



Article

Near-Ultraviolet to Near-Infrared Band Thresholds Cloud Detection Algorithm for TANSAT-CAPI

Ning Ding^{1,2}, Jianbing Shao¹, Changxiang Yan^{1,3,*}, Junqiang Zhang^{1,4}, Yanfeng Qiao¹, Yun Pan^{1,2}, Jing Yuan^{1,2}, Youzhi Dong^{1,2} and Bo Yu^{1,2}

- ¹ Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China; dingning17@mails.ucas.ac.cn (N.D.); shaojianbing@ciomp.ac.cn (J.S.); zhangjq@ciomp.ac.cn (J.Z.); qiaoyf@ciomp.ac.cn (Y.Q.); panyun15@mails.ucas.ac.cn (Y.P.); yuanjing16@mails.ucas.ac.cn (J.Y.); dongyouzhi18@mails.ucas.ac.cn (Y.D.); yubo19@mails.ucas.ac.cn (B.Y.)
- ² University of the Chinese Academy of Sciences, Beijing 100049, China
- ³ Center of Materials Science and Optoelectrics Engineering, University of Chinese Academy of Sciences, Beijing 100049, China
- ⁴ Ysense Information Technology and Equipment (Qingdao) Inc., Qingdao 266000, China
- * Correspondence: yanxc@ciomp.ac.cn; Tel.: +86-186-4307-5317



Citation: Ding, N.; Shao, J.; Yan, C.; Zhang, J.; Qiao, Y.; Pan, Y.; Yuan, J.; Dong, Y.; Yu, B. Near-Ultraviolet to Near-Infrared Band Thresholds Cloud Detection Algorithm for TANSAT-CAPI. *Remote Sens.* **2021**, *13*, 1906. <https://doi.org/10.3390/rs13101906>

Academic Editor: Beatriz M. Funatsu

Received: 15 April 2021

Accepted: 9 May 2021

Published: 13 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Cloud and aerosol polarization imaging detector (CAPI) is one of the important payloads on the China Carbon Dioxide Observation Satellite (TANSAT), which can realize multispectral polarization detection and accurate on-orbit calibration. The main function of the instrument is to identify the interference of clouds and aerosols in the atmospheric detection path and to improve the retrieval accuracy of greenhouse gases. Therefore, it is of great significance to accurately identify the clouds in remote sensing images. However, in order to meet the requirement of lightweight design, CAPI is only equipped with channels in the near-ultraviolet to near-infrared bands. It is difficult to achieve effective cloud recognition using traditional visible light to thermal infrared band spectral threshold cloud detection algorithms. In order to solve the above problem, this paper innovatively proposes a cloud detection method based on different threshold tests from near ultraviolet to near infrared (NNDT). This algorithm first introduces the 0.38 μm band and the ratio of 0.38 μm band to 1.64 μm band, to realize the separation of cloud pixels and clear sky pixels, which can take advantage of the obvious difference in radiation characteristics between clouds and ground objects in the near-ultraviolet band and the advantages of the band ratio in identifying clouds on the snow surface. The experimental results show that the cloud recognition hit rate ($\text{POD}_{\text{cloud}}$) reaches 0.94 (ocean), 0.98 (vegetation), 0.99 (desert), and 0.86 (polar), which therefore achieve the application standard of CAPI data cloud detection. The research shows that the NNDT algorithm replaces the demand for thermal infrared bands for cloud detection, gets rid of the dependence on the minimum surface reflectance database that is embodied in traditional cloud recognition algorithms, and lays the foundation for aerosol and CO_2 parameter inversion.

Keywords: cloud detection; TANSAT-CAPI; NNDT; near ultraviolet band 380 nm; thresholds

1. Introduction

CO_2 is one of the greenhouse gases on earth, and the increase in its concentration over the last century seriously affects the environment of human survival [1]. The IPCC report in 2014 shows that the radiative forcing effect of aerosols is the largest source of uncertainty in climate change assessment. Monitoring and evaluating the spatial distribution and parameters of both CO_2 and aerosols have become a matter of great concern for scientists. Therefore, research and improvement of aerosol and CO_2 parameter retrieval schemes have been carried out all over the world. Large-scale observation of the earth can be achieved by satellite remote sensing, which has become one of the main ways to detect CO_2 and aerosol [2].

The China Carbon Dioxide Observation Satellite (TANSAT) project is the first global atmospheric carbon dioxide observation scientific experimental satellite developed entirely by China. It was successfully launched by the CZ-2D carrier rocket at 3:22 on 22 December 2016. TANSAT is a sun-synchronous polar orbiting satellite with an orbital altitude of 712 km. There are two key instruments on the satellite: carbon dioxide sensor (CDS) and cloud and aerosol polarization imager (CAPI) [3]. CAPI aims to identify cloud and aerosol interference in the atmospheric to more accurately retrieve CO₂ parameters. CAPI has a spatial resolution of 250 m/1000 m, and a calibration accuracy of better than 5%. There are two modes of sub-satellite point observation and flare observation on it, between which the sub-satellite point observation mode is not affected by the solar flare. CAPI is equipped with five near ultraviolet to near infrared bands (0.38, 0.67, 0.87, 1.375, and 1.64 μm), as shown in Table 1, and realizes the small-scale lightweight design. Similar to the cloud and aerosol imager (TANSO-CAI) on the greenhouse gas observation satellite (GOSAT) [4], CAPI has an additional 1.375 μm band for cirrus detection. CAPI adds polarization channels at 0.67 and 1.64 μm , respectively [5]. Its swath width is 400 km, which provides a large enough area for detecting aerosol spatial distribution and cloud cover.

Table 1. Wavelength and wavelength range of each CAPI channel.

Channel		Center Wavelength	Wavelength Coverage
		μm	μm
Near-ultraviolet	Band1	0.38	0.365–0.408
	Visible	Band2	0.67 (P)
Near-infrared		Band3	0.87
	Band4	1.375	1.36–1.39
	Band5	1.64 (P)	1.628–1.654

Note: (P) indicates that the band is polarized.

However, most remote sensing images contain clouds [6]. Cloud is an interference factor for atmospheric retrieval, especially the thin cirrus, which is difficult to detect, and its existence will greatly affect the retrieval accuracy of aerosol and carbon dioxide parameters [7]. Therefore, in order to minimize the impact, accurate cloud detection plays an important role in image preprocessing.

Since the advent of the satellite era, many cloud recognition algorithms have been developed. The cloud recognition algorithms are mainly divided into three categories: machine learning method [8–18], cloud texture and spatial features method [19–23], and spectral thresholds method [24–30]. The research status of these methods is classified and summarized in Table 2.

With the increasing application of machine learning in the field of remote sensing, it provides another effective means for cloud detection. However, because the machine learning method usually requires a large number of training samples for model construction, the workload is large, and the universality of the model is often not effectively guaranteed. The method of cloud texture and spatial features performs cloud detection based on the spatial information features of the image. The cloud has a unique texture feature that distinguishes it from the background, which is reflected in the spatial variation of the spectral brightness value. Due to the complexity of this algorithm and the large amount of calculation, it is difficult to meet the needs of high-efficiency calculations. The basic principle of the spectral thresholds method of cloud recognition is to rely on the spectrum difference between the cloud and the earth's surface in the remote sensing image. It is a physical method to distinguish between the cloud and the earth's surface by setting the radiance (reflectivity) thresholds. Due to its simple model and fast calculation speed, this method can meet the needs of batch image processing. Therefore, this research is

based on the spectral thresholds method to design a cloud detection algorithm suitable for CAPI data.

Table 2. Summary of three classes of cloud detection methods.

Optical Model	Author	Key Results
Machine learning method	Yu et al. (2006)	Automatic detection of cloud was realized using a clustering method according to texture features.
	Wang et al. (2010)	A K-means clustering algorithm was used to classify the clustering feature of the thresholds for reflectivity data.
	Jin et al. (2016)	A back propagation was established in a neural network for MODIS data, which has efficient learning ability.
	Yu Oishi et.al (2017)	Analyzed the impact of different support vectors on GOSAT-2 CAI-2 L2 cloud discrimination.
	Fu et al. (2019)	The cloud detection on FY meteorological satellite images was carried out using Random Forest Method.
Texture and spatial features method	Lee et al. (1990)	Texture cues were utilized to recognize clouds by their high spatial heterogeneity.
	Liu (2007)	The cloud detection on MODIS data was carried out using a classification matrix and dynamic clustering algorithm of shortest distance.
	Sun et al. (2017)	Gauss filtering was used to generate a cloud mask for Landsat.
	Liu et al. (2017)	A gray level co-occurrence matrix was used to extract the texture features of clouds.
Spectral thresholds method	Saunders et al. (1987)	A series of spectral thresholds was used to process the NOAA AVHRR sensor data.
	Ishida et al. (2009)	An unbiased cloud detection algorithm, which referred to as CLAUDIA, was developed for visible-to-infrared imagers and introduced the clear confidence level (CCL).
	Wang et al. (2017)	The scheme for CAPI classifies each spectral test as either clear- or cloud-conservative according to the CCL value derived from the individual test, and final CCL is determined by combining the results from these two groups.
	Quan et al. (2020)	Developed a cloud detection approach based on hybrid multispectral features with dynamic thresholds for GF-1 remote sensing images.
	This study	Proposed a cloud detection approach based on near-ultraviolet to near-infrared band with thresholds for different underlying surfaces (NNDT).

However, there are a few researches on cloud recognition algorithm for CAPI at present. Only one scheme based on Chinese FengYun-3A Polar-orbiting Meteorological Satellite data was proposed by Wang et al. before satellite launch [30]. Although the cloud and aerosol unbiased decision intellectual algorithm (CLAUDIA) has been proved to be applicable to all remote sensing instrument data with corresponding visible to thermal infrared bands [27], it also has limitations. For CAPI, due to design limitation, the channels that can be applied to the CLAUDIA method only contain four bands from visible to near-infrared, which is difficult to achieve the accuracy of cloud detection algorithms using thermal infrared bands. In addition, using 0.67 and 0.87 μm bands for cloud detection needs

the support of minimum surface reflectance database to ensure the inversion accuracy, while the use of the minimum surface reflectance database has its own limitations such as causing image interpretation errors.

To solve the above problems, this paper proposes a more effective cloud detection method for CAPI, which is a thresholds detection method based on different underlying surface thresholds from near ultraviolet to near infrared (NNDT). For the first time, 0.38 μm and the ratio of 0.38 to 1.64 μm bands are used in cloud detection. First, the algorithm designed unique band test combinations for the four surface types of ocean, vegetation, desert, and polar regions. Subsequently, CAPI data have been collected over the four surface types to establish threshold values for the cloud detection spectral tests, and the corresponding fixed thresholds are determined. Finally, in order to verify and evaluate the effectiveness of the NNDT algorithm, the cloud recognition results of the algorithm and the official cloud recognition results of the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Second-generation Global Imager (SGLI) are compared visually and quantitatively, respectively. Experiments prove that the algorithm can overcome the above limitations and can effectively identify the cloud.

This paper is organized as follows. In Section 2, the NNDT algorithm is introduced, including the characteristics analysis of wavelength and wavelength combination and the introduction of algorithm flow (including: preprocessing process, algorithm design for different underlying surfaces, and the statistics and determination of each threshold). Section 3 gives the verification results of cloud detection. Finally, the experimental results are discussed in Section 4, and in Section 5, the conclusions are given and the future work is prospected.

2. Materials and Methods

The spectral thresholds method is based on the radiation difference between cloud and underlying surfaces reflectance in satellite images. The spectral characteristics used by NNDT algorithm include: (a) reflectance of solar radiance, (b) dependence of reflectance on wavelength, and (c) reflectance of solar radiance at the absorption wavelength. The algorithm uses one near ultraviolet (0.38 μm), one visible (0.67 μm), and three near-infrared (0.87, 1.375, and 1.64 μm) CAPI channels. The following section will explain the detailed radiation characteristics and how to apply them to cloud/clear sky discrimination in CAPI images.

2.1. Spectral Reflection Characteristics of Cloud and Underlying Surface

2.1.1. Reflectance of Solar Radiance

In the non-absorption visible and near-infrared bands, the clouds have higher reflectance value than that of clear sky surfaces [30]. Based on this fact, 0.67 μm reflectance is used for vegetated land and coastal regions, and 0.87 μm reflectance is used over water scenes in many cloud detection algorithms [31,32]. However, in the 0.67 and 0.87 μm bands, the reflectivity of different land surface types varies greatly in different seasons, so the reflectivity thresholds of these bands also change [33]. In order to ensure the effectiveness of cloud recognition in these two bands, it is usually necessary to count the minimum surface reflectance varying with seasons or even months, while CAPI lacks the minimum surface reflectance with timeliness and full coverage. If an external database is used, such as MODIS minimum albedo product MCD43A3 or the minimum albedo map proposed by Ishida and Nakajima [26], the observation geometry matching and spectral response function matching of data need to be considered. In addition, if the minimum albedo is affected by optically thin clouds, the actual optically thin clouds will not be recognized.

