Estimation of cloud top parameters from Himawari-8/AHI measurements with infrared spectral bands using the Random Forest method

Xu Ri (1), Husi Letu (1), Shi Chong (2), Huazhe Shang (1)

[1] State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China
[2] Japan Aerospace Exploration Agency, Earth Observation Research Center, 2-1-1, Sengen, Tsukuba, Ibaraki, Japan
Email: xuri916924@163.com; husiletu@radi.ac.cn; shanghz@ radi.ac.cn.

KEY WORDS: Random Forests, cloud top parameters, CALIPSO, Advanced Himawari Imager, machine learning

ABSTRACT: We have developed a rapid simplified algorithm to estimate cloud top properties from infrared bands of Himawari based on machine learning. The new-generation geostationary satellite of Himawari-8 with Advanced Himawari Imager (AHI) provides high temporal (every 10 min) and high spatial resolution. CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations) provides cloud top parameters with high accuracy, but with limited temporal-spatial resolution. This paper reports on a study to derive the cloud top properties from combined AHI and CALIPSO using Random forests (RFs) algorithm, an advanced machine learning (ML) method with better accuracy than that from the traditional physical algorithms. Further sensitivity and validation analyses help determine the optimal RF classification and regression models for predicting process. The selected RF regression model is found to predict cloud top properties with highly consistent with CALIPSO observations (correlation coefficients are 0.89、0.89、0.90 for CTH, CTP, CTT respectively). New algorithm provides a robust and rapid algorithm of cloud top properties and we find significant accuracy improvements compared to AHI. Based on the accuracy evaluation of the model estimation results, the characteristics of cloud top properties in time and space are analyzed, and a typical case is selected to study. The application of the algorithm and error analysis are carried out to evaluate the estimation ability of cloud top parameters. The new approach could be used to process data from advanced geostationary imagers for climate and weather applications.

INTRODUCTION

Various cloud top properties are able to influence the balance of the earth atmosphere radiation, and have a significant regulatory effect on incoming and outgoing shortwave and thermal radiation [1]. Parameters such as cloud top height, cloud top temperature and cloud top pressure are of great practical significance for Atmospheric Physics, meteorological support and of particular importance for determining longwave radiation at the surface and aviation safety [2]. Therefore, accurate acquisition of cloud top properties information is of great significance for quantitatively describing the radiation budget of the earth-atmosphere system and studying climate change [3].

Satellite measurements are the most common and effective way to retrieve cloud top parameters because they offer continuous global information (e.g., Himawari, Fengyun-4, GEOS) with high accuracy. The easiest way to obtain the CTH is using a single IR channel mostly in the 11 μm assuming that the observed radiation is completely from the cloud. It was works well for optically thick clouds and large enough to fill the satellite view and the terrestrial radiation increases the observed IR value for thin clouds and partially fills the view. And finally causes CTH underestimated [4][5].

The most common way to retrieve the CTP, furthermore CTH and CTT is the CO_2-slicing method originally invented by Wylie and Menzel 1989[6].The method select two adjacent CO_2 absorption channels in the 15 μm because the atmosphere becomes more opaque as the wavelength approaches the selected bands leading observed to be sensitive to a different layer in the atmosphere [7]. And the method mentioned above also requires a radiative transfer model (RTM) simulation in the two selected channels for clear-sky radiances and additionally needs the atmospheric humidity profiles and temperature as inputs for a RTM. usually an RTM in cloudy skies has large uncertainties [8][9] with limited accuracy. It highly depends on the atmospheric temperature/humidity profiles and the spectral difference of cloud emissivity [10].

With the development of computer technology, a wide spectrum of advanced machine learning (ML) techniques, such as Knearest-neighbor (KNN), random forests (RF), support vector machines(SVM), artificial neural network (ANN), deep learning (DL, one kind of complex ANN algorithm), etc., offers a possible solution to some nonlinear issues in remote sensing and geoscience fields [11][12][13]. It has been successfully and extensively applied in cloud products in recent years. The computing efficiencies of ML techniques have been much improved, offering us unprecedented opportunities to process large-volume data sets in near real-time systems. use advanced ML algorithms for model training and high efficiency prediction. A previous study [14] used a neural network algorithm to train and CTP and CTH for several passive sensors in polar-orbit. RFs, as the high-accurate and promising ML algorithms, have received increasing attention for remote sensing applications. Including bagging ensemble classification and regression technique, the RF algorithm can easily run in a parallel computing mode and capture nonlinear or complex relationships between predictor and predictand [15].

