atmosphere

MDPI

Article

Estimation of Surface Downward Longwave Radiation and Cloud Base Height Based on Infrared Multichannel Data of Himawari-8

Jiangqi Shao [1,2], Husi Letu [1,2,*], Xu Ri [1,3], Gegen Tana [4], Tianxing Wang [5] and Huazhe Shang [1]

[1] State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China
[2] University of Chinese Academy of Sciences, Beijing 100049, China
[3] College of Geography and Environmental Science, Northwest Normal University, Lanzhou 730070, China
[4] National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China
[5] School of Geospatial Engineering and Science, Sun Yat-sen University, Zhuhai 519082, China
* Correspondence: husiletuw@hotmail.com

Abstract: Surface downward longwave radiation (SDLR) is significant with regard to surface energy budgets and climate research. The uncertainty of cloud base height (CBH) retrieval by remote sensing induces the vast majority of SDLR estimation errors under cloudy conditions; reliable CBH observation and estimation are crucial for determining the cloud radiative effect. This study presents a CBH retrieval methodology built from 10 thermal spectral data from Himawari-8 (H-8) observations, utilizing the random forest (RF) algorithm to fully account for each band's contribution to CBH. The algorithm utilizes only infrared band data, making it possible to obtain CBH 24 h a day. Considering some factors that can significantly affect the CBH estimation, RF models are trained for different clouds using inputs from multiple H-8 channels together with geolocation information to target CBH derived from CloudSat/CALIPSO combined measurements. The validation results reveal that the new methodology performs well, with a root-mean-square error (RMSE) of only 1.17 km for all clouds. To evaluate the effect of CBH on SDLR estimation, an all-sky SDLR estimation algorithm based on previous CBH predictions is proposed. The new SDLR product not only has a resolution that is noticeably higher than that of benchmark products of the SDLR, such as the Clouds and the Earth's Radiant Energy System (CERES) and the next-generation reanalysis (ERA5) of the European Centre for Medium-Range Weather Forecasts (ECMWF), but it also has greater accuracy, with an RMSE of 21.8 W m^{-2} for hourly surface downward longwave irradiance (SDLI).

Keywords: cloud base height; surface downward longwave radiation; Himawari-8; machine learning; surface energy budget

Citation: Shao, J.; Letu, H.; Ri, X.; Tana, G.; Wang, T.; Shang, H. Estimation of Surface Downward Longwave Radiation and Cloud Base Height Based on Infrared Multichannel Data of Himawari-8. *Atmosphere* **2023**, *14*, 493. https://doi.org/10.3390/atmos14030493

Academic Editor: Filomena Romano

Received: 10 February 2023
Revised: 27 February 2023
Accepted: 28 February 2023
Published: 2 March 2023

1. Introduction

Surface downward longwave radiation (SDLR) is a crucial component of the surface energy budget and one of the significant drivers of climate change, as well as a critical forcing parameter in many hydrological and land surface models [1–4]. Therefore, accurate SDLR retrieval is essential for understanding Earth's energy cycle and in climate research [5,6].

Remote sensing significantly benefits SDLR estimates due to its global/regional coverage, regular monitoring, and a high spatial resolution as compared to ground-based observations and reanalysis data [7,8]. SDLR estimation algorithms based on remote sensing have been created for clear-sky and cloudy-sky situations. The clear-sky SDLR are frequently derived from air temperature and vapor pressure or relative humidity in much of the literature [9–11]. However, these parameters are difficult to obtain from satellite signals. Wang et al. [12] demonstrate the availability of the general parameterizations that rely on

the land surface temperature (LST) and the precipitable water vapor (PWV) in retrieving clear-sky SDLR; both parameters can be easily estimated from space [13–19]. Contrary to clear-sky data, it is challenging and complex to retrieve cloudy-sky SDLR from space. The thermal contribution for cloudy-sky SDLR is divided into two components: besides the thermal radiation from the atmosphere under the cloud layers, the radiation from the cloud layers is another important factor that cannot be disregarded. Numerous research studies have shown that SDLR estimates of cloudy skies are controlled by the temperature and height of the cloud base [5,20–22]. Moreover, the uncertainty of cloud base height (CBH) retrieval from the satellites induces a large number of SDLR estimation errors under cloudy conditions [23,24]. Some researchers often employ surrogated variables, such as cloud fraction [25,26], ice water path (IWP) or liquid water path (LWP) [27,28], and cloud top temperature (CTT) [12] to quantify the cloud contribution, whereas few studies have used CBH directly. This practice will inevitably cause some additional errors in SDLR estimations [7]. The accurate acquisition of CBH is the key to estimating SDLR [12].

Effective methods for retrieving CBH include ground-based observations, polar-orbiting observations, and geostationary satellite observations. Ground-based observations often provide precise CBH, but these are restricted to a small region around the observation site [29,30]. Polar-orbiting observations can be active or passive, among which active observation satellites carry radar and lidar (CPR/CloudSat and CALIOP/CALIPSO) to explore the cloud profile globally [31]. Still, narrow scanning strips limit the observation coverage. Passive polar-orbiting observations, such as NPP/VIIRS, Aqua/MODIS, and Terra/MODIS, may yield CBH estimates with extensive horizontal coverage [32] and can measure targets with high spatial and spectral precision, but cannot be utilized to identify stationary areas continually due to the polar orbit. Recently, new geostationary satellites with much better spatial, spectral, and temporal resolutions, such as FY-4, Himawari-8, and GOES-R, have opened up new possibilities for accessing the CBH [33–35]. Using upstream cloud products from the Advanced Himawari Imager (AHI) on Himawari-8 and the Advanced Geostationary Radiation Imager (AGRI) on Fengyun-4A (FY-4A), Tang et al. [36,37] recently utilized a random forest algorithm to retrieve CBH. Additionally, the Advanced Baseline Imager (ABI) level-1b data from the Geostationary Operational Environmental Satellite-R Series (GOES-16), together with the fifth-generation reanalysis (ERA5) data from the European Centre for Medium-Range Weather Forecasts (ECMWF), were used by Lin et al. [38] to estimate CBH. However, because these all rely on visible/near-infrared data, they could not estimate CBH at night, which limits their applications during nighttime. Due to the fact that the majority of presently used algorithms are based on cloud property products or reanalysis data, the errors from upstream retrieval will accumulate in CBH estimation [39]. In addition, most of those algorithms target daytime situations, and few research studies have committed to deriving the all-day CBH.

In recent years, machine learning (ML) methods have been proven effective for many nonlinear problems in remote sensing and meteorology, such as the estimates of CTH or precipitation [40–42]. The correlation between CTH and brightness temperatures (BTs) from infrared (IR) bands is apparent physically. However, CBH is complicated by the strong absorption of the cloud in IR bands. ML methods are beneficial for solving such a problem. The random forest (RF) technique of ML methods is used in this study. RF constructs multiple learners to complete learning tasks, which is widely applied for solving nonlinear fitting problems [36,43,44]. It is suitable for dealing with the complex relationship between nonlinear inputs and prediction, and can significantly improve learning ability and generalization in both sampling and features.

In this study, we present an ML approach for CBH retrieval using Himawari-8 level-1b infrared data based on a RF algorithm, which was developed for all-day concentration on single-layer clouds. Moreover, we try to see if the accuracy of CBH retrieval can be improved and further applied to the estimation of SDLR. The Japan Meteorological Agency (JMA) launched Himawari-8, the country's first new-generation geostationary meteorological satellite, on 7 October 2014, and it started operating on 7 July 2015. The Advanced

Himawari Imager (AHI) aboard Himawari-8 offers greater capabilities and better performance than earlier Japanese geostationary meteorological satellites due to considerable increases in spatial, temporal, and spectral resolutions [33]. Random forest (RF) has been used extensively in remote sensing research in recent years [40,45] and is well suited to handle the complex relations between nonlinear input and prediction. Cloud absorption and scattering effects on electromagnetic waves make it more challenging to obtain CBH information directly from satellite observations. An advanced RF algorithm can help to solve this problem. The advantages of all-day infrared observations and the geostationary satellite's ability to continuously observe a large area at all times provided the CBH estimation obtained in this study with the benefits of a high spatial and temporal resolution, extensive spatial coverage, and all-day observation ability. Further SDLR estimations based on this CBH retrieval will have the same characteristics. Compared to previous studies, this research was dedicated to obtaining a reliable all-day CBH retrieval method with high spatiotemporal resolution and less reliance on upstream retrieval. Further, the CBH estimations were further applied directly to the SDLR retrieval to improve the accuracy of SDLR estimations. The data used in this research and the CBH and SDLR estimation methods are presented in Section 2. The estimated CBH and SDLR results are presented in Section 3. Finally, conclusions are given in Section 4.

2. Materials and Methods

For the estimation of CBH and SDLR of the new-generation geostationary satellites, the data used in this study include Himawari-8 (H-8) level-1B (L1B) infrared observation data and level-2 cloud products (L2 CLP), CBH data from the combined measurements of CloudSat and CALIPSO; LST and PWV products from the ERA5 reanalysis; SDLR data from in situ radiation stations of four networks; and auxiliary data (latitude and elevation). The H8 L2 CLP used in this study include both cloud type product and CTT data. The cloud type product was utilized to classify the training data for CBH retrieval to build RF models. Furthermore, CTT data was used as a surrogate parameter for CBH, and the resultant accuracy increase from directly applying CBH to SDLR estimation was evaluated. The following subsections present detailed data descriptions.