Based on the above situation, this article uses 0.38 μm near-UV band of CAPI as an alternative. Before this, there was no research using 0.38 μm band for cloud recognition. Yu Oishi et al. analyzed the annual variation of reflectivity of each band of CAI load for several typical land surface types [13], indicating that the reflectivity of 0.38 μm band remains at a low value and basically unchanged throughout the year except that it increases due to the

influence of ice and snow. From the spectral reflectance of cloud and underlying surface in Figure 1, it can be seen that the surface of soil, desert, and vegetation has lower reflectance in the 0.38 μm band than in the 0.67 and 0.87 μm band, the reflectance of water is still at a low level, and the reflectance of cloud is still at a high level. Moreover, the reflectivity of aerosol is very low in the 0.38 μm band [34]. In terms of cloud characteristics, the edge details of low and middle clouds in the 0.38 μm satellite image are very obvious. Figure 2 shows the CAPI grayscale images of the 0.38, 0.67, and 0.87 μm bands in the Australian semi-vegetated and desert. Therefore, the reflectance of 0.38 μm band is more beneficial to distinguish between cloud and clear sky surface on soil, desert, and vegetation underlying surface. In general, the reflectance of water and vegetation in the 0.38 μm band is at extremely low level of less than 0.05, while for the dry soil (i.e., bare soil) and desert surface, the reflectance is higher, between 0.1 and 0.3, there is a possibility of misidentification as a cloud. Because ice and snow have the same high reflectivity as cloud at 0.38 μm , simply using 0.38 μm cannot distinguish the cloud from the snow and ice surface.

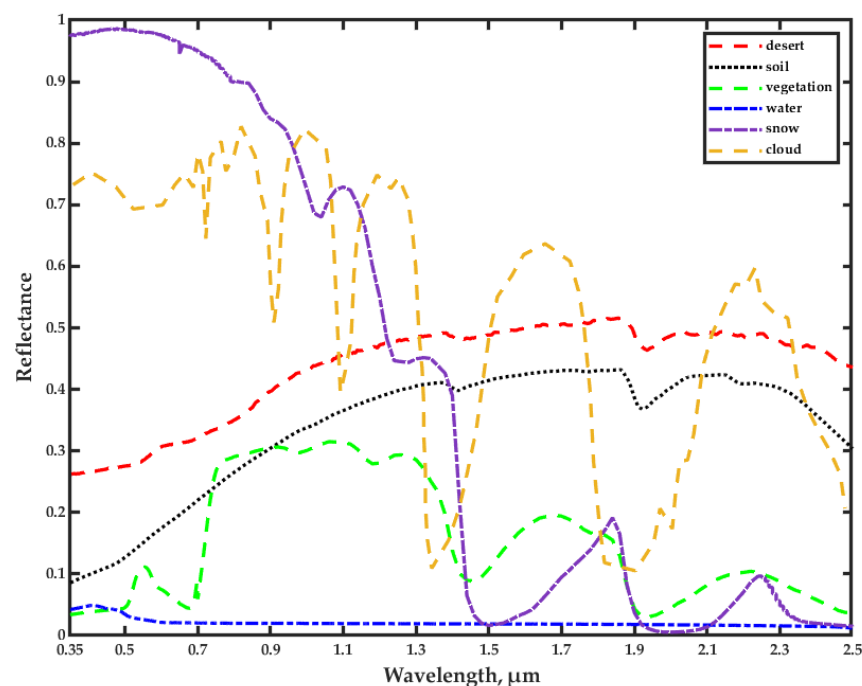


Figure 1. Reflectance curve of cloud and underlying surface. The underlying surface spectrum comes from ENVI spectral library [35], and the cloud spectrum comes from Airborne Visible Infrared Imaging Spectrometer.

2.1.2. Dependence of Reflectance on Wavelength

From visible to near-infrared in the solar radiation region, without considering the absorption valley region, cloud reflectivity does not change much with wavelength. On the contrary, the reflectance of several kinds of ground surface changes with the change of wavelength. Therefore, the difference (or ratio) of the reflectance at different wavelengths can well indicate whether the pixel is cloud or not [26]. In deserts or areas with sparse vegetation, the reflectivity in visible light region changes very little. In the near-infrared region, the reflectivity of deserts tends to increase with increasing wavelength [36]. Therefore, the reflectance ratio of 0.87 and 1.64 μm can better solve the problem of cloud identification over desert areas, and the ratio is expressed as:

$$R_f(0.87 \mu\text{m}) / R_f(1.64 \mu\text{m}) \quad (1)$$

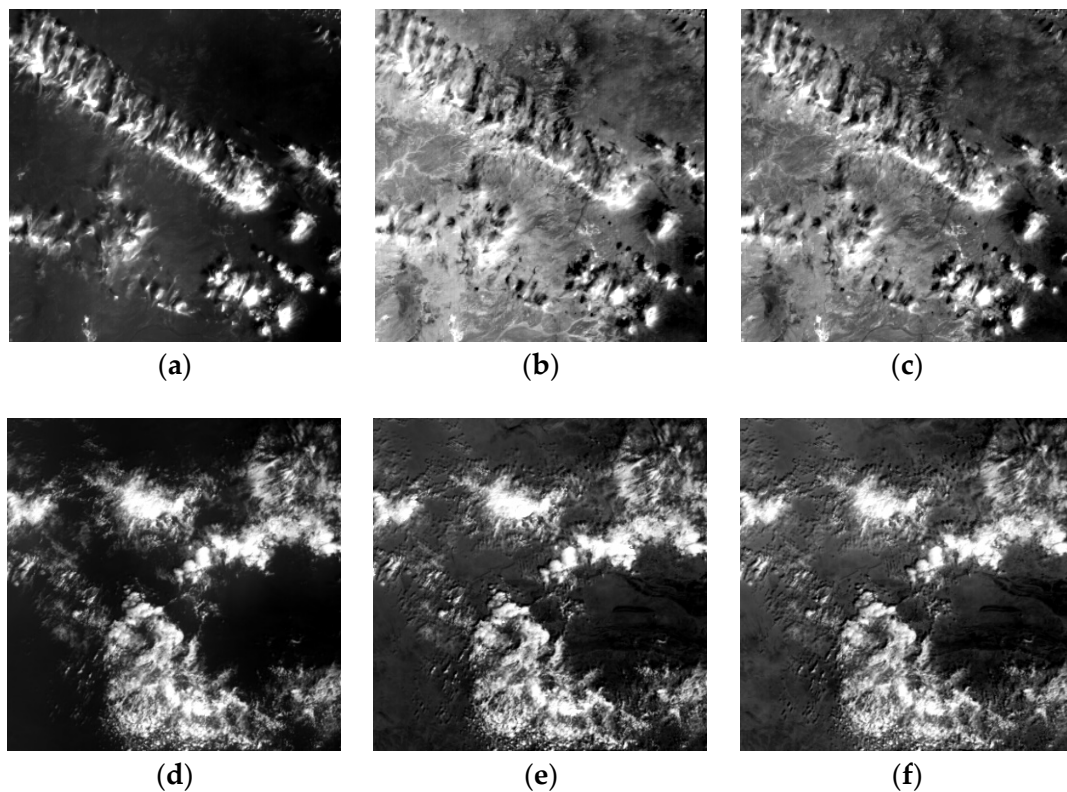


Figure 2. CAPI images over Australia on 29 March 2017. (a–c): Images of semi-vegetated areas in southern Australia; (d–f): Images of desert areas in central Australia. (a,d): Images of 0.38 μm band; (b,e): Images of 0.67 μm band; (c,f): Images of 0.87 μm band.

But for bright surfaces, such as snow, it is not that effective.

Normalized difference snow index (NDSI) is used to distinguish snow cover areas [37–39]. The NDSI, which has a larger value on snow surface than for cloud, is determined by the reflectance in 0.67 μm bands and 1.64 μm bands of CAPI [30]:

$$\text{NDSI} = \frac{R_f(0.67 \mu\text{m}) - R_f(1.64 \mu\text{m})}{R_f(0.67 \mu\text{m}) + R_f(1.64 \mu\text{m})} \quad (2)$$

In the visible and infrared regions, cloud shadow has very low reflectivity, while in the visible region, cloud reflective bright edge has very high reflectivity. As a result, the NDSI values in the cloud shadow area are very close to those in the cloud area. In addition, in the cloud reflective bright edge, the NDSI values are very close to those in the snow surface area. In this way, it is easy to mistakenly identify the cloud shadow as cloud, and mistakenly identify the cloud reflective bright edge as clear sky. For this reason, this paper proposes a new method of using the reflectance ratio of 0.38 μm band to 1.64 μm band to make up for this fault, as shown in Equation (3).

$$R_f(0.38 \mu\text{m}) / R_f(1.64 \mu\text{m}) \quad (3)$$

The values of cloud shadow and cloud reflective bright edge in 0.38 μm band are not too low or too high. Therefore, the ratio of 0.38 μm band to 1.64 μm band in cloud shadow is higher than that of snow surface, and in cloud reflective bright edge, it is equivalent to that of cloud pixel. The ratio of 0.38 μm band to 1.64 μm band is lower at cloud pixel than in snow pixel. By setting threshold, cloud on snow surface can be effectively identified.

2.1.3. Reflectance of Solar Radiance at The Absorption Wavelength

High clouds, such as cirrus, can be detected by solar radiation at absorption wavelengths [40]. The $1.375 \mu\text{m}$ channel is in the water vapor absorption zone, and the radiation reflected by the surface or low altitude clouds is almost completely absorbed by the abundant water vapor beneath the cloud. When high altitude clouds and cirrus clouds mainly composed of ice crystals exist, the reflected radiation intensity at $1.375 \mu\text{m}$ band will increase. However, $1.375 \mu\text{m}$ band detection is not applicable at high altitude areas and polar regions, where the atmosphere is thin, absorbing less and reflecting more radiation from the underlying surface. This band is therefore only designated to detect clouds at surface altitudes of less than 2000 meters.

2.2. CAPI Spectral Image Cloud Detection Process

Figure 3 shows the flow chart of NNDT algorithm using CAPI data. There are four steps, including solar altitude angle and sun-earth distance correction, snow surface predetection, determination of land cover type, and cloud detection for four underlying surfaces. The first three of them are pre-processing work. The fourth step is the key part of the NNDT algorithm.

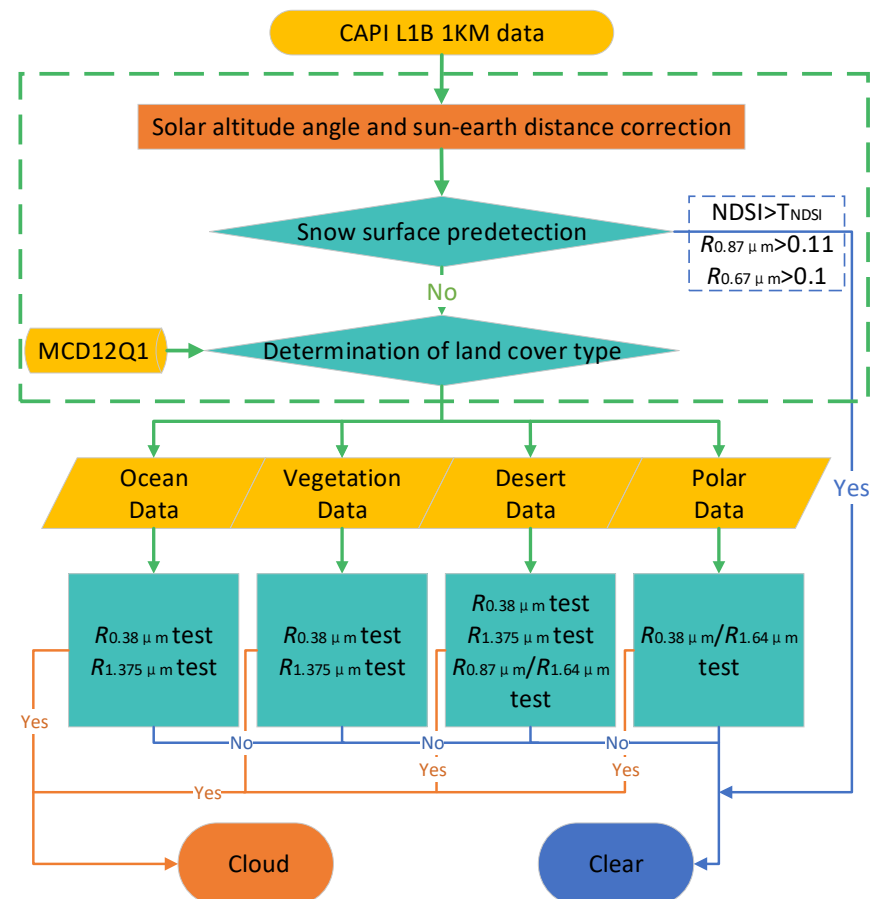


Figure 3. Flow chart of NNDT algorithm using CAPI data.