The primary objective of this paper is to develop a rapid and unified retrieval algorithm for cloud top parameters from real-time Himawari-8/AHI (H08/AHI) observations, using RFs. The implementation of this rapid and unified cloud top algorithm is expected to improve the accuracy of cloud top products in East Asia, and to promote wider application of GEO meteorological satellite data in nowcasting applications. Cloud properties are derived in the same way for both day- and nighttime data because the IR method is independent of solar illumination. Cloud top pressure, as an important indicator of cloud top height, is a good indicator for cloud dynamics. The significance of studying cloud top height or cloud top pressure lies not only in its representativeness of dynamic characteristics, but also in its close relationship with thermodynamic factors. However, the retrieval of cloud top pressure is still one of the difficulties in retrieving cloud properties from satellite remote sensing. This is not only related to the spatial and temporal resolution of satellite data, but also to the signal-to-noise ratio (SNR) performance of remote sensing instruments for satellite observation and detection, cloud amount in the field of view and other atmospheric parameters (such as atmospheric temperature and humidity profile, underlying surface temperature).

The remainder of this paper is organized as follows, Section 2 briefly introduces the satellite and ancillary data used for training RF model. Section III presents the algorithm in detail, including RFs introduction and RF regression models. In Section IV, major results of cloud top parameters based on the ML algorithm are presented, which are further validated against CALIOP and MODIS followed by characteristics analyzation of cloud top parameters in time and space and Finally, Section V provides a short summary.

DATA AND METHODS

The new-generation geostationary meteorological satellite Himawari-8 carrying Advanced Himawari Imager (AHI), was successfully launched into geosynchronous orbit by Japan Meteorological Agency (Japan Aerospace Exploration Agency JAXA) on October7, 2014, and the observation data began to be released on July 7, 2015[16]. AHI has 16 (VIS-4, NIR-2 and IR-10) bands with diverse spatial resolutions ranging from 0.5 (visible band) to 2.0 km(IR band) and a full-disk observation frequency of 10min [17] At present, Himawari-8 satellite has been applied in the fields of surface and sea surface temperature inversion, cloud and haze detection, aerosol data assimilation and forest fire detection. In this investigation, we use one year (from January to December of 2018) of continuous H8/AHI data for developing the new algorithm and conducting validation and further analysis. Table 1 list the Himavari-8/AHI specification [18][19].

In this investigation, to train the model in RFs algorithm, the global cloud observations with 333m resolution from The Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations(CALIPSO) Version-4.20 dataset are used to create the parameters training dataset and to provide a benchmark for validation, Note that, in this paper, the matching algorithm calculates the minimum distance between CALIPSO footprint and Himawari-8/AHI pixel to determine the specific matching point within a time distance of about±5 min. CALIPSO has flown in formation with the NASA A-train constellation of satellites launched at May 2006. Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) is a two-wavelength polarization lidar (the primary instrument on the satellite) that performs global profiling of aerosols and clouds in the troposphere and lower stratosphere [20][21]. It is the first Cloud-Aerosol Lidar with two channels at 532 and 1064 nm. Thus, it can detect more complete vertical structure of clouds are reliable for determining the cloud-top properties of thin, high-level clouds. The global, multiyear dataset obtained from CALIOP provides a new view of the earth's atmosphere and will lead to an improved understanding of the role of aerosols and clouds in the climate system. A CALIOP represents a major advance in space-based active remote sensing of clouds and aerosols, and the CALIOP algorithms have many unique aspects designed to take advantage of these new capabilities. There are three basic types of level 2 data products: layer products, profile products, and the vertical feature mask (VFM). Layer products provide layer-integrated or layer-averaged properties of detected aerosol and cloud layers.

