

Unsupervised Clustering of Geostationary Satellite Cloud Properties for Estimating Precipitation Probabilities of Tropical Convective Clouds

DOYI KIM^a, HEE-JAE KIM,^b AND YONG-SANG CHOI^a

^a Department of Climate and Energy Systems Engineering, Ewha Womans University, Seoul, South Korea

^b Department of Electronic and Electrical Engineering, Ewha Womans University, Seoul, South Korea

(Manuscript received 3 November 2022, in final form 7 April 2023, accepted 3 July 2023)

ABSTRACT: Understanding the growth of tropical convective clouds (TCCs) is of vital importance for the early detection of heavy rainfall. This study explores the properties of TCCs that can cause them to develop into clouds with a high probability of precipitation. Remotely sensed cloud properties, such as cloud-top temperature (CTT), cloud optical thickness (COT), and cloud effective radius (CER) as measured by a geostationary satellite are trained by a neural network. First, the image segmentation algorithm identifies TCC objects with different cloud properties. Second, a self-organizing map (SOM) algorithm clusters TCC objects with similar cloud microphysical properties. Third, the precipitation probability (PP) for each cluster of TCCs is calculated based on the proportion of precipitating TCCs among the total number of TCCs. Precipitating TCCs can be distinguished from nonprecipitating TCCs using Integrated Multi-Satellite Retrievals for Global Precipitation Measurement precipitation data. Results show that SOM clusters with a high PP ($>70\%$) satisfy a certain range of cloud properties: $\text{CER} \geq 20 \mu\text{m}$ and $\text{CTT} < 230 \text{ K}$. PP generally increases with increasing COT, but COT cannot be a clear cloud property to confirm a high PP. For relatively thin clouds ($\text{COT} < 30$), however, CER should be much larger than $20 \mu\text{m}$ to have a high PP. More importantly, these TCC conditions associated with a $\text{PP} \geq 70\%$ are consistent across regions and periods. We expect our results will be useful for satellite nowcasting of tropical precipitation using geostationary satellite cloud properties.

SIGNIFICANCE STATEMENT: We aim to identify the properties of tropical convective clouds (TCCs) that have a high precipitation probability. We designed a two-step framework to identify TCC objects and the conditions of cloud properties for TCCs to have a high precipitation probability. The TCCs with a precipitation probability $> 70\%$ tend to have a low cloud-top temperature and a cloud particle effective radius $\geq 20 \mu\text{m}$. Cloud optical thicknesses are distributed over a wide range, but thinning requires a particle radius larger than $20 \mu\text{m}$. These conditions of cloud properties appear to be unchanged under various spatial–temporal conditions over the tropics. This important observational finding advances our understanding of the cloud–precipitation relationship in TCCs and can be applied to satellite nowcasting of precipitation in the tropics, where numerical weather forecasts are limited.

KEYWORDS: Convective clouds; Precipitation; Cloud microphysics; Classification; Deep learning


1. Introduction

Tropical convective clouds (TCCs) are of great interest in atmospheric science as well as hydrological risk management. TCCs are found across the tropics in the 20°N – 20°S regions where large populations are exposed to flood risks (Rentschler and Salhab 2020). In such vulnerable regions, early detection of precipitation from TCCs would be an effective way to minimize losses due to flooding. However, early detection remains challenging because TCCs are developed in as little as 3 to 6 h before producing heavy rainfalls (Chen and Houze 1997). Forecasters, therefore, need to distinguish discernible indicators of precipitation potential from the optical and microphysical properties of developing TCCs.

Satellite observations are now indispensable to weather forecasting. A wide range of TCCs is being perpetually monitored

by geostationary satellites. In the early days of satellite observations, detecting precipitating clouds was based on the statistical relationship between infrared brightness temperature T_B at $10.4 \mu\text{m}$ and surface rainfall rates. The T_B is generally known as an “atmospheric window channel,” where it is less affected by the absorption of gas molecules and can identify thick clouds among high-altitude clouds. The T_B threshold method assumes that rainfall rates on the ground are proportional to the decrease in T_B , as indicators of the growth stage of sufficiently thick clouds, such as mature TCCs (Spencer et al. 1983; Yuter and Houze 1998; Yuan and Houze 2010). For example, Arkin (1979) proposed that the strongest correlation between cold cloud cover (above 10-km height) and precipitation over the tropics would appear for a $T_B \leq 235 \text{ K}$.

However, the T_B snapshots capture entire clouds, including rainless areas, rather than precipitation-band boundaries. Mapes and Houze (1993) introduced adjusted T_B thresholds to distinguish precipitable areas. From 180 to 265 K were commonly used to identify the area of TCCs. In the TCC area, below 200 K indicates cumulonimbus cores. The core region of TCCs can produce heavy precipitation, but cold cloud tops do not necessarily imply precipitation (Xu et al. 1999).

 Denotes content that is immediately available upon publication as open access.

Corresponding author: Yong-Sang Choi, ysc@ewha.ac.kr

DOI: 10.1175/JAMC-D-22-0175.1

© 2023 American Meteorological Society. This published article is licensed under the terms of the default AMS reuse license. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

Unauthenticated | Downloaded 04/27/24 06:23 PM UTC

Moreover, these empirical T_B thresholds depend on the region and seasonal conditions. If upper thin-ice clouds overlay the initial state of TCCs, the observed T_B may be lower than would be expected under single-layer conditions. In other words, the method of detecting the precipitating area of TCCs with only T_B thresholds involves large uncertainties.

Recent studies have used satellite retrieval of cloud properties to estimate precipitation probabilities in growing TCCs (Lin and Rossow 1996, hereinafter LR96; Hong et al. 2007; Rosenfeld et al. 2008; Mecikalski et al. 2011; Senf et al. 2015; Mecikalski et al. 2016; van Diedenhoven et al. 2016). Experimental and theoretical studies have shown that cloud properties, including cloud-top temperature (CTT), cloud optical thickness (COT), and cloud effective radius (CER), correlate with rainfall rates (Rosenfeld and Lensky 1998). CTT and COT are directly related to changes in the vertical growth of clouds. CER has been used to indicate the magnitude of the updraft and glaciation, which are prerequisites for the formation of raindrops in convective clouds. Using these properties, it is possible to estimate the probability of precipitation in TCCs. LR96 compared the amount of liquid water path between nonprecipitating and precipitating clouds using COT and CER. They discovered that precipitating clouds have more than 40 mg cm^{-2} in the liquid water path. Rosenfeld et al. (2008) revealed that, in CTT–CER relationships, the probability of precipitation increases rapidly when the CER exceeds $14 \text{ }\mu\text{m}$, and conversion from cloud water to rainwater becomes frequent at $\text{CER} \geq 25 \text{ }\mu\text{m}$. Mecikalski et al. (2011) and Senf et al. (2015) showed that both COT and precipitation probability increased as the cloud deepened. These studies provided sufficient potential to infer precipitation using cloud properties as discernible indicators.

A problem with this approach is that the collected TCC properties also include features of areas that may not be directly relevant to precipitation. The pixel-averaged cloud properties of an entire TCC differ from the state appearing in the convective core region of the TCC. This is because the convective core, representing the latest growth state from a vigorous updraft, occupies only a portion of the TCC total area. For this reason, previous studies have elaborated on how to obtain the properties that belong only to the core of the TCC. These tasks rely on human labor to extract features from TCC objects. However, various studies that have attempted to detect individual TCC objects by manual tracking (Rosenfeld and Lensky 1998; Vila et al. 2008; Feng et al. 2018) are hampered by time and space constraints. Recent studies attempted to effectively segment cloud masks from satellite images using deep learning methods (Nailussa'ada et al. 2018; Morales et al. 2018; Gonzales and Sakla 2019), but there is almost no research to classify individual cloud objects depending on cloud microphysics. Consequently, previous observational results on cloud–precipitation relationships have rarely been stable or reproduced.

This study aims to explore the cloud properties of TCCs that indicate a high probability of precipitation by training a neural network with daytime satellite imagery. We assumed that most of the TCCs are the mesoscale convective systems (MCS) over the tropics. MCS consists of cumulonimbus, nimbostratus, and anvil cirrus clouds. They are generally located in the tropics and account for most of the rainfall. Also, they

have a significant role in heat and energy transport (Chen and Houze 1997; Houze 2014). We propose an efficient two-step framework to identify conditions for cloud properties based on individual TCC objects with various stages of growth. In the proposed network, we segment South Korean geostationary satellite images to detect TCC objects in tropical Pacific regions. Next, we use a clustering method to classify TCCs with similar cloud properties. We also analyze the cloud properties of clusters with a high probability of precipitation. Section 2 describes the dataset used, followed by an explanation of the image segmentation and clustering method in section 3. Section 4 explains the clustering results and cloud property conditions. Section 5 discusses the specific conditions associated with a high probability of precipitation.

