Cloud, Atmospheric Radiation and Renewal Energy Application (CARE) Version 1.0 Cloud Top Property Product From Himawari-8/AHI: Algorithm Development and Preliminary Validation

Xu Ri, Gegen Tana, Chong Shi, Takashi Y. Nakajima, Jiancheng Shi, Fellow, IEEE, Jun Zhao, Jian Xu, Member, IEEE, and Husi Letu

Abstract—Investigations of the effects of clouds on Earth’s radiation budget demand accurate representations of cloud top parameters, which can be efficiently obtained by large-scale satellite remote sensing approaches. However, the insufficient utilization of multiband information is one of the major sources of uncertainty in cloud top products derived from geostationary satellites. In this study, we developed a new algorithm to estimate Cloud, Atmospheric Radiation and renewal Energy application (CARE) version 1.0 cloud top properties [cloud top height (CTH), cloud top pressure (CTP), and cloud top temperature (CTT)]. The algorithm is constructed from ten thermal spectral measurements in Himawari-8 observations by using the random forest (RF) method to comprehensively consider the contribution of each band to the cloud top parameters. We chose the highly accurate Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) products in 2018 as the true values. The sensitivity analysis demonstrated that the products can be fully reproduced by using multiple Himawari-8 channels with the addition of the digital elevation model (DEM) data. The validation results of the 2019 CALIOP data confirm that the new algorithm shows an effective performance, with correlation coefficients (R) of 0.89, 0.89, and 0.90 for CTH, CTP, and CTT, respectively. Moreover, a significant improvement in the ice cloud estimation is achieved, in which the CTT R value increased from 0.46 to 0.70, as well as an improvement in the sea area, where the CTT R value increased from 0.71 to 0.84 compared with the Himawari-8 products of the Japan Aerospace Exploration Agency (JAXA) P-tree system. The further analyses performed here capture the diurnal cycle of cloud top parameters well in different temporal scales over the Asia–Pacific region.

Index Terms—Advanced Himawari Imager (AHI), Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation satellite (CALIPSO), cloud top parameters, ice cloud, remote sensing, random forests (RFs).

I. INTRODUCTION

Large uncertainties remain in the thermal radiation budgets represented by climate models and atmospheric products, thus requiring an enhanced understanding of cloud top parameters [1]. Therefore, there is a pressing need to better quantify cloud top properties to quantitatively describe the thermal cloud radiative forcing and climate change monitor [2]. The influence of cloud top layer on the top of atmosphere fluxes enhances the reflected solar flux, whereas clouds reduce the outgoing longwave flux relative to clear skies. This importance is reflected in many fields, such as aeronautical meteorological support and numerical weather prediction (NWP) [3].

Satellite remote sensing is one of the most effective ways to obtain cloud top information given its ability to provide continuous global data. Cloud products are retrieved from spaceborne passive and active measurements with unique advantages and limitations corresponding to specific sensors and algorithms [4]. Passive satellites, such as the High-Resolution Infrared Sounder (HIRS) onboard the global National Oceanic and Atmospheric Administration (NOAA), the Moderate Resolution Imaging Spectroradiometer (MODIS) [5] onboard the Aqua and the Terra satellites, Advanced Baseline Imager (ABI) [6] onboard the Geostationary Operational Environment Satellite (GOES)-R, and the Advanced Himawari Imager (AHI) [7] onboard the Himawari-8 geostationary satellite, have been routinely used to retrieve cloud top parameters in a large-coverage area with repeated observations. Compared with passive satellite retrieval, active sensors, such as the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) [8], the primary instrument on the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation satellite (CALIPSO), and the Cloud-Profiling Radar (CPR) [9] onboard the...
CloudSat satellite, can provide more accurate cloud top parameter estimations; however, their spatial coverage and repeat cycles are limited.

In regard to the retrieval algorithms used to operate passive satellite sensors, HIRS used a single infrared (IR) channel to obtain cloud top height (CTH) measurements based on the assumption that the observed satellite signal comes entirely from clouds [10]; this mechanism performs well for optically thick clouds and large clouds covering the satellite view, but nevertheless, the CTHs of thin clouds are underestimated with this method due to the neglect of the influence of terrestrial radiation in the satellite observed radiance [11]. In contrast to the single-channel method used by HIRS, MODIS uses the CO2-slicing method to infer cloud top pressure (CTP) measurements (Collection 6) based on the partial absorption of CO2 channels in IR bands approximately of 13.3–15 μm for mid-to-high-level clouds [12]. In addition, the MODIS CTP measurements are converted to CTH and cloud top temperature (CTT) measurements through the use of gridded meteorological products generated by the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) [5]. However, the CO2 solution results in large errors in the lower troposphere, where the instrument noise and uncertainty affect the GFS data, causing the measurements to miss approximately half of thin cirrus clouds [5].

As for gestational satellites, GOES-R/ABI provides CTT measurements by using IR observations with an IR radiative transfer model (RTM) based on an optimal estimation method. CTP and CTH measurements are then obtained from this CTT product via the atmospheric temperature profile provided by NWP data, where a mechanism is implemented to account for the large surface biases in the NWP data [6]. Level-2 CTT and CTP products have been released on the Japan Aerospace Exploration Agency (JAXA) homepage as a “P-Tree system” (ftp.ptree.jaxa.jp) from 2016 to the present [2]. The cloud microphysical properties algorithm used in the P-Tree system was developed by Nakajima et al. [13] and has been employed to retrieve cloud top properties using single-channel thermal IR data and the profile information of reanalysis data. Nevertheless, this process may cause some uncertainty that arises from the auxiliary data.

In contrast to passive satellites, for which it is difficult to retrieve vertical cloud information, active satellites (such as CALIOP and CPR) can derive global cloud profile information by using spaceborne lidar and millimeter-wave radar measurements, thus providing reliable cloud top parameters at a high spatial resolution with high precision [8], [9]. However, these instruments are nadir-viewing with low revisit times and cover very limited areas, thus leading to extreme scarceness in many areas of the world, such as over the oceans, inland lakes, mountainous areas, remote areas, and sparsely populated areas [8].