We employed 60 days of continuous quantitative level 2 gridded CALIPSO (version of V4.20) cloud top parameters data in the 2018 as the training set. The layer data have a time interval of an hour and covers the whole area between the latitudes of 60°S and 60°N within a spatial resolution of 0.03° × 0.03°. Considering the potential effect of the angle of field-of-view (FOV) of GEO satellite imaging sensor on the cloud top properties retrieval, we abandon the samples of H8/AHI at high latitudes (large satellite angles) and only focus on the samples bracketed from 70°E to 180°E and −70°S to 70°N. The first-class products of Himawari-8, observed by IR bands (Bands 9–16) of H08/AHI from 7.3 to 13.2 μm (Table I) as input features to predict or estimate cloud top parameters. Level 2 gridded CALIPSO cloud top properties data collocated by taking the target output. The data are utilized to train and develop the ML prediction model. Furthermore, collocated sample data on the 10nd and 16th days of every month in 2019 are randomly chosen to be an independent validation dataset (these become the testing dataset in next paragraph).

Finally, to further validate cloud top properties retrieval algorithms from GEO satellite measurements, the accuracy of predicted products is evaluated by comparing with CALIPSO data and we also use the previous algorithms for satellite passive imagers employ spatial-temporally matched temperature, pressure, and height data of latest Moderate Resolution Imaging Spectroradiometer (MODIS) official Collection-6.1 level-2 products for inter-comparison [22]. MODIS have widely been used to illustrate the global distributions. Offering various physical products for clouds, aerosols, sea surfaces, wildfires, photosynthetically active radiation, and so on. Cloud mask, types, microphysical and optical properties, and cloud life cycles, these products are widely used for weather analysis and forecasting, short-term climate prediction, and environmental and disaster monitoring in Asia of cloud fraction, cloud mask, cloud optical thickness (COT), top height (CTH), and temperature (CTT) for both ice and liquid water clouds.

Random forest RFs is a machine learning algorithm developed in recent ten years by Leo Breiman and Adele Cutler. which has been widely used for both classification and regression problems without much hyperparameter tuning. Its basic unit is decision tree, and its essence belongs to Ensemble Learning, a major branch of machine learning. Ensemble learning is to use a series of learners to learn, and integrate various learning methods through a specific rule. RFs has its unique advantages and characteristics and is a classifier that contains multiple decision trees. By integrating multiple decision trees to form the whole forest in a random way, the algorithm results are obtained. There are many decision trees in the forest, and there is no correlation between these decision trees. An algorithm that integrates multiple trees with the idea of ensemble learning is suitable for dealing with the complex relationship between predicted values and non-linear inputs. averaged to improve the predicted accuracy and reduce overfitting. It can well capture nonlinear association patterns between predictor and predictand variables. Despite the afore mentioned advantages, the RF algorithm is still lack of interpretability and mathematical theory by nature, making it almost impossible or uneasily to demonstrate how the predictions or decisions are made.

In this investigation, the free, simple, and efficient scikitlearn toolkit [13], a well-known Python module for ML, has been used to implement the training, parameter adjustment, and prediction within this RF algorithm. A range of typical classification, regression, and clustering algorithms are integrated into this Python ML toolkit, including RFs, SVMs, and k-means, among others. The entire RF model flowchart is shown in Fig. 1.

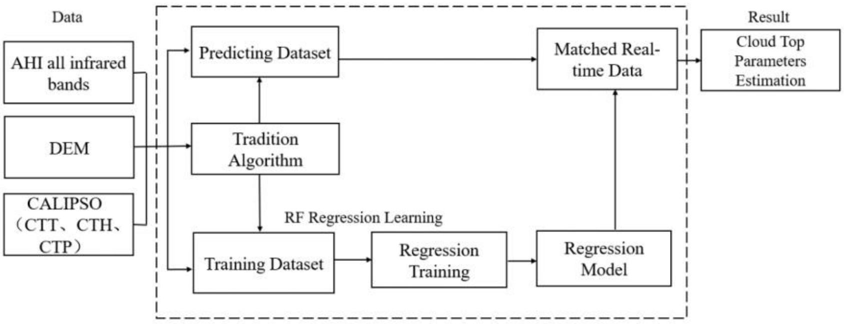

Fig. 1 The flowchart of the cloud top parameters estimation algorithm

TABLE I

Coefficient of determination, RMSE and MBE as a function of different independent variables. (variable 1: CTT, variable 2: CTH and variable 3: CTP)

