Development and Evaluation of AMSU-A Cloud Detection over the Tibetan Plateau

Jiawen Wu, Zhengkun Qin, Juan Li and Zhiwen Wu

School of Atmospheric Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China; 20198300040@nuist.edu.cn
Centre of Data Assimilation for Research and Application, Nanjing University of Information Science and Technology, Nanjing 210044, China; 20191201084@nuist.edu.cn
Earth System Modeling and Prediction Centre, China Meteorological Administration, Beijing 100081, China; lj@cmo.gov.cn
State Key Laboratory of Severe Weather, Institute of Meteorological, China Meteorological Administration, Beijing 100081, China
* Correspondence: qzk_0@nuist.edu.cn

Abstract: Advanced Microwave Sounding Unit-A (AMSU-A) and Microwave Humidity Sounder (MHS) data have been widely assimilated in operational forecasting systems. However, effective distinction between cloudy and clear-sky data is still an essential prerequisite for the assimilation of microwave observations. Cloud detection over the Tibetan Plateau has long been a challenge owing to the influence of low temperatures, terrain height, surface vegetation, and inaccurate background fields. Based on the variations in the response characteristics of different channels of AMSU-A to clouds, five AMSU-A window and low-peaking channels (channels 1–4 and 15) are chosen to establish a cloud detection index. Combined with the existing MHS cloud detection index, a cloud detection scheme over the Tibetan Plateau is proposed. Referring to VISSR-II (Stretched Visible and Infrared Spin Scan Radiometer-II) and CALIPSO (The Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation) cloud classification products, the detection rate of cloudy data and the rejection rate of clear-sky data under different cloud index thresholds are evaluated. Results show that the new cloud detection scheme can identify more than 80% of cloudy data on average, but this decreases to 72% for area with terrain higher than 5 km, and the false deletion rate remains stable at 45%. The detection rates of mixed clouds and cumulonimbus are higher than 90%, but it is lower than 50% for altostratus with an altitude of about 7–8 km. Comparative analysis shows that the new method is more suitable for areas with terrain higher than 700 m. Based on the cloud detection results, the effects of terrain height on the characteristics of observation error and bias are also discussed for AMSU-A channels 5 and 6.

Keywords: AMSU-A; MHS; Tibetan Plateau; cloud detection

1. Introduction

The Tibetan Plateau is a sensitive area in terms of global climate, and its surroundings also serve as a key precursor-signal area regarding the occurrence of major disastrous events in China. Therefore, meteorological research in the Tibetan Plateau region is of critical importance. Multiple reanalysis datasets provide spatial and temporal continuity for studying the characteristics of changes in weather and climate over the Tibetan Plateau. However, from the numerous results published on the applicability of reanalysis data to the Tibetan Plateau, it can be seen that, although reanalysis data incorporate many ground-based and airborne observations, the errors of these reanalysis data are still significantly higher in the Tibetan Plateau region than in surrounding areas [1–10]; plus, there are considerable differences among different reanalysis datasets, and no single reanalysis dataset can generally be regarded as better than other data sources.
As a sensitive region for weather forecasting in China, the accuracy of the initial field in numerical models in the Tibetan Plateau region inevitably affects the level of numerical forecasting of high-impact catastrophic weather in the downstream region of the plateau. In the 1990s, Eyre et al. directly assimilated the data from satellite-based vertical atmospheric sounders within the framework of variational data assimilation, which led to a new era of applying satellite data in numerical forecasting [11,12]. This helped to address the problem of a lack of observational data in data assimilation studies, and the assimilation of various types of satellite observations has since significantly improved the accuracy of the initial field in numerical forecasting. Among many types of satellite data that have been assimilated, microwave-sounding data have the longest history, and specifically, AMSU-A (Advanced Microwave Sounding Unit-A) data have become the most influential microwave-sounding observations in all operational forecasting systems [13–15]. However, the improvement and extended application of assimilation techniques for microwave-sounding data is still a key priority in assimilation research. Currently, such research focuses on developing more detailed quality-control methods [16,17], how to assimilate microwave-sounding surface channels [18–20], and assimilation techniques for satellite data in cloudy areas [21–24]. This latter topic is aimed at quantitatively assessing the variability of the simulation bias in cloudy areas and clarifying the error covariance characteristics of the fast radiative transfer pattern in cloudy areas, which is one of the main difficulties in the assimilation of microwave-sounding data in cloudy areas [25–28].

Since satellites observe the Earth from outer space, AMSU-A can penetrate non-precipitating clouds, but the instrument is still inevitably affected by cloud radiation, scattering, and reflection of cloud radiation to the ground surface. Therefore, scientists have developed different assimilation methods for clear-sky data and cloudy data. However, first, for both cloudy and clear-sky data assimilation, it is necessary to effectively distinguish between cloudy and clear-sky data, so an efficient cloud detection method is very important toward improving the assimilation of AMSU-A data.