2. Data

a. GEO-KOMPSAT-2A products

GEO-KOMPSAT-2A (GK2A) is a geostationary satellite managed by the Korean Meteorological Administration (Choi and Ho 2015). Considering the field of view of GK2A, we decided on the study area between 80° – 175°E and 20°N – 20°S , known as the Indo-Pacific warm pool, where the sea surface temperature is relatively high throughout the year (Takayabu 1994). TCCs frequently occur in this domain, directly influencing tropical Pacific Ocean rainfall patterns (Sassen et al. 2008; De Decker 2016). We concentrated on the latitudes between 20°N and 20°S instead of the 30°N and 30°S zone (known as the tropical belt) to minimize the influence of convective clouds from the subtropics (Schumacher and Houze 2003).

The GK2A satellite uses the Advanced Meteorological Imager (AMI), which provides high-resolution imagery in 16 visible and infrared bands. This study focuses on AMI-retrieved cloud properties (CTT, COT, and CER) data at 2-km and 10-min resolutions. All of the retrieved AMI cloud properties are formed based on the varying reflectance and emittance of the liquid- and ice-phase clouds. The physical meaning and retrieval methods of the three cloud properties are described in detail in the following section.

The CTT is the temperature corresponding to the emission level near cloud tops. In the GK2A algorithm, the CTT was calculated considering the actual cloud emission rate from the double-infrared spectrum test. The liquid clouds assume an emissivity of 1 on a $10.4\text{-}\mu\text{m}$ band and estimate the CTT from the amount of observed radiation (Kim et al. 2019). For ice- and uncertain-phase clouds, additional emissivity calculated in the $11.2\text{-}\mu\text{m}$ band was used to estimate the range of CTTs, considering the various emission rates of ice particles (Strabala et al. 1994; Kim et al. 2019). In this study, we used CTT retrievals rather than T_B to consider the multilayered and thin-ice clouds. Choi et al. (2007) showed that a significant linear relationship between T_B and CTT is valid for optically thick clouds (i.e., $\text{COT} > 10$). For optically thin clouds, T_B is usually higher than CTT.

The basic principle involved in COT and CER retrieval is based on a bispectral solar reflectance method described by Nakajima and King (1990). Since then, retrieval methods have continued to develop, and the GK2A algorithm calculates the

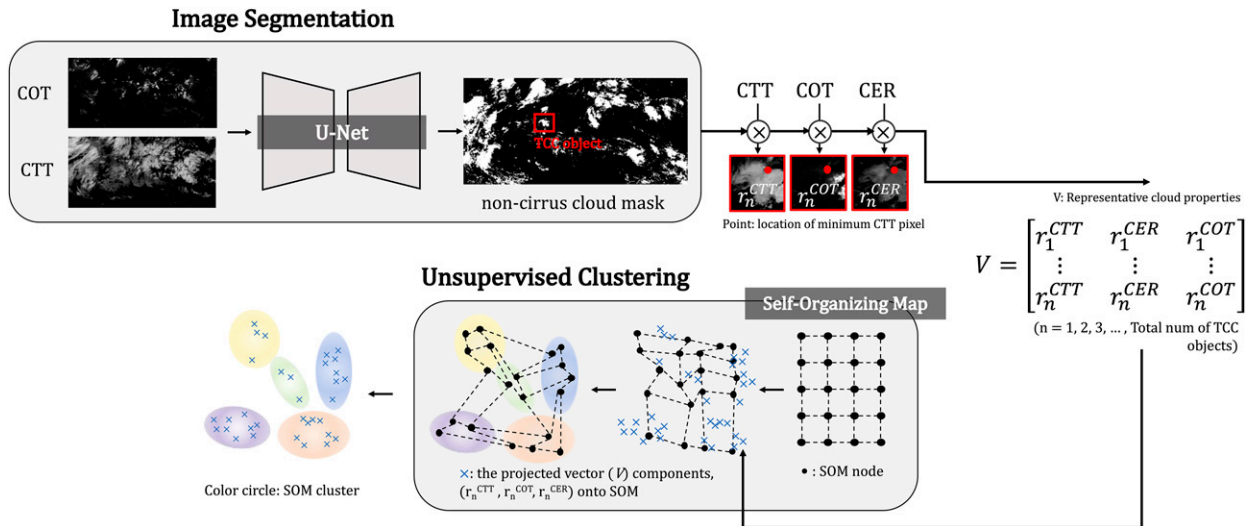


FIG. 1. Overview of the proposed PCCS network, including an image segmentation module and an unsupervised clustering module. In the image segmentation module, the CTT and COT of AMI/GK2A are used as input data for the U-Net model for noncirrus cloud pixel masking. The cloud mask from the image segmentation module is multiplied by each cloud property image to distinguish TCC objects [section 3a(1)]. Next, we detect each TCC object and extract cloud properties. In the unsupervised clustering module with a self-organizing map (SOM), we use a vector component (r_n^{CTT} , r_n^{COT} , r_n^{CER}) as the input dataset of SOM. After each TCC object is allocated to each SOM cluster index, we calculate a precipitation probability for each cluster [section 3a(2)].

solar radiation along various particle phases, such as the cloud optical properties algorithm of MODIS Collection 6 (Platnick et al. 2015). In the GK2A algorithm, the reflection function of clouds at a nonabsorbing visible band (0.64 and 0.87 μm) is primarily used as a COT function. The COT is defined as the optical thickness in a vertical profile from the cloud bottom to its maximum height and is often used as cloud optical depth in MODIS products. COT is quantified by the magnitude of extinction when light passes vertically through the clouds due to scattering and absorption by cloud droplets. The absorbing band in the near-infrared (1.6 μm) wavelength is used as a function of the CER, which represents the radiative characteristics of whole clouds composed of particles of several different sizes, taking into account the size of each cloud particle in a column. The CER is the ratio of the third to second moments of the particle-size distribution. In practice, an optimal estimation technique is used to determine the COT and CER with cloud reflectivity (Yang et al. 2013). The reflection, transmission, and absorption of clouds depend entirely on these two parameters, COT and CER.

The COT and CER are only available during the daytime because accuracy depends mainly on visible radiation. Accordingly, AMI cloud data retrieved during the day–night transition in the tropics were flagged as low quality. To deal with the data as clearly as possible, we used hourly AMI cloud products between 0200 and 0400 UTC (1100 and 1300 Korea standard time) from August 2020 through December 2021.

b. GPM products

This study also uses the Global Precipitation Measurement (GPM) product, intercalibrated with precipitation data from estimated infrared of partner satellites and gauge analysis. The GPM

was launched in early 2014 by the National Aeronautics and Space Administration and the Japan Aerospace Exploration Agency in advance of the Tropical Rainfall Measuring Mission (TRMM) satellite. In comparison with the TRMM, the GPM products provide more accurate rainfall estimates with greater spatiotemporal resolutions (Prakash et al. 2016). Among the Integrated Multi-Satellite Retrievals for GPM (IMERG) products, we used the “final” run product (IMERG-F), which was released approximately 3.5 months after observation time through the climatological correction with Global Precipitation Climatology Centre precipitation gauge analysis (Hou et al. 2014; Huffman et al. 2015).

The resolution of IMERG-F products was rescaled with the nearest-interpolation method to match the 2-km resolution of the AMI. IMERG-F provides *precipitationCal* data that include the precipitation rate (mm h^{-1}) at a resolution of $0.1^\circ \times 0.1^\circ$ (approximately 10 km) on a half-hourly time scale and a quasi-global coverage (60°N – 60°S latitude band) (Huffman et al. 2015). We calculated the average and maximum rain rate for each TCC object using IMERG-F products. TCC objects with an average rain rate of more than 0.1 mm h^{-1} were defined as precipitating clouds.