(band number)	variable 1			variable 2			variable 3		
Independent Variables	**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	183.16
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
16	0.536	0.70	29.38	0.434	0.15	5.17	0.503	15.77	264.63

The effects of various kinds of group of input parameters compared and showed in tableI,the RF model involving all the input IR bands and the DEM value provided the best performance. There was a 32% increase in the RMSE than when applied single band. The table probably implies that the channel data itself mostly contains the determinant information for estimating, which suggests the fact that the infrared channel detected by AHI can help the model better estimate cloud properties and should be all considered in RF retrieval algorithms. In order to get the optimal RF model, we debug the structure and parameters of the model many times, compare the error of the model after each debugging, and select the model with the highest stability and the smallest error.

TABLE II
Predictor variables in the RF Model and their corresponding rankings
(variable 1: CTT, variable 2: CTH and variable 3: CTP).

	Variable 1		Variable 2		Variable 3	
Satellite measurement	Importance score for variable1	Ranking	Importance score for variable2	Ranking	Importance score for variable3	Ranking
TBB7	0.0409711	6	0.0332308	7	0.0332308	7
TBB8	0.0381229	7	0.0345255	6	0.0327108	8
TBB9	0.1304547	3	0.0819584	4	0.0696068	4
TBB10	0.0376878	8	0.0317636	8	0.0287540	9
TBB11	0.0546763	5	0.0485053	5	0.0478789	5
TBB12	0.1506291	2	0.2507723	2	0.2083979	2
TBB13	0.0151776	10	0.0127667	11	0.0125791	11
TBB14	0.0159162	9	0.0140720	10	0.0146956	10
TBB15	0.3849212	1	0.3041278	1	0.4129626	1
TBB16	0.1170605	4	0.1595683	3	0.1027405	3
DEM	0.0143821	11	0.0255693	9	0.0364426	6

Table II shows all the predictor variables and their rankings for this optimal RF regression model. As one of the key parameters in RF algorithm, here represents the weighting coefficient of every predictor in the fitting prediction model. Among others, the TBB observed by Himawari-8/AHI band15 get the top rankings, indicating the importance of cloud top parameters from space. It is known that this IR wavelength are generally characterized with relatively high atmospheric transmission and weak atmospheric absorption. In addition, the TBBs observed by band 12 and band 16 show relatively high rankings. It is very interesting to see that band 9 representing low atmospheric layer information, also get a high importance. Overall, the results illustrate that band 13 and band14 real-time Himawari-8/AHI observation data are not significant for prediction pixels.

VALIDATION AND RESULTS

In this study, we developed an algorithm for cloud top parameters estimation using a machine learning method. The RF algorithm model can quickly and accurately calculate the CTT, CTH and CTP by directly connect the satellite in a simple way rather than the complex scheme based on the radiance ratioing method. To verify the effectiveness of the algorithm,12 days (each month of day 10 and 16 in 2019) of independent and spatial-temporally CALIPSO product matched data are used here for validating the performances of the RF regression models and utilized to assess the performance of the cloud top properties produced from RF method. In addition, the validation of estimated cloud top properties against MODIS cloud top product indicates the further confirmation of the accuracy of the RF algorithm model.

To quantitatively assess the accuracy of estimated properties derived from AHI high temporal-resolution observations, we assume that the CALIPSO measurements as the truth data. This high spatial-temporal resolution of Himawari 8 level 1 product provide us opportunity to study cloud top parameters in depth. Therefore, we select the optimum RF model with high stability by tested multiple experiments. In total 205,670 sample points are adopted to evaluate the performance of the selected model. It can be seen overall that the instantaneous cloud top parameters agree well with Calipso product for all cases. The validation is limited in the following several aspects: 1) The Clay measurements are taken in orbit and cannot provide validation over large-area regions. 2) The horizontal resolution of the Clay data at three different elevations ranged from 333 m to 1667 m. We sampled the Clay data to a horizontal resolution of 333m using only the layer top product, which might have introduced bias into the resampled Clay result.

