Dear Editor,

We are grateful for the careful reading, helpful comments, and constructive suggestions of the reviewers, which have allowed us to clarify and improve the manuscript. The three reviewers have provided valuable insights from different perspectives. Reviewer #1 focused on the quality and application of the dataset, particularly the selection and presentation of the data. Reviewer #2 emphasized the quality of the training, validation and testing datasets, as well as the reliability assessment of the nighttime results. Reviewer #3 concentrated on the model's training and validation, as well as the application of the cloud dataset.

In response to their comments, we made the following main changes and improvements in our manuscript:

1. We have conducted a sensitivity experiment comparing the TIR-CNN-based COT with the MODIS daytime COT, and the results are now included in the revised manuscript to strengthen our cloud dataset.

2. All the NumPy format files were converted to HDF format to enhance the cross-platform and cross-software accessibility of our dataset.

3. We have added a new sub-section in Section 2 to describe the dataset contents, including variables, their calculation methods, and the file nomenclature, along with an additional table to summarize.

4. We validated the representativeness of our training dataset by examining the differences in the probability density functions (PDFs) of thermal radiance, cloud optical thickness (COT), and cloud morphology between our training dataset and a global full-year dataset. Additionally, we also compared the PDFs of these variables between our training dataset and a nighttime dataset to further verify the reliability of our nighttime results.

5. The detailed descriptions of the validation method and model performance metrics were added in the manuscript, including precision, F1-score, and recall, to enhance the reliability and robustness of our model.

Although some of the improvements suggested by the reviewers are not reflected in this paper, many of the directions they pointed out are being actively pursued by other members of our team. This cloud dataset serves as a foundational work, and we hope that, under its guidance, a series of studies will be conducted in this direction, encouraging the community to pay more attention to the research in this field.

Below we addressed the reviewers' comments, with the reviewer comments in black and our response in blue. The revised sentences in the manuscript are indicated in italics.

Yannian Zhu

On behalf of the co-authors

Reviewer #1

Wu et al. present a global daytime and nighttime low cloud morphology dataset classified based on deep learning methods. The work builds on the algorithms of Yuan et al (2020; for deep learning) and Wang et al (2022; for nighttime COT retrieval) to expand the range of cloud classification, for the first time, to nighttime retrievals. The dataset is novel, unique, and of high quality, with significant potential for use in cloud and climate studies. However, I find that the manuscript lacks some important information, particularly regarding their choice of testing data, sensitivity studies, and data screening techniques. Moreover, the data could be presented in a better way for a wider scientific community. Therefore, I recommend a major revision of the manuscript/dataset by addressing the following comments before it can be considered for potential acceptance in ESSD.

We appreciate your recognition of the novelty and quality of our dataset. Regarding to your concerns, we realize the need for further clarification and improvement. In the revised manuscript, we have provided additional details on the selection criteria for the testing data, conducted sensitivity studies to evaluate the robustness of our model, and further clarified our data processing and screening methods. Furthermore, we have improved the presentation of our dataset by providing additional descriptions and modifying its storage format, which significantly enhances its accessibility. We believe these revisions will address your concerns and improve the quality of our manuscript. Please see the response below for further details.

Comments:

1. While the TIR-CNN-based retrieval of cloud properties in Wang et al (2022) could be a better one compared to the TIR-based algorithm, it cannot replace the standard daytime retrieval algorithm in MODIS. Therefore, to justify using the TIR-CNN-based COT used in training the MCC-classification, I suggest the authors include an additional validation in this study, which includes comparing their MCC classification with that from the outputs of a CNN trained on MODIS daytime COT. Even though I agree that the choice of TIR-CNN-based COT is methodologically justified to stay consistent in their application to both daytime and nighttime retrievals, the inclusion of this additional validation/sensitivity study will strengthen their results and the MCC classification dataset.

Thank you for your valuable suggestion. We agree that an additional validation comparing the algorithms based on TIR-CNN-based COT and MODIS daytime COT would be beneficial. In response, we have performed a sensitivity experiment by training a CNN model using MODIS COT and the original three infrared channels. The results indicate that the accuracy of the CNN model based on MODIS COT is 91.3%, which is almost identical to the 91.5% accuracy achieved by the model using the COT retrieved from Wang et al. (2022). The new training results based on MODIS

COT are shown in Figure R1. This demonstrates that the TIR-CNN-based COT can effectively serve as a reliable alternative to MODIS COT for cloud classification, which further reinforce the reliability of our model and the MCC classification dataset. Therefore, we have added some supplementary statements regarding this sensitivity experiment in the '2.2 Data' Section, as follows: "*To validate the reliability of using TIR-CNN-based COT as a replacement for MODIS COT, we conducted a sensitivity experiment: comparing our classification with the outputs of a CNN trained on MODIS daytime COT. The results (Fig. S3) showed that the accuracy of both models is nearly identical, indicating that TIR-CNN-based COT is a reliable alternative to MODIS COT.*" (Lines 142-145).

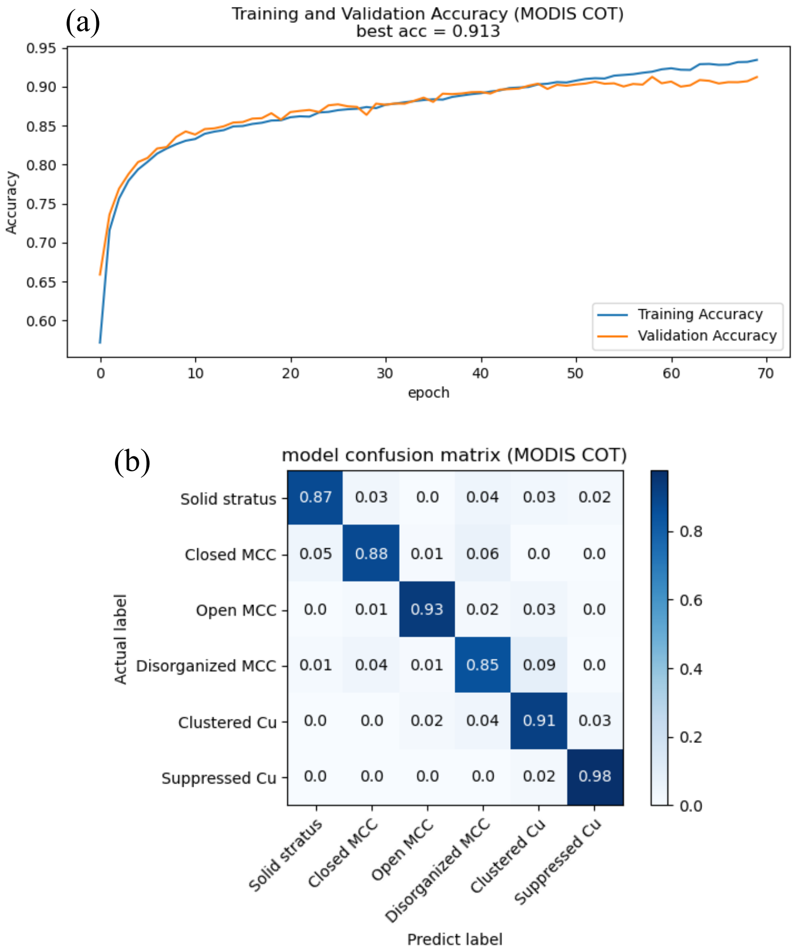


Figure R1. Model training results based on MODIS COT. (a) the model's accuracy on the training and validation datasets. (b) the confusion matrix of the model.

2. The authors use different numbers of samples in each MCC category to train the model.
For example, (0.6 times) 9,900 labeled suppressed Cu are used compared to just (0.6 times) 3,548 solid stratus samples. Shouldn't this disparity impact the performance of the classification for different MCC categories? Can the authors comment on this?

We agree that the disparity in the number of samples across different cloud types can potentially affect the model's performance. When there is an imbalance in the training data, the model might become biased towards the categories with larger sample sizes and perform less accurately on the underrepresented categories.

However, since each category in our training set contains a sufficient number of samples and the sample size ratio between categories does not exceed 4:1, the impact of the sample imbalance on our training should be relatively small.

To approve this, we conducted an experiment in which 2,000 samples were randomly selected from each category to train a new CNN model. The training result (Figure R2) shows a prediction accuracy of 88.8%, which is slightly lower than our previous result. This may be due to the relatively small sample size. Despite this, it indicates that sample imbalance has little impact on our training process. Therefore, we have added some explanations in Section '2.2 Data': *"Despite the disparity in sample sizes within our training dataset, it is capable of yielding superior model performance compared to a balanced dataset (Fig. S2)."* (Lines 132-133)

In the future, we plan to add more manually labeled samples to the category with the fewest samples to explore whether a larger balanced training dataset can lead to a better performance.

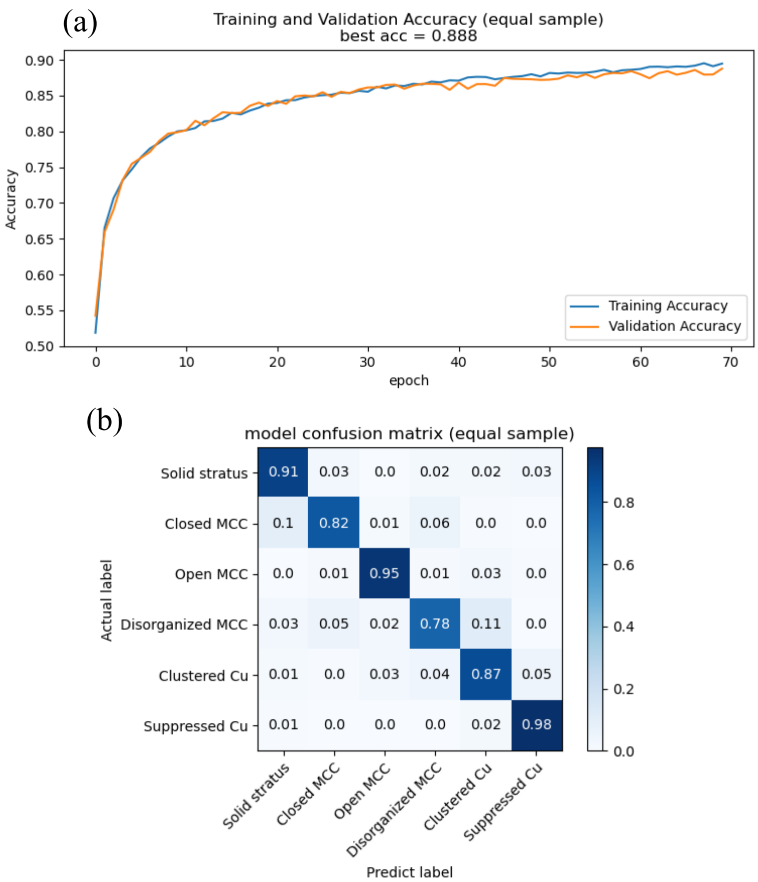


Figure R2. Model training results with 2,000 samples for each category. (a) the model's accuracy on the training and validation datasets. (b) the confusion matrix of the model.

3. Why do the authors interpolate the data within a $1^{o} \times 1^{o}$ scene to 128×128 pixels? Even if I consider the finest resolution of 1km, the number of pixels within a $1^{o} \times 1^{o}$ scene would be less than 128×128 pixels, leading to extrapolation-related truncation error. Also, did the authors perform any sensitivity test regarding the size of the scene considered in training the model except from $1^{o} \times 1^{o}$? Wood and Hartmann (2006) use native MODIS 256×256 pixels in their classification. Increasing the grid size may reduce the probability of misclassifications (Fig. 3). For instance, considering a smaller domain may result in misclassification of edges of open cells into clustered Cu. In case you achieve a better classification, the resulting dataset can be resampled to a finer grid easily for future use in conjunction with other climate and weather-related datasets.

Thank you for pointing out these valuable issues. For the first question, we did not perform extrapolation; rather, we performed refinement, where the 128×128 pixels images obtained through interpolation are included within a 1-degree grid. We have made the following revisions in the manuscript to clarify the translation: "*To align with conventional climate datasets, we developed a standard 1° gridded datasets by applying the trained model to 1°-resolution images, where the 1°×1° satellite images were interpolated and refined to 128 × 128 pixels.*" (Lines 175-176)

Although this may still lead to issues near the cloud edges, it avoids interference from other grid pixels. Particularly at higher latitudes, a 128×128 pixels scene will cover multiple grids, thus the classification of center grid may be interfered by the cloud features in neighboring grids. By directly classifying the cloud pixels within a 1° grid, this type of error can be minimized.

For the secondary question regarding to the sensitivity test of the scene size, first, the size of the scene can influence the definition of cloud types. A larger scene may necessitate a redefinition of cloud types, such as the Sugar, Gravel, Fish, and Flowers categories defined by Stevens et al. (2020) ($10^{o} \times 10^{o}$). Secondly, Yuan et al. (2020) also noted that larger scene sizes increase the probability of multiple different cloud types appearing within a single scene, while smaller scenes may lack sufficient contextual information for effective classification. Beyond the consideration of Yuan et al. (2020), we took additional factors into account and ultimately decided to adopt a standardized grid scene to better address the issue of pixel stretching at high latitudes. Moreover, standardized cloud classification datasets are more convenient for the community to use, as most current studies utilize 1-degree grid datasets for analysis.

Thirdly, we agree that using a larger scene size helps to constrain and reduce the misclassifications caused by smaller domains. Therefore, a more reasonable approach would be classifying the integration of scenes of different sizes using the automatic unsupervised learning, which is an area we plan to explore in our future work.

4. Information is missing regarding why channels 29, 31, and 32 were particularly

used in training and classification when multiple other cloud-top-related channels (33-36) are available in MODIS.

Channels 29, 31, and 32 were chosen because they effectively represent cloud properties and cloud-top temperature, which are critical for cloud classification. Specifically, channels 29 (8.7µm) is sensitive to water-vapour absorption, while Channels 31 (10.8µm) and 32 (12.0µm) provide valuable information on cloud-top temperature. In contrast, Channels 33-36 are more focused on cloud-top altitude and other related properties. We previously conducted an experiment using all six channels (29, 31, 32, 33, 34, and 35) for model training, the results were very similar to those obtained only using channels 29, 31, 32 (unfortunately, the experimental data from that trial was not properly archived). In order to reduce the amount of data, we ultimately chose the three channels 29, 31, 32 as inputs.

Therefore, we added explanations in the manuscript to justify our selection of these three channels: *"Thermal infrared (TIR) channels 29 (8.7µm), 31 (10.8µm) and 32 (12.0µm) were specifically chosen as they most effectively represent the cloud properties and cloud-top temperature."* (Lines 135-136)

5. What are the parallel yellow and red lines in the panels of Figure 4? Are these physical and being used in classification or graphics-related artefacts?

The striped patterns (yellow and red lines) visible in Figure 4 are graphical artifacts generated by satellite sensors. Although these artifacts may create some visual noise, they have minimal effect on our pattern identification process since our model relies on qualitative pattern recognition rather than quantitative analysis. Furthermore, the CNN model can filter the noise out, so we opted not to eliminate the striped noise during the training and classification processes.

Nevertheless, we also tried several methods to eliminate this noise and improve the visual quality, such as mean filtering, Fourier transform, and directional filtering. However, the stripe noise in our data is not traditionally horizontal or vertical, and there are no significant numerical characteristics, so none of these methods were effective in removing it. When the noise was removed through these methods, the image became very blurry, as shown below. Perhaps AI-based methods could help eliminate it, and we will continue to try that.

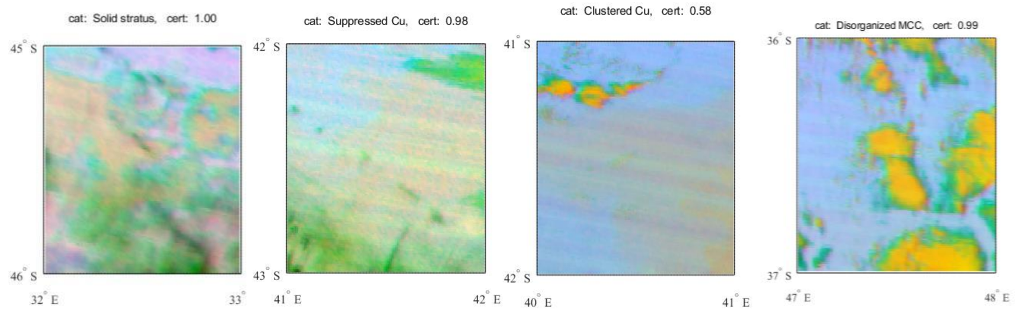


Figure R3. The image processed with directional filtering.

6. Regarding the dataset, I highly recommend using standard data formats used in atmospheric sciences like netCDF and HDF for easy cross-platform and cross-software accessibility. Not all users will be accustomed to the Python-specific NumPy format.

We fully agree that using standard data formats would enhance cross-platform and cross-software accessibility. In response, we have converted all the files previously in NumPy format to HDF files.

7. Since this is a data-descriptor paper, some important information on the contents (variables and how they are calculated) and the file nomenclature should be included in the manuscript. It may be presented as a separate sub-section within Section 2 and summarized using an additional table. This information is currently missing from the manuscript.

Indeed, our manuscript lacks a description of the dataset-related content. We have added the following sub-section in Section 2 to explain the contents of our dataset:

"2.3 Marine Low-cloud Mesoscale Morphology Dataset
Our cloud dataset provides global classifications of daytime and nighttime marine low-cloud mesoscale morphology for the years 2018-2022, with a spatial resolution of 1° × 1° and a temporal resolution of 5 minutes. The dataset is provided in two kinds of files: those prefixed with "day" store the daytime classification results for each year, while files with the prefix "night" contain the nighttime classification results for each year. Both sets of files include the same variables. Table 3 provides an overview of the variables and their associated information. The key variables in the dataset include 'date' (representing the time of the 1° × 1° scene, formatted as the MODIS granule date), 'lon' and 'lat' (indicating the central longitude and latitude), and 'cat' (assigned cloud category, the values from 0 to 5 correspond to 'Solid Stratus', 'Closed MCC', 'Open MCC', 'Disorganized MCC', 'Clustered Cu', and 'Suppressed Cu', respectively). Additionally, 'cert' represents the model certainty, quantifying the probability that the cloud morphology belongs to the assigned category. 'low_cf' denotes the low cloud fraction, and 'COT_CNN', 'CER_CNN', and 'LWP_CNN' provide the in-cloud average cloud optical thickness, effective radius, and liquid

water path respectively, as derived from the TIR-CNN model from Wang et al. (2022). The 'Sensor_zenith' variable indicates the scene average sensor zenith angle." (Lines 191-203)

Table 3 Variables of the Daytime and Nighttime Global Marine Low-cloud Mesoscale Morphology Dataset

Variable Name	*Description*	*Source*	*Spatial Resolution*	*Temporal Resolution*	*Units*
date	*Time of the 1°×1° grid point, formatted as 'YYYYDDD.HHHH'*	*MODIS MYD021*	*1°×1°*	*5 minutes*	*-*
lon	*Central longitude (-180,180)*	*MODIS MYD021*	*1°×1°*	*5 minutes*	*degree (°)*
lat	*Central latitude (-60,60)*	*MODIS MYD021*	*1°×1°*	*5 minutes*	*degree (°)*
cat	*Category of the cloud morphology: 0-Solid stratus, 1-Closed MCC, 2-Open MCC, 3-Disorganized MCC, 4-Clustered Cu, 5-Suppressed Cu*	*Cloud Classification Model*	*1°×1°*	*5 minutes*	*-*
cert	*Model certainty*	*Cloud Classification Model*	*1°×1°*	*5 minutes*	*-*
low_cf	*Cloud fraction of low clouds*	*MODIS MYD06*	*1°×1°*	*5 minutes*	*-*
COT_CNN	*In-cloud average cloud optical thickness (COT)*	*TIR-CNN model of Wang et al. (2022)*	*1°×1°*	*5 minutes*	*-*
CER_CNN	*In-cloud average cloud effective radius (CER)*	*TIR-CNN model of Wang et al. (2022)*	*1°×1°*	*5 minutes*	*µm*
LWP_CNN	*In-cloud average cloud liquid water path (LWP)*	*Calculated from COT_CNN and CER_CNN*	*1°×1°*	*5 minutes*	*g/m^2*
Sensor_ze	*Scene average*	*MODIS*	*1°×1°*	*5 minutes*	*degre*

nith	*sensor zenith angle*	*MYD021*			*e (°)*

Minor comments:

1. Line 49-51: More recently Goren et al (2019) showed a similar delay in closed-to-open transition using LES.

Yes, references have been added.