Since most currently used algorithms are based solely on 1–2 IR channels or the oxygen A-band and highly depend on the atmospheric reanalysis data and some auxiliary data to construct the cloud top parameters, they may induce uncertainty due to the lower use of the effective multiband information of satellite observations. In addition, most schemes are developed using the lookup table (LUT) method via inter/extrapolation or the complex optimal estimation approach based on the LUT-based RTM, which may decrease the calculation efficiency and availability when the interpolated parameters of the LUT are large. Advanced machine learning (ML) technology is promising for cloud retrieval tasks with high dimensionality by modeling the complex relationship of multiple parameters in nonlinear systems in the field of Earth sciences [14], [15]. Min et al. [16] primarily used four classical ML algorithms to obtain a CTH prediction model by directly using satellite observation data and NWP model data. However, the real-time NWP data of the GFS may introduce errors and require more operation time when input into the algorithm [16]. To avoid the problems mentioned above, we make full use of the advantages of ML methods in multichannel satellite data information mining and develop a new algorithm to retrieve cloud top parameters by using only satellite IR channels without depending on auxiliary data.

The purpose of this study is to develop a fast and convenient cloud top parameter retrieval algorithm using the random forest (RF) method to serve the development of the Cloud, Atmospheric Radiation and renewal Energy application (CARE) version 1.0 products. CARE products were developed within the framework of the CARE international symposium (http://www.slrss.cn/care/). The new product is expected to have an improved cloud top product accuracy derived from geostationary satellites compared to the accuracy of products based on traditional physics-based retrieval algorithms, thus furthering the wide application of Earth meteorological satellite data in nowcasting applications. We also conducted the sensitivity analysis of the thermal IR channel to cloud top parameters to make full use of the effective information of IR multichannel for retrieval. Moreover, our method can integrate the advantages of passive geostationary sensors to provide high spatiotemporal observations and active sensors to achieve accurate retrievals by selecting CALIOP products as the benchmark and compiling AHI observation data to train the regression model. In addition, the algorithm avoids the sophisticated RTM calculation procedure and the reliance on solar illumination, in which both daytime and nighttime data are handled in the same way.

This article is organized as follows. The satellite and auxiliary data used to train the model are introduced briefly in Section II. Section III presents the details of the new algorithms. In Section IV, the retrieval results of cloud top parameters based on the ML algorithm are given and verified by CALIOP and MODIS followed by characteristics analyses at different temporal and spatial scales. The summary is shown in Section V.

II. DATA

A. Himawari-8/AHI Data

Himawari-8 is a next-generation geostationary meteorological satellite that was successfully launched into geosynchronous orbit by the Japan Meteorological Agency (JMA) on October 7, 2014; observation data began to be published on July 7, 2015 [17]. The AHI onboard the satellite has 16 bands...
for different spatial resolutions ranging from 0.5 (visible) to 2.0 km (IR) with an observation frequency of 10 min [18]. At present, the satellite has been applied in the fields of surface and sea surface temperature retrieval, cloud and haze detection, aerosol data assimilation, and forest fire detection (Table I lists the Himawari-8/AHI specification) [19], [20]. The high spatiotemporal resolution of the Himawari-8/AHI product provides us an opportunity to study the cloud top parameters in depth.

B. CALIPSO/CALIOP Data

CALIPSO is the first polarization lidar with two wavelengths (at 532 and 1064 nm) carried on the CALIPSO satellite and flown in formation with the National Aeronautics and Space Administration (NASA) A-train constellation of satellites launched in May 2006. CALIPSO provides the global vertical structure of aerosols and clouds in the troposphere and lower stratosphere [21], [8]. The global multiyear dataset obtained from CALIPSO provides a new perspective on the Earth’s atmosphere that can help us better understand the role of aerosols and clouds in the climate system (https://subset.larc.nasa.gov/calipso/login.php).

C. MODIS Data

MODIS is a 36-band radiometer that covers the spectral range from 0.42 to 14.24 μm and is currently onboard the NASA Terra/Aqua satellites. Terra was launched into a descending orbit in December 1999, with an equator-crossing local solar time (LT) of 1030 h. Aqua was launched into an ascending orbit in May 2006, with an equator-crossing LT of 1330 h. MODIS has a spatial resolution varying from 250 m to 1 km depending on the spectral band and is widely used to illustrate the global distributions with measurements every 5-min covering an area of 2330 km × 2030 km. It provides cloud, aerosol, sea surface, wildfire, photosynthetically active radiation, and other physical product information that is widely used for weather analysis and forecasting, short-term climate prediction, and environmental and disaster monitoring [5] (https://ladsweb.modaps.eosdis.nasa.gov/search/).

III. METHODOLOGY

A. Physical Algorithm

The classical CO₂-slicing method is widely used in passive satellites as the official algorithm for determining cloud top parameters. The algorithm calculates radiative transfer with a single-layer cloud [5]. For a given cloud element in a given field of view (FOV), the radiance observed \( R(\nu) \) in a spectral band \( \nu \) can be written as follows [5]:

\[
R(\nu) = (1 - NE)R_{cl}(\nu) + NE[R_{obc}(\nu, P_c)]
\]  

(1)

where \( R_{cl}(\nu) \) is the clear-sky radiance, \( R_{obc}(\nu, P_c) \) is the opaque (black) cloud radiance from pressure level \( P_c, N \) is the fraction of the FOV covered with clouds, and \( E \) is the cloud emissivity. \( R_{obc}(\nu, P_c) \) can be calculated from the following function [5]:

\[
R_{obc}(\nu, P_c) = R_{cl}(\nu, P_c) - \int_{P_c}^{P_0} \tau(\nu, p) \frac{dB[v, T(p)]}{dp} dp. \tag{2}
\]

