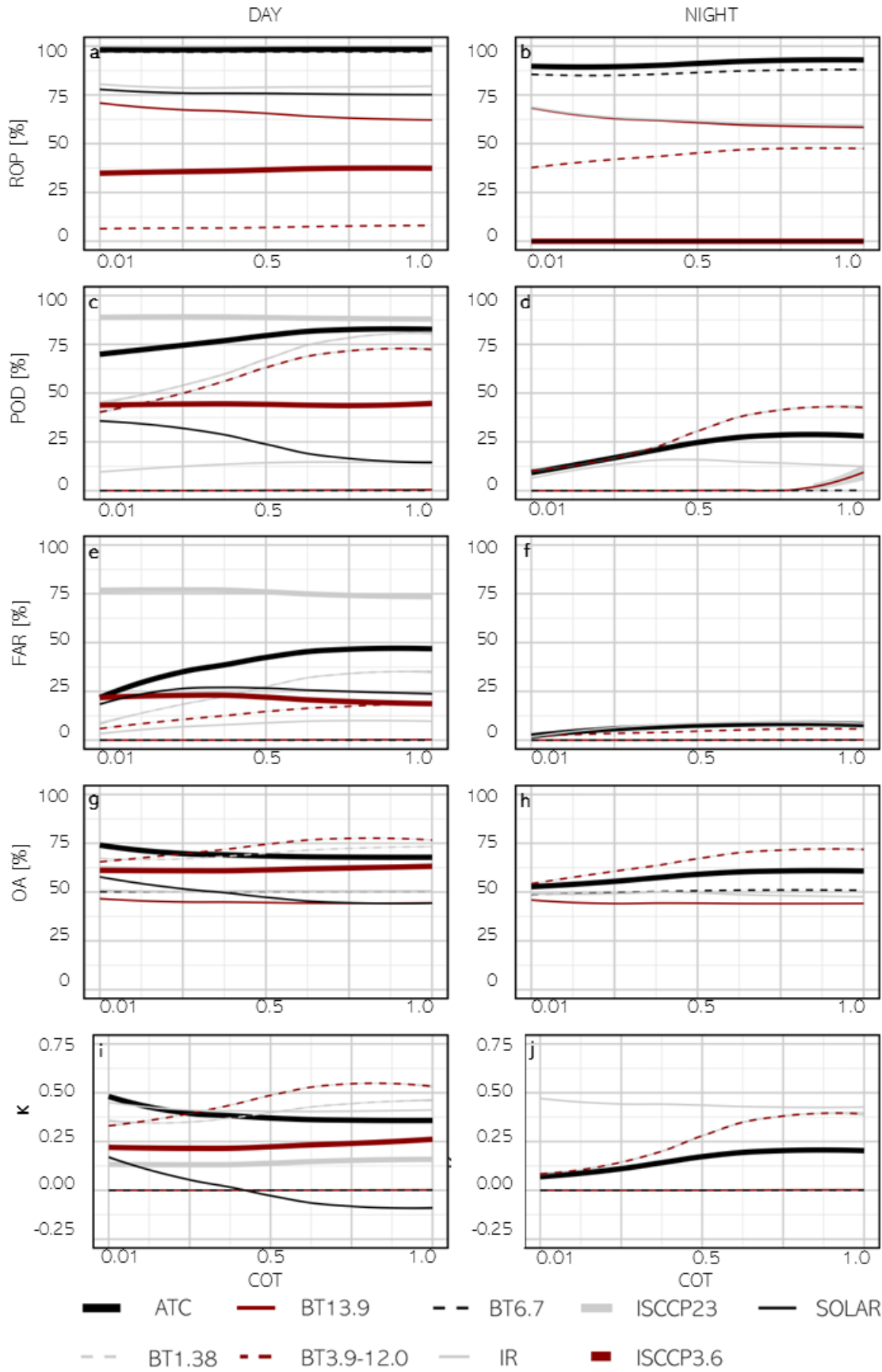
As observed in Fig. 10., there are no significant changes within the COT range of 0.1 to 1.0, and even up to 10.0. The most noticeable changes occur at COT values close to 10, though these may be influenced by the sample size, as the occurrence of cirrus clouds with a COT near 10 is limited or may represent a misclassification by CALIOP. At such high COT values, the lidar signal tends to be significantly attenuated, making accurate retrievals increasingly uncertain. This is because as the optical thickness increases, the lidar backscatter signal becomes weaker and may become too weak for precise measurements. Therefore, clouds with such high optical thickness may not be reliably detected by CALIOP, leading to potential misclassifications or missing data (Winker et al., 2024). In some cases, high COT values assigned to cirrus clouds may actually correspond to the cirrus-like top of a strong cumulonimbus cloud, which can be misclassified as cirrus due to the limitations of the CALIOP classification algorithm under conditions of strong signal attenuation. 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 κkappa 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. κKappa improves at night for all methods as COT increases but remains lower than daytime values. For daytime, κKappa 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