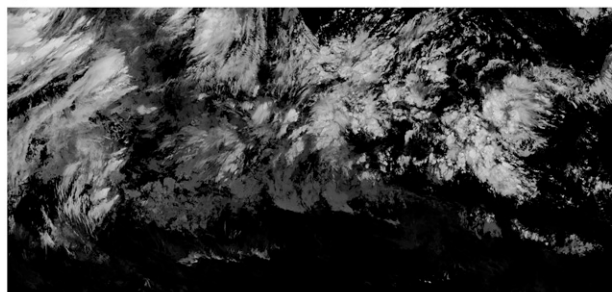
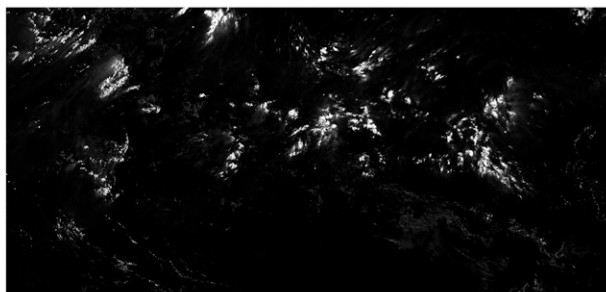
3. Methods

a. Network architecture

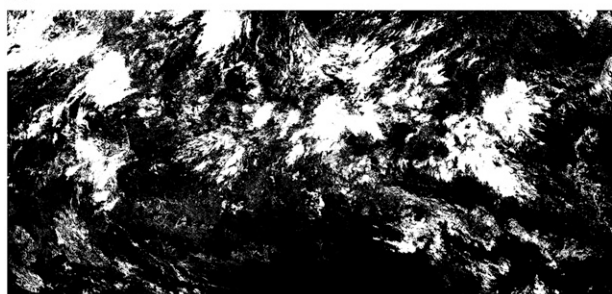
This section explains the details of the network architecture, as illustrated in Fig. 1.

1) DISTINGUISHING TCC OBJECTS FROM SEGMENTED SATELLITE IMAGERY

The image segmentation module in the “precipitating cloud classification using a self-organizing map (SOM)” (PCCS)

(a) CTT (at 8.6, 10.4, 11.2, 12.4, and 13.3 μm)(b) COT (at 0.64, 0.86, and 1.61 μm)

(c) Reference data for U-Net



(d) Predicted mask from U-Net

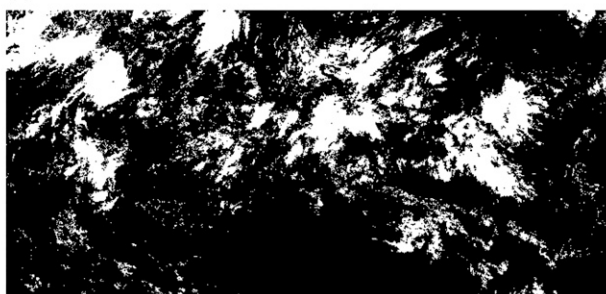


FIG. 2. The input data of U-Net: (a) CTT and (b) COT snapshots from *GK2A* (bands in parentheses are used) over the tropical warm pool (80° – 175°E and 20°N – 20°S) at 0300 UTC 4 Oct 2020. Bright pixels in (a) and (b) represent the highly grown cloud region. (c) Reference data for segmentation training; the cirrus clouds have been removed from the CTT image based on the ISCCP cloud classification thresholds. (d) The result of image segmentation by U-Net. In (c) and (d), the white cloud area and the black background area are shown as binary types (1, 0).

network distinguishes cloud pixels from the background and identifies the location of a single object. Each TCC object can consist of a single cloud cell or multiple cloud cells, reflecting varying sizes and growth states of TCCs. We divided input satellite imagery into multiple TCC objects and collected cloud properties used to discriminate among precipitation conditions in the subsequent clustering module. CTT, COT, and CER values at the location of the minimum CTT pixel were extracted as cloud properties of each TCC object. Here, the minimum CTT is assumed to be the latest physical property to indicate the TCC growth because the cloud growth is centered on the location where the updraft is the strongest.

Several studies have adopted U-Net architecture for satellite imagery segmentation (Gonzales and Sakla 2019; Guo et al. 2020; Trebing et al. 2021). We adopted U-Net (Ronneberger et al. 2015) architecture to generate a cloud mask map at the pixel level. As in Ronneberger et al. (2015), the network consisted of four encoder and four decoder blocks connected via bridges (e.g., skip connection) to avoid loss of information from layer to layer. The sigmoid activation function, added to a backend, estimated the predicted cloud mask classified at the pixel level.

Before segmentation, we generated preprocessing data that filtered the cirrus from CTT imagery based on International Satellite Cloud Climatology Project (ISCCP) cloud classification (Rossow and Schiffer 1999). Cirrus clouds, which appear in the upper atmosphere, are thin and comprise primarily ice particles. This type of extensive cloud cover can significantly

affect Earth's radiation but is unlikely to evolve into precipitating clouds. According to ISCCP-based thresholds, cirrus pixels with a cloud-top pressure (CTP) < 440 hPa and a COT < 3.6 were eliminated from CTT imagery. The preprocessed image was assigned a pixelwise class label (cloud and non-cloud) and used as the reference data for image segmentation.

For U-Net training, we chose CTT and COT images as “input” data and cirrus-free CTT images as “reference” data. The network was trained from scratch with a learning rate of 10^{-4} using an Adam optimizer and batch size 8 as described by Gonzales and Sakla (2019). We trained the network for 10 000 iterations using binary cross-entropy loss and dice loss functions. Both loss terms measure pixelwise differences between reference data and our predicted cloud mask. Figures 2a and 2b are input satellite imagery data, while Fig. 2c is used as the reference data for network training. The test result of the segmentation by the U-Net model achieved an average pixel accuracy of 97% in predicting cloud pixels, as shown in Fig. 2d. The resulting accuracy was superior to that reported by Gonzales and Sakla (2019).

2) UNSUPERVISED CLUSTERING OF GEOSTATIONARY SATELLITE CLOUD PROPERTIES

Machine learning requires having data labeled as the correct answer to learn. However, it is impossible to label all the clouds that appear in satellite imagery because we do not know the correct boundary of overlapped clouds. Learning on such unlabeled datasets is called unsupervised learning. Thus,

we exploited unsupervised learning for the tasks of this study. Clustering is one of the unsupervised machine learning approaches that automatically groups data samples with similar characteristics. *K*-means is a simple and popular clustering algorithm for intuitively retrieving data structure. However, the *K*-means method is rarely applied to satellite-retrieved cloud data, given the high computational costs and difficulties in visualizing topological data. In addition, we applied the *K*-means algorithm to our dataset. The results showed that the separation ability for dense data regions is worse than SOM (the result is not shown here). We, therefore, adopted SOM developed by Kohonen (1990) using the MiniSom library (Vettigli 2019), which has proven highly successful in clustering high-dimensional data without requiring prior knowledge of the features of the input data and the class labels of the output data (Zhang et al. 2018). In addition, the mapping process of SOM, which maps high-dimensional data to planes, can preserve the topology, meaning the distances between the original data. This allows the network to analyze the cloud properties of each cluster by tracing back the corresponding features in the original satellite product.

The values of cloud properties (CTT, COT, and CER) were extracted at the minimum CTT pixel position of each TCC object. We generated a SOM input vector consisting of the three cloud properties from the total number of all extracted TCC objects n . The dimension of the input data was $n \times 3$. Although sufficient TCC objects can be extracted from a single satellite image, we used monthly images to avoid overfitting the SOM results. We randomly selected the date of each month. Using these input vectors, each SOM first initializes the weight vector of each node in two-dimensional coordinates ($x \times y$) in the lattice. The best matching unit in any set of weight vectors in the training data was selected in every iteration. The network adjusted the weights of nearby nodes to be close to the input data point.

We set the sizes of the SOM nodes on the x and y axes of the target discretized space to 4 and 5, respectively. These sets are empirical values. As the number increases, the data become overly fragmented, making it difficult to find general characteristics. We set the sigma at 0.2 and performed 10 000 iterations. As a result, the SOM projected all TCC objects into 20 different clusters based on their unique characteristics extracted in minutes. In our network, the SOM was trained with a learning rate of 10^{-4} on all input data as an unsupervised learning technique.

b. Calculation of precipitation probability

We proposed a precipitation probability (PP) index that indicates the probability that a TCC in a cluster with specific cloud properties will have a rainfall rate $\geq 0.1 \text{ mm h}^{-1}$. Some studies (Hong et al. 2004; Kidder et al. 2005; Risyanto et al. 2019) have already defined and used various indices similar to PP (e.g., rain probability, probability of rain, tropical rainfall potential). Where there are differences, PP is calculated at the classified cluster level based on similar cloud properties learned by the SOM rather than individual TCC objects.