The upper and lower limits of the integral sign are the surface pressure and cloud pressure, respectively, and \( B[v, T(p)] \) is the Planck radiance at frequency \( \nu \). The temperature is represented as \( T(p) \), and \( \tau(\nu, p) \) represents the fractional transmittance of radiation at frequency \( \nu \) emitted from the pressure level \( p \) arriving at the top of the atmosphere \( (p = 0) \); the decrease in radiation from clear conditions is expressed in the second term on the right of the equation. CTP is estimated from the ratio of radiance between two spectral bands of wavenumbers \( \nu_1 \) and \( \nu_2 \) [22], [23]

\[
\frac{R(\nu_1) - R_{cl}(\nu_1)}{R(\nu_2) - R_{cl}(\nu_2)} = \frac{NE_1 \int_{P_c}^{P_0} \tau(\nu_1, p) \frac{dB[v_1, T(p)]}{dp} dp}{NE_2 \int_{P_c}^{P_0} \tau(\nu_2, p) \frac{dB[v_2, T(p)]}{dp} dp}. \tag{3}
\]

Within the FOV, assuming that \( E_1 \) and \( E_2 \) are almost equal, it allows us to specify the cloud pressure when the atmospheric temperature and transmittance profiles of the two spectral bands are known [16], [24].

B. CARE Cloud Top Property Algorithm

In this article, the RF method was introduced into the CARE cloud top parameter estimation algorithm. RF is a highly effective ML algorithm that was proposed by Buehlmann and Hothorn [24]. It has been widely used in classification and regression problems by integrating randomly containing multiple decision trees without the need for extensive hyperparameter adjustment [24]. The RF method has strong applicability because of its unique advantages and characteristics, such as its ability to estimate feature importance and achieve good performance even when there are many unknown features and noise in the dataset (http://scikit-learn.org/stable/).

The continuous first-class Himawari-8/AHI data for one year (January to December 2018) observed in IR bands (7–16) from 7.3 to 13.2 μm (see Table I) and digital elevation model (DEM) data were used as input features. Fig. 1 shows the flowchart of the cloud top properties estimation algorithm. The level-2 gridded CALIOP cloud top products dataset was collocated by taking the target output for the training, parameter adjustment, and prediction within this new algorithm to finally perform the cloud top parameter (CTT, CTH, and CTP) calculations. It should be noted that in this article, the matching algorithm calculates the minimum distance between the CALIOP footprint and the Himawari-8/AHI pixel and determines a specific matching point in approximately ± 5 min. We removed high-latitude Himawari-8/AHI samples (with large satellite angles) and focused only on the samples from 80 °E to 180 °E and from −50 °S to 50 °N, considering the potential influence of the FOV on cloud top parameter retrieval. To improve the prediction accuracy of low-value situations, we increased the number of samples. Based on the compiled Himawari-8/AHI level-1 satellite data from February to December 2018, 805 980 training samples were matched with CALIOP data. In addition, the collocated sample data
TABLE I
HIMAWARI-8/AHI SPECIFICATIONS (SST: SEA SURFACE TEMPERATURE)

| Channel & Infrared | Band(μm) | Spatial Resolution(km) | Calibration Accuracy | Primary Application |
|-------------------|----------|------------------------|----------------------|---------------------|
| Visible           | 1        | 0.47                   | 1.0                  | 2.63%               | Aerosol |
| & Near-           | 2        | 0.51                   | 1.0                  | 2.53%               | Vegetation |
| Infrared          | 3        | 0.64                   | 0.5                  | 2.55%               | Vegetation |
|                   | 4        | 0.86                   | 1.0                  | 2.39%               | Cirrus |
|                   | 5        | 1.61                   | 2.0                  | 2.73%               | Cloud, Snow |
|                   | 6        | 2.25                   | 2.0                  | 2.82%               | Cloud, Aerosol |
| Shortwave         | 7        | 3.88                   | 2.0                  | 0.42%               | Fire, Land and surface |
| Infrared          | 8        | 6.24                   | 2.0                  | 0.34%               | WV |
|                   | 9        | 6.94                   | 2.0                  | 0.29%               | WV |
|                   | 10       | 7.35                   | 2.0                  | 0.24%               | WV |
| Longwave          | 11       | 8.60                   | 2.0                  | 0.20%               | WV, Cloud |
| Infrared          | 12       | 9.63                   | 2.0                  | 0.21%               | Ozone |
|                   | 13       | 10.4                   | 2.0                  | 0.23%               | Cloud |
|                   | 14       | 11.24                  | 2.0                  | 0.22%               | SST, Cloud |
|                   | 15       | 12.38                  | 2.0                  | 0.20%               | SST, Cloud |
|                   | 16       | 13.28                  | 2.0                  | 0.22%               | Cloud |

Fig. 1. Flowchart of the cloud top properties estimation algorithm.

on the 10th and 16th of each month in 2019 were randomly selected as the independent verification datasets. In the training process, to avoid the phenomenon of overfitting and improve the convergence speed of the training, the proportion of the training, testing, and validation data was 9/13, 1/13, and 3/13 of the total sample points, respectively. After multiple debugging of the model parameters, the optimal structure was determined based on the comprehensive consideration of the stability and prediction accuracy of the model.

We constructed a sensitivity analysis of different groups of input parameters to show their effects on the cloud top parameters. The results are compared and shown in Table II. Several important indices, including the average deviation (MBE), mean square root error (RMSE), and determination coefficient ($R^2$), were used to evaluate the accuracy of the RF model in this study. We can see that the RF model involving all ten input IR bands and the DEM value delivered the optimum performance with a 32% increase in the RMSE compared to those when only two bands were applied. Table II also indicates that the channel data itself mainly contains decisive information for estimation, confirming that the IR channels detected by AHI can help the model better estimate the cloud properties and should all be considered in the retrieval algorithm.

To investigate the performance of the new algorithm on cloud top retrieval, we use the $K$-fold cross-validation technique [25] to preliminarily verify the accuracy of the selected RF model in this section. The principle of this technique is to divide the test data into $K$ equal subsets; then, one subset is used to validate the RF model and the remaining $K - 1$ data are adopted to train the model. In this case, a total of ten nonrepetitive training and validation steps are carried out for the RF model. The average error of the $K$-times validation can be regarded as the final accuracy of the RF model. One of the advantages of this technique is that all sample points are used to evaluate the performance of the selected model. Through many experiments, the best model with high stability and small errors is selected. The $K$-fold cross validation of cloud top parameters shown in Table III indicates that the selected model has a stable prediction accuracy and that the absolute value of MBE was very small in the new algorithm for estimating cloud top parameters.