741   outperforms other approaches across all evaluated metrics, making it a reliable choice for cirrus

742   detection.

743
744 *Figure 11. Cirrus detection accuracy with respect to the COT (0-1) (letters (a, …, j) used to facilitate reference in the text)*

745

6. Summary

This study assessed the applicability of the MODIS operational cloud mask product (MYD35) for generating a cirrus cloud mask. To evaluate the accuracy of cirrus detection using the cloud mask tests, we employed a dataset comprising 136 million CALIOP lidar observations from the year 2015 as a reference. The analysis considered: six existing MODIS cloud tests (already reported in the Cloud Mask product), their combination (the ATC test, introduced by this study), and two methods originating from the ISCCP cloud classification scheme.

The key findings was that the ATC test is the most effective for detecting cirrus clouds:

- during daytime, it achieved a moderate reliability, confirmed by 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 κkappa coefficient of 0.46.

- at nighttime, its showed a low reliability, as proved by an overall accuracy of 59.50%, with a POD of 25.46%, FAR of 6.90%, and low κkappa 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.

Although this study is based on one year of data, the large sample size ensures statistical relevance. The ATC test demonstrates relatively high detection capability during daytime and acceptable agreement with CALIOP, but with noted limitations at night and for optically thin cirrus. While MODIS data are often used in cirrus climatologies due to their long-term consistency and global coverage, our findings suggest that cirrus detection within the MODIS cloud mask should be used with caution. The accuracy and reliability observed in this study indicate that the product's applicability to long-term trend analysis may be limited, depending on the specific requirements of the study. In the context of climate studies, the key consideration is not only the absolute accuracy of individual detections, but also the consistency of detection biases over time. Despite its limitations, for the daytime 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

779   The research has been supported by the National Science Center of Poland [grant number

780   2021/41/N/ST10/02274] and a grant from the Priority Research Area ("Anthropocene") under the

781   Strategic Programme Excellence Initiative at Jagiellonian University.

782   We gratefully acknowledge Poland's high-performance Infrastructure PLGrid ACK Cyfronet AGH for

783   providing computer facilities and support within computational grant no PLG/2024/016949.

784

785   Ackerman, S. A., Liou, K.-N., Valero, F. P. J., and Pfister, L.: Heating Rates in Tropical Anvils, J. Atmos.

786   Sci., 45, 1606–1623, 1988.

787   Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., and Gumley, L. E.:

788   Discriminating clear sky from clouds with MODIS, J. Geophys. Res. Atmos., 103, 32141–32157,

789   https://doi.org/10.1029/1998JD200032, 1998.

790   Ackerman, S. A., Holz, R. E., Frey, R., Eloranta, E. W., Maddux, B. C., and McGill, M.: Cloud detection

791   with MODIS. Part II: Validation, J. Atmos. Ocean. Technol., 25, 1073–1086,

792   https://doi.org/10.1175/2007JTECHA1053.1, 2008.

793   Amato, U., Antoniadis, A., Cuomo, V., Cutillo, L., Franzese, M., Murino, L., and Serio, C.: Statistical

794   cloud detection from SEVIRI multispectral images, Remote Sens. Environ., 112, 750–766,

795   https://doi.org/10.1016/j.rse.2007.06.004, 2008.

796   Baum, B. A., Menzel, W. P., Frey, R. A., Tobin, D. C., Holz, R. E., Ackerman, S. A., Heidinger, A. K., and

797   Yang, P.: MODIS cloud-top property refinements for collection 6, J. Appl. Meteorol. Climatol., 51,

798   1145–1163, https://doi.org/10.1175/JAMC-D-11-0203.1, 2012.