Over ocean areas, since the surface emissivity of the ocean surface is significantly lower than that of clouds, various cloud parameter inversion algorithms have been developed to perform cloud detection based on the inversion results [29–31]. However, the situation over land areas is more complicated. The surface emissivity of the land surface is closer to that of clouds, making it difficult to determine whether anomalous changes in the observed brightness temperature (BT) are caused by surface emissivity or by clouds. Several major previously developed cloud detection methods for microwave sounding data over land are given in Table 1. Scientists developed many empirical statistical methods for AMSU-A and MHS (Microwave Humidity Sounder) cloud detection methods over land areas [32–35], but these approaches demand background field information, and the detection results are highly dependent on the accuracy of the pattern background field. Moreover, it is difficult to obtain accurate model background fields for the Tibetan Plateau region, where short-term convective systems are prevalent. Aires et al. [36] used the MSG-SEVIRI (Meteosat Second Generation–Spinning Enhanced Visible and Infrared Imager) cloud product as a reference to train AMSU-A/B observations with a neural network algorithm and proposed a land and ocean cloud classification method. Neural network algorithms rely on a vast number of accurately labeled datasets to train microwave information. This method is prone to overfitting and increases the computational burden, leading to poor feasibility in business applications. Given the low accuracy of the model simulation of BT, many scientists have proposed various schemes to invert the surface temperature and surface emissivity, which have markedly improved the clear-sky assimilation of terrestrial observations [37,38]. However, most of these schemes rely on cloud inversion products from other satellite-based instruments (e.g., MODIS cloud). Thus, in summary, how to eliminate the influence of surface emissivity in cloud detection remains one of the key challenges in research on terrestrial cloud detection by AMSU-A and MHS.

In the Tibetan Plateau region, with its high altitude, the process of cloud detection also faces the influence of several other problems, such as the lower surface temperature and
ground snow, resulting in more-complicated variations in surface emissivity. Moreover, short-term convective systems are commonplace in the Tibetan Plateau region, which requires the development of fast cloud detection algorithms that only depend on satellite observations. Wu et al. [39] developed an AMSU-A cloud index for the plain area based on the differences in the response characteristics of different channels of AMSU-A to clouds and selected five AMSU-A windows and low-peaking channels (channels 15 and 1–4). The results showed that the AMSU-A cloud index can detect most convective clouds but misses cirrus and some cirrostratus clouds; however, after adding the matched MHS cloud index, most of the clouds missed by the AMSU-A index could be compensated for. The new method only depends on the observation data, which can well avoid the influence of background field error on the cloud detection method. Using the data of a single observation point can well eliminate the influence of surface emissivity, which is difficult to estimate accurately. The calculation is simple and efficient and also ensures that the method can effectively meet the needs of data assimilation. Therefore, this method has a good application prospect in the study of plateau data assimilation. However, the effectiveness of this cloud detection method in areas with low surface temperatures, such as high latitudes or glacial surfaces, was not evaluated. Accordingly, this paper attempts to apply the cloud detection method to the Tibetan Plateau region to test the effectiveness of the AMSU-A cloud detection index over the Tibetan Plateau region. Based on the accurate identification of clear-sky data, the error and bias characteristics of the AMSU-A mid-peaking channels in the Tibetan Plateau region are evaluated for different surface types and terrain heights, and a foundation is laid for establishing a more effective bias-correction scheme.

The structure of the paper is as follows: Following this introduction, Section 2 introduces the study area and data; Section 3 describes the cloud detection method for the Tibetan Plateau region; Section 4 assesses the cloud detection results and evaluates the error and bias characteristics of the AMSU-A mid-peaking channels under clear-sky conditions; and Section 5 provides a summary and some further discussion.

Table 1. Similar cloud detection techniques and main parameters.

<table>
<thead>
<tr>
<th>Author</th>
<th>Region</th>
<th>Method</th>
<th>Application</th>
<th>Other Supplementary Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ferraro et al.</td>
<td>The West Coast of the United States</td>
<td>BTs differences between AMSU-A channel 1 and 15 and between AMSU-B channel 1 and 2.</td>
<td>AMSU-A/B</td>
<td>A monthly rainfall climatology [41] based on 11 years of SSM/I measurements (1987–1998)</td>
</tr>
<tr>
<td>Bennartz et al.</td>
<td>The Baltic region</td>
<td>BTs differences between AMSU-A channel 1 and AMSU-B channel 2.</td>
<td>AMSU-A/B</td>
<td>None</td>
</tr>
<tr>
<td>Geer et al.</td>
<td>60°N to 60°S</td>
<td>Differences between the observed (O) and simulated (B) BTs of AMSU-A channel 4.</td>
<td>AMSU-A</td>
<td>Simulated BTs</td>
</tr>
<tr>
<td>Aires et al.</td>
<td>120°W–60°E 50°N–50°S</td>
<td>Relationship between the cloud product and observed BTs of all of AMSU-A/B channels</td>
<td>AMSU-A/B</td>
<td>MSG-SEVIRI (Meteosat Third Generation–Spinning Enhanced Visible and Infrared Imager)</td>
</tr>
</tbody>
</table>