2. Line 60-61: The cloud morphology dataset by Wood and Hartmann (2006) has been expanded to more than a decade of MODIS observations, the Morphology Identification Data Aggregated over the Satellite-era (MIDAS), by McCoy et al (2023).

Thank you for reminding us! We have revised the sentence to: *"Their work was pioneering and has since been extended to more than a decade of MODIS observations by McCoy et al. (2023)"* (Line 62)

3. Line 64: Abbreviation VGG not defined!

The abbreviation *"*VGG*"* has already been defined as *"*Visual Geometry Group*"* in the manuscript.

4. Line 64: "… for daytime scenes …". All the morphology datasets discussed prior to this point correspond to daytime observations, don't they?

Yes, they did. We have removed the original sentence *"Their dataset has higher spatial resolution, at 128×128 pixel, but also only includes classifications for daytime scenes."*

5. Line 102-103: "Disorganized MCC … larger droplets and lower optical thickness." Can the authors cite studies that have demonstrated this fact?

The citations for this fact have been added, and the original sentence has been checked and revised to: *"Disorganized MCC are a mix of convective elements and extensive stratiform clouds, marked by smaller droplets and lower optical thickness (Yuan et al., 2020; Liu et al., 2024)."* (Lines 104-105)

6. Line 106: Citation missing!

References have been added.

7. Line 116: "spatial resolution of 1 km" This resolution is for nadir pixels. It changes with sensor zenith angle.

Thank you for pointing that! The original sentence has been further clarified as: *"The primary observation data utilized in this study were derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard NASA's Aqua satellite,* *including the Level-1B radiance product MYD021KM and the Level-2 cloud product MYD06 (Platnick et al., 2017), both with a spatial resolution of 1 km at nadir point."* (Lines 115-117)

8. Line 121: The authors state that they filter out scenes with more than 10% high clouds or ice clouds. How do the authors deal with ice/high cloud pixels in scenes where they are less than 10%? Are they set to missing values and not used in either training or classification steps?

Thank you for pointing out this oversight! We initially overlooked this detail and directly used the remaining ice/high cloud pixels in both training and classification steps, which may introduce noise into the model. In future model iterations, we will exclude these pixels by setting them to missing values, ensuring they do not interfere with our training and classification process.

9. Line 172: How is the reanalysis data co-located? Do you select the nearest timestamp or interpolated the data to MODIS observations?

We interpolated the ERA5 data temporally to match the MODIS observation time. We added a detailed explanation in the manuscript to clarify our co-location method: "*For the purpose of investigating the influence of meteorological conditions on low-cloud morphologies, we conducted some statistical analyses utilizing the co-located ERA5 reanalysis data (1° × 1°, 1-hourly) from European Centre for Medium-Range Weather Forecasts (ECMWF). The co-location is achieved by spatially selecting the nearest ERA5 grid point to each MODIS observation and temporally interpolating the ERA5 data to match the exact time of the MODIS observations. This ensures accurate alignment between the two datasets in both space and time."* (Lines 177-181)

10. A link to the classification dataset is missing in the "Data Availability" section.

The classification dataset (training, validation, and test sets) has been added to the same link as our cloud classification product dataset, along with descriptions of the variables included. We informed in the manuscript that: *"Daytime and nighttime cloud classification datasets as well as our training dataset are accessible on the https://doi.org/10.5281/zenodo.13801408 (Wu et al., 2024)."* (Lines 426-427)

11. No information on the file "example.xlsx" in the data repository.

Thank you for noting that. The description for the file "example.xlsx" has been added into the data repository, which is: "*File 'example.xlsx': A sample of the variable data from our cloud classification dataset, showcasing the classification results of a MODIS granule captured on January 1, 2018, at 00:25 UTC. This sample is provided to help users better understand the content of our dataset."*

Language-related suggestions:

Line 21: Abbreviation RFO defined in abstract is not used.

Fixed.

Line 84: dependent?

Yes, fixed.

Line 91: Prior to "Section 2 intro…", perhaps insert an introductory sentence like "The manuscript is organized as follows."

Done.

Line 184: Abbreviation ML is not defined

Fixed.

Line 210: Remove underscore after Yuan et al (2020)

Done.

Line 383: Consider changing the word "worse"

The sentence has been changed to: *"which largely explains the performance gap between our nighttime model and the daytime model proposed by Yuan et al. (2020)."* (Lines 409-410)

Line 409: "… nightly …"Do you mean nighttime?

Yes, fixed.

Reviewer #2

Marine low-clouds cover the majority of the ocean, and play an important role on the Earth's radiation budget. Due to a lack of local or ground-based observations, satellites become powerful tools for MLC measurement, while satellite observations over nighttime are still relatively limited. Thus, this study by Wu et al. introduced a deep-learning based method for the classification of MLC and their mesoscale morphology using MODIS observations, and a global dataset is developed as well. Both all-day model and day-time model were developed and evaluated. It is interesting to find some differences on the daytime and nighttime MLC, and distinct seasonal variations were also noticed for different MLCs. The new method as well as the resulting dataset is an important addition for the community, and the paper is well organized and presented. The paper could be considered for publication after considering following suggestions.

Thank you very much for acknowledging our work and for your valuable comments and suggestions. We fully agree your concern on the construction of our dataset, the representativeness of training data, as well as the model training and validation process. In the revised version, we have further clarified the construction of our training, validation, and testing datasets, evaluated the data representativeness, and explained the reliability of our nighttime results. Please refer to the response below for further details.

1. The quality of the training and testing dataset has been essential for DL-based models, so the datasets for the training should be carefully constructed. The 2.2 Data session gave some information on the dataset, while missed some as well. For example, Figure 1 gave some examples of MLCs of different kinds, and how was the original training dataset classified? The independency of training and testing dataset is also important, so I would suggest to introduce the testing and evaluation dataset at the Data session as well.

We apologize for any confusion caused by the lack of detailed information. Our training, validation, and test datasets are all sourced from the same manually annotated dataset. While they originate from the same data pool, they have been randomly partitioned into mutually independent subsets to ensure robust evaluation and model generalization. To clarify further, we have made some modifications to the original text, and it currently reads like: *"A total of 38,756 labeled daytime scenes were obtained, including 3,548 scenes of solid stratus, 6,277 of closed MCC, 3,345 of open MCC, 6,739 of disorganized MCC, 8,947 of clustered Cu and 9,900 of suppressed Cu. These scenes were then randomly partitioned into three mutually independent datasets for training, validation, and testing, with a distribution ratio of 3:1:1 respectively."* (Lines 129-132)

2. Cloudy and atmospheric properties show clear seasonal variations. For example, surface and atmospheric temperatures may significantly different from season to season, and this is also true for clouds. It is mentioned that only the results over the first half of 2014 were used for data training. Would such choice of results from half a year influence the DL performance?

Your suggestions are highly valuable. Indeed, the formation and development of clouds are significantly influenced by meteorological conditions, such as sea surface temperature and lower tropospheric stability (LTS), which differ from season to season. Nevertheless, as the ultimate manifestation of meteorological conditions, cloud patterns exhibit certain similarities across different regions and seasons. That is, the cloud patterns in a specific region resemble those on a global scale, and the cloud patterns in the first half of the year are similar to those throughout the entire year. Therefore, the cloud patterns contained in our dataset can largely represent the all-year and global clouds.
To validate our hypothesis, we examined the differences in the probability density functions (PDFs) of thermal radiance, cloud optical thickness (COT), and cloud morphology between our training dataset and the global full-year dataset. As shown in the Figure R4. The results revealed a substantial overlap between the two PDFs, suggesting that the training data we selected is relatively representative and can be used to substitute the global full-year dataset. Therefore, we have added a statement to the article: "*The representativeness of this dataset was validated as the probability density functions (PDFs) of thermal radiance data and cloud optical thickness show large overlap with those of the global and full-year dataset (Fig. S1).*" (Lines 120-122)

Despite this, in the future, we will also re-label the global dataset for all seasons, both day and night, and update our model and products in subsequent iterations.

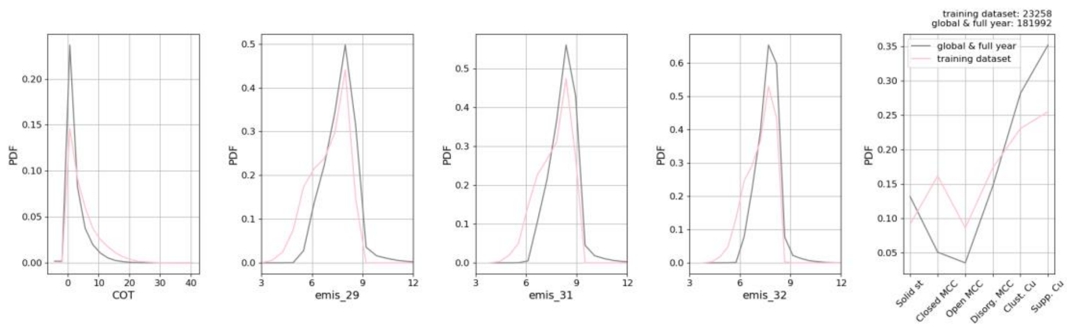


Figure R4. The comparison of probability density functions (PDFs) between our training dataset and a global full-year dataset. (a) PDFs of cloud optical thickness (COT); PDFs of radiance data from infrared channels: (b) 29, (c) 31, (d) 32; (e) PDFs of cloud morphology.

3. Would it be possible to include the exact variables of input for different models in the flowchart of figure 2? This would be very helpful to better understood the details of the model efficiently.

Thank you for your valuable suggestion. We have included the exact input variables in Figure 2a, as also shown in the following Figure R5, to enhance the reader's understanding of the model details.

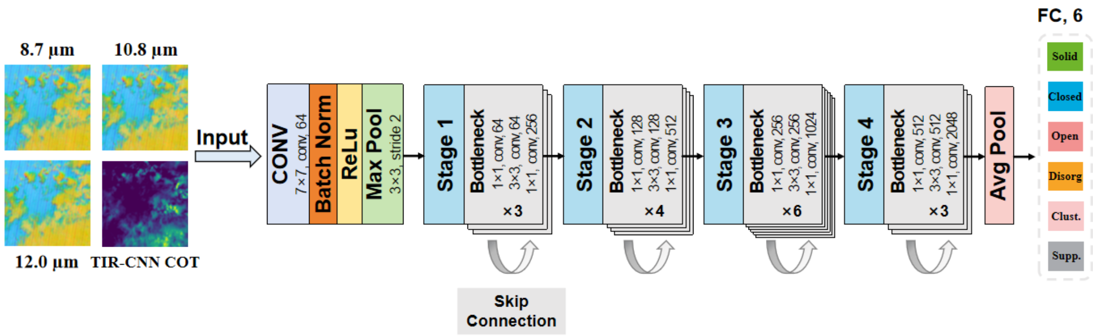


Figure R5. The revised ResNet-50 model structure.

4. The example of solid stratus show relatively regular linear structure, and are such structures natural? Please double check.

As mentioned in the responds to Comments #5 from Reviewer #1, these linear structures are strip noise caused by the components of satellite sensor. This type of noise is present in the classification, but it can be identified and filtered out by the model, so it does not have a significant impact on the classification process. We have attempted conventional elimination methods, but these approaches have proven unsuccessful so far. AI-based removal techniques appear promising, but additional time is required to fully master it.

5. The training model based on daytime results is extend to nighttime observations. This is essential for the work, and could be tricky. The validation of the model for nighttime observations is very important, while only some examples were shown in Figure 4. Would it be possible to improve the validation to ensure the reliability of the results for nighttime?

As we responded to the Comments #1 from Reviewer #1, the difficulty in nighttime predictions arises from the lack of cloud thickness information. However, the COT retrieval method by Wang et al. (2022) has been proven reliable and can effectively replace MODIS COT for accurate nighttime classification. Moreover, since cloud classification requires pattern recognition rather than quantitative values, the differences in infrared radiance and COT between day and night do not significantly affect the classification.

In addition, in our response to the Major comments #1 from Reviewer #3, we examined the probability density functions (PDFs) of COT and thermal radiance between our training dataset (daytime) and nighttime dataset (Figure R7), which present large overlap. The significant similarity between daytime and nighttime input data indicates less extrapolation by the model and ensures the reliability of our nighttime results.

6. Figures 7 indicates clear day and time differences between RFO of different MLCs. Could the authors give some discussions on the reasons for the differences?

Thank you for pointing us in this direction. Your suggestion actually aligns with our own thoughts, and it is part of our next steps.

In this study, we have conducted a statistical analysis of six meteorological conditions in the article but found that the variations in these meteorological factors between day and night were not significant. Therefore, other factors, such as cloud-top radiative cooling, might be responsible for the observed differences. We plan to further investigate the controlling factors behind the day-night changes of cloud morphology in future work.

Furthermore, given that the primary focus of this article is to introduce the cloud dataset and the machine learning method, we feel that including an analysis of cloud-controlling factors might shift the focus away from the main theme.

Reviewer #3

Review of "A Global Classification Dataset of Daytime and Nighttime Marine Low-cloud Mesoscale Morphology Based on Deep Learning Methods" by Wu et al [MS number: essd-2024-536]

This study produces a global dataset of daytime and nighttime low-cloud mesoscale morphologies (categorized into six types) using a convolutional neural network through a combination of MODIS infrared radiance data and machine-learning-retrieved cloud optical thickness. Leveraging this novel dataset, the authors analyzed the day-night contrast in climatology, seasonal cycles, and cloud properties of cloud morphologies. One of notable findings is the significant diurnal variation in the occurrence frequency of closed MCC and suppressed Cu. The primary contribution of this work lies in the generation of nighttime low-cloud morphology data, which complements the well-established daytime morphology datasets from prior studies. This advancement would inspire and enable more downstream research like understanding the diurnal cycle of cloud morphology and cloud-longwave-radiation-climate feedback. The manuscript is overall well-written and well-organized, with nice presentation of figures. However, my major concern pertains to the limitations in the model's training and validation processes, which could impact the dataset's reliability. Addressing these issues would significantly strengthen the study's contribution to the marine low cloud research community. I'd like to recommend a major revision before this manuscript is considered for publication in ESSD.

Thank you for your valuable comments and suggestions. We are pleased that the value of our nighttime dataset has been recognized. As you mentioned, this dataset can be used for various follow-up studies, such as cloud morphology diurnal cycles, the controlling factors of cloud type transitions, and cloud-longwave-radiation-climate feedbacks. Beyond these, we can also investigate the impact of climate change and anthropogenic emissions reductions on the long-term trends of cloud morphology, providing a more comprehensive evaluation of cloud radiative feedbacks associated with type transitions. We hope this research will encourage greater attention to the indirect climate effects of marine low-cloud morphology, which is also the primary purpose of our dataset: to be widely used by the community and to offer valuable insights.

Regarding the limitations you pointed in the model's training and validation process, we fully agree with your concerns. In the revised manuscript, we have further clarified the model's training and validation processes, validated the global applicability of the regionally trained model, and fixed some minor issues. Please see the response below for further details.

At last, your thoughts actually align closely with ours. Many of the suggestions you mentioned, such as investigating the physical reasons behind the diurnal variation of cloud morphology and the impact of morphology transitions on shortwave and longwave radiation at the TOA, are already ongoing within our group. However, since this article is more foundational, many aspects were not fully presented. We look forward to sharing new results with you soon.

Major comments:

1. One of my primary concerns is the validity of applying a regionally trained deep learning (DL) model to global predictions. In this study, the authors developed their model using data from the SEP region only and then applied it to generate a global dataset. While the model demonstrates relatively high prediction accuracy over SEP (Figure 3), it is unclear whether this performance extends to global applications. Regarding this issue, the authors should first clarify the rationale for selecting SEP as the training region rather than using a global or other regional dataset. Was this choice subject to the limited availability of the data, or is there a similarity in morphology climatology between SEP and the global scale? If SEP is your best choice at the moment, it would be essential to evaluate whether using a regionally trained model for global predictions is reasonable. One approach to examine this would be to generate a global map of prediction accuracy for each cloud morphology type to check the model's global performance. Additionally, the authors could examine the differences in the PDFs of thermal radiance, COT, and cloud morphology between SEP and the global dataset. A smaller difference or larger overlaps would indicate less extrapolation by the model, enhancing the credibility of the global dataset.

Similarly, the authors would have to be careful when extending the daytime-trained model to nighttime predictions, as this may also introduce potential extrapolation issues. The authors provided only a single example to illustrate the model's success at nighttime, which is insufficient to establish its statistical reliability. To address this concern, additional cases should be analyzed to validate the model's nighttime performance. Alternatively, examining the differences in the PDFs of thermal radiance, COT, and cloud morphology between daytime and nighttime could help assess the extent of extrapolation and ensure the robustness of the predictions.

We totally agree with your ideas. Applying a model trained on regional datasets to global dataset need careful validation.

First, we give the rationale for selecting the SEP region as the training dataset:

1) We believe that SEP region encompass all the cloud types around the world, and can provide sufficient samples for each type.

2) Meanwhile, we are limited by the available training data. We only labeled this portion of the data initially, and re-labeling the global dataset and regenerating the cloud dataset would take a considerable amount of time.