799   Behrangi, A., Nguyen, H., and Granger, S.: Probabilistic seasonal prediction of meteorological drought

800   using the bootstrap and multivariate information, J. Appl. Meteorol. Climatol., 54, 1510–1522,

801   https://doi.org/10.1175/JAMC-D-14-0162.1, 2015.

802   Bo-Cai Gao and Kaufman, Y. J.: Selection of the 1.375-µm MODIS channel for remote sensing of cirrus

803   clouds and stratospheric aerosols from space, J. Atmos. Sci., 52, 4231–4237,

804   https://doi.org/10.1175/1520-0469(1995)052<4231:sotmcf>2.0.co;2, 1995.

805   Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo,

806   Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S. K., Sherwood, S., Stevens, B., and Zhang, X.-Y.: 2013:

807   Clouds and Aerosols, in: Climate Change 2013: The Physical Science Basis. Contribution of Working

808   Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, vol.

809   9781107057, 571–658, https://doi.org/10.1017/CBO9781107415324.016, 2013.

810   Campbell, J. R., Lolli, S., Lewis, J. R., Gu, Y., and Welton, E. J.: Daytime cirrus cloud top-of-the-

811   atmosphere radiative forcing properties at a midlatitude site and their global consequences, J. Appl.

812   Meteorol. Climatol., 55, 1667–1679, https://doi.org/10.1175/JAMC-D-15-0217.1, 2016.

813   Chen, P. Y., Srinivasan, R., Fedosejevs, G., and Narasimhan, B.: An automated cloud detection method

814   for daily NOAA-14 AVHRR data for Texas, USA, Int. J. Remote Sens., 23, 2939–2950,

815   https://doi.org/10.1080/01431160110075631, 2002.

816   Cohen, J.: 1960: A Coefficient of Agreement for Nominal Scales. Educational and Psychological

817   Measurement, 20(1), 37-46. https://doi.org/10.1177/001316446002000104

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

819   Feng, X., Delsole, T., and Houser, P.: Bootstrap estimated seasonal potential predictability of global

820   temperature and precipitation, Geophys. Res. Lett., 38, 1–6, https://doi.org/10.1029/2010GL046511,

821   2011.

822   Frey, R. A., Ackerman, S. A., Liu, Y., Strabala, K. I., Zhang, H., Key, J. R., and Wang, X.: Cloud detection

823   with MODIS. Part I: Improvements in the MODIS cloud mask for Collection 5, J. Atmos. Ocean.

824   Technol., 25, 1057–1072, https://doi.org/10.1175/2008JTECHA1052.1, 2008.

825   Frey, R. A., Ackerman, S. A., Holz, R. E., Dutcher, S., and Griffith, Z.: The continuity MODIS-VIIRS cloud

826   mask, Remote Sens., 12, 1–18, https://doi.org/10.3390/rs12203334, 2020.

827   Gu, L., Ren, R., and Zhang, S.: Automatic cloud detection and removal algorithm for MODIS remote

828   sensing imagery, J. Softw., 6, 1289–1296, https://doi.org/10.4304/jsw.6.7.1289-1296, 2011.

829   Guenther, B., Xiong, X., Salomonson, V. V., Barnes, W. ., and Young, J.: On-orbit performance of the

830   Earth Observing System Moderate Resolution Imaging Spectroradiometer; first year of data, Remote

831   Sens. Environ., 83, 16–30, 2002.

832   Heidinger, A. K. and Pavolonis, M. J.: Gazing at cirrus clouds for 25 years through a split window. Part

833   I: Methodology, J. Appl. Meteorol. Climatol., 48, 1100–1116,

834   https://doi.org/10.1175/2008JAMC1882.1, 2009.

835   Holz, R. E., Ackerman, S. A., Nagle, F. W., Frey, R., Dutcher, S., Kuehn, R. E., Vaughan, M. A., and

836   Baum, B.: Global Moderate Resolution Imaging Spectroradiometer (MODIS) cloud detection and

837   height evaluation using CALIOP, J. Geophys. Res. Atmos., 114, 1–17,

838   https://doi.org/10.1029/2008JD009837, 2009.

839   Hutchinson, K. D. and Hardy, K. R.: Threshold functions for automated cloud analyses of global

840   meteorological satellite imagery, Int. J. Remote Sens., 16, 3665–3680,

841   https://doi.org/10.1080/01431169508954653, 1995.

