remote sensing

MDPI

Article

A Machine-Learning-Based Study on All-Day Cloud Classification Using Himawari-8 Infrared Data

Yashuai Fu [1,2,†], Xiaofei Mi [3,†], Zhihua Han [1,2], Wenhao Zhang [1,2,*], Qiyue Liu [1,2], Xingfa Gu [1,3] and Tao Yu [1,3]

[1] School of Remote Sensing and Information Engineering, North China Institute of Aerospace Engineering, Langfang 065000, China; fuys@stumail.nciae.edu.cn (Y.F.); hanzhihua@stumail.nciae.edu.cn (Z.H.); liuqy_bhht@nciae.edu.cn (Q.L.); guxingfa@radi.ac.cn (X.G.); yutao@radi.ac.cn (T.Y.)
[2] Hebei Collaborative Innovation Center for Aerospace Remote Sensing Information Processing and Application, Langfang 065000, China
[3] Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; mixf@aircas.ac.cn
* Correspondence: zhangwh@radi.ac.cn
† These authors contributed equally to this work.

Abstract: Clouds are diverse and complex, making accurate cloud type identification vital in improving the accuracy of weather forecasting and the effectiveness of climate monitoring. However, current cloud classification research has largely focused on daytime data. The lack of visible light data at night presents challenges in characterizing nocturnal cloud attributes, leading to difficulties in achieving continuous all-day cloud classification results. This study proposed an all-day infrared cloud classification model (AInfraredCCM) based on XGBoost. Initially, the latitude/longitude, 10 infrared channels, and 5 brightness temperature differences of the Himawari-8 satellite were selected as input features. Then, 1,314,275 samples were collected from the Himawari-8 full-disk data and cloud classification was conducted using the CPR/CALIOP merged cloud type product as training data. The key cloud types included cirrus, deep convective, altostratus, altocumulus, nimbostratus, stratocumulus, stratus, and cumulus. The cloud classification model achieved an overall accuracy of 86.22%, along with precision, recall, and F1-score values of 0.88, 0.84, and 0.86, respectively. The practicality of this model was validated across all-day temporal, daytime/nighttime, and seasonal scenarios. The results showed that the AInfraredCCM consistently performed well across various time periods and seasons, confirming its temporal applicability. In conclusion, this study presents an all-day cloud classification approach to obtain comprehensive cloud information for continuous weather monitoring, ultimately enhancing weather prediction accuracy and climate monitoring.

Keywords: all-day cloud classification; XGBoost; CPR/CALIOP; Himawari-8; AInfraredCCM

check for updates

Citation: Fu, Y.; Mi, X.; Han, Z.; Zhang, W.; Liu, Q.; Gu, X.; Yu, T. A Machine-Learning-Based Study on All-Day Cloud Classification Using Himawari-8 Infrared Data. *Remote Sens.* **2023**, *15*, 5630. https://doi.org/10.3390/rs15245630

Academic Editors: Manuel Antón and Ismail Gultepe

Received: 14 October 2023
Revised: 17 November 2023
Accepted: 1 December 2023
Published: 5 December 2023

1. Introduction

Cloudiness, a consequence of the presence of minuscule water vapor particles or ice crystals in the atmosphere, is closely intertwined with numerous climatic phenomena [1]. According to the International Satellite Cloud Climatology Project [2,3], the global annual average cloud cover covers two-thirds of the Earth's surface. As the primary regulators of the Earth's radiation balance, the water cycle, and biogeochemical cycles [4,5], distinct cloud types produce varying radiative effects. Therefore, the precise categorization of cloud types and understanding of their distribution patterns hold significant practical importance [6,7]. The pursuit of effective and precise cloud classifications remains a prominent focus in meteorological research [8].

In past decades, cloud classification has predominantly been carried out using traditional methods and machine-learning-based techniques. Traditional methods include thresholds and statistical mathematical methods. However, these methods rely heavily on empirical knowledge when dealing with data, resulting in low classification efficiency. With continuous developments in computer science, several machine learning methods have emerged. Machine learning methods do not require extensive prior knowledge and offer high computational efficiency and excellent classification performance [9]. Therefore, conventional machine learning methods such as Random Forest (RF) and Support Vector Machine (SVM) approaches have been widely used in cloud classification research [10]. Wohlfarth et al. [11] used data from three visible channels and one infrared (IR) channel obtained from the Landsat-8 satellite. They classified clouds into nine different cloud classes and four subsurface classes using the SVM algorithm, achieving an impressive classification accuracy of up to 95.4%. Yu et al. [12] proposed a cloud classification method based on an RF algorithm for FY-4A. This method was combined with CloudSat's 2B-CLDCLASS cloud product to classify 8 single-layer cloud types and 12 multilayer cloud types. As the volume of data continues to increase, deep learning models such as deep neural networks (DNN) and artificial neural networks (ANN) are increasingly being used in cloud classification research. Cai et al. [13] employed FY-2C's infrared channel 1 (10.3–11.5 μm) data in combined with a Convolutional Neural Network (CNN) model for cloud classification. Their classification divided clouds into 5 types, achieving an impressive average recognition rate of 90.6%. In addition, Gorooh et al. [14] proposed the Deep Neural Network Cloud-Type Classification (DeepCTC) model that can classify clouds into 8 different types with an overall classification accuracy of 85%. Machine learning techniques have surpassed traditional cloud classification methods in terms of both classification speed and accuracy. However, it is important to note that these methods are generally applicable to daytime cloud classification. Nighttime visible light data are often noisy and cannot be used for nighttime cloud classification. Therefore, one of the research questions addressed in this study is how to achieve nighttime cloud classification.

Researchers have commenced extensive studies using infrared data to more accurately monitor cloud evolution at night, eliminating the need for visible data and achieving successful research outcomes [15]. Tan et al. [16] proposed a nighttime cloud classification algorithm based on Himawari-8 satellite channel data and machine learning algorithms. They utilized data from 5 infrared channels, 3 brightness temperature difference (BTD) datasets, and latitude/longitude information as training features. By employing the RF algorithm, this approach achieved an overall accuracy of 0.79, classifying clouds into clear-sky, single-layer, and multilayer clouds. Li et al. [17] used infrared data from the Himawari-8 satellite to classify clouds. Their classification scheme included five types: clear, single-ice clouds, single-mixed clouds, single-water clouds, and multilayer clouds. They achieved an impressive overall classification accuracy of 0.81. Despite significant progress in nighttime cloud classification research, a critical analysis of existing research results revealed a limited number of nighttime cloud classifications. Typically, cloud classification is constrained to primary cloud categories without considering the finer distinctions between clouds.

This study addresses the challenges of nighttime cloud classification and the limited number of classifiers using the latitude/longitude, five brightness temperature differences (BTD), and ten IR channels from the Himawari-8 satellite. The combined Cloud-Profiling Radar (CPR) and Cloud-Aerosol LIDAR with Orthogonal Polarization (CALIOP) cloud-type product 2B-CLDCLASS-LIDAR were used as labels, which classified clouds into nine types: clear-sky (Clear), cirrus (Ci), deep convective (Dc), altostratus (As), altocumulus (Ac), nimbostratus (Ns), stratocumulus (Sc), stratus (St), and cumulus (Cu). By comparing various models using various metrics, such as overall accuracy, precision, recall, and F1-scores, this study investigated the viability of all-day cloud classification and the selection of optimal parameters. The goal was to improve both the quantity and accuracy of cloud classification, thus providing valuable additional reference data for nighttime meteorological observations.

Section 2 introduces the experimental data and methodology used in this study. Section 3 presents the experimental results and their implications. Finally, Sections 4 and 5 present the discussion and conclusion, respectively.

2. Materials and Methods

2.1. Data Collection

This study used Level-1 infrared channel data from the Himawari-8 satellite and the cloud type product 2B-CLDCLASS-LIDAR, which is a joint project of CloudSat and CALIPSO. Additionally, longitude, latitude, and solar zenith angle data were employed for support.

(1) Himawari-8 data

The Japan Meteorological Agency's Geostationary Meteorological Satellite, Himawari-8, was launched on 7 July 2015. The data from this satellite can be accessed through the Japan Aerospace Exploration Agency (JAXA) (http://www.eorc.jaxa.jp/ptree/index.html, accessed on 15 May 2023) [18]. Compared with MTSAT-2 (Himawari-7), Himawari-8 is equipped with an Advanced Himawari Imager (AHI) sensor that has expanded from the original 5 bands to 16 bands, including 3 visible light bands, 3 near-infrared bands, and 10 infrared bands [19,20]. It covers the area from 60°S to 60°N and from 80°E to 200°E (Figure 1). The observation frequency of Himawari-8 has been increased to once every 10 min, providing abundant data for meteorological studies and continuous cloud observations [21]. This study used 10 infrared bands from the Himawari-8 satellite. Table 1 displays the parameters of the 10 infrared bands of the Level-1 products of Himawari-8 and their main applications.

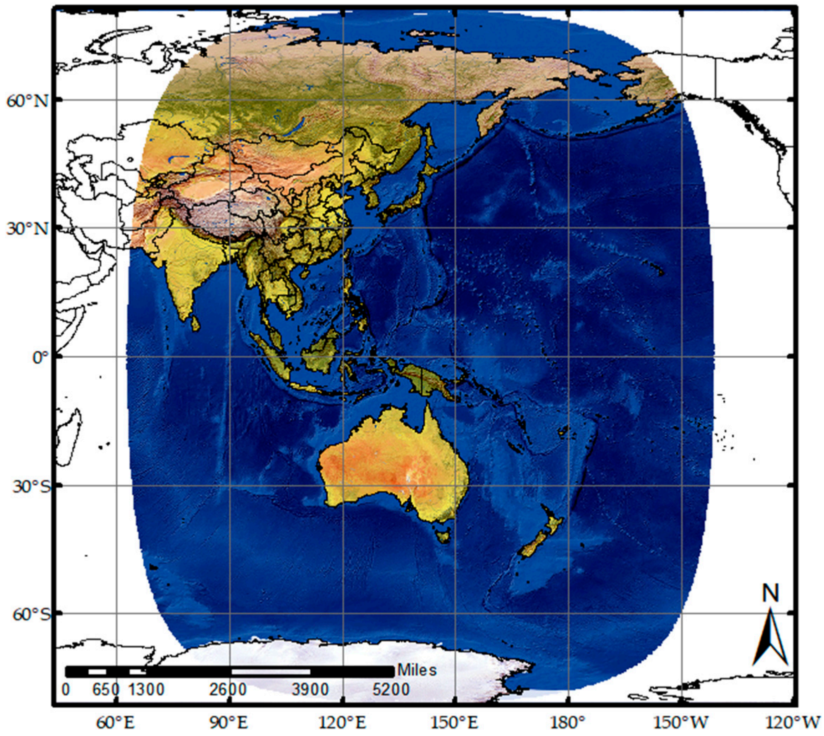


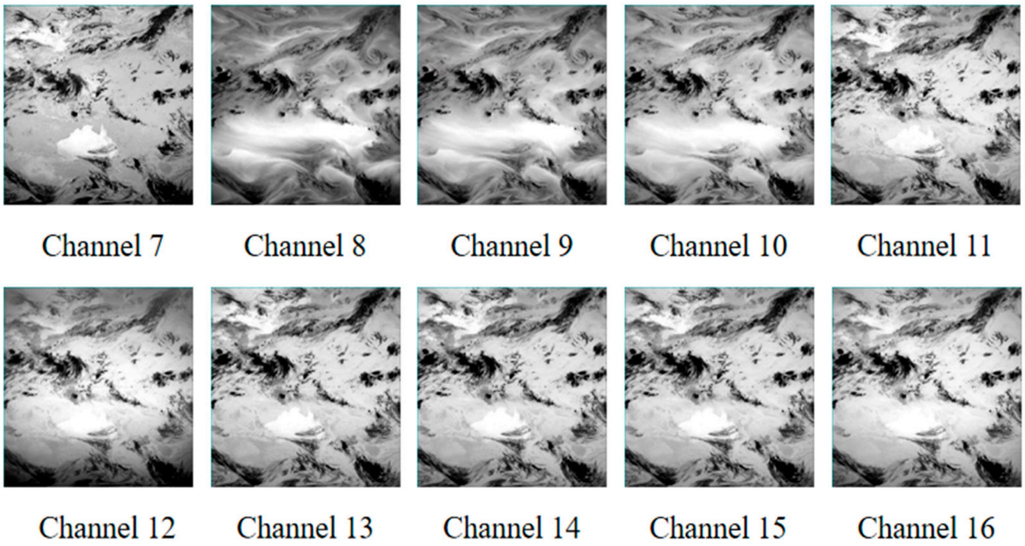
Figure 1. Illustration of the observation area of the Himawari-8 satellite.

The Level-2 (L2) cloud product of the Himawari-8 satellite provides a comprehensive set of cloud-related parameters, such as cloud type (CLTYPE), cloud top height, cloud top temperature, and cloud optical thickness. However, it is important to note that this product solely provides cloud information for observations made during the daytime period. The spatial resolution is 5 km and the temporal resolution is 10 min. A more detailed description of AHI Level-1/2 data is available at https://www.eorc.jaxa.jp/ptree/userguide.html, accessed on 20 June 2023.

Table 1. Himawari-8 band parameters and applications.