However, your suggestions provided effective methods for our model validation. Following your suggestion, we generated the plots of probability density functions (PDFs) of COT, thermal radiance, and cloud morphology. One plot compares our training dataset with the global full-year dataset (Figure R6), while the other compares the nighttime dataset (Figure R7).

Both figures show large overlap in the PDFs of COT and thermal radiance. So, we have added some statements in the manuscript to validate the reliability of our model and the nighttime results: "*The representativeness of this dataset was validated as the probability density functions (PDFs) of thermal radiance data and cloud optical thickness show large overlap with those of the global and full-year dataset (Fig. S1)*" (Lines 120-122)

"In addition, we further examined the differences in the PDFs of the thermal radiance data and the TIR-CNN-based COT between our training dataset (daytime) and nighttime dataset. As depicted in Fig. S4, these PDFs nearly overlapped, which means less extrapolation will be introduced when the model is generalized to nighttime data. And it also illustrates the credibility of our nighttime classification results." (Lines 145-148)

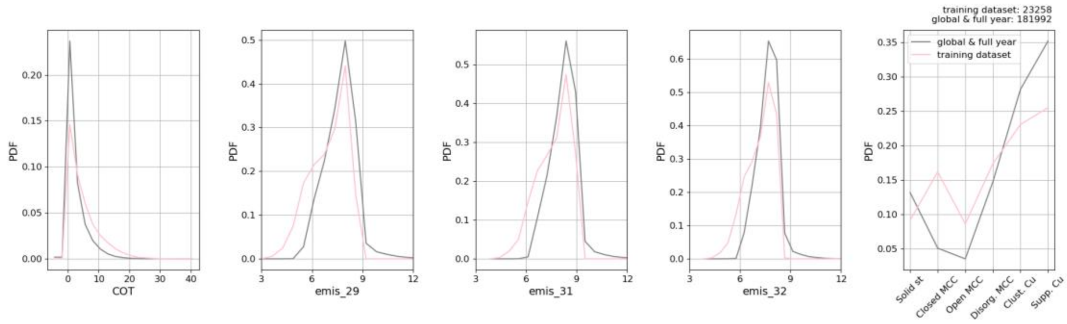


Figure R6. The comparison of probability density functions (PDFs) between our training dataset and a global full-year dataset. (a) PDFs of cloud optical thickness (COT); PDFs of radiance data from infrared channels: (b) 29, (c) 31, (d) 32; (e) PDFs of cloud morphology.

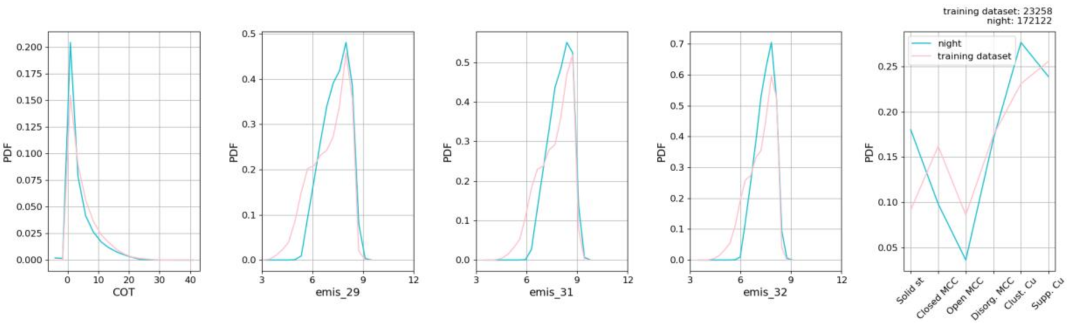


Figure R7. The comparison of probability density functions (PDFs) between our training dataset and a nighttime dataset. (a) PDFs of cloud optical thickness (COT); PDFs of radiance data from infrared channels: (b) 29, (c) 31, (d) 32; (e) PDFs of cloud morphology.

2. Regarding the model training, validation, and testing, the data-splitting strategy is unclear. For instance, was the dataset split randomly or manually into the 6:2:2 ratio? Furthermore, the validation method used to assess the model's predictions has not been described. The authors should clarify these aspects to improve the robustness of their results.

We fully agree that clarifying our dataset splitting method and model validation method is necessary. Indeed, the original annotated dataset was randomly split according to a 3:1:1 ratio, so we clarified it in the manuscript: *"These scenes were then randomly partitioned into three mutually independent datasets for training, validation, and testing, with a distribution ratio of 3:1:1 respectively."* (Lines 131-132)

Based on your suggestions, we have added the following content regarding to the validation method in the "2.4 Method" Section: *"After each training epoch, the validation dataset was used to evaluate the trained model's performance, which allows us to monitor the model's success. When the training process was completed, the test dataset was used for the final evaluation of the model's performance. We used the optimal model to predict each sample in the test dataset, compared the model's predictions with the true labels, and assessed the accuracy using metrics such as accuracy, F1-score, and recall."* (Lines 229-232)

And we have also added the content regarding model test results in Section '3.1 Model Performance': *"Our model achieves an average precision of approximately 91%, an F1-score of 90.6%, and a recall of 90.8%, demonstrating its strong generalization capability and robustness."* (Lines 241-243)

3. Given the critical role of cloud morphologies in Earth's radiation budget, the authors could consider including a climatological analysis of shortwave and longwave radiation at the TOA for the six cloud morphology types. Adding such an analysis would significantly enhance the insights and scientific value of this study.

Thank you for your insightful and constructive suggestions! Your idea aligns perfectly with our thoughts. We are currently conducting a long-term trend study on cloud morphology using this cloud dataset, as well as exploring the impact of aerosols and climate change on these long-term trends. We aim to conduct a systematic study of cloud morphology and cloud type transitions. Based on this, our next step will be to investigate radiative effects. However, some issues are existing in the current radiative datasets: there is a lack of instantaneous clear sky albedo. We need to address some fundamental issues to make the radiative data more solid, which will take time. Moreover, since our cloud scenes are instantaneous rather than monthly, matching the five years of radiative data will be slow, and this could eventually become a separate paper.

In earlier research, Mohrmann et al. (2021) have assessed the radiative properties of the six cloud types, using data from the Clouds and the Earth's Radiant Energy

System (CERES), specifically SYN 1-degree hourly data (daytime). They analyzed the net cloud radiative effect (CRE) for each cloud type at different spatial scales and found that Solid MCC and Disorganized MCC exhibit the strongest climatological average cloud radiative effects (Figure R8c, square symbol, about -48 Wm^{-2}). However, they did not account for global radiative impacts associated with long-term changes in cloud morphology, which is the question we aim to address in future work.

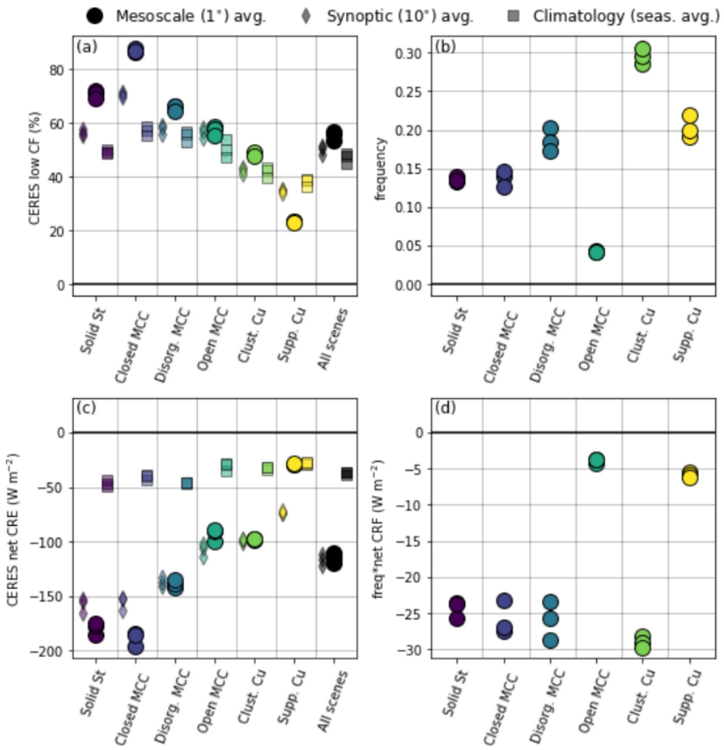


Figure R8. Cloud radiative properties by cloud type from Mohrmann et al. (2021): (a) CERES cloud fraction, (b) cloud frequency of occurrence, (c) average CERES net CRE per cloud type, (d) frequency-weighted net CRE. Each set of three symbols is for the 3 years (2014-2016) used. For panels (a) and (c), the mesoscale, synoptic, climatological averages are shown using circular, diamond, and square symbols, respectively.

Minor comments:

L30: longwave warming effects are more significant for high clouds, which might not be so for low clouds.

The original sentence was changed to: *"They exert a strong radiative cooling on the planet as the residual of a larger cooling effect and a positive warming effect (Klein and Hartmann, 1993; Eytan et al., 2020)."* (Lines 30-32)

L67: What is the major difference between the six-type classification of this study and the four type one here?

The major differences are: (1) Scale: The four types here require a larger scale to be fully observed, at least a 10° by 10° field of view. (2) Classification Basis: Our six-type classification emphasizes the underlying physical processes, while their four-type classification focuses primarily on the external appearance and morphological features.

We have incorporated some key information into the original sentence, and it now reads as follows: *"Moreover, Schulz et al. (2021) developed an object detection model to classify four larger scale ($10^o \times 10^o$) cloud morphologies in trade wind regions of North Atlantic. These morphologies were vividly named as "sugar," "gravel," "flowers," and "fish" mainly based on their visual appearances." (Lines 66-69)*

L81: Do you mean the decline in the *long-term* trend?

Yes, fixed.

L83: "how much they contribute to … remain unclear" to "how nighttime cloud cover varies under different cloud morphology types remain unclear."

Done.

L90: Please clarify the temporal and spatial resolution.

Done.

L97: "created" to "driven"

Done.

L119: Please clarify the temporal resolution of the training dataset.

Done.

L121-122: Have you excluded middle clouds (i.e., those situated between 3 and 6 km)? These clouds are prevalent over midlatitude oceans, and they also contaminate low cloud observations.

We are sorry for not eliminating the influence of middle clouds in our training process. In future work, we plan to refine our methodology by re-screening middle clouds to improve the accuracy of our model.

L174: Please clarify the level of the divergence used.

Done.

L199: I'd suggest labeling the input variables (three channels and COT) and the output variables (six cloud morphology types) in Figure 2a to improve its clarity and readability.

Done.

L210: It looks like the improvement is limited. Have you examined the COT retrieval uncertainty? If it is greater than the improved accuracy, it would be unnecessary to include the COT into the predictors.

Thank you for your suggestion. As mentioned in Comments #1 from Reviewer #1, the model using retrieved COT demonstrates comparable predictive accuracy to the model based on MODIS COT, confirming its reliability.

L210: Typo: "Yuan et al (2020)_due to" to "Yuan et al (2020) due to"

Fixed.

L212: Which is it relative to?

We have completed the sentence as follows: *" Although it is a bit less accurate compared to the visible light model from Yuan et al. (2020), it is undeniable that this model has achieved a relatively high accuracy level when compared to other TIR* *model (Lang et al., 2022), and can effectively accomplish the classification tasks we proposed."* (Lines 237-240)

L219: "clustered Cu" to "clustered Cu or closed MCC"

Done.

L250: "n denotes" to "with n denoting"

Done.

L305: "its seasonal variation" to "the peak in summer"

Done.

L332: do you mean "decrease by 2 microns on average"?

Yes, fixed.

L333: Please clarify whether the LWP mentioned here represents the in-cloud value or the grid box mean value.

Sorry for the lack of information provided. We have added the following clarification in Lines 356-357: *"All of the cloud microphysical properties represent the in-cloud mean value within a 1° grid."*

L349: Why is there a westward shift at night? Also, for stratocumulus clouds, LTS is usually higher at night. Why does it decline for closed MCC at night?

Sorry for the confusion. We originally hypothesize that the observed decline in LTS at night may be related to the movement of clouds. Since our statistical analysis covers a relatively large spatial scale, the meteorological conditions associated with specific cloud types could change as the clouds move, potentially leading to the observed decrease in LTS. However, we apologize for the previous implication that this movement is necessarily westward. So, we removed the mention of "westward" in the revised text. We believe that further investigation is required to fully understand the reasons behind the LTS decline in closed MCCs at night.

L351: It would be more interesting to discuss their physical reason.

Thank you for pointing the direction for us, it is the next part of our work. It is difficult to determine the key cloud-controlling factors for each cloud type based solely on the statistical analysis in this study, and many other environmental conditions have not been included. Therefore, we plan to conduct a more detailed and comprehensive investigation into their physical causes in our next work. Furthermore, since the main focus of this article is to introduce a cloud dataset, we feel that including an analysis of cloud-controlling factors here would dilute the main theme.

L367: Why are the results shown here only for SEP, while Figure 10 presents global results?

Cloud morphology is controlled by multiple meteorological factors (Liu et al., 2024). When we study the influence of one controlling factor, cloud morphologies can be affected by the variability in other factors if the study is conducted in a global scale. For instance, while exist within the same LTS environment, clouds in mid-to-high latitude regions and those near the equator have distinctly different sea surface temperatures. Therefore, by restricting the region, we can facilitate the day-night comparison of the primary controlling factors while excluding the interference from other variables. Clouds properties are the final result of all meteorological conditions, and their properties show little differences across different regions.

References

Eytan, E., Koren, I., Altaratz, O., Kostinski, A. B., and Ronen, A.: Longwave radiative effect of the cloud twilight zone, Nature Geoscience, 13, 669-673, https://doi.org/10.1038/s41561-020-0636-8, 2020.

Klein, S. A. and Hartmann, D. L.: The Seasonal Cycle of Low Stratiform Clouds, Journal of Climate, 6, 1587-1606, https://doi.org/10.1175/1520-0442(1993)006<1587:TSCOLS>2.0.CO;2, 1993.

Lang, F., Ackermann, L., Huang, Y., Truong, S. C. H., Siems, S. T., and Manton, M. J.: A climatology of open and closed mesoscale cellular convection over the Southern Ocean derived from Himawari-8 observations, Atmos. Chem. Phys., 22, 2135-2152, https://doi.org/10.5194/acp-22-2135-2022, 2022.

Liu, J., Zhu, Y., Wang, M., and Rosenfeld, D.: Cloud Susceptibility to Aerosols: Comparing Cloud-Appearance vs. Cloud-Controlling Factors Regimes, EGU General Assembly 2024, Vienna, Austria, 14–19 Apr 2024, EGU24-4059, https://doi.org/10.5194/egusphere-egu24-4059, 2024.

McCoy, I. L., McCoy, D. T., Wood, R., Zuidema, P., and Bender, F. A.-M.: The Role of Mesoscale Cloud Morphology in the Shortwave Cloud Feedback, Geophysical Research Letters, 50, e2022GL101042, https://doi.org/10.1029/2022GL101042, 2023.

Mohrmann, J., Wood, R., Yuan, T., Song, H., Eastman, R., and Oreopoulos, L.: Identifying meteorological influences on marine low-cloud mesoscale morphology using satellite classifications, Atmos. Chem. Phys., 21, 9629-9642, https://doi.org/10.5194/acp-21-9629-2021, 2021.

Platnick, S., Meyer, K. G., King, M. D., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Holz, R. E., Yang, P., Ridgway, W. L., and Riedi, J.: The MODIS Cloud Optical and Microphysical Products: Collection 6 Updates and Examples From Terra and Aqua, IEEE Transactions on Geoscience and Remote Sensing, 55, 502-525, https://doi.org/10.1109/TGRS.2016.2610522, 2017.

Schulz, H., Eastman, R., and Stevens, B.: Characterization and Evolution of Organized Shallow Convection in the Downstream North Atlantic Trades, Journal of Geophysical Research: Atmospheres, 126, e2021JD034575, https://doi.org/10.1029/2021JD034575, 2021.

Stevens, B., Bony, S., Brogniez, H., Hentgen, L., Hohenegger, C., Kiemle, C., L'Ecuyer, T. S., Naumann, A. K., Schulz, H., Siebesma, P. A., Vial, J., Winker, D. M., and Zuidema, P.: Sugar, gravel, fish and flowers: Mesoscale cloud patterns in the trade winds, Quarterly Journal of the Royal Meteorological Society, 146, 141-152, https://doi.org/10.1002/qj.3662, 2020.

Wang, Q., Zhou, C., Zhuge, X., Liu, C., Weng, F., and Wang, M.: Retrieval of cloud properties from thermal infrared radiometry using convolutional neural network, Remote Sensing of Environment, 278, 113079, https://doi.org/10.1016/j.rse.2022.113079, 2022.

Wu, Y., Liu, J., Zhu, Y., Zhang, Y., Cao, Y., Huang, K.-E., Zheng, B., Wang, Y., Wang, Q., Zhou, C., Liang, Y., Wang, M., and Rosenfeld, D.: Global Classification Dataset of Daytime and Nighttime Marine Low-cloud Mesoscale Morphology [dataset], https://doi.org/10.5281/zenodo.13801408, 2024.

Yuan, T., Song, H., Wood, R., Mohrmann, J., Meyer, K., Oreopoulos, L., and Platnick, S.: Applying deep learning to NASA MODIS data to create a community record of marine low-cloud mesoscale morphology, Atmos. Meas. Tech., 13, 6989-6997, https://doi.org/10.5194/amt-13-6989-2020, 2020.

A Global Classification Dataset of Daytime and Nighttime Marine Low-cloud Mesoscale Morphology Based on Deep Learning Methods

Yuanyuan Wu[1], Jihu Liu[2,3], Yannian Zhu[1,2,3,*], Yu Zhang[2], Yang Cao[2,3], Kang-En Huang[2], Boyang Zheng[2], Yichuan Wang[2], Yanyun Li[2], Quan Wang[2], Chen Zhou[2], Yuan Liang[4], Jianning Sun[2], Minghuai Wang[2,3], Daniel Rosenfeld[2,5]

[1]Nanjing-Helsinki Institute in Atmospheric and Earth System Sciences, Nanjing University, Nanjing, 210023, China
[2]School of Atmospheric Sciences, Nanjing University, Nanjing, 210023, China
[3]Joint International Research Laboratory of Atmospheric and Earth System Sciences & Institute for Climate and Global Change Research, Nanjing, 210023, China
[4]TianJi Weather Science and Technology Company, Beijing, 100000, China
[5]Institute of Earth Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel
Correspondence to: Yannian Zhu (yannian.zhu@nju.edu.cn)

Abstract. Marine low clouds tend to organize into larger mesoscale patterns with distinct morphological appearances over the ocean, referred to as mesoscale morphology. While prior studies have mainly examined the fundamental characteristics and shortwave radiative effects of these mesoscale morphologies, their behaviour in the nighttime marine boundary layer (MBL) remains underexplored due to limited observations. To address this, we ~~created~~ established a global classification dataset of daytime and nighttime marine ~~mesoscale morphologies of marine~~ low~~-clouds~~ cloud morphology using a deep residual network model and infrared radiance data of 1°×1° resolution from the Moderate Resolution Imaging Spectroradiometer (MODIS) ~~infrared radiance data~~, with machine-learning-retrieved all-day cloud optical thickness aiding in model training. We analysed day-night contrasts in climatology, seasonal cycles, and cloud properties of different cloud morphology types in this study. Results show that relative frequency of occurrence ~~(RFO)~~ of closed mesoscale cellular convection (MCC) significantly increase at night, while that of suppressed cumulus (Cu) shows a remarkable decrease. Disorganized MCC and clustered Cu display a slight frequency increase during night. In addition, solid stratus and three MCC types exhibit distinct seasonal variations, whereas two cumuliform types show no clear seasonal cycle. Our dataset extends the study of mesoscale cloud morphologies from daytime to nighttime and 1° × 1° resolution makes it better match with other climate datasets. It will provide an ~~vital~~ important foundation for further research on the interactions between cloud morphology and climate processes. Our dataset is open-access and available at https://doi.org/10.5281/zenodo.13801408~~https://doi.org/10.5281/zenodo.13990646~~ (Wu et al., 2024).