2. Materials

2.1. Study Area

The Tibetan Plateau, often referred to as the “roof of the world”, is the largest plateau in China and the highest in the world. It is located south of the Himalayas, north of the Kunlun and Altun Mountains, west of the Pamir Plateau, east of the Qinling Mountains, and connects with the Loess Plateau. Its length is ~2800 km, its width is ~300–1500 km, and the total area is ~2.5 million km². According to the terrain, it can be divided into
six parts: Qiangtang Plateau, Southern Tibet Valley, Qaidam Basin, Qilian Mountains, Qinghai Plateau, and Sichuan–Tibet Plateau Valley. It is an important task to study and monitor the meteorology of the Tibetan Plateau. In this paper, the Tibetan Plateau and the nearby highlands above 500 m were selected as the study area, in order to maintain topographic coherence; scattered regions located north of 35°N that were lower than 500 m were also included as the study area. The topographic distribution of the study area is given in Figure 1, and the elevation data were 2-Minute Gridded Global Relief Data (ETOPO2) v2 from the NCEI (National Centers for Environmental Information).

![Figure 1. Topographical map of the study area.](image_url)

### 2.2. AMSU-A and MHS

In this paper, the observed BTs of AMSU-A and MHS carried by NOAA19 were taken as the study data, and the time range was from August 1 to 31, 2019. The NOAA-19 satellite was launched on 6 February 2009 aboard the Delta-2 launch vehicle 7320-10C from Vandenberg Air Force Base Launch Complex 2. NOAA-19 carries a range of instruments that provide data for climate and weather prediction. Similar to previous satellites in the series, NOAA-19 provides global images of cloud and surface characteristics and the vertical distribution of atmospheric temperature and humidity for application in numerical prediction models of weather and oceans. In addition, the distribution of ozone in the upper atmosphere and near-Earth space is also measured, which is crucial in the fields of oceanography, aviation, power generation, and agriculture. This paper makes use of data from the AMSU-A and MHS instruments onboard NOAA-19. Table 2 details each channel of AMSU-A and MHS.

AMSU-A has 15 channels: channels 1, 2, 3, and 15 are four window channels corresponding to 23.8, 31.4, 50.30, and 89.0 GHz, respectively, mainly used for detecting rain clouds and providing surface temperature and emissivity information; channels 4–14 are 11 oxygen absorption channels with a frequency distribution of 50–60 GHz, used for detecting atmospheric temperature profiles from the troposphere to the stratosphere. MHS has five channels: channels 1 and 2 are two window channels (89 and 157 GHz), and channels 3–5 are three water vapor channels, corresponding to frequencies of around 181.3 GHz, mainly observing water vapor and providing information on rain clouds in the lower and middle troposphere, which are more sensitive to clouds, especially those with ice particles.
Table 2. NOAA-19 AMSU-A and MHS channel characteristics.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Channel</th>
<th>Central Frequency (GHz)</th>
<th>Polarization</th>
<th>Height of Peak Energy Contribution (hPa)</th>
<th>NEΔT (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMSU-A</td>
<td>1</td>
<td>23.80</td>
<td>V</td>
<td>Window</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>31.40</td>
<td>V</td>
<td>Window</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>50.30</td>
<td>V</td>
<td>Window</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>52.80</td>
<td>V</td>
<td>Window</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>53.59 ± 0.115</td>
<td>H</td>
<td>850</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>54.40</td>
<td>H</td>
<td>700</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>54.94</td>
<td>V</td>
<td>400</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>55.50</td>
<td>H</td>
<td>270</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>f₀ = 57.29</td>
<td>H</td>
<td>180</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>f₀ ± 0.217</td>
<td>H</td>
<td>90</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>f₀ ± 0.322 ± 0.048</td>
<td>H</td>
<td>50</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>f₀ ± 0.322 ± 0.022</td>
<td>H</td>
<td>25</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>f₀ ± 0.322 ± 0.010</td>
<td>H</td>
<td>12</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>f₀ ± 0.322 ± 0.004</td>
<td>H</td>
<td>5</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>89.00</td>
<td>H</td>
<td>Window</td>
<td>1.20</td>
</tr>
<tr>
<td>MHS</td>
<td>1</td>
<td>89.00</td>
<td>V</td>
<td>Window</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>157.0</td>
<td>V</td>
<td>Window</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>183.31 ± 1.0</td>
<td>V</td>
<td>300</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>183.31 ± 3.0</td>
<td>H</td>
<td>500</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>190.0</td>
<td>V</td>
<td>800</td>
<td>0.60</td>
</tr>
</tbody>
</table>

AMSU-A and MHS are both cross-orbit scans with sub-stellar point resolutions of around 48 and 17 km, with the latter having approximately three times the resolution of the former. In one scan line, there are 30 and 90 fields of view (FOVs), respectively, which is comparable to one AMSU-A FOV translating to around nine MHS FOVs. The width of AMSU-A is around 2226.8 km, whereas the width of MHS is approximately 2348 m. The gap between the two widths is negligible when compared to the observation height of more than 800 km, and similar widths can practically assure the overlap of the two FOVs under different scanning angles. Moreover, both AMSU-A and MHS are installed on the same polar-orbiting satellite. Thus, the time deviation between them is negligible, so the FOVs of both are approximately overlapping.