IV. RESULTS AND DISCUSSION

A. Validation

To develop a scheme to retrieve CARE cloud top products from multispectral measurements in the Himawari-8 satellite, we need to understand cloud processes with sufficient accuracy and test the algorithm statistically. We adopted the CALIOP cloud top products as a benchmark for the evaluation of the CTT and CTH measurements derived from the new algorithm. Fig. 2 shows the scatter plots of CALIOP cloud top parameters (CTT and CTH) against CARE correspondence and AHI level-2 products, with their associated correlation coefficients ($R$) and joint histogram of the CALIOP-CTT/CTH difference. Overall, the instantaneous cloud top parameters agree well with the CALIOP product for all cases. The CARE CTT and CTH values showed higher consistency with CALIOP than the
corresponding AHI values, with $R$-values of 0.90 and 0.89 for CTT and CTH, respectively.

The new method for estimating cloud top parameters can significantly improve the accuracy of the resulting products and is comparable to other spaceborne lidar measurements. Further characterization of the CTT and CTH biases is shown in the joint histogram results. The CTT and CTH difference for the CARE product with CALIOP is primarily below 10 K and 4 km, respectively. In total, 205 670 sample points were used to test the performance of the results, and the validation was limited to the following aspects: First, the CALIOP measurements are conducted in orbit and thus cannot provide validation over a large area. Second, the horizontal resolution of the Clay data varies at three different altitudes from 333 to 1667 m. We use only the top product to sample the data to a horizontal resolution of 333 m, which may have introduced bias in the resampled result.

1) Validation for Each Area: The validation of the CTT and CTH retrieval algorithms was conducted using the current official AHI products, with collocated CALIOP data used as the benchmark for the evaluations over the ocean and land, in which great improvements were found. Fig. 3 shows the scatter plots of AHI and CARE CTT quantitatively compared against the CALIOP product as well as their sample numbers and the $R$ values derived for the ocean and land areas in 2019. As we can see from Fig. 3, each parameter value has two peaks, assumed to be the reason for different cloud phases. We can also see that the $R$-value over land is 0.79 with an MBE of 2.17 K and a reasonable RMSE of 14.87 K for the CARE CTT results in the land area. Compared to the AHI CTT, these results reveal an improvement in precision obtained through the introduction of the new method. The $R$ value of 0.84 with MBE of 0.2 K and the RMSE of 23.04 K for CARE CTT in the sea area demonstrate substantial improvements compared to AHI cloud top products. As shown in Fig. 3, we also investigated the histogram of CTT differences for the

---

### Table II

| Different groups (band number) | Variable 1 (K) | Variable 2 (km) | Variable 3 (hpa) |
|-------------------------------|---------------|----------------|-----------------|
|                               | $R$    | MBE | RMSE | $R$    | MBE | RMSE | $R$    | MBE | RMSE |
| All IR BAND, DEM             | 0.857  | 0.428 | 17.69 | 0.865 | 0.18 | 2.98 | 0.857 | 13.67 | 164.37 |
| All IR BAND                  | 0.847  | 0.726 | 18.49 | 0.854 | 0.21 | 3.0  | 0.841 | 17.26 | 166.18 |
| 8, 9, 10, 11, 12, 13, 14, 15, 16 | 0.843 | 0.982 | 18.75 | 0.850 | 0.17 | 3.03 | 0.839 | 15.27 | 166.82 |
| 9, 10, 11, 12, 13, 14, 15, 16 | 0.843 | 1.015 | 18.73 | 0.850 | 0.15 | 3.03 | 0.839 | 13.68 | 166.42 |
| 10, 11, 12, 13, 14, 15, 16   | 0.836 | 0.843 | 19.10 | 0.843 | 0.16 | 3.09 | 0.835 | 14.75 | 168.61 |
| 11, 12, 13, 14, 15, 16       | 0.835 | 0.845 | 19.17 | 0.842 | 0.17 | 3.10 | 0.832 | 15.98 | 169.99 |
| 12, 13, 14, 15, 16           | 0.819 | 0.46 | 19.95 | 0.825 | 0.24 | 3.25 | 0.814 | 19.86 | 178.33 |
| 13, 14, 15, 16               | 0.794 | 0.87 | 21.17 | 0.803 | 0.10 | 3.42 | 0.801 | 12.92 | 185.16 |
| 14, 15, 16                   | 0.758 | 0.94 | 22.68 | 0.771 | 0.09 | 3.66 | 0.774 | 13.45 | 193.62 |
| 15, 16                       | 0.584 | 0.43 | 28.32 | 0.503 | 0.22 | 4.98 | 0.572 | 19.96 | 252.36 |