To validate the RF-retrieved results, the scatter plots of Calipso cloud top parameters against AHI measurements and Himawari level-2 products corresponds, as well as their concomitant correlation coefficients (R) are shown in Fig. 2. Predicted value were compared with collocated CALIPSO observations. Collocation is accomplished by matching the CALIPSO latitude and longitude to those of 5 km×5 km H8 cloud top products. We can easily see that RF-derived cloud top products showed high consistency with Calipso than the Himawari holding the determination coefficient of 0.90, 0.89, 0.89 for CTT, CTH and CTP respectively. We also found the RF values estimated from AHI are slightly smaller than the radar-measured results, the introduction of RF predicted cloud top parameters can significantly improve the accuracy of products and comparable to other spaceborne lidar and radar measurements. The performance of current study is generally better than that of other studies. The evaluation also showed that the RF model has a better performance in reproducing CTP and supplements the absence of Himawari CTP product.

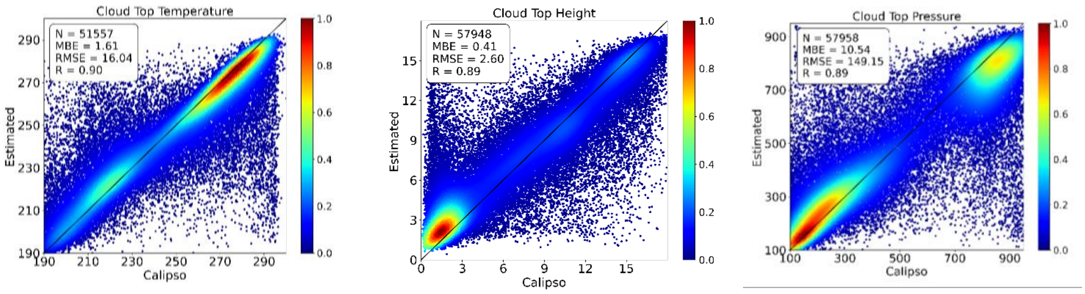

Fig.2 Comparison between RF-derived against CALIOP: (1) CTT, (2) CTH and (3) CTP.

The scatter plots of RF-derived cloud top parameters quantitatively compared against CALIOP product as well as their sample numbers (N), concomitant correlation coefficients (R) in 2019 are shown in Fig.3.CALIOP was regard as true value in this study. the overall coefficients of determination are high for most of the RF-derived parameters. As we also can see from the graph, each parameter value has two peaks, which is assumed to the reason of different cloud phases. Thus, the validation was followed by the cloud top parameters evaluation of the for each cloud phase.

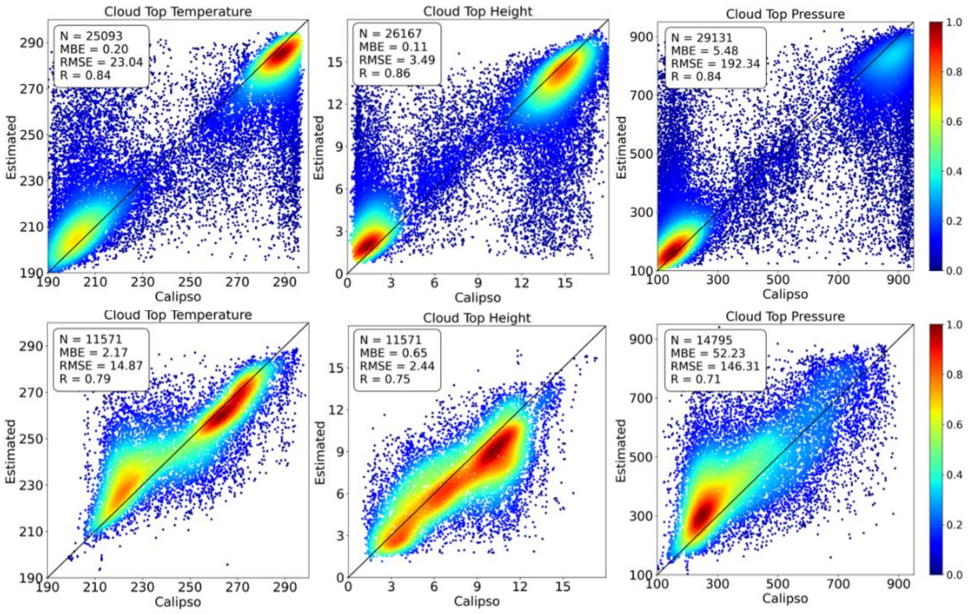

Fig.3 Cloud top parameters evaluation of the RF model retrieval for each area in 2019.
each row represents: (1) Sea area and (2) Land area.