Bands	Channel Type	Center Wavelength (μm)	Spatial Resolution (km)	Main Applications
7	Midwave IR	3.9	2	Natural disasters, low cloud (fog) observation
8	Water vapor	6.2	2	Observation of water vapor volume in the upper and middle layers
9		6.9	2	Observations of water vaporization in the mesosphere
10		7.3	2	
11	Longwave IR	8.6	2	Cloud phase identification and SO2 detection
12		9.6	2	Measurement of total ozone
13		10.4	2	Observation of cloud images and cloud top conditions
14		11.2	2	Observation of cloud images and sea surface water temperature
15		12.4	2	Observation of cloud images and sea surface water temperature
16		13.3	2	Measurement of cloud height

This study used 10 infrared bands as basic data, ranging from band 7 to band 16. In addition, auxiliary data, including longitude, latitude, sun zenith angle, and CLTYPE, were incorporated. Figure 2 displays the cloud images of Himawari-8 satellite for its 10 infrared channels at UTC 03:20 on 1 July 2019.


Channel 7 Channel 8 Channel 9 Channel 10 Channel 11

Channel 12 Channel 13 Channel 14 Channel 15 Channel 16

Figure 2. Cloud data for 10 channels of AHI.

(2) CloudSat data

CloudSat and CALIPSO are part of the A-train that crosses the equator in the afternoon [22]. Both satellites provide near global views of clouds from sun-synchronous orbits and are available at https://www.cloudsat.cira.colostate.edu, accessed on 23 May 2023. CloudSat's onboard CPR and CALIOP (a vision and near-infrared LIDAR) are 2 powerful onboard active instruments [23]. Currently, these are the only instruments capable of accurately detecting the vertical structure of clouds [24]. Owing to the different working wavelengths of CPR and CALIOP, they exhibit different sensitivities to different types of clouds. The millimeter-wave radar CPR has the advantage of detecting optically thick clouds and precipitation systems [25]. Combining their advantages, a joint product of CPR/CALIOP was developed to provide the most reliable vertical cloud information; this was named 2B-CLDCLASS-LIDAR. This product provides information on up to 10 cloud layers. These layers are characterized by a horizontal resolution of 1.4 × 1.8 km and a vertical resolution of 0.24 km. This dataset contains valuable details, including cloud layers, cloud heights, and cloud phases [26]. The CloudLayerType band elements range from 0 to 8 and represent different cloud types: Clear, Ci, As, Ac, St, Sc, Cu, Ns, and Dc. In this study, CloudLayerType band data were utilized as the labeling criteria. For a more comprehensive understanding of 2B-CLDCLASS-LIDAR, please refer to the CloudSat product brochure [27].

(3) Cloud type of this study

In this study, 130 days of data (Table A1) were selected from November 2018, January 2019, March 2019, June 2019, and July 2019. The Himawari-8 CLTYPE products encompass 10 types (0–9 represent Clear, Ci, Cs, Dc, Ac, As, Ns, Cu, Sc, and St, respectively) and CPR/CALIOP products consist of only 9 types. To ensure consistency, the clouds were finally classified into 9 types based on the one-to-one mapping between the CPR/CALIOP and Himawar-8 clouds. Notably, in Himawari-8, Ci and Cs were merged into a one type called Ci clouds. The classification criteria are listed in Table 2.

Table 2. Cloud type of this study.

Cloud Label	Label of CPR/CALIOP	Label of CLTYPE	Name of Cloud
0	0 (Clear)	0 (Clear)	Clear
1	1 (Ci)	1, 2 (Ci, Cs)	Ci (Ci, Cs)
2	8 (Dc)	3 (Dc)	Dc
3	3 (Ac)	4 (Ac)	Ac
4	2 (As)	5 (As)	As
5	7 (Ns)	6 (Ns)	Ns
6	6 (Cu)	7 (Cu)	Cu
7	5 (Sc)	8 (Sc)	Sc
8	4 (St)	9 (St)	St

2.2. Method

In this study, 10 different IR bands from Himawari-8 and 5 BTDs were used to select potential donor pixels. Longitude and latitude were used as additional constraints. In the following section, the proposed algorithm is referred to as the All-day Infrared Cloud Classification Model (AInfraredCCM). Figure 3 shows a conceptual diagram of the method used in this study and its comparison.

(1) Data collection and processing: This study utilized various sources, such as CloudLayerType, Himawari-8 Level-1 IR, latitude and longitude, and solar zenith angle data from the 2B-CLDCLASS-LIDAR product [26]. To obtain BTD information, pairwise differences were calculated between all the infrared channels. Subsequently, feature selection was performed using the feature analysis function within the machine learning module [28]. The final input feature set included 5 BTDs, 10 infrared channels, and latitude and longitude data. BTD (11.2–7.3 µm) can be employed to detect high- and mid-level clouds over land during the night. BTD (3.9–11.2) is known as useful BTD to detect low-level clouds. BTD (11.2–12.4 µm) is useful for distinguishing thin clouds from clear sky. BTD (12.4–10.4 µm) can serve as a substitute for visible light to describe cloud optical thickness. BTD (7.3–10.4 µm) is utilized for nighttime cloud detection by differencing the water vapor channel and the infrared channel [10,29]. After obtaining the BTD data, a series of data preprocessing steps were applied to both the Himawari-8 and CPR/CALIOP datasets. These steps include spatiotemporal matching, label extraction, data resampling, and cropping. Consequently, a training dataset consisting of 1,182,688 samples and a testing dataset consisting of 131,587 samples were created following a random 9:1 split ratio. The specific features contained in the dataset are listed in Table 3.

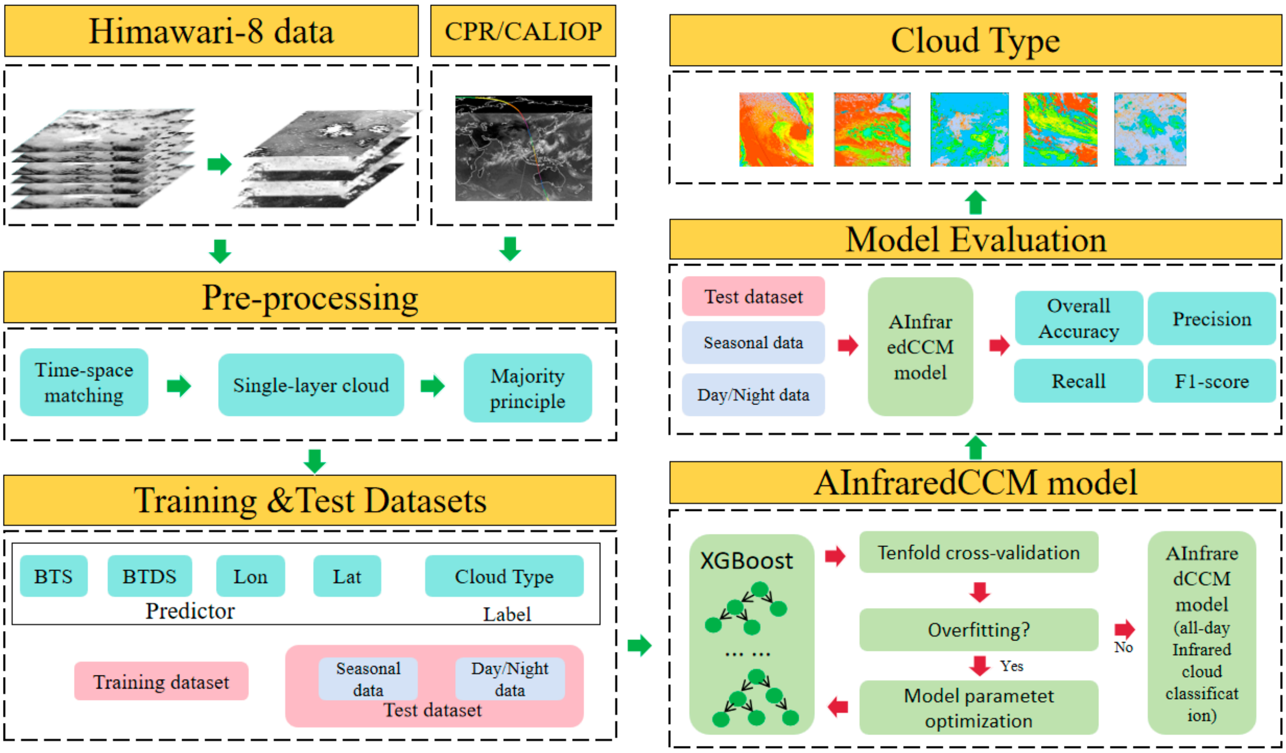


Figure 3. General flow diagram.

Table 3. Information of dataset.

Dimension	Number	Variables
Predictor	BTs (10)	BT (3.9 µm), BT (6.2 µm), BT (6.9 µm), BT (7.3 µm), BT (8.6 µm), BT (9.6 µm), BT (10.4 µm), BT (11.2 µm), BT (12.4 µm), and BT (13.3 µm)
	BTDs (5)	BTD (11.2–7.3 µm), BTD (3.9–11.2 µm), BTD (11.2–12.4 µm), BTD (12.4–10.4 µm), and BTD (7.3–10.4 µm)
	Auxiliary data (2)	Latitude and Longitude
Prediction	1	Cloud label from 2B-CLDCLASS-LIDAR and CLTYPE

(2) Label extraction: To obtain cloud type samples for training, Himawari-8 and 2B-CLDCLASS-LIDAR products were subjected to temporal and spatial matching. Each pixel in the CloudLayerType dataset contained information on 10 cloud types, and each Himawari-8 pixel corresponded to a cloud type. During data matching, the first step was to compress the 10-layer data. If the CPR point corresponded to multilayer clouds with different cloud types in each layer, it was defined as a multilayer cloud; otherwise, it corresponded to the cloud type in the first layer. The second step involved a temporal and spatial matching process due to the difference in the spatial and temporal resolutions between AHI and CPR. This entailed selecting AHI data points within a 5 min time window of the CPR scan point and within a 5 km radius of the CPR point location [30,31]. In the third step, considering that the AHI observations did not include multilayer cloud types, the CPR points corresponding to multilayer clouds were eliminated. Finally, cloud type labels were assigned by majority rule based on the matching of AHI and CPR points [32,33].

(3) Model building: The processed dataset was used to train the XGBoost model, and the optimal model parameters were determined by Bayesian optimization. This step selection was made by sampling using the Bayesian optimization algorithm after estimating the distribution of the objective function in the parameter space through a Gaussian process model. It evaluated the overall classification accuracy of related models while searching for the best parameter combination across a finite number of iterations. Finally, the parameter

combination with the highest overall classification accuracy was obtained and represented as the optimal model parameters.

(4) Model evaluation: The overall accuracy (OA), precision, recall, and F1-score were calculated based on ten-fold cross-validation [34]. If the final model achieved sufficient accuracy, it was selected; otherwise, the parameters were modified.

This study used the parameters shown in Table 4 as an example to better explain the metrics used for accuracy assessment. Assume that a dataset C consists of T samples; here, AS is the number of samples for class A clouds and BS is the number of samples for class B clouds. Table 4 shows the specific classification results from dataset C, which were used for model classification prediction.

Table 4. Classification results of the model on dataset C.

		Number of Ever Category		**Total Number**
		Number of category A clouds	Number of category B clouds	
Model classification result	Number of category A clouds	TA	FB	T1
	Number of category B clouds	FA	TB	T2
Total number		AS	BS	T

The overall accuracy (OA) is a crucial measure, representing the proportion of correctly classified samples to the total number of samples [12]. A higher overall accuracy indicates a more reliable classification outcome. The overall accuracy is calculated as follows:

$$OA = \frac{TA + TB}{TA + FB + FA + TB} \tag{1}$$

where TB represents an accurately classified cloud type B.

Precision is a metric that quantifies the proportion of correctly identified clouds in the target class, indicating fewer classification errors. Precision can be expressed as follows:

$$Precision = \frac{TA}{TA + FA} \tag{2}$$

where TA denotes correctly identified class A clouds and FA denotes the misdiagnosis of class A clouds as class B clouds.

The recall pertains to the proportion of correctly identified class A clouds within the total number of identified class A clouds [35]. Recall can be represented as follows:

$$Recall = \frac{TA}{TA + FB} \tag{3}$$

where TA denotes the correct identification of class A clouds and FB denotes the misdiagnosis of class B clouds as class A clouds.

The $F1$-score combines precision and recall, with values ranging from 0 to 1. A higher score indicated that the model was more accurate. Compared with that using only precision or recall metrics, this score provides a more comprehensive evaluation of model accuracy [36]. The F1-score can be written as follows:

$$F1 = \frac{2PrecisionRecall}{Precision + Recall} \tag{4}$$

(5) The AInfraredCCM was conducted through all-day cloud classification for the entire research area using the test data.

XGBoost is an ensemble learning method that employs boosting principles. Unlike traditional boosting methods, XGBoost uses tree models (decision trees) as weak learners. It combines multiple weak learners to create a strong learner, gradually improving the model's performance. XGBoost is adaptive; in each iteration, it introduces a new weak classifier that attempts to correct the errors of the previous round until the stopping condition is satisfied. The key characteristic of XGBoost is the integration of multiple weak learners into a strong learner, gradually improving the model accuracy through iterative error correction [37].

Bayesian optimization approach was employed to tune the model and identify the optimal parameters of the AInfraredCCM model. The n_estimators parameter was set within a range from 100 to 300, the learning_rate parameter was set within a range from 0.01 to 0.999, the max_depth parameter was set within a range from 10 to 100, and the min_child_weight parameter ranges from 1 to 10 during optimization. After 100 iterations, the AInfraredCCM was developed, and the model parameters are displayed in Table 5.