In the image segmentation module of PCCS, we detected individual TCC objects from the contour lines of adjacent

cloud mask pixels (Here, we used the Contour function of OpenCV, Python). First, we determined the average rainfall rate of each TCC object. We interpolated the IMERG-F data using the nearest-interpolation method. The closest IMERG-F data to the cloud pixel were considered the rainfall rate value of that pixel. The average rainfall rate was the sum of the rainfall rate of whole pixels in the individual TCC objects divided by the area. We assumed the TCC object had precipitation when the value exceeded 0.1 mm h^{-1} (Lau and Wu 2003; Tao et al. 2016).

We then calculated the PP for each cluster using Eq. (1). All TCC objects in the study domain are specified as one of the given numbers of SOM nodes. For clusters in each node, the PP was calculated as the ratio of the total number N of precipitating TCC objects to the total number of TCC objects. If the PP of the cluster appeared to be high, we presumed that TCC properties are highly correlated with precipitating. However, the number of SOM clusters and the criteria for separating them is set with empirical values and is, therefore, variable. We analyzed only the 25th and 75th ranges of physical properties appearing in each cluster to reduce the uncertainty:

$$\text{PP} = \frac{N(\text{Precipitating TCCs in a cluster})}{N(\text{TCCs in a cluster})} \times 100 \text{ (\%)}. \quad (1)$$

We attempted to determine the PP threshold, which the cluster was very likely precipitating. For the same purpose, LR96 used the cloud water path (WP, defined as the weight of the water droplets in the cloud), which was estimated from the COT and CER as follows: $\text{WP} = (0.6292 \times \text{COT} \times \text{CER})/10$. According to LR96, more than one-half of the precipitating clouds had WP values exceeding $40\text{--}50 \text{ mg cm}^{-2}$. Note that they only assumed liquid water particles in clouds.

We were able to obtain a new WP range with a specific COT and CER range (in Table 1) that was revealed by SOM clustering. As the PP threshold that can distinguish precipitating TCCs, a $\text{PP} \geq 70\%$ closely resembles the WP results of LR96 for precipitating clouds. We used this PP threshold in this study. Substituting the COT and CER values of the clusters satisfying PP_{70} into LR96's equation, the WP of TCCs with a precipitation probability of 70% or more appeared to be $40\text{--}300 \text{ mg cm}^{-2}$. This new WP range is wider than the results of LR96. This is probably because our data are not limited to liquid water but consider various water phases in TCCs.

4. Results

From the PCCS network, each TCC object was assigned to 20 cluster indices from 20 SOM nodes. In comparison with the original input data (Fig. 3a), the index-colored cloud masks in Fig. 3b are more discernable for each TCC object. In Fig. 3b, identically colored objects have the same cluster index, which means they have similar cloud properties. This simple method reduced the ambiguity of cloud cognition from a wide range of thin upper clouds. At the same time, small TCC objects that can develop into precipitating clouds were preserved. Approximately 700–800 TCC objects were detected in a single satellite image.

TABLE 1. Cloud properties and precipitation probabilities (PP) of SOM clusters in TCC samples extracted from 15 randomly selected satellite images each month from August 2020 to December 2021 over the tropical warm pool. The boldface numbers are for the clusters of high PP ($\geq 70\%$). The averaged max rainfall rate (mm h^{-1}) is obtained from IMERG-F and is defined as the mean of the maximum rain rate of TCCs in the same cluster. This was calculated with a bilinear interpolation process because the IMERG-F data have a coarser spatiotemporal resolution than the GK2A data. The range of cloud properties is between the 25th and 75th percentile values of each cluster. Cloud properties include cloud optical thickness (COT; no units), cloud effective radius (CER; μm), and cloud-top temperature (CTT; K).

Cluster index	Avg max rain rate (mm h^{-1})	PP (%)	COT range		CER range (μm)		CTT range (K)	
			25th	75th	25th	75th	25th	75th
1	0.34	24	29.8	39.7	10.9	17.6	279.2	286.1
2	2.06	73	16.2	21.7	31.1	37.8	212.8	225.3
3	1.18	52	4.3	9.3	46.6	54.7	217.5	235.6
4	2.72	75	18.8	27.3	42.7	52.9	215.8	235.6
5	0.71	44	4.3	8.1	24.1	28.6	216.2	229.6
6	0.42	31	0.0	0.1	0.0	0.4	249.8	271.4
7	2.73	67	24.2	30.4	25.1	29.3	208.9	279.2
8	0.65	43	4.3	8.1	32.1	28.5	217.1	228.7
9	0.71	50	14.9	21.1	18.7	24.7	238.8	283.2
10	0.29	26	3.1	7.5	7.1	11.3	240.8	276.2
11	0.25	21	13.1	19.9	8.8	13.1	263.7	285.7
12	7.22	89	46.6	63.1	28.2	37.8	203.1	223.1
13	4.91	81	93.4	113.5	28.3	37.6	208.4	238.1
14	5.24	85	32.3	39.2	27.3	34.3	204.5	218.4
15	0.49	37	54.7	72.4	10.2	20.5	274.1	284.2
16	0.52	38	3.1	6.8	16.1	20.5	221.2	254.6
17	3.31	73	3.1	9.9	65.1	75.6	217.1	235.2
18	14.91	96	130.3	149.4	30.2	42.9	197.3	222.3
19	2.06	53	85.9	110.2	13.1	27.8	250.8	284.4
20	10.41	83	70.5	97.2	46.5	69.4	199.6	239.7

During the study period, the total number of TCC objects obtained from 15 randomly selected images in each month was approximately 10 500 samples. Cloud property values were extracted from these TCC samples, input to SOM, and then assigned to 20 clusters. The COT–CER scatterplot for each clustered index is shown in Fig. 4.

In Fig. 4, each marker represents each TCC object. This graph only shows the relationship between COT and CER, but the CTT values also contributed an identical weight to the SOM clustering process (see Table 1). The same-colored marker represents the TCC objects assigned to the same cluster. Of the 20 SOM clusters, 8 clusters for a PP $\geq 70\%$ are denoted by the filled circles, and the “×” markers indicate 12 clusters with a PP lower than that. Most clusters with a high PP had shown more than 20 μm for CER values (clusters 2, 4, 12, 13, 14, 17, 18, and 20). These high-PP clusters sometimes had increased PP as COT increased, reaching 96% in the rightmost cluster (cluster 18). Also, clusters 12, 13, 14, 18, and 20, on the right side of the high-PP clusters, had a PP in excess of 80%. However, if the COT values of the 75th quantile of clusters fell below 30, the CER criteria for clusters meeting high PP should be progressively larger. On the left side, clusters 2, 4, and 17 with relatively thin (COT < 30) TCCs were inversely proportional between CER and COT, as reported earlier by Nakajima and King (1990). These clusters represent relatively low PP (<80%).

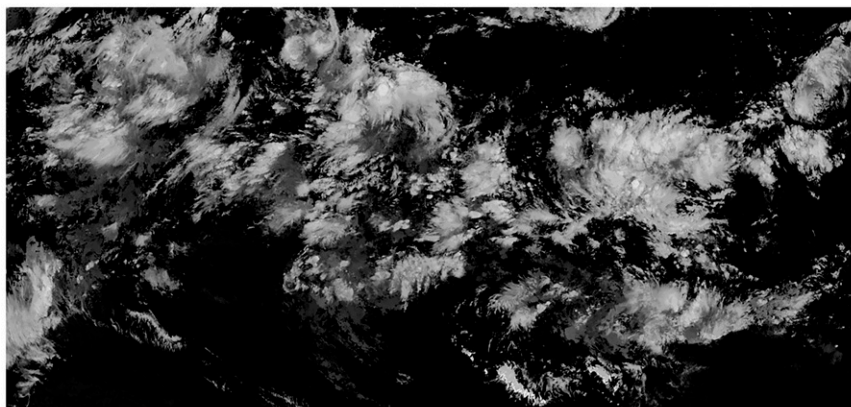
Statistics results for each cluster are provided in Table 1. The range of cloud properties was between the 25th and 75th

percentile of total TCC values corresponding to the same cluster. The averaged maximum rainfall rate was defined as the average of the maximum rainfall rate of each TCC object in the same cluster, and the values were calculated from IMERG-F. The clusters in bold in Table 1 indicate high-PP ($\geq 70\%$) clusters. In general, and with few exceptions, precipitating clouds have lower CTT ranges when compared with nonprecipitating clouds. Low-PP clusters fall in a CTT range of 215 to 235 K; however, the high-PP clusters are associated with 190 to 230 K. If the cloud CER was not large enough (<20 μm), the PP was small regardless of other cloud property values. For example, clusters 1 and 15 have sufficiently large COTs but with small CERs and high CTTs, indicating a low probability of precipitation.