842   Iida, Y., Kubota, T., Iguchi, T., and Oki, R.: Evaluating sampling error in TRMM/PR rainfall products by

843   the bootstrap method: Estimation of the sampling error and its application to a trend analysis, J.

844   Geophys. Res. Atmos., 115, 1–14, https://doi.org/10.1029/2010JD014257, 2010.

845   Jolliffe, I. T.: Uncertainty and inference for verification measures, Weather Forecast., 22, 637–650,

846   https://doi.org/10.1175/WAF989.1, 2007.

847   Kärcher, B.: Formation and radiative forcing of contrail cirrus, Nat. Commun., 9, 1–17,

848   https://doi.org/10.1038/s41467-018-04068-0, 2018.

849   Kinne, S. and Liou, K.-N.: The Effects of the Nonsphericity and Size Distribution of Ice Crystals on the

850   Radiative Properties of Cirrus Clouds, Atmos. Res., 24, 273–284, 1989.

851   Kotarba, A. Z.: Regional high-resolution cloud climatology based on MODIS cloud detection data, Int.

852   J. Climatol., 36, 3105–3115, https://doi.org/10.1002/joc.4539, 2016.

853   Kotarba, A. Z.: Calibration of global MODIS cloud amount using CALIOP cloud profiles, Atmos. Meas.

854   Tech., 13, 4995–5012, https://doi.org/10.5194/amt-13-4995-2020, 2020.

855   Kotarba, A. Z.: Impact of the revisit frequency on cloud climatology for CALIPSO, EarthCARE, Aeolus,

856   and ICESat-2 satellite lidar missions, Atmos. Meas. Tech., 15, 4307–4322,

857   https://doi.org/10.5194/amt-15-4307-2022, 2022.

858   Kotarba, A. Z. and Nguyen Huu, Ż.: Accuracy of Cirrus Detection by Surface-Based Human Observers,

859   J. Clim., 35, 3227–3241, https://doi.org/10.1175/JCLI-D-21-0430.1, 2022.

860   Liu, Y., Key, J. R., Frey, R. A., Ackerman, S. A., and Menzel, W. P.: Nighttime polar cloud detection with

861   MODIS, Remote Sens. Environ., 92, 181–194, https://doi.org/10.1016/j.rse.2004.06.004, 2004.

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

863   and Hostetler, C.: The CALIPSO lidar cloud and aerosol discrimination: Version 2 algorithm and initial

864   assessment of performance, J. Atmos. Ocean. Technol., 26, 1198–1213,

865   https://doi.org/10.1175/2009JTECHA1229.1, 2009.

866   Lolli, S., Campbell, J. R., Lewis, J. R., Gu, Y., Marquis, J. W., Chew, B. N., Liew, S. C., Salinas, S. V., and

867   Welton, E. J.: Daytime top-of-the-atmosphere cirrus cloud radiative forcing properties at Singapore, J.

868   Appl. Meteorol. Climatol., 56, 1249–1257, https://doi.org/10.1175/JAMC-D-16-0262.1, 2017.

869   Macke, A., Francis, P. N., Mcfarquhar, G. M., and Kinne, S.: The role of ice particle shapes and size

870   distributions in the single scattering properties of cirrus clouds, J. Atmos. Sci., 55, 2874–2883,

871   https://doi.org/10.1175/1520-0469(1998)055<2874:TROIPS>2.0.CO;2, 1998.

872   McGill, M. J., Vaughan, M. A., Trepte, C. R., Hart, W. D., Hlavka, D. L., Winker, D. M., and Kuehn, R.:

873   Airborne validation of spatial properties measured by the CALIPSO lidar, J. Geophys. Res. Atmos.,

874   112, 1–8, https://doi.org/10.1029/2007JD008768, 2007.

875   Menzel, W. P., Frey, R. A., and Baum, B. A.: Cloud Top Properties and Cloud Phase Algorithm

876   Theoretical Basis Document Collection 006 Update, 73, 2015.

877   Minnis, P., Trepte, Q. Z., Sun-Mack, S., Chen, Y., Doelling, D. R., Young, D. F., Spangenberg, D. A.,

878   Miller, W. F., Wielicki, B. A., Brown, R. R., Gibson, S. C., and Geier, E. B.: Cloud detection in nonpolar

879   regions for CERES using TRMM VIRS and Terra and Aqua MODIS data, IEEE Trans. Geosci. Remote

880   Sens., 46, 3857–3884, https://doi.org/10.1109/TGRS.2008.2001351, 2008.

881   Mishchenko, M. I., Rossow, W. B., Macke, A., and Lacis, A. A.: Sensitivity of cirrus cloud albedo,

882   bidirectional reflectance and optical thickness retrieval accuracy to ice particle shape, J. Geophys.