Table 5. Parameters of the AInfraredCCM.

Parameter	Meaning	Value
n_estimators	Number of trees	204
learning_rate	Magnitude of the iterative model update	0.2122
max_depth	Maximum tree depth	26
min_child_weight	Minimum number of samples required in a leaf node	3

3. Results

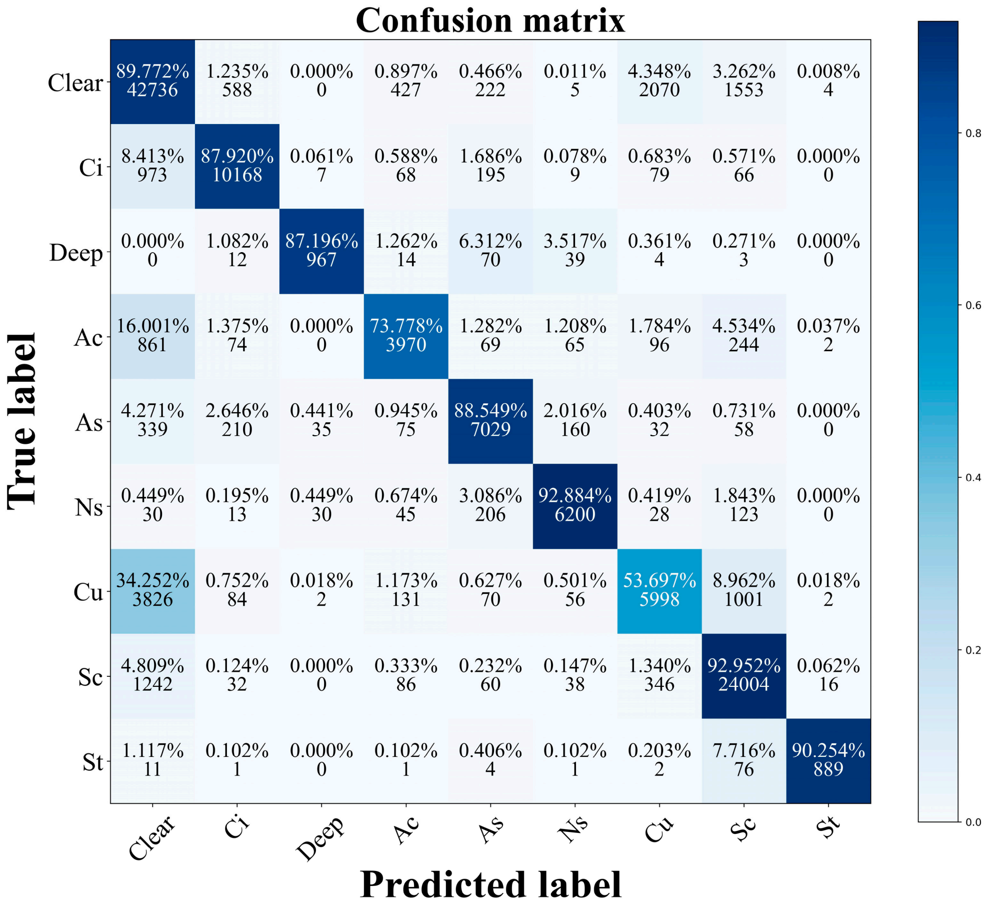
3.1. Comparison with Other Methods

The results are based on an independent test dataset of approximately 130,000 pixels, which is divided into clear sky and eight different cloud forms. Table 6 shows the precision, recall, and F1-score of the AInfraredCCM for each form of cloud in the training dataset. The confusion matrix for various types of clouds is shown in Figure 4, wherein the vertical coordinate represents the true value and the horizontal coordinate represents the predicted value.

Table 6. Precision, recall, and F1-score of the AInfraredCCM.

Cloud Type	Precision	Recall	F1-Score
Clear	0.85	0.89	0.87
Ci	0.90	0.88	0.89
Dc	0.93	0.87	0.90
Ac	0.82	0.74	0.78
As	0.89	0.88	0.89
Ns	0.95	0.93	0.94
Cu	0.68	0.57	0.60
Sc	0.88	0.93	0.91
St	0.98	0.90	0.94

Based on the model's classification results for the all-day data, the average overall accuracy for all cloud types was 86.22%, with a precision of 0.88, a recall of 0.84, and an F1-score of 0.86. Notably, Cu clouds exhibited considerably lower identification rates than the other cloud types, and both recall and F1-score are significantly lower than those of other cloud types (Table 6). The confusion matrix shown in Figure 4 graphically displays that most of Cu is misclassified as clear sky (34.252%) and Sc (8.962%). These phenomena can be attributed to the unique properties of Cu, which include two possibilities: one resembles Cu in clear weather and appears as thin, fragmented clouds that closely resemble clear in terms of reflectance, making them easy to misidentify; the alternative scenario is continuous Cu, which is denser and more vertically developed than clear-sky Cu and is frequently categorized as Sc.



	Clear	Ci	Deep	Ac	As	Ns	Cu	Sc	St
Clear	89.772% 42736	1.235% 588	0.000% 0	0.897% 427	0.466% 222	0.011% 5	4.348% 2070	3.262% 1553	0.008% 4
Ci	8.413% 973	87.920% 10168	0.061% 7	0.588% 68	1.686% 195	0.078% 9	0.683% 79	0.571% 66	0.000% 0
Deep	0.000% 0	1.082% 12	87.196% 967	1.262% 14	6.312% 70	3.517% 39	0.361% 4	0.271% 3	0.000% 0
Ac	16.001% 861	1.375% 74	0.000% 0	73.778% 3970	1.282% 69	1.208% 65	1.784% 96	4.534% 244	0.037% 2
As	4.271% 339	2.646% 210	0.441% 35	0.945% 75	88.549% 7029	2.016% 160	0.403% 32	0.731% 58	0.000% 0
Ns	0.449% 30	0.195% 13	0.449% 30	0.674% 45	3.086% 206	92.884% 6200	0.419% 28	1.843% 123	0.000% 0
Cu	34.252% 3826	0.752% 84	0.018% 2	1.173% 131	0.627% 70	0.501% 56	53.697% 5998	8.962% 1001	0.018% 2
Sc	4.809% 1242	0.124% 32	0.000% 0	0.333% 86	0.232% 60	0.147% 38	1.340% 346	92.952% 24004	0.062% 16
St	1.117% 11	0.102% 1	0.000% 0	0.102% 1	0.406% 4	0.102% 1	0.203% 2	7.716% 76	90.254% 889

Figure 4. Confusion matrix of the AInfraredCCM.

To enhance the credibility of the model, the cross-validation results were compared with those obtained using various other models, including RF, LightGBM, AdaBoost, and GradientBoost. RF is a conventional machine learning model known for its versatility in handling both classification and regression tasks. LightGBM has been optimized for large datasets, offering high processing speed and low memory usage. This makes it particularly suitable for scenarios involving substantial amounts of data and high dimensional features [38]. Gradient Boosting is a boosting-based integrated learning technique known for its robustness to outliers and good performance in a variety of classification and regression problems. Similarly, AdaBoost is an integrated learning method that employs weighted voting to combine several weak classifiers. The weights are updated after each iteration based on the outcome of the previous round, thereby reducing overfitting and improving the management of sample imbalances. The best parameter combinations for each model were obtained after running several Bayesian optimizations. Table 7 provides more information about the optimal parameter combinations for the four models.

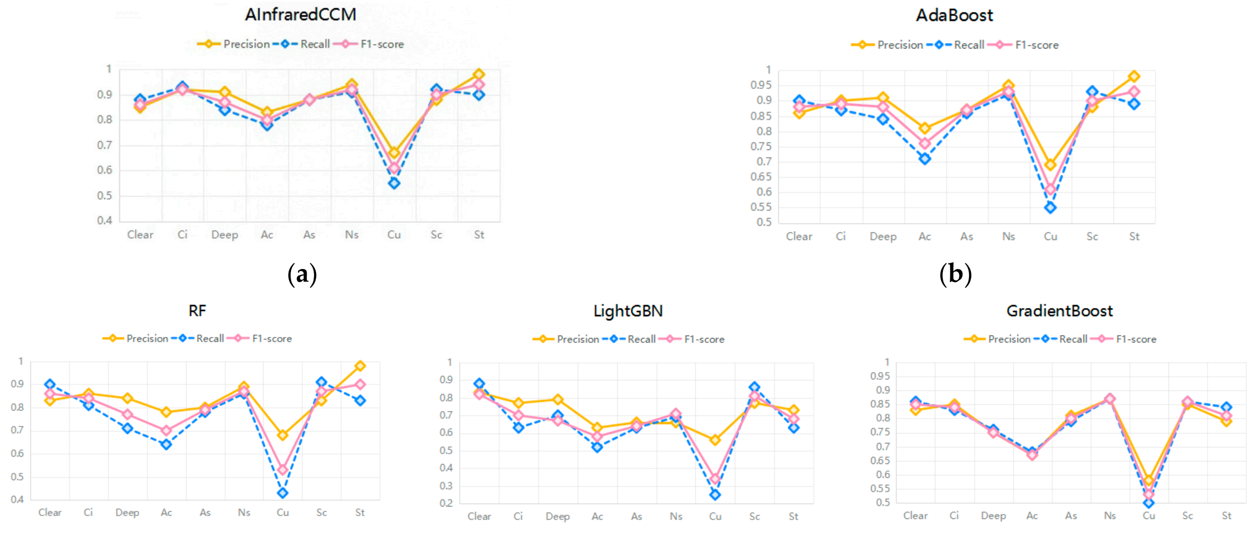
All models were trained using the selected parameters and tested for the classification of the training data. As demonstrated in Table 8 and Figure 5, the AInfraredCCM achieved the highest overall accuracy of 86.22% for the test datasets. Following was the AdaBoost classification model, which attained an accuracy of 85.83%. However, the AInfraredCCM outperformed AdaBoost across all three criteria, i.e., precision, recall, and the F1-score. The four indices of overall accuracy outside precision, recall, and F1-score were highest for the AInfraredCCM in the full analysis (OA—86.22%; precision—0.88; recall—0.84; F1-score—0.86). Table 8 lists the categorization metrics of the test datasets for the models.

Table 7. Optimal parameter combinations for the model.

Algorithm	Parameted Range
Random Forest	1. max_depth = 73 2. n_estimators = 280
LightGBM	1. learning_rate = 0.095 2. max_depth = 22 3. n_estimators = 252 4. num_leaves = 35
AdaBoost	1. learning_rate = 0.4224 2. max_depth = 74 3. n_estimators = 458 4. min_samples_leaf = 1
GradientBoost	1. learning_rate = 0.4749 2. max_depth = 37 3. n_estimators = 10

Table 8. Evaluation for different models.

Algorithm	Accuracy	Precision	Recall	F1-Score
Random Forest	82.53%	0.83	0.76	0.79
LightGBM	74.60%	0.70	0.64	0.66
GradientBoost	80.96%	0.78	0.77	0.78
AdaBoost	85.83%	0.87	0.83	0.85
AInfraredCCM	**86.22%**	**0.88**	**0.84**	**0.86**

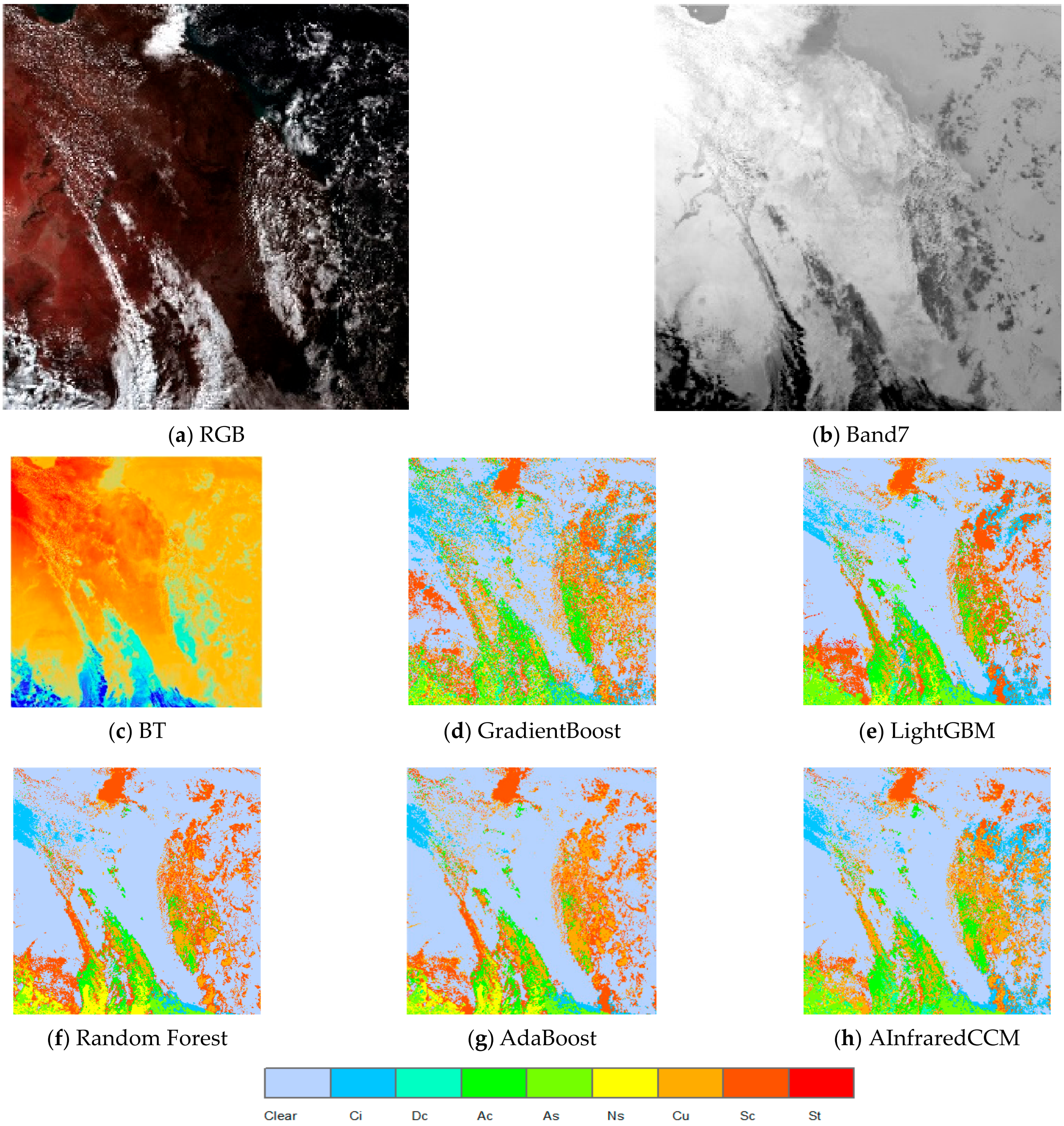


(**a**) (**b**)

(**c**) (**d**) (**e**)

Figure 5. Classification performance graphs of different models. (**a**–**e**) Graphs of the precision, recall, and F1-scores of classification results of 5 models in different cloud types.