2.3. VISSR-II Cloud Classification Product

In order to verify the effectiveness of the proposed cloud detection method, the cloud product data of geostationary satellites were selected as the verification data. Specifically, this paper selected the cloud classification products of the Stretched Visible and Infrared Spin Scan Radiometer-II (VISSR-II), which is onboard China’s FengYun-2H satellite. VISSR-II includes one visible light and four infrared channels, which can provide real-time cloud images and dozens of remote-sensing products, such as atmospheric radiation, cloud wind, and dust, thereby providing reference data for weather forecasting, disaster warning, environmental monitoring, etc. It can also enrich the data sources for global numerical weather forecasting. The spatial resolution of the cloud classification products is 5 km, and the temporal resolution is 30 min. The cloud classification products are effective at identifying cumulonimbus, stratocumulus, mid-height cloud, cirrus, and other cloud types with meteorological significance. Table 3 describes the VISSR-II cloud classification products.

Because of the differences in the temporal and spatial resolution of the observed BT of AMSU-A and cloud classification products of VISSR-II, it is necessary to match the two in time and space. Cloud types, unlike air pressure and temperature, are not continuous, and there is, frequently, more than one type of cloud in the same AMSU-A FOV. As a result, the cloud type that occurs the most in a scanned AMSU-A FOV was selected as the representative cloud type for this FOV, and the diameter of the FOV was chosen as 25 km.
based on the size of the scanned AMSU-A FOV. In terms of time, the closest time between the two observations was selected, and the maximum time error was within 3 h to ensure that the cloud system did not change dramatically within the time error. From Figure 2a,b, it can be seen that the matching method only ignored some broken clouds in the clear-sky area, and the main cloud systems were well matched.

Table 3. Description of 2VISSR-II cloud classification products.

<table>
<thead>
<tr>
<th>Cloud Class Definition</th>
<th>Cloud Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear ocean</td>
<td>Clear</td>
</tr>
<tr>
<td>Clear land</td>
<td>Clear</td>
</tr>
<tr>
<td>Mixed pixels</td>
<td>Mixed</td>
</tr>
<tr>
<td>Nimbostratus or altostratus</td>
<td>Ns and As</td>
</tr>
<tr>
<td>Cirrostratus</td>
<td>Cs</td>
</tr>
<tr>
<td>Cirrus dens</td>
<td>Ci</td>
</tr>
<tr>
<td>Cumulonimbus</td>
<td>Cb</td>
</tr>
<tr>
<td>Stratocumulus or altocumulus</td>
<td>Sc and Ac</td>
</tr>
</tbody>
</table>

Figure 2. The (a) VISSR-II cloud classification products, (b) matched cloud classification products, and (c) spatial distribution of O–B (unit: K) of AMSU-A channel 3 over the Tibetan Plateau region at 1200 UTC 12 August 2019.

2.4. CALIPSO LIDAR Level 2 Version 4.2, 1 km Cloud Layer

The Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) is an Earth-observing satellite project initiated and executed by NASA’s Langley Research Center and the National Space Research Center of France in 2006. It orbits the Earth
at 705 km altitude, 1.55°, and 96 min orbital parameters, respectively, for 16 days to retrieve data on a three-dimensional distribution of the cloud and aerosol layers around the globe [42]. The CALIOP lidar is a key instrument of the CALIPSO satellite as it provides polarization backscattering vertical profile data for clouds and aerosols between south and north latitude 82° on a 532 nm channel and attenuation scattering vertical profile data on a 1064 nm channel [43,44]. CALIOP can be used to accurately reconstruct the vertical distribution of clouds and aerosols as well as to quantify their size, irregularity, and type of aerosol particles, by considering vertically resolved backscattering coefficients with high resolution. This study used the CALIPSO LIDAR Level 2 Version 4.2, 1 km Cloud Layer and matched it to AMSU-A using the same space–time matching method as the VISSR-II Cloud Classification Product.

3. Methods

This study used the Community Radiative Transfer Model (CRTM) developed by the Joint Center for Satellite Data Assimilation to simulate the BT of each channel of AMSU-A, with FNL (final analysis) data as the background field, and the NPOESS (National Polar-orbiting Operational Environmental Satellite System) dataset as the land-surface type of each FOV. The BT of each channel of AMSU-A was simulated against the background of clear-sky conditions, and the distribution of the difference between the observed and simulated BT (O–B) of each channel was calculated. Affected by the scattering and absorption properties of water and ice particles within the clouds, the irradiance emitted from the surface is weakened to different degrees when passing through clouds, while the surface emissivity is generally around 0.9, which is significantly higher than that of clouds, resulting in the irradiance received by the satellite in cloudy regions being significantly lower than the simulated clear-sky irradiance. As a result, the O–B in cloudy regions will show negative values but with large absolute values. Comparing Figure 2a,c, it can be seen that, over the Tibetan Plateau region, the O–B of window channel 3 did not show negative values with large absolute values over the cloudy region, nor did it show significantly smaller absolute values over the clear-sky region. This was because the simulation error of the model was substantial over the plateau region, due to many complex factors such as surface temperature and vegetation type. Therefore, cloud detection methods that depend heavily on the accuracy of the background field information are not applicable over the plateau region.