2.2.1. Preprocessing

(1) Solar altitude and sun-earth distance correction

The data used in this paper are obtained from CAPI L1B (1000 M) V2.0 sub-satellite point observation mode (tracking the main plane) scientific observation products and geometric positioning products. The L1B data of CAPI is reflectance data, which is generated from the radiometric calibrated DN (digital number) value. Before the data can be used

for cloud identification, the sun to earth distance correction and the solar altitude angle correction are also needed.

$$R^* = (\pi \times L) / E_{\text{sun}} \quad (4)$$

$$R = (R^* \times d^2) / \cos A. \quad (5)$$

R^* is L1B data, R is reflectivity data, L is the radiance data after radiometric calibration, E_{sun} is the solar constant, d is the distance between the sun and the earth, $\cos A$ is the cosine of the solar zenith angle (unit: radians). The sun-earth distance d and the solar zenith angle A are stored in the geometric data of L1B.

(2) Snow surface pre-detection

Scattered snow cover may be misidentified as cloud, therefore identifying snow cover in advance is required before cloud detection. Identifying the snow surface in advance can simplify the subsequent cloud detection procedure and reduce the probability of misjudging the snow as a cloud. The following tests are used:

$$NDSI > T_{NDSI} \quad (6)$$

$$R_f(0.87 \mu\text{m}) > 0.11 \quad (7)$$

$$R_f(0.67 \mu\text{m}) > 0.1. \quad (8)$$

T_{NDSI} is the detection threshold. Based on [41,42], we mark a pixel as snow when the NDSI is higher than 0.48 (0.6) during the Northern Hemisphere warm (cold) season, i.e., April to September (October to March); notice that warm and cold months are switched for the Southern Hemisphere (SH), with the reflectivity of 0.87 and 0.67 μm higher than 0.11 and 0.10, respectively.

(3) Determination of land cover type

Another important step is to determine the land cover type. The land classification data were derived from the MODIS global land cover data (MCD12Q1) with a resolution of 500 m, which is matched to CAPI geometric data according to the longitude and latitude coordinates. In this paper, all underlying surface types are divided into four types: ocean, vegetation, desert, and polar, so as to implement the subsequent targeted cloud detection for different underlying surface types. Different from the traditional classification method, the inland water is classified as ocean in this paper, and the desert and bare surface are taken out as an independent category from the land category, and the rest types in the land category are collectively referred to as vegetation.

2.2.2. Cloud Recognition Algorithms for Different Surface Types

Cloud/clear discrimination is done by comparing the reflectance from the CAPI to the thresholds, which defines the boundary between cloud and clear pixels. NNDDT algorithm consists of different threshold tests to ensure detection accuracy, in consideration that a threshold test that is effective for a certain surface type may not be appropriate for another type. Different cloud detection schemes 1–4 are designed for different underlying surface types, as shown in Formulas (9)–(16). T is the threshold of corresponding detection. For ocean and vegetation surface cloud recognition, we use the union of $R_f(0.38 \mu\text{m})$ and $R_f(1.375 \mu\text{m})$ threshold detection results. On the desert surface, we use the union of $R_f(1.375 \mu\text{m})$ threshold detection results with the intersection of $R_f(0.38 \mu\text{m})$ and $R_f(0.87 \mu\text{m})/R_f(1.64 \mu\text{m})$ threshold detection results. In polar regions, we only use $R_f(0.38 \mu\text{m})/R_f(1.64 \mu\text{m})$ threshold detection for cloud recognition. As shown in Table 3, different thresholds are used for 0.38 and 1.375 μm band detection on different underlying surfaces.

Table 3. Thresholds for cloud screening with CAPI.

Threshold Test	Thresholds			
	Ocean	Vegetation	Desert	Polar
$R_f(0.38 \mu\text{m})$	0.08	0.15	0.25	–
$R_f(1.38 \mu\text{m})$	0.011	0.019	0.030	–
$R_f(0.87 \mu\text{m})/R_f(1.64 \mu\text{m})$	–	–	0.95	–
$R_f(0.38 \mu\text{m})/R_f(1.64 \mu\text{m})$	–	–	–	4.25

1. Ocean

$$R_f(0.38 \mu\text{m}) > T_{\text{ocean-0.38}} \quad (9)$$
 or

$$R_f(1.375 \mu\text{m}) > T_{\text{ocean-1.375}} \quad (10)$$
2. Vegetation

$$R_f(0.38 \mu\text{m}) > T_{\text{veg-0.38}} \quad (11)$$
 or

$$R_f(1.375 \mu\text{m}) > T_{\text{veg-1.375}} \quad (12)$$
3. Desert

$$R_f(0.38 \mu\text{m}) > T_{\text{desert-0.38}} \quad (13)$$
 and

$$R_f(0.87 \mu\text{m})/R_f(1.64 \mu\text{m}) > T_{\text{desert-0.87-1.64}} \quad (14)$$
 or

$$R_f(1.375 \mu\text{m}) > T_{\text{desert-1.375}} \quad (15)$$
4. Polar

$$R_f(0.38 \mu\text{m})/R_f(1.64 \mu\text{m}) > T_{\text{polar-0.38-1.64}} \quad (16)$$

When the above conditions are met, the pixel is identified as a cloud.

The thresholds used in all cloud tests in this study are obtained by CAPI data statistics, rather than the thresholds set by other cloud recognition algorithms for the band characteristics of their sensors [43]. The reflectance data of the Indian Ocean (35°–40° S, 120°–130° E), southern Africa (7°–10° S, 23°–26° E), Sahara Desert (15°–18° N, 12°–17° E), and Antarctica (68°–72° S, 148°–158° E) in 2017, which as representative areas of the four surface types, were collected for threshold statistics. All the data are corrected by solar altitude angle and sun earth distance. Considering whether the thresholds change with the season, we select all the images in the same area and at the same time with an 18-day revisit cycle from March 3 in a year to carry out visual cloud recognition by setting the detection thresholds. Because clouds with different heights, types, and optical thicknesses have different apparent reflectance [44,45]. For the purpose of identifying all possible clouds, a threshold is selected between the surface reflectance of clear sky and the reflectance of the cloud with the lowest optical thickness, which is close to the surface reflectance of clear sky. All threshold results are shown in Table 4.

Table 4. The statistical results of cloud detection thresholds every 18 days from March to December 2017 for the representative regions of four land surface types.

Data	Thresholds							
	Ocean		Vegetation		Desert		Polar	
	0.38	1.375	0.38	1.375	0.38	1.375	0.87/1.64	0.38/1.64
3 March	0.09	0.012	0.15	0.019	0.26	0.032	0.94	4.25
21 March	0.09	0.011	0.15	0.020	0.26	0.031	0.94	4.24
8 April	0.08	0.011	0.16	0.019	0.25	0.031	0.93	4.23
26 April	0.09	0.013	0.16	0.020	0.25	0.031	0.94	4.24
14 May	0.09	0.012	0.20	0.025	0.29	0.036	0.92	4.14
7 July	0.08	0.012	0.16	0.021	0.26	0.032	0.95	4.19
25 July	0.08	0.013	0.16	0.021	0.26	0.032	0.95	4.20
12 August	0.08	0.013	0.16	0.021	0.26	0.031	0.95	4.22
30 August	0.09	0.012	0.20	0.024	0.26	0.031	0.94	4.22
5 October	0.08	0.011	0.15	0.020	0.22	0.030	0.92	4.24
23 October	0.08	0.012	0.1	0.021	0.25	0.026	0.95	4.50
10 November	0.08	0.011	0.15	0.020	0.26	0.030	0.95	4.23
28 November	0.09	0.011	0.12	0.019	0.25	0.031	0.94	4.25
16 December	0.07	0.008	0.15	0.017	0.24	0.028	0.93	4.19

Note: CAPI did not have satellite imagery in June and September.

It can be seen from the thresholds listed in the table that there is no significant change in the reflectance thresholds of the same land surface type in a year; but for different land surface types, the thresholds are quite different. Therefore, after excluding the obvious deviation value (irregular maximum or minimum) in the thresholds of the same land surface type, the smallest (large) value in the remaining data is used as the final cloud detection threshold of the corresponding land surface type, as shown in Table 3.

3. Results

Verifying the effectiveness of cloud recognition algorithms is difficult. Two important steps in validation are visual comparison and quantitative analysis [46–48]. In visual comparison, an analyst conducts a validation through visual inspection of the spectral, spatial, and temporal features in a set of images. Visual inspection is an important first step in validating any cloud mask algorithm. The analyst uses knowledge and experience of cloud and land surface spectral properties to identify obvious problems. However, visual comparison provides poor quantitative evaluation [25]. More quantitative validation can be attained through directly comparing the two results pixel by pixel. This section provides two schemes for verifying NNDT algorithm, which are visual comparison with MODIS cloud identification products and quantitative comparison with SGLI cloud identification products. Some validation examples will be given at the same time.

3.1. Visual Comparison with MODIS Cloud Detection Production

NNDT algorithm was evaluated by visually comparing cloud/clear discrimination results (referred to as the CAPI cloud flag) with the CAPI composited RGB image and the MODIS cloud-mask product (MYD35). MYD35 is a cloud-mask product for solar synchronous orbit data from MODIS (onboard Aqua) at 13:30 local solar time, close to TANSAT local solar time, making it possible to find images of CAPI and MODIS in the same place on certain dates with small transit intervals. However, their swath and revisit cycles are different: the CAPI has a swath of 400 km and a revisit cycle of 18 days. The

Aqua/MODIS has a swath of 2330 km and a revisit cycle of 16 days. The orbit of Aqua overlaps with that of CAPI only intermittently.

MYD35 uses clear confidence level (CCL) to evaluate cloud detection results, including four confidence levels, which are: cloudy, probably cloudy, probably clear, and confidently clear [49,50]. In order to correspond with the output cloud flag “either cloud or clear” of the NNDT algorithm, the “confidently clear” and “probably clear” of MYD35 are classified as clear, and the other confidence levels are classified as cloud. CAPI and MODIS data, as well as cloud flag, are projected onto grids with a spatial resolution of 1 km × 1 km.

To clarify the characteristics of CAPI Cloud Flag, five scenes were selected (Figures 4–8) with different cloud properties and ground surface characteristics. The locations are the northwest coast of Australia and Indian ocean (15°–17.5° S, 118°–123° E, ocean), the southwest part of Australia (27°–32° S, 122°–127° E, bare area), the islands of Indonesia (2.5°–7° N, 114.5°–119° E, vegetation), the Sahara Desert (21°–24.5° N, 12°–15° E, desert), and the Antarctic continent (76°–81° S, 112°–123° W, polar). The Antarctic continental scene experiment used CAPI image at 00:30 and MODIS image at 01:15 on 1 March 2017. Other scenarios used data from 26 April 2017, when the track of two instruments crossed within 5 min of each other. In order to facilitate comparative observation, ENVI software is used to clip MODIS and CAPI images after geometric correction to obtain the overlapping area images. In polar regions, the geometrically corrected satellite images have a very small width, which is not conducive to observation. Therefore, images without geometric correction are used, and the approximate geographic range is marked.

Each scene has four panels: (a) the CAPI false color image synthesized by the reflectance data of the 0.87 μm (Red), 0.67 μm (Green), 0.38 μm (Blue) channels for the scene; (b) the cloud flag image of CAPI derived by applying NNDT algorithm to the CAPI reflectance data; (c) the MODIS true-color image combining bands 1, 3, and 4 of multiple MODIS level-1b products overlapping with the CAPI scene; and (d) the cloud flag of MYD35 with only “cloud” and “clear” results.

As can be seen from the cloud identification of the northwest coast of Australia and Indian ocean area in Figure 4, the cloud area contours of CAPI cloud flag and MYD35 are very similar. Small clouds around the cloud cluster in the CAPI RGB image are also well recognized. Compared with MYD35, the cloud regions identified by NNDT algorithm are more coherent at the edge of the clouds and the hole in the center of the cloud cluster, and these regions are identified as clear in MYD35.

Figure 5 shows images taken on Indonesian islands, including clouds at different heights and optical thicknesses over both surfaces of the ocean and vegetation. NNDT algorithm can recognize all kinds of cloud as cloud pixels, such as small scattered clouds, thick clouds, and optically thin clouds. The cloud area of CAPI cloud flag and MYD35 is basically the same.