Figure 6 illustrates the classification results of different models on 29 June 2019. A comparison between the images in the figure reveals that Gradient Boosting yields better classification results but with notable inconsistencies. The classification results of LightGBM, AdaBoost, and RF contained a relatively large number of errors and were ineffective for accurately classifying the Ci clouds. While all these models possess certain advantages, they failed to fully harness the dataset's features for precise cloud classification. In contrast, the classification results derived from the AInfraredCCM exhibited relatively smooth boundaries and accurate classification results.



(**a**) RGB (**b**) Band7

(**c**) BT (**d**) GradientBoost (**e**) LightGBM

(**f**) Random Forest (**g**) AdaBoost (**h**) AInfraredCCM

Figure 6. Classification results of different models at 05:00 UTC on June 29, 2019. (**a**) Visible data, (**b**) thermal infrared data, (**c**) the brightness temperature map; (**d**–**h**) sequentially present the classification results of GradientBoost, LightGBM, RF, AdaBoost, and AInfraredCCM, respectively.

Table 9 provides a statistical summary of the cloud classification models, featuring columns for the employed model, the number and specific types of cloud classifications, features, time, overall accuracy, sample size, and references. When comparing this study with those of other researchers, the AInfraredCCM demonstrated superior cloud classification performance (Table 9). While the CNN and the Backpropagation Neural Network (BP) achieved higher overall accuracy (0.95 and 0.86, respectively), it is noteworthy that these models were exclusively designed for daytime data and did not incorporate nighttime data. Furthermore, the CNN and BP models covered only eight and six cloud types (including clear sky), respectively, whereas nine cloud type classifications were retrieved in this study. Considering both daytime and nighttime data, this study achieved an overall accuracy of 0.86. In contrast, Yu et al. also categorized clouds into nine types, but their approach was exclusively applicable to the daytime, with a significantly lower overall accuracy compared to this study. Similarly, Tan and Li conducted research throughout the day; however, their approach did not exceed the overall accuracy or number of cloud types in this study. In summary, the AInfraredCCM has notable advantages in cloud classification.

Table 9. Cloud classification model statistical table.

Model	Category	Feature	Time	OA	Sample	Reference
RF	Dc, Ns, Cu, Sc, St, Ac, As, Ci, and multi	REF and BT of 13 channels, cloud top height, cloud optical thickness, cloud effective radius	Day	0.67	272414	Yu et al. (2021) [12]
BP	Clear, low cloud, middle cloud, thick cirrus clouds, thin cirrus cloud, deep convective	IR1 (10.3–11.3), IR2 (11.5–12.5), WV (6.3–7.6), IR1-IR2, IR1-WV, IR2-WV	Day	0.86	2449	Zhang et al. (2012) [39]
CNN	Clear, Ci, Ac, As, Sc, Dc, Ns, Cu	All channel of FY-4A	Day	0.95	15780	Wang et al. (2023) [40]
RF	Clear, low cloud, middle cloud, thin cloud, thick cloud, multilayer cloud, cumulonimbus	R (0.64), R (1.6), BT (11.2 µm), BTD (11.2–3.9 µm), BTD (11.2–7.3 µm), BTD (11.2–8.6 µm), BTD (11.2–12.3 µm)	Day	0.88	127192	Wang et al. (2023) [41]
RF	Clear, low cloud, middle cloud, thin cloud, thick cloud, multilayer cloud, cumulonimbus	BT (11.2 µm), BTD (11.2–3.9 µm), BTD (11.2–7.3 µm), BTD (11.2–8.6 µm), BTD (11.2–12.3 µm)	Night	0.79	72934	Wang et al. (2023) [41]
RF	Clear, single, multi	BT (3.9 um), BT (7.3 m), BT (8.6 µm), BT (11.2 µm), BT (12.4 µm), BTD (3.9–11.2 µm), BTD (8.6–11.2 µm), BTD (11.2–12.4 µm), latitude, longitude	Day and night	0.79	12553889	Tan et al. (2022) [16]
DNN	Clear, single-ice, single-mixed, single-water, multi	BT (3.9–13.3 µm), cosine of satellite zenith angle, simulated clear-sky radiances	Day and night	0.81	1114591	Li et al. (2022) [17]
AInfraredCCM	**Clear, Ci, Dc, Ac, As, Ns, Cu, Sc, St**	**BT (3.9–13.3 µm), BTD (11.2–9.6 µm), BTD (3.9–11.2 µm), BTD (11.2–12.4 µm), BTD (12.4–10.4 µm), BTD (7.3–10.4 µm), latitude, longitude**	**Day and night**	**0.86**	**1314275**	**This study**

3.2. Comparison with Himawari-8 Cloud Classification Production

In this study, a comparative analysis approach was employed to compare the cloud classification results of AInfraredCCM with CLTYPE using test data. To facilitate this comparative analysis in this study, the CLTYPE data were mapped to those of 2B-CLDCLASS-LIDAR. Notably, since Himawari-8 products officially cover cloud daytime evaluations, comparison and evaluation were limited to the daytime scenarios. The comparison was conducted in three scenarios, and the results are summarized in Table 10.

Table 10. Accuracy rate of Himawari-8 CLTYPE and AInfraredCCM.

	Himawari-8 CLTYPE	AInfraredCCM
Full area	0.48	**0.86**
Cloudy area	0.36	**0.87**
Clear sky	0.77	**0.85**

The three scenarios (all-sky, cloudy, and clear-sky) produced correct identification probabilities of 0.48, 0.36, and 0.77, respectively. The results of AInfraredCCM are presented in Tables 7 and 8. The model demonstrated superior performance, with overall accuracies of 86.22%, 87%, and 85% for all-sky, cloudy, and clear sky, respectively, outperforming the CLTYPE. The comparison data results are presented in Table 10; for further reference, Figures S1–S130 illustrate the disk classification results in contrast to those of the CLTYPE.

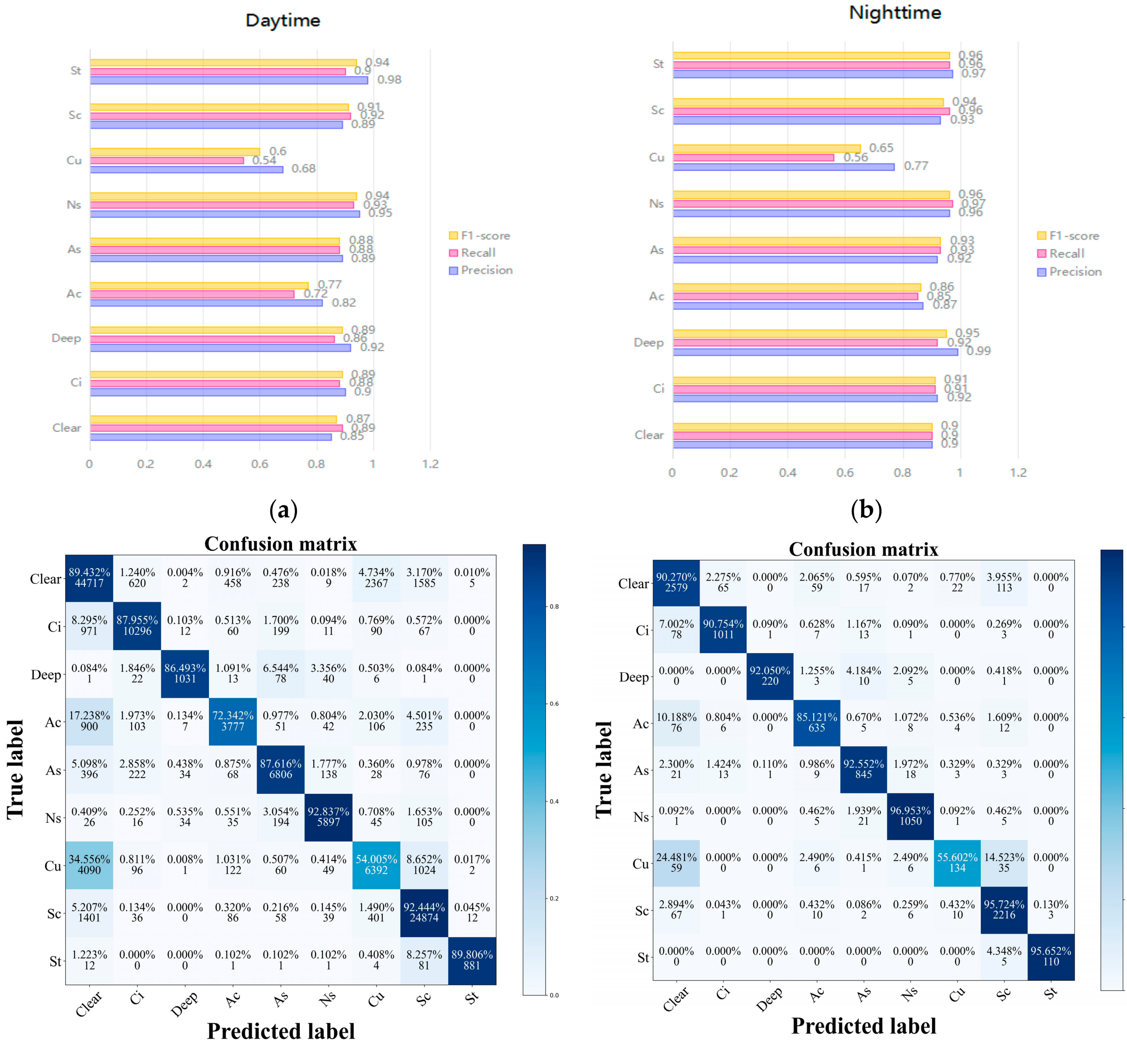
3.3. Effects of Day and Night on Cloud Classification

The results were meticulously examined individually under both daytime and nighttime conditions to assess the classification performance. An analysis of the efficacy of the AInfraredCCM during both the diurnal and nocturnal phases was carried out by assessing the overall precision, recall, and F1-score. Pixels in the test dataset with a solar zenith angle exceeding 80° were designated as nighttime, whereas the rest were designated as daytime [42]. Table 11 displays cloud classification results of the AInfraredCCM for daytime and nighttime.

Table 11. Cloud classification results of the AInfraredCCM for daytime and nighttime.

Time	Cloud Type	Clear	Ci	Dc	Ac	As	Ns	Cu	Sc	St
Accuracy = 85.82%										
Daytime	Precision	0.85	0.90	0.92	0.82	0.89	0.95	0.68	0.89	0.98
	Recall	0.89	0.88	0.86	0.72	0.88	0.93	0.54	0.92	0.90
	F1-score	0.87	0.89	0.89	0.77	0.88	0.94	0.60	0.91	0.94
Accuracy = 91.45%										
Nighttime	Precision	0.90	0.92	0.99	0.87	0.92	0.96	0.77	0.93	0.97
	Recall	0.90	0.91	0.92	0.85	0.93	0.97	0.56	0.96	0.96
	F1-score	0.90	0.91	0.95	0.86	0.93	0.96	0.65	0.94	0.96

Figure 7 presents the bar charts and provides a visual representation of the evalumetrics for the AInfraredCCM in daytime and nighttime. During the daytime, the AInfraredCCM achieved an average accuracy of 85.82% for all cloud types. All cloud classifications, except Cu, had accuracy levels above 0.8, accompanied by high recall and F1 scores, which indicated that the classification results were reliable. The AInfraredCCM performed better in nighttime than in daytime, attaining an impressive accuracy of 91.45%. In contrast, Cu had an overall accuracy of 0.77, whereas Dc, Ac, and As exhibited classification accuracies above 0.87. The confusion matrices presented in Figure 7 indicate that Cu was the cloud type with the weakest classification performance during both the daytime and the nighttime. This trend can be attributed to the presence of various types of Cu that exhibit complex characteristics. Some fragmented Cu clouds possess limited thickness and area along with high radiation transmittance, so their reflectance is very similar to that of the ground surface. These properties make them easy to be misclassified as clear sky, thus resulting in relatively low recognition accuracy for Cu throughout both during the daytime and nighttime evaluations.