Although microwave radiation can penetrate non-precipitating clouds, the radiation from clouds and the scattering and reflection of ground-based radiation from clouds still have a significant effect on the observed BT of AMSU-A. The frequency of each channel of AMSU-A is different, and therefore, the response of each channel to clouds also differs. In the window channels of AMSU-A, the weighting function peak heights are located at the ground, and the observed radiation intensity of the channels depends mainly on the surface emitted radiation, that is, it depends mainly on the surface temperature and is affected by both the surface emissivity and atmospheric transmittance. For deep cloud systems, when the peaking height of the channel’s weighting function is lower than the height of the cloud top, the observed radiation intensity of the channel mainly comes from the thermal radiation of the cloud itself, so these channels mainly detect the information of the cloud top.

The window channels of AMSU-A are more sensitive to the presence of clouds and precipitation [45]. In areas where convection is strong, the water content in the cloud is high, and other solid precipitation particles may exist. Due to the strong scattering and low emissivity of the cloud, the observed BT in cloudy areas will be significantly reduced. As shown in Figure 3a–e, in the cloudy area, the observed BTs of AMSU-A channels showed obvious low-value centers. Based on the different responses of channel 15 and channel 3 to clouds (channel 15 responds more strongly than channel 3), Wu et al. [39] selected five
low-peaking channels (channels 1–4 and 15) of AMSU-A to construct an AMSU-A cloud index, and the AMSU-A cloud index was defined as follows:

\[
A_{\text{index}} = \frac{T_{\text{normalized}}}{\frac{T_{b,3} - 200}{50}}
\]

\[
T_{\text{normalized}} = \frac{T_{b,3} - \mu}{\sigma};
\]

\[
\mu = \frac{1}{5} \sum_{i=1}^{5} T_{b,i}, \quad \sigma = \sqrt{\frac{1}{5} \sum_{i=1}^{5} (T_{b,i} - \mu)^2}.
\]

where \(T_{b,i}\) is the observed BT of the \(i\)th channel of the five channels 1–4 and 15 of AMSU-A.

Figure 3. Spatial distribution of the observed BT of AMSU-A (a) channel 15 and (b–e) channels 1–4 (unit: K) over the Tibetan Plateau region at 1200 UTC 12 August 2019.

Comparing Figure 3a,e, it can be seen that the observed BT of AMSU-A channel 15 was significantly lower than that of channel 3 over the cloudy region, indicating that the sensitivity of channel 15 to clouds was also stronger than that of channel 3 over the Tibetan Plateau region, meaning this AMSU-A cloud index is also applicable to the Tibetan Plateau region. Wu et al. [39] showed that the AMSU-A cloud index can detect most of the deep convective clouds but misses cirrus and some cirrostratus clouds; however, after adding
the matched MHS cloud index, most of the clouds missed by the AMSU-A cloud index could be compensated for, and most of the cloudy observations could be eliminated. The MHS cloud index is not easily affected by the low temperature of the plateau but reflects more of the water vapor change. Thus, the MHS cloud index can still play an important role over the Tibetan Plateau region. Referring to the study of Wu et al. [39], this paper attempted to use the AMSU-A cloud index and MHS cloud index for cloud detection over the Tibetan Plateau region. The MHS cloud index is defined as follows:

\[
M_{\text{index}} = \frac{T_{\text{normalized}}^{b,1}}{1 \times (T_{b,2} - T_{\text{norm}}^{b,1})}
\]

\[
T_{\text{normalized}}^{b,1} = \frac{T_{b,1} - \mu}{\sigma},
\]

\[
\mu = \frac{1}{5} \sum_{i=1}^{5} T_{b,i}, \quad \sigma = \sqrt{\frac{1}{5} \sum_{i=1}^{5} (T_{b,i} - \mu)^2}.
\]

where \(T_{b,i}\) is the observed BT of the \(i\)th channel of the five channels, 1–5, of MHS. The MHS cloud index was matched to the FOV of AMSU-A according to the instrumental characteristics of one MHS FOV corresponding to nine AMSU-A FOVs.

Figure 4 shows the cloudy and clear-sky areas identified based on the old thresholds of the two cloud indices. It is clear that the old thresholds were not effective in detecting clouds over the Tibetan Plateau region. The AMSU-A cloud index was established by the BTs of the window and low-peaking channels. The overall temperature of the Tibetan Plateau region is significantly lower than that of the plains; thus, the observed BT of channel 15 in the denominator of the AMSU-A cloud index is also significantly lower. However, the numerator of the AMSU-A cloud index (the normalized observed BT of channel 3) does not differ significantly between the plains and the plateau region. Therefore, it is the low temperature in the Tibetan plateau region that causes the old threshold to become inapplicable and require modification. The MHS cloud index is barely affected by the low temperature of the plateau, and the old threshold can still detect some cloud systems, so only a small adjustment of the threshold of the MHS cloud index is needed. Therefore, it is necessary to provide reasonable thresholds for the cloud indices based on the practical circumstances of the plateau, and this study utilized the following two quantitative assessment criteria for the threshold adjustment analysis:

\[
Pd = \frac{Nd}{Ncld} \times 100\%;
\]

\[
Pclr = \frac{Nclrm}{Nclr} \times 100\%;
\]

where \(Pd\) is the “detection rate”, i.e., the proportion of cloud data detected by the AMSU-A and MHS cloud indices, \(Pclr\) is the “rejection rate”, i.e., the proportion of clear-sky data rejected by both cloud indices, \(Nd\) represents the number of collocated FOVs determined as cloudy by both the method and VISSR-II, \(Nclrm\) represents the number of collocated FOVs classified as cloudy by the method but clear by VISSR-II, \(Ncld\) and \(Nclr\) represent the number of collocated FOVs that are, respectively, identified as cloudy and clear by VISSR-II.

