A Cloud Detection Algorithm Over Land Based on the Polarized Characteristics Difference Between Cloudless and Cloud Targets

Xin Gao1,2, Xiuqing Hu1,2, Wei Fang2, and Dekui Yin4

1Department of Optical Engineering, University of Chinese Academy of Sciences, Beijing, China, 2Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, China, 3China Meteorological Administration, Beijing, China, 4Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai, China

Abstract The accurate identification of cloud over land is one of the key issues of the satellite data processing and the product retrievals. This paper describes a new cloud detection algorithm based on Level 1 data of Polarization and Directionality of Earth’s Reflectance. The simulation of multiangular normalized polarized reflectance is done for cloudless targets over land before the cloud identification processing. First, the Normalized Difference Vegetation Index over land and reflectance of 670 nm are used as two initial criterions for the cloud mask. Then, the difference between the simulation and Polarization and Directionality of Earth’s Reflectance observation of polarized reflectance is used as the third criterion to distinguish cloudless pixels from cloud ones. And this algorithm is proved to be more convenient and effective. This algorithm is also applied to cloud mask processing of the Multi-Angular Polarization Imager (onboard the Tiangong-2) observation. The results show that this algorithm can effectively detect cloud targets over land, and its consistency with Polarization and Directionality of Earth’s Reflectance official cloud mask products is about 90%. This algorithm can provide reliable cloud mask products for the retrieval of optical and physical properties of land aerosol using Multi-Angular Polarization Imager data.

Plain Language Summary Cloud detection is a crucial step to study clouds and other products generation for the satellite application, the result of which affects further researches. This study uses the multiangular polarized data to achieve cloud detection. A simulation was done for cloudless pixels with the scattering phase function theory before the cloud identification. And this paper shows the cloud detection results with this new algorithm, as well as comparisons with reference cloud mask from the Polarization and Directionality of Earth’s Reflectance operational products. The consistency is about 90%.

1. Introduction

Cloud is one of the important regulators in the land-atmosphere coupling system. It covers about two thirds of the Earth’s surface and is closely related to the Earth’s water cycle, radiation budget, and climate change. In order to retrieve the optical and physical properties of land aerosol, cloud detection and removal should be completed first, so as to ensure the accuracy and reliability of the retrieval products.

With the development of remote sensing technology, a large amount of remote sensing data has been obtained from various satellite payloads. From these data sets, researchers have proposed and improved many cloud detection algorithms. Key and Barry (1989) proposed a cloud detection method for Arctic data using the Advanced Very High Resolution Radiometer. Derrien et al. (1993) proposed an automatic cloud detection algorithm, and its automation came from the computation of the 11-μm infrared threshold from a monthly sea surface temperature climatology over the oceans and from air temperature (near the surface) forecast by NWP over land. Chen et al. (2002) presented an automatic cloud detection method for Advanced Very High Resolution Radiometer, which includes three major indicators: top-of-the-atmosphere reflectance of channel 1, the temperature difference of channels 3 and 4, and a combination of ratio of channel 2 to channel 1 and temperature in channel 4. The objects of the above algorithms are all oriented to traditional remote sensing data, without polarized or multiangular information.

As a new remote sensing method in recent years, multiangular polarized remote sensing is more suitable in cloud and aerosol retrievals than traditional optical remote sensing. The French Centre National d’Etudes
Spatiales has launched three Polarization and Directionality of Earth’s Reflectance (POLDER) payloads successively from 1996 to 2004, recording a large amount of data between 1996 and 2013. The Multi-Angular Polarization Imager (MAPI), developed by the Shanghai Institute of Technical Physics of the Chinese Academy of Sciences, was onboard the Tiangong-2 space laboratory, which was launched in 2016. The acquisition of polarization information provided a new study direction for cloud detection. Buriez et al. (1997) did a series of sequential tests applied to each individual pixel, including the top-of-atmosphere reflectance of 670 and 865 nm, apparent pressure (using mainly the R763/R765 ratio), and labeled every pixel as clear, cloudy or undetermined. Cheng et al. (2007) detected clouds in China regions via the combination of POLDER data and Moderate Resolution Imaging Spectroradiometer (MODIS) data, based on the spectral, multiangular, and polarized characteristics of cloud. Chen et al. (2018) established a spatial fusion rule for multiangular cloud detection results and proposed a fast cloud detection method. Desmons et al. (2017) described an algorithm based on the POLDER observations to identify monolayered and multilayered cloudy situations along with a confidence index. Li et al. (2018) proposed an improved POLDER cloud detection method, using radiation transfer model and setting new dynamic thresholds to improve the accuracy of the final cloud detection.