Next, because the results of Fig. 4 and Table 1 are the average values of all input data, we identified the difference in cloud properties that appear in the high-PP clusters under various spatiotemporal conditions. Because countries within the study domain have different monsoon and dry seasons due to the intertropical convergence zone (ITCZ), we identified clusters satisfying the high-PP thresholds dependent on the location of ITCZ.

Figure 5a shows the results for the ITCZ facing south from December to March (DJFM), and Fig. 5b shows the results for the ITCZ facing north from June to September (JJAS). To test the SOM, 15 scenes were randomly selected from a given period of each seasonal condition. Coincidentally, the

(a) CTT at 2021/09/20 0300 UTC



(b) Cloud objects with cluster index at 2021/09/20 0300 UTC

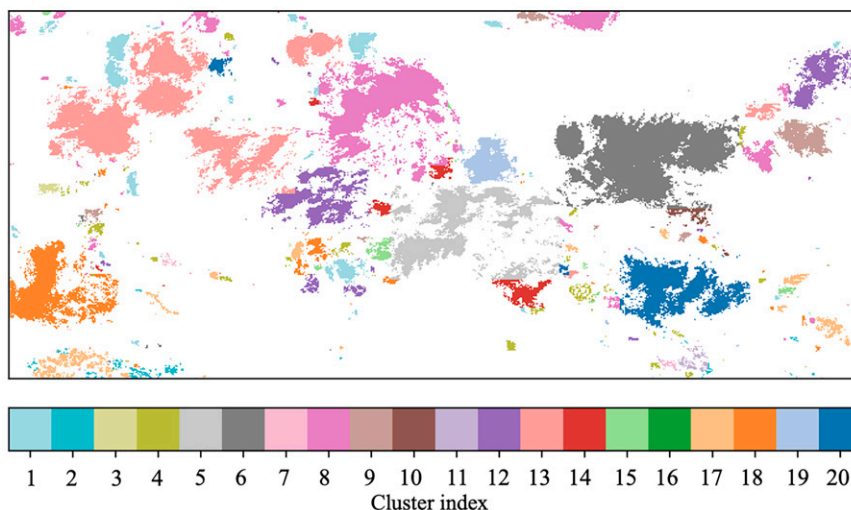


FIG. 3. (a) Normalized grayscale CTT imagery from *GK2A* (from 0 to 255) over the tropical warm pool at 0300 UTC 20 Sep 2021. (b) TCC objects to which the 20 SOM cluster indices are applied to the cloud mask of (a). The color means the TCC object is assigned to the same SOM node, in which TCCs have similar cloud microphysical properties. The cloud area in (b) has been removed from cirrus clouds as a result of the image segmentation module of PCCS.

pattern in which the cluster has a high PP ($\geq 70\%$) under both conditions has a shape similar to that in Fig. 4. Also, the overall scattered distribution of TCC objects is identical. Under seasonal constraints, the highest-PP clusters appear at the far right of the distribution. In the DJFM season, cluster 3 has an 89% PP and 11.35 mm h^{-1} of averaged maximum rainfall, and cluster 1 in the JJAS season has a 96% PP and 16.26 mm h^{-1} of averaged maximum rainfall.

We also compared cloud properties according to where the clouds formed, finding distributions similar to previous results. Takayabu et al. (1994) and Rosenfeld and Lensky (1998) demonstrated that continental and maritime clouds have different microphysical precipitation-forming processes. Park et al. (2007) reported that cases of lightning and low- T_B clouds were detected more frequently over land than over the ocean. To confirm the difference at both locations, TCC objects were

sorted into continental and maritime clouds using a land-sea mask of *GK2A*. The SOM test results for the continental-maritime dataset are shown in Fig. 6. Because the cases of clouds over land and ocean were not considered here, the number of available TCC samples for clustering was lower than that of previous experiments.

Under both locational conditions, the pattern of high-PP ($\geq 70\%$) clusters appearing as CER is larger than $20 \mu\text{m}$, similar to previous results. However, the highest-PP clusters had a relatively small COT in cluster 19 (Figs. 6a,b). For continental clouds (Fig. 6a), the PP was 93% for cluster 19, where the COT was between 40 and 50. However, cluster 19, with 82% PP in maritime clouds (Fig. 6b), had a COT of 30–40, a slightly lower COT than the continental cloud. The CER variance of the high-PP clusters was larger in maritime clouds than in continental clouds.

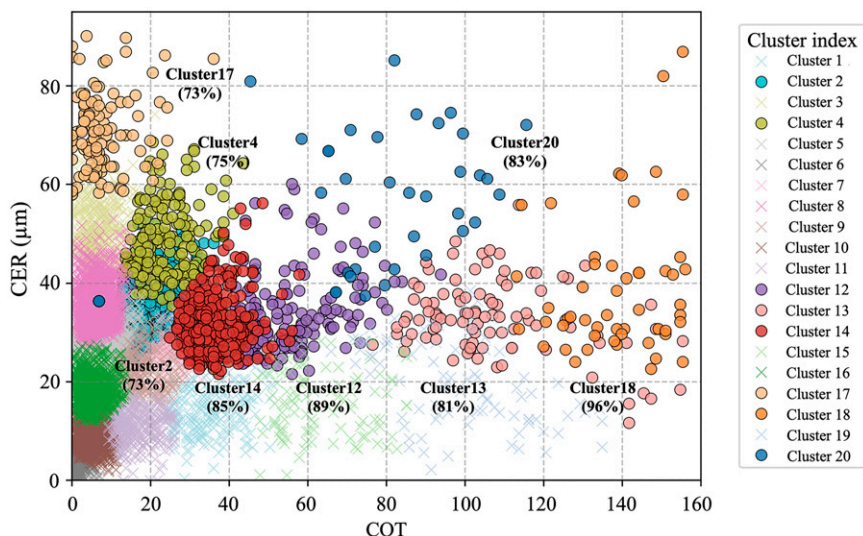


FIG. 4. A scatterplot of COT and CER from the monthly random dataset for October 2020 to December 2021 in the tropical warm pool. Each symbol means each TCC object, located at the extracted COT and CER values from a minimum CTT pixel. The TCC objects in a cluster that meets the PP criterion of 70% are denoted by the colored circles [we also marked the cluster number and PP (%)]. The “x” symbols are TCC objects of the cluster for PP < 70%. Each color represents a specific cluster index, meaning the same-colored TCCs have similar microphysical properties.

5. Conclusions and discussion

This study revealed the relationship among the cloud properties (CTT, COT, and CER) of precipitating TCCs at various growth stages. Although we collected TCC objects under different spatiotemporal conditions, TCCs with high precipitation probabilities appeared in the specific condition of cloud properties. First, as shown in Figs. 4–6, high-precipitation probability ($\text{PP} \geq 70\%$) clusters commonly satisfy $\text{CER} \geq 20 \mu\text{m}$ and $\text{TCC} \leq 230 \text{ K}$ conditions regardless of COT. However, relatively thin TCCs ($\text{COT} < 30$) certainly have much larger CER than $20 \mu\text{m}$ in order to have a high PP. On the other hand, relatively thick TCCs ($\text{COT} \geq 30$) showed higher PP than thin TCCs in general. This result corresponds with a report by van Diedenhoven et al. (2014), who noted that convective clouds over the tropics have CER values of $28\text{--}36 \mu\text{m}$. In addition, several studies have shown that one-half of the cases keep or decrease the CER during the continuous growth of convective clouds (Mecikalski et al. 2011; Senf et al. 2015; van Diedenhoven et al. 2016; van Diedenhoven 2021). The decrease and congestion of CER with increasing COT in Fig. 4 can be accounted for by ice splinter from riming or particle sedimentation from fallout in the cloud top. This indicates that the high-PP ($\geq 70\%$) clusters are either actively growing with vigorous convection or have already reached the mature stage of TCCs.

Furthermore, when the COT or CER satisfies the $\text{CTT} \leq 230 \text{ K}$ condition at the same time, that cluster shows the highest probability of rain. Although few objects with $\text{PP} \geq 70\%$ were considered a precipitating TCC by the LR96 equation, tracking results showed that these TCCs occupied more than half of the total cloud area over the tropics and were

accompanied by a heavy rainfall rate, with maximum rates exceeding 6 mm h^{-1} (Yuan and Houze 2010). However, the precipitating area is just a few parts of the giant TCC. When using only the T_B threshold method, it was difficult to distinguish the actual precipitation area from the cold clouds. In this context, the microphysical conditions revealed in this study can help determine precipitation areas strictly between TCCs covered with extensive anvil cirrus clouds.