2.1. Input Data

In the CBH retrieval study, the training data of the RF model were first classified using a cloud type product, the H-8 L2 CLP, and then the CBH was estimated from H-8 L1B infrared data. The H-8 satellite started its operation in 2015 in geosynchronous orbit at approximately 140° E. The Advanced Himawari Imager (AHI), which offers observations with 16 different channels (Table 1), with center wavelengths ranging from 0.47 to 13.3 µm and spatial resolutions ranging from 0.5 to 2 km, is the main instrument onboard H-8. At 10 min interval, the AHI scans the whole disk (a satellite's view of the planet). To achieve 24 h continuous estimation, 10 infrared bands out of the 16 channels were used as input data in this study. The most recent standard H-8 L1B data are available on www.eorc.jaxa.jp/ptree/index.html (accessed on 15 February 2023) in HSD and netCDF formats. The netCDF format (60° S–60° N, 80° E–160° W) is rearranged to 0.02° and 0.05° with significantly poorer coverage due to margin cutting, whereas the HSD format keeps the previous resolution. Since the H-8 L2 CLP is only available for daytime observation, we used the CARE (Cloud, Atmospheric Radiation and Renewal Energy Application) version 1.0 cloud type product obtained from H-8/AHI data for our study [44]; the 24 h cloud top temperature product for the surrogated parameter algorithm used in the SDLR estimation is also provided by the CARE.

Table 1. Bands of H-8/AHI.

	Channel	Band (μm)	Spatial Resolution (km)	Use or Not
Visible	1	0.46	1	
	2	0.51	1	
	3	0.64	0.5	
Near−Infrared	4	0.86	1	
	5	1.6	2	
	6	2.3	2	
Infrared	7	3.9	2	√
	8	6.2	2	√
	9	7.0	2	√
	10	7.3	2	√
	11	8.6	2	√
	12	9.6	2	√
	13	10.4	2	√
	14	11.2	2	√
	15	12.3	2	√
	16	13.3	2	√

We used the LST and PWV products from the ERA5 reanalysis to derive clear-sky SDLR. The CBH estimated in this study was added to the inputs with LST and PWV to derive cloudy-sky SDLR, and the surrogated parameter algorithm employed for the comparison of results used the CARE-CTT described previously. The LST and PWV data were provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) next-generation reanalysis (ERA5) (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview, accessed on 15 February 2023), which have a 1 h temporal resolution and a 25 km spatial resolution. The spatial resolution and temporal resolution of our CBH estimate and CARE-CTT were 5 km and 10 min, respectively. We used bilinear interpolation to resample the spatial resolution of the ERA5 LST and PWV data to 5 km to retain consistency in the spatial and temporal resolution of the inputs in the SDLR estimation [46]. The CBH estimation and CARE-CTT were averaged on an hourly scale to obtain the hourly products.

2.2. Validation Data

We employed the combined measurements from CloudSat and CALIPSO, which are members of the A-Train constellation and have equator-crossing periods in the afternoon (1330, local time), to determine the intended CBH output. Except for precipitating clouds and optically thin cirrus, CloudSat's 94 GHz cloud profiling radar (CPR) can identify vertical cloud boundaries [47]. The Cloud-Aerosol LIDAR with Orthogonal Polarization (CALIOP) aboard CALIPSO has a greater vertical resolution and is more sensitive to thin cirrus [48]. The most well-known cloud profile product worldwide is the combined 2B-GEOPROF-LIDAR (geometric profile, GEOPROF; light detection and ranging, LIDAR) product of CPR and CALIOP, which offers more specific information on cloud vertical structure [49]. The 2B-GEOPROF-LIDAR product has a vertical resolution of 60 m and a horizontal spatial resolution of 1.1 km. It contains information on cloud boundaries (layer top and layer base) for a maximum of five layers. Our spatial matching method exclusively utilizes H-8 pixels closest to CloudSat/CALIPSO profiles in terms of position. The objective CBH in the RF training procedure is the lowest height of the cloud layer nearest to the ground.

A total of thirty-five in situ radiation stations from four networks situated in H-8 full-disk zones were utilized to verify the predicted SDLR in this study: thirteen sites from the Australian Governments Bureau of Meteorology (BOM), four sites from the Nam Co Watershed (NMC), seven sites from the Global Moored Buoy Array (GTMBA), and eleven sites from the Baseline Surface Radiation Network (BSRN) [50,51]. The World Climate Research Programme created the BSRN, which seeks to provide high-accuracy

validation data at 1 min intervals. For the worldwide SDLR, the measurement uncertainties are about 10 W/m^{-2} [52]. The Second Tibetan Plateau Scientific Expedition and Research Program (STEP) initiative constructed the NMC, a ground observation network, and the accompanying measures include surface downward shortwave radiation (SDSR), SDLR, and net radiation. These four stations are also all located at elevations of more than 4.5 km. Given that certain BOM sites are also members of the BSRN, the properties of BOM data should be equivalent to those of BSRN. Regarding the GTMBA, daily and weekly programs are used to carry out data quality control formally [53]. High-frequency SDLR observations of one to three minutes are offered by these stations. These full-disk regions of H-8 sites from various networks serve as examples of SDLR validation. When the files had quality flags, all of the high-quality in situ data were included in the research. Low-quality data were excluded after manually checking non-quality highlighted data.

2.3. Algorithm Description

The flowchart in Figure 1 was followed for the construction of the algorithms that produce the CBH and all-sky SDLR products. First, RF models for CARE cloud types are generated separately by training using the 10 infrared channel data of H-8 L1, the digital elevation model (DEM), latitude (Lat), and the CBH of 2B-GEOPROF-LIDAR. Second, RF models for cloudy-sky and clear-sky SDLR are trained using the CBH estimation, ERA5 LST and PWV products, and ground-based SDLR observations. Third, validation of the estimated CBH and SDLR is performed using data from active observation satellites and in situ radiation stations.

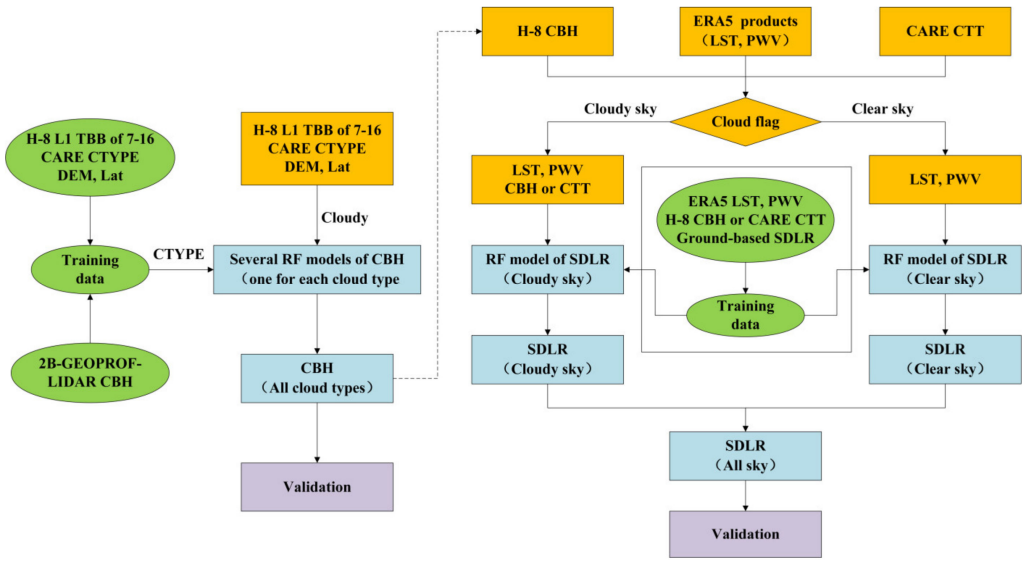

Figure 1. Flowchart for the estimation of CBH and the all-sky SDLR based on H-8.

2.4. RF-CBH Model Development

The RF algorithm [54] is a machine learning technique that constructs multiple learners to complete learning tasks. It is suitable for dealing with the complex relationship between nonlinear inputs and prediction, and can significantly improve learning ability and generalization in both sampling and features. Although the RF algorithm is attempted here, readers may use other machine learning techniques instead, such as support vector machines and extremely randomized trees. We fully considered our choice of inputs for the RF model. To retrieve the all-day CBH, we used all of the 10 infrared channel data of H-8 L1 as input parameters, which carry, more or less, cloud base information. The scatter plot of the H-8 infrared band vs. cloud geometric thickness (Figure 2) also confirms a correlation between all infrared bands and cloud geometric thickness. The cloud types and geographical characteristics (e.g., latitude, altitude, etc.) have been shown to affect CBH estimates [49,55]. To demonstrate the impact of several sets of input parameters on

the CBH, we built a sensitivity analysis of those parameters (Table 2). We observed that, when compared to all other groups, the performance of the RF model with all 10 input IR bands, DEM, latitude, and cloud types was the best.