The cloud detection algorithm for POLDER official cloud mask product was described by Nadal and Bréon (1999). Four tests were applied to the measurements. The first one was a threshold on the 0.44-μm reflectance after atmospheric correction. The second one was similar but with a smaller threshold and was applied only over targets with significant spectral variation. The third one compared the surface pressure to an estimate derived from two POLDER channels centered on an oxygen absorption band. The fourth one made use of POLDER polarization capabilities and sought the presence of a rainbow generated by water clouds.

We can see that most of these methods use the difference or ratio of reflectance between several channels for cloud detection, but this paper proposes a new cloud detection algorithm focusing more on the characteristics of the multiangular polarized reflectance of one channel, which uses POLDER Level 1 data and MAPI Level 1 data, and uses only three thresholds to make the algorithm more convenient but also effective. The simulation of multiangular normalized polarized reflectance is done for cloudless targets over land before the cloud identification processing. First, the Normalized Difference Vegetation Index (NDVI) over land and reflectance of 670 nm are used as two initial criterions for the cloud mask. Then, the difference between the simulation and POLDER observation of polarized reflectance is used as the third criterion to distinguish cloudless pixels from cloud ones.

2. Data and Method

2.1. Data and Instruments

2.1.1. POLDER

Three POLDER payloads were developed by the French Centre National d’Etudes Spatiales, which were applied to Earth observation missions from November 1996 to June 1997, April 2003 to October 2003, and December 2004 to December 2013. POLDER payload is a polarized, multiangular, and multispectral imaging spectrometer, consisting of nine spectral bands (Table 1). Among them, 490, 670, and 865 nm are polarized channels, and the corresponding Stokes parameters I, Q, and U are obtained through three polarized subchannels. At the same time, it also records abundant observation information such as cloud mask, landsea mask, and observation geometry (i.e., solar zenith angle, solar azimuth angle, satellite zenith angle, satellite azimuth angle, relative azimuth angle). POLDER Level 1 data are available for download from the ICARE Data and Services Center (France). Here I, Q, and U are given in units of “normalized radiance”: the radiance (W m⁻² sr⁻¹) has been multiplied by \( \pi / E_\odot \), where \( E_\odot \) is the extraterrestrial solar flux accounting for the variations of Sun-Earth distance. The “normalized radiance” data can be converted to reflectance by a simple division by the cosine of the solar zenith angle (Bréon, 2003).

2.1.2. MAPI

The MAPI was developed by the Shanghai Institute of Technical Physics, Chinese Academy of Sciences. It was boarded on the Tiangong-2 space laboratory, which was launched in September 2016 and has been performing the Earth observation mission and application experiment ever since then. MAPI is also a polarized, multiangular, and multispectral imaging spectrometer, which contains six spectral bands (Table 1). Among them, 565, 670, and 865 nm are the polarization channels, and three polarized subchannels are set to obtain Stokes parameters I, Q, and U, which are also in the units of normalized radiance.
2.1.3. MODIS Cloud Mask Products

The MODIS Level 1B data set contains calibrated and geolocated at-aperture radiances for 36 discrete bands located in the 0.4- to 14.4-μm region of the electromagnetic spectrum. The resolution of channels 1 and 2 is 250 m, channels 3 through 7 are 500-m resolution, and the rest are 1-km resolution. There are two MODIS Level 1B data product files: MOD021KM, containing data collected from the Terra platform, and MYD021KM, containing data collected from the Aqua platform. The MODIS Cloud Mask product is a Level 2 product generated at 1-km and 250-m (at nadir) spatial resolutions. The algorithm employs a series of visible and infrared threshold and consistency tests to specify confidence that an unobstructed view of the Earth's surface is observed. There are two MODIS Cloud Mask data product files: MOD35_L2, containing data collected from the Terra platform, and MYD35_L2, containing data collected from the Aqua platform. All those data sets can be downloaded from https://modis.gsfc.nasa.gov/.