### Table III

| Variable1 (K) | Variable2 (km) | Variable3 (hpa) |
|---------------|----------------|-----------------|
| MBE | RMSE | MBE | RMSE | MBE | RMSE | R
| 1   | 0.15 | 23.64 | 0.84 | 0.22 | 3.32 | 0.84 | 14.75 | 144.18 | 0.84 |
| 2   | 0.17 | 19.59 | 0.83 | 0.26 | 3.15 | 0.83 | 13.89 | 146.82 | 0.82 |
| 3   | 0.21 | 22.96 | 0.82 | 0.35 | 2.93 | 0.85 | 15.68 | 144.18 | 0.85 |
| 4   | 0.24 | 24.25 | 0.82 | 0.46 | 4.17 | 0.82 | 12.75 | 156.82 | 0.83 |
| 5   | 0.18 | 30.54 | 0.83 | 0.22 | 3.69 | 0.83 | 15.89 | 149.58 | 0.83 |
| 6   | 0.16 | 22.94 | 0.84 | 0.38 | 1.93 | 0.86 | 16.68 | 154.76 | 0.81 |
| 7   | 0.10 | 18.95 | 0.85 | 0.42 | 2.64 | 0.84 | 14.75 | 153.68 | 0.82 |
| 8   | 0.23 | 17.28 | 0.82 | 0.34 | 3.14 | 0.85 | 12.89 | 131.92 | 0.85 |
| 9   | 0.18 | 19.20 | 0.83 | 0.21 | 2.96 | 0.85 | 13.76 | 144.18 | 0.83 |
| 10  | 0.20 | 21.64 | 0.84 | 0.19 | 3.45 | 0.83 | 13.57 | 152.77 | 0.83 |
CARE and AHI with CALIOP. The results show a tendency for AHI products to overestimate the CTT in land area. The scatter plots of AHI and CARE CTH quantitatively compared against the CALIOP product as well as their sample numbers and $R$ values for sea area and land area in 2019 are shown in Fig. 4. An obvious CTH improvement, with $R$-values ranging from 0.72 to 0.86, was observed by comparison with the AHI products. The CARE CTH values obtained over the land area had an $R$ value of 0.75 with 0.65-km MBE and 2.44-km RMSE compared to CALIOP. The CARE CTH coefficients shown in Fig. 4 yielded the highest values, indicating that the cloud top parameters derived in this research are reliable over the ocean. However, the lowest $R$ value is only 0.64 for the CTH of the AHI products, and the corresponding MBE and RMSE values are 1.71 and 3.33 km, respectively, for the land site. The validation was followed by the cloud top parameter evaluation for each cloud phase.

2) Validation for Each Cloud Phase: Fig. 5 shows the comparison results of the instantaneous CARE CTT and AHI CTT from a total of 51,185 pixels against the CALIOP product for different cloud phases in 2019. The results demonstrate that the overall $R$-values of CARE CTT are high for both sources; CTT yielded an $R$ of 0.74 with an MBE of 2.34 K for water clouds and an $R$ of 0.70 with a higher MBE of 0.39 K for ice clouds. Fig. 5 also shows a tendency for the AHI product and CARE CTT to underestimate the CTT of water clouds and to overestimate the CTT of ice clouds compared to CALIOP. Interestingly, the $R$ value obtained for the ice water CTT increased obviously from 0.46 for AHI to 0.70 for CARE CTT compared to the CALIOP products. Fig. 6 further confirms the significant improvement in the ice pattern. Fig. 6 shows the similar scatter and histogram plots but for the CTH validation, with the $R$ value of 0.71 carrying an MBE of 0.41 km and RMSE of 3.0 km for the CARE CTH results in the water cloud. The AHI values tend to overestimate the CTH values for water clouds, and this overestimation can be attributed to temperature inversions near the cloud tops, as shown in a previous study [26]. In addition, the $R$-value of 0.79 with a small MBE of 0.21 km and an RMSE of 2.57 km derived for the CARE CTH estimations in the ice cloud case also shows a remarkable improvement in CTH accuracy compared to the AHI products. The results obtained from the new algorithm for cloud top parameters and CALIOP measurements agree well with each other; the increased accuracy when ice cloud situations occur can be attributed to the fact that CALIOP provides reliable heights of thin cirrus clouds containing small ice particles [27].

Table IV shows the statistical evaluation of CTP for different areas and cloud phases. This is a way to gain more insight into the capability of the new algorithm; the $R$ value is 0.89 with an MBE of 11.4 hpa and 149.01-hpa RMSE for the CARE CTP results in the full Asia–Pacific area. This evaluation confirms that the CARE CTP estimations perform well when reproducing the CTP, thus supplementing the absence of the CTP product on the Himawari-8/AHI official cloud products. For the different areas and different cloud phases, the corresponding $R$ values are also reliable for the CARE CTP results.

3) Intercomparison Validation: To further investigate the performance of the new algorithm, we conducted an intercomparison validation of the CARE cloud top properties...
TABLE IV
STATISTICAL SUMMARY OF CTP EVALUATION FOR EACH AREA AND DIFFERENT CLOUD PHASES

|        | Full Disk | Land | Sea | Water | Ice |
|--------|-----------|------|-----|-------|-----|
| N      | 46357     | 14795| 29131| 37175 | 18120|
| R      | 0.89      | 0.71 | 0.84 | 0.74  | 0.61 |
| MBE    | 11.4      | 52.23| 5.48 | 26.64 | 5.31 |
| RME    | 149.01    | 146.31| 192.34 | 185.61 | 131.10|

by comparing them to the collocated pixels from different platforms. An instantaneous MODIS granule acquired on October 3, 2018, with a resolution of 1 km is selected for validation. Fig. 7 presents the observed cloud top properties image for MODIS, CARE, and AHI [(a)–(c) for CTT, (d)–(f) for CTH, and (g) and (h) for CTP] on October 3, 2018. From Fig. 7, we can see that the zonal distributions of CARE products are comparable with other platforms. Interestingly, the results demonstrate that the spatial distribution of CARE cloud top products is generally similar to those of MODIS more than the AHI product. The lines of CALIOP products are overlapped on each image where the visibility of line is used to indicate the consistency of parameters with CALIOP data.

Fig. 8(a) and (b) shows a joint histogram of CTT and CTH difference between MODIS, CARE, and AHI with CALIOP. Due to the absence of CTP product from AHI, Fig. 8(c) only compared the MODIS and CARE CTP with CALIOP, where it shows that the CARE products have good consistency with validation samples of MODIS cloud top parameters. Fig. 8 also shows that the AHI observation CTT and CTT are slightly lower than that from MODIS correspondents. However, this study offers the validation results with MODIS only for limited cases, it is still necessary to fully verify the cloud characteristics on a near-global scale in the future. To verify the effectiveness of the new algorithm in the estimation of cloud top parameters, the independent and spatial-temporally MODIS product matched data are used here for validating the performances of the CARE products. Fig. 9 shows the MODIS cloud top properties overlaid with the CARE correspondences along with the CALIPSO overpass (20°S–6°S) on October 3, 2018. CALIOP, MODIS, and CARE are represented by black, brown, and red dots, respectively. The overall performance shows good agreement for CARE cloud top products compared to CALIOP. However, the deviation shown in MODIS CTH compared to CALIOP was relatively higher than CARE CTH in the whole research region, especially for the region between 10°S and 6°S. These results may relate to the sensor characteristics and the algorithms based on different IR bands.