Regardless of spatiotemporal conditions, cloud properties for high-PP TCCs are shown as general characteristics over the tropical warm pool region. However, the differences in the graph under various spatiotemporal conditions are noteworthy, and two of them warrant discussion.

First, there were differences in the cloud microphysical properties of high-PP clusters depending on seasonal conditions. In the DJFM season shown in Fig. 5a, the CER decreased slightly as the COT increased in high-PP clusters. The cluster with the largest PP among all clusters was in the most extensive COT range. The CER range for cluster 3 during the DJFM season is $25.26\text{--}33.56 \mu\text{m}$, lower than the range for cluster 1 of the JJAS season ($31.26\text{--}41.33 \mu\text{m}$). In other words, TCCs of the DJFM have smaller CER values when compared with JJAS under the same COT conditions. Rosenfeld et al. (2008) revealed that a small CER value indicates the presence of a strong updraft in clouds. We can therefore presume that the DJFM season has a stronger updraft than the JJAS season.

Second, cloud characteristics over land may involve aerosol effects. Previous studies have shown that aerosols may cause other physical processes of precipitation formation in both regions (Rosenfeld and Woodley 2003). In general, land has a higher concentration of aerosols released from more pollutants

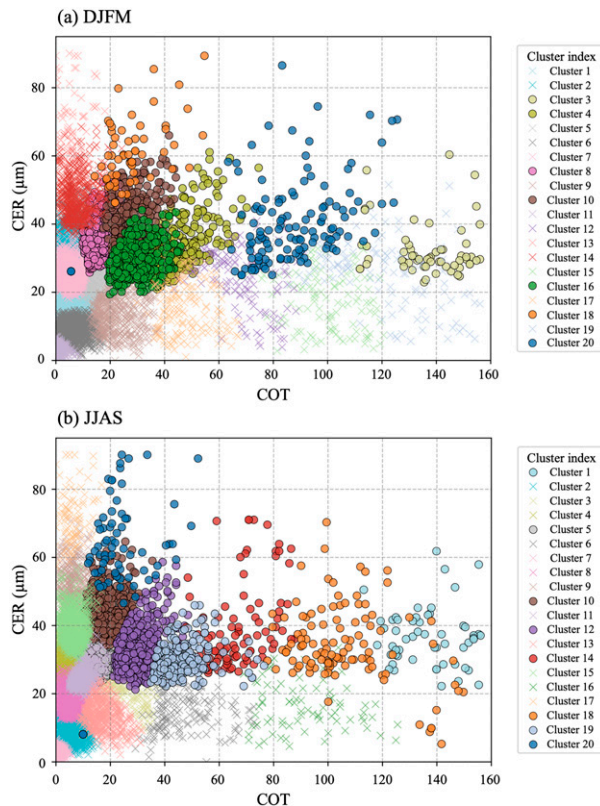


FIG. 5. As in Fig. 4, but for the seasonal condition: scatterplots of the random dataset obtained (a) during the DJFM 2020 and 2021 period and (b) for the JJAS 2020 and 2021 period. More than eight clusters indicated by filled circles satisfy the high-PP ($\geq 70\%$) condition.

than maritime air, and continental air has a higher cloud condensation nuclei concentration. Continental clouds are made up of many small droplets, and conversely, maritime clouds are generated with relatively large and fewer droplets. Continental clouds, therefore, induce glaciation at a lower temperature in combination with updrafts caused by strong convection on land. Because the clusters—a total of 19 in Figs. 6a and 6b—of the highest PP have a relatively small CER, we can infer that the convection on land is relatively strong (Rosenfeld et al. 2008).

Meanwhile, note the uncertainty of data and model. We used the IMERG-F products to address the lack of observation equipment in the tropics. Several studies have evaluated the performance of the IMERG-F products in comparison with other observation products over the tropics (Liu 2016; Sunilkumar et al. 2019). Results showed significantly superior performance when compared with ground-gauge data, except for underestimated extreme precipitation (Xu et al. 2017). Ramadhan et al. (2022) showed that IMERG-F products tended to smooth out extreme rainfall values at an hourly time scale. However, we did not focus on extreme precipitation and only evaluated the least precipitation in the TCC. This limitation had little impact on the conclusions of this study. In addition, about our PCCS network, the accuracy of clustering depends on the accuracy of TCC detection. We used a contour

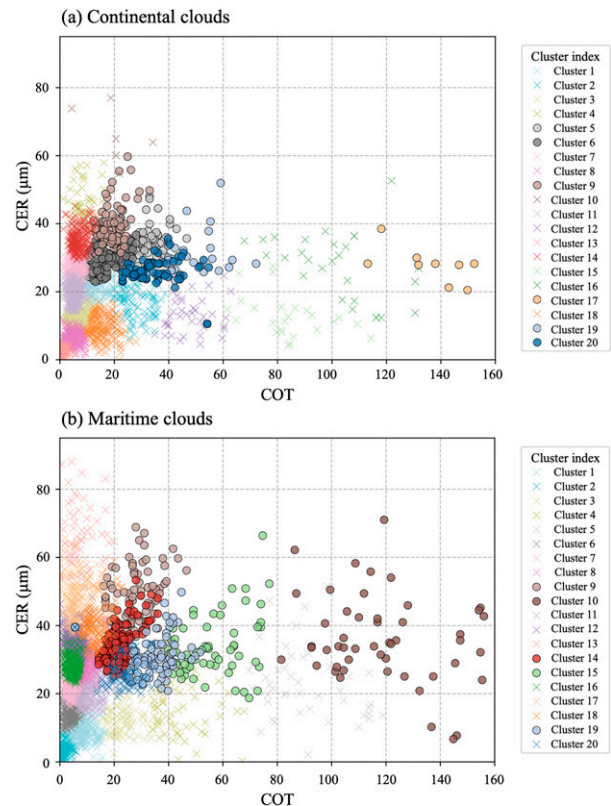


FIG. 6. As in Fig. 4 but for topological conditions: (a) continental clouds and (b) maritime clouds during the study period. Clouds that span continents and oceans are not included.

line-based classification method relying on pixel contiguity. However, this approach may result in errors with overlapping clouds. In a future study, we need more elaborate detection methods and reasonable classification policies for ambiguous cloud boundaries.

The results are expected to help detect and predict precipitation in tropical regions by clarifying the conditions of cloud properties for high precipitation probabilities (Platt 1989; Hu and Stamnes 2000; Grabowski et al. 2019). Because TCCs directly impact tropical monsoons and can cause flooding, nowcasting to provide early warning is required. Through this study, TCCs satisfying the conditions of SOM clusters with high PPs can be detected in the target region. Because the experimental method is simple and insensitive, it can be reproduced in any region with such satellite images with the same cloud products (CTT, COT, and CER). Our results could be used to improve the initial conditions of numerical weather prediction (NWP) models or other models that forecast tropical rainfall. Since rainfall rates can be turned into latent heat rates, NWP models could benefit other routinely operated forecast algorithms.

Acknowledgments. This study was supported by a National Research Foundation of Korea (NRF) grant funded by the South Korean government (MSIT) (2021R1A2C1093402). The

first author was supported by the Hyundai Motor Chung Mong-Koo Foundation.

Data availability statement. This study uses publicly available datasets. The GK2A products used during this study are available from KMA NMSC (<http://datasvc.nmsc.kma.go.kr/datasvc/html/main/main.do?lang=en>). The GPM IMERG datasets used in this study are available from NASA (<https://gpm.nasa.gov/data/directory>).