From Figure 5a, it can be seen that, as the two thresholds increased, the detection rate and rejection rate decreased. The new thresholds need to meet the high detection rate and low rejection rate. When the old threshold of 0.35 was used for the MHS cloud index, the rejection rate exceeded 60%, while the detection rate was high (detection rate greater than 80%), resulting in too much clear-sky data being rejected. When the old threshold of 0.1 was used for the AMSU-A cloud index, the detection rate basically exceeded 80%, but the rejection rate even exceeded 70%. In order to reduce the rejection rate, the MHS cloud index threshold needs to be lowered and the AMSU-A cloud index threshold needs to be raised. When the MHS cloud index threshold was reduced to 0.3 and the AMSU-A cloud index threshold was increased to 1.0, the detection rate of cloudy data was 79.20%, which is close to 80%; the rejection rate of clear-sky data was merely 46.57%. These thresholds can detect...
most of the cloudy data and reject only a small amount of the clear-sky data. If the MHS cloud index threshold continued to be lowered, the rejection rate of clear-sky data became higher with the same AMSU-A cloud index threshold. Similarly, when the MHS cloud index threshold remained unchanged, continuing to increase the AMSU-A cloud index threshold resulted in a lower detection rate of cloudy data. In summary, 0.3 and 1.0 were chosen here as the new thresholds of the cloud indices, i.e., when the MHS cloud index > 0.3 or the AMSU-A cloud index > 1.0 in the FOV, the FOV was considered as a cloudy FOV.

**Figure 4.** Clear-sky FOVs (white circles) and cloudy FOVs detected by the AMSU-A cloud index only (black circles), by the MHS cloud index only (blue circles), and by both cloud indices (black dots) along with the matched VISSR-II cloud classification products at the same moment—1200 UTC 12 August 2019 (AMSU-A cloud index > 0.10 or MHS cloud index > 0.35 is the threshold to identify cloudy FOVs).

**Figure 5.** Detection rate of cloudy data (contours) and rejection rate of clear-sky data (colors) with (a) and without (b) channel 4 under different cloud index thresholds from 1 to 31 August 2019. The horizontal axis is the threshold of the AMSU-A cloud index, and the vertical axis is the threshold of the MHS cloud index.

Channel 4, an oxygen absorption channel, was also included here, and to demonstrate the importance of channel 4 to the AMSU-A cloud index, we also calculated the index without channel 4. Figure 5b shows the detection rate of cloudy data and the rejection rate of clear sky data without channel 4. Comparing Figure 5a,b, it can be found that when the MHS index was greater than 0.4, that is, the MHS index has little detection effect,
if channel 4 were included, the detection effect basically did not change with the MHS index. However, if we did not include channel 4, the detection effect rapidly decreased with the increase of the MHS index threshold, which means that removing channel 4 will significantly reduce the independent detection ability of the AMSU-A index. Therefore, it was necessary to include channel 4 in the calculation of the AMSU-A cloud index.

4. Results

4.1. Effectiveness of the Cloud Detection Method

From Figure 6a,b, it can be seen that the AMSU-A cloud index can detect most of the cloudy areas—for instance, in Figure 6a, the banded deep convective cloud system over the southern part of the central Tibetan Plateau and the stratiform cloud system over the west; and in Figure 6b, the deep convective cloud system over the central Tibetan Plateau and north of the plateau. However, the AMSU-A cloud index also had some missed detections—for example, the cumulus clouds at the edge of the cloud system are clearly missed. However, the addition of the MHS cloud index made up for the missing detection of the AMSU-A cloud index. As shown in Figure 6a, the cumulus clouds scattered to the north of the Tianshan Mountains and the southern Tarim Basin, and the cumulonimbus clouds in the southern Himalayas; and in Figure 6b, the stratocumulus and high cumulus clouds distributed near the Qilian Mountains, north of the Tianshan Mountains, and in the southern Himalayas, were all detected by the MHS cloud index. In both cases, the detection rate of cloudy data surpassed 80%, the rejection rate of clear-sky data was around 40%, and the cloud index detection was reliable. It is, therefore, demonstrated that, with the new threshold, the AMSU-A cloud index and MHS cloud index can detect major cloudy data, and the structure and edges of the cloud system can be detected more accurately.