B. Spatiotemporal Analysis of Cloud Top Parameters

1) Diurnal Variations in Cloud Top Parameters: As an application, the algorithm used to investigate the diurnal cycle of clouds in this article is based on data recorded at a 10-min
temporal resolution data. The region spans from 80°E to 145°E longitude and from 15°N to 55°N latitude, covering part of the Asian continent and the surrounding oceanic areas. The convective activities over the land in the afternoon and over the ocean in the morning are ordinarily strengthened by many mechanisms, such as sea–land breezes, convective forced gravity wave propagation in remote regions, and direct radiative convection [28], [29].

Fig. 10 shows the CTP, CTT, and CTH time series retrieved by the new algorithm every 4 h from 00:00 UTC. High clouds predominantly appear over the ocean throughout the day. The proportion of high-temperature coverage over the land is highest at night. Small-scale convective cloud growth occurs from 12:00 to 16:00 UTC, the cell volume increases, and the system is subsequently enlarged. Notably, low clouds do not appear when high clouds cover low clouds. One limitation of passive remote sensing is that it cannot show the characteristics of lower clouds.

2) Seasonal Analysis in Cloud Top Parameters: The diurnal characteristics of the CTP values described above can be assessed quantitatively by calculating the average latitudinal patterns in different seasons, as shown in Fig. 11. CTP was calculated every 10 min (superior to the temporal resolution of other passive satellites) in UTC and averaged over 30 days. Fig. 11 shows the temporal variations in the CTP in zonal distribution patterns from 60°N to 60°S latitude in different seasons, where January, April, July, and October represent winter, spring, summer, and autumn, respectively. From Fig. 11, we can see that the spatial distribution of CTP differs significantly in different latitudes for each month and the equatorial low-pressure zone occurred in all four seasons. The airflow rises and cools in the equatorial area attributed to the high temperature, strong evaporation, and high water vapor content in the air. To investigate the seasonal variation of CTP in diurnal scale over the different situation, CTP is calculated separately over the land and ocean. Fig. 12 shows the temporal variations in CTP patterns over the ocean and land areas in different seasons. The average CTP value is smallest at approximately 12:00–14:00 UTC, with the maximum peaks occurring at approximately 02:00–03:00 UTC for both cases (see Fig. 12). Subsequently, the CTP experiences a sharp decrease, while the CTP increases rapidly from 00:00 to 02:00 UTC. The CTP decreases slowly from 03:00 to 15:00 UTC over the land area. The diurnal cycle of the CTP is less pronounced over land areas than over ocean areas. This result shows the typical diurnal cycle of the cloud system in the region, where the analysis period was continued for 30 days.
3) Annual Analysis in Cloud Top Parameters: The characteristics of the high-spatial-scale and high-temporal-scale cloud top properties in the Asia–Pacific region in 2018 are analyzed in this section. Due to differences in natural and climatic environments, the spatial distribution of cloud top parameters differs significantly in each region. The results shown in Fig. 13 display the annual average distribution of the cloud top parameters over the Asia–Pacific region and the cloud top parameter percentage ranges and annual average statistics are shown in Fig. 14.

The spatial distributions of cloud top properties are visually similar to each other, though large discrepancies exist over many regions for each parameter. The results indicate that the annual average CTP, CTT, and CTH values are generally between 200 and 800 hpa, 220 and 280 K, and 2 and 14 km, respectively, in each region. A similar trend is seen from the zonal averages of CTT and CTP for different latitudes in the research area. Both algorithms have similar performances over the Asia–Pacific region. While the CTH has an opposite appearance to the pressure showing the highest values on the horizontal level and generally lower values than those of other regions at latitudes from 60 °S to 20 °S with a peak at approximately 30 °S, in its spatial distribution, the largest CTT in the Northern Hemisphere occurs at approximately 280 K, the largest Southern Hemisphere occurs at 30 °N, and the largest CTP in the Northern Hemisphere occurs at 800 hpa. These properties both decreased to the minimum values over the equatorial regions. Large biases were found between the results of CTT and CTP in the Southern Hemisphere through both methods. The highest CTT areas are distributed over the southeastern part of the study area, and the minimum CTT appears on the equatorial region.

V. Conclusion

The remote sensing technique offers an effective way to estimate cloud top properties, which plays an essential role in the radiation budget and the energy exchange process between the Earth and the atmosphere. The new generation of the Himawari-8 satellite has a greatly improved temporal and spatial resolution and the CALIOP cloud layer products are widely acknowledged to have a reliable accuracy. This product thus offers a great opportunity to construct an ML-based algorithm to rapidly estimate all cloud top parameters simultaneously with high accuracy at a high spatial–temporal scale. Thus, here, we developed a simple, yet efficient algorithm, that can make full use of multi-IR bands without the reliance of atmospheric reanalysis profiling data (concerned to have large uncertainty). In addition, the new algorithm provides 24-h products on a 10-min scale, covering missing nighttime data and obtaining CTP products from AHI.

The CALIOP data were also introduced in this research to validate the accuracy of the new algorithm. It is worth reemphasizing that the values estimated from the AHI data in 2019 showed good consistency with the corresponding CALIPSO product samples. A validation study of each cloud phase and different areas compared with those of AHI products confirms the significant accuracy samples over the ocean area and in the ice cloud patterns, reflecting the fact that the CALIOP product provides reliable altitudes of thin cirrus clouds composed mainly of small-sized ice particles. Moreover, the distribution of estimated cloud top properties against the MODIS cloud top products further confirms the performance of the estimation process.