1 Introduction

Marine low clouds cover the vast area of oceans~~cover the vast majority of the ocean~~ and have a pronounced impact on the Earth's radiation budget. They exert a strong radiative cooling on the planet as the residual of a larger cooling effect and a positive warming effect ~~Their daytime shortwave cooling effect and night-time longwave warming effect are essential in modulating the climate variability~~ (Klein and Hartmann, 1993; Eytan et al., 2020). These radiative effects are known to be sensitive to ~~the~~ cloud types due to their different cloud properties, such as cloud fraction and albedo. Traditional ground-based observations have ~~historically~~ classified individual marine low clouds using the cloud types defined by the World Meteorological Organization (WMO~~)~~), ~~cloud types~~ such as cumulus (Cu), stratocumulus (Sc) and stratus (St) (Zhang et al., 2018; Li et al., 2022; Guzel et al., 2024). However, satellite imagery shows that these individual clouds tend to organize into larger mesoscale patterns with distinct morphological features that are not easily discernible from the limited perspective of ground observation instruments. These mesoscale cloud patterns, referred to as cloud mesoscale morphologies, have been shown to exert different radiative effects on climate (McCoy et al., 2017; McCoy et al., 2023; Mohrmann et al., 2021), and also reflect the intricate physical processes of the underlying marine boundary layer (MBL) (Wood, 2012; Bony et al., 2020; Eastman et al., 2022; Liu et al., 2024; Mohrmann et al., 2021).

Previous studies have identified several critical environmental factors that influence the evolution of marine low-cloud morphologies. ~~In the mid-latitudes, o~~Open and closed mesoscale cellular convection (MCC) clouds are both affected by cloud-top longwave radiation cooling (Wood, 2012), but the surface fluxes dominate the open MCC when there is a strong cold advection such as polar outbreak. As a result, the passage of mid-latitude cold air outbreaks serve as key triggers for the transition from closed to open MCC (McCoy et al., 2017; Tornow et al., 2021). In the subtropics, precipitation promotes the organization and sustainment of open cell structure and dominates the transformation of closed to open MCC clouds (Savic-Jovcic and Stevens, 2008; Feingold et al., 2010; Yamaguchi and Feingold, 2015; Eastman et al., 2022). In contrast, closed MCC tend to evolve into more disorganized cumulus under conditions of warmer sea surface temperature and increased entrainment of dry air at the cloud top (Eastman et al., 2022; McCoy et al., 2023). Apart from meteorological influences, aerosols are another key cloud-controlling factor (Cao et al., 2024) and can ~~also~~ initiate these transitions by modulating precipitation~~(Cao et al., 2024)~~. High aerosol concentration suppresses the precipitation and favors the maintenance of closed MCCs, while the scarcity of aerosols promotes the generation of widespread precipitation, leading to the conversion toward open MCCs (Stevens et al., 2005; Rosenfeld et al., 2006; Petters et al., 2006; Xue et al., 2008; Goren et al., 2019). With global warming and emission reductions, there is a high likelihood that meteorological factors and aerosols will change accordingly. This raises several important questions regarding low-cloud feedback, such as whether the mesoscale morphology of low clouds will change as the climate warms and how these changes will affect radiation.

An objective classification of mesoscale morphology from satellite observations is essential for facilitating a more systematic investigation of these questions. In recent years, deep learning methods, especially those based on convolutional neural networks (CNNs), have proven particularly effective in the objective classification of mesoscale cloud morphology in satellite images~~, enabling the subsequent generation of potentially informative marine low cloud datasets for further study~~. By using ~~Wood and Hartmann (2006) trained~~ a three-layer neural network, ~~to~~ Wood and Hartmann (2006) ~~classify~~ classified daytime cloud morphology ~~in~~ at a resolution of 256 × 256 pixels ~~scenes~~ into four categories: no MCC, closed MCC, open MCC, and cellular but disorganized MCC. Their work was pioneering~~,~~ and has since been extended to more than a decade of MODIS observations by McCoy et al. (2023). ~~but limited to only subtropical regions for 2 months.~~ ~~Subsequently,~~ Yuan et al. (2020) then subdivided the cellular but disorganized category into disorganized MCC, clustered Cu and suppressed Cu for 128 × 128 scenes, and developed a global dataset of these six cloud types ~~-~~using a fine-tuned Visual Geometry Group 16-layer (VGG-16) network~~VGG-16 model~~. Subsequently, ~~Their dataset has higher spatial resolution, at 128 × 128 pixel, but also only includes classifications for daytime scenes.~~ Watson-Parris et al. (2021) employed a pre-trained CNN model to detect pockets of open cells (POCs) (224 × 224 pixel) in three main marine stratocumulus regions during daytime. Moreover, Schulz et al. (2021) developed~~trained~~ an object detection model to ~~identify~~ classify four larger scale (10°×10°) cloud morphologies in ~~the trades~~ trade wind regions of North Atlantic~~,~~. These morphologies were vividly named as "sugar," "gravel," "flowers," and "fish" mainly based on their visual appearances.~~categorizing them into types including "sugar", "gravel", "flowers", and "fish".~~

The datasets mentioned above have been utilized for various downstream tasks, such as quantifying shortwave cloud radiative effects and identifying key controlling factors of different cloud morphologies (Bony et al., 2020; Mohrmann et al., 2021; Watson-Parris et al., 2021), quantifying shortwave cloud feedbacks resulting from changes in morphology (McCoy et al., 2023), and investigating aerosol-cloud interactions across different morphologies (Liu et al., 2024). However, most current studies focus on the role of morphology in daytime shortwave radiation, with a notable lack of understanding regarding longwave radiation, particularly nighttime longwave radiation, primarily due to the scarcity of nighttime observations of cloud morphologies. Although ~~there are~~ a few geostationary satellite-based studies ~~that~~ give a nighttime morphological classification ~~during night~~, they are also limited to regional scales and lack a global-scale classification dataset (Lang et al., 2022; Segal Rozenhaimer et al., 2023).

~~Datasets~~ The studies on nighttime cloud morphology are ~~scarce~~limited~~, yet~~ but ~~it is~~ essential for investigating cloud-climate feedback. Closed MCC clouds have been shown to peak at night (Lang et al., 2022) and the subsequently increased cloud cover could lead to a rise in surface temperature by enhancing downward longwave radiation (Dai et al., 1999), which would further reduce the diurnal temperature range and affect sea breeze-like circulations (Vose et al., 2005; Davy et al., 2017; Cox et al., 2020). Climate models suggest~~s~~ that, compared to daytime, the slower decline in the long-term trend ~~in~~ of night~~-~~time cloud cover could raise the global temperature and amplify climate warming (Luo et al., 2024). However, how these cloud morphology types behave under the influence of the nighttime MBL regime and how nighttime cloud cover varies under different cloud morphology types remain unclear.~~how much they contribute to nighttime cloud cover variation remain unclear.~~ In addition, marine precipitation is more frequent at night (Dai, 2001; Dai et al., 2007), with its intensity strongly dependent on cloud morphology types (Muhlbauer et al., 2014). Therefore, comparing the differences in cloud morphology between daytime and nighttime may help explain the uneven distribution of precipitation, as well as improve our understanding and prediction of global precipitation changes against the backdrop of climate warming.

Motivated by the aforementioned issues, a new 1° × 1° classification dataset of daytime and nighttime marine low-cloud mesoscale morphology was generated in this study using a residual network model. In contrast to previous cloud classification datasets, our dataset provides global coverage with a temporal resolution of 5 minutes and a spatial resolution of 1° ~~over a 5-year period (~~from 2018~~–~~ to 2022~~)~~, which makes it ~~and can~~ better integrate with other ~~reanalysis~~ standard-grid datasets to deliver~~offer~~ more precise information about the meteorological conditions and ~~environmental~~ aerosols. The manuscript is organized as follows. Section 2 introduces the datasets and methods. Section 3 presents the training results and the contents of our dataset. The advantages and limitations of this dataset are discussed in Section 4. Section 5 states the data availability and Section 6 concludes.

2 Data and methods

2.1 Cloud type classifications

We adopted the classification scheme in Yuan et al. (2020) for mesoscale morphological classification of marine low clouds~~,~~. ~~and~~ The examples of each morphology classification are shown in Fig. 1. Solid stratus clouds are driven~~created~~ by cloud-top radiative cooling method and have a flat and uniform surface. Closed MCC clouds are stratocumulus driven by longwave radiative cooling and surface fluxes and display distinctive honeycomb-like structures with clear and descending edges. Open MCCs have a clear descending region in the center, which is surrounded by several active shallow convective clouds. They appear in more unstable environment and typically generate heavier drizzle, lower shortwave reflectance, and greater transmissivity than closed MCC (Wang and Feingold, 2009; Muhlbauer et al., 2014). Disorganized MCC are a mix of convective elements and extensive stratiform clouds, marked by ~~larger~~ smaller droplets and lower optical thickness (Yuan et al., 2020; Liu et al., 2024). They tend to occur in a drier troposphere and over warmer oceans (Wyant et al., 1997; Bretherton et al., 2019). Clustered Cu refers to the aggregation of shallow, vigorous convective elements, while suppressed Cu consists of individual, scattered cumulus clouds that occasionally form linear or branched patterns. Both of them are frequently ~~and extensively~~ observed over warm tropical oceans (Yuan et al., 2020; Mohrmann et al., 2021).

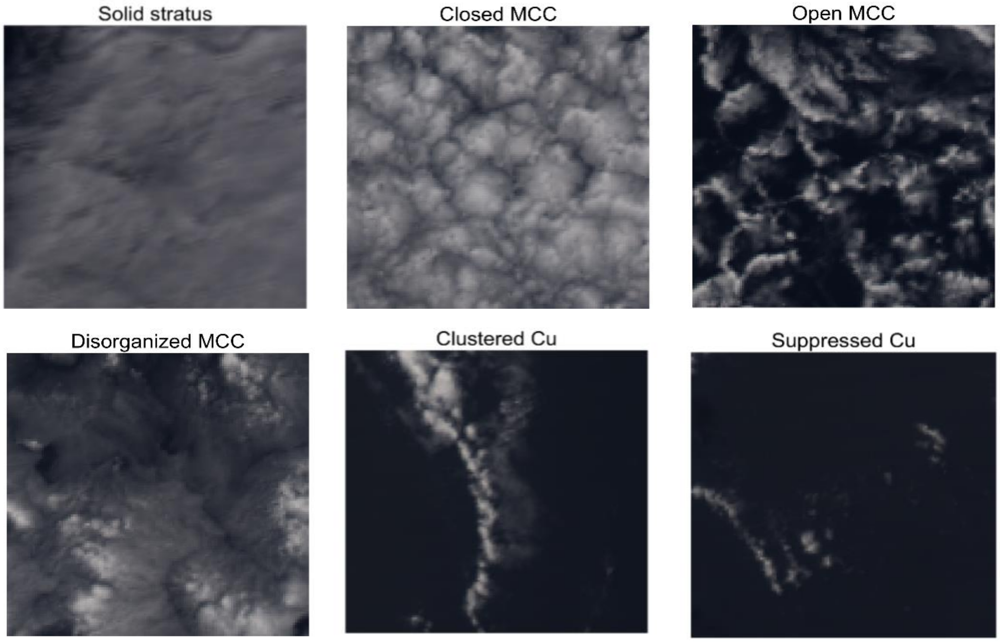


Figure 1: Example scenes of six cloud morphological types: solid stratus, closed mesoscale cellular convection, open mesoscale cellular convection, disorganized mesoscale cellular convection, clustered cumulus and suppressed cumulus. They are visible light images composed of channels 1, 4, 3 and the spatial resolution is 1° ~~o~~ × 1° ~~o~~.

2.2 Data

The primary observation data utilized in this study were derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard NASA's Aqua satellite, including the Level-1B radiance product MYD021KM and the Level-2 cloud product MYD06 (Platnick et al., 2017), both with a spatial resolution of 1 km at nadir point. ~~.~~ The thermal infrared radiance data from MYD021KM were used for model training and testing, while the cloud properties from MYD06 were utilized for quality control, low clouds filtering and statistical analyses~~, as explained below~~.

First, we selected daytime MODIS images ~~-~~over the Southeast Pacific (SEP) from January to June ~~the first half~~ of 2014 to create our ~~training~~classification dataset. The representativeness of this dataset was validated as the probability density functions (PDFs) of thermal radiance data and cloud optical thickness show large overlap with those of the global and full-year dataset (Fig. S1). After ~~breaking~~ dividing them into 128 × 128 pixels scenes, we filtered out the cloudless scenes (cloud fraction less than 1%) and scenes containing a large amount of high clouds or ice clouds (high/ice cloud ~~not~~ exceeding 10%). High clouds are defined as those with cloud top height above 6km, ice clouds are those with cloud top temperature below

273.15K. In addition, the severe stretching at the edge of MODIS granules has been avoided by filtering scenes with sensor zenith angle greater than 45°. Ultimately, these eligible scenes are manually classified ~~as~~ into one of the ~~solid stratus, closed MCC, open MCC, disorganized MCC, clustered Cu and suppressed Cu~~six categories. ~~We used t~~The cloud properties from MYD06 product, such as cloud top height (CTH), cloud liquid water path (CLWP), cloud optical thickness (COT) and cloud effective radius (CER)~~, to~~ are used to help ~~us~~ label and check~~ed our~~ the results with the cloud dataset ~~from~~ (Mohrmann et al., 2021). ~~As a consequence, we obtained a~~A total of 38,756 labeled daytime scenes were obtained, including 3,548 scenes of solid stratus, 6,277 of closed MCC, 3,345 of open MCC, 6,739 of disorganized MCC, 8,947 of clustered Cu and 9,900 of suppressed Cu. These scenes were then randomly partitioned into three mutually independent datasets for training, validation, and testing, with a distribution ratio of 3:1:1 respectively.~~These scenes were then divided into training, validation, and test dataset in a 0.6:0.2:0.2 ratio.~~ Despite the disparity in sample sizes within our training dataset, it is capable of yielding superior model performance compared to a balanced dataset (Fig. S2). ~~These scenes were then divided into training, validation, and test dataset in a 0.6:0.2:0.2 ratio.~~

In order to~~To~~ classify daytime and nighttime morphological types ~~only~~ using one model only, we utilized daytime radiance data ~~from thermal infrared channels 29 (8.7μm), 31 (10.8μm) and 32 (12.0μm)~~ to train our model. Thermal infrared (TIR) channels 29 (8.7μm), 31 (10.8μm) and 32 (12.0μm) were specifically chosen as they most effectively represent the cloud properties and cloud-top temperature. Notably, ~~due~~ owing to the subtle temperature variations on the cloud top, our model is unable to comprehensively discern convective cellular structures within the clouds by ~~only depend~~using ~~on~~ radiance data only ~~that reflects temperature~~. ~~Thus, i~~Incorporating COT can better address the model's shortcomings in studying these cellular structures by providing ~~more information about the~~cloud thickness ~~of clouds~~information. Considering that there are no nighttime COT in the Level-2 cloud product MYD06, we used the COT data retrieved by Wang et al. (2022) as the fourth channel input for our model. Their all-day COT products, obtained using a thermal infrared CNN model, have shown a good consistency with both ~~MODIS~~ daytime products from MODIS and ~~active sensors'~~ all-day products from active sensors. To validate the reliability of using TIR-CNN-based COT as a replacement for MODIS COT, we conducted a sensitivity experiment: comparing our classification with the outputs of a CNN trained on MODIS daytime COT. The results (Fig. S3) showed that the accuracy of both models is nearly identical, indicating that TIR-CNN-based COT is a reliable alternative to MODIS COT. In addition, we further examined the differences in the PDFs of the thermal radiance data and the TIR-CNN-based COT between our training dataset (daytime) and nighttime dataset. As depicted in Fig. S4, these PDFs nearly overlapped, which means less extrapolation will be introduced when the model is generalized to nighttime data. And it also illustrates the credibility of our nighttime classification results. In summary, ~~the training of our model was based on daytime radiance data and COT. Once the model is well-trained, it can be generalized to nighttime classification.~~ all ~~The~~ the variables and datasets used in this study are outlined in Table 1.

Table 1 Summary of datasets used in the study.

Dataset	Count	Channels	Day/Night	Size	Period
Training dataset	23,254	29,31,32, COT	Daytime	128×128	January–June 2014
Validation dataset	7,751	29,31,32, COT	Daytime	128×128	January–June 2014
Testing dataset	7,751	29,31,32, COT	Daytime	128×128	January–June 2014
Application dataset	18 million	29,31,32, COT	Daytime and Nighttime	1°×1° (128×128)	2018–2022

Before starting training, we first converted the radiance data into Brightness Temperature (BT) according to the inverse Planck Function shown in Eq. (1):

$$\mathrm{BT}(\lambda, \mathrm{L}) = \frac{\mathrm{C}_2}{\lambda \ln(\mathrm{C}_1/\lambda^5 \mathrm{L}+1)}, (1)$$

where λ is the wavelength (μm), L represents the radiance (W/m^2·sr·μm), ~~and~~ C_1 = 1.191042×10^8(W/m^2·sr·μm^{-4}), and C_2 =1.4387752×10^4 (K·μm). ~~Then w~~We combined BT data from the three thermal infrared channels according to the Day and Night colour scheme proposed by Lensky and Rosenfeld (2008) (Table 2). Table 2 includes both the original Day and Night color scheme and our modified scheme used in this study. Different from the original Day and Night scheme, we did not clip each scene's data to a fixed range of maximum and minimum values. We think ~~because~~ clipping might lead to the loss of important information, such as convective cell characteristics, which will~~thereby~~ affect~~ing~~ model performance. In our scheme,~~Although~~ ~~this behavior~~the data of each scene may ~~lead to the data from each scene being~~be compressed into different ranges and cause slight ~~variations in the~~ color variations ~~of~~ in each scene image (Fig. 4), it has little impact on the model's judgment capabilit~~y~~ies since the convolutional neural network primarily focuses on the statistical relationships between adjacent pixels in satellite images (Goodfellow, 2016). Moreover, after multiple rounds of practical training adjustments, we decided to use a factor of 2 to stretch the green channel to achieve a better model prediction outcome~~s~~. ~~The original Day and Night colour scheme and the modified scheme ultimately used in this study are both shown in Table 2. In the end, t~~To enhance the training efficiency and accuracy, the combined BT data and COT are normalized using Min-Max normalization following Eq. (2):

$$x_{\mathrm{n}} \cancel{x'} = \frac{x - min(x)}{max(x) - min(x)}, (2)$$

where x represents the input data, x_n ~~x'~~ represents the data after normalization, min(x) and max(x) represent the minimum and maximum values of the input data, respectively.