Figure 6 illustrates the results of cloud detection over bare surface. Like MYD35, the NNDT algorithm does not determine the locally highlighted sandbanks, which are easily identified as clouds. MYD35 sometimes recognizes inland intermittent waters as clouds, and the NNDT algorithm avoids this situation well. By comparing the CAPI cloud flag with the CAPI pseudo color image, it can be seen that NNDT algorithm can determine the distribution position of small scattered clouds. For the recognition of small scattered clouds, the cloud shape of CAPI cloud flag is more discrete, while that of MYD35 is more continuous.

Figure 7 shows a typical desert scene, Sahara Desert scene. Its clear sky surface pixels have high reflectance, which can be easily identified as clouds. Nevertheless, the CAPI cloud flag is reasonable compared with the CAPI RGB image and MYD35.

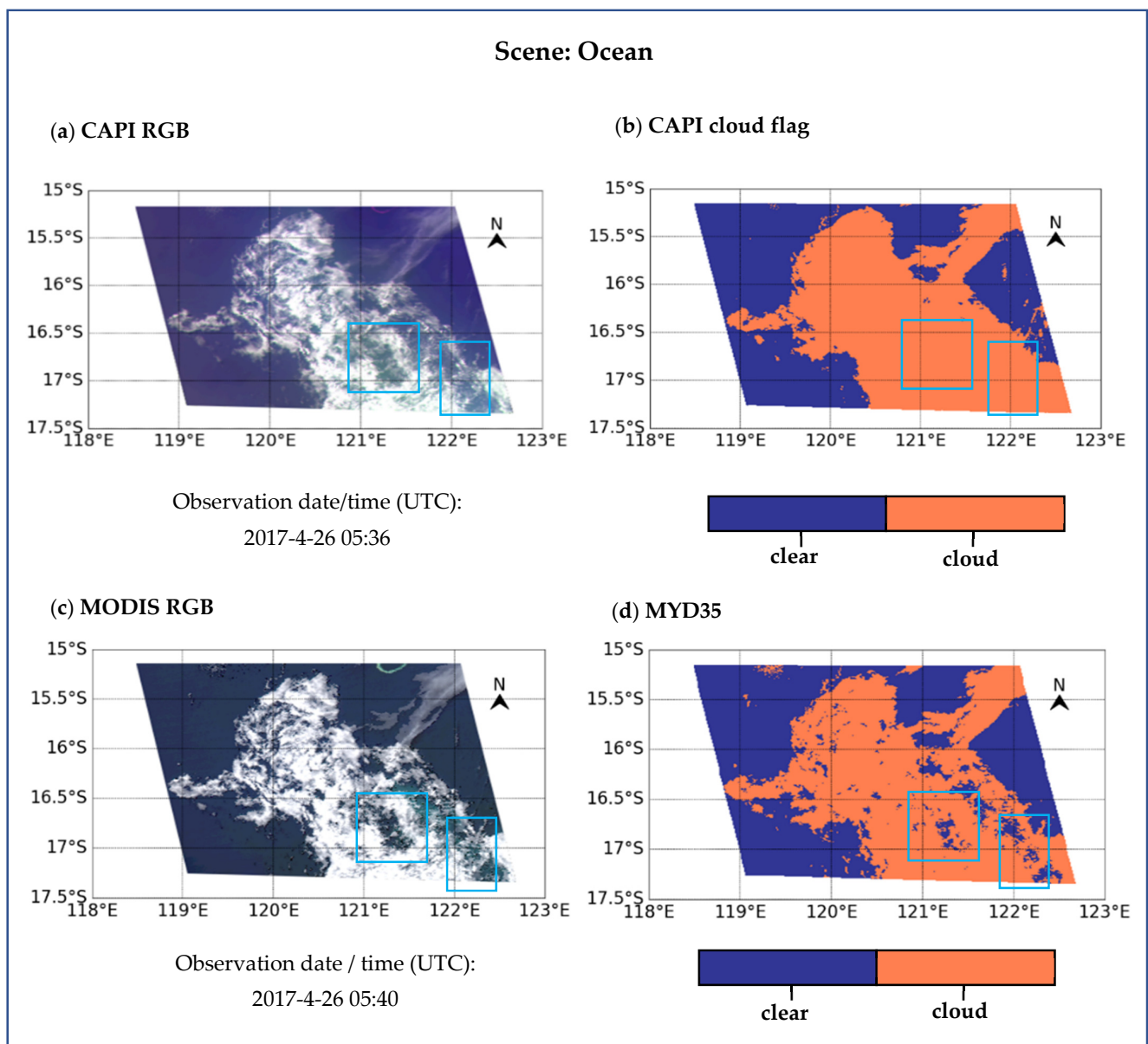


Figure 4. Case over the northwest coast of Australia and Indian ocean at 05:36 UTC 26 April 2017 for CAPI and at 05:40 UTC 26 April 2017 for MODIS. The “time” is the start time when the imager took the image of this scene. (a) The CAPI false color image, (b) the cloud flag image of CAPI derived by applying NNDT algorithm, (c) the MODIS true-color image overlapping with the CAPI scene, and (d) the cloud flag of MYD35 with only “cloud” and “clear” results.

The cloud detection results for a scene over the Antarctic continent are shown in Figure 8. For pixels covered with snow and/or sea ice, it is generally difficult to distinguish clear sky and clouds because they have a visual feature and spectrum similar to clouds, namely, white with high reflectance. Comparison with the CAPI RGB image suggests that our algorithm can accurately distinguish between clear sky and cloudy areas over the bright snow and ice surface. The NNDT algorithm identifies the cloud anti bright edge as cloud and the cloud shadow as clear sky. It shows the sensitivity and accuracy of NNDT algorithm in distinguishing cloud anti bright edge from surface and cloud shadow from cloud. In addition, for pixels with high and thin cloud distribution, MYD35 is identified as clear sky, while NNDT algorithm can accurately identify them as clouds.

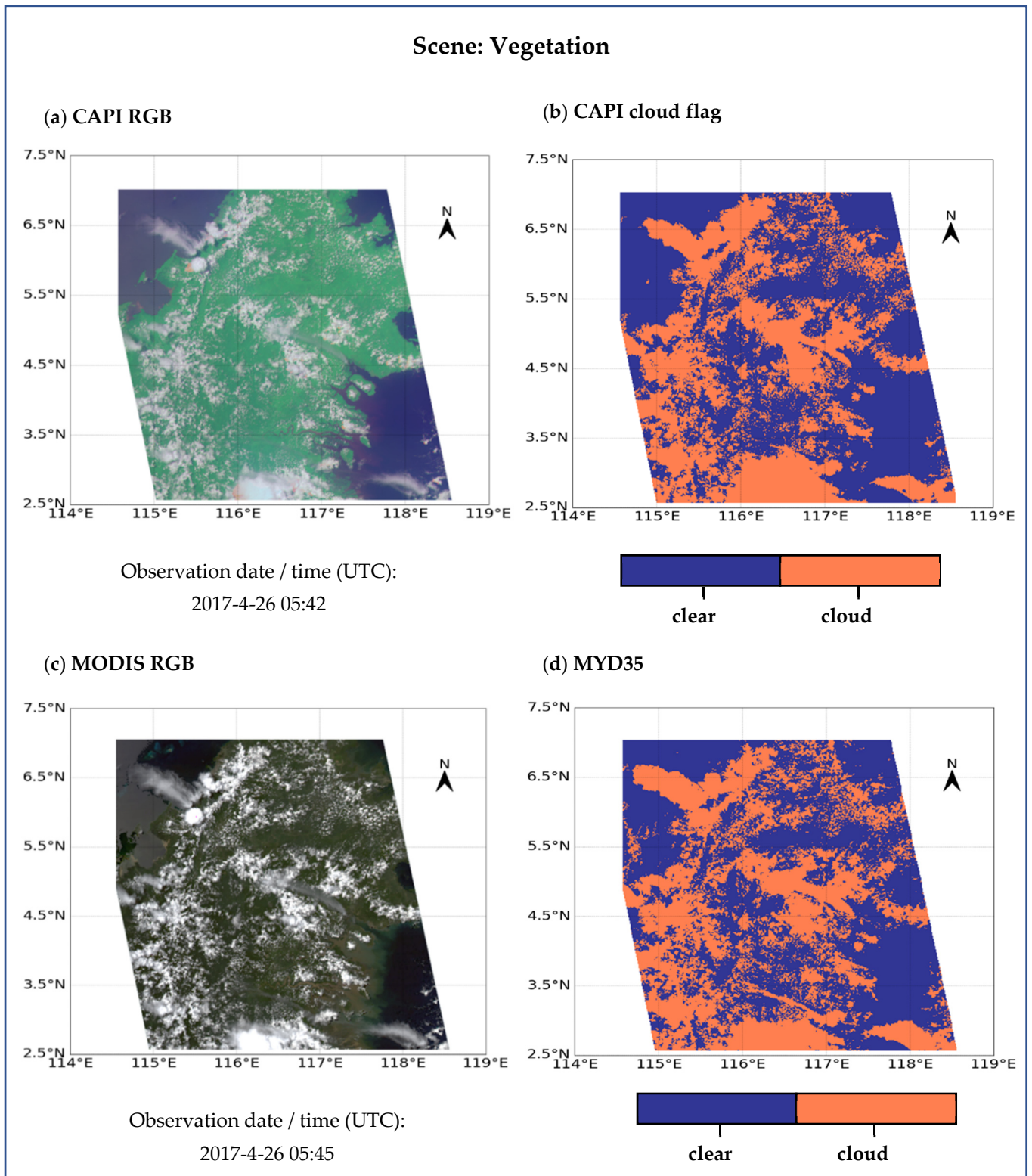


Figure 5. Same as Figure 4, but over the islands of Indonesia at 05:42 UTC 26 April 2017 for CAPI and at 05:45 UTC 26 April 2017 for MODIS.

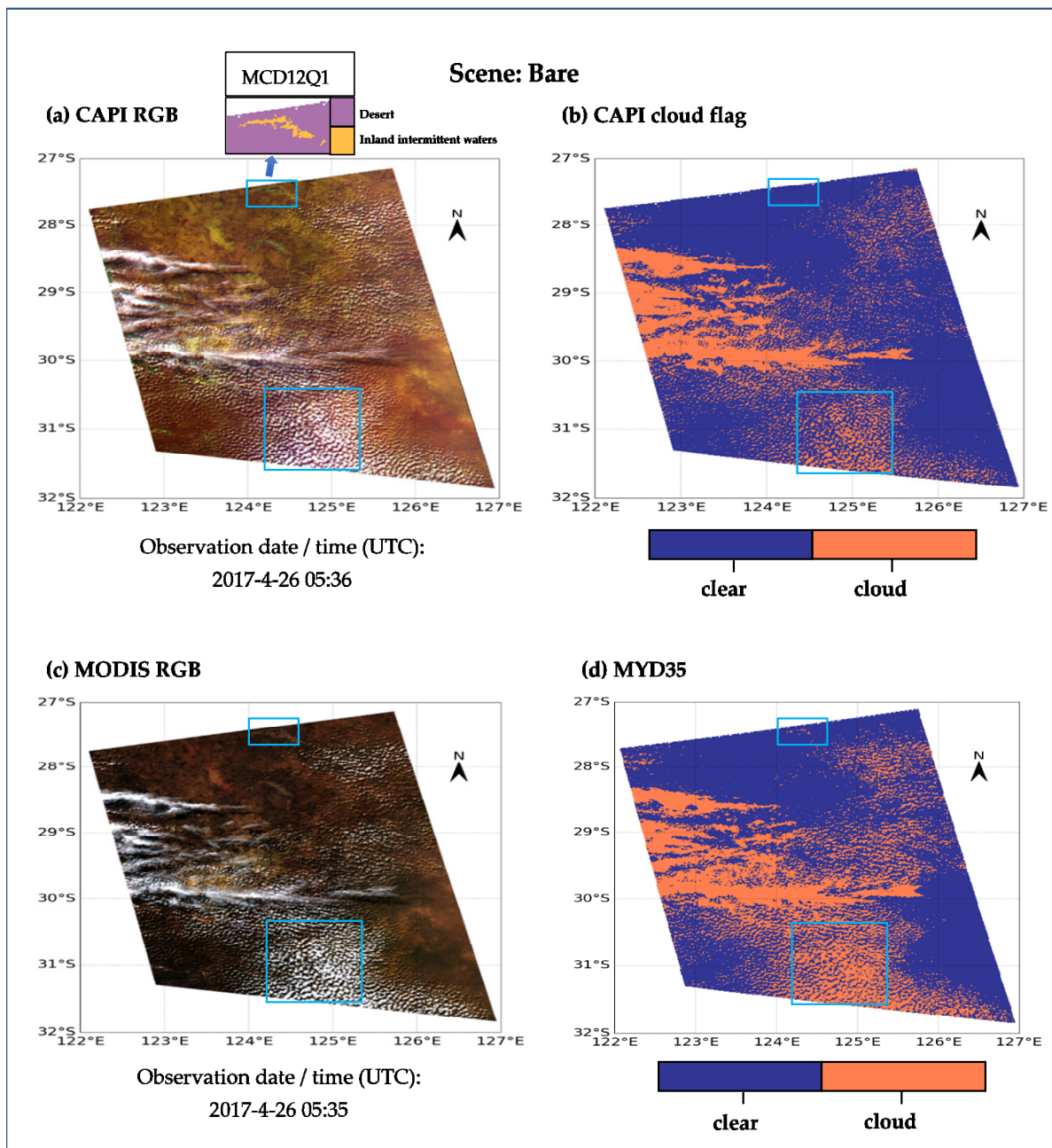


Figure 6. Same as Figure 4, but over the southwest part of Australia at 05:36 UTC 26 April 2017 for CAPI and at 05:35 UTC 26 April 2017 for MODIS.