Based on the accuracy evaluation process, the analyses of CTT, CTH, and CTP were performed on their corresponding spatial and temporal patterns. The application of the annual average distribution of the cloud top parameters over the Asia–Pacific region shows an expected similar pattern between CTP and CTT and a reverse distribution for CTH. We also analyzed the seasonal data for 30 days in 2019 to investigate the diurnal cycle of the cloud system over the research region. Both over land and ocean, the average CTP value is smallest at approximately 12:00–14:00 UTC and exhibits the highest value at approximately 02:00–03:00 UTC. These results seem to demonstrate the typical diurnal cycle of the cloud system in this area.

This work validates that the new algorithm is effective and practical with high accuracy for estimating cloud top parameters at high speeds, making it suitable for use in near real-time cloud top parameter estimation applications. However, it is worth noting that there may be uncertainty in the invalidation or analysis of the spatiotemporal features of the analyzed parameters since we used a dataset covering only one year. More validation on daily or even hourly time scales is still needed.

ACKNOWLEDGMENT

The authors would like to thank NASA for freely providing the MODIS and CALIOP data online. The Himawari-8 cloud property products used in this article were supplied by the P-Tree System, JAXA. They would also like to thank the team of Letu et al. from the State Key Laboratory of the Science and Remote Sensing, Institute of Remote Sensing and Digital Earth, for providing the cloud mask data.
REFERENCES

[1] Y. Huang, S. Siems, M. Manton, A. Protat, and H. Nguyen, “Evaluating Himawari-8 cloud products using shipborne and CALIPSO observations: Cloud-top height and cloud-top temperature,” J. Atmos. Ocean. Technol., vol. 36, no. 5, pp. 2327–2347, 2019.

[2] H. Letu et al., “A new benchmark for surface radiation products over the East Asia–Pacific region retrieved from the Himawari-8/ACHI next-generation geostationary satellite,” Bull. Amer. Meteorol. Soc., vol. 10, no. 3, pp. E873–E881, 2021.

[3] R. M. A. Vidale, M. T. Chahine, and M. Khairoutdinov, “Estimation of surface shortwave radiation from Himawari-8 satellite data based on a combination of radiative transfer and deep neural network,” IEEE Trans. Geosci. Remote Sens., vol. 58, no. 8, pp. 5304–5316, Aug. 2020.

[4] H. Iwabuchi, M. Saito, Y. Tokoro, N. S. Putri, and M. Sekiguchi, “Retrieval of radiative and microphysical properties of clouds from multispectral infrared measurements,” Prog. Earth Planet. Sci., vol. 3, no. 1, p. 32, Oct. 2016.

[5] W. P. Menzel et al., “MODIS global cloud-top pressure and amount estimation: Algorithm description and results,” J. Appl. Meteorol. Climatol., vol. 47, no. 4, pp. 1175–1198, Apr. 2008.

[6] A. K. Heidinger et al., “Using CALIPSO to explore the sensitivity to cirrus height in the infrared observations from NPOESS/VIIRS and GOES-R/ABI,” J. of Geophys. Res., vol. 115, 2010.

[7] K. Bessho et al., “An introduction to Himawari-89—Japan’s new-generation geostationary meteorological satellites,” J. Meteorol. Soc. Jpn., vol. 94, no. 2, pp. 151–183, 2016.

[8] D. M. Winker, M. A. Vaughan, A. Omar, Y. Hu, and K. A. Powell, “Overview of the CALIPSO mission and CALIOP data processing algorithms,” J. Atmos. Ocean. Technol., vol. 26, no. 11, pp. 2310–2323, 2009.

[9] G. Stephens, D. Vane, and R. J. Boain, “The CloudSat mission and the A-train: A new dimension of space-based observations of clouds and precipitation,” Bull. Amer. Meteorol. Soc., vol. 83, no. 12, pp. 1771–1790, 2002.

[10] W. P. Menzel, D. P. Wylie, and K. I. Strabala, “Seasonal and diurnal changes in cirrus clouds as seen in four years of observations with the VAS,” J. Appl. Meteorol., vol. 31, no. 4, pp. 370–385, Apr. 1992.

[11] D. P. Wylie, D. Santek, and D. O. Starr, “Cloud-top heights from GOES-8 and GOES-9 stereoscopic imagery,” J. Appl. Meteorol. Climatol., vol. 37, no. 4, pp. 405–413, 2010.

[12] B. A. Baum et al., “MODIS cloud top property refinements for collection 6,” J. Appl. Meteor. Clim., vol. 51, pp. 1145–1163, 2012.

[13] T. Y. Nakajima et al., “Theoretical basis of the algorithms and early phase results of the GCOM-C (Shikisai) SG1 cloud products,” Prog. Earth Planet. Sci., vol. 6, no. 1, pp. 1–25, 2019.

[14] C. Shi, M. Hashimoto, K. Shiono, and T. Nakajima, “Development of an algorithm to retrieve aerosol optical properties over water using an artificial neural network radiative transfer scheme: First result from GOSAT-2/CAI-2,” IEEE Trans. Geosci. Remote Sens., vol. 59, no. 12, pp. 9861–9872, Dec. 2021.

[15] M. Min et al., “Estimating summertime precipitation from Himawari-8 and global forecast system based on machine learning,” IEEE Trans. Geosci. Remote Sens., vol. 57, no. 5, pp. 2557–2570, May 2019.

[16] M. Min, J. Li, F. Wang, Z. Liu, and W. P. Menzel, “Retrieval of cloud top properties from advanced geostationary satellite imager measurements based on machine learning algorithms,” Remote Sens. Environ., vol. 239, Mar. 2020, Art. no. 111616.

[17] H. Letu et al., “Ice cloud properties from Himawari-8/ACHI next-generation geostationary satellite: Capability of the AHI to monitor the DC cloud generation process,” IEEE Trans. Geosci. Remote Sens., vol. 57, no. 6, pp. 3229–3239, Jun. 2019.

[18] H. Letu et al., “High-resolution retrieval of cloud microphysical properties and surface solar radiation using Himawari-8/ACHI next-generation geostationary satellite,” Remote Sens. Environ., vol. 239, pp. 1–16, Mar. 2020, 111583.

[19] H. Shang et al., “Diurnal haze variations over the north China plain using measurements from Himawari-8/AHI,” Atmos. Environ., vol. 210, pp. 1040–1059, Aug. 2019.