In the sea area, coefficients of RF-derived CTT yielded the highest value of 0.84 which indicated that the cloud top parameters in this research are reliable over the sea land. However, the CTHs estimated from RF and are smaller than the CAISPSO results while the CTP overestimated over the land area. We also found the CTP values show large inconsistent variations with the CALIPSO. At the Land site, the highest R is 0.75 for CTH of RF-derived and the corresponding MBE and RMSE are 0.65 and 2.44, respectively. Table VI is another way to gain more insight into the capability of RF algorithm products.

Based on the observation data of Himawari-8 satellite level-1 data, combined with artificial intelligence machine learning method and CALIPSO cloud top product data, an efficient algorithm for estimating CTT, CTH and CTP was developed. On this basis, the characteristics of high spatial and temporal scales of cloud top properties in Asia-Pacific region in 2018 are analyzed. Himavari-8/AHI covers Asia-Pacific region, where the spatial distribution of cloud top properties is significantly diverse due to the difference of natural and climatic environments for each region. In this case, we analyze the annual scale of cloud top parameters. The results in fig.4 shows the annual average distribution of the cloud top parameters over Asia-Pacific region: (1) CTP, (2) CTT and (3) CTT. The spatial distribution of cloud top parameters is visually similar with each other and large discrepancies exits over many regions for each parameter.

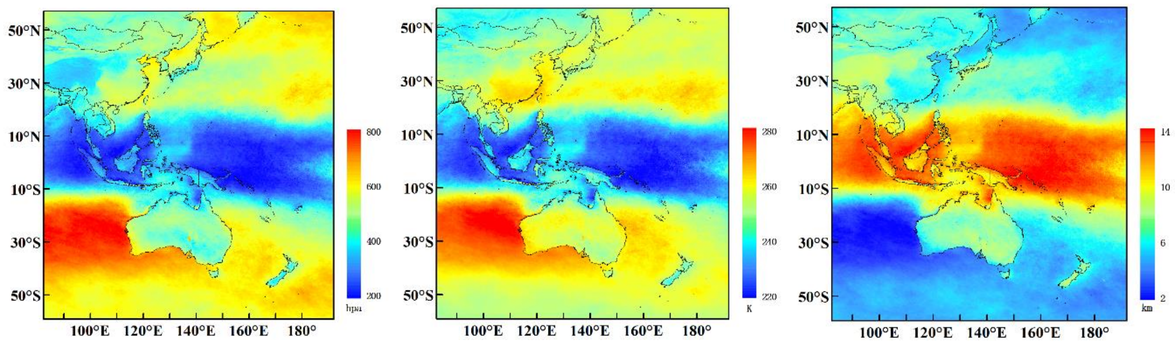

Fig.4 Annual average distribution of the cloud top parameters over Asia-Pacific region (1) CTP, (2) CTT and (3) CTT.

The results indicate that the annual average CTP, CTT and CTH are generally between 200-800 hpa, 220-280 K and 2-14km in each region. Similar trend is seen from the zonal averages of cloud top temperature and pressure for different latitudes in research area. Both the algorithms of the two properties have similar performance over the Asia-Pacific region. While the cloud top height has an opposite appearance to the pressure showing highest value on the horizontal level and generally lower than that of other regions on the latitude from -60° to -20° with a peak at around -30°, in spatial distribution the largest CTT in north hemisphere occurs at around 280K, the largest south hemisphere occurs at 30°, the largest CTP in north hemisphere occurs at 800 hpa. The properties both decreased to the minimum value at equator regions. Large bias is found between the results of CTT and CTP in and south hemisphere through both two methods. The highest CTT areas are distributed over the southeastern part of the disk and the minimum CTT appears on the equator region. The cloud top height is generally less than 7km in Asia. and there are some differences in its spatial variation. The rainy area with more low cloud cover distributes lower CTP. When winter comes, the high precipitation area moves southward, at the same time the sub-high value area of CTP moves to North area, which is due to precipitation in North area in winter.