Table 2 The original (adapted from Lensky and Rosenfeld (2008)) and modified Day and Night color schemes

~~**Table 2 The original (adapted from Lensky and Rosenfeld (2008)) and modified Day and Night color schemes**~~

Color scheme	Red				Green				Blue			
	Channel	Min	Max	Stretch	Channel	Min	Max	Stretch	Channel	Min	Max	Stretch
Original Day and Night	IR12.0–IR10.8	−4 K	2 K	Linear	IR10.8–IR8.7	0 K	6 K	Γ =1.2	IR10.8	248K	303K	Linear
Modified Day and Night	IR12.0–IR10.8	min	max	Linear	IR10.8–IR8.7	min	max	Γ = 2	IR10.8	min	max	Linear

To align with conventional climate datasets, we developed a standard 1° ×1° gridded datasets by applying the trained model to 1°-resolution images, where the 1° × 1° satellite images were interpolated and refined to 128 × 128 pixels.~~To better match conventional climate datasets, we produced standard 1° gridded datasets by applying the trained model to a standard 1° scene, where one-degree-grid-sized satellite images were interpolated to 128 × 128 pixel.~~

For the purpose of investigating the influence of meteorological conditions on low-cloud morphologies, we conducted some statistical analyses utilizing the co-located hourly ERA5 reanalysis data (1° × 1°) from European Centre for Medium-Range Weather Forecasts (ECMWF). The co-location is achieved by spatially selecting the nearest ERA5 grid point to each MODIS observation and temporally interpolating the ERA5 data to match the exact time of the MODIS observations. This ensures accurate alignment between the two datasets in both space and time. ~~We further conduct some statistical analysis of meteorological conditions that may affect low cloud morphology using the co-located ERA5 reanalysis data (1° × 1°, 1-hourly) from the European Centre for Medium-Range Weather Forecasts (ECMWF).~~ Several variables, such as sea surface temperature (SST), relative humidity (RH), vertical velocity (ω) and divergence (1000 hPa, 700 hPa), can be directly obtained from the reanalysis data, while lower tropospheric stability (LTS) needs to be calculated using the following equation (3):

$$LTS = \theta_{700\mathrm{hPa}} - \theta_{1000\mathrm{hPa}} \, , (3)$$

where θ is the potential temperature.

Furthermore, we ~~also~~ retrieved 5-year daytime and night-time CER (r_e) and COT (τ) using the TIR-CNN model from Wang et al. (2022) for subsequent statistical analysis ~~of~~ for cloud properties. This approach will ensure the consistency of data range by using the same cloud detection algorithm. ~~Based on them, w~~We can further calculate the liquid water path (LWP) utilizing Eq. (4):

$$\mathrm{LWP} = \frac{2}{3}\rho_w \tau r_e \text{ , (4)}$$

with ρ_w the density of liquid water.

2.3 Marine Low-cloud Mesoscale Morphology Dataset

Our cloud dataset provides global classifications of daytime and nighttime marine low-cloud mesoscale morphology for the years 2018-2022, with a spatial resolution of 1° × 1° and a temporal resolution of 5 minutes. The dataset is provided in two kinds of files: those prefixed with "day" store the daytime classification results for each year, while files with the prefix "night" contain the nighttime classification results for each year. Both sets of files include the same variables. Table 3 provides an overview of the variables and their associated information. The key variables in the dataset include 'date' (representing the time of the 1° × 1° scene, formatted as the MODIS granule date), 'lon' and 'lat' (indicating the central longitude and latitude), and 'cat' (assigned cloud category, the values from 0 to 5 correspond to 'Solid Stratus', 'Closed MCC', 'Open MCC', 'Disorganized MCC', 'Clustered Cu', and 'Suppressed Cu', respectively). Additionally, 'cert' represents the model certainty, quantifying the probability that the cloud morphology belongs to the assigned category. 'low_cf' denotes the low cloud fraction, and 'COT_CNN', 'CER_CNN', and 'LWP_CNN' provide the in-cloud average cloud optical thickness, effective radius, and liquid water path respectively, as derived from the TIR-CNN model from Wang et al. (2022). The 'Sensor_zenith' variable indicates the scene average sensor zenith angle.

Table 3 Variables of the Daytime and Nighttime Global Marine Low-cloud Mesoscale Morphology Dataset

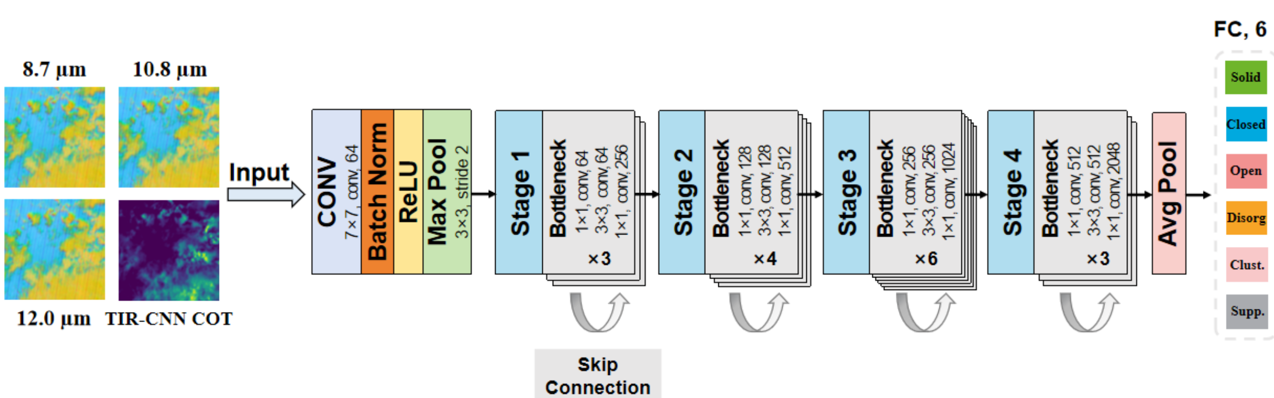
Variable Name	Description	Source	Spatial Resolution	Temporal Resolution	Units
date	Time of the 1°×1° grid point, formatted as 'YYYYDDD.HHHH'	MODIS MYD021	1° × 1°	5 minutes	-
lon	Central longitude (-180,180)	MODIS MYD021	1° × 1°	5 minutes	degree (°)
lat	Central latitude (-60,60)	MODIS MYD021	1° × 1°	5 minutes	degree (°)
cat	Category of the cloud morphology: 0-Solid stratus, 1-Closed MCC, 2-Open MCC, 3-Disorganized MCC, 4-Clustered Cu, 5-Suppressed Cu	Cloud Classification Model	1° × 1°	5 minutes	-
cert	Model certainty	Cloud Classification Model	1° × 1°	5 minutes	-
low_cf	Cloud fraction of low clouds	MODIS MYD06	1° × 1°	5 minutes	-
COT_CNN	In-cloud average cloud optical thickness	TIR-CNN model of	1° × 1°	5 minutes	-

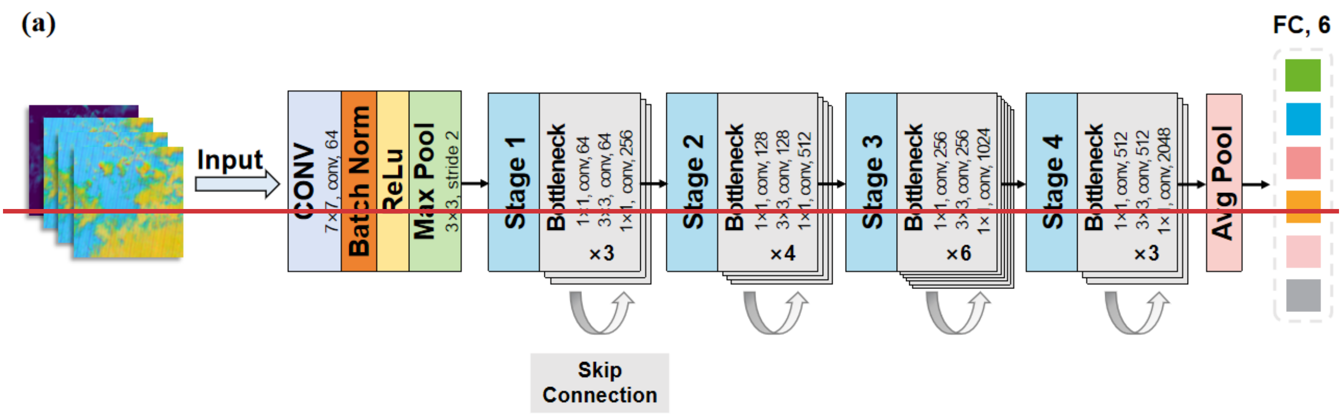
	(COT)	Wang et al. (2022)			
CER_CNN	In-cloud average cloud effective radius (CER)	TIR-CNN model of Wang et al. (2022)	1° × 1°	5 minutes	μm
LWP_CNN	In-cloud average cloud liquid water path (LWP)	Calculated from COT_CNN and CER_CNN	1° × 1°	5 minutes	g/m^2
Sensor_zenith	Scene average sensor zenith angle	MODIS MYD021	1° × 1°	5 minutes	degree (°)

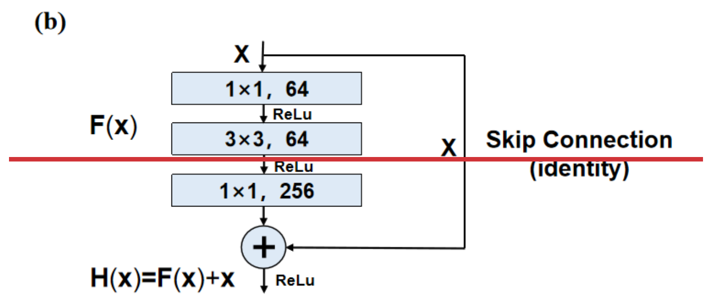
2.~~3~~ 4 Method

In this study, a~~n~~ machine learning (ML) model ResNet-50 (Koonce, 2021) was chosen as our model architecture. It is a deep CNN model which employs a residual learning framework to construct a network with 50 convolutional layers. Despite a fairly deep convolutional layer, the incorporation of residual units in ResNet-50 enables direct signal transmission from earlier to later layers, ensuring high computational efficiency in deep architectures and markedly boosting both accuracy and the speed of convergence. We made some adjustments to the overall architecture of ResNet50 to better suit our datasets and the fine-tuned model structure is presented in Fig. 2a. The number of input channels was set to 4 to include the additional COT channel. Then, we configured the output dimension of the final fully connected layer to 6 to produce a probability distribution over the 6 output classes for each scene via a softmax activation function. The internal structure of ResNet50 remains unchanged, consisting of a preprocessing layer, four stages, and a global average pooling. The preprocessing layer includes a convolutional layer, a batch normalization (BN) layer, a ~~ReLU~~ ReLU activation function, and a Max Pooling layer. Each stage contains several residual blocks and is connected by skip connections (Fig. 2b).

(a)







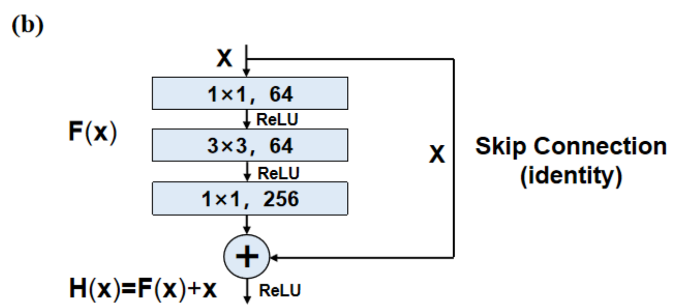


Figure 2: (a) The fine-tuned ResNet50 model architecture. (b) The skip connection structure of the residual blocks in the model.

Throughout the training process, we employed the Adaptive Moment Estimation (Adam) optimizer for gradient descent calculation and utilized cross-entropy as the loss function. Given the substantial size of our training dataset, we chose a batch size of 256 to enhance memory utilization and expedite the training process. In ~~Addition~~additional~~ly~~, to counteract the tendency for overfitting due to the increased noise in the radiance data of thermal infrared channels, we applied random rotation augmentation to the training images with a 50% probability. L2 regularization was also introduced to further decay the weights and prevent overfitting.

After each training epoch, the validation dataset was used to evaluate the trained model's performance, which allows us to monitor the model's success. When the training process was completed, the test dataset was used for the final evaluation of the model's performance. We used the optimal model to predict each sample in the test dataset, compared the model’s predictions with the true labels, and assessed the accuracy using metrics such as accuracy, F1-score, and recall.

3 Results

3.1 Model performance

In this section, our model performance is evaluated and a nighttime classification example is given.~~we present the performance evaluation of our model, as well as a classification example at night.~~

Our model begins to show signs of convergence around the 50th epoch, with its maximum validation accuracy reaching around 92% by the 65th epoch (Fig. ~~S1~~S5). Compared to training without the incorporation of COT, the model's accuracy improved by 4%. Although it is a bit less accurate compared to the visible light model from Yuan et al. (2020)~~ due to the limitations of the brightness temperature data in describing cloud morphology~~, it is undeniable that this model has ~~already reached~~achieved a relatively high accuracy level when compared to other TIR model (Lang et al., 2022), and ~~it~~ can effectively accomplish the classification tasks we proposed. The optimal model is subsequently evaluated on an independent test dataset, yielding the confusion matrix illustrated in Fig. 3. Elements on the diagonal represent the model's prediction accuracy for each type. Our model achieves an average precision of approximately 91%, an F1-score of 90.6%, and a recall of 90.8%, demonstrating its strong generalization capability and robustness. ~~The average precision across all types is approximately 91%. Compared with the infrared model from Lang et al. (2022), our model yields the same average prediction accuracy for closed and open MCC clouds, with closed MCC accuracy being 5% lower but open MCC accuracy being 5% higher than theirs.~~

In our model, closed MCC is more likely to be misidentified as solid stratus or disorganized MCC, while disorganized MCC tends to be misclassified as clustered Cu or closed MCC (Fig. 3). As observed in some classification samples, these misclassifications are partly attributed to the existence of mixed and transitional scenes (Fig. ~~S2~~S6). In addition, considering the similarity between the morphology of these clouds, the misclassifications may be related to the model's limited capacity to distinguish between stratiform structures and convective cells due to the small temperature difference on the cloud top. This can also be reflected in the sample images (Fig. ~~S3~~S7).

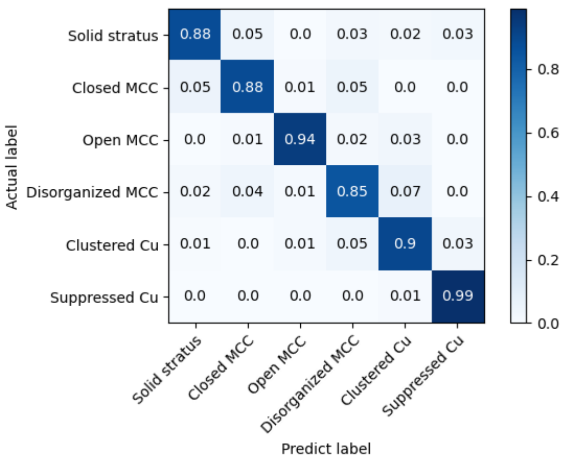


Figure 3: The confusion matrix of the model's predictions on the test dataset. The rows of confusion matrix represent actual categories, columns represent predicted categories. The elements on the diagonal indicate the proportion of samples correctly classified by the model in each category.

Training by daytime infrared data, this model can be applied to nighttime scenarios, as shown in Fig. 4~~By training on daytime infrared data, this model can be applied to nighttime scenarios~~. ~~Here is a successful example shown in Fig. 4.~~ Circles with different colors represent different cloud categories within the 1°×1°~~°~~ grid. Grids without circles indicate that they do not meet the criteria outlined in Section 2. The pseudo-RGB images are composed of thermal infrared channels 29, 31, 32 following the modified Day and Night color scheme (Table 2). This scheme provides a clearer visual distinction between different cloud types. However, as the data range for each image ~~was~~ is not fixed, the colors of different cloud types will vary in different situations. For example, in the left granule image, light yellow represents low water clouds, green ~~indicates~~ indicating thin cirrus~~,~~ and dark yellow signif~~i~~ying~~es~~ thick cumulonimbus clouds. In the sample scenes of open MCC and suppressed Cu, the yellow of low clouds becomes lighter and the green indicates small cumulus cloud. As for the remaining four cloud types, the surrounding thin cirrus appears in green while the stratiform clouds and shallow cumulus convections are both depicted in a brighter yellow due to their similar temperatures. Thus, it is challenging to discern convective cells among the yellow background of stratiform clouds. That~~'~~ is why we incorporated COT to assist in model predictions.