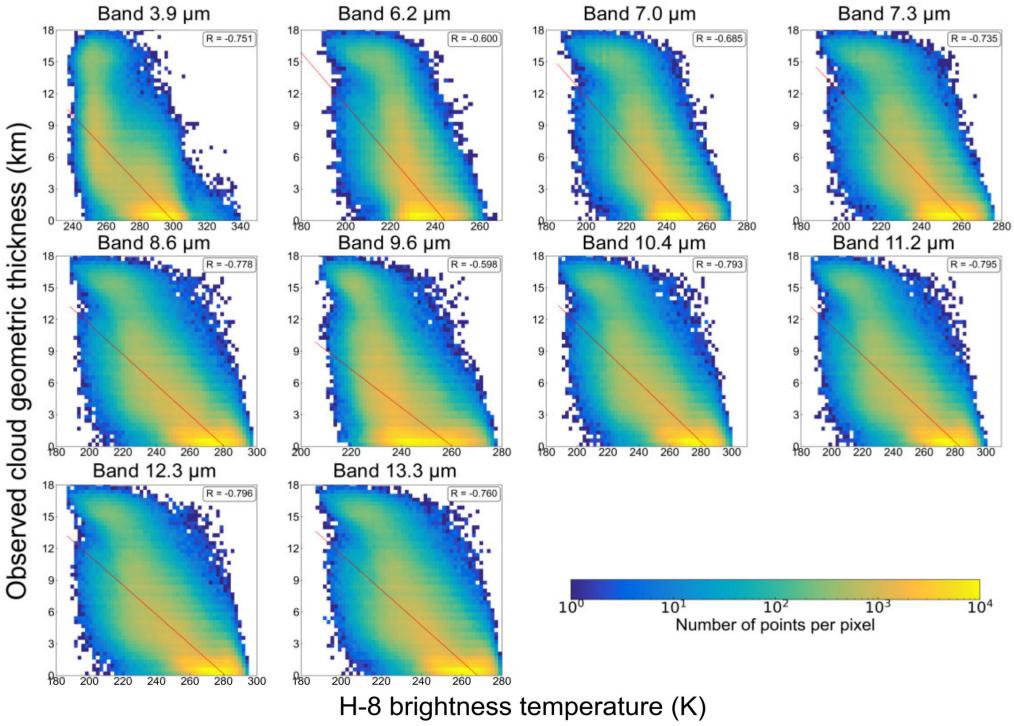

Figure 2. Scatter plots of H-8 infrared band vs. cloud geometric thickness.

Table 2. Different input groups and their validation of CBH retrieval.

Different Groups (Channel Number)	Cloud Base Height (km)		
	RMSE	MBE	R
All IR bands, DEM, Lat, cloud types	1.17	−0.02	0.92
All IR bands, DEM, Lat	1.53	0.08	0.82
All IR bands, DEM	1.61	0.17	0.81
All IR bands	1.63	0.21	0.80
8,9,10,11,12,13,14,15,16	1.67	0.19	0.80
9,10,11,12,13,14,15,16	1.67	0.18	0.79
10,11,12,13,14,15,16	1.69	0.16	0.79
11,12,13,14,15,16	1.77	0.16	0.77
12,13,14,15,16	1.82	0.17	0.75
13,14,15,16	1.88	0.13	0.73
14,15,16	2.07	0.18	0.67
15,16	2.96	0.29	0.35

The infrared channels of H-8 may support the model's ability to predict the CBH more accurately, and they should all be considered in relation to the retrieval technique, according to Table 2. Therefore, we extracted training data according to cloud types and produced RF models suitable for different cloud types while inputting latitude and altitude data to the models to optimize the training effect.

The construction of the RF-CBH models starts with the preparation of the training data. In this work, we used brightness temperature data in bands 7 to 16 of H-8 L1, CARE cloud types, and the CBH obtained from 2B-GEOPROF-LIDAR, DEM, and latitude data that covered two whole years, from 2016 to 2017, as the training dataset in order to reduce the effects of seasonal variation. Although the spatiotemporal matching strategies for all data have been described before, we would like to state some key points again. The CARE

product and the H-8 data have the same spatiotemporal resolution. The only H-8 pixels that are spatially matched are those whose positions are closest to the CloudSat/CALIPSO profiles, and the time intervals between the matches are limited to 5 min. All clouds are regarded as single-layered to streamline the CBH retrieval procedure. The goal CBH in the RF training process is the lowest height of the CloudSat/CALIPSO cloud layer closest to the ground.

Finally, we obtained 5,508,920 matchups divided according to different cloud types. We allocated 20% of these filtered samples as a test dataset for algorithm validation and 80% as a training dataset for RF model training. During the training phase, we set the number of trees to 10, which incrementally increased by 10 until 500 was reached, and all other settings were left at their default values. Finally, the number of trees that performed optimally on the test dataset was selected. The optimal number of trees for all RF-CBH models ranged from 220 to 290, with the minimum number of cumulus (Cu) being 220 and the maximum number of deep convective clouds (Dc) and altocumulus (Ac) being 290.

2.5. RF-SDLR Model Development

The SDLR estimate methods were created for two different sky conditions: clear skies and cloudy skies. With the use of LST and PWV as inputs, we derived clear-sky SDLR; with CBH, LST, and PWV as inputs, we derived cloudy-sky SDLR. The surrogated parameter algorithm of Wang et al. [12], which is based on CTT, LST, and PWV to derive cloudy-sky SDLR, was also attempted in this study.

In the preparation of the training data, we used the LST and PWV of the ERA5 reanalysis, the CBH estimation derived from our study, and the SDLR of in situ radiation stations that was obtained for a whole year in 2017. The spatiotemporal matching strategies for ERA5 reanalysis and CBH are described in Section 2.1. The geographical and temporal resolutions of the inputs to RF models were 5 km and 1 h, respectively. Radiation stations offer high-frequency observations lasting one to three minutes, which were resampled to 1 h. The CBH pixels whose locations were nearest the radiation stations were matched spatially.

Finally, we obtained 206,136 matchups, comprising 125,720 records for clear skies and 80,416 records for cloudy skies. We allocated 20% of these filtered samples as a test dataset for algorithm validation and 80% as a training dataset for RF model training. The RF-SDLR model training method is the same as that for RF-CBH: set the number of trees to 10, incrementally increase by 10 until 500 is reached, and leave all other settings at their default values. Finally, the number of trees that performed optimally on the test dataset was selected. In the end, the optimal number of trees for the clear-sky SDLR model was 290, and for the cloudy-sky model was 260.

3. Results and Discussion

3.1. Spatial Features of the Cloud Geometric Properties and SDLR

Figure 3 shows the spatial features of the H-8 full-disk cloud geometric properties and SDLR. Both CTH and CBH increase as the latitude decreases, and clouds with a high CBH are mainly generated near the equator. The subfigure of CGT shows that thick clouds also often form close to the equator, and most of them are deep convective clouds with significant convection phenomena. High latitudes will have more cloud coverage than low latitudes. The spatial distribution of 35 radiation stations from four networks is shown in Figure 3f.

Because clouds have the ability to block longwave radiation transmitted to space from the surface and lower atmosphere, as well as emit longwave radiation to the surface, ultimately warming the Earth–atmosphere system, cloud cover tends to raise the SDLR.

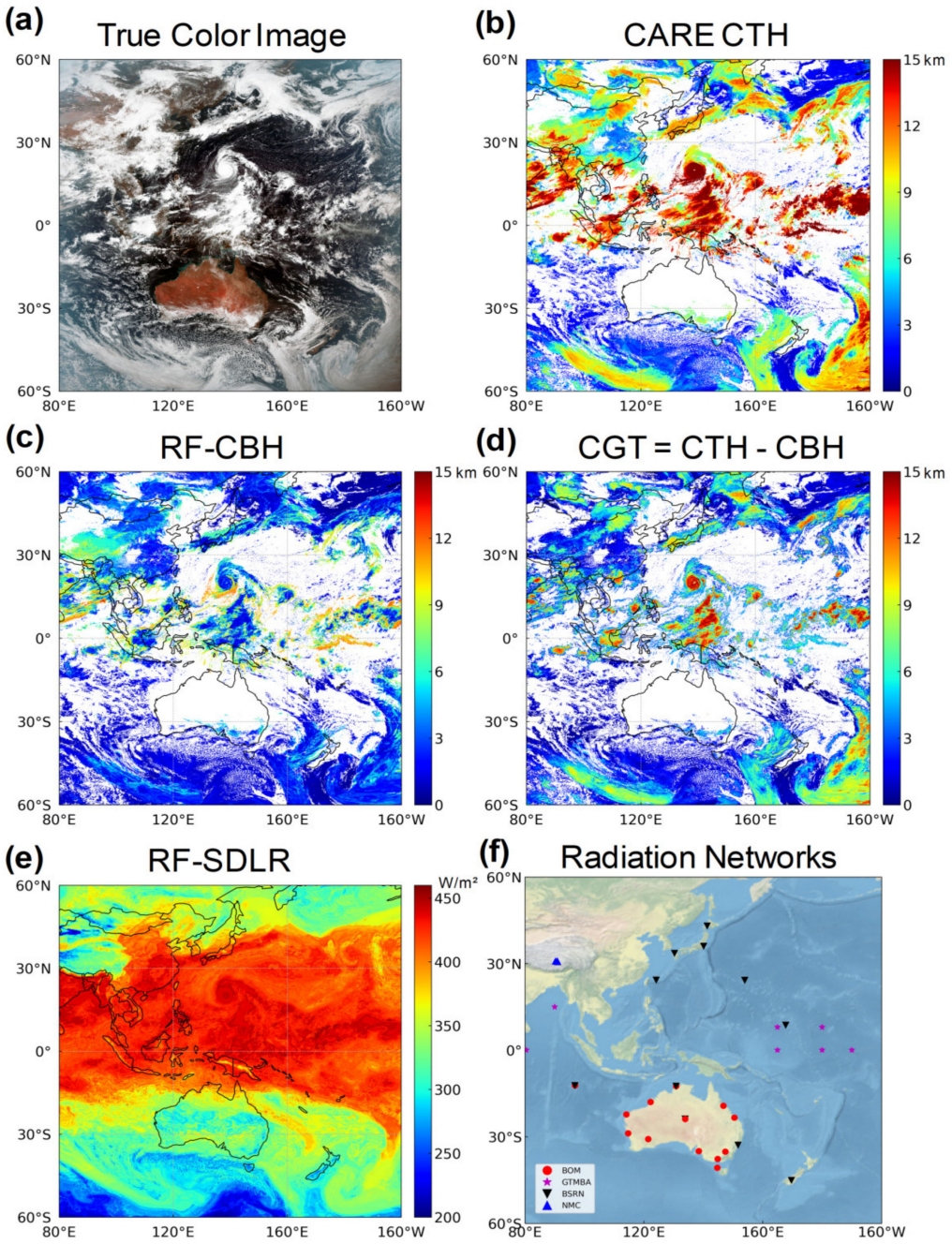

Figure 3. Spatial features of the H-8 full-disk cloud geometric properties and SDLR on September 1, 2018: (**a**) bands 3, 2, and 1 for RGB of the H-8 true color image; (**b**) CARE cloud top height product; (**c**) H-8 cloud base height derived using the RF models; (**d**) cloud geometry thickness derived from the difference between CTH and CBH; (**e**) H-8 SDLR derived using the RF models; and (**f**) spatial distribution of radiation stations from four networks (BOM, GTMBA, BSRN, and NMC) used to validate the performance of the estimated SDLR. Different symbols stand for various networks.