883   Res., 101, 16973–16985, 1996.

884   Murino, L., Amato, U., Carfora, M. F., Antoniadis, A., Huang, B., Menzel, W. P., and Serio, C.: Cloud

885   detection of modis multispectral images, J. Atmos. Ocean. Technol., 31, 347–365,

886   https://doi.org/10.1175/JTECH-D-13-00088.1, 2014.

887   Musial, J. P., Hüsler, F., Sütterlin, M., Neuhaus, C., and Wunderle, S.: Daytime low stratiform cloud

888   detection on AVHRR imagery, Remote Sens., 6, 5124–5150, https://doi.org/10.3390/rs6065124,

889   2014.

890   Nguyen Huu, Ż. and Kotarba, A. Z.: Reliability of visual detections of cirrus over Poland, Theor. Appl.

891   Climatol., 14, 1–11, 2021.

892   Noel, V., Chepfer, H., Chiriaco, M., and Yorks, J.: The diurnal cycle of cloud profiles over land and

893   ocean between 51° S and 51° N, seen by the CATS spaceborne lidar from the International Space

894   Station, Atmos. Chem. Phys., 18, 9457–9473, https://doi.org/10.5194/acp-18-9457-2018, 2018.

895   Oreopoulos, L., Cho, N., and Lee, D.: New Insights about Cloud Vertical Structure from CloudSat and

896   CALIPSO observations, J. Geophys. Res. Atmos., 122, 9280–9300,

897   https://doi.org/10.1002/2017JD026629.New, 2017.

898   Orlowsky, B., Bothe, O., Fraedrich, K., Gerstengarbe, F. W., and Zhu, X.: Future climates from bias-

899   bootstrapped weather analogs: An application to the Yangtze River basin, J. Clim., 23, 3509–3524,

900   https://doi.org/10.1175/2010JCLI3271.1, 2010.

901   Rossow, W. B. and Schiffer, R. A.: Advances in Understanding Clouds from ISCCP, Bull. Am. Meteorol.

902   Soc., 80, 2261–2287, https://doi.org/10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2, 1999.

903   Sassen, K., Wang, Z., and Liu, D.: Global distribution of cirrus clouds from CloudSat/cloud-aerosol

904   lidar and infrared pathfinder satellite observations (CALIPSO) measurements, J. Geophys. Res.

905   Atmos., 114, 1–12, https://doi.org/10.1029/2008JD009972, 2008.

906   Stanski, H. R., Wilson, L. J., and Burrows, W. R.: Survey of Common Verification Methods in

907   Meteorology, World weather watch technical report no 8, WMO/TD No 358, 1989.

908   Stephens, G. L. and Webster, P. J.: Clouds and Climate: Sensitivity of Simple Systems, J. Atmos. Sci.,

909   38, 235–247, 1981.

910   Stephens, G. L., Tsay, S. C., Stackhouse, P. W., and Flatau, P. J.: The relevance of the microphysical

911   and radiative properties of cirrus clouds to climate and climatic feedback,

912   https://doi.org/10.1175/1520-0469(1990)047<1742:trotma>2.0.co;2, 1990.

913   Stephens, G. L., Winker, D., Pelon, J., Trepte, C., Vane, D., Yuhas, C., L'Ecuyer, T., and Lebsock, M.:

914   Cloudsat and calipso within the a-train: Ten years of actively observing the earth system, Bull. Am.

915   Meteorol. Soc., 99, 569–581, https://doi.org/10.1175/BAMS-D-16-0324.1, 2018.

916   Stubenrauch, C. J., Cros, S., Guignard, A., and Lamquin, N.: A 6-year global cloud climatology from the

917   Atmospheric InfraRed Sounder AIRS and a statistical analysis in synergy with, Atmos. Chem. Phys.

918   Discuss., 15, https://doi.org/10.5194/acp-10-7197-2010, 2010.

919   Sun-Mack, S., Minnis, P., Chen, Y., Kato, S., Yi, Y., Gibson, S. C., Heck, P. W., and Winker, A. D. M.:

920   Regional apparent boundary layer lapse rates determined from CALIPSO and MODIS data for cloud-

921   height determination, J. Appl. Meteorol. Climatol., 53, 990–1011, https://doi.org/10.1175/JAMC-D-

922   13-081.1, 2014.