2.2. Principle of Cloud Detection Algorithm

The cloud detection algorithm over land can be divided into three steps. First, all the pixels in the remote sensing data are identified as water and land pixels, and the land pixels are extracted, while the water ones are not involved in the cloud detection process. Then, NDVI is calculated by 670- and 865-nm data, and reflectance of 670 nm is extracted. Due to the high reflection of cloud, water, and snow to visible bands, part of cloud pixels could be identified, so NDVI and \( R_{670} \) are taken as two initial criterions. Finally, for the remaining pixels, a new criterion based on the scattering phase function theory is adopted to effectively distinguish cloudless pixels from cloud ones. The diagram of algorithm is shown in Figure 1.

2.2.1. Land-Water Flag

In this study, a database of the global land-water flag is established with the MODIS Land Cover Type product (short name: MCD12). Then, by searching the longitude and latitude with the nearest matching method, we can assign a land-water flag for every pixel of POLDER/MAPI data according to the corresponding cover type flag number. The spatial resolution of POLDER/MAPI/MCD12 is approximately 6 km, 3 km, and 250 m, respectively, which shows that MCD12 has enough ability to provide accurate land-water flags for POLDER and MAPI.

| Band | Central wavelength (nm) POLDER/MAPI | Band width (nm) POLDER/MAPI | Polarization POLDER/MAPI |
|------|-------------------------------------|-----------------------------|-------------------------|
| 443  | 443.9/-                             | 13.5/-                      | No/-                    |
| 490  | 491.5/-                             | 16.5/-                      | Yes/-                   |
| 565  | 563.9/565                           | 15.5/20                     | No/Yes                  |
| 670  | 669.9/670                           | 15.0/20                     | Yes/Yes                 |
| 763  | 762.8/763                           | 11.0/10                     | No/No                   |
| 765  | 762.5/765                           | 38.0/40                     | No/No                   |
| 865  | 863.4/865                           | 33.5/40                     | Yes/Yes                 |
| 910  | 906.9/910                           | 21.0/20                     | No/No                   |
| 1020 | 1019.4/-                            | 17.0/-                      | No/-                    |

Table 1. Characteristics of the POLDER/MAPI Bands

![Figure 1. The diagram of the cloud detection algorithm.](image-url)
2.2.2. NDVI and Reflectance Criterions

The so-called “multiangular” means that there are 12–13 effective layers in POLDER data sets according to the observation geometry, and 8–9 effective layers in MAPI data sets. Both for POLDER and MAPI, NDVI and $R_{670}$ used in this study are all calculated by the data of layer 5.

NDVI is defined as (Nadal & Bréon, 1999)

$$\text{NDVI} = \frac{R_{865} - R_{670}}{R_{865} + R_{670}}$$

where $R$ is the normalized reflectance. And there are equations as

$$I = \frac{\pi L_I}{E_\lambda} \quad Q = \frac{\pi L_Q}{E_\lambda} \quad U = \frac{\pi L_U}{E_\lambda}$$

$$R = \frac{I}{\cos \delta_s} \quad R_p = \frac{\sqrt{Q^2 + U^2}}{\cos \delta_s}$$

where $L_I, L_Q, and L_U$ are the Stokes parameters in units of radiance ($\text{Wm}^2/\text{sr}$). $I, Q, and U$ are the same but in units of “normalized radiance.” $E_\lambda$ is the extraterrestrial solar flux ($\text{Wm}^2$). $\delta_s$ is the solar zenith angle. Therefore, there is the equation:

$$\text{NDVI} = \frac{I_{865} - I_{670}}{I_{865} + I_{670}}$$

The effective value range of NDVI is $[-1, 1]$. Theoretically, the value in $[-1, 0]$ represents that the surface is covered by clouds, water, snow, etc., which are highly reflective to visible bands; the value 0 represents rocks or bare soil; and the value in $(0, 1]$ represents that the surface is covered by vegetation, and the higher the value is, the higher the vegetation coverage is.

However, the above classification standards are not accurate, so we identify the pixels with NDVI $\geq 0.1$ and $R_{670} < 0.3$ as cloudless ones. In addition, because of the high reflectivity of the cloud target itself, the pixels with reflectance $R_{670} > 0.3$ or NDVI $\leq -0.1$ are directly judged as clouds. The snow has a large reflectance with a spectral signature like that of the clouds in the POLDER spectral range. This may lead to some uncertainties for the algorithm. And the pixels with NDVI in $(-0.1, 0.1)$ and $R_{670}$ less than 0.3 should be examined in the next step. The selections of these thresholds will be discussed in section 3.