CONCLUSION

Cloud top parameters plays a significant role in the radiation budget and energy exchange in the earth-atmosphere system. Satellite remote sensing provides a unique way to estimate the variation on a high spatial-temporal scale. The new generation of Himawari-8 satellites has great improvement in time and space resolution. This paper developed a high-temporal for AHI cloud top products, and constructed a high-precision random forest model for cloud top parameters, which can rapidly develop cloud top products based on the combination of real time AHI IR bands observation and CALIPSO cloud layer products. A simple, yet efficient and accurate algorithm is presented to estimate the cloud top parameters.

We also used CALIPSO to validate the accuracy of RF model algorithm seen CALIPSO as true value. As a result, it is worth reemphasizing that, estimated values from Himawari-8 data in 2019 shows good consistency with validation samples of CALIPSO cloud top parameters, holding the determination coefficient of 0.90, 0.89 and 0.89 respectively. The inter-comparison results were different for the case study and the small area region because of the lack of sample quantities and the differ sensitivities to the cloud particles. However, the retrieval precision of our algorithm was generally within the scope of expected theoretical accuracy.

Finally, based on the accuracy evaluation of the RF-derived results, the characteristics of cloud top temperature and pressure in time and space are analyzed, and a typical case is selected to study. The application of the algorithm and error analysis are carried out to evaluate the estimation ability of cloud top parameters. It should be notable that we only use one-year data to analyze the seasonal character, so there may be uncertainty in analyzing the spatial-temporal features of parameters. In addition, more analysis in daily or even hourly time scales are needed. This study not only validates that the RF model is effective and practical but also proposed an alternative method for estimating cloud top parameters with a high accuracy. RF model could be used to retrieval the further cloud variables such as the optical cloud thickness and effective radii that are associated with the absence of high accuracy products.

REFERENCE

[1]Weisz, E., Li, J., Menzel, W.P., Heidinger, A.K., Kahn, B.H., Liu, C.Y., 2007. Comparison of AIRS, MODIS, CloudSat and CALIPSO cloud top height retrievals. Geophys. Res. Lett.34, 1–5

[2] Holz, R.E., Ackerman, S.A., Nagle, F.W., Frey, R., Dutcher, S., Kuehn, R.E., Vaughan, M., Baum, B.A., 2008. Global Moderate resolution Imaging Spectroradiometer (MODIS)cloud detection and height evaluation using CALIOP. J. Geophys. Res. 113.

[3]Ma, R., Letu, H., Yang, K., Wang, T.X., Shi, C., Xu, j., Shi, J.C., Shi, C., & Chen, L.F. (2020). Estimation of Surface Shortwave Radiation From Himawari-8 Satellite Data Based on a Combination of Radiative Transfer and Deep Neural Network. IEEE Transactions on Geoscience and Remote Sensing, 1-13

[4]Menzel, W.P.; Wylie, D.P.; Strabala, K.I. Seasonal and diurnal changes in cirrus clouds as seen in four years of observations with the VAS. J. Appl. Meteor. 1992, 31, 370–385.

[5]Wylie, D.P.; Santek, D.; Starr, D.O.C. Cloud-top heights from GOES-8 and GOES-9 stereoscopic imagery. J. Appl. Meteor. 1998, 37, 405–413.

[6]Wylie, D.P.; Menzel, W.P. Two years of cloud cover statistics using VAS. J. Clim. 1989, 2, 380–392. 8

[7] Menzel, W.P.; Frey, R.A.; Zhang, H.; Wylie, D.P.; Moeller, C.C.; Holz, R.E.; Maddux, B.; Baum, B.A.; Strabala, K.I.; Gumley, L.E. MODIS global cloud-top pressure and amount estimation: Algorithm description and results. J. Appl. Meteorol. Cliim. 2008, 47, 1175–1198.

[8] Li, J., Li, Z., Wang, P., Schmit, T.J., Bai, W., Atlas, R., 2017. An efficient radiative transfer model for hyperspectral IR radiance simulation and applications under cloudy sky conditions. J. Geophys. Res. 122, 7600–7613.