923   Tang, H., Yu, K., Hagolle, O., Jiang, K., Geng, X., and Zhao, Y.: A cloud detection method based on a

924   time series of MODIS surface reflectance images, Int. J. Digit. Earth, 6, 157–171,

925   https://doi.org/10.1080/17538947.2013.833313, 2013.

926   Thorsen, T. J., Fu, Q., Comstock, J. M., Sivaraman, C., Vaughan, M. A., Winker, D. M., and Turner, D.

927   D.: Macrophysical properties of tropical cirrus clouds from the CALIPSO satellite and from ground-

928   based micropulse and Raman lidars, J. Geophys. Res. Atmos., 118, 9209–9220,

929   https://doi.org/10.1002/jgrd.50691, 2013.

930   Vaughan, M. A., Powell, K. A., Kuehn, R. E., Young, S. A., Winker, D. M., Hostetler, C. A., Hunt, W. H.,

931   Liu, Z., Mcgill, M. J., and Getzewich, B. J.: Fully automated detection of cloud and aerosol layers in the

932   CALIPSO lidar measurements, J. Atmos. Ocean. Technol., 26, 2034–2050,

933   https://doi.org/10.1175/2009JTECHA1228.1, 2009.

934   Wang, C., Luo, Z. J., and Huang, X.: Parallax correction in collocating CloudSat and Moderate

935   Resolution Imaging Spectroradiometer (MODIS) observations: Method and application to convection

936   study, J. Geophys. Res. Atmos., 116, 1–9, https://doi.org/10.1029/2011JD016097, 2011.

937   Wang, T., Fetzer, E. J., Wong, S., Kahn, B. H., and Yue, Q.: Validation of MODIS cloud mask and

938   multilayer fl ag using CloudSat-CALIPSO cloud profiles and a cross-reference of their cloud

939   classifications, J. Geophys. Res., 121, https://doi.org/10.1038/175238c0, 2016.

940   Wilks, D. S., Neumann, C. J., and Lawrence, M. B.: Statistical extension of the National Hurricane

941   Center 5-day forecasts, Weather Forecast., 24, 1052–1063,

942   https://doi.org/10.1175/2009WAF2222189.1, 2009.

943   Winker, D., Cai, X., Vaughan, M., Garnier, A., Magill, B., Avery, M., and Getzewich, B.: A Level 3

944   monthly gridded ice cloud dataset derived from 12 years of CALIOP measurements, Earth Syst. Sci.

945   Data, 16, 2831–2855, https://doi.org/10.5194/essd-16-2831-2024, 2024.

946   Winker, D. M., Hostetler, C. A., Vaughan, M., and Omar, A.: CALIOP Algorithm Theoretical Basis

947   Document Part 1 : CALIOP Instrument, and Algorithms Overview, 2006.

948   WMO: International Cloud Atlas, Volume I: Manual on the Observation of Clouds and Other Meteors,

949   180 pp., https://doi.org/10.2307/1550553, 1977.

950   Xie, Y., Qu, J. J., and Xiong, X.: Improving the CALIPSO VFM product with Aqua MODIS measurements,

951   Remote Sens. Lett., 1, 195–203, https://doi.org/10.1080/01431161003720387, 2010.

952   Zhang, Y., Laube, M., and Raschke, E.: Numerical simulations of cirrus properties, Contrib. to Atmos.

953   Phys., 67, 109–120, 1994.

954 Zhang, Y., MacKe, A., and Albers, F.: Effect of crystal size spectrum and crystal shape on stratiform

955 cirrus radiative forcing, Atmos. Res., 52, 59–75, https://doi.org/10.1016/S0169-8095(99)00026-5,

956 1999.

957 Zou, L., Griessbach, S., Hoffmann, L., Gong, B., and Wang, L.: Revisiting global satellite observations of

958 stratospheric cirrus clouds, Atmos. Chem. Phys., 20, 9939–9959, https://doi.org/10.5194/acp-20-

959 9939-2020, 2020.

960