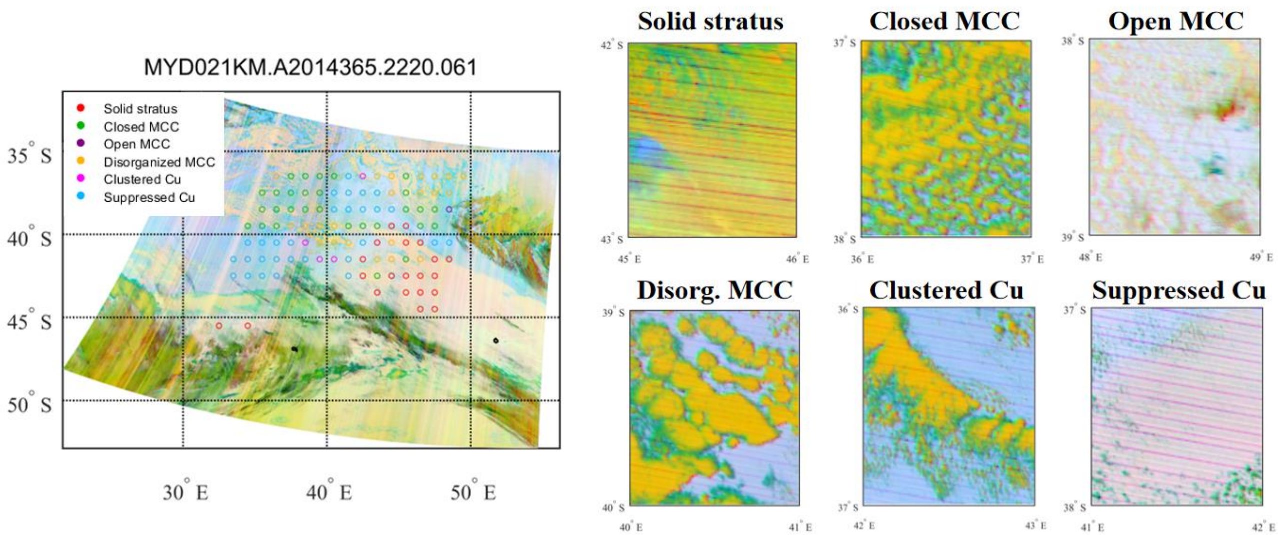


Figure 4: A nighttime classification example for MODIS image taken at 22:20 UTC on December 31, 2014. The pseudo-RGB images were generated from the combination of 29, 31, 32 thermal infrared channels while the classification results ~~are~~ were derived by incorporating the retrieved Cloud Optical Thickness (COT) data.

3.2 Climatology of morphological types

Using the well-trained ResNet-50 model, we classified nearly 18 million 1° MODIS scenes and recorded the occurrence counts of different cloud types. The occurrence counts of each cloud type were divided by the total occurrences of the six cloud types within each grid to calculate their relative frequency of occurrence (RFO). The daytime climatology of RFO for the six cloud types is presented in Fig. 5. Each subplot's upper-right corner displays n and a percentage, with n denoting ~~n denotes~~ the total number of occurrences for each cloud type during the daytime over the five-year period, while the percentage indicates the proportion of each cloud type's five-year occurrences (n) relative to the five-year total occurrences of all six cloud types, which called total relative frequency.

Solid stratus predominantly distributes in nearly symmetrical latitude bands between 40°N–60°N and 50°S–60°S with a total relative frequency of 14%. In the mid to high latitudes of the Southern Hemisphere, its RFO exceeds 90%, which is higher than that in the Northern Hemisphere. Additionally, a substantial presence of solid stratus is also observed along the western coasts of continents in tropical and subtropical regions. Closed MCCs mainly appear in the cold eastern subtropical and mid-latitude oceans, with marked peak along the western coasts of the North America, South America and Africa. Their total relative frequency during the daytime is relatively low, accounting for only 5%. Disorganized MCCs exhibit a distribution pattern similar to closed MCC but are typically located farther offshore. They cover more extensive area and occur more frequently. The total relative frequency of disorganized MCC during the daytime is 15%, three times higher than closed MCC.

In addition, it is worth noting that, the peak areas of disorganized MCC ~~are found~~appear in the west of closed MCC. This may be related to the transition between these two cloud types. The occurrence of open MCCs is least frequent over the global ocean, accounting for only 3%. In the ~~waters~~ water areas west of Peru, there is a minor frequency peak of open MCC, which may be attribute to the fragmentation of closed MCC caused by strong winds and precipitation (Rosenfeld et al., 2006; Eastman et al., 2022). Clustered Cu and suppressed Cu are primarily observed in tropical and subtropical regions. They have the highest overall relative frequencies, both around 30%. However, in terms of their spatial distribution, clustered cumulus is more prevalent over central and western oceans, while suppressed Cu commonly peaks in coastal waters near continents. We have compared the daytime climatology with the results from Yuan et al. (2020) using visible light channels and consistent results are obtained ~~We compared the daytime climatology with the results from Yuan et al. (2020) using visible light model and found consistent outcomes~~ (Fig. ~~S4~~S8).

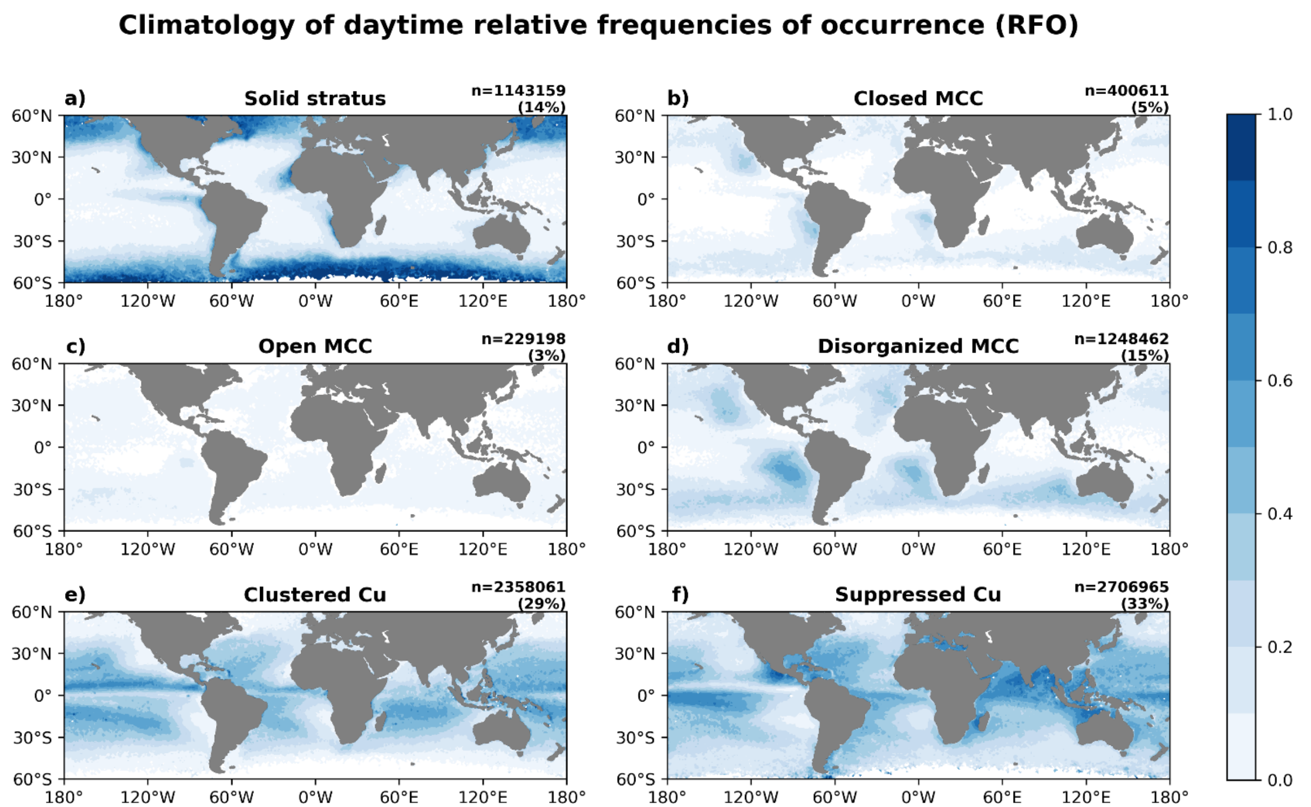


Figure 5: The climatology of daytime relative frequencies of occurrence (RFO) for six categories from 2018 to 2022. N represents the total number of occurrences for each cloud type during the day over the five-year period, while the percentage indicates the proportion of each cloud type's five-year occurrences relative to the five-year total occurrences of all six cloud types.

~~Although~~ In Fig. 6, the spatial distribution of these six cloud types remains largely unchanged at night, but their RFO show notable variations~~ in contrast to daytime (Fig. 6)~~. Figure 7 presents the nighttime–daytime contrast in RFO for each morphological type (nighttime minus daytime). The total nocturnal frequency of solid stratus clouds is 15%, which is similar to daytime. At night, they occur more frequently over mid-latitude oceans and less frequently in low-latitude regions~~ (Fig. 7)~~. The RFO for closed MCC shows a pronounced increase at night, reaching approximately twice the levels of the daytime~~ (Fig. 6)~~. Some of the increase occurs over mid-latitude oceans, while the most significant rise is observed over the eastern subtropical ocean, particularly in the Southern Hemisphere (Fig. 7). The overall frequency of open MCC remains relatively unchanged at night, while the total frequency of disorganized MCC and clustered Cu slightly increase~~ (Fig. 6)~~. At nighttime, the RFO of all these three cloud types decrease over the colder eastern subtropical and mid-latitude oceans, while increase over the warmer sea surface at lower latitudes (Fig. 7). Notably, westward from the continents, the night-time frequency pattern of disorganized MCC exhibits an initial decrease followed by an increase. This opposite trend is most pronounced along the western coast of South America. Among six cloud types, only the total frequency of suppressed Cu experiences a marked decline at night, with a total decrease of 11~~%.~~ %. A statistical analysis of some meteorological conditions will be conducted in Section 3.5. Exploration of the critical cloud-controlling factors contributing to these diurnal variations will be done in the future. ~~Explorations of the critical cloud-controlling factors contributing to these diurnal variations will be left for future work and we will only conduct a simple statistical analysis of several meteorological conditions in Section 3.5.~~

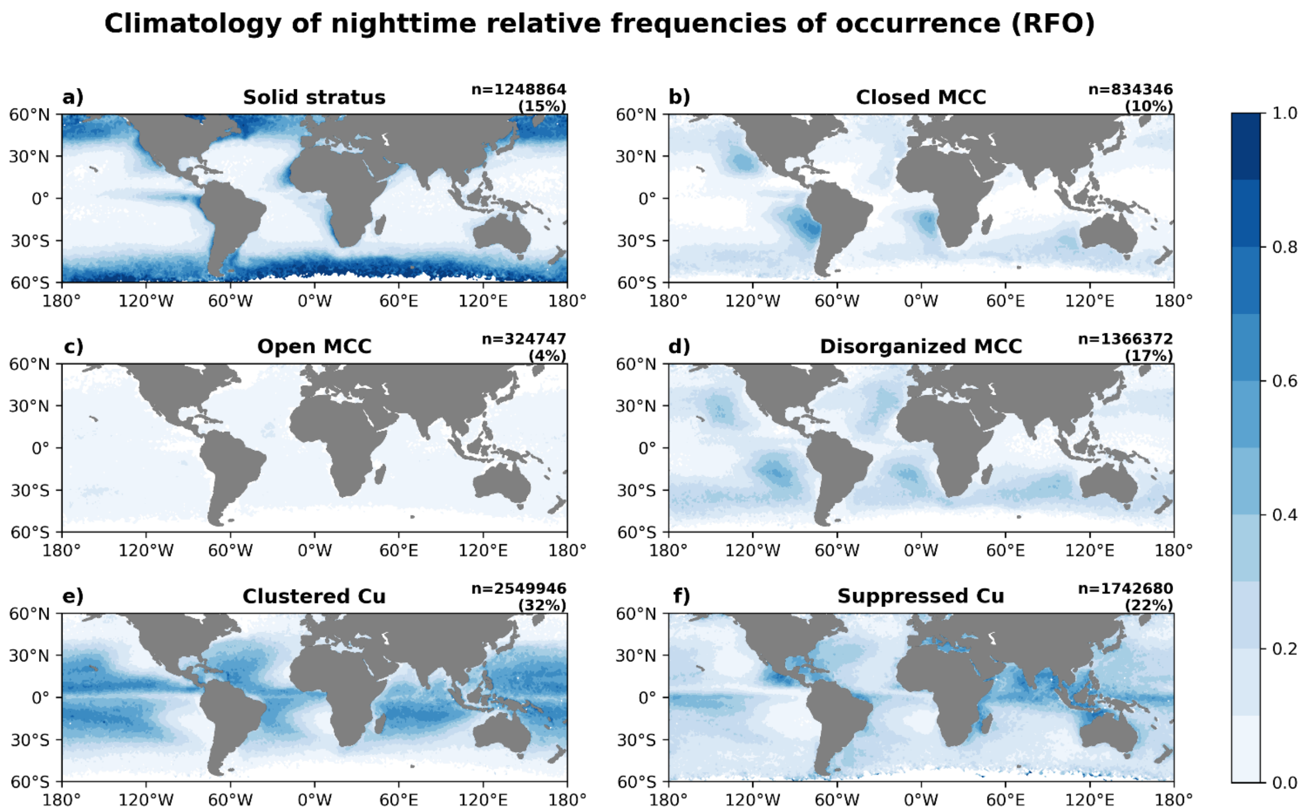


Figure 6: The climatology of nighttime relative frequencies of occurrence (RFO) for six categories from 2018 to 2022.

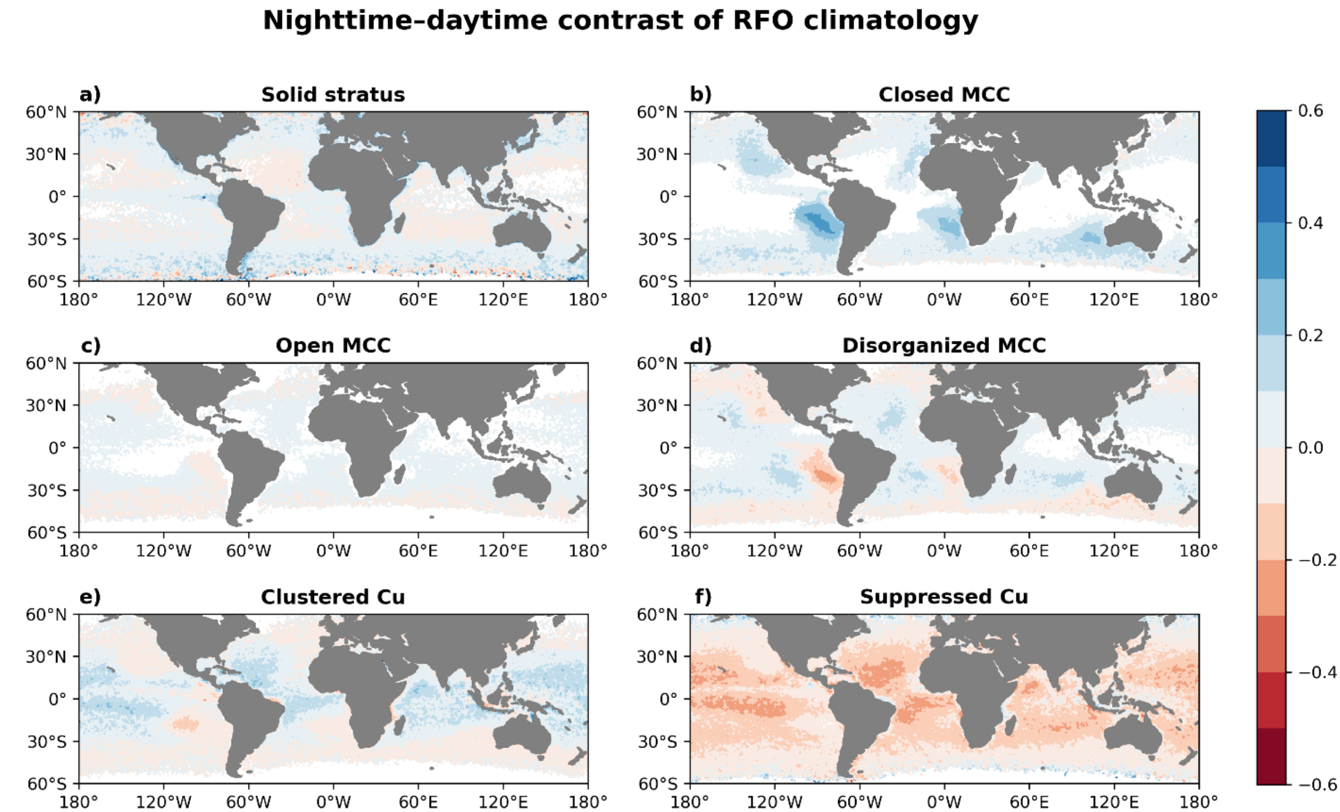


Figure 7: The difference between daytime and night-time RFO for each morphological type (nighttime minus daytime).

3.3 Seasonal variations in morphological types

We further classified the RFO of different cloud ~~morphology~~ morphologies by season. Figure 8 presents the seasonal variation of daytime RFO while Fig. 9 shows the nighttime situation. It can be seen from the two figures that the RFO of these six cloud types exhibit similar seasonality ~~during~~ in both day and night. At mid latitudes, solid stratus clouds usually peak during the summer of the respective hemisphere (JJA for the Northern Hemisphere, DJF for the Southern Hemisphere) and have the lowest occurrence during the winter (DJF for the Northern Hemisphere, JJA for the Southern Hemisphere). They show equal RFO during spring and autumn in both hemispheres (MAM and SON). The peak occurrence of solid stratus in mid-latitude regions aligns with the latitudinal shift of solar insolation. Thus it can be inferred that the increased temperature and enhanced moisture ~~availability~~ available from ~~melting~~ melted sea ice may contribute to the peak in summer ~~its seasonal variation~~ (Herman and Goody, 1976). The RFO of closed MCC notably increases during the winter (JJA) and ~~Spring~~ spring (SON) in the Southern Hemisphere, particularly in the southeast Pacific (SEP) and southeast Atlantic (SEA) regions. McCoy et al. (2017)

suggest that the seasonal cycle of closed MCC in such regions correlates well with estimated inversion strength (EIS). In contrast to solid stratus, open MCC demonstrates an opposite seasonal cycle in mid-latitudes, with the highest frequency occurring in the winter of respective hemisphere (DJF for the Northern Hemisphere and JJA for the Southern Hemisphere). Previous work suggests that its seasonality is more likely associated with cold air outbreaks in mid-latitude oceanic regions (McCoy et al., 2017). This may also explain why open MCC exhibits zonal frequency peak over the Southern Pacific during the winter of Southern Hemisphere (JJA) (Fig. 8 and Fig. 9). Disorganized MCC clouds occur more frequently over ~~the~~ warmer ocean surface western of the continents during the summer of respective hemisphere (JJA in the Northern Hemisphere and DJF in the Southern Hemisphere) and occur less frequently during the winter of respective hemisphere (DJF in the Northern Hemisphere and JJA in the Southern Hemisphere). Thus, the sea surface temperature may be one of the controlling factors of its seasonal variation. All the MCC types show distinct seasonal cycles while the clustered Cu and suppressed Cu do not show marked seasonal variations during both day and night.