2.2.3. Simulation of Multiangular Normalized Polarized Reflectance

For a cloudless pixel over land, the top-of-atmosphere reflectance (apparent reflectance) received by the detector consists of three parts: molecule reflectance, aerosol reflectance, and surface reflectance. They can be calculated by the following formulas (Deuzé et al., 2001):

$$R_{mol} = \frac{\tau_{mol} Q_{mol}}{4 \cos \delta_s \cos \delta_v}$$

$$R_{aer} = \frac{\tau_{aer} Q_{aer}}{4 \cos \delta_s \cos \delta_v}$$

$$R_{surf} = f_s(\Omega)$$

The parameters $\tau_{mol, aer}$ are the optical thickness of molecule and aerosol, respectively. $Q_{mol, aer}$ are the corresponding scattering polarized phase function.

The scattering characteristics of the molecule follow the Rayleigh scattering theory, and the scattering polarized phase function can be given by the empirical formula:

$$Q_{mol}(\Theta) = \frac{3}{4} (1 - \cos^2 \Theta)$$

where $\Theta$ is the scattering angle, defined as

$$\Theta = \arccos(-\cos \delta_s \cos \delta_v - \sin \delta_s \sin \delta_v \cos \varphi)$$

$\delta_s, \delta_v,$ and $\varphi$ stand for the solar zenith angle, the satellite zenith angle, and the relative azimuth angle.
The molecular optical thickness $\tau_{\text{mol}}$ depends on the empirical formula (Bodhaine et al., 1999):

$$\tau_{\text{mol}} = 0.008569\lambda^{-4}\left(1 + 0.0113\lambda^{-2} + 0.00013\lambda^{-4}\right)\frac{P}{P_0} \tag{10}$$

where $\lambda$ is the wavelength, $P$ is the atmospheric pressure at the target point, and $P_0$ is the standard atmospheric pressure.

The scattering characteristics of aerosol are more sophisticated. In addition to the complexity of the scattering characteristics of aerosol particles, multiple scattering also needs to be considered. In order to simplify the operation and improve the efficiency of the algorithm, the hypothesis of spherical particles and single scattering is adopted in this paper. According to the Mie scattering theory, the scattered and incident waves are bridged by scattering phase matrix $P$ (Li et al., 2004) in the form

$$\begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \frac{\Omega_{\text{eff}}}{4\pi} \begin{bmatrix} P_{11}(\Theta) & P_{12}(\Theta) & 0 & 0 \\ P_{12}(\Theta) & P_{22}(\Theta) & 0 & 0 \\ 0 & 0 & P_{33}(\Theta) & P_{34}(\Theta) \\ 0 & 0 & -P_{34}(\Theta) & P_{44}(\Theta) \end{bmatrix} \begin{bmatrix} I_0 \\ Q_0 \\ U_0 \\ V_0 \end{bmatrix} \tag{11}$$

$\Omega_{\text{eff}}$ denotes the effective solid angle associated with scattering. $P_{11}(\Theta)$ and $P_{12}(\Theta)$ stand for the aerosol scattering phase function $p_{\text{aer}}(\Theta)$ and polarized phase function $q_{\text{aer}}(\Theta)$ for single spherical particle. The polarized phase function for aerosol in the atmosphere is calculated as

$$Q_{\text{aer}}(\Theta) = \frac{\int_{r_{\text{min}}}^{r_{\text{max}}} q_{\text{aer}}(\Theta) C_{\text{sca}}(\lambda, r, m) n(r) dr}{\int_{r_{\text{min}}}^{r_{\text{max}}} C_{\text{sca}}(\lambda, r, m) n(r) dr} \tag{12}$$

$C_{\text{sca}}(\lambda, r, m)$ is the scattering efficiency factor as given by the Mie theory. And $Q_{\text{aer}}(\Theta)$ is associated with $m$ (the complex refractive index, $m = m_r - im_i$) and $n(r)$ (the columnar radius distribution) of aerosol, which is modeled as a lognormal distribution described by the median radius $r_g$ and the standard deviation $\sigma$:

$$n(r) = \frac{d N}{d \ln r} = \frac{N_0}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{(\ln r - \ln r_g)^2}{2\sigma^2} \right] \tag{13}$$

In this paper, a database for $Q_{\text{aer}}(\Theta)$, which were calculated by different $m$ and $n(r)$, was established to represent different types of aerosol, listed as Table 2. Using the look-up-table method, the optimal set of $r_{\text{aer}}$ and $Q_{\text{aer}}$ was selected to calculate the aerosol polarized reflectance $R_{\text{aer}}$ for every pixel.