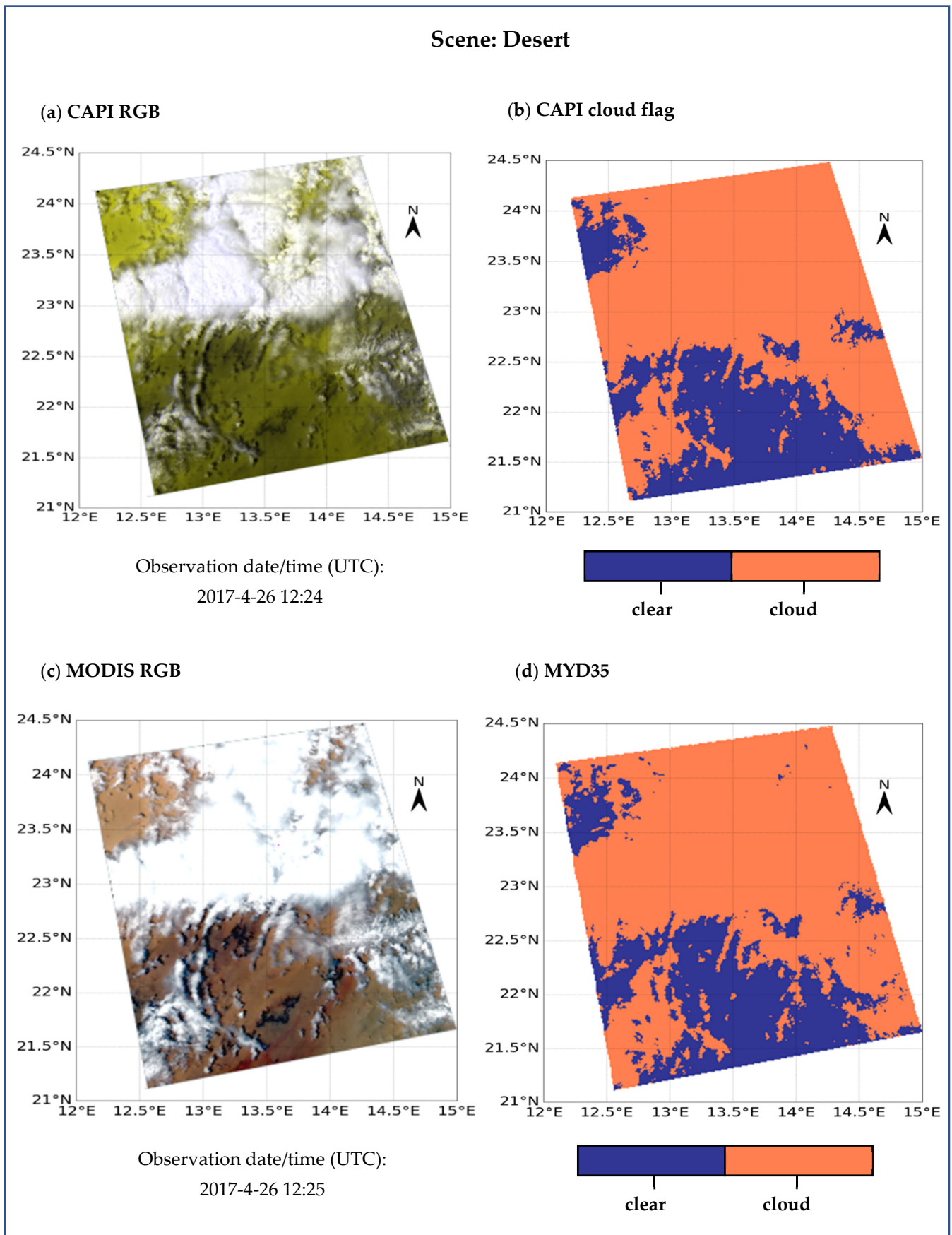


Figure 7. Same as Figure 4, but over the Sahara Desert at 12:24 UTC 26 April 2017 for CAPI and at 12:25 UTC 26 April 2017 for MODIS.

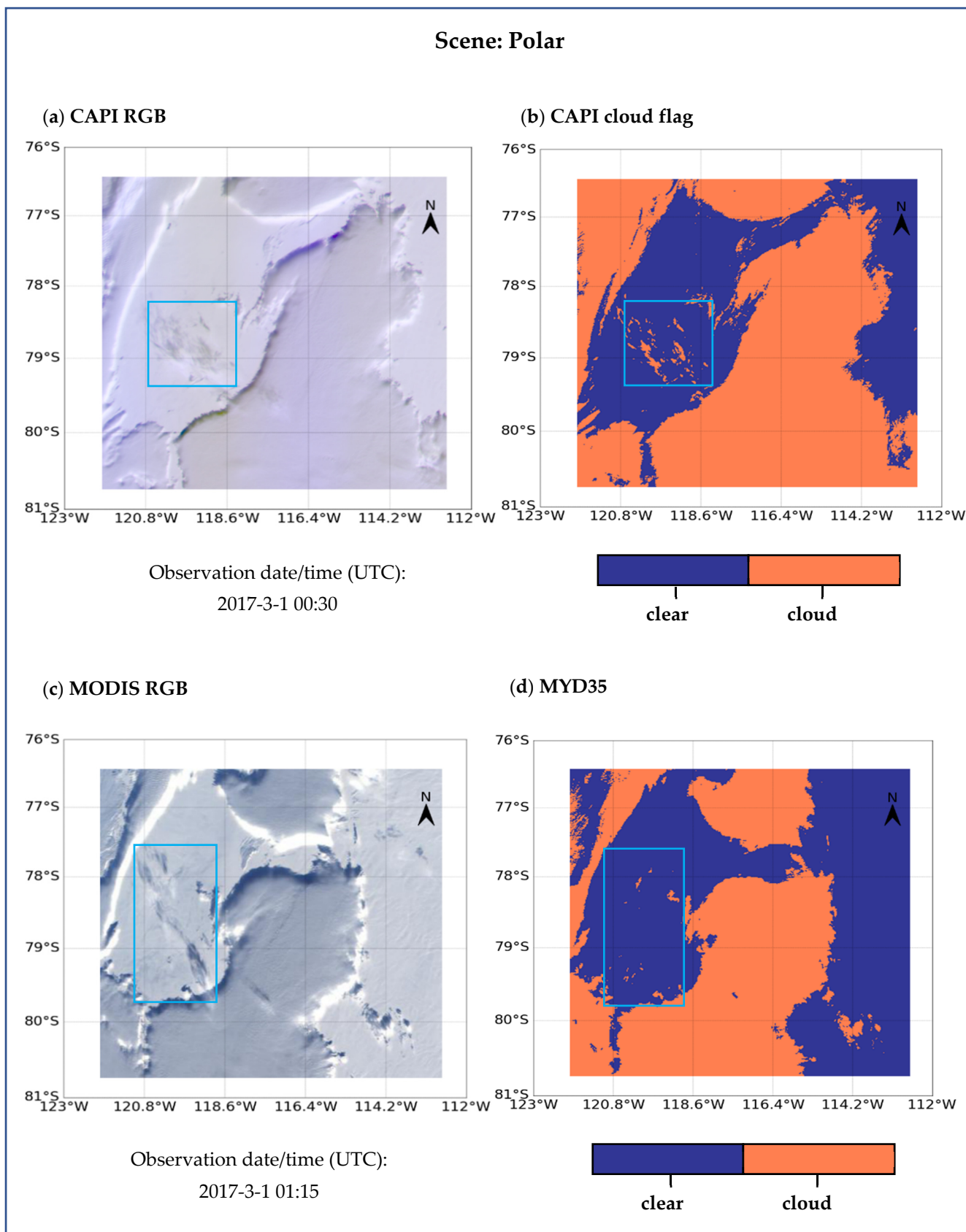


Figure 8. Same as Figure 4, but over the Antarctic continent at 00:30 UTC 1 March 2017 for CAPI and at 01:15 UTC 1 March 2017 for MODIS.

Table 5 lists the cloud cover (the percentage of cloud pixels in the total pixels) identified from MODIS and CAPI images in these five scenes. It can be seen from the cloud cover of the two results that the two values are close, with only 1–5% difference. Therefore, it intuitively proves the rationality of the NNDT algorithm in this paper.

Table 5. The cloud cover identified from MODIS and CAPI images.

Scenes	Ocean	Vegetation	Bare	Desert	Polar
MODIS	51.88%	43.70%	32.32%	71.65%	55.95%
CAPI	56.54%	41.72%	27.46%	70.19%	59.89%
Difference	4.66%	−1.98%	−4.86%	−1.46%	3.94%

3.2. Quantitative Comparison with SGLI Cloud Detection Production

The verification experiment in Section 3.1 cannot provide quantitative comparison verification, so for the NNDT algorithm, this paper designs a pixel-by-pixel verification scheme based on SGLI radiance data at the top of atmosphere (LTOA) and SGLI official cloud recognition products (SGLI-CLFG). SGLI (the Second-generation GLOBal Imager aboard the GCOM-C satellite, launched in December 2017) has 19 channels covering ultraviolet to thermal infrared spectra and two polarization and bidirectional channels, and contains all the channels used in the NNDT algorithm which are listed in Table 6. The thresholds of the tests are taken directly from the NNDT algorithm, considering the similar corresponding spectral bands between CAPI and SGLI.

Table 6. Wavelength and wavelength range of SGLI channel, which are used in NNDT algorithm and true color composite image.

Channel	Center Wavelength	Band Width
	VNR, SWI: nm	
VN1	380	10
VN3	443	10
VN5	530	20
VN8	673.5	20
VN11	868.5	20
SW2	1380	20
SW3	1630	200

SGLI official cloud recognition product (SGLI-CLFG) is similar to MYD35, using the cloudy/clear discrimination algorithm (CLAUDIA) [31], which includes abundant cloud tests from visible to thermal infrared band. CCL is represented by a 3-bit binary number and is divided into 8 levels. In order to facilitate the comparison, this article simplifies the CLFG levels and merges them into two categories—cloud and clear. Specific applications are shown in Table 7. Please refer to the user’s Manual for the specific application process [51,52].

In this paper, the algorithm verification experiments based on four scenes with different cloud attributes and surface features are selected as examples. The scene locations are shown in Table 8. The cloud screening cases over the four scenes are shown in Figures 9 and 10. Each scene has two panels: (a) The SGLI true color image of this scene synthesized from the VN8 (red), VN5 (green), and VN3 (blue) channel reflectance data and (b) The pixel-by-pixel comparison results of NNDT cloud recognition results (NNDT CLFG) based on SGLI observation data and SGLI official cloud recognition products (SGLI CLFG).

Table 7. Overall CCLs (Overall clear confidence levels) in SGLI-CLFG and simplified application results in this article.

CCLs	Bit	Application Results
0.00	000	Cloud
0.00–0.17	001	Cloud
0.17–0.33	010	Cloud
0.33–0.50	011	Cloud
0.50–0.67	100	Cloud
0.67–0.83	101	Cloud
0.83–1.00	110	Clear
1.00	111	Clear

Table 8. The sites and locations of the four scenes for this verification experiment.

Scenes	Site	Location
Ocean	East China Sea	(25~27° N, 121~123° E)
Vegetation	Shandong China	(34~36° N, 114~116° E)
Desert	Taklimakan desert	(37~39° N, 79~82° E)
Snow	Antarctic region	(66~69° S, 72~75° E)

As shown in Figures 9 and 10, the comparison results between NNDT-CLFG and SGLI-CLFG record pixels in four cases: A. NNDT-CLFG and SGLI-CLFG are both cloudy, B. both are clear, C. NNDT-CLFG are cloudy and SGLI-CLFG are clear, and D. NNDT-CLFG are clear and SGLI-CLFG are cloudy. It is apparent that the NNDT-CLFG provides relatively good agreement with both the SGLI-CLFG and the composite true-color images over the vegetation, desert, and ocean scenes. The cloud areas marked by the two results are basically the same, and only the determination of a very few pixels at the edge of the cloud has a deviation, which is within a reasonable range in the thresholds cloud recognition. However, for some optically thin clouds with extremely thin visual effects in the desert scene, both algorithms have the phenomenon of missing detection. The larger deviation between NNDT-CLFG and SGLI-CLFG appears in the polar scene, where a large number of cloud pixels at the edge of the cloud cluster marked by NNDT-CLFG are marked as clear sky by SGLI-CLFG. But by visually comparing NNDT-CLFG and true color images, it can be found that the two are highly consistent. These prove the effectiveness of the NNDT algorithm.