(**a**) (**b**)

Confusion matrix

True label \ Predicted label	Clear	Ci	Deep	Ac	As	Ns	Cu	Sc	St
Clear	89.432% 44717	1.240% 620	0.004% 2	0.916% 458	0.476% 238	0.018% 9	4.734% 2367	3.170% 1585	0.010% 5
Ci	8.295% 971	87.955% 10296	0.103% 12	0.513% 60	1.700% 199	0.094% 11	0.769% 90	0.572% 67	0.000% 0
Deep	0.084% 1	1.846% 22	86.493% 1031	1.091% 13	6.544% 78	3.356% 40	0.503% 6	0.084% 1	0.000% 0
Ac	17.238% 900	1.973% 103	0.134% 7	72.342% 3777	0.977% 51	0.804% 42	2.030% 106	4.501% 235	0.000% 0
As	5.098% 396	2.858% 222	0.438% 34	0.875% 68	87.616% 6806	1.777% 138	0.360% 28	0.978% 76	0.000% 0
Ns	0.409% 26	0.252% 16	0.535% 34	0.551% 35	3.054% 194	92.837% 5897	0.708% 45	1.653% 105	0.000% 0
Cu	34.556% 4090	0.811% 96	0.008% 1	1.031% 122	0.507% 60	0.414% 49	54.005% 6392	8.652% 1024	0.017% 2
Sc	5.207% 1401	0.134% 36	0.000% 0	0.320% 86	0.216% 58	0.145% 39	1.490% 401	92.444% 24874	0.045% 12
St	1.223% 12	0.000% 0	0.000% 0	0.102% 1	0.102% 1	0.102% 1	0.408% 4	8.257% 81	89.806% 881

(**c**)

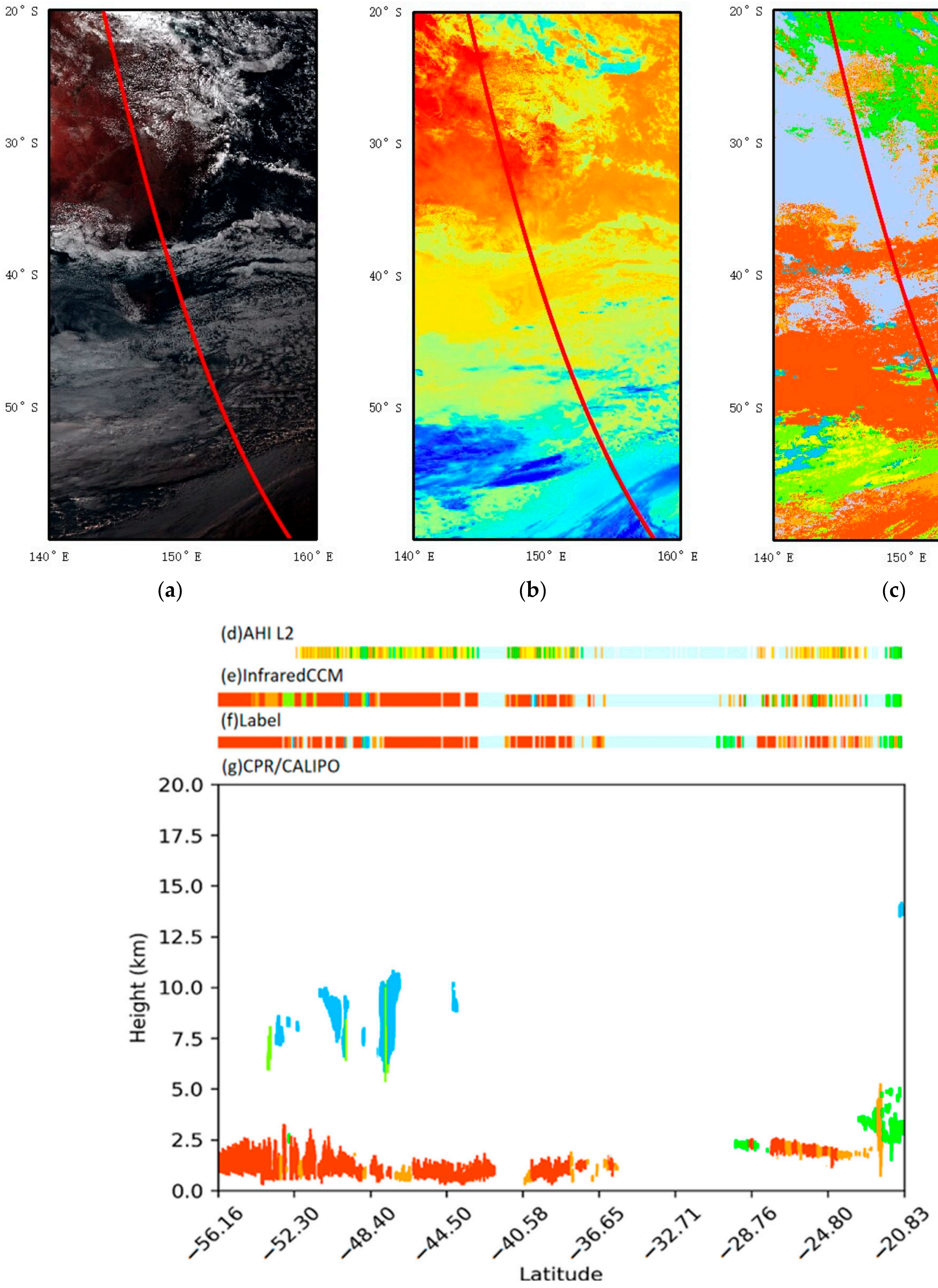
Confusion matrix

True label \ Predicted label	Clear	Ci	Deep	Ac	As	Ns	Cu	Sc	St
Clear	90.270% 2579	2.275% 65	0.000% 0	2.065% 59	0.595% 17	0.070% 2	0.770% 22	3.955% 113	0.000% 0
Ci	7.002% 78	90.754% 1011	0.090% 1	0.628% 7	1.167% 13	0.090% 1	0.000% 0	0.269% 3	0.000% 0
Deep	0.000% 0	0.000% 0	92.050% 220	1.255% 3	4.184% 10	2.092% 5	0.000% 0	0.418% 1	0.000% 0
Ac	10.188% 76	0.804% 6	0.000% 0	85.121% 635	0.670% 5	1.072% 8	0.536% 4	1.609% 12	0.000% 0
As	2.300% 21	1.424% 13	0.110% 1	0.986% 9	92.552% 845	1.972% 18	0.329% 3	0.329% 3	0.000% 0
Ns	0.092% 1	0.000% 0	0.000% 0	0.462% 5	1.939% 21	96.953% 1050	0.092% 1	0.462% 5	0.000% 0
Cu	24.481% 59	0.000% 0	0.000% 0	2.490% 6	0.415% 1	2.490% 6	55.602% 134	14.523% 35	0.000% 0
Sc	2.894% 67	0.043% 1	0.000% 0	0.432% 10	0.086% 2	0.259% 6	0.432% 10	95.724% 2216	0.130% 3
St	0.000% 0	0.000% 0	0.000% 0	0.000% 0	0.000% 0	0.000% 0	0.000% 0	4.348% 5	95.652% 110

(**d**)

Figure 7. Evaluation of classification metrics of the AInfraredCCM for daytime/nighttime. (**a**,**b**) Bar graphs of precision, recall, and F1-score of different types of cloud classification results; (**c**,**d**) confusion matrix of different types of cloud classification in daytime and nighttime.

On 15 May 2019, at 04:20 UTC, a cloud map covering southeastern Australia and the South Pacific Ocean was classified using the AInfraredCCM. The data corresponding to the CPR/CALIOP orbit are visually depicted in Figure 8a. The brightness temperature data collected from the 10.4 µm channel of the CPR/CALIOP orbit are displayed in Figure 8b. The AInfraredCCM classification results for this specific region are presented in Figure 8c. The radar track for this scenario extends from the lower-right corner to the upper-left corner of the cloud map, encompassing both daytime and nighttime hours. In this case, the CLTYPE encompasses both daytime and nighttime data, with data in the latitude range from −52.45°S to −56.56°S representing the nighttime region, as depicted in Figure 8d. As part of this comprehensive study, nighttime data, in addition to daytime data, underwent cloud classification. The classification results (Figure 8e) exhibit a significant degree of alignment with the single-layer cloud data (Figure 8f) obtained from the 2B-CLDCLASS-LIDAR. Figure 8f illustrates the outcomes of the combined cloud products of the CPR/CALIOP and Himawari-8 data. Figure 8g provides an insight into the vertical profile of the cloud types along the orbit for CPR/CALIOP.



(**a**) (**b**) (**c**)

Figure 8. Plot of classification results at 04:20 UTC on May 15, 2019. (**a**) RGB; (**b**) bright temperature; (**c**) AInfraredCCM results; (**d**,**e**) CLTYPE, AInfraredCCM results, and label; (**f**) the combined cloud products of the CPR/CALIOP and Himawari-8 data; (**g**) vertical profiles of the cloud types along the orbit of the CPR/CALIOP.

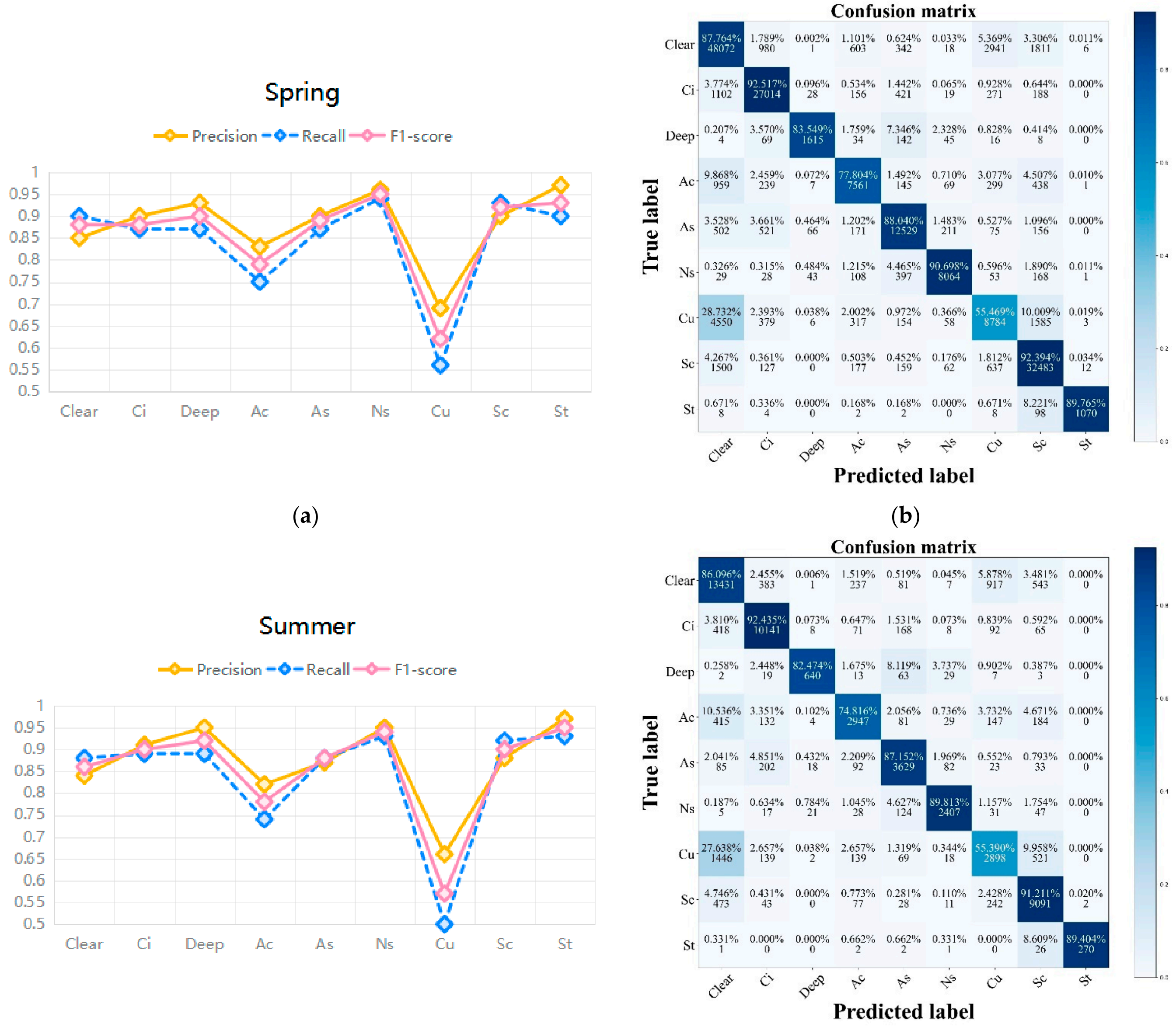
3.4. Effects of Different Seasons on Cloud Classification

The model was evaluated across all four seasons to determine its suitability. Table 12 presents the performance of AInfraredCCM in different seasons. Figure 9 shows the confusion matrices for the four seasons. The test dataset incorporates data from the four seasons and includes samples from daytime and nighttime observations as well as diverse subsurface information. During spring, summer, autumn, and winter, the overall accuracy was consistently >85% (86.61%, 85.60%, 85.87%, and 87.27% for spring, summer, autumn, and winter, respectively), which is consistent with the overall accuracy of the model. Ci, Dc, Ns, and St exhibited accuracies greater than 90%, coupled with recall and F1-scores of approximately 0.9. However, the classification of Cu remains challenging. Based on the characteristics of Cu, in subsequent experiments, we divided it into continuous and fragmented forms to improve the detection rate of Cu.