3.2. Validation of the CBH Estimation

We chose a suitable case to show the performance of the new CBH product. Super Typhoon Meranti approached the southeastern coast of China on 13 September 2016. In particular, we focused on three areas, as shown in Figure 4: the super typhoon (domain B), the high clouds over the Chinese Bohai Rim (domain C), and the low clouds over the

southern Australian coast (domain D). This example was selected because it covers a variety of clouds, including deep convective, low, and high clouds. This enabled us to develop a better comprehension of the RF-CBH performance for various cloud types.

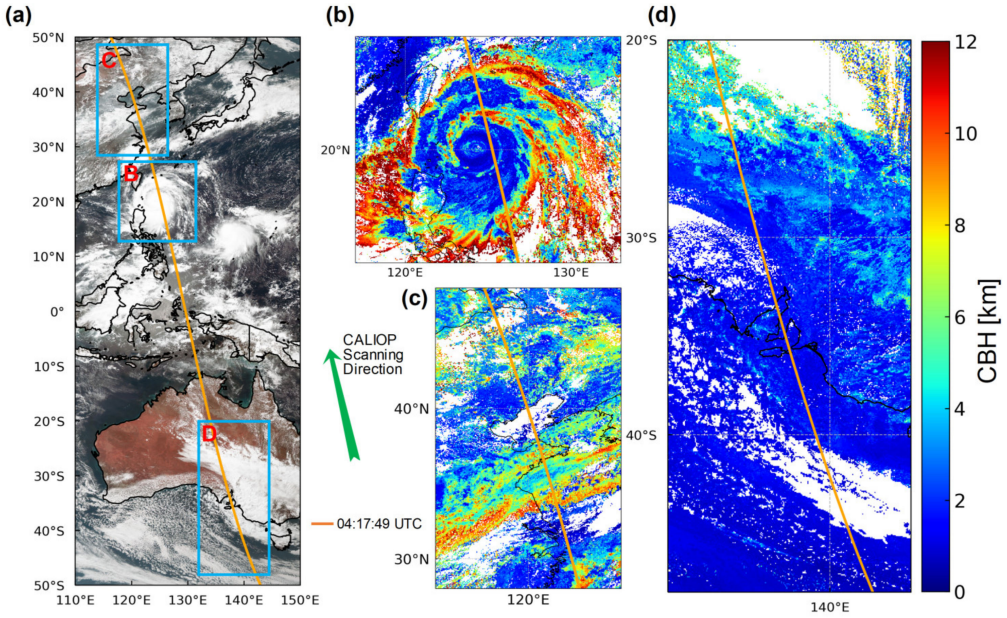

Figure 4. Case study of Super Typhoon Meranti and low/high clouds on 13 September 2016: (**a**) AHI Natural True Color images starting at 05:00:41 UTC, with the start time and direction indicated in the right-side legends; the orange lines represent the CALIOP scanning tracks. The CBH calculated by the RF-CBH corresponding to areas B, C, and D in panel (**a**) is shown in panels (**b**–**d**).

Corresponding to Figure 4, the cross sections along CALIOP scanning tracks for Super Typhoon Meranti and low/high cloud case study were shown in Figure 5. For all cloud types, the CBH estimations (black upward triangles) and the CALIOP observations (red squares) correspond pretty well, as shown in the cross-section view in Figure 5a–c. The majority of CBH estimations, particularly for low clouds, have uncertainties of less than 1 km. The CBH estimates from RF-CBH are highly accurate. For the CBH of Typhoon Meranti, which is the most interesting to us, it can be seen that the typhoon's center has a high CTH and very low CBH in CALIOP observations. Even so, the RF-CBH works well in estimating CBH. The RF-CBH accurately reveals that the closer to the edge of the typhoon, the higher the CBH. It should be noted that the estimation errors increase significantly in multilayer clouds (green downward triangles), and the CBH estimates will be higher than the observations. In this study, all clouds are considered single-layered, and the influence of multilayer clouds on CBH has not been studied. Therefore, multilayer clouds need to be considered in the future to improve the accuracy of the RF-CBH. In this example study, all cloud types exhibit strong CBH estimation using the RF-CBH with respect to accuracy. Although estimating CBH based on geostationary satellites is challenging, the RF-CBH successfully obtains those CBH estimates.

We chose the CBH of 2B-GEOPROF-LIDAR for two whole years, from 2018 to 2019, to compare and statistically assess the performance of new CBH products. The results for different cloud types are shown in Figure 6, where the RMSE of stratus (St) is the smallest at 0.47 km, and that of altocumulus (Ac) is the largest at 1.56 km. The general RMSE of all cloud types is 1.17 km, with a mean bias error (MBE) of −0.02 km and a correlation coefficient (R) of 0.92. The performance of RF-CBH for different cloud types is shown in Table 3. The comparison demonstrates that the CBH estimations produced from the H-8 achieve good performance.

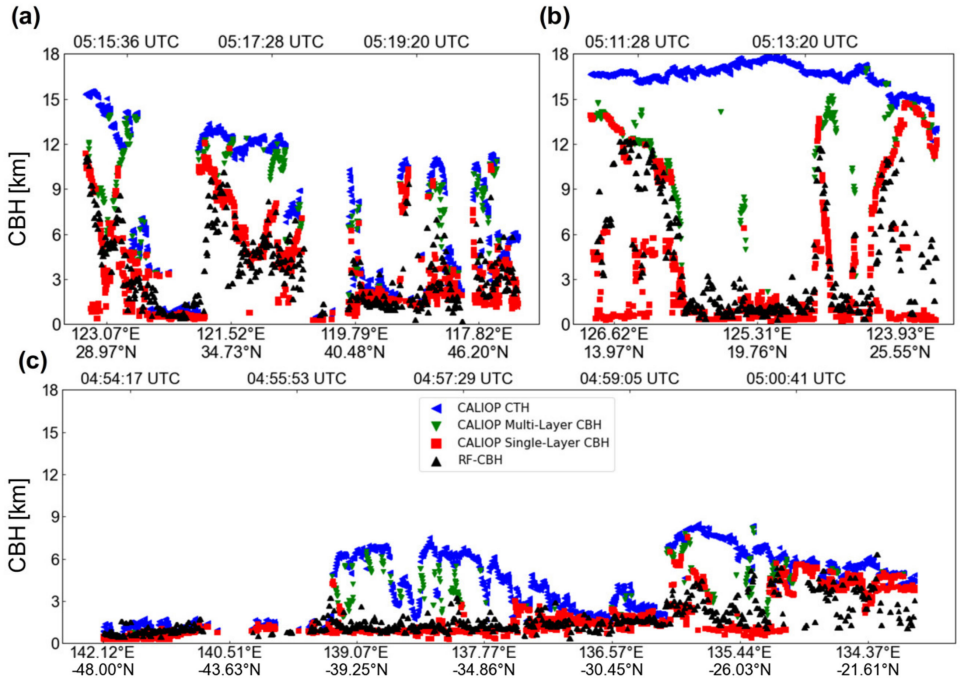

Figure 5. The cross sections along CALIOP scanning tracks for Super Typhoon Meranti and low/high cloud case study: (**a**–**c**) The cross sections refer to the orange dot lines in Figure 4b–d. The black upward triangle markers represent CBH estimates, and the red square, green downward triangle, and blue leftward triangle markers are the CBH of the lowest clouds, the CBH of the topmost clouds, and the CTH of the topmost clouds from CALIOP observations, respectively. Note that the lowest cloud can be either single- or multilayered, and the green downward triangle markers only exist in multilayer clouds.