Four validation scores are used here for the quantitative analysis: the probability of detection (POD), the false-alarm ratio (FAR), the hit rate (HR), and Kuiper's skill score (KSS) [53]. The definitions are:

$$\text{POD}_{\text{clear}} = d / (c + d) \quad (17)$$

$$\text{POD}_{\text{cloud}} = a / (a + b) \quad (18)$$

$$\text{FAR}_{\text{clear}} = b / (b + d) \quad (19)$$

$$\text{FAR}_{\text{cloud}} = c / (a + c) \quad (20)$$

$$\text{HR} = (a + d) / (a + b + c + d), \text{ where } 0 < \text{HR} < 1 \quad (21)$$

$$\text{KSS} = ad - cb / (a + b)(c + d), \text{ where } -1 < \text{HR} < 1 \quad (22)$$

where a and d, respectively, represent the number of pixels where NNDT-CLFG and SGLI-CLFG are both "cloud" and "clear sky", b represents the number of pixels where

NNDT-CLFG are “clear” but SGLI-CLFG are “cloud”, and c represents the number of pixels where NNDT-CLFG are “cloud” but SGLI-CLFG are “clear”. These six scores all take SGLI-CLFG as standard, which means that SGLI-CLFG is correct.

The POD and FAR scores are the measure of the efficiency of the cloud identification algorithm in determining either cloud or clear events [54]. The POD values are supposed to be as close to 1 as possible. Conversely, the FAR values are supposed to be as close to 0 as possible. The HR values estimate the overall efficacy of the cloud detection algorithm. In addition, the KSS index is a complementary measurement method, reflecting the misclassifications to some extent [54]. It is used to evaluate the performance of the algorithm in separating cloud events from clear events. A value of 1.0 represents a perfect discrimination, while a value of -1.0 describes a complete discrimination failure.

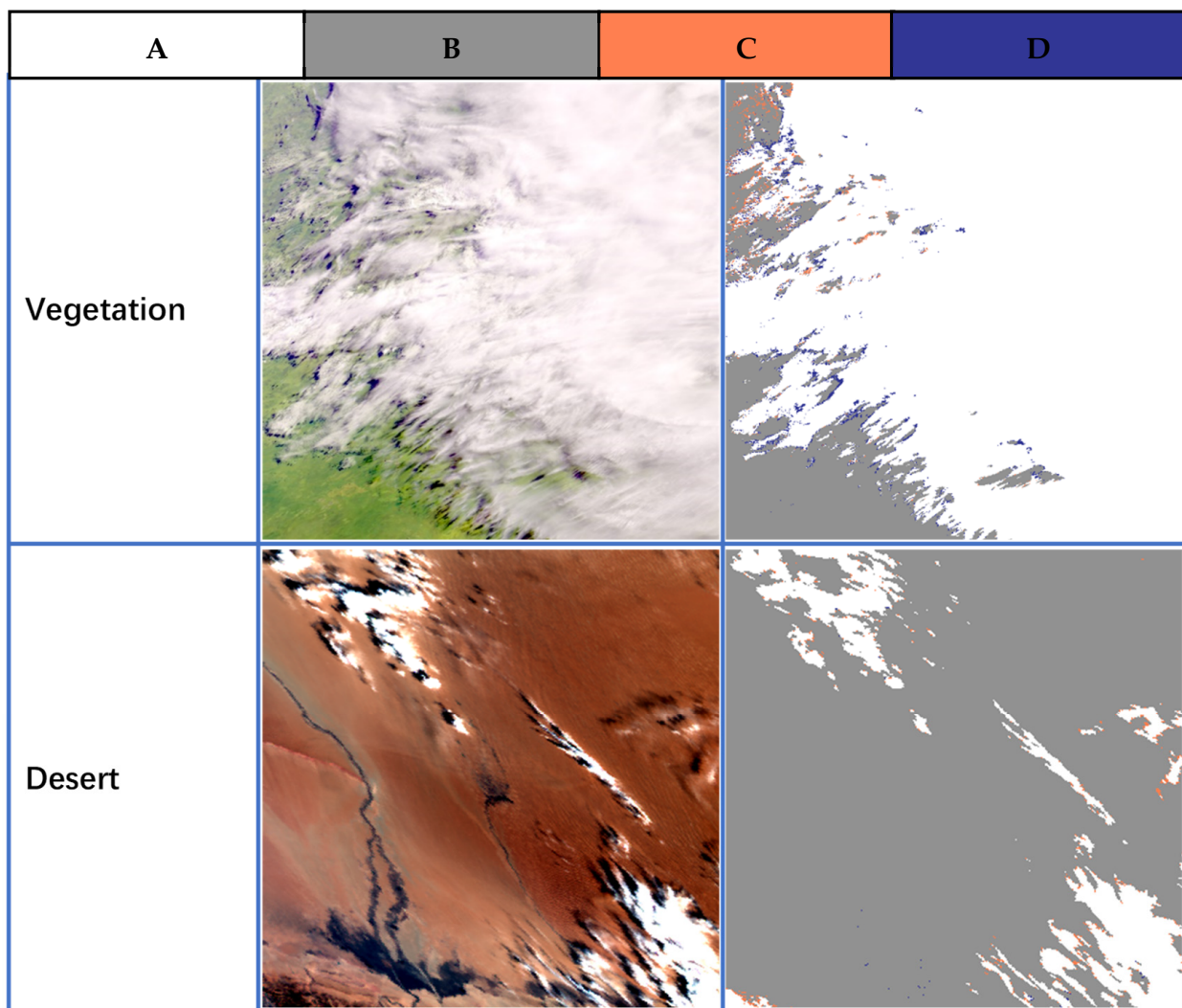


Figure 9. Per-pixel comparison of NNDT cloud identification results (NNDT-CLFG) based on SGLI observation data and SGLI-CLFG. The pixels represent (A). NNDT-CLFG and SGLI-CLFG are both cloudy (white), (B). both are clear (gray), (C). NNDT-CLFG are cloudy and SGLI-CLFG are clear (orange), and (D). NNDT-CLFG are clear and SGLI-CLFG are cloudy (blue). Column 1: Composite true-color images of SGLI. Column 2: Per-pixel comparison image for cloud-screening results obtained by NNDT algorithm and SGLI-CLFG. Line 1: Case over vegetation at T0527 26 April 2018 for SGLI. Line 2: Case over desert at T0524 10 March 2018 for SGLI.

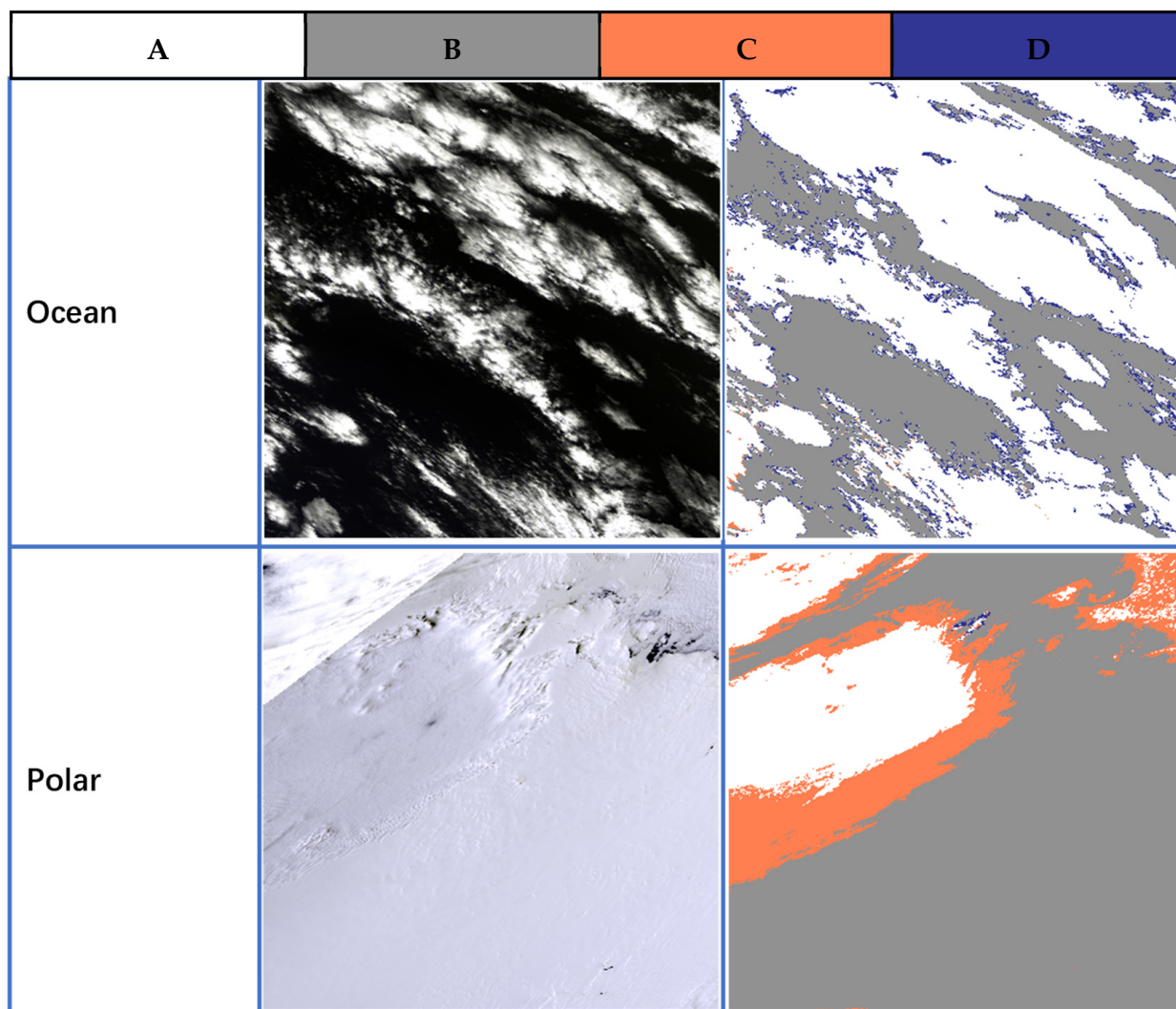


Figure 10. Same as Figure 9, but Line 1: Case over ocean at T0629 10 March 2018 for SGLI. Line 2: Case over snow at T1520 25 December 2018 for SGLI.

The scores of POD, FAR, HR, and KSS for NNDT algorithm, based on the cloud flag dataset of all seasons, are illustrated in Figure 11. Using the NNDT algorithm proposed in this paper, except for the polar region (the POD_{clear} , HR, and KSS values of this region are about 0.82, 0.86, and 0.82, respectively), the POD and HR scores of cloud pixels and clear pixels reach a relatively high value of more than 0.90 or close to 1. Except for the region above the polar (the FAR_{cloud} in this region is about 0.44), the FARs of other regions remain low or even close to zero. Overall, the results are encouraging.