The polarized reflectance of surface is obtained by the BPDF semiempirical model proposed by Nadal et al. (1999):

$$F_{\text{surf}}^L(\Omega) = \rho \left[ 1 - \exp\left(-\beta \frac{F_p(\alpha)}{\cos \delta_i + \cos \delta_f}\right) \right] \tag{14}$$

$F_p$ is the Fresnel reflection coefficient, $\rho$ and $\beta$ are the empirical coefficients adjusted for different types of surface, and $\Omega$ is the incidence angle relative to the ground for the Sun (i.e., $\Omega = (\pi - \Theta)/2$).

Finally, the apparent reflectance can be calculated according to the atmospheric radiation transfer equation (Vermote et al., 2006):

$$R' = T_g \left( R_{\text{mol}} + R_{\text{aer}} + \frac{T^T R_{\text{aer}} T}{1 - SR_{\text{surf}}} \right) \tag{15}$$

where $T_g$ is the molecular transmittance, $T^T, T$ are the upward and downward transmittances of aerosol, respectively, and $S$ is the spherical albedo.

| Parameter | Values | Number of values |
|-----------|--------|-----------------|
| $m_r$     | 1.4, 1.5, 1.6 | 3               |
| $m_i$     | 0.01, 0.05, 0.1 | 3               |
| $r_g$     | 0.001, 0.005, 0.01, 0.02, 0.05, 0.1, 0.5, 1 | 8               |
| $r_{\text{aer}}$ | 0.00, 0.05, 0.10, 0.20, 0.40, 0.80, 1.20, 1.60, 3.00, 5.00 | 10              |
So far, the simulation of the multiangular normalized polarized reflectance of cloudless pixels is done. Analyzing the model, we find that the curves of reflectance with scattering angle for cloudless pixels should be smooth and predictable, while the curves for cloud pixels should be very different because of the more complicated scattering and the more diverse shapes or sizes of cloud particles. In order to describe this regularity, a fitting error between measured polarized reflectance and simulated one was used as the third criterion in this paper, which is defined as:

$$
Fitting\; error = \frac{1}{m} \sum_{i=1}^{m} \frac{|R_{\text{sim},i} - R_{\text{meas},i}|}{R_{\text{meas},i}}
$$

(16)

where $i$ is the number of effective data layers and $m$ is 12 for POLDER and 8 for MAPI. The greater the difference between measured polarized reflectance and simulated one is, the bigger the fitting error is, and the more likely the target is to be cloudy. This is the main basis of this cloud detection algorithm.

We should also note that, because of different physical process and properties between ice crystals and liquid water drops, liquid clouds are different from ice clouds on the polarization features. Many researchers have done such simulations and detailed research on this aspect (Chepfer et al., 2001; Goloub et al., 2000). The results showed that the polarization of liquid clouds exhibits a strong maximum at about 140° from the incoming incident direction, which makes it easily detectable. While for ice clouds, the main feature is the general decrease of the polarization for increasing scattering angle, which may be similar with some cloudless pixels.

### 3. Discussion of the Thresholds for $R_{670}$, NDVI, and Fitting Error

To demonstrate the effects of the three criterions in cloud detecting, pixels labeled as cloud/cloudless in POLDER level 1 data were used to analyze the pairwise relationships between them, as shown in Figure 2. Figures 2a, 2e, and 2i show that, although there are some intersections, cloud and cloudless pixels obviously separate from each other in these three features.

Figures 2b, 2c, and 2f reveal the pairwise relationships between $R_{670}$, NDVI, and fitting error. There is an approximate inverse proportional relationship between $R_{670}$ and NDVI. While the correlation between
fitting error and NDVI/R\textsubscript{670} is not definite. Furthermore, it also exhibits different cluster features between cloud and cloudless pixels in these subgraphs, which can be used to detect cloud.