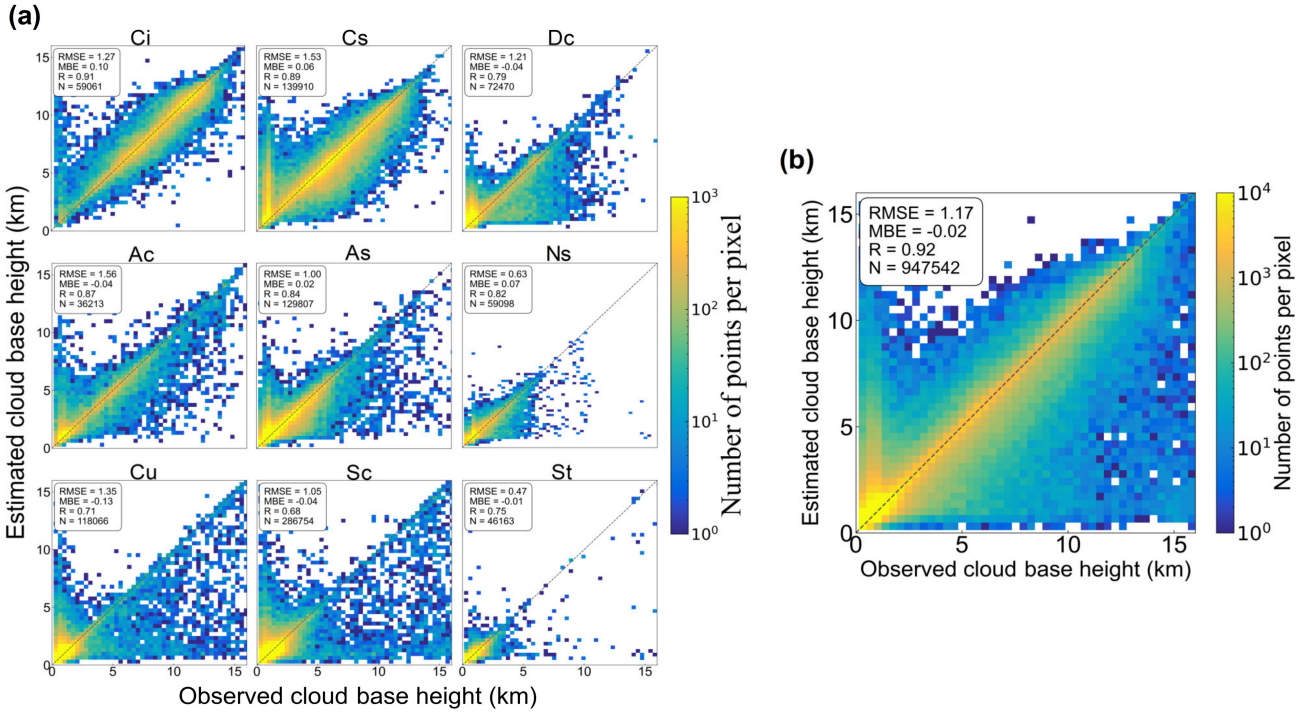

Figure 6. RF-CBH performance: (**a**) the performance of RF-CBH for different cloud types. Ci: cirrus, Cs: cirrostratus, Dc: deep convection, Ac: altocumulus, As: altostratus, Ns: nimbostratus, Cu: cumulus, Sc: stratocumulus, St: stratus; (**b**) the performance of RF-CBH for all cloud types.

Table 3. The performance of RF-CBH for different cloud types.

Cloud Type	RMSE (km)	MBE (km)	R	N
Ci	1.27	0.10	0.91	59,061
Cs	1.53	0.06	0.89	139,910
Dc	1.21	−0.04	0.79	72,470
Ac	1.56	−0.04	0.87	36,213
As	1.00	0.02	0.84	129,807
Ns	0.63	0.07	0.82	59,098
Cu	1.35	−0.13	0.71	118,066
Sc	1.05	−0.04	0.68	286,754
St	0.47	−0.01	0.75	46,163
All types	1.17	−0.02	0.92	947,542

Regarding the comparison of CBH results, a relatively large deviation from the 1:1 line occurs with some cloud types, especially for the overestimation of high clouds such as Cs, Ac, and As and the underestimation of low clouds such as Cu and Sc. Overestimation is mainly concentrated in low observations due to neglecting the influence of precipitation clouds on CBH estimation. The radar echo signal of precipitation clouds is powerful until it reaches the surface, and thus, it retrieves a very small CBH in observations. As for the RF-CBH, the estimations of CBH are only determined by the model inputs, and all clouds are considered non-precipitation clouds, leading to overestimation compared to the results obtained through observation. Underestimation is due to the misclassification of cloud types. Since the RF-CBH is based on separate models for different cloud types, there will be significant differences between the models. When a high cloud is misclassified as a low cloud, the RF-CBH estimate will be underestimated. The errors caused by misclassification should exist for all cloud types, i.e., low clouds misclassified as high clouds will also be overestimated by the RF-CBH, and thus, the Cu and Sc clouds are misclassified more easily in this study. We believe that the CBH estimation accuracy will be improved if precipitation clouds are considered and better cloud type products are used.

The errors of the RF-CBH may also come from two other aspects. One is the errors carried out by 2B-GEOPROF-LIDAR itself. It is still challenging to obtain the CBH for clouds with strong convective effects and a high optical thickness by depending only on satellite data, such as for deep convective clouds, even though the combination of radar and lidar can identify the cloud profile effectively and produce a pretty accurate CBH. The other is the errors of the RF model training. There are still errors in the CBH estimates of the RF-CBH models compared with the results obtained with active observations, and the optimization of the ML algorithm chosen in this study can improve the estimation accuracy.

3.3. Validation of the SDLR Estimation

Figure 7 shows the spatial features of the H-8 full-disk SDLR and RF inputs on 1 September 2018. The spatial resolution of ERA5 SDLR is coarse at 25 km, and the product has noticeable regular stripes, which is unreasonable. The RF-SDLR estimates closely match the ERA5 SDLR spatially with a 5 km spatial resolution and show more details in the graph. Comparing the differences between these two SDLR products under cloudy-sky and clear-sky conditions, as shown in Figure 7c, and the spatial features of CBH, LST, and PWV, as shown in Figure 7d–f, respectively, can lead us to some interesting conclusions. Firstly, the regions where the SDLR difference is close to 0 tend to have a high PWV, and the difference increases when the PWV decreases. Second, the RF-SDLR estimations are higher in the Northern Hemisphere than those of ERA5 and lower in the Southern Hemisphere under clear-sky conditions. The differing ground-based observation networks used in the Northern and Southern Hemispheres may be to blame for this performance. The main networks used in the Northern Hemisphere are the BSRN and NMC, where NMC is a network established to observe the Tibetan Plateau with reliable SDLR observations in the region. In a further study of the comparison between the ERA5 SDLR and in situ radiation observations, as shown in Table 4, the ERA5 SDLR is

seriously underestimated compared with NMC observations. The employment of the NMC observations can correct the underestimation of SDLR estimations in the plateau. The main network used in the Southern Hemisphere is the BOM. The RF-SDLR generated by the BOM, which the Australian government provides, should reasonably describe the truly SDLR in the Southern Hemisphere. Finally, for cloudy-sky conditions, high values of SDLR differences tend to include a high CBH and low PWV, and low difference values tend to include a low CBH and low PWV, which is related to the radiation effects of different CBHs. According to Stephens et al. [23], the underlying water vapor that leads to the highest cloud sensitivity for the SDLR happens in the colder, drier parts of the Earth, and dramatically modulates the SDLR. On the contrary, when PWV is high, SDLR is insensitive to clouds. A high CBH results in a lower cloud base temperature, substantial longwave radiation absorption and emission from the cloud base to the surface, and a significant warming effect of clouds, which increases SDLR. In contrast, when the CBH is low, longwave radiation from the surface is less blocked; thus, more longwave radiation is sent into space, and the SDLR is reduced.