Table 12. Results of the four seasonal classifications.

Season	Cloud Type	Clear	Ci	Dc	Ac	As	Ns	Cu	Sc	St
					Accuracy = 86.61%					
Spring	Precision	0.85	0.90	0.93	0.83	0.90	0.96	0.69	0.90	0.97
	Recall	0.90	0.87	0.87	0.75	0.87	0.94	0.56	0.93	0.90
	F1-score	0.88	0.88	0.90	0.79	0.89	0.95	0.62	0.92	0.93
					Accuracy = 85.60%					
Summer	Precision	0.84	0.91	0.95	0.82	0.87	0.95	0.66	0.88	0.97
	Recall	0.88	0.89	0.89	0.74	0.88	0.93	0.50	0.92	0.93
	F1-score	0.86	0.90	0.92	0.78	0.88	0.94	0.57	0.90	0.95
					Accuracy = 85.87%					
Autumn	Precision	0.85	0.90	0.93	0.81	0.88	0.95	0.68	0.88	0.99
	Recall	0.89	0.87	0.90	0.74	0.88	0.93	0.54	0.93	0.90
	F1-score	0.87	0.89	0.91	0.77	0.88	0.94	0.60	0.90	0.94
					Accuracy = 87.27%					
Winter	Precision	0.87	0.91	0.86	0.84	0.91	0.95	0.69	0.89	0.99
	Recall	0.91	0.89	0.88	0.75	0.88	0.93	0.55	0.93	0.87
	F1-score	0.89	0.90	0.87	0.79	0.90	0.94	0.60	0.91	0.92

Figure 10 shows a comparative analysis of the results for all four seasons, with individual cases (a)–(d) corresponding to spring, summer, autumn, and winter, respectively. The three columns represent the brightness temperature for each case and the classification results of the AInfraredCCM and the CLTYPE. In general, the brightness temperature of a cloud is inversely proportional to its altitude, with lower brightness temperatures indicating higher cloud cover.



Confusion matrix

	Clear	Ci	Deep	Ac	As	Ns	Cu	Sc	St
Clear	87.764% 48072	1.789% 980	0.002% 1	1.101% 603	0.624% 342	0.033% 18	5.369% 2941	3.306% 1811	0.011% 6
Ci	3.774% 1102	92.517% 27014	0.096% 28	0.534% 156	1.442% 421	0.065% 19	0.928% 271	0.644% 188	0.000% 0
Deep	0.207% 4	3.570% 69	83.549% 1615	1.759% 34	7.346% 142	2.328% 45	0.828% 16	0.414% 8	0.000% 0
Ac	9.868% 959	2.459% 239	0.072% 7	77.804% 7561	1.492% 145	0.710% 69	3.077% 299	4.507% 438	0.010% 1
As	3.528% 502	3.661% 521	0.464% 66	1.202% 171	88.040% 12529	1.483% 211	0.527% 75	1.096% 156	0.000% 0
Ns	0.326% 29	0.315% 28	0.484% 43	1.215% 108	4.465% 397	90.698% 8064	0.596% 53	1.890% 168	0.011% 1
Cu	28.732% 4550	2.393% 379	0.038% 6	2.002% 317	0.972% 154	0.366% 58	55.469% 8784	10.009% 1585	0.019% 3
Sc	4.267% 1500	0.361% 127	0.000% 0	0.503% 177	0.452% 159	0.176% 62	1.812% 637	92.394% 32483	0.034% 12
St	0.671% 8	0.336% 4	0.000% 0	0.168% 2	0.168% 2	0.000% 0	0.671% 8	8.221% 98	89.765% 1070

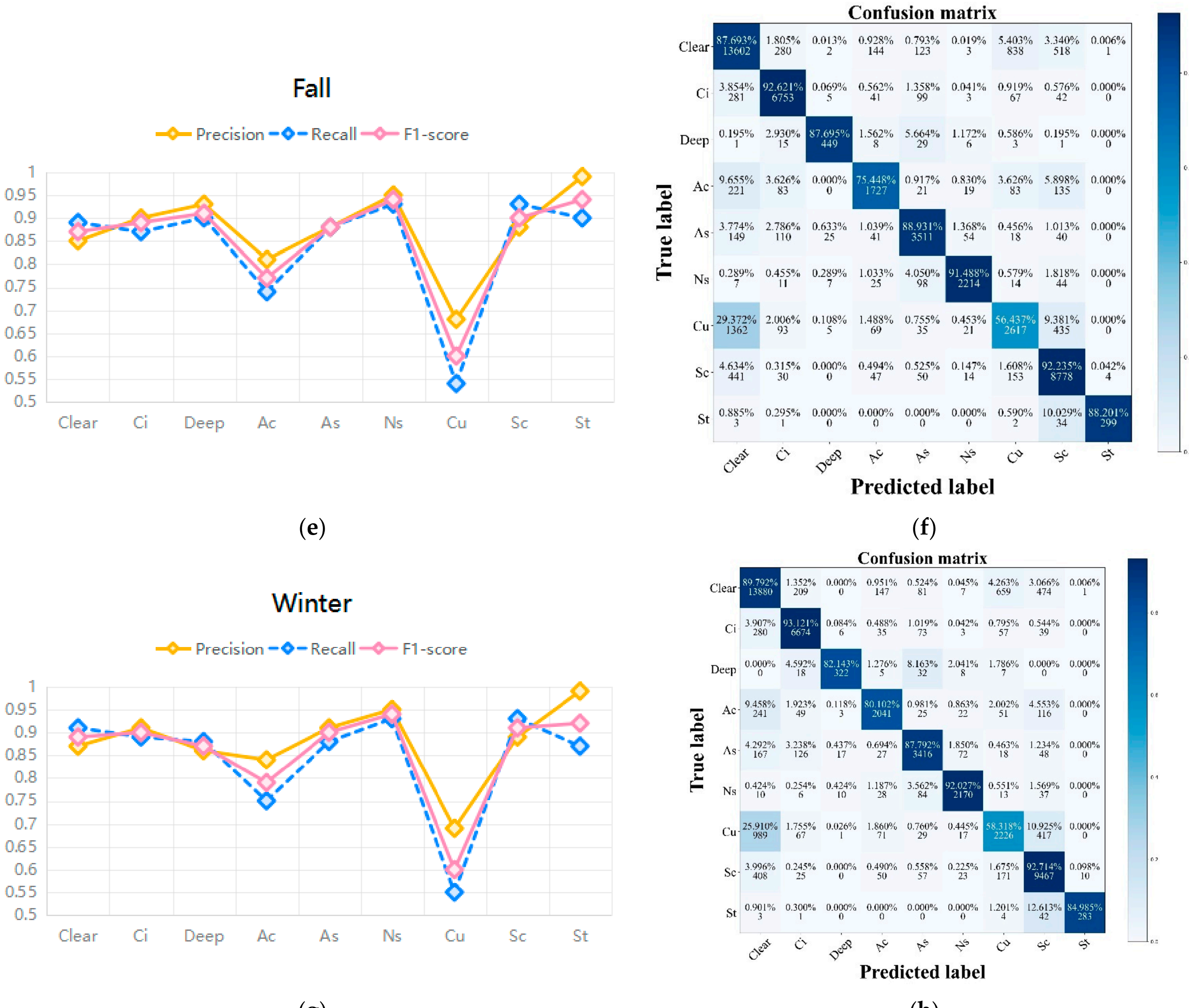
(**a**) (**b**)

Confusion matrix

	Clear	Ci	Deep	Ac	As	Ns	Cu	Sc	St
Clear	86.096% 13431	2.455% 383	0.006% 1	1.519% 237	0.519% 81	0.045% 7	5.878% 917	3.481% 543	0.000% 0
Ci	3.810% 418	92.435% 10141	0.073% 8	0.647% 71	1.531% 168	0.073% 8	0.839% 92	0.592% 65	0.000% 0
Deep	0.258% 2	2.448% 19	82.474% 640	1.675% 13	8.119% 63	3.737% 29	0.902% 7	0.387% 3	0.000% 0
Ac	10.536% 415	3.351% 132	0.102% 4	74.816% 2947	2.056% 81	0.736% 29	3.732% 147	4.671% 184	0.000% 0
As	2.041% 85	4.851% 202	0.432% 18	2.209% 92	87.152% 3629	1.969% 82	0.552% 23	0.793% 33	0.000% 0
Ns	0.187% 5	0.634% 17	0.784% 21	1.045% 28	4.627% 124	89.813% 2407	1.157% 31	1.754% 47	0.000% 0
Cu	27.638% 1446	2.657% 139	0.038% 2	2.657% 139	1.319% 69	0.344% 18	55.390% 2898	9.958% 521	0.000% 0
Sc	4.746% 473	0.431% 43	0.000% 0	0.773% 77	0.281% 28	0.110% 11	2.428% 242	91.211% 9091	0.020% 2
St	0.331% 1	0.000% 0	0.000% 0	0.662% 2	0.662% 2	0.331% 1	0.000% 0	8.609% 26	89.404% 270

(**c**) (**d**)

Figure 9. *Cont.*



Confusion matrix

	Clear	Ci	Deep	Ac	As	Ns	Cu	Sc	St
Clear	87.693% 13602	1.805% 280	0.013% 2	0.928% 144	0.793% 123	0.019% 3	5.403% 838	3.340% 518	0.006% 1
Ci	3.854% 281	92.621% 6753	0.069% 5	0.562% 41	1.358% 99	0.041% 3	0.919% 67	0.576% 42	0.000% 0
Deep	0.195% 1	2.930% 15	87.695% 449	1.562% 8	5.664% 29	1.172% 6	0.586% 3	0.195% 1	0.000% 0
Ac	9.655% 221	3.626% 83	0.000% 0	75.448% 1727	0.917% 21	0.830% 19	3.626% 83	5.898% 135	0.000% 0
As	3.774% 149	2.786% 110	0.633% 25	1.039% 41	88.931% 3511	1.368% 54	0.456% 18	1.013% 40	0.000% 0
Ns	0.289% 7	0.455% 11	0.289% 7	1.033% 25	4.050% 98	91.488% 2214	0.579% 14	1.818% 44	0.000% 0
Cu	29.372% 1362	2.006% 93	0.108% 5	1.488% 69	0.755% 35	0.453% 21	56.437% 2617	9.381% 435	0.000% 0
Sc	4.634% 441	0.315% 30	0.000% 0	0.494% 47	0.525% 50	0.147% 14	1.608% 153	92.235% 8778	0.042% 4
St	0.885% 3	0.295% 1	0.000% 0	0.000% 0	0.000% 0	0.000% 0	0.590% 2	10.029% 34	88.201% 299

(e) (f)

Confusion matrix

	Clear	Ci	Deep	Ac	As	Ns	Cu	Sc	St
Clear	89.792% 13880	1.352% 209	0.000% 0	0.951% 147	0.524% 81	0.045% 7	4.263% 659	3.066% 474	0.006% 1
Ci	3.907% 280	93.121% 6674	0.084% 6	0.488% 35	1.019% 73	0.042% 3	0.795% 57	0.544% 39	0.000% 0
Deep	0.000% 0	4.592% 18	82.143% 322	1.276% 5	8.163% 32	2.041% 8	1.786% 7	0.000% 0	0.000% 0
Ac	9.458% 241	1.923% 49	0.118% 3	80.102% 2041	0.981% 25	0.863% 22	2.002% 51	4.553% 116	0.000% 0
As	4.292% 167	3.238% 126	0.437% 17	0.694% 27	87.792% 3416	1.850% 72	0.463% 18	1.234% 48	0.000% 0
Ns	0.424% 10	0.254% 6	0.424% 10	1.187% 28	3.562% 84	92.027% 2170	0.551% 13	1.569% 37	0.000% 0
Cu	25.910% 989	1.755% 67	0.026% 1	1.860% 71	0.760% 29	0.445% 17	58.318% 2226	10.925% 417	0.000% 0
Sc	3.996% 408	0.245% 25	0.000% 0	0.490% 50	0.558% 57	0.225% 23	1.675% 171	92.714% 9467	0.098% 10
St	0.901% 3	0.300% 1	0.000% 0	0.000% 0	0.000% 0	0.000% 0	1.201% 4	12.613% 42	84.985% 283

(g) (h)

Figure 9. Evaluation metrics of the model across 4 seasons; (**a**–**h**) are the line graphs and confusion matrix of precision, recall, and F1-score in 4 seasons.

Taking the spring scenario (a) as an example, the first image is a brightness temperature map captured on 15 May 2019 at 02:50 UTC. In this figure, the AInfraredCCM effectively classifies the entire cloud image into regions with lower brightness temperature values. However, CLTYPE struggles to identify nighttime areas. In the summer scenario (b) taken at the typhoon center at UTC 05:00 on 5 August 2017, the CLTYPE identified Ci and Dc within the core of the typhoon but encountered difficulties in distinguishing these clouds from clear skies over most of the region. This study is able to finely classify clouds based on their brightness and temperature characteristics. It also categorized fill areas in CLTYPE, thereby enhancing the coherence of the classification results. Owing to the complex cloud composition during typhoons, the classification results of AInfraredCCM are consistent with the actual conditions. In contrast, the CLTYPE exhibited inaccuracies such as mislabeling as Ac in autumn (c). Similarly, the winter scenario (d) resulted in Cu being erroneously classified as clear, with many small cloud structures being misclassified. This study employed a pixel-level cloud classification method that facilitated the identification of more fragmented clouds and cloud pixels that resemble clear skies.