REFERENCES

- Arkin, P. A., 1979: The relationship between fractional coverage of high cloud and rainfall accumulations during GATE over the B-scale array. *Mon. Wea. Rev.*, **107**, 1382–1387, [https://doi.org/10.1175/1520-0493\(1979\)107<1382:TRBFCO>2.0.CO;2](https://doi.org/10.1175/1520-0493(1979)107<1382:TRBFCO>2.0.CO;2).
- Chen, S. S., and R. A. Houze Jr., 1997: Diurnal variation and life-cycle of deep convective systems over the tropical Pacific warm pool. *Quart. J. Roy. Meteor. Soc.*, **123**, 357–388, <https://doi.org/10.1002/qj.49712353806>.
- Choi, Y.-S., and C.-H. Ho, 2015: Earth and environmental remote sensing community in South Korea: A review. *Remote Sens. Appl. Soc. Environ.*, **2**, 66–76, <https://doi.org/10.1016/j.rsase.2015.11.003>.
- , —, M.-H. Ahn, and Y.-M. Kim, 2007: An exploratory study of cloud remote sensing capabilities of the Communication, Ocean and Meteorological Satellite (COMS) imagery. *Int. J. Remote Sens.*, **28**, 4715–4732, <https://doi.org/10.1080/0143160701264235>.
- De Deckker, P., 2016: The Indo-Pacific warm pool: Critical to world oceanography and world climate. *Geosci. Lett.*, **3**, 20, <https://doi.org/10.1186/s40562-016-0054-3>.
- Feng, Z., L. R. Leung, R. A. Houze Jr, S. Hagos, J. Hardin, Q. Yang, B. Han, and J. Fan, 2018: Structure and evolution of mesoscale convective systems: Sensitivity to cloud microphysics in convection-permitting simulations over the United States. *J. Adv. Model. Earth Syst.*, **10**, 1470–1494, <https://doi.org/10.1029/2018MS001305>.
- Gonzales, C., and W. Sakla, 2019: Semantic segmentation of clouds in satellite imagery using deep pre-trained U-Nets. *2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, Washington, DC, Institute of Electrical and Electronics Engineers, 1–7, <https://doi.org/10.1109/AIPR47015.2019.9174594>.
- Grabowski, W. W., H. Morrison, S.-I. Shima, G. C. Abade, P. Dziekan, and H. Pawlowska, 2019: Modeling of cloud microphysics: Can we do better? *Bull. Amer. Meteor. Soc.*, **100**, 655–672, <https://doi.org/10.1175/BAMS-D-18-0005.1>.
- Guo, Y., X. Cao, B. Liu, and M. Gao, 2020: Cloud detection for satellite imagery using attention-based U-Net convolutional neural network. *Symmetry*, **12**, 1056, <https://doi.org/10.3390/sym12061056>.
- Hong, G., P. Yang, H.-L. Huang, B. A. Baum, Y. Hu, and S. Platnick, 2007: The sensitivity of ice cloud optical and microphysical passive satellite retrievals to cloud geometrical thickness. *IEEE Trans. Geosci. Remote Sens.*, **45**, 1315–1323, <https://doi.org/10.1109/TGRS.2007.894549>.
- Hong, Y., K.-L. Hsu, S. Sorooshian, and X. Gao, 2004: Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *J. Appl. Meteor.*, **43**, 1834–1853, <https://doi.org/10.1175/JAM2173.1>.
- Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, <https://doi.org/10.1175/BAMS-D-13-00164.1>.
- Houze, R. A., Jr., 2014: *Cloud Dynamics*. Academic Press, 496 pp.
- Hu, Y., and K. Stamnes, 2000: Climate sensitivity to cloud optical properties. *Tellus*, **52B**, 81–93, <https://doi.org/10.3402/tellusb.v52i1.16084>.
- Huffman, G. J., D. T. Bolvin, E. J. Nelkin, and J. Tan, 2015: Integrated Multi-satellite Retrievals for GPM (IMERG) technical documentation. NASA/GSFC Code 612 Tech. Doc., 48 pp., http://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf.
- Kidder, S. Q., J. A. Knaff, S. J. Kusselson, M. Turk, R. R. Ferraro, and R. J. Kuligowski, 2005: The tropical rainfall potential (TRaP) technique. Part I: Description and examples. *Wea. Forecasting*, **20**, 456–464, <https://doi.org/10.1175/WAF860.1>.
- Kim, H.-S., B. A. Baum, and Y.-S. Choi, 2019: Use of spectral cloud emissivities and their related uncertainties to infer ice cloud boundaries: Methodology and assessment using CALIPSO cloud products. *Atmos. Meas. Tech.*, **12**, 5039–5054, <https://doi.org/10.5194/amt-12-5039-2019>.
- Kohonen, T., 1990: The self-organizing map. *Proc. IEEE*, **78**, 1464–1480, <https://doi.org/10.1109/5.58325>.
- Lau, K. M., and H. T. Wu, 2003: Warm rain processes over tropical oceans and climate implications. *Geophys. Res. Lett.*, **30**, 2290, <https://doi.org/10.1029/2003GL018567>.
- Lin, B., and W. B. Rossow, 1996: Seasonal variation of liquid and ice water path in nonprecipitating clouds over oceans. *J. Climate*, **9**, 2890–2902, [https://doi.org/10.1175/1520-0442\(1996\)009<2890:SVOLAI>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<2890:SVOLAI>2.0.CO;2).
- Liu, Z., 2016: Comparison of Integrated Multisatellite Retrievals for GPM (IMERG) and TRMM Multisatellite Precipitation Analysis (TMPA) monthly precipitation products: Initial results. *J. Hydrometeorol.*, **17**, 777–790, <https://doi.org/10.1175/JHM-D-15-0068.1>.
- Mapes, B. E., and R. A. Houze Jr., 1993: Cloud clusters and superclusters over the oceanic warm pool. *Mon. Wea. Rev.*, **121**, 1398–1416, [https://doi.org/10.1175/1520-0493\(1993\)121<1398:CCASOT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1993)121<1398:CCASOT>2.0.CO;2).
- Mecikalski, J. R., P. D. Watts, and M. Koenig, 2011: Use of Meteosat Second Generation optimal cloud analysis fields for understanding physical attributes of growing cumulus clouds. *Atmos. Res.*, **102**, 175–190, <https://doi.org/10.1016/j.atmosres.2011.06.023>.
- , D. Rosenfeld, and A. Manzato, 2016: Evaluation of geostationary satellite observations and the development of a 1–2 h prediction model for future storm intensity. *J. Geophys. Res. Atmos.*, **121**, 6374–6392, <https://doi.org/10.1002/2016JD024768>.
- Morales, G., S. G. Huamán, and J. Telles, 2018: Cloud detection in high-resolution multispectral satellite imagery using deep learning. *Artificial Neural Networks and Machine Learning—ICANN 2018*, V. Kůrková et al., Eds., Lecture Notes in Computer Science, Vol. 11141, Springer, 280–288, https://doi.org/10.1007/978-3-030-01424-7_28.
- Nailussa'ada, T. Harsono, and A. Basuki, 2018: Cloud satellite image segmentation using Meng Hee Heng *k*-means and DBSCAN clustering. *2018 Int. Electronics Symp. on Knowledge Creation and Intelligent Computing (IES-KCIC)*, Bali, Indonesia, Institute of Electrical and Electronics Engineers, 367–371, <https://doi.org/10.1109/KCIC.2018.8628523>.
- Nakajima, T., and M. D. King, 1990: Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory. *J. Atmos. Sci.*,