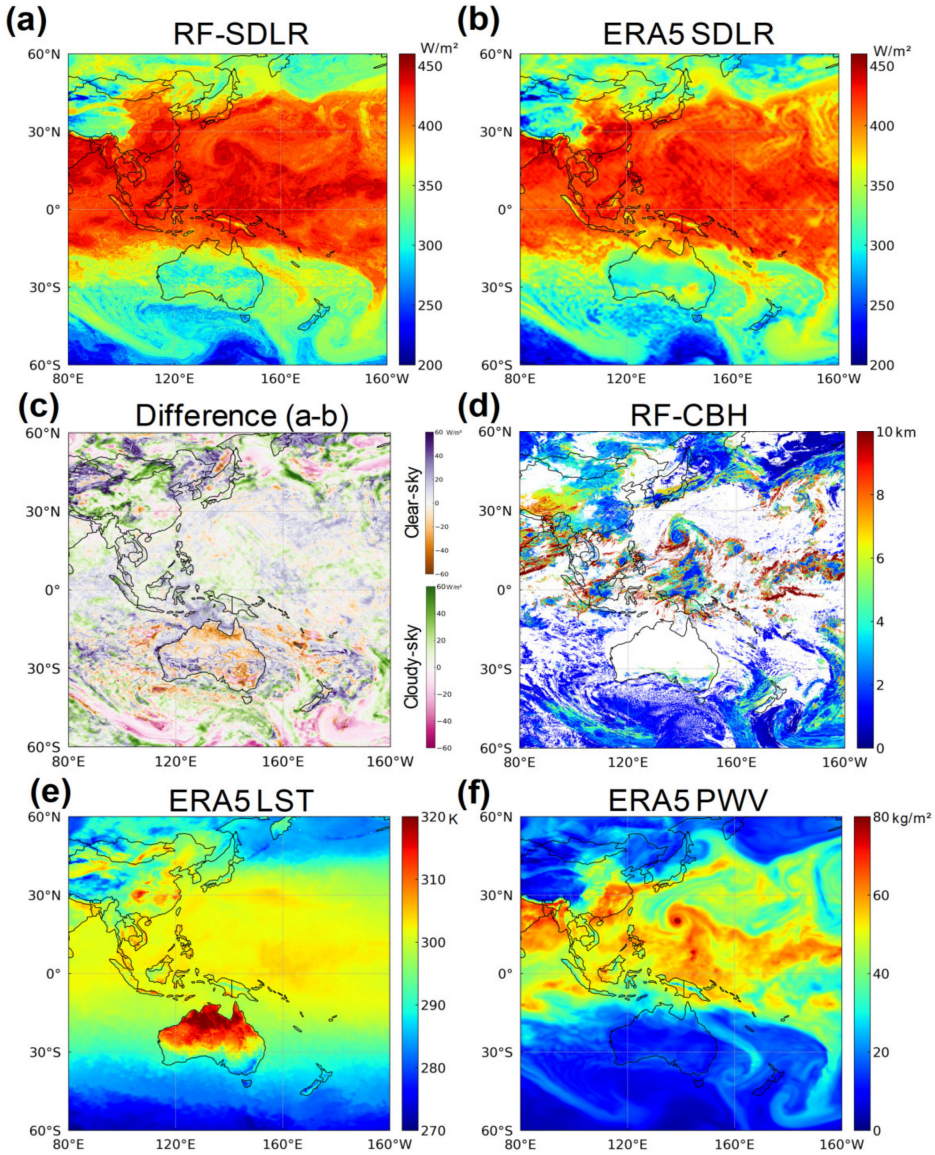

Figure 7. Spatial features of the H-8 full-disk SDLR and RF inputs on 1 September 2018: (**a**) H-8 SDLR derived using RF models; (**b**) ERA5 reanalysis SDLR product; (**c**) the difference between H-8 SDLR and ERA5 reanalysis SDLR product under cloudy-sky and clear-sky conditions; (**d**) H-8 CBH derived using the RF models; (**e**) ERA5 reanalysis LST product; and (**f**) ERA5 reanalysis PWV product.

Table 4. The validation results of RF-SDLR, the surrogated parameter algorithm, ERA5 reanalysis, and CERES against different ground networks. Note that there are three stations used in this study that belong to both the BOM and BSRN networks.

		RF(CBH)−SDLR			RF(CTT)−SDLR			ERA5			CERES			N
		RMSE	MBE	R	RMSE	MBE	R	RMSE	MBE	R	RMSE	MBE	R	
BOM	Clear−sky	23.6	−1.1	0.88	23.6	−1.1	0.88	28.9	−7.7	0.85	33.8	−3.5	0.77	101,775
	Cloudy−sky	19.5	0.8	0.92	26.8	2.6	0.88	25.1	−7.1	0.89	28.5	0.5	0.83	30,861
	All−sky	22.7	−0.6	0.89	24.3	−0.2	0.88	28.1	−7.6	0.86	32.6	−2.6	0.79	132,636
BSRN	Clear−sky	19.4	2.1	0.94	19.4	2.1	0.94	22.1	−2.9	0.92	26.6	3.1	0.89	67,303
	Cloudy−sky	17.0	−0.8	0.97	25.2	3.7	0.95	22.2	−7.7	0.95	23.6	−1.5	0.93	41,257
	All−sky	18.5	1.0	0.95	21.6	2.7	0.94	22.1	−4.7	0.94	25.5	1.3	0.91	108,560
GTMBA	Clear−sky	11.5	0.9	0.82	11.5	0.9	0.82	13.0	−2.6	0.79	12.0	−1.4	0.80	19,850
	Cloudy−sky	9.2	−0.3	0.68	12.4	0.7	0.52	12.7	−6.1	0.54	12.0	−5.8	0.57	9686
	All−sky	10.8	0.5	0.84	11.8	0.8	0.82	12.9	−3.8	0.80	12.0	−2.8	0.81	29,536
NMC	Clear−sky	31.0	1.3	0.79	31.0	1.3	0.79	43.8	−32.6	0.84	40.2	−17.9	0.76	11,006
	Cloudy−sky	29.7	5.1	0.86	34.2	7.4	0.88	37.6	−22.1	0.87	38.4	−1.0	0.79	14,929
	All−sky	30.2	3.5	0.83	32.8	4.8	0.85	40.3	−26.5	0.86	39.2	−8.1	0.77	25,935
All	Clear−sky	22.5	0.9	0.93	22.5	0.9	0.93	27.1	−7.2	0.92	30.5	−2.2	0.89	161,160
	Cloudy−sky	20.5	0.9	0.97	25.9	3.0	0.95	26.1	−10.0	0.96	28.0	−1.1	0.94	85,515
	All−sky	21.8	0.9	0.95	23.7	1.7	0.94	26.8	−8.2	0.94	29.6	−1.8	0.91	246,675

We chose the SDLR of in situ radiation data for a whole year in 2018 to compare and statistically assess the performance of new SDLR products. In this work, we compared the validation results of the RF-SDLR, the surrogated parameter algorithm, ERA5 reanalysis, and CERES. The RF-SDLR uses CBH as the input parameter, whereas the surrogated parameter method uses CTT. The SDLR of ERA5 reanalysis is discussed in Section 2, and the Clouds and the Earth's Radiant Energy System (CERES) is one of the few missions that delivers global all-sky surface SDLR outputs. The ERA5 and CERES SDLR products are widely verified and determined to perform well [11,23]. The validation results are displayed in Figure 8.

The RMSE of the RF-SDLI against the ground observations is 21.8 W m^{-2} for the full-disk H-8, which is better than 26.8 W m^{-2} for the ERA5 reanalysis, 29.6 W m^{-2} for the CERES, and 23.7 W m^{-2} for the surrogated parameter algorithm. The MBE of the RF-SDLI in this study is only 0.9 W m^{-2}, whereas the SDLR products of CERES and ERA5 show an apparent underestimation. This is especially true for the ERA5, with MBE values of −1.8 W m^{-2} and −8.2 W m^{-2}. Table 4 shows the performance of the four SDLR products for validation against different ground networks.

It can be seen in Table 4 that the RF-SDLR has a better performance compared to the other SDLR products in the comparison of each ground network for all indicators such as RMSE, MBE, and R. In addition, the performance of all SDLR products on the ocean is significantly better than that on land, as the accuracy of the SDLR observations on land will be influenced by the differences in the background.

More interestingly, the CERES performs the worst of the four SDLR products in the BOM and BSRN networks, and ERA5 performs poorly in the GTMBA and NMC networks; we believe that ERA5, as a worldwide reanalysis data system, references the World Radiation Monitoring Center's Baseline Surface Radiation Network (WRMC-BSRN), making it perform better for the BSRN and BSRN-like BOM. Furthermore, the SDLR of ERA5 reanalysis shows a severe underestimation of MBE values for all networks in almost all cloud-sky, clear-sky, and all-sky situations, which is consistent with many studies [24,56].

In summary, the RF-SDLR created in this work performed well in the validation compared to the ground observations, which can also show the effectiveness of the RF-CBH for the estimation of CBH. The SDLR estimates have the advantages of extensive coverage

and high spatiotemporal resolution. We think the conceptual framework and research results might inspire fresh approaches to Earth radiation balance.

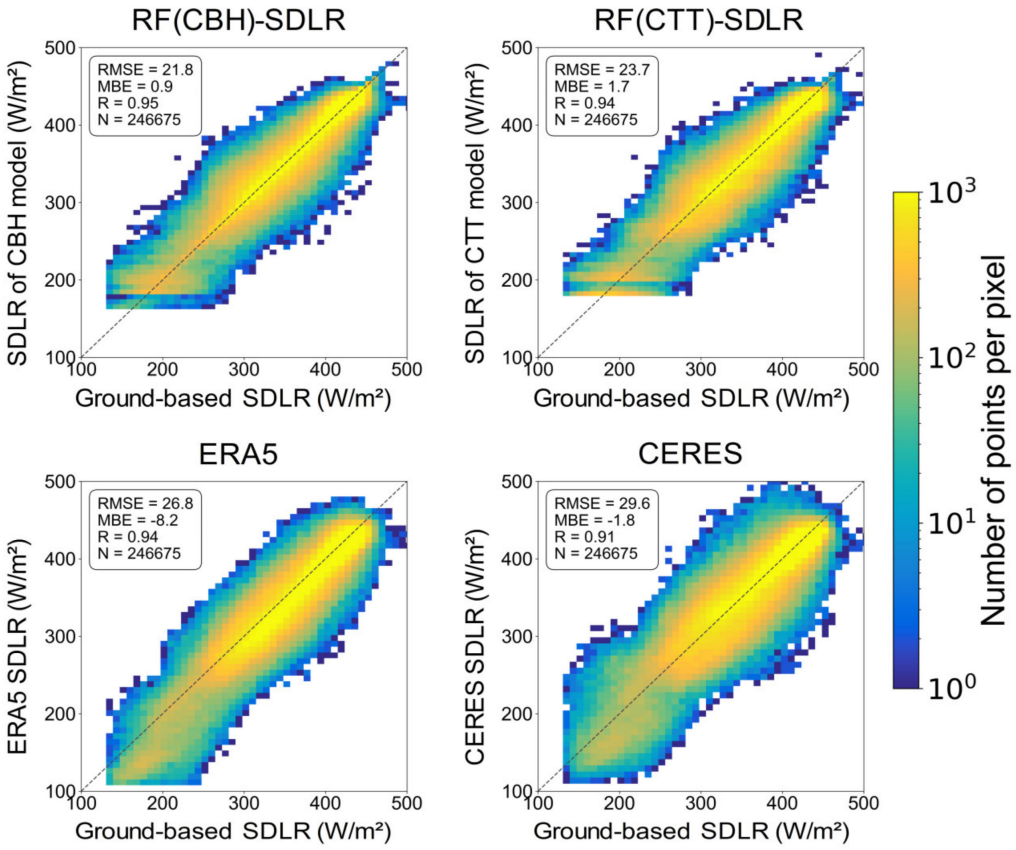

Figure 8. The validation results of RF-SDLR, the surrogated parameter algorithm, ERA5 reanalysis, and CERES against the ground observations.