![Figure 6.](image)

Table 4 shows the detection rates of different cloud types over the study area in August 2019, and it can be seen from the figure that the detection rate of mixed clouds was the largest with over 95%. These were followed by cumulonimbus, stratocumulus/altocumulus, with detection rates of 92.3% and 85.8%, respectively. The cloud detection method worked worst for cirrostratus, followed by cirrus, nimbostratus/altostratus. The cirrostratus and cirrus clouds are mainly composed of ice crystals, the nimbostratus and altostratus are mainly ice-water mixed clouds. Microwave radiation can easily bypass these small ice particles, which may be the reason for the low effect of cloud detection.
Table 4. The detection rates of different cloud types over the study area in August 2019.

<table>
<thead>
<tr>
<th>Cloud Abbreviation</th>
<th>Mixed</th>
<th>Ns and As</th>
<th>Cs</th>
<th>Ci</th>
<th>Cb</th>
<th>Sc and Ac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pd (%)</td>
<td>95.3</td>
<td>76.4</td>
<td>72.5</td>
<td>75.6</td>
<td>92.3</td>
<td>85.8</td>
</tr>
</tbody>
</table>

Figure 7 shows the cloud detection results of the new and old thresholds of the two cloud indices above different topographic heights in East Asia during August 2019. It can be seen that the rejection rate of the clear-sky data for the old threshold remained around 70% before the topographic height increased to 700 m. When the altitude was higher than 700 m, the rejection rate of the old threshold began to increase significantly, this proves that the high detection rate of the old threshold was achieved by eliminating a large number of useful data. On the contrary, the rejection rate of the new threshold remained stable at about 45%. Although the detection rate of cloudy data decreased with the increase of terrain height due to the fact that there mainly exist cloud types with lower detection rates, such as cirrus, in high-terrain areas, especially the region above 5 km. From the above statistical results, the new threshold may be more suitable for the observations in areas with topographic height above 700 m.

Figure 8 shows the detection of cloudy data at different cloud base altitudes and cloud top altitudes over the study area in August 2019. As shown in Figure 8, the cloud indices were generally effective in detecting low and high clouds but poor in detecting cloud types with cloud base altitude and cloud top altitude located at 7–8 km, and the detection rates were lower than 50%. The statistical results show that most of the clouds appearing at this height were altostratus, which is mainly composed of ice crystals. The radiation in the microwave band can easily bypass these small ice particles, resulting in low detection rate, but it also shows that these clouds may not significantly affect the observed BTs of those microwave channels, nor will they affect the assimilation application of these BTs.

4.2. Evaluation of Bias and Error Characteristics

Different from the pure cloud detection research, the cloud detection method proposed in this paper mainly aimed to provide the assimilation system with fast and accurate clear sky data in the process of data assimilation, so that the data assimilation system can correct the bias effectively and set the appropriate observation error for the observation data. In order to verify the effect of the cloud detection method, the bias and error characteristics of clear-sky data identified by the new cloud detection method were also analyzed.
The substantial increase in terrain height in the Tibetan Plateau region caused more of the original high-altitude channels to become near-ground channels; plus, the simulation error increased, and the characteristics of the increase in error also varied with the difference in peak height of the weighting function and the terrain height. This required a re-evaluation of the error and deviation characteristics of the observed and the simulated BTs of different channels in the Tibetan Plateau region, providing a basis for establishing a more effective topographic deviation revision scheme.

The BT is mainly affected by the surface emissivity, surface temperature, atmospheric temperature, water vapor profile, various trace gases, and clouds. To minimize the effects of surface emissivity and surface temperature, this study selected the mid-peaking channels (channels 5 and 6) of AMSU-A for simulation error analysis. In addition to the influence of clouds, terrain height and vegetation type are also major sources of AMSU-A simulation errors. Therefore, this study analyzed the simulation error characteristics of each mid-peaking channel of AMSU-A, for each vegetation type, at different terrain heights, based on the accurate identification of clear-sky data.

This study used the NPOESS land surface classification scheme to determine the land vegetation type of each FOV and selected the five vegetation types with the largest number of observed samples in the study area—namely, grass, sand, pine forest, broadleaf pine forest, and scrub—while other vegetation types were not considered because of the small number of observed samples. The highest elevation of the samples in the study area was around 5800 m, and the topographic heights of the samples were classified into 500–1000, 1000–2000, 2000–3000, 3000–4000, 4000–5000, and above 5000 m. Observations with a sample size of less than 100 in each classification were excluded.

From Figure 9a,b, it can be seen that the simulated BT deviations of both channels were positive at all heights except 500–1000 m. In channel 5, the simulated BT deviation was within 1.5 K for all vegetation types except pine forest distributed at 3000–4000 m. The deviation of the simulated BT in channel 6 was relatively larger, and the deviation of the simulated BT in pine forest distributed at 3000–4000 m was the largest, exceeding 4.0 K. The deviation of the simulated BT in the remaining vegetation types was within 3.0 K. The deviation of the simulated BT in the height range below 2000 m was within 1.0 K. The deviation exhibited an increasing trend with increasing terrain height, especially when the terrain height was higher than 2000 m. In the GSI (Gridpoint Statistical Interpolation) assimilation system, the terrain height of 2000 m was also taken as the dividing line, and different deviations and error coefficients were given to the observations above this height, which is consistent with the findings of this paper [46].
Figure 9. Histograms of (a,b) the average O–B and (c,d) standard deviation of clear-sky observations for AMSU-A (a,c) channels 5 and (b,d) channel 6 for five vegetation types at different topographic heights over the Tibetan Plateau region on 1–31 August 2019.