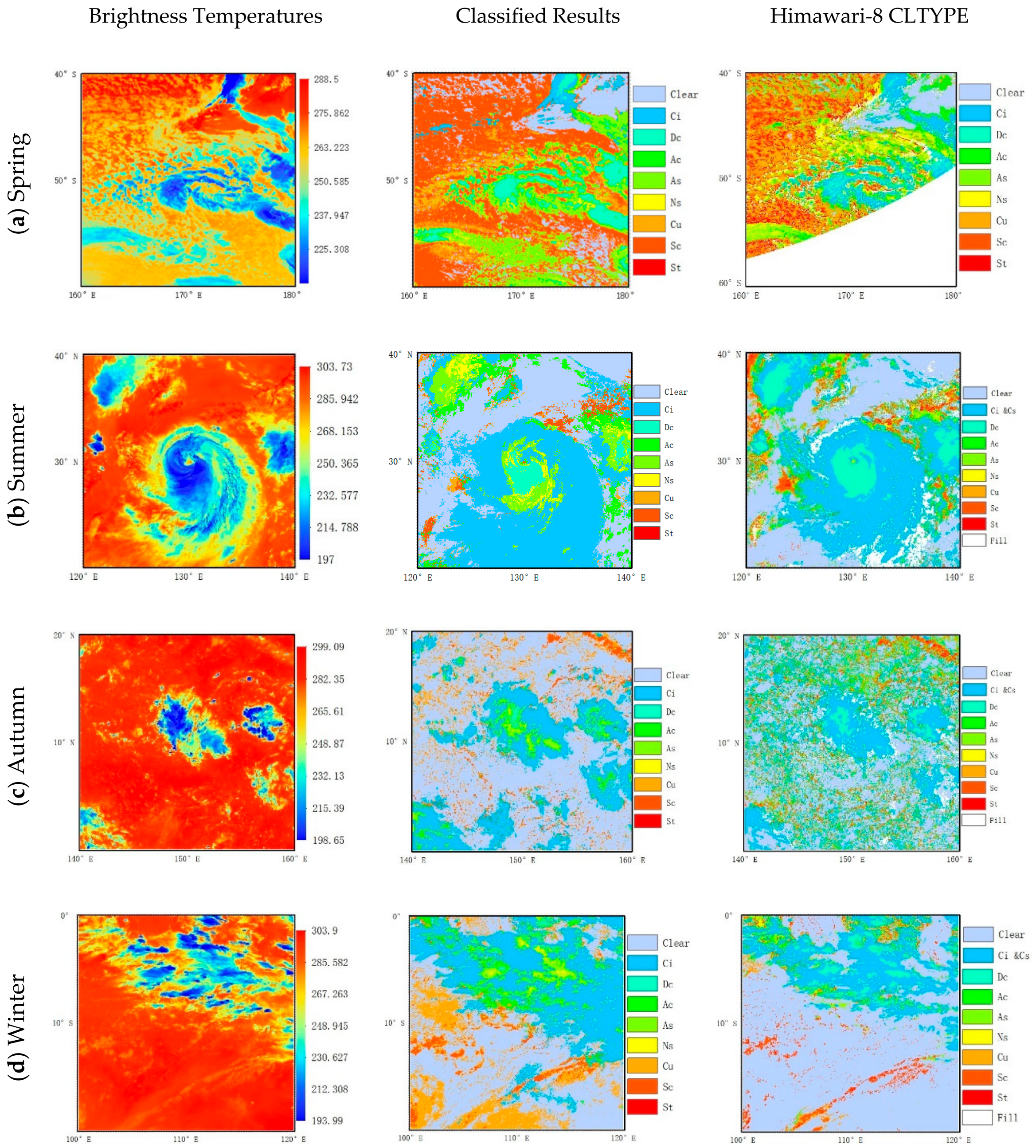


Figure 10. Bright temperatures (K), AInfraredCCM classification results, and Himawari-8 CLTYPE for different seasons. (**a**–**d**) Cloud map moments in the following order: 15 May 2019 UTC 02:50; 5 August 2017 UTC 05:00; 10 October 2018 UTC 03:30; and 1 February 2019 UTC 06:40, respectively.

4. Discussion

This study utilized Himawari-8 satellite infrared data and CPR/CALIOP cloud products for conducting comprehensive research on all-day cloud classification. The input features included latitude, longitude, brightness temperature, and brightness temperature difference data derived from Himawari-8. Cloud type labels were extracted from the CloudSat/CALIOP joint secondary product 2B-CLDCLASS-LIDAR. Following this, a cloud classification model was developed to handle all-day observations by leveraging satellite detection. Five machine learning models were employed for model development. The results highlight the superior performance of the AInfraredCCM based on the XGBoost model compared to that of the other models. It achieved impressive overall accuracy, precision, recall, and F1 scores of 86.22%, 0.88, 0.84, and 0.86, respectively.

The AInfraredCCM surpassed many previous models in terms of overall accuracy and cloud classification diversity (Table 9). Unlike previous studies that primarily focused on daytime data due to limited nighttime cloud product data, this study included both daytime and nighttime data. The model performed well at night, yielding an overall classification accuracy of 91.45%. Previous studies on nighttime cloud classification often had a limited range of cloud types. In this study, satellite data were categorized into nine types—clear, Ci, Dc, Ac, As, Ns, St, Sc, and Cu. Compared with other nighttime classification models, this model achieved classification of a broader range of cloud types while maintaining a good overall classification performance. However, due to the limited amount of nighttime data, it may result in a more favorable classification performance. In future research, it is advisable to increase the volume of nighttime data. Through the AInfraredCCM, continuous cloud classification was realized throughout the day with a time resolution of up to 10 min. These high-resolution and continuous cloud classification results provide a richer dataset for exploring various cloud parameters and support all-day meteorological monitoring.

After conducting a comparative and analytical evaluation of the classification results across all-day, daytime, nighttime, and four-season scenarios, it was evident that the classification performance for Cu clouds was relatively subpar. The confusion matrix (Figures 4, 7 and 9) clearly depicts that Cu clouds are predominantly misclassified as clear skies and Sc. This phenomenon is closely related to the developmental stages of Cu. Cu clouds were classified into two types: fragmented and continuous Cu. Fragmented Cu are usually small, which makes their detection using satellite observations challenging. These clouds are easily dispersed by wind and have a short lifespan. Consequently, satellite-based systems may mistakenly identify clear skies as broken Cu clouds. Additionally, Cu clouds, as a convective system forming in the lower atmosphere, can transition into Sc clouds. Cu exhibited distinct characteristics when the convection intensified. If their convection development is suppressed, then Cu clouds may transform into Sc, leading to frequent misidentifications between the two cloud types.

There are some limitations that need to be further investigated. First, the cloud type labels used in this research were derived from the 2B-CLDCLASS-LiDAR product, which is based on data from CPR/CALIOP. During the label screening process, only pixels with a single layer of clouds were selected for comparison with the CLTYPE. Moreover, the majority principle was employed for cloud type screening during filtering labels. It is important to note that this screening method may result in the omission of certain data with limited feature information, particularly in the case of fragmented Cu and optically thin Ci. To improve the quality of the dataset, more refined label-generation techniques should be explored in future studies. Second, errors can arise when observing the same object using the Himawari-8 and CloudSat satellites from different observation positions. Spatiotemporal matching errors may be more pronounced, especially in regions with large observation angles [43,44]. Therefore, error correction is vital. In future research, the cloud positions in the AHI data could be computed using cloud information, including longitude, latitude, and cloud top height, collected by CALIPSO satellites. The corrected cloud positions can then be used to create a more accurate dataset [34,45].

5. Conclusions

This study describes an algorithm for all-day cloud classification, referred to as AInfraredCCM, which is based on machine learning and utilizes Himawari-8 data. AHI and CPR/CALIOP data collected in November 2018, January 2019, March 2019, and June–July 2019 were used to develop the algorithm. After preprocessing, the data were randomly divided according to a ratio of 9:1. Of the data, 90% were allocated to the training and validation datasets for algorithm development, while the remaining 10% were reserved for testing. The predictors employed in this algorithm include 10 IR channels, 5 BTDs channels, and latitude and longitude information. The conclusions of this study are detailed here:

(1) The overall accuracy, precision, recall, and F1-score of AInfraredCCM cloud classification were 86.22%, 0.88, 0.84, and 0.86, respectively. Notably, the here-proposed model

outperformed the other models selected for this study (Table 8) and those proposed by other researchers (Table 9). These results indicate that it is an efficient all-day cloud classification method.

(2) The model performed well when used for all-day cloud classification or when used separately for daytime and nighttime classification, which suggests that the AInfraredCCM provides continuous data for cloud classification research throughout the day.

(3) The model was applied to both day and night scenarios as well as to four seasons and produced good classification results. In addition to Cu, this study demonstrated efficacy in classifying other cloud types. More emphasis should be laid on Cu in future studies.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs15245630/s1, Figures S1–S130 represent the thematic map of Himawari-8's official cloud-type products and the classification results generated by the AInfraredCCM at UTC 4:00 for 130 days.

Author Contributions: The authors' contributions are as follows: conceptualization, methodology, validation, formal analysis, and data curation, Y.F. and W.Z.; writing—original draft preparation, Y.F. and X.M.; writing—review and editing, all authors; supervision, Z.H. and Q.L.; funding acquisition, W.Z. and X.M.; and project administration, X.G. and T.Y. All authors have read and agreed to the published version of the manuscript.

Funding: The Major Project of High-Resolution Earth Observation System, No. 30-Y60B01-9003-22/23 and No. 30-Y30F06-9003-20/22; North China Institute of Aerospace Engineering Foundation of Doctoral Research, BKY-2021-31; Science and Technology Research Projects of Higher Education Institutions in Hebei Province, ZD2021303; Hebei Province Graduate Student Innovation Ability Training Funding Project, No. CXZZSS2023166, North China Institute of Aerospace Engineering's University-level Innovation Project, No. YKY-2022-58.

Data Availability Statement: The data in this article can be found online at https://github.com/tpmao/cloud-classification-data, accessed on 13 October 2023.

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

Appendix A

Table A1 shows 130 images of the whole area of Himawari-8 from 2018 to 2019, which is used to compare and verify the model classification results with Himawari-8 CLTYPE in Section 3.2. The data are categorized into four folders based on the seasons.

Table A1. The statistics of AInfraredCCM classification results for the four seasons and corresponding CLTYPE of Himawari-8.

Season	No.	Data ID	No.	Data ID
Spring	1	20190301_0400	17	20190317_0400
	2	20190302_0400	18	20190318_0400
	3	20190303_0400	19	20190319_0400
	4	20190304_0400	20	20190320_0400
	5	20190305_0400	21	20190321_0400
	6	20190306_0400	22	20190322_0400
	7	20190307_0400	23	20190323_0400
	8	20190308_0400	24	20190324_0400
	9	20190309_0400	25	20190325_0400

Table A1. *Cont.*

Season	No.	Data ID	No.	Data ID
	10	20190310_0400	26	20190326_0400
	11	20190311_0400	27	20190327_0400
	12	20190312_0400	28	20190328_0400
	13	20190313_0400	29	20190329_0400
	14	20190314_0400	30	20190330_0400
	15	20190315_0400	31	20190331_0400
	16	20190316_0400		
Summer	32	20190601_0400	51	20190620_0400
	33	20190602_0400	52	20190622_0400
	34	20190603_0400	53	20190623_0400
	35	20190604_0400	54	20190624_0400
	36	20190605_0400	55	20190625_0400
	37	20190606_0400	56	20190626_0400
	38	20190607_0400	57	20190627_0400
	39	20190608_0400	58	20190628_0400
	40	20190609_0400	59	20190630_0400
	41	20190610_0400	60	20190701_0400
	42	20190611_0400	61	20190702_0400
	43	20190612_0400	62	20190703_0400
	44	20190613_0400	63	20190704_0400
	45	20190614_0400	64	20190705_0400
	46	20190615_0400	65	20190706_0400
	47	20190616_0400	66	20190707_0400
	48	20190617_0400	67	20190708_0400
	49	20190618_0400	68	20190709_0400
	50	20190619_0400	69	20190710_0400
Autumn	70	20181101_0400	85	20181116_0400
	71	20181102_0400	86	20181117_0400
	72	20181103_0400	87	20181118_0400
	73	20181104_0400	88	20181119_0400
	74	20181105_0400	89	20181120_0400
	75	20181106_0400	90	20181121_0400
	76	20181107_0400	91	20181122_0400
	77	20181108_0400	92	20181123_0400
	78	20181109_0400	93	20181124_0400
	79	20181110_0400	94	20181125_0400
	80	20181111_0400	95	20181126_0400
	81	20181112_0400	96	20181127_0400
	82	20181113_0400	97	20181128_0400
	83	20181114_0400	98	20181129_0400
	84	20181115_0400	99	20181130_0400

Table A1. *Cont.*

Season	No.	Data ID	No.	Data ID
Winter	100	20190601_0400	116	20190617_0400
	101	20190602_0400	117	20190618_0400
	102	20190603_0400	118	20190619_0400
	103	20190604_0400	119	20190620_0400
	104	20190605_0400	120	20190621_0400
	105	20190606_0400	121	20190622_0400
	106	20190607_0400	122	20190623_0400
	107	20190608_0400	123	20190624_0400
	108	20190609_0400	124	20190625_0400
	109	20190610_0400	125	20190626_0400
	110	20190611_0400	126	20190627_0400
	111	20190612_0400	127	20190628_0400
	112	20190613_0400	128	20190629_0400
	113	20190614_0400	129	20190630_0400
	114	20190615_0400	130	20190631_0400
	115	20190616_0400		