- 47, 1878–1893, [https://doi.org/10.1175/1520-0469\(1990\)047<1878:DOTOTA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1990)047<1878:DOTOTA>2.0.CO;2).
- Park, M.-S., Y.-S. Choi, C.-H. Ho, C.-H. Sui, S. K. Park, and M.-H. Ahn, 2007: Regional cloud characteristics over the tropical northwestern Pacific as revealed by Tropical Rainfall Measuring Mission (TRMM) Precipitation Radar and TRMM Microwave Imager. *J. Geophys. Res.*, **112**, D05209, <https://doi.org/10.1029/2006JD007437>.
- Platnick, S., and Coauthors, 2015: MODIS cloud optical properties: User guide for the Collection 6 Level-2 MOD06/MYD06 product and associated Level-3 datasets, version 1.0. NASA, 145 pp., https://modis-images.gsfc.nasa.gov/_docs/C6MOD06OPUserGuide.pdf.
- Platt, C. M. R., 1989: The role of cloud microphysics in high-cloud feedback effects on climate change. *Nature*, **341**, 428–429, <https://doi.org/10.1038/341428a0>.
- Prakash, S., A. K. Mitra, D. S. Pai, and A. AghaKouchak, 2016: From TRMM to GPM: How well can heavy rainfall be detected from space? *Adv. Water Resour.*, **88**, 1–7, <https://doi.org/10.1016/j.advwatres.2015.11.008>.
- Ramadhan, R., and Coauthors, 2022: Evaluation of GPM IMERG performance using gauge data over Indonesian Maritime Continent at different time scales. *Remote Sens.*, **14**, 1172, <https://doi.org/10.3390/rs14051172>.
- Rentschler, J., and M. Salhab, 2020: People in harm's way: Flood exposure and poverty in 189 countries. The World Bank Policy Research Working Paper 9447, 28 pp., <https://documents1.worldbank.org/curated/en/669141603288540994/pdf/People-in-Harms-Way-Flood-Exposure-and-Poverty-in-189-Countries.pdf>.
- Risyanto, F. Lasmono, and G. Nugroho, 2019: Identification of rainfall area in Indonesia using infrared channels of Himawari-8 Advance Himawari Imager (AHI). *IOP Conf. Ser. Earth Environ. Sci.*, **303**, 012057, <https://doi.org/10.1088/1755-1315/303/1/012057>.
- Ronneberger, O., P. Fischer, and T. Brox, 2015: U-Net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, N. Navab et al., Eds., Lecture Notes in Computer Science, Vol. 9351, Springer, 234–241, https://doi.org/10.1007/978-3-319-24574-4_28.
- Rosenfeld, D., and I. M. Lensky, 1998: Satellite-based insights into precipitation formation processes in continental and maritime convective clouds. *Bull. Amer. Meteor. Soc.*, **79**, 2457–2476, [https://doi.org/10.1175/1520-0477\(1998\)079<2457:SBIIIPF>2.0.CO;2](https://doi.org/10.1175/1520-0477(1998)079<2457:SBIIIPF>2.0.CO;2).
- , and W. L. Woodley, 2003: Closing the 50-year circle: From cloud seeding to space and back to climate change through precipitation physics. *Cloud Systems, Hurricanes, and the Tropical Rainfall Measuring Mission (TRMM)*, Meteor. Monogr., No. 51, Amer. Meteor. Soc., 59–80.
- , —, A. Lerner, G. Kelman, and D. T. Lindsey, 2008: Satellite detection of severe convective storms by their retrieved vertical profiles of cloud particle effective radius and thermodynamic phase. *J. Geophys. Res.*, **113**, D04208, <https://doi.org/10.1029/2007JD008600>.
- Rossow, W. B., and R. A. Schiffer, 1999: Advances in understanding clouds from ISCCP. *Bull. Amer. Meteor. Soc.*, **80**, 2261–2288, [https://doi.org/10.1175/1520-0477\(1999\)080<2261:AIUCFI>2.0.CO;2](https://doi.org/10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2).
- Sassen, K., Z. Wang, and D. Liu, 2008: Global distribution of cirrus clouds from CloudSat/Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) measurements. *J. Geophys. Res.*, **113**, D00A12, <https://doi.org/10.1029/2008JD009972>.
- Schumacher, C., and R. A. Houze Jr., 2003: Stratiform rain in the tropics as seen by the TRMM precipitation radar. *J. Climate*, **16**, 1739–1756, [https://doi.org/10.1175/1520-0442\(2003\)016<1739:SRITTA>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<1739:SRITTA>2.0.CO;2).
- Senf, F., F. Dietzsch, A. Hünnerbein, and H. Deneke, 2015: Characterization of initiation and growth of selected severe convective storms over central Europe with MSG-SEVIRI. *J. Appl. Meteor. Climatol.*, **54**, 207–224, <https://doi.org/10.1175/JAMC-D-14-0144.1>.
- Spencer, R. W., D. W. Martin, B. B. Hinton, and J. A. Weinman, 1983: Satellite microwave radiances correlated with radar rain rates over land. *Nature*, **304**, 141–143, <https://doi.org/10.1038/304141a0>.
- Strabala, K. I., S. A. Ackerman, and W. P. Menzel, 1994: Cloud properties inferred from 8–12- μ m data. *J. Appl. Meteor.*, **33**, 212–229, [https://doi.org/10.1175/1520-0450\(1994\)033<0212:CPIFD>2.0.CO;2](https://doi.org/10.1175/1520-0450(1994)033<0212:CPIFD>2.0.CO;2).
- Sunilkumar, K., A. Yatagai, and M. Masuda, 2019: Preliminary evaluation of GPM-IMERG rainfall estimates over three distinct climate zones with APHRODITE. *Earth Space Sci.*, **6**, 1321–1335, <https://doi.org/10.1029/2018EA000503>.
- Takayabu, Y. N., 1994: Large-scale cloud disturbances associated with equatorial waves. Part I: Spectral features of the cloud disturbances. *J. Meteor. Soc. Japan*, **72**, 433–449, https://doi.org/10.2151/jmsj1965.72.3_433.
- , T. Nitta, and W. K. M. Lau, 1994: Large-scale cloud systems associated with equatorial waves—Climatology and TOGA-COARE. *Proc. Int. Conf. on Monsoon Variability and Prediction*, Trieste, Italy, International Centre for Theoretical Physics, WMO, 296–303, https://library.wmo.int/doc_num.php?explnum_id=11231.
- Tao, Y., X. Gao, A. Ihler, K. Hsu, and S. Sorooshian, 2016: Deep neural networks for precipitation estimation from remotely sensed information. *2016 IEEE Congress on Evolutionary Computation (CEC)*, Vancouver, BC, Canada, Institute of Electrical and Electronics Engineers, 1349–1355, <https://doi.org/10.1109/CEC.2016.7743945>.
- Trebing, K., T. Stańczyk, and S. Mehrkanon, 2021: SmaAt-UNet: Precipitation nowcasting using a small attention-UNet architecture. *Pattern Recognit. Lett.*, **145**, 178–186, <https://doi.org/10.1016/j.patrec.2021.01.036>.
- van Diedenhoven, B., 2021: Variation of ice microphysical properties with temperature and humidity at tops of convective clouds. *Geophys. Res. Lett.*, **48**, e2021GL093673, <https://doi.org/10.1029/2021GL093673>.
- , A. M. Fridlind, B. Cairns, and A. S. Ackerman, 2014: Variation of ice crystal size, shape, and asymmetry parameter in tops of tropical deep convective clouds. *J. Geophys. Res. Atmos.*, **119**, 11 809–11 825, <https://doi.org/10.1002/2014JD022385>.
- , —, —, —, and J. E. Yorks, 2016: Vertical variation of ice particle size in convective cloud tops. *Geophys. Res. Lett.*, **43**, 4586–4593, <https://doi.org/10.1002/2016GL068548>.
- Vettigli, G., 2019: MiniSom: minimalistic and NumPy-based implementation of the self organizing map. Release 2.1.5. 2019, GitHub, accessed on 4 August 2021, <https://github.com/JustGlowing/minisom>.
- Vila, D. A., L. A. T. Machado, H. Laurent, and I. Velasco, 2008: Forecast and Tracking the Evolution of Cloud Clusters (ForTraCC) using satellite infrared imagery: Methodology and validation. *Wea. Forecasting*, **23**, 233–245, <https://doi.org/10.1175/2007WAF2006121.1>.

- Xu, L., S. Sorooshian, X. Gao, and H. V. Gupta, 1999: A cloud-patch technique for identification and removal of no-rain clouds from satellite infrared imagery. *J. Appl. Meteor.*, **38**, 1170–1181, [https://doi.org/10.1175/1520-0450\(1999\)038<1170:ACPTFI>2.0.CO;2](https://doi.org/10.1175/1520-0450(1999)038<1170:ACPTFI>2.0.CO;2).
- Xu, R., F. Tian, L. Yang, H. Hu, H. Lu, and A. Hou, 2017: Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network. *J. Geophys. Res. Atmos.*, **122**, 910–924, <https://doi.org/10.1002/2016JD025418>.
- Yang, P., L. Bi, B. A. Baum, K.-N. Liou, G. W. Kattawar, M. I. Mishchenko, and B. Cole, 2013: Spectrally consistent scattering, absorption, and polarization properties of atmospheric ice crystals at wavelengths from 0.2 to 100 μm . *J. Atmos. Sci.*, **70**, 330–347, <https://doi.org/10.1175/JAS-D-12-039.1>.
- Yuan, J., and R. A. Houze Jr., 2010: Global variability of mesoscale convective system anvil structure from A-Train satellite data. *J. Climate*, **23**, 5864–5888, <https://doi.org/10.1175/2010JCLI3671.1>.
- Yuter, S. E., and R. A. Houze Jr., 1998: The natural variability of precipitating clouds over the western Pacific warm pool. *Quart. J. Roy. Meteor. Soc.*, **124**, 53–99, <https://doi.org/10.1002/qj.49712454504>.
- Zhang, W., J. Wang, D. Jin, L. Oreopoulos, and Z. Zhang, 2018: A deterministic self-organizing map approach and its application on satellite data based cloud type classification. *2018 IEEE Int. Conf. on Big Data (Big Data)*, Seattle, WA, Institute of Electrical and Electronics Engineers, 2027–2034, <https://doi.org/10.1109/BigData.2018.8622558>.