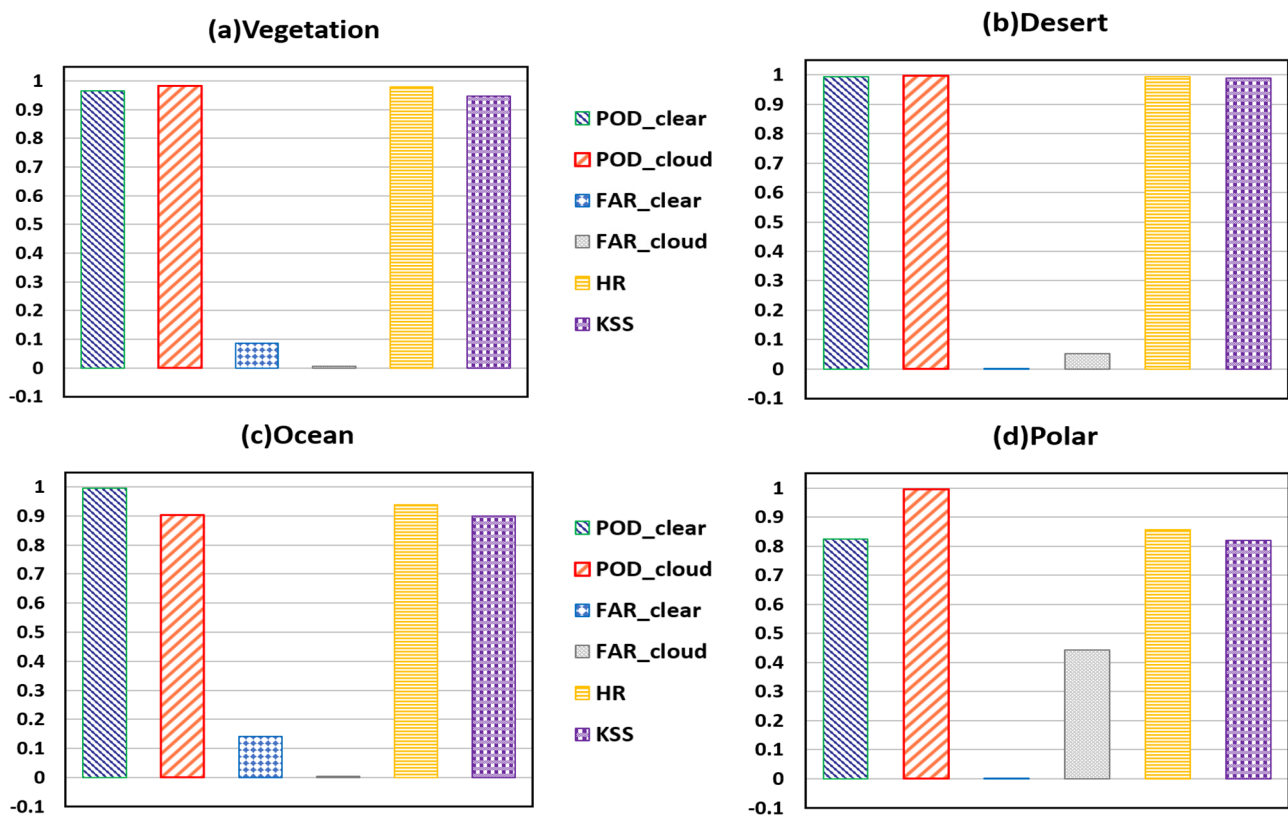


Figure 11. POD, FAR, HR, and KSS scores of NNDT algorithm in four scenes: (a) vegetation (b) desert (c) ocean, and (d) polar.

4. Discussion

From the visual comparison results in Section 3.1, it can be seen that the NNDT algorithm cloud recognition results and MYD35 have a relatively consistent cloud distribution range in the five scenes, but there are differences in local details. Moreover, there is a value difference between 1% and 5% of the cloud amount between the two results. The common reasons that caused the cloud cover difference are that the cloud displacement caused during the transit time difference between CAPI and MODIS as well as the cloud distortion caused by the observation geometric difference and geometric correction error [55]. Other reasons will be discussed separately for different scenarios.

In the Pacific Ocean scene, compared with MYD35, the cloud area identified by the NNDT algorithm is more consistent with the distribution of clouds. In most cases, the spatial distribution of cloud optical thickness is gradual. The distribution areas of optically thin clouds include the edge of the cloud, the hole in the center of the cloud cluster, and the middle of two close cloud clusters, which are difficult to be seen subjectively in RGB images. Due to the sensitivity of $0.38 \mu\text{m}$ band to the edge details of middle and low clouds and $1.375 \mu\text{m}$ band to optically thin clouds, the existence of these thin clouds can be objectively determined by the reflectivity thresholds of these two bands. Therefore, for the phenomenon that the pixels of optically thin clouds or edge clouds are recognized as clouds by NNDT algorithm, but not in MYD35, it shows the advantages of NNDT algorithm in the recognition of optically thin clouds and cloud edge pixels on the ocean surface, excluding the cloud displacement during this time difference and the cloud distortion caused by the geometric difference between the two instruments.

In the cloud detection results of bare ground in Figure 6, for the recognition of small scattered clouds, the cloud shape of CAPI cloud flag is more discrete, while the cloud shape of MYD35 is more continuous. The reason is that while the NNDT algorithm uses the threshold test of the band combination $R_f(0.87 \mu\text{m})/R_f(1.64 \mu\text{m})$ to suppress the

recognition of the highlighted sandbank as a cloud, it also causes a part of the cloud pixels with a gray value similar to that of the sandbank to be recognized as clear. Therefore, the edges of some small scattered clouds are missed. The MODIS cloud recognition algorithm includes cloud detection in the thermal infrared band [32], which can make up for the omission of point cloud detection. Point clouds are thick clouds. However, for high and thin clouds, where the $1.375\ \mu\text{m}$ test is not limited by the $R_f(0.87\ \mu\text{m})/R_f(1.64\ \mu\text{m})$ band combination, there will be no similar missed detection.

From the six verification scores of the four scenarios in Section 3.2, it can be seen that the POD values are all greater than 0.8, most of which are close to 1, indicating that the NNDT algorithm cloud recognition result basically conforms to the official SGLI algorithm result. The $\text{FAR}_{\text{clear}}$ values are all below 0.1, even close to 0, except in the ocean scene, indicating that the NNDT algorithm rarely distinguishes the pixels marked as cloud by SGLI CLFG as clear. In the ocean scene, the $\text{FAR}_{\text{clear}}$ value is slightly larger than in other scenes, which is 0.14. The reason is that the tests use the thresholds based on CAPI data statistics, which cannot be fully applicable to SGLI data, so the NNDT algorithm misses some cloud pixels, which is most obvious on the underlying surface of the ocean. The $\text{FAR}_{\text{cloud}}$ has good values close to 0 in other scenes, but it is 0.44 in the polar scene. However, the high consistency of NNDT-CLFG with true color images proves the effectiveness of the NNDT algorithm. The HR values of all scenes are between 0.86 and 0.99, which indicates that the NNDT algorithm has a high cloud recognition hit rate and further proves that the NNDT algorithm has a high cloud recognition accuracy.

Through the verification examples and verification scores in Sections 3.1 and 3.2, it can be seen that, in general, the NNDT algorithm cloud recognition results have a more consistent range and cloud amount on the four underlying surfaces compared with MYD35 and SGLI CLFG. The NNDT algorithm has excellent performance in the ocean, vegetation, and snow underlying surface, especially in the detection of thin clouds and cloud edges, which are difficult to detect. In scenes affected by sand with high reflectivity, the NNDT algorithm can identify most of the clouds, although the recognition of pixels at the edge of the cloud and thin cloud pixels is insufficient. Moreover, the algorithm can effectively avoid identifying inland waters as clouds. These verifications can prove that the $0.38\ \mu\text{m}$ band test can be used as an effective alternative to the 0.67 and $0.87\ \mu\text{m}$ band tests in the cloud recognition algorithm and does not require the support of the minimum surface reflectance database of the corresponding band. The band test of $R_f(0.38\ \mu\text{m})/R_f(1.64\ \mu\text{m})$ can be effectively used to identify cloud pixels on the polar surface. Based on the principle of identifying all cloud pixels in the image as much as possible, the NNDT algorithm proposed in this paper is more concise than the CLAUDIA algorithm and other algorithms that need to calculate the confidence level. It shows that the NNDT algorithm is reasonable, efficient, and suitable for cloud recognition of CAPI data.

5. Conclusions

Based on the requirements of TANSAT satellite sensor aerosol and CO_2 inversion for cloud detection, an innovative cloud recognition algorithm—NNDT (near-ultraviolet to near-infrared band with threshold tests for different underlying surfaces), suitable for CAPI data, is proposed. The algorithm uses unique band test combinations for four types of underlying surface, which are ocean, vegetation, desert, and polar regions, respectively. From the CAPI 2017 data, the four underlying surfaces are counted to obtain the thresholds used in all cloud recognition tests of this algorithm. Subsequently, two comparison methods are used, namely, a visual comparison method based on CAPI data and MODIS official cloud recognition results, and a quantitative comparison method based on SGLI data and SGLI official cloud recognition results, to verify and evaluate the effectiveness of the NNDT algorithm. Both verification methods randomly selected typical target areas representing four surface types (ocean, vegetation, desert, polar) and performed experimental verification. It can obtain good cloud recognition effects in all scenarios, with a hit rate between 0.86 and 0.99. Combined with the analysis of the spectral characteristics, it preliminarily shows that

the 0.38 μm band can well distinguish water, vegetation, and desert surface from clouds, and the ratio of 0.38 to 1.64 μm band has an excellent distinction ability between clouds and polar snow in the CAPI data.

This algorithm uses the combined threshold detections of near-ultraviolet to near-infrared bands to get rid of the drawbacks introduced using the 0.67 and 0.87 μm minimum reflectance database in traditional cloud detection algorithm. Moreover, the introduction of near-ultraviolet band to cloud detection algorithm enables it to obtain an effect comparable to that with thermal infrared band support, even without using the thermal infrared band. Based on the principle of identifying cloud pixels as much as possible, classifying all image pixels into cloud and clear categories can greatly improve the efficiency and accuracy of data utilization in aerosol and CO_2 inversion. This algorithm applies the 0.38 μm and the ratio of 0.38 to 1.64 μm to cloud recognition unprecedentedly, providing a theoretical basis and new ideas for future research.

In this paper, the underlying surface types are divided into four categories. The experimental regions selected for algorithm validation are different from those selected for threshold statistics. However, the application scenarios of this article cannot cover all the more detailed and special surface types in the world. In the future, we will focus on exploring and studying the differences of thresholds in different regions. According to the spectral characteristics in different regions, the underlying surfaces can be grouped in more detail, and the corresponding band detection schemes and thresholds can be set. Furthermore, cloud shadow is also a challenge in cloud detection. The surface data covered by cloud shadows will also affect the use of data by researchers [56–58]. In this paper, this method cannot solve this problem. This algorithm recognizes all cloud shadow areas as clear, which may affect the application effect of data in aerosol inversion. In future work, we will consider some bands and combinations that can detect cloud shadows, as well as corresponding threshold statistics.

Author Contributions: Conceptualization, N.D.; methodology, N.D.; software, N.D.; validation, N.D. and C.Y.; formal analysis, N.D. and J.S.; investigation, N.D.; resources, C.Y., J.S. and J.Z.; data curation, N.D.; writing—original draft preparation, N.D.; writing—review and editing, J.S., C.Y., J.Z., Y.Q., Y.P., J.Y., Y.D. and B.Y.; visualization, N.D. and B.Y.; supervision, C.Y.; project administration, C.Y.; funding acquisition, C.Y. and J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 61805235, in part by the National Key Research and Development Program of China under Grant 2016YFF0103603, in part by the Technology Development Program of Jilin Province, China, under Grant 20180201012GX, in part by CIOMP-Fudan University Joint Fund under Grant Y9R633A190, in part by the National Natural Science Foundation of China under Grant 61627819, Grant 61727818 and Grant 61875192 and in part by the STS Project of Chinese Academy of Sciences under Grant KFJ-STC-SCYD-212, Grant KFJ-STC-ZDTP-049, and Grant KFJ-STC-ZDTP-057.

Institutional Review Board Statement: In this section, you should add the Institutional Review Board Statement and approval number, if relevant to your study. You might choose to exclude this statement if the study did not require ethical approval. Please note that the Editorial Office might ask you for further information. Please add “The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of NAME OF INSTITUTE (protocol code XXX and date of approval).” OR “Ethical review and approval were waived for this study, due to REASON (please provide a detailed justification).” OR “Not applicable” for studies not involving humans or animals.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Publicly available datasets were analyzed in this study. The MODIS/Aqua Cloud Mask and Spectral Test Results 5-Min L2 Swath 1km dataset was acquired from the Level-1 and Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC), located in the Goddard Space Flight Center in Greenbelt, Maryland (<https://ladsweb.nascom.nasa.gov/>, accessed on 15 April 2021). Both LTOA and CLFG data are provided by Japan Aerospace Exploration

Agency, which can be downloaded from the official website of G-Portal (<https://gportal.jaxa.jp/gpr/information/beginner1>, accessed on 15 April 2021).