[20] F. Sun, M. Min, D. Qin, F. Wang, and J. Hu, “Refined typhoon geometric center derived from a high spatiotemporal resolution geostationary satellite imaging system,” IEEE Geosci. Remote Sens. Lett., vol. 16, no. 4, pp. 499–503, Apr. 2019.

[21] Z. Liu et al., “Discriminating between clouds and aerosols in the CALIOP version 4.1 data products,” Atmos. Meas. Techn., vol. 12, no. 1, pp. 703–734, Feb. 2019.

[22] M. T. Chahine, “Remote sounding of cloudy atmospheres. I. The single cloud layer,” J. Atmos. Sci., vol. 31, no. 1, pp. 233–243, Jan. 1974.

[23] W. L. Smith and C. M. R. Platt, “Comparison of satellite-deduced cloud heights with indications from radiosonde and ground-based laser measurements,” J. Appl. Meteorol., vol. 17, no. 12, pp. 1796–1802, Dec. 1978.

[24] P. Bühmann and T. Hothorn, “Boosting algorithms: Regularization, prediction and model fitting,” Stat. Sci., vol. 22, no. 4, Nov. 2007.

[25] Z. Peng et al., “Estimation of shortwave solar radiation using the artificial neural network from Himawari-8 satellite imagery over China,” J. Quant. Spectrosc. Radiat. Transf., vol. 240, Jan. 2020, Art. no. 106672.

[26] H. Iwabuchi et al., “Cloud property retrieval from multiband infrared measurements by Himawari-8,” J. Meteorol. Soc. Japan, vol. 96, pp. 27–42, Mar. 2018.

[27] S.-W. Kim, E.-S. Chung, S.-C. Yoon, B.-J. Sohn, and N. Sugimoto, “Intercomparisons of cloud-top and cloud-base heights from ground-based lidar, CloudSat and CALIPSO measurements,” Int. J. Remote Sens., vol. 32, no. 4, pp. 1179–1197, Feb. 2011.

[28] S. S. Chen and R. A. Houze, “Diurnal variation and life-cycle of deep convective systems over the tropical Pacific warm pool,” Quart. J. Roy. Meteorol. Soc., vol. 123, no. 538, pp. 357–388, 1997.

[29] G.-Y. Yang and J. Slingo, “The diurnal cycle in the tropics,” Monthly Weather Rev., vol. 129, no. 4, pp. 784–801, Apr. 2001.
Takashi Y. Nakajima received the D.Sc. degree in Earth and planetary physics from The University of Tokyo, Tokyo, Japan, in 2002. He joined the Japan Aerospace Exploration Agency, Tsukuba, Japan, in 1994, and moved to Tokai University, Tokyo, in 2005, where he is currently a Professor. He was a Visiting Associate Professor with Chiba University, Chiba, Japan, from 2004 to 2007, a Visiting Researcher with the National Institute of Environmental Study, Tsukuba, from 2005 to 2011, a Visiting Scientist with Colorado State University, Fort Collins, CO, USA, from 2008 to 2009, and a part-time Lecturer with Kyushu University, Fukuoka, Japan, in 2013. His research interests include climate change, remote sensing of clouds and aerosols, and theory of light scattering by nonspherical particles and calculating terrestrial renewable energy. Dr. Nakajima is a member of JpGU, JMS, RSSJ, AGU, AMS, and OSA. He received the SICE Award in 1998, the Matsumae Shigeyoshi Award, the Horiuchi Award of JMS, and the Remote Sensing Society of Japan Award in 2011.

Jiancheng Shi (Fellow, IEEE) is currently a Senior Research Scientist and the Director of the Earth Science Division, National Space Science Center, Chinese Academy of Sciences (CAS), Beijing, China. He is also the Director of the State Key Laboratory for Remote Sensing Science, Beijing. He has published more than 500 papers with nearly 12,000 citations. His research interests mainly include remote sensing theory and techniques, remote sensing of cryospheric components, water cycle components, and radiation energy balance, development of new satellite missions, and synergy of remote sensing observations and Earth process models for hydrology and climatic change. Dr. Shi is a fellow of SPIE.

Jun Zhao received the B.S. and M.S. degrees in geography from Northwest Normal University, Lanzhou, China, in 1984 and 1992, respectively, and the Ph.D. degree in pratacultural science from Gansu Agricultural University, Lanzhou, in 2007. He is currently a Professor with Northwest Normal University. His research interests include remote sensing of ecological environment and application of GIS. Dr. Zhao is a member of GSC. He received the Scientific and Technological Progress Award from the Chinese Academy of Sciences (CAS) in 1993.

Jian Xu (Member, IEEE) received the B.E. degree in geographic information systems from Hohai University, Nanjing, Jiangsu, China, in 2004, and the M.S. degree in Earth-oriented space science and technology and the Ph.D. degree in atmospheric remote sensing from Technische Universität München, Munich, Germany, in 2009 and 2015, respectively. From 2010 to 2021, he was with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, Germany. He was involved in the development of atmospheric retrieval algorithms for ESA’s Sentinel-5P and Sentinel-4 satellite missions. He is currently with the National Space Science Center, Chinese Academy of Sciences, Beijing, China. His research interests include remote sensing of planetary atmosphere, radiative transfer modeling, and ill-posed inverse problems.

Husi Letu received the B.S. and M.S. degrees in geography from Inner Mongolia Normal University, Hohhot, China, in 1999 and 2002, respectively, and the Ph.D. degree in geosciences and remote sensing from the Center for Environmental Remote Sensing, Chiba University, Chiba, Japan, in 2010. He was with the Research and Information Center, Tokai University, Tokyo, Japan. He is involved in the algorithm development and validation of the ice cloud product for JAXA’s GCOM-C and Himawari-8 satellite missions. He is currently a Professor with the State Key Laboratory of the Science and Remote Sensing, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. His research interests include atmospheric radiative transfer simulation, light scattering calculation, cloud remote sensing, and ice cloud property retrievals.