4. Conclusions

The uncertainty of CBH retrieval by remote sensing induces a large number of SDLR estimation errors under cloudy conditions. This paper proposes a CBH retrieval methodology based on 10 thermal spectral data from H-8 observations utilizing the RF approach and using models trained individually for distinct cloud types to obtain more precise CBH data with a high spatiotemporal resolution. The algorithm can estimate the CBH 24 h a day, and has less reliance on upstream retrieval. After the models are trained, as long as there are appropriate inputs, the CBH can be independently estimated. The validation results of the RF-CBH show an effective performance, with the RMSE of stratus being the smallest at 0.47 km, and that of altocumulus being the largest at 1.56 km. The general RMSE value of all cloud types is 1.17 km.

After deriving a good CBH product, this study estimates all-sky SDLR based on the RF-CBH, and ERA5 reanalysis LST and PWV products using the RF approach. We innovatively utilized the CBH directly to the SDLR retrieval, so that the accuracy of SDLR estimations can be improved. Hourly RF-SDLI validation findings reveal an RMSE value of 21.8 W m^{-2} compared to ground observations, which is significantly better than 26.8 W m^{-2} for ERA5 reanalysis, 29.6 W m^{-2} for CERES, and 23.7 W m^{-2} for the surrogated parameter algorithm. The RF-SDLR also has a better performance compared to other SDLR products, as determined through the comparison of each ground network for all indicators, such as RMSE, MBE, and R, proving that using CBH to estimate the SDLR is an effective idea.

There are some limitations of the RF-CBH and RF-SDLR. For RF-CBH, all clouds are regarded as single-layered to streamline the CBH retrieval procedure. In reality, the multilayer clouds are complex with multiple cloud bases and signal reduction, leading to a potential error for CBH estimates. Precipitation clouds with strong radar signals will cause the wrong CBH of the training data. Precipitation clouds should be filtered out in further research to adopt a suitable method to estimate CBH. Finally, RF-CBH is also influenced by the cloud types products, and accurate cloud classification can improve the accuracy of CBH estimations. For RF-SDLR, the ERA5 data used in this study are not available in real-time or near real-time and have some limitations on the applications with high timeliness requirements.

The RF method used in this study can provide high-accuracy retrieval for CBH and SDLR, with high-speed performance after training the RF models. The algorithm is applicable to new-generation geostationary meteorological satellites. We must acknowledge, however, that the current iteration of this new approach still has much room for improvement. For instance: (1) the RF method is one of the more commonly used ML methods, and further attempts can be made using other methods such as the LightGBM method or extremely randomized trees to select the optimal method for estimation; (2) we ignored the influence of precipitation clouds on CBH estimation, as it retrieves a very small CBH in radar observations, but precipitation clouds should be filtered out in further research to adopt a suitable method to estimate CBH; (3) in the extraction and matching of model training data, the inputs and targets come from different satellites, and there are differences in the altitudes and observation angles of the satellites at the time of observation, and thus, we can add a parallax correction for the matching of training data to strengthen the training dataset; and (4) we can find more accurate cloud types products and superior LST and PWV products with higher spatiotemporal resolution to improve CBH and SDLR estimations, respectively.

Author Contributions: Conceptualization, H.L.; Data curation, J.S. and X.R.; Funding acquisition, H.L.; Methodology, J.S., H.L. and T.W.; Project administration, H.L. and G.T.; Software, J.S.; Supervision, H.L. and T.W.; Validation, J.S. and X.R.; Visualization, J.S.; Writing—original draft, J.S.; Writing—review and editing, H.L., G.T. and H.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (No. 42025504) and Second Tibetan Plateau Scientific Expedition and Research Program (No. 2019QZKK0206).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to still being used in a proceeding project.

Acknowledgments: The authors would like to thank the BSRN, the GTMBA, the NMC, and the Australian Government's BOM for providing surface radiation measurement data. They would also like to thank NASA, JAXA, and ECMWF for providing the CloudSat product, the Himawari-8 satellite data, and the ERA5 reanalysis data, respectively.

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

References

1. Wang, K.; Dickinson, R.E. Global atmospheric downward longwave radiation at the surface from ground-based observations, satellite retrievals, and reanalyses. *Rev. Geophys.* **2013**, *51*, 150–185. [CrossRef]
2. Yang, K.; Rasmy, M.; Rauniyar, S.; Koike, T.; Taniguchi, K.; Tamagawa, K.; Koudelova, P.; Kitsuregawa, M.; Nemoto, T.; Yasukawa, M. Initial CEOP-based review of the prediction skill of operational general circulation models and land surface models. *J. Meteorol. Soc. Jpn.* **2007**, *85*, 99–116. [CrossRef]
3. Iacono, M.J.; Mlawer, E.J.; Clough, S.A.; Morcrette, J.J. Impact of an improved longwave radiation model, RRTM, on the energy budget and thermodynamic properties of the NCAR community climate model, CCM3. *J. Geophys. Res. Atmos.* **2000**, *105*, 14873–14890. [CrossRef]

4. Wang, K.; Liang, S. Global atmospheric downward longwave radiation over land surface under all-sky conditions from 1973 to 2008. *J. Geophys. Res. Atmos.* **2009**, *114*, D19101. [CrossRef]
5. Wild, M.; Ohmura, A.; Gilgen, H.; Morcrette, J.-J.; Slingo, A. Evaluation of downward longwave radiation in general circulation models. *J. Clim.* **2001**, *14*, 3227–3239. [CrossRef]
6. Yang, K.; He, J.; Tang, W.; Qin, J.; Cheng, C.C. On downward shortwave and longwave radiations over high altitude regions: Observation and modeling in the Tibetan Plateau. *Agric. For. Meteorol.* **2010**, *150*, 38–46. [CrossRef]
7. Yang, F.; Cheng, J. A framework for estimating cloudy sky surface downward longwave radiation from the derived active and passive cloud property parameters. *Remote Sens. Environ.* **2020**, *248*, 111972. [CrossRef]
8. Jiao, Z.-H.; Mu, X. Global validation of clear-sky models for retrieving land-surface downward longwave radiation from MODIS data. *Remote Sens. Environ.* **2022**, *271*, 112903. [CrossRef]
9. Prata, A. A new long-wave formula for estimating downward clear-sky radiation at the surface. *Q. J. R. Meteorol. Soc.* **1996**, *122*, 1127–1151. [CrossRef]
10. Bisht, G.; Bras, R.L. Estimation of net radiation from the MODIS data under all sky conditions: Southern Great Plains case study. *Remote Sens. Environ.* **2010**, *114*, 1522–1534. [CrossRef]
11. Wang, T.; Shi, J.; Yu, Y.; Husi, L.; Gao, B.; Zhou, W.; Ji, D.; Zhao, T.; Xiong, C.; Chen, L. Cloudy-sky land surface longwave downward radiation (LWDR) estimation by integrating MODIS and AIRS/AMSU measurements. *Remote Sens. Environ.* **2018**, *205*, 100–111. [CrossRef]
12. Wang, T.; Shi, J.; Ma, Y.; Letu, H.; Li, X. All-sky longwave downward radiation from satellite measurements: General parameterizations based on LST, column water vapor and cloud top temperature. *ISPRS J. Photogramm. Remote Sens.* **2020**, *161*, 52–60. [CrossRef]
13. Duan, S.-B.; Li, Z.-L.; Tang, B.-H.; Wu, H.; Tang, R. Direct estimation of land-surface diurnal temperature cycle model parameters from MSG–SEVIRI brightness temperatures under clear sky conditions. *Remote Sens. Environ.* **2014**, *150*, 34–43. [CrossRef]
14. Gao, B.C.; Kaufman, Y.J. Water vapor retrievals using Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared channels. *J. Geophys. Res. Atmos.* **2003**, *108*, 4389. [CrossRef]
15. Li, Z.-L.; Becker, F. Feasibility of land surface temperature and emissivity determination from AVHRR data. *Remote Sens. Environ.* **1993**, *43*, 67–85. [CrossRef]
16. Ma, H.; Liang, S.; Xiao, Z.; Shi, H. Simultaneous inversion of multiple land surface parameters from MODIS optical–thermal observations. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 240–254. [CrossRef]
17. Wan, Z.; Dozier, J. A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 892–905.
18. Wu, Y.; Zhang, F.; Wu, K.; Min, M.; Li, W.; Liu, R. Best water vapor information layer of Himawari-8-based water vapor bands over east Asia. *Sensors* **2020**, *20*, 2394. [CrossRef]
19. Dong, S.; Cheng, J.; Shi, J.; Shi, C.; Sun, S.; Liu, W. A Data Fusion Method for Generating Hourly Seamless Land Surface Temperature from Himawari-8 AHI Data. *Remote Sens.* **2022**, *14*, 5170. [CrossRef]
20. Kato, S.; Rose, F.G.; Sun-Mack, S.; Miller, W.F.; Chen, Y.; Rutan, D.A.; Stephens, G.L.; Loeb, N.G.; Minnis, P.; Wielicki, B.A. Improvements of top-of-atmosphere and surface irradiance computations with CALIPSO-, CloudSat-, and MODIS-derived cloud and aerosol properties. *J. Geophys. Res. Atmos.* **2011**, *116*, D19209. [CrossRef]
21. Zhang, Y.; Rossow, W.B.; Lacis, A.A.; Oinas, V.; Mishchenko, M.I. Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data. *J. Geophys. Res. Atmos.* **2004**, *109*, D19105. [CrossRef]
22. Jiang, Y.; Tang, B.-H.; Zhao, Y. Estimation of Downwelling Surface Longwave Radiation with the Combination of Parameterization and Artificial Neural Network from Remotely Sensed Data for Cloudy Sky Conditions. *Remote Sens.* **2022**, *14*, 2716. [CrossRef]
23. Stephens, G.L.; Wild, M.; Stackhouse, P.W.; L'Ecuyer, T.; Kato, S.; Henderson, D.S. The global character of the flux of downward longwave radiation. *J. Clim.* **2012**, *25*, 2329–2340. [CrossRef]
24. Yu, S.; Li, L.; Cao, B.; Zhang, H.; Zhu, L.; Xin, X.; Liu, Q. Surface downward longwave radiation estimation from new generation geostationary satellite data. *Atmos. Res.* **2022**, *276*, 106255. [CrossRef]
25. Carmona, F.; Rivas, R.; Caselles, V. Estimation of daytime downward longwave radiation under clear and cloudy skies conditions over a sub-humid region. *Theor. Appl. Climatol.* **2014**, *115*, 281–295. [CrossRef]
26. Lhomme, J.-P.; Vacher, J.J.; Rocheteau, A. Estimating downward long-wave radiation on the Andean Altiplano. *Agric. For. Meteorol.* **2007**, *145*, 139–148. [CrossRef]
27. Crawford, T.M.; Duchon, C.E. An improved parameterization for estimating effective atmospheric emissivity for use in calculating daytime downwelling longwave radiation. *J. Appl. Meteorol. Climatol.* **1999**, *38*, 474–480. [CrossRef]
28. Duarte, H.F.; Dias, N.L.; Maggiotto, S.R. Assessing daytime downward longwave radiation estimates for clear and cloudy skies in Southern Brazil. *Agric. For. Meteorol.* **2006**, *139*, 171–181. [CrossRef]
29. Gaumet, J.; Heinrich, J.; Cluzeau, M.; Pierrard, P.; Prieur, J. Cloud-base height measurements with a single-pulse erbium-glass laser ceilometer. *J. Atmos. Ocean. Technol.* **1998**, *15*, 37–45. [CrossRef]
30. Takano, T.; Yamaguchi, J.; Abe, H.; Futaba, K.I.; Yokote, S.I.; Kawamura, Y.; Takamura, T.; Kumagai, H.; Ohno, Y.; Nakanishi, Y. Development and performance of the millimeter-wave cloud profiling radar at 95 GHz: Sensitivity and spatial resolution. *Electron. Commun. Jpn.* **2010**, *93*, 42–49. [CrossRef]