In addition, we have discussed how to select the thresholds for R\textsubscript{670}, NDVI, and fitting error to improve the cloud detection. The accuracy of the cloud detection is defined as the proportion of the number of the pixels, which are both declared as cloud or cloudless by POLDER official cloud mask product and the algorithm in this paper, in all pixels. An experimental area (called area A) was extracted from one of POLDER Level 1 files (2011.03.28 05:50:29) as an example.

When NDVI\textsubscript{threshold} and Fitting error\textsubscript{threshold} were fixed and R\textsubscript{670,threshold} changed from small to large, the cloud detection accuracy showed the trend of underdetection to overdetection, which was unimodal. So there should be an optimal value for R\textsubscript{670,threshold}. The same analysis had been applied to the other two

Figure 3. The accuracy varied with the change of NDVI\textsubscript{threshold}, Fitting error\textsubscript{threshold}, and R\textsubscript{670,threshold}; the indicator = 0 stands for cloudless, 100 for cloud, and 50 for undetermined.
thresholds, and similar conclusions were drawn, as shown in Figure 3. Moreover, for the sake of symmetry, a negative value of NDVI\text{threshold} was used for the lower bound.

Finally, the pixels of Fitting error in \((0, 0.98\times \text{Fitting error}_{\text{threshold}})\) are declared as cloudless pixels, and those of \([0.98\times \text{Fitting error}_{\text{threshold}}, \text{Fitting error}_{\text{threshold}}]\) are declared as undetermined pixels, \((\text{Fitting error}_{\text{threshold}}, \text{Fitting error}_{\text{max}}]\) are declared as cloud pixels.

However, we should note that one single value for every threshold cannot be strictly optimal for all situations, which may be very different in atmosphere conditions, surface types, and data qualities. Therefore, the thresholds given in this paper (NDVI\text{threshold} = \pm 0.1, R_{670}\text{threshold} = 0.3, \text{Fitting error}_{\text{threshold}} = 2.0) are empirical values which are relatively suitable for POLDER Level 1 data.

For threshold selection of MAPI, the process is the same as above, but due to different sensors, the values are adjusted to NDVI\text{threshold} = \pm 0.1, R_{670}\text{threshold} = 0.3, and Fitting error\text{threshold} = 2.3.

4. Verification of Cloud Detection Results

4.1. Cloud Detection Results of POLDER Data and Comparisons With MYD35

Two experimental areas (area A, mentioned in section 3, and area B) were extracted from one of POLDER Level 1 files (2011.03.28 05:50:29), and repeated experiments were conducted. The reflectance of 670-, 565-, and 443-nm channels were used as RGB channels to form true color remote sensing images, as shown in Figures 4a and 5a, respectively. Then, four corresponding MYD35 files are used to verify the accuracy of cloud detection results for both area (2011.03.28 04:45, 04:50, 06:25, 06:30 for area A, 2011.03.28 04:50, 04:55, 06:25, 06:30 for area B). Because the transit times of POLDER and MODIS are not entirely uniform, so the shapes of clouds cannot be very consistent, which can be visually seen in the true color image Figures 4a, 4d, 5a, and 5d.
The 17 points were taken out of the experimental area A, and the curves of polarized reflectance (670-nm channel) with scattering angle of these points were plotted. The model of cloud detection algorithm was used to fit these curves (Figure 6).

We can see that the simulation can well fit the multiangular reflectance curves of most cloudless pixels, while the fitting effect is greatly reduced for cloud pixels, which means that for cloudless ones, the Fitting errors are much smaller than cloud ones. The Fitting errors for cloudless pixel nos. 0, 1, and 2 equal to 0.45998, 0.61462, and 0.29356, respectively, while for cloud pixel nos. 7, 8, and 9 reach 2.9153, 5.1665, and 5.0309, respectively. The results are consistent with previous analysis.

The cloud detection results of the experimental areas are shown in Figure 4c/Figure 5c. Comparing with the RGB images, this algorithm can effectively distinguish cloud/cloudless pixels and capture cloud targets in remote sensing data. The accuracy can reach 92.3915% and 91.4740% for area A and area B, separately.

And comparing Figures 4c, 4d, and 4e/ Figures 5c, 5d, and 5e, we can see that some obviously cloudless pixels are misjudged as cloud pixels in MYD35, but in POLDER cloud mask product and the result of this algorithm, there is no such kind of misjudgment. This also proves that the multiangular polarization method is more suitable than the traditional remote sensing method in the field of cloud detection.