[9] Li, J., Yi, Y.H., Stamnes, K., Ding, X.D., Wang, T.H., Jin, H.C., Wang, S.S., 2013. A new approach to retrieve cloud base height of marine boundary layer clouds. Geophys.Res. Lett. 40, 4448–4453.

[10] Hamann, U.; Walther, A.; Baum, B.; Bennartz, R.; Bugliaro, L.; Derrien, M.; Francis, P.N.; Heidinger, A.; Joro, S.; Kniffka, A.; et al. Remote sensing of cloud top pressure/height from SEVIRI: Analysis of ten current retrieval algorithms. Atmos. Meas. Tech. 2014, 7, 2839–2867. 4

[11] Kühnlein, M., Appelhans, T., Thies, B., Nauß, T., 2014a. Precipitation estimates from MSGSEVIRI daytime, nighttime, and twilight data with random forests. J. Appl. Meteorol. Climatol. 53, 2457–2480.

[12] Kühnlein, M., Appelhans, T., Thies, B., Nauss, T., 2014b. Improving the accuracy of rainfall rates from optical satellite sensors with machine learning—a random forestsbased approach applied to MSG SEVIRI. Remote Sens. Environ. 141, 129–143.

[13] Min, M., Bai, C., Guo, J., Sun, F., Liu, C., Wang, F., Xu, H., Tang, S., Li, B., Di, D., Dong, L., Li, J., 2019. Estimating summertime precipitation from Himawari-8 and global forecast system based on machine learning. IEEE Trans. Geosci. Remote Sens. 57,2557–2570.

[14] Håkansson, N., Adok, C., Thoss, A., Scheirer, R., Hiörnquist, S., 2018. Neural network cloud top pressure and height for MODIS. Atmos. Meas. Tech. 11, 3177–3196.

[15] Breiman, L., 2001. Random forests. In: Machine Learning. 45. pp. 5–32.

[16] Letu H, Nagao TM, Nakajima TY, Riedi J, Ishimoto H, Baran AJ, et al. Ice Cloud Properties from Himawari-8/AHI Next-Generation Geostationary Satellite: Capability of the AHI to Monitor the DC Cloud Generation Process. IEEE Transactions on Geoscience and Remote Sensing. 2019;

57:3229-39.
[17] Husi, L., Nagao, T.M., Nakajima, T.Y., Riedi, J., Ishimoto, H., Baran, A.J., Shang, H., Sekiguchi, M., Kikuchi, M., 2019. Ice cloud properties from Himawari-8/AHI next generation geostationary satellite: capability of the AHI to monitor the DC cloud generation process. IEEE Trans. Geosci. Remote Sens. 57, 3229–3239.
[18] Min, M., Wu, C., Li, C., Liu, H., Xu, N., Wu, X., Chen, L., Wang, F., Sun, F., Qin, D., Wang, X., Li, B., Zheng, Z., Cao, G., Dong, L., 2017. Developing the science product algorithm testbed for Chinese next-generation geostationary meteorological satellites: Fengyun-4 series. J. Meteorol. Res. 31, 708–719.
[19] Sun, F., Min, M., Qin, D., Wang, F., Hu, J., 2019. Refined typhoon geometric center derived from a high spatiotemporal resolution geostationary satellite imaging system. IEEE Geosci. Remote Sens. Lett. 16, 499–503.
[20]Min, M., Zhang, Z., 2014. On the influence of cloud fraction diurnal cycle and sub-grid cloud optical thickness variability on all-sky direct aerosol radiative forcing. J. Quant. Spectrosc. Radiat. Transf. 142, 25–36.
[21] Winker, D.M., Vaughan, M.A., Omar, A., Hu, Y., Powell, K.A., Liu, Z., Hunt, W.H., Young, S.A., 2009. Overview of the CALIPSO mission and CALIOP data processing algorithms. J. Atmos. Ocean. Technol. 26, 2310–2323.
[22] Baum, B., Menzel, W.P., Frey, R., Tobin, D., Holz, R., Ackerman, S., 2012a. MODIS cloud top property refinements for Collection 6. J. Appl. Meteorol. Climatol. 51,1145–1163.