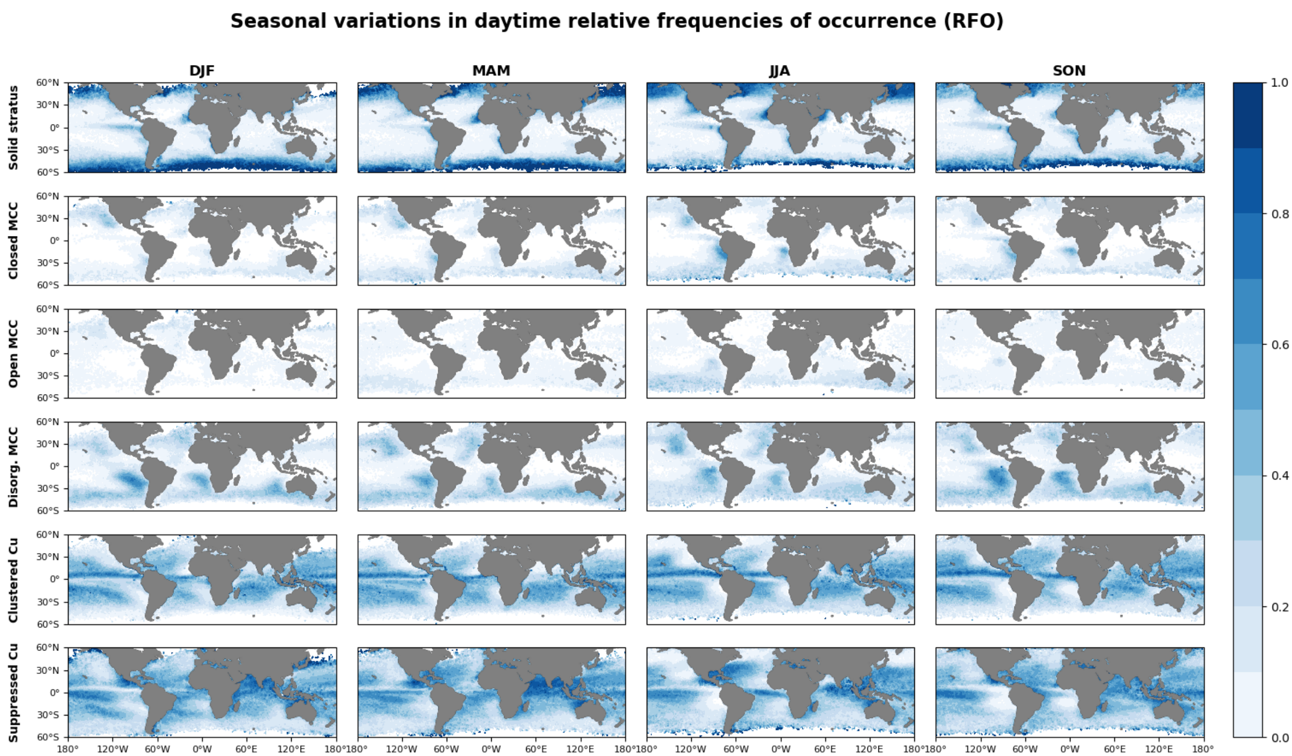


Figure 8: Seasonal variations ~~in~~ of daytime relative frequencies of occurrence (RFO).

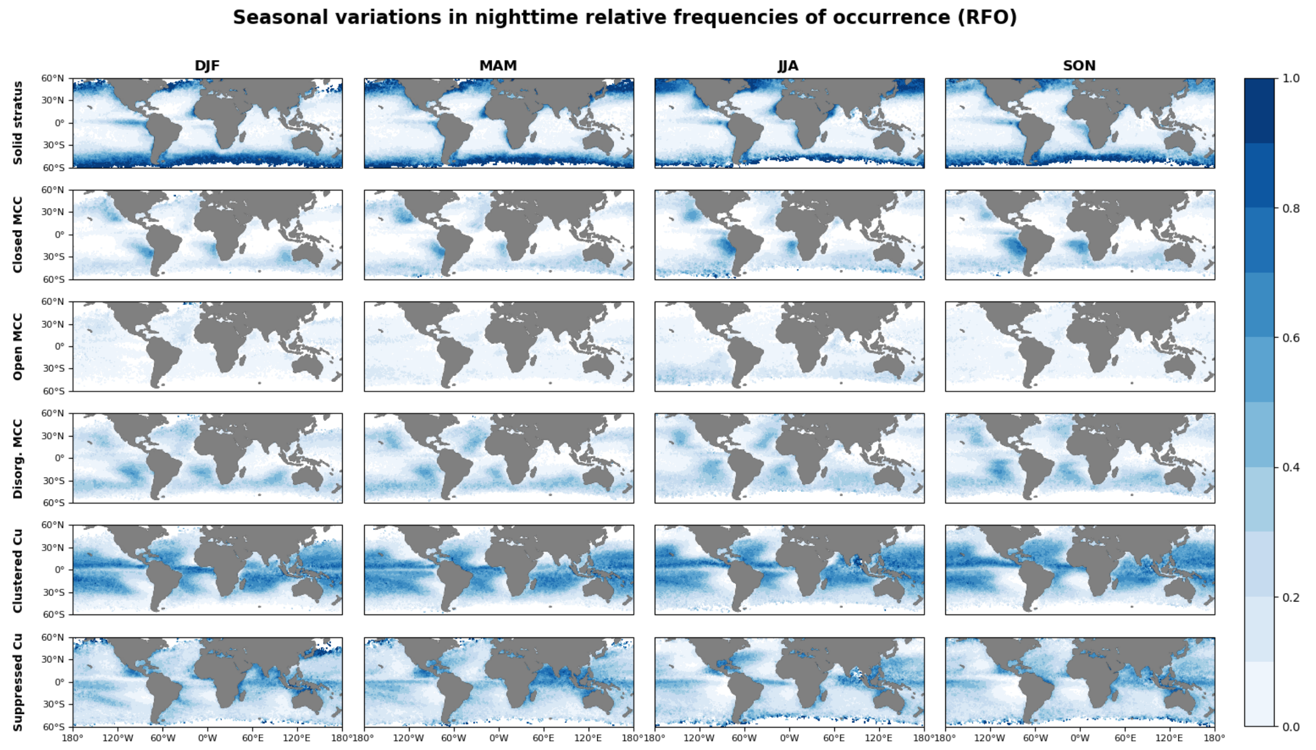


Figure 9: Seasonal variations ~~in~~ of nighttime relative frequencies of occurrence (RFO).

3.4 Cloud properties

Different clouds types exhibit different radiative effects due to their unique physical characteristics. In Fig. 10, we compared the physical properties of each cloud type ~~during~~ in both day and night, including cloud fraction (CF), cloud effective radius (CER), cloud liquid water path (LWP), and cloud optical thickness (COT). The CF is derived from the cloud mask in the Level-2 cloud product MYD06, while CER and COT are both retrieved using the method from Wang et al. (2022). LWP is calculated from CER and COT as mentioned in Section 2.2. All of the cloud microphysical properties represent the in-cloud mean value within a 1°×1° grid.

Solid stratus and closed MCC possess the highest CF, therefore the increase in their nocturnal frequency may account for a major portion of the overall rise in cloud cover. Open MCC possesses the largest CER and it will decrease by 2 ~~μm~~microns on average at night. ~~During~~ In the daytime, closed MCC clouds exhibit the highest values ~~for~~ of LWP and COT. At nighttime, their CER, LWP and COT increase further substantially, with a substantial magnitude. The four cloud properties of disorganized MCC also show a slight increase at nighttime.

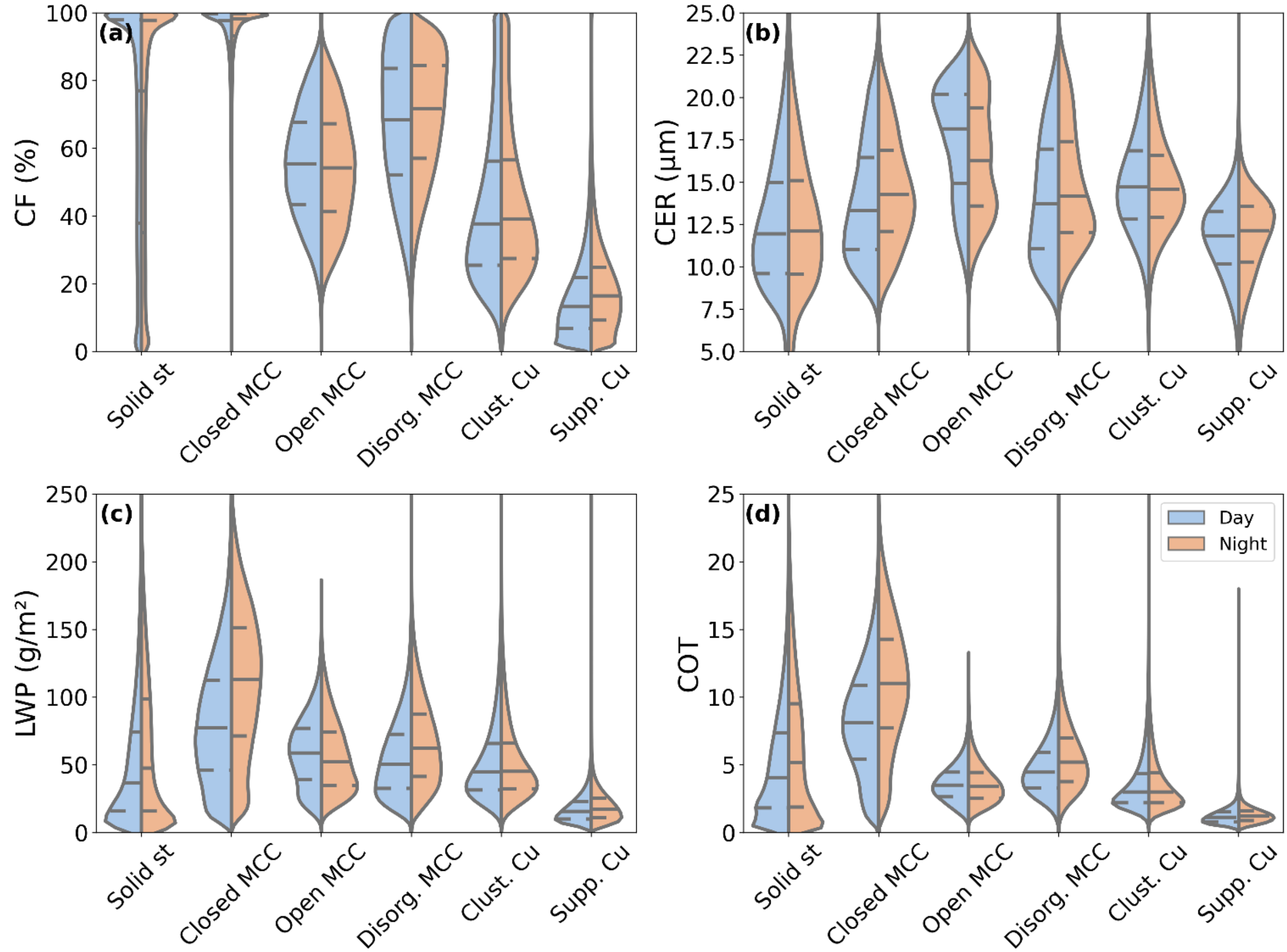


Figure 10: A violin plot of the properties of six cloud types over the global oceans from 2018 to 2022. (a) cloud fraction, (b) retrieved cloud effective radius, (c) cloud liquid water path, (d) retrieved cloud optical thickness. Blue represents the daytime data and yellow represents the nighttime data. The central long dashed line in each plot represents the median of the distribution, and the short dashed line indicates the interquartile range. The shape of the violin plots suggests the density distribution of the values, with wider sections indicating a higher frequency of data points.

3.5 Large-scale meteorological condition

The statistics of several meteorological factors which may control the marine low cloud morphology in the Southeast Pacific (SEP) region (0 – 30°S, 80°W – 120°W) are shown in Figure 11. The lower tropospheric stability (LTS) for the six cloud types is shown in Figure 11a. A higher LTS indicates a more stable lower troposphere. Closed MCC have the highest LTS, implying the significance of tropospheric stability in their formation. The two cumulus types have the lowest LTS because an unstable troposphere is conducive to cumulus activity. The LTS of six cloud types show different degrees of decline at nighttime, which may be due to the ~~westward~~ shift in their geographical locations. Sea surface temperature (SST) is the lowest for closed MCC during both day and night (Fig. 11b). Open MCC and disorganized MCC exhibit higher SST compared with closed MCC, which corresponds to their geographical positions. Two cumulus types have the highest SST. At nighttime, the increase ~~in~~ of SST ~~of~~ for different cloud types may also be attributed to their ~~westward~~ movements. Figure 11c and 11d show the relative humidity (RH) at 700hPa and 1000hPa respectively. Throughout the day and night, solid stratus clouds exhibit the highest RH at both 700hPa and 1000hPa. At the 700hPa level (Fig. 11c), the RH values for two cumulus types are higher than those for MCC clouds, while at the 1000hPa level (Fig. 11d), the difference is minimal. At night and at 700 hPa, the RH of solid stratus and two cumulus types increases, while that of closed MCC decreases. Due to lower temperatures at nighttime, the ~~relative humidity~~RH over sea surface for all six cloud types increases by a similar magnitude. Figure 11e indicates that all cloud types are associated with large-scale subsidence. Open MCC experiences the strongest upper-level subsidence, while solid stratus has the weakest vertical motion. At nighttime, the subsidence for all six cloud types weakens and closed MCC exhibits a more pronounced reduction. Figure 11f presents the boundary layer anomaly divergence which is calculated by subtracting the divergence at 700 hPa from the surface divergence. This index has been proven effective in distinguishing between the two cumulus types (Mohrmann et al., 2021). Suppressed Cu shows the largest boundary layer anomaly divergence, indicating that strong surface divergence favors the maintenance of suppressed Cu. Clustered Cu has the smallest anomaly divergence, with weaker surface divergence. Therefore, the weakening of surface divergence at nighttime may be the reason for the reduction of suppressed Cu.

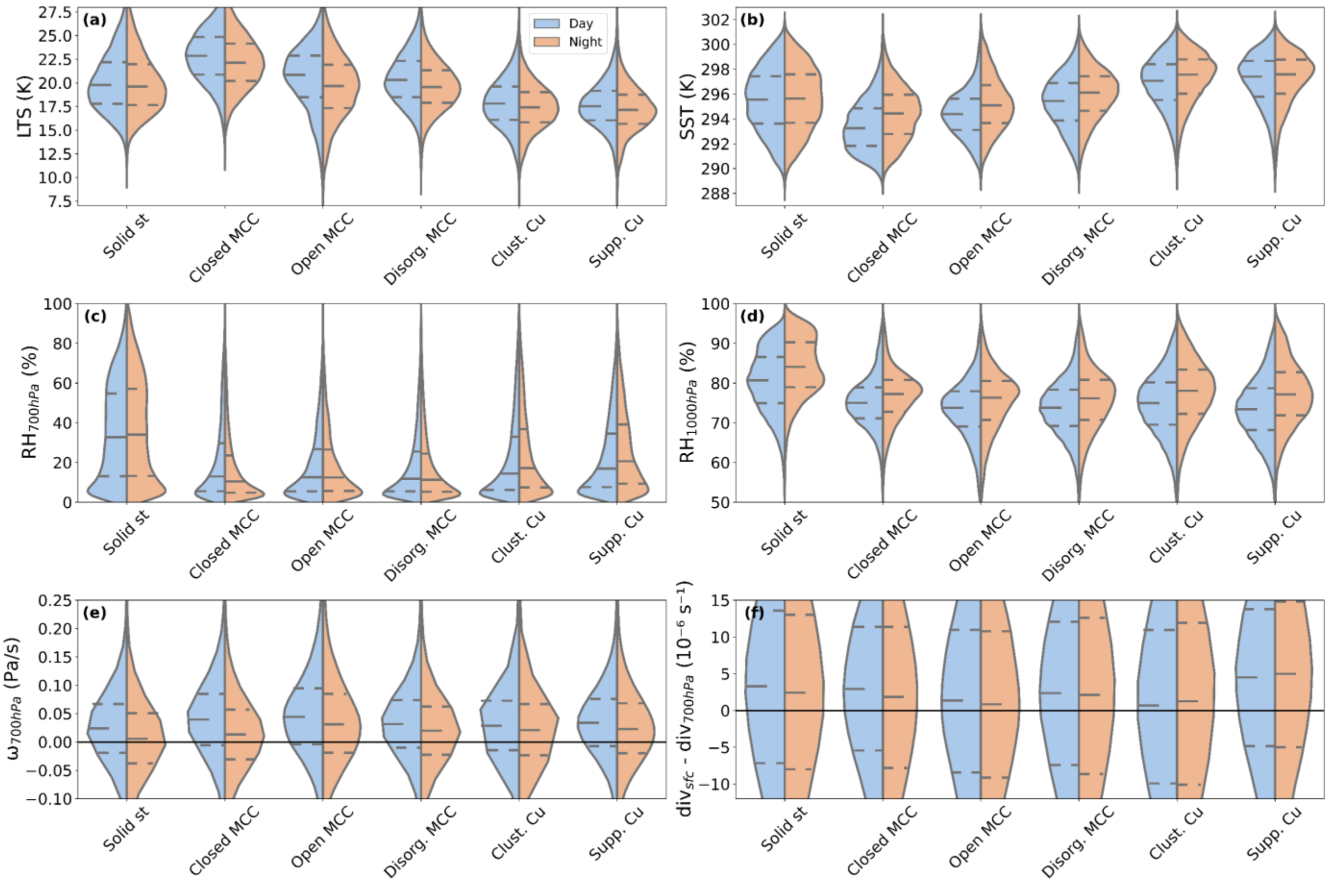


Figure 11: Same as Fig. 10, but shows day-night comparison of meteorological conditions in the Southeast Pacific (SEP) region (0–30°S, 80°W–120°W) from 2018 to 2022. All matched from ERA5 reanalysis data. (a) Lower tropospheric stability, (b) sea surface temperature, (c) 700hPa relative humidity, (d) 1000hPa relative humidity, (e) 700hPa vertical velocity, (f) boundary layer anomaly divergence.

4 Discussion

The mesoscale cloud morphology dataset presented in this paper enables a comparative investigation of cloud morphology during both daytime and nighttime. Its 1° × 1° resolution allows ~~for~~ a better alignment with other gridded datasets, facilitating further studies on driving factors, precipitation efficiency, and radiative effects (shortwave and longwave).

Although our model has achieved a high prediction accuracy and performed well in the classification tasks，there is still room for improvement. In future ~~research~~work, ~~model performance can be optimized through these two methods: replacing~~ the ~~classification~~ CNN model can be considered replaced and ~~improving~~ the quality of our training dataset need to be improved. For the former one, novel deep CNN models can be applied to cloud morphology classification through transfer learning. For example, the Xception model, which achieved an accuracy of 97.66% in classifying traditional cloud types (Guzel et al., 2024), could be considered. For the latter goal, removing the mixed and mislabeled scenes from the training dataset, along with adding more ~~representative~~ global and multi-seasonal scenes will improve the model performance in identifying these cloud morphological types. The ~~Limitations~~ limitations of brightness temperature in capturing cloud-top morphology significantly ~~impact~~ constrain the model's accuracy, which ~~is a key reason why~~largely explains the ~~our nighttime model's~~ performance gap between our nighttime model ~~is worse than that~~and the daytime model proposed by ~~of~~ Yuan et al. (2020) ~~daytime model~~. Nevertheless~~However~~, a 4% improvement~~increase~~ in model accuracy achieved by incorporating COT indicates~~with the inclusion of COT suggests~~ that integrating~~incorporating~~ additional cloud property channels represent a promising avenue~~is another feasible approach~~ for further enhancing ~~our~~ the ~~model~~ performance of our model.