Figure 6. Different characteristics of multiangular polarized reflectance curve and fitting results for cloudless pixels (nos. 0, 1, 2) and cloud pixels (nos. 7, 8, 9) over land.

Figure 7. The cloud detection result for scene 1: 2011.04.16. (a) The real color image. (b) The cloud detection result. 0 for clear, 50 for undetermined, and 100 for cloudy.
In addition, cloud detection results of two typical scenes in China region on 2011.04.16 and 2012.10.01 were shown in Figures 7 and 8. The accuracy has decreased in larger scenes as we can see, and little part of cloud pixels were misjudged, that is because a single set of thresholds cannot fit in all cases, just as discussed in section 3. In short, most cloud pixels were identified, and the accuracy and applicability of our algorithm were further verified.

4.2. Cloud Detection Results of MAPI Data

Two experimental scenes were selected from one of MAPI Level 1 files (2017.02.24 13:37:14, 2017.02.20 14:33:15). Because the detector lacks blue band, the reflectance of 670-, 565-, and 765-nm channels in the data are used as color channels to form pseudo-color remote sensing image, shown in Figures 9a and 10a. However, pseudo-color images do not affect eyes to directly distinguish the cloudless/cloud pixels.

The cloud detection results are shown in Figures 9b and 10b. Similarly, cloudless/cloud pixels are effectively distinguished and most of the cloud targets in the images are captured. However, as we can see, some cloud pixels in the area were misjudged. According to the phase characteristics of ice cloud (section 2.2.3), these pixels may be in the ice cloud area, and the features are so similar with cloudless pixels that it leads to some misjudgments.

Overall, two groups of experiments for POLDER and MAPI data verify the reliability and applicability of the cloud detection algorithm for multiangular polarized remote sensing data, and this algorithm can produce cloud mask products for Tiangong-2 MAPI Level 1 data.

But at the same time, we also find some limitations of this algorithm. Some pixels in the ice cloud area may be misjudged and missed because of the similar polarized characteristics as the cloudless ones, which increase the difficulty for the selection and adjustment of thresholds. And this algorithm may fail in the case of snow-covered surfaces. In addition, Figures 4b and 4c/Figures 5b and 5c reveal that a few discrete cloudless pixels can be declared as cloudy, which is due to the fluctuation of the recorded reflectance of these points for some uncertain reasons. From Figure 7–10, it seems that our algorithm underestimate part of
cloud targets in large scenes, which maybe because the thresholds are not suitable enough for all situations. These are the future efforts for further research and progress.

5. Conclusions

This paper describes a new cloud detection algorithm over land based on the scattering polarized phase function theory of the TOA cloud and cloudless targets, using satellite multangular polarized remote sensing data. The main innovation lies in incorporating the polarized scattering simulation of cloudless targets over land into cloud detection algorithm, making use of the different characteristics between cloud and cloudless targets, which is rarely seen in other algorithms. This method mainly utilizes the unique information of the multangle polarized observation in the cloud mask processing at first time.

Applying this algorithm to the cloud mask processing of POLDER data, the results has a good consistency of about 90% with official cloud mask products, proving the accuracy of this algorithm which uses only three thresholds to make the algorithm more convenient but also effective. The result from this algorithm is comparable to POLDER's official cloud detection algorithm. In addition, this method is applied to MAPI Level 1 data to correctly capture cloud targets in MAPI images, and provide the valuable utilization for other products generation from the MAPI Level 1 data.

References

Bodhaine, B. A., Wood, N. B., Dutton, E. G., & Slusser, J. R. (1999). On Rayleigh optical depth calculations. Journal of Atmospheric and Oceanic Technology, 16(11), 1854–1861. https://doi.org/10.1175/1520-0426(1999)016<1854:ORODCA>2.0.CO;2

Bréon, F.‐M. (Ed.) (2003). POLDER level‐1 product data format and user manual.