References

1. Tapakis, R.; Charalambides, A.G. Equipment and methodologies for cloud detection and classification: A review. *Sol. Energy* **2013**, *95*, 392–430. [CrossRef]
2. Stubenrauch, C.J.; Rossow, W.B.; Kinne, S.; Ackerman, S.; Cesana, G.; Chepfer, H.; Di Girolamo, L.; Getzewich, B.; Guignard, A.; Heidinger, A.; et al. Assessment of Global Cloud Datasets from Satellites: Project and Database Initiated by the GEWEX Radiation Panel. *Bull. Am. Meteorol. Soc. Bull. Am. Meteorol. Soc.* **2013**, *94*, 1031–1049. [CrossRef]
3. Rossow, W.B.; Mosher, F.; Kinsella, E.; Arking, A.; Desbois, M.; Harrison, E.; Minnis, P.; Ruprecht, E.; Seze, G.; Simmer, C. ISCCP cloud algorithm intercomparison. *J. Appl. Meteorol. Clim.* **1985**, *24*, 877–903. [CrossRef]
4. Zhuang, Z.H.; Wang, M.; Wang, K.; Li, S.; Wu, J. Research progress of ground-based cloud classification technology based on deep learning. *J. Nanjing Univ. Inf. Sci. Technol. (Nat. Sci. Ed.)* **2022**, *14*, 566–578. [CrossRef]
5. Zhao, C.; Garrett, T.J. Effects of Arctic haze on surface cloud radiative forcing. *Geophys. Res. Lett.* **2015**, *42*, 557–564. [CrossRef]
6. Liu, Y.; Xia, J.; Shi, C.-X.; Hong, Y. An Improved Cloud Classification Algorithm for China's FY-2C Multi-Channel Images Using Artificial Neural Network. *Sensors* **2009**, *9*, 5558–5579. [CrossRef]
7. Chen, D.; Guo, J.; Wang, H.; Li, J.; Min, M.; Zhao, W.; Yao, D. The Cloud Top Distribution and Diurnal Variation of Clouds Over East Asia: Preliminary Results From Advanced Himawari Imager. *J. Geophys. Res. Atmos.* **2018**, *123*, 3724–3739. [CrossRef]
8. Astafurov, V.G.; Skorokhodov, A.V. Using the results of cloudclassification based on satellite data for solving climatological andmeteorological problems. *Russ. Meteorol. Hydrol.* **2021**, *46*, 839–848. [CrossRef]
9. Toğaçar, M.; Ergen, B. Classification of cloud images by using super resolution, semantic segmentation approaches and binary sailfish optimization method with deep learning model. *Comput. Electron. Agric.* **2022**, *193*, 106724. [CrossRef]
10. Zhang, C.; Zhuge, X.; Yu, F. Development of a high spatiotemporal resolution cloud-type classification approach using Himawari-8 and CloudSat. *Int. J. Remote Sens.* **2019**, *40*, 6464–6481. [CrossRef]
11. Wohlfarth, K.; Schröer, C.; Klaß, M.; Hakenes, S.; Venhaus, M.; Kauffmann, S.; Wilhelm, T.; Wohler, C. Dense Cloud Classification on Multispectral Satellite Imagery. In Proceedings of the 2018 10th IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS), Beijing, China, 19–20 August 2018; pp. 1–6. [CrossRef]
12. Yu, Z.; Ma, S.; Han, D.; Li, G.; Gao, D.; Yan, W. A cloud classification method based on random forest for FY-4A. *Int. J. Remote Sens.* **2021**, *42*, 3353–3379. [CrossRef]
13. Cai, K.; Wang, H. Cloud classification of satellite image based on convolutional neural networks. In Proceedings of the 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 24–26 November 2017; pp. 874–877. [CrossRef]
14. Afzali Gorooh, V.; Kalia, S.; Nguyen, P.; Hsu, K.-l.; Sorooshian, S.; Ganguly, S.; Nemani, R.R. Deep Neural Network Cloud-Type Classification (DeepCTC) Model and Its Application in Evaluating PERSIANN-CCS. *Remote Sens.* **2020**, *12*, 316. [CrossRef]
15. Olesen, F.-S.; Grassl, H. Cloud detection and classification over oceans at night with NOAA-7. *Int. J. Remote Sens.* **1985**, *6*, 1435–1444. [CrossRef]

16. Tan, Z.; Liu, C.; Ma, S.; Wang, X.; Shang, J.; Wang, J.; Ai, W.; Yan, W. Detecting Multilayer Clouds From the Geostationary Advanced Himawari Imager Using Machine Learning Techniques. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4103112. [CrossRef]
17. Li, W.; Zhang, F.; Lin, H.; Chen, X.; Li, J.; Han, W. Cloud Detection and Classification Algorithms for Himawari-8 Imager Measurements Based on Deep Learning. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4107117. [CrossRef]
18. Letu, H.; Nagao, T.M.; Nakajima, T.Y.; Riedi, J.; Ishimoto, H.; Baran, A.J.; Shang, H.; Sekiguchi, M.; Kikuchi, M. Ice Cloud Properties from Himawari-8/AHI Next-Generation Geostationary Satellite: Capability of the AHI to Monitor the DC Cloud Generation Process. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 3229–3239. [CrossRef]
19. Min, M.; Bai, C.; Guo, J.; Sun, F.; Liu, C.; Wang, F.; Xu, H.; Tang, S.; Li, B.; Di, D.; et al. Estimating Summertime Precipitation from Himawari-8 and Global Forecast System Based on Machine Learning. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 2557–2570. [CrossRef]
20. Min, M.; Wu, C.; Li, C.; Liu, H.; Xu, N.; Wu, X.; Chen, L.; Wang, F.; Sun, F.; Qin, D.; et al. Developing the science product algorithm testbed for Chinese next-generation geostationary meteorological satellites: Fengyun-4 series. *J. Meteorol. Res.* **2017**, *31*, 708–719. [CrossRef]
21. Bessho, K.; Date, K.; Hayashi, M.; Ikeda, A.; Imai, T.; Inoue, H.; Kumagai, Y.; Miyakawa, T.; Murata, H.; Ohno, T.; et al. An Introduction to Himawari-8/9—Japan's New-Generation Geostationary Meteorological Satellites. *J. Meteor. Soc. Jpn.* **2016**, *94*, 151–183. [CrossRef]
22. Stephens, G.; Winker, D.; Pelon, J.; Trepte, C.; Vane, D.; Yuhas, C.; L'Ecuyer, T.; Lebsock, M. CloudSat and CALIPSO within the A-Train: Ten Years of Actively Observing the Earth System. *Bull. Am. Meteorol. Soc. Bull. Am. Meteorol. Soc.* **2018**, *99*, 569–581. [CrossRef]
23. Winker, D.M.; Vaughan, M.A.; Omar, A.; Hu, Y.; Powell, K.A.; Liu, Z.; Hunt, W.H.; Young, S.A. Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms. *J. Atmos. Ocean. Technol.* **2009**, *26*, 2310–2323. [CrossRef]
24. Sassen, K.; Wang, Z.; Liu, D. Global distribution of cirrus clouds from CloudSat/Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) measurements. *J. Geophys. Res.* **2008**, *113*, D00A12. [CrossRef]
25. Stephens, G.L.; Vane, D.G.; Boain, R.J.; Mace, G.G.; Sassen, K.; Wang, Z.; Illingworth, A.J.; O'Connor, E.J.; Rossow, W.B.; Durden, S.L.; et al. THE CLOUDSAT MISSION AND THE A-TRAIN: A New Dimension of Space-Based Observations of Clouds and Precipitation. *Bull. Am. Meteorol. Soc.* **2002**, *83*, 1771–1790. [CrossRef]
26. Unglaub, C.; Block, K.; Mülmenstädt, J.; Sourdeval, O.; Quaas, J. A new classification of satellite-derived liquid water cloud regimes at cloud scale. *Atmos. Chem. Phys.* **2020**, *20*, 2407–2418. [CrossRef]
27. Wang, Z. *CloudSat 2B-CLDCLASS-LIDAR Product Process Description and Interface Control Document*; Process Description and Interface Control Document (PDICD) P1_R05; NASA: Washington, DC, USA, 2019; Volume 33.
28. Zhang, A.; Fu, Y. Life Cycle Effects on the Vertical Structure of Precipitation in East China Measured by Himawari-8 and GPM DPR. *Mon. Weather Rev.* **2018**, *146*, 2183–2199. [CrossRef]
29. Strabala, K.I.; Ackerman, S.A.; Menzel, W.P. Cloud Properties inferred from 8–12 µm Data. *J. Appl. Meteor. Climatol.* **1994**, *33*, 212–229. [CrossRef]
30. Chen, S.; Cheng, C.; Zhang, X.; Su, L.; Tong, B.; Dong, C.; Wang, F.; Chen, B.; Chen, W.; Liu, D. Construction of Nighttime Cloud Layer Height and Classification of Cloud Types. *Remote Sens.* **2020**, *12*, 668. [CrossRef]
31. Yue, Q.; Fetzer, E.J.; Kahn, B.H.; Wong, S.; Manipon, G.; Guillaume, A.; Wilson, B. Cloud-State-Dependent Sampling in AIRS Observations Based on CloudSat Cloud Classification. *J. Clim.* **2013**, *26*, 8357–8377. [CrossRef]
32. Berry, E.; Mace, G.G. Cloud properties and radiative effects of the Asian summer monsoon derived from A-Train data. *J. Geophys. Res. Atmos.* **2014**, *119*, 9492–9508. [CrossRef]
33. Behrangi, A.; Kubar, T.; Lambrigtsen, B. Phenomenological Description of Tropical Clouds Using CloudSat Cloud Classification. *Mon. Weather Rev.* **2012**, *140*, 3235–3249. [CrossRef]
34. Yang, Y.; Sun, W.; Chi, Y.; Yan, X.; Fan, H.; Yang, X.; Ma, Z.; Wang, Q.; Zhao, C. Machine learning-based retrieval of day and night cloud macrophysical parameters over East Asia using Himawari-8 data. *Remote Sens. Environ.* **2022**, *273*, 112971. [CrossRef]
35. Liu, C.; Yang, S.; Di, D.; Yang, Y.; Zhou, C.; Hu, X.; Sohn, B.-J. A Machine Learning-based Cloud Detection Algorithm for the Himawari-8 Spectral Image. *Adv. Atmos. Sci.* **2022**, *39*, 1994–2007. [CrossRef]
36. Tan, Y.; Zhang, W.; Yang, X.; Liu, Q.; Mi, X.; Li, J.; Yang, J.; Gu, X. Cloud and Cloud Shadow Detection of GF-1 Images Based on the Swin-UNet Method. *Atmosphere* **2023**, *14*, 1669. [CrossRef]
37. Fan, X.; Kong, J.L.; Zhong, Y.L.; Jiang, Y.Z.; Zhang, J.Y. Cloud Detection of Remote Sensing Images based on XGBoost Algorithm. *Remote Sens. Technol. Appl.* **2023**, *38*, 156–162.
38. Mommert, M. Cloud Identification from All-sky Camera Data with Machine Learning. *Astron. J.* **2020**, *159*, 178. [CrossRef]
39. Jiang, Y.; Cheng, W.; Gao, F.; Zhang, S.; Wang, S.; Liu, C.; Liu, J. A Cloud Classification Method Based on a Convolutional Neural Network for FY-4A Satellites. *Remote Sens.* **2022**, *14*, 2314. [CrossRef]
40. Wang, B.; Zhou, M.; Cheng, W.; Chen, Y.; Sheng, Q.; Li, J.; Wang, L. An Efficient Cloud Classification Method Based on a Densely Connected Hybrid Convolutional Network for FY-4A. *Remote Sens.* **2023**, *15*, 2673. [CrossRef]
41. Wang, Y.; Hu, C.; Ding, Z.; Wang, Z.; Tang, X. All-Day Cloud Classification via a Random Forest Algorithm Based on Satellite Data from CloudSat and Himawari-8. *Atmosphere* **2023**, *14*, 1410. [CrossRef]
42. Cermak, J.; Bendix, J. A novel approach to fog/low stratus detection using Meteosat 8 data. *Atmos. Res.* **2008**, *87*, 279–292. [CrossRef]

43. Guo, Q.; Feng, X.; Yang, C.; Chen, B. Improved Spatial Collocation and Parallax Correction Approaches for Calibration Accuracy Validation of Thermal Emissive Band on Geostationary Platform. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 2647–2663. [CrossRef]
44. Kim, H.-W.; Yeom, J.-M.; Shin, D.; Choi, S.; Han, K.-S.; Roujean, J.-L. An assessment of thin cloud detection by applying bidirectional reflectance distribution function model-based background surface reflectance using Geostationary Ocean Color Imager (GOCI): A case study for South Korea. *J. Geophys. Res. Atmos.* **2017**, *122*, 8153–8172. [CrossRef]
45. Vicente, G.A.; Davenport, J.C.; Scofield, R.A. The role of orographic and parallax corrections on real time high resolution satellite rainfall rate distribution. *Int. J. Remote Sens.* **2002**, *23*, 221–230. [CrossRef]