31. Marchand, R.; Mace, G.G.; Ackerman, T.; Stephens, G. Hydrometeor detection using CloudSat—An Earth-orbiting 94-GHz cloud radar. *J. Atmos. Ocean. Technol.* **2008**, *25*, 519–533. [CrossRef]
32. Noh, Y.-J.; Forsythe, J.M.; Miller, S.D.; Seaman, C.J.; Li, Y.; Heidinger, A.K.; Lindsey, D.T.; Rogers, M.A.; Partain, P.T. Cloud-base height estimation from VIIRS. Part II: A statistical algorithm based on A-Train satellite data. *J. Atmos. Ocean. Technol.* **2017**, *34*, 585–598. [CrossRef]
33. Bessho, K.; Date, K.; Hayashi, M.; Ikeda, A.; Imai, T.; Inoue, H.; Kumagai, Y.; Miyakawa, T.; Murata, H.; Ohno, T. An introduction to Himawari-8/9-Japan's new-generation geostationary meteorological satellites. *J. Meteorol. Soc. Japan. Ser. II* **2016**, *94*, 151–183. [CrossRef]
34. Letu, H.; Nakajima, T.Y.; Wang, T.; Shang, H.; Ma, R.; Yang, K.; Baran, A.J.; Riedi, J.; Ishimoto, H.; Yoshida, M. A new benchmark for surface radiation products over the East Asia-Pacific region retrieved from the Himawari-8/AHI next-generation geostationary satellite. *Bull. Am. Meteorol. Soc.* **2022**, *103*, E873–E888. [CrossRef]
35. Yang, J.; Zhang, Z.; Wei, C.; Lu, F.; Guo, Q. Introducing the new generation of Chinese geostationary weather satellites, Fengyun-4. *Bull. Am. Meteorol. Soc.* **2017**, *98*, 1637–1658. [CrossRef]
36. Tan, Z.; Huo, J.; Ma, S.; Han, D.; Wang, X.; Hu, S.; Yan, W. Estimating cloud base height from Himawari-8 based on a random forest algorithm. *Int. J. Remote Sens.* **2021**, *42*, 2485–2501. [CrossRef]
37. Tan, Z.; Ma, S.; Han, D.; Gao, D.; Yan, W. Estimation of cloud base height for FY-4A satellite based on random forest algorithm. *J. Infrared Millim. Waves* **2019**, *38*, 381–388.
38. Lin, H.; Li, Z.; Li, J.; Zhang, F.; Min, M.; Menzel, W.P. Estimate of daytime single-layer cloud base height from advanced baseline imager measurements. *Remote Sens. Environ.* **2022**, *274*, 112970. [CrossRef]
39. Hutchison, K. The retrieval of cloud base heights from MODIS and three-dimensional cloud fields from NASA's EOS Aqua mission. *Int. J. Remote Sens.* **2002**, *23*, 5249–5265. [CrossRef]
40. Min, M.; Li, J.; Wang, F.; Liu, Z.; Menzel, W.P. Retrieval of cloud top properties from advanced geostationary satellite imager measurements based on machine learning algorithms. *Remote Sens. Environ.* **2020**, *239*, 111616. [CrossRef]
41. Dong, Y.; Sun, X.; Li, Q. A Method for Retrieving Cloud-Top Height Based on a Machine Learning Model Using the Himawari-8 Combined with Near Infrared Data. *Remote Sens.* **2022**, *14*, 6367. [CrossRef]
42. Tang, T.; Jiao, D.; Chen, T.; Gui, G. Medium-and long-term precipitation forecasting method based on data augmentation and machine learning algorithms. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 1000–1011. [CrossRef]
43. Wang, G.; Wang, H.; Zhuang, Y.; Wu, Q.; Chen, S.; Kang, H. Tropical overshooting cloud-top height retrieval from himawari-8 imagery based on random forest model. *Atmosphere* **2021**, *12*, 173. [CrossRef]
44. Ri, X.; Tana, G.; Shi, C.; Nakajima, T.Y.; Shi, J.; Zhao, J.; Xu, J.; Letu, H. Cloud, Atmospheric Radiation and Renewal Energy Application (CARE) Version 1.0 Cloud Top Property Product From Himawari-8/AHI: Algorithm Development and Preliminary Validation. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–11. [CrossRef]
45. Håkansson, N.; Adok, C.; Thoss, A.; Scheirer, R.; Hörnquist, S. Neural network cloud top pressure and height for MODIS. *Atmos. Meas. Tech.* **2018**, *11*, 3177–3196. [CrossRef]
46. Gribbon, K.T.; Bailey, D.G. A novel approach to real-time bilinear interpolation. In Proceedings of the DELTA 2004, Second IEEE international workshop on electronic design, test and applications, Perth, WA, Australia, 28–30 January 2004; pp. 126–131.
47. 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. 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]
48. 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]
49. Miller, S.D.; Forsythe, J.M.; Partain, P.T.; Haynes, J.M.; Bankert, R.L.; Sengupta, M.; Mitrescu, C.; Hawkins, J.D.; Haar, T.H.V. Estimating three-dimensional cloud structure via statistically blended satellite observations. *J. Appl. Meteorol. Climatol.* **2014**, *53*, 437–455. [CrossRef]
50. Masumoto, Y.; Murty, V.; Nobre, P.; Ravichandran, M.; Vialard, J.; Vousden, D.; Yu, W. The global tropical moored buoy array. In Proceedings of the OceanObs09: Sustained Ocean Observations and Information for Society; Miloslavich, P., Klein, E., Daz, J.M., Hernndez, C.E.S., Bigatti, G., Eds.; Campos: Venice, Italy, 2010.
51. Ohmura, A.; Dutton, E.G.; Forgan, B.; Fröhlich, C.; Gilgen, H.; Hegner, H.; Heimo, A.; König-Langlo, G.; McArthur, B.; Müller, G. Baseline Surface Radiation Network (BSRN/WCRP): New precision radiometry for climate research. *Bull. Am. Meteorol. Soc.* **1998**, *79*, 2115–2136. [CrossRef]
52. Driemel, A.; Augustine, J.; Behrens, K.; Colle, S.; Cox, C.; Cuevas-Agulló, E.; Denn, F.M.; Duprat, T.; Fukuda, M.; Grobe, H. Baseline Surface Radiation Network (BSRN): Structure and data description (1992–2017). *Earth Syst. Sci. Data* **2018**, *10*, 1491–1501. [CrossRef]
53. Yu, Y.; Shi, J.; Wang, T.; Letu, H.; Yuan, P.; Zhou, W.; Hu, L. Evaluation of the himawari-8 shortwave downward radiation (swdr) product and its comparison with the ceres-syn, merra-2, and era-interim datasets. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *12*, 519–532. [CrossRef]
54. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

55. 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. Atmos.* **2008**, *113*, D00A12. [CrossRef]
56. Tang, W.; Qin, J.; Yang, K.; Zhu, F.; Zhou, X. Does ERA5 outperform satellite products in estimating atmospheric downward longwave radiation at the surface? *Atmos. Res.* **2021**, *252*, 105453. [CrossRef]