Acknowledgments: The authors would like to thank the editor and reviewers for their instructive comments that helped to improve this manuscript. Besides, they would also like to thank the National Satellite Meteorological Center data service platform and thank the NASA LAADS DAAC and G-Portal science team for providing the excellent and accessible data products that made this paper possible.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chen, L.F.; Zhang, Y.; Zou, M.M. Overview of atmospheric CO₂ remote sensing from space. *J. Remote Sens.* **2015**, *19*, 1–11. [[CrossRef](#)]
2. Zhao, L.R. Retrieval of Atmospheric CO₂ and CH₄ Using GOSAT. Master's Thesis, Jilin University, Jilin, China, December 2017.
3. Cai, Z.N.; Liu, Y.; Yang, D.X. Analysis of XCO₂ retrieval sensitivity using simulated Chinese Carbon Satellite (TanSat) measurements. *Sci. China Earth Sci.* **2014**, *57*, 1919–1928. [[CrossRef](#)]
4. Kuze, A.; Suto, H.; Nakajima, M.; Hamazaki, T. Thermal and near infrared sensor for carbon observation Fourier transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring. *Appl. Opt.* **2009**, *48*, 6716–6733. [[CrossRef](#)] [[PubMed](#)]
5. Shi, G.M.; Li, C.C.; Ren, T. Sensitivity analysis of single-angle polarization reflectance observed by satellite. *Chin. Sci. Bull.* **2014**, *59*, 1519–1528. [[CrossRef](#)]
6. Li, H.Y.; Zheng, H. Onboard Spectral and Spatial Cloud Detection for Hyperspectral Remote Sensing Images. *Remote Sens.* **2018**, *10*, 152. [[CrossRef](#)]
7. Uchino, O.; Kikuchi, N. Influence of aerosols and thin cirrus clouds on the GOSAT-observed CO₂: A case study over Tsukuba. *Atmos. Chem. Phys.* **2012**, *12*, 3393–3404. [[CrossRef](#)]
8. Jin, Z.; Zhang, L.; Liu, S.; Yi, F. Cloud Detection and Cloud Phase Retrieval Based on BP Neural Network. *Opt. Optoelectron. Technol.* **2016**, *14*, 74–77.
9. Yu, W.; Cao, X.; Xu, L.; Bencherkei, M. Automatic cloud detection for remote sensing image. *Chin. J. Sci. Instrum.* **2006**, *27*, 2184–2186.
10. Wang, W.; Song, W.; Liu, S.; Zhang, Y.; Zheng, H.; Tian, W. A Cloud Detection Algorithm for MODIS Images Combining Kmeans Clustering and Multi-Spectral Thershold Method. *Spectrosc. Spectr. Anal.* **2010**, *31*, 1061–1064.
11. Gao, J.; Wang, K.; Tian, X.Y.; Chen, J. A BP-NN based cloud detection method for FY-4 remote sensing images. *J. Infrared Millim. Waves* **2018**, *4*, 477–486.
12. Yuan, J.; Wang, X.; Yan, C.X. Soil Moisture Retrieval Model for Remote Sensing Using Reflected Hyperspectral Information. *Remote Sens.* **2019**, *11*, 366. [[CrossRef](#)]
13. Oishi, Y.; Ishida, H.; Nakajima, T.Y. The Impact of Different Support Vectors on GOSAT-2 CAI-2 L2 Cloud Discrimination. *Remote Sens.* **2017**, *9*, 1236. [[CrossRef](#)]
14. Yu, J.C.; Li, Y.C.; Zheng, X.X. An Effective Cloud Detection Method for Gaofen-5 Images via Deep Learning. *Remote Sens.* **2020**, *12*, 2106. [[CrossRef](#)]
15. Hollstein, A.; Segl, K.; Guanter, L. Ready-to-Use Methods for the Detection of Clouds, Cirrus, Snow, Shadow, Water and Clear Sky Pixels in Sentinel-2 MSI Images. *Remote Sens.* **2016**, *8*, 666. [[CrossRef](#)]
16. Bulgin, C.E.; Mittaz, J.P.D.; Embury, O. Bayesian Cloud Detection for 37 Years of Advanced Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC) Data. *Remote Sens.* **2018**, *10*, 97. [[CrossRef](#)]
17. Stöckli, R.; Bojanowski, J.S. Cloud Detection with Historical Geostationary Satellite Sensors for Climate Applications. *Remote Sens.* **2019**, *11*, 1052. [[CrossRef](#)]
18. Fu, H.L.; Shen, Y.; Liu, J.; He, G.J. Cloud Detection for FY Meteorology Satellite Based on Ensemble Thresholds and Random Forests Approach. *Remote Sens.* **2019**, *11*, 44. [[CrossRef](#)]
19. Lee, J.; Weger, R.C.; Sengupta, S.K.; Welch, R.M. A neural network approach to cloud classification. *IEEE Trans. Geosci. Remote Sens.* **1990**, *28*, 846–855. [[CrossRef](#)]
20. Liu, Z.; Li, Y.; Huang, F. Cloud Detection of MODIS Satellite Images Based on Dynamical Cluster. *Remote. Sens. Inf.* **2007**, *4*, 33–35.
21. Sun, S. A multi-spectral remote sensing imagery cloud detection algorithm based on spectral angle principle. *Microcomput. Its Appl.* **2017**, *36*, 16–18.
22. Liu, Z.; Han, L.; Zhou, P.; Wang, X.; Wu, T. A Method for Cloud Interpretation in ZY-3 Satellite Imagery and Its Application. *Remote Sens. Inf.* **2017**, *32*, 41–46.
23. Wu, W.; Luo, J.C.; Hu, X.D.; Yang, H.P. A Thin-Cloud Mask Method for Remote Sensing Images Based on Sparse Dark Pixel Region Detection. *Remote Sens.* **2018**, *10*, 617. [[CrossRef](#)]
24. Saunders, R.W.; Kriebel, K.T. An improved method for detecting clear sky and cloudy radiances from AVHRR data. *Int. J. Remote Sens.* **1988**, *9*, 123–150. [[CrossRef](#)]

25. Ackerman, S.A.; Strabala, K.I.; Menzel, W.P.; Frey, R.A.; Moeller, C.C.; Gumley, L.E. Discriminating clear sky from clouds with MODIS. *J. Geophys. Res. Atmos.* **1998**, *103*, 32141–32157. [[CrossRef](#)]
26. Ishida, H.; Nakajima, T.Y. Development of an unbiased cloud detection algorithm for a spaceborne multispectral imager. *J. Geophys. Res. Atmos.* **2009**, *114*, D07206. [[CrossRef](#)]
27. Nakajima, T.Y.; Tsuchiya, T.; Ishida, H.; Matsui, T.N.; Shimoda, H. Cloud detection performance of spaceborne visible-to-infrared multispectral imagers. *Appl. Opt.* **2011**, *50*, 2601–2616. [[CrossRef](#)]
28. Xiong, Q.; Wang, Y.; Liu, D.Y. A Cloud Detection Approach Based on Hybrid Multispectral Features with Dynamic Thresholds for GF-1 Remote Sensing Images. *Remote Sens.* **2020**, *12*, 450. [[CrossRef](#)]
29. Parmes, E.; Rauste, Y.; Molinier, M. Automatic Cloud and Shadow Detection in Optical Satellite Imagery Without Using Thermal Bands—Application to Suomi NPP VIIRS Images over Fennoscandia. *Remote Sens.* **2017**, *9*, 806. [[CrossRef](#)]
30. Wang, X.; Guo, Z. A Cloud Detection Scheme for the Chinese Carbon Dioxide Observation Satellite (TANSAT). *Adv. Atmos. Sci.* **2017**, *34*, 16–25. [[CrossRef](#)]
31. Nakajima, T.Y.; Ishida, H. Theoretical basis of the algorithms and early phase results of the GCOM-C (Shikisai) SGLI cloud products. *Prog. Earth Planet. Sci.* **2019**, *6*, 52. [[CrossRef](#)]
32. Liu, R.; Liu, Y. Generation of new cloud masks from MODIS land surface reflectance products. *Remote Sens. Environ.* **2013**, *133*, 21–37. [[CrossRef](#)]
33. Oishi, Y.; Sawada, Y.; Kamei, A. Impact of Changes in Minimum Reflectance on Cloud Discrimination. *Remote Sens.* **2018**, *10*, 693. [[CrossRef](#)]
34. Fukuda, S.; Nakajima, T.Y. New approaches to removing cloud shadows and evaluating the 380nm surface reflectance for improved aerosol optical thickness retrievals from the GOSAT/TANSO-Cloud and Aerosol Imager. *J. Geophys. Res. Atmos.* **2013**, *118*, 13520–13531. [[CrossRef](#)]
35. Baldridge, A.; Hook, S.; Grove, C. The ASTER spectral library version 2.0. *Remote Sens. Environ.* **2009**, *113*, 711–715. [[CrossRef](#)]
36. Irish, R. Landsat-7 automatic cloud cover assessment. *Proc. SPIE Int. Soc. Opt. Eng.* **2000**, *4049*, 348–355.
37. Hall, D.K.; Riggs, G.A.; Salomonson, V.V. Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data. *Remote Sens. Environ.* **1995**, *54*, 127–140. [[CrossRef](#)]
38. Hall, D.K.; Riggs, G.A.; Salomonson, V.V. MODIS snow-cover products. *Remote Sens. Environ.* **2002**, *83*, 181–194. [[CrossRef](#)]
39. Hall, D.K.; Riggs, G.A. *Normalized-Difference Snow Index (NDSI)*. *Encyclopedia of Snow, Ice and Glaciers*; Singh, V.P., Haritashya, U.K., Eds.; Springer: Dordrecht, The Netherlands, 2011; pp. 779–780.
40. Gao, B.C.; Goetz, A.F.H.; Wiscombe, W.J. Cirrus cloud detection from Airborne Imaging Spectrometer data using the 1.38 μm water vapor band. *Geophys. Res. Lett.* **1993**, *20*, 301–304. [[CrossRef](#)]
41. Vogel, S.W. Usage of high-resolution Landsat 7 band 8 for single-band snow-cover classification. *Ann. Glaciol.* **2002**, *34*, 53–57. [[CrossRef](#)]
42. Riggs, G.A.; Hall, D.K. Snow and cloud discrimination factors in the MODIS snow algorithm. In Proceedings of the IGARSS 2004, 2004 IEEE International Geoscience and Remote Sensing Symposium, Anchorage, AK, USA, 20–24 September 2004; IEEE: Piscataway, NJ, USA, 2004; pp. 3714–3716.
43. Hutchison, K.D.; Iisager, B.D.; Hauss, B. The use of global synthetic data for pre-launch tuning of the VIIRS cloud mask algorithm. *Int. J. Remote Sens.* **2012**, *33*, 1400–1423. [[CrossRef](#)]
44. Han, J.; Yang, L.K. Research on Algorithm of Cloud Detection for HJ-1B Image Based on Dynamical Thresholding. *Remote Sens. Land Resour.* **2012**, *6*, 2.
45. Zhao, X. Study on Cloud Detection Method for High Resolution Satellite Remote Sensing Image. Master's Thesis, Harbin Institute of Technology University, Harbin, China, June 2013.
46. Ackerman, S.A.; Cox, S.K. Comparison of satellite and all-sky camera estimates of cloud cover during GATE. *J. Appl. Meteorol.* **1981**, *20*, 581–587. [[CrossRef](#)]
47. Rossow, W.B.; Garder, L.C. Cloud detection using satellite measurements of infrared and visible radiances for ISCCP. *J. Clim.* **1993**, *6*, 2341–2369. [[CrossRef](#)]
48. Baum, B.A.; Uttal, T.; Poellot, M.; Ackerman, T.P.; Alvarez, J.M. Satellite remote sensing of multiple cloud layers. *J. Atmos. Sci.* **1995**, *52*, 4210–4230. [[CrossRef](#)]
49. Ackerman, S.; Frey, R.; Strabala, K.; Liu, Y.H. *Discriminating Clear-Sky from Cloud with MODIS Algorithm Theoretical Basis Document (Mod35)*; Version 6.1.; Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin: Madison, WI, USA, 2010.
50. Ackerman, S. *MODIS Atmosphere L2 Cloud Mask Product*; NASA MODIS Adaptive Processing System, Goddard Space Flight Center: Greenbelt, MD, USA, 2017. [[CrossRef](#)]
51. *GCOM-C “SHIKISAI” Data Users Handbook*, 1st ed.; Japan Aerospace Exploration Agency: Chofu City, Tokyo, Japan, 2018.
52. *G-Portal User's Manual Revision A*; Japan Aerospace Exploration Agency: Chofu City, Tokyo, Japan, 2019.
53. Musial, J.P.; Husler, F. Daytime Low Stratiform Cloud Detection on AVHRR Imagery. *Remote Sens.* **2014**, *6*, 5124–5150. [[CrossRef](#)]
54. Karlsson, K.G.; Dybbroe, A. Evaluation of Arctic cloud products from the EUMETSAT Climate Monitoring Satellite Application Facility based on CALIPSO-CALIOP observations. *Atmos. Chem. Phys.* **2010**, *10*, 1789–1807. [[CrossRef](#)]

-
55. Taylor, T.E. Orbiting Carbon Observatory-2 (OCO-2) cloud screening algorithms: Validation against collocated MODIS and CALIOP data. *Atmos. Meas. Tech.* **2016**, *9*, 973–989. [[CrossRef](#)]
 56. Zhu, Z.; Woodcock, C.E. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* **2012**, *118*, 83–94. [[CrossRef](#)]
 57. Wang, B.; Ono, A.; Muramatsu, K.; Fujiwara, N. Automated detection and removal of clouds and their shadows from Landsat TM images. *IEICE Trans. Inf. Syst.* **1999**, *82*, 453–460.
 58. Zhu, Z.; Woodcock, C.E. Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sens. Environ.* **2014**, *152*, 217–234. [[CrossRef](#)]