The standard deviation of the two channels’ O–B varied more obviously with terrain height. The standard deviation of channel 5 increased remarkably with terrain height for all vegetation types, but especially grass, and the standard deviation increased with terrain height from 0.8 to 2.0 K. In contrast, the standard deviation of channel 6 was relatively smaller, basically maintaining at around 1.0 K. Moreover, the standard deviation of channel 6 had a clear difference with terrain changes. The standard deviation did not show regular changes for grass and scrub, but it appeared to increase with topography for both pine forest and broadleaf pine forest. This also indicates that, despite the different observation weights set in the GSI assimilation system with the height of 2000 m as the dividing line, the actual assimilation may still need to be further adjusted according to the vegetation type and topography to optimize the assimilation of AMSU-A data in plateau areas.

5. Discussion and Conclusions

Short-term convective systems are commonplace over the Tibetan Plateau, and the spatial and temporal variability of surface emissivity is also influenced by factors such as terrain height and surface type, meaning cloud detection using AMSU-A and MHS data has been a challenging task in this region. In this paper, the cloud detection method established by Wu et al. [39], which relies only on the observations by merging the AMSU-A data and MHS data, was applied to the Tibetan Plateau region, but the original cloud index thresholds led to severe over-detection due to the low surface temperature in the Tibetan Plateau region. This study collected NOAA19 AMSU-A and MHS observations continuously for one month and evaluated the detection rate of cloudy data and the rejection rate of clear-sky data under different cloud index thresholds with reference to the VISSR-II cloud classification product, to verify the new cloud index thresholds relatively objectively. With the new thresholds, the AMSU-A and MHS cloud indices could distinguish more than 80% of cloudy data on average, but the detection rate decreased with the increase of terrain height, and the detection rate was about 72% in the area of terrain higher than 5 km, but
the false deletion rate basically did not change with the terrain height and remained stable at about 45%.

Cases revealed that the AMSU-A cloud index identified most convective cloud systems but missed the cumulus clouds obviously, the addition of the MHS cloud index compensated for the missed detection by the AMSU-A cloud index. Result of the batch test showed that the detection rates of mixed clouds and cumulonimbus were higher than 90%, but the new thresholds had a low detection rate for cirrostratus owing to the fact that these clouds are mainly composed of ice particles, which have little effect on microwaves. The new thresholds performed much better than the old thresholds in areas where the terrain height was higher than 700 m, so it is recommended to apply the new thresholds in areas higher than 700 m.

With the accurate identification of clear-sky data, this study evaluated the error and bias characteristics of the AMSU-A mid-peaking channels in the Tibetan Plateau region for different vegetation types and terrain heights. The results demonstrated that the deviation of simulated BT of channel 5 was basically within 1.5 K, and the standard deviation within 2 K. For the simulated BT of channel 6, the deviation is basically within 3 K, and the standard deviation within 1 K. The deviations and standard deviations of the simulated BT of the AMSU-A mid-peaking channels basically exhibited an increasing pattern with increasing terrain height. Among the five vegetation types analyzed, the standard deviation and deviation of the mid-peaking channels were found to be the smallest for sand and the largest for pine forest at 3000–4000 m. For the standard deviation of the simulated BT, channel 5 was larger for grass and broadleaf pine forest above 3000 m, and channel 6 was largest for broadleaf pine forest at 3000–4000 m.

It is important to acknowledge that the present cloud detection method has a certain limitation insofar as the experimental data were concentrated in summer, so the scheme may generate bias in other seasons. In addition, this paper only studied the Tibetan Plateau region, with no comprehensive analysis having been conducted yet for other high-terrain regions. Therefore, more data from different seasons and regions are needed in subsequent studies, considering the combined effects of each index. In terms of the capability of the data, the new threshold of the cloud index was confirmed by referring to the VISSR-II cloud classification product in this paper, but the time difference between the AMSU-A, MHS, and VISSR-II instruments will also influence the test results. Furthermore, the validity of the new cloud detection method needs to be examined by assimilation experiments.

**Author Contributions:** Conceptualization, Z.Q.; data curation, J.W.; formal analysis, J.W.; funding acquisition, Z.Q. and J.L.; investigation, J.W.; methodology, J.W. and Z.W.; project administration, Z.Q. and J.L.; resources, J.W.; software, Z.Q. and J.W.; validation, J.W. and Z.Q.; visualization, J.W.; writing—original draft, J.W.; writing—review and editing, Z.Q., J.W. and J.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was jointly funded by the National Key R&D Program of China (Grant 2018YFC1507302), Fengyun Application Pioneering Project (FY-APP-2021.0201), and the NUIST Students’ Platform for Innovation and Entrepreneurship Training Program (Project 202110300090Y).

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to acknowledge the Numerical Weather Prediction Center of CMA for providing the satellite data and supercomputer computing resources.

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

**References**


41. Ferraro, R.R. Special sensor microwave imager derived global rainfall estimates for climatological applications. *J. Geophys. Res. Atmos.* 1997, 102, 16715–16735. [CrossRef]