Brière, J. C., Vanhouce, C., Parol, F., Goloub, P., Herman, M., Bonnel, B., et al. (1997). Cloud detection and derivation of cloud properties from POLDER. International Journal of Remote Sensing, 18(13), 2785–2813. https://doi.org/10.1080/014316917217332

Chen, P. Y., Srinivasan, R., Fedosejevs, G., & Narasimhan, B. (2002). An automated cloud detection method for daily NOAA‐14 AVHRR data for Texas, USA. International Journal of Remote Sensing, 23(15), 2939–2950. https://doi.org/10.1080/01431160110075631

Chen, Z., Sun, X., & Qiao, Y. (2018). Cloud detection over ocean from PARASOL/POLDER3 satellite data. Journal of Remote Sensing, 22(6), 996–1004. https://doi.org/10.1134/jrs.20187366

Cheng, T., Gu, X., Chen, L., & Yu, T. (2007). Cloud detection based on the spectral, multi‐angular, and polarized characteristics of cloud. IEEE International Geoscience & Remote Sensing Symposium. IEEE. https://doi.org/10.1109/IGARSS.2007.4423555

Chepfer, H., Goloub, P., Riedi, J., Haan, J. F. D., Hovenier, J. W., & Flamant, P. H. (2001). Ice crystal shapes in cirrus clouds derived from POLDER/ADEOS‐1. Journal of Geophysical Research, 106(D8), 7955–7966. https://doi.org/10.1029/2000JD000285

Derrien, M., Farki, B., Harang, L., LeGléau, H., Noyalet, A., Pochic, D., & Sairouni, A. (1993). Automatic cloud detection applied to NOAA‐11/AVHRR imagery. Remote Sensing of Environment, 46(3), 246–267. https://doi.org/10.1016/00344257(93)90046‐z

Desmons, M., Ferlay, N., Parol, F., Riedi, J., & Thieuleux, F. (2017). A global multilayer cloud identification with POLDER/PARASOL. Journal of Applied Meteorology and Climatology, 56(4), 1121–1139. https://doi.org/10.1175/JAMC‐D‐16‐0159.1

Deuzé, J. L., Bréon, F. M., Devaux, C., Goloub, P., Herman, M., Lafrance, B., et al. (2003). Remote sensing of aerosols over land surfaces from POLDER‐ADEOS‐1 polarized measurements. Journal of Geophysical Research, 108(D5), 4913–4926. https://doi.org/10.1029/2000JD000364

Goloub, P., Herman, M., Chepfer, H., Riedi, J., Brogniez, G., Couvert, P., & Sée, G. (2000). Cloud thermodynamical phase classification from the polder spaceborne instrument. Journal of Geophysical Research, 105(D11), 14,747–14,759. https://doi.org/10.1029/1999JD9001183

Key, J., & Barry, R. G. (1989). Cloud cover analysis with arctic AVHRR data: 1. Cloud detection. Journal of Geophysical Research, 94(D15), 18,521–18,535. https://doi.org/10.1029/JD094iD15p18521

Li, C., Ma, J., Yang, P., & Li, Z. (2018). Detection of cloud cover using dynamic thresholds and radiative transfer models from the polarization satellite image. Journal of Quantitative Spectroscopy and Radiative Transfer, 222–223, 196–214. https://doi.org/10.1016/j.jqsrt.2018.10.026

Li, Z., Goloub, P., Devaux, C., Gu, X., Qiao, Y., Zhao, F., & Chen, H. (2004). Aerosol polarized phase function and single-scattering albedo retrieved from ground-based measurements. Atmospheric Research, 71(4), 233–241. https://doi.org/10.1016/j.atmosres.2004.06.001

Acknowledgments

This research was funded by the National Key R&D Program of China (grants 2018YFB0504900, 2018YFB0504901) and the National Natural Science Foundation of China (grants 41871249, 41471302). POLDER Level 1 data are obtained from ICARE Data and Services Center (France, http://www.icare.unil-lille1.fr/). MCD12 data sets are all downloaded from http://ladsweb.nascom.nasa.gov/. MCD12 data sets are all downloaded from http://ladsweb.nascom.nasa.gov/.

MAPI Level 1 data are provided by the Shanghai Institute of Technical Physics, Chinese Academy of Sciences.

Figure 10. (a) Pseudo‐color image of the scene 2: 2017.02.20 14:33:15. (b) Cloud detection result of this scene.
Nadal, F., & Bréon, F. M. (1999). Parameterization of surface polarized reflectance derived from POLDER spaceborne measurements. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3), 1709–1718. https://doi.org/10.1109/36.763292

Vermote, E., Tanré, D., Deuzé, J. L., Herman, M., Morcrette, J. J., & Kotchenova, S. Y. (2006). Second Simulation of a Satellite Signal in the Solar Spectrum–Vector (6SV), from http://6s.ltdri.org/pages/manual.html