As we apply the model trained on 128 × 128 pixels scenes to the interpolated 1° × 1° cloud scenes, the issue regarding to scene area require further discussion. For a given latitude, when the satellite zenith angle changes, the area of 128 × 128 pixel will vary due to the pixel stretching, while the area of the 1º grid remains constant. This is an advantage of the 1° × 1° grid dataset ~~since~~ as a ~~the~~ larger the area ~~is, the~~has a larger ~~more~~ possibility to cover multiple cloud types in one scene. However, at the same satellite zenith angle, the size of 1°× 1° grid will change with latitude, whereas the 128 × 128 pixel scene area remains undistorted, which is a limitation of the 1° × 1° grid dataset. Moreover, in some cases, especially with stratocumulus and cumulus clouds, interpolating images into a 1º× 1° grid may smooth or blur small-scale cloud features and introduce unrealistic structures that do not exist in the original images, which could lead to potential misclassifications of the model. Therefore, testing the model on a labeled, standard-grid dataset will be necessary in future work.

The six cloud types examined in this study are the most common and representative types over the ocean~~;~~. ~~however~~However, they are not exhaustive. In future work, we will explore the overall low-cloud morphological types over the global land and ocean, and gradually extend to mid and high-level clouds.

5 Data Availability

Daytime and nighttime cloud classification datasets as well as our training dataset are accessible on the https://doi.org/10.5281/zenodo.13801408~~https://doi.org/10.5281/zenodo.13990646~~ (Wu et al., 2024). The model and the code related to this article are available at https://github.com/YuanyuanWu-NJU/Cloud-morphology-dataset. MODIS data can be downloaded from NASA Official Website (https://ladsweb.modaps.eosdis.nasa.gov/). ERA5 reanalysis data are provided by ECMWF (https://cds.climate.copernicus.eu/datasets). The cloud property retrieval model of Wang et al. (2022) can be found at (https://github.com/WgQuan/cloud-property-retrievals).

6 Conclusion

In this study, approximately 40,000 MODIS daytime low-cloud scenes were manually labeled to train a deep residual network model, ResNet50. By using this model, we developed a new global standard-grid classification dataset (2018–2022) of marine low-cloud mesoscale morphology, encompassing classifications for both daytime and nighttime. Compared to the 128 × 128 pixel dataset of Yuan et al. (2020) and Mohrmann et al. (2021), our standard-grid dataset offers more uniform and widely applicable cloud morphology data, and more importantly, extending the dataset to ~~nightly~~nighttime. This dataset can integrate more easily with other climate and surface datasets, thus will provide a solid data foundation for future research on understanding cloud dynamics and their impact on climate.

The climatology of cloud morphologies is also ~~documented~~analyzed. The results reveal that solid stratus dominates within the 50°–60° latitude bands ~~in mid-latitude regions,~~ and that closed MCC is most commonly found in the cold eastern subtropical and mid-latitude oceans. Disorganized MCC occurs on the warmer ocean surfaces west of closed MCC, with a much higher frequency. Open MCC is more evenly distributed across the global oceans but with the lowest frequency. In regions with higher sea surface temperatures, such as the tropics and the trade wind zone~~s~~, clustered Cu and suppressed Cu are the primary types of marine low clouds.~~, with~~ ~~clustered~~ Clustered Cu is more prevalent over oceans and suppressed Cu is more concentrated along continental coasts.

When comparing the daytime and nighttime climatology, we found that there is a pronounced increase in the RFO of closed MCC during night, whereas the occurrence of suppressed Cu undergoes a significant decline. The frequencies of disorganized MCC and clustered Cu exhibit minor variation between day and night. From the perspective of different seasons, solid stratus and all MCC types exhibit clear seasonal cycles while two cumulus types do not show notable seasonality.

Although we have statistically analyzed~~did a rough statistical analysis~~ ~~of~~ the meteorological factors that may affect low cloud morphology, identifying the specific dominant factors for each cloud type remains challenging~~,~~. ~~and it is left for future work that could combine the dataset we proposed. Furthermore, i~~In the context of global warming, the long-term trends of these cloud types during ~~the~~ daytime and nighttime may also exhibit significant differences. The changes in Earth's radiation budget caused by~~-~~ low-cloud ~~them~~ morphology transitions may have a substantial impact on climate sensitivity, which will be a topic of our future research.

~~are an important component of the low cloud feedback, thus we intend to release a 20-year product for such research in the near future.~~

Acknowledgments

This work is supported by the National Nature Science Foundation of China (42475089, 41925023, 42075093, U2342223, and 91744208).

Author contributions

Y.Z. designed the study, Y.W., J.L., Y.C., Y.Z. and D.R. wrote and revised the manuscript. J.L., Y.L., Y.C. and Q.W. collected the data. J.L. and Y.Z. contributed to the modeling and data process. Q.W. and C.Z. provided the cloud property retrieval model and Y.W. retrieved the cloud property product. K.H., B.Z., Y.W., Y.L. and C.Z. contributed to interpreting results and discussions. The entire study was conducted under the supervision of Y.Z., M.W., J.S. and D.R.

Competing interests

The authors declare no competing interests.

References

Bony, S., Schulz, H., Vial, J., and Stevens, B.: Sugar, Gravel, Fish, and Flowers: Dependence of Mesoscale Patterns of Trade-Wind Clouds on Environmental Conditions, Geophysical Research Letters, 47, e2019GL085988, https://doi.org/10.1029/2019GL085988, 2020.

Bretherton, C. S., McCoy, I. L., Mohrmann, J., Wood, R., Ghate, V., Gettelman, A., Bardeen, C. G., Albrecht, B. A., and Zuidema, P.: Cloud, Aerosol, and Boundary Layer Structure across the Northeast Pacific Stratocumulus–Cumulus Transition as Observed during CSET, Monthly Weather Review, 147, 2083-2103, https://doi.org/10.1175/MWR-D-18-0281.1, 2019.

Cao, Y., Zhu, Y., Wang, M., Rosenfeld, D., Zhou, C., Liu, J., Liang, Y., Huang, K.-E., Wang, Q., Bai, H., Wang, Y., Wang, H., and Zhang, H.: Improving Prediction of Marine Low Clouds Using Cloud Droplet Number Concentration in a Convolutional Neural Network, Journal of Geophysical Research: Machine Learning and Computation, 1, e2024JH000355, https://doi.org/10.1029/2024JH000355, 2024.

Cox, D. T. C., Maclean, I. M. D., Gardner, A. S., and Gaston, K. J.: Global variation in diurnal asymmetry in temperature, cloud cover, specific humidity and precipitation and its association with leaf area index, Global Change Biology, 26, 7099-7111, https://doi.org/10.1111/gcb.15336, 2020.

Dai, A.: Global Precipitation and Thunderstorm Frequencies. Part II: Diurnal Variations, Journal of Climate, 14, 1112-1128, https://doi.org/10.1175/1520-0442(2001)014<1112:GPATFP>2.0.CO;2, 2001.

Dai, A., Lin, X., and Hsu, K.-L.: The frequency, intensity, and diurnal cycle of precipitation in surface and satellite observations over low- and mid-latitudes, Climate Dynamics, 29, 727-744, https://doi.org/10.1007/s00382-007-0260-y, 2007.

Dai, A., Trenberth, K. E., and Karl, T. R.: Effects of Clouds, Soil Moisture, Precipitation, and Water Vapor on Diurnal Temperature Range, Journal of Climate, 12, 2451-2473, https://doi.org/10.1175/1520-0442(1999)012<2451:EOCSMP>2.0.CO;2, 1999.

Davy, R., Esau, I., Chernokulsky, A., Outten, S., and Zilitinkevich, S.: Diurnal asymmetry to the observed global warming, International Journal of Climatology, 37, 79-93, https://doi.org/10.1002/joc.4688, 2017.

Eastman, R., McCoy, I. L., and Wood, R.: Wind, Rain, and the Closed to Open Cell Transition in Subtropical Marine Stratocumulus, Journal of Geophysical Research: Atmospheres, 127, e2022JD036795, https://doi.org/10.1029/2022JD036795, 2022.

Eytan, E., Koren, I., Altaratz, O., Kostinski, A. B., and Ronen, A.: Longwave radiative effect of the cloud twilight zone, Nature Geoscience, 13, 669-673, https://doi.org/10.1038/s41561-020-0636-8, 2020.

Feingold, G., Koren, I., Wang, H., Xue, H., and Brewer, W. A.: Precipitation-generated oscillations in open cellular cloud fields, Nature, 466, 849-852, https://doi.org/10.1038/nature09314, 2010.

Goodfellow, I.: Deep Learning, MIT Press2016.

Goren, T., Kazil, J., Hoffmann, F., Yamaguchi, T., and Feingold, G.: Anthropogenic Air Pollution Delays Marine Stratocumulus Breakup to Open Cells, Geophysical Research Letters, 46, 14135-14144, https://doi.org/10.1029/2019GL085412, 2019.

Guzel, M., Kalkan, M., Bostanci, E., Acici, K., and Asuroglu, T.: Cloud type classification using deep learning with cloud images, PeerJ Computer Science, 10, e1779, https://doi.org/10.7717/peerj-cs.1779, 2024.

Herman, G. and Goody, R.: Formation and Persistence of Summertime Arctic Stratus Clouds, Journal of Atmospheric Sciences, 33, 1537-1553, https://doi.org/10.1175/1520-0469(1976)033<1537:FAPOSA>2.0.CO;2, 1976.

Klein, S. A. and Hartmann, D. L.: The Seasonal Cycle of Low Stratiform Clouds, Journal of Climate, 6, 1587-1606, https://doi.org/10.1175/1520-0442(1993)006<1587:TSCOLS>2.0.CO;2, 1993.

Koonce, B.: ResNet 50, in: Convolutional Neural Networks with Swift for Tensorflow: Image Recognition and Dataset Categorization, Apress, Berkeley, CA, 63-72, https://doi.org/10.1007/978-1-4842-6168-2_6, 2021.

Lang, F., Ackermann, L., Huang, Y., Truong, S. C. H., Siems, S. T., and Manton, M. J.: A climatology of open and closed mesoscale cellular convection over the Southern Ocean derived from Himawari-8 observations, Atmos. Chem. Phys., 22, 2135-2152, https://doi.org/10.5194/acp-22-2135-2022, 2022.

Lensky, I. M. and Rosenfeld, D.: Clouds-Aerosols-Precipitation Satellite Analysis Tool (CAPSAT), Atmos. Chem. Phys., 8, 6739-6753, https://doi.org/10.5194/acp-8-6739-2008, 2008.

Li, X., Qiu, B., Cao, G., Wu, C., and Zhang, L.: A Novel Method for Ground-Based Cloud Image Classification Using Transformer, Remote Sensing, 14, 3978, https://doi.org/10.3390/rs14163978, 2022.

Liu, J., Zhu, Y., Wang, M., and Rosenfeld, D.: Cloud Susceptibility to Aerosols: Comparing Cloud-Appearance vs. Cloud-Controlling Factors Regimes, EGU General Assembly 2024, Vienna, Austria, 14–19 Apr 2024, EGU24-4059, https://doi.org/10.5194/egusphere-egu24-4059, 2024.

Luo, H., Quaas, J., and Han, Y.: Diurnally asymmetric cloud cover trends amplify greenhouse warming, Science Advances, 10, eado5179, https://doi.org/10.1126/sciadv.ado5179, 2024.

McCoy, I. L., Wood, R., and Fletcher, J. K.: Identifying Meteorological Controls on Open and Closed Mesoscale Cellular Convection Associated with Marine Cold Air Outbreaks, Journal of Geophysical Research: Atmospheres, 122, 11,678-611,702, https://doi.org/10.1002/2017JD027031, 2017.

McCoy, I. L., McCoy, D. T., Wood, R., Zuidema, P., and Bender, F. A.-M.: The Role of Mesoscale Cloud Morphology in the Shortwave Cloud Feedback, Geophysical Research Letters, 50, e2022GL101042, https://doi.org/10.1029/2022GL101042, 2023.

Mohrmann, J., Wood, R., Yuan, T., Song, H., Eastman, R., and Oreopoulos, L.: Identifying meteorological influences on marine low-cloud mesoscale morphology using satellite classifications, Atmos. Chem. Phys., 21, 9629-9642, https://doi.org/10.5194/acp-21-9629-2021, 2021.

Muhlbauer, A., McCoy, I. L., and Wood, R.: Climatology of stratocumulus cloud morphologies: microphysical properties and radiative effects, Atmos. Chem. Phys., 14, 6695-6716, https://doi.org/10.5194/acp-14-6695-2014, 2014.

Petters, M. D., Snider, J. R., Stevens, B., Vali, G., Faloona, I., and Russell, L. M.: Accumulation mode aerosol, pockets of open cells, and particle nucleation in the remote subtropical Pacific marine boundary layer, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/10.1029/2004JD005694, 2006.

Platnick, S., Meyer, K. G., King, M. D., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Holz, R. E., Yang, P., Ridgway, W. L., and Riedi, J.: The MODIS Cloud Optical and Microphysical Products: Collection 6 Updates and Examples From Terra and Aqua, IEEE Transactions on Geoscience and Remote Sensing, 55, 502-525, https://doi.org/10.1109/TGRS.2016.2610522, 2017.

Rosenfeld, D., Kaufman, Y. J., and Koren, I.: Switching cloud cover and dynamical regimes from open to closed Benard cells in response to the suppression of precipitation by aerosols, Atmos. Chem. Phys., 6, 2503-2511, https://doi.org/10.5194/acp-6-2503-2006, 2006.

Savic-Jovcic, V. and Stevens, B.: The Structure and Mesoscale Organization of Precipitating Stratocumulus, Journal of the Atmospheric Sciences, 65, 1587-1605, https://doi.org/10.1175/2007JAS2456.1, 2008.

Schulz, H., Eastman, R., and Stevens, B.: Characterization and Evolution of Organized Shallow Convection in the Downstream North Atlantic Trades, Journal of Geophysical Research: Atmospheres, 126, e2021JD034575, https://doi.org/10.1029/2021JD034575, 2021.

Segal Rozenhaimer, M., Nukrai, D., Che, H., Wood, R., and Zhang, Z.: Cloud Mesoscale Cellular Classification and Diurnal Cycle Using a Convolutional Neural Network (CNN), Remote Sensing, 15, 1607, https://doi.org/10.3390/rs15061607, 2023.

Stevens, B., Vali, G., Comstock, K., Wood, R., van Zanten, M. C., Austin, P. H., Bretherton, C. S., and Lenschow, D. H.: POCKETS OF OPEN CELLS AND DRIZZLE IN MARINE STRATOCUMULUS, Bulletin of the American Meteorological Society, 86, 51-58, https://doi.org/10.1175/BAMS-86-1-51, 2005.

Tornow, F., Ackerman, A. S., and Fridlind, A. M.: Preconditioning of overcast-to-broken cloud transitions by riming in marine cold air outbreaks, Atmos. Chem. Phys., 21, 12049-12067, https://doi.org/10.5194/acp-21-12049-2021, 2021.

Vose, R. S., Easterling, D. R., and Gleason, B.: Maximum and minimum temperature trends for the globe: An update through 2004, Geophysical Research Letters, 32, https://doi.org/10.1029/2005GL024379, 2005.

Wang, H. and Feingold, G.: Modeling Mesoscale Cellular Structures and Drizzle in Marine Stratocumulus. Part I: Impact of Drizzle on the Formation and Evolution of Open Cells, Journal of the Atmospheric Sciences, 66, 3237-3256, https://doi.org/10.1175/2009JAS3022.1, 2009.

Wang, Q., Zhou, C., Zhuge, X., Liu, C., Weng, F., and Wang, M.: Retrieval of cloud properties from thermal infrared radiometry using convolutional neural network, Remote Sensing of Environment, 278, 113079, https://doi.org/10.1016/j.rse.2022.113079, 2022.

Watson-Parris, D., Sutherland, S. A., Christensen, M. W., Eastman, R., and Stier, P.: A Large-Scale Analysis of Pockets of Open Cells and Their Radiative Impact, Geophysical Research Letters, 48, e2020GL092213, https://doi.org/10.1029/2020GL092213, 2021.

Wood, R.: Stratocumulus Clouds, Monthly Weather Review, 140, 2373-2423, https://doi.org/10.1175/MWR-D-11-00121.1, 2012.

Wood, R. and Hartmann, D. L.: Spatial Variability of Liquid Water Path in Marine Low Cloud: The Importance of Mesoscale Cellular Convection, Journal of Climate, 19, 1748-1764, https://doi.org/10.1175/JCLI3702.1, 2006.

Wu, Y., Liu, J., Zhu, Y., Zhang, Y., Cao, Y., Huang, K.-E., Zheng, B., Wang, Y., Wang, Q., Zhou, C., Liang, Y., Wang, M., and Rosenfeld, D.: Global Classification Dataset of Daytime and Nighttime Marine Low-cloud Mesoscale Morphology [dataset], https://doi.org/10.5281/zenodo.13801408, 2024.

Wyant, M. C., Bretherton, C. S., Rand, H. A., and Stevens, D. E.: Numerical Simulations and a Conceptual Model of the Stratocumulus to Trade Cumulus Transition, Journal of the Atmospheric Sciences, 54, 168-192, https://doi.org/10.1175/1520-0469(1997)054<0168:NSAACM>2.0.CO;2, 1997.

Xue, H., Feingold, G., and Stevens, B.: Aerosol Effects on Clouds, Precipitation, and the Organization of Shallow Cumulus Convection, Journal of the Atmospheric Sciences, 65, 392-406, https://doi.org/10.1175/2007JAS2428.1, 2008.

Yamaguchi, T. and Feingold, G.: On the relationship between open cellular convective cloud patterns and the spatial distribution of precipitation, Atmos. Chem. Phys., 15, 1237-1251, https://doi.org/10.5194/acp-15-1237-2015, 2015.

Yuan, T., Song, H., Wood, R., Mohrmann, J., Meyer, K., Oreopoulos, L., and Platnick, S.: Applying deep learning to NASA MODIS data to create a community record of marine low-cloud mesoscale morphology, Atmos. Meas. Tech., 13, 6989-6997, https://doi.org/10.5194/amt-13-6989-2020, 2020.

Zhang, J., Liu, P., Zhang, F., and Song, Q.: CloudNet: Ground-Based Cloud Classification With Deep Convolutional Neural Network, Geophysical Research Letters, 45, 8665-8672, https://doi.org/10.1029/2018GL077787, 2018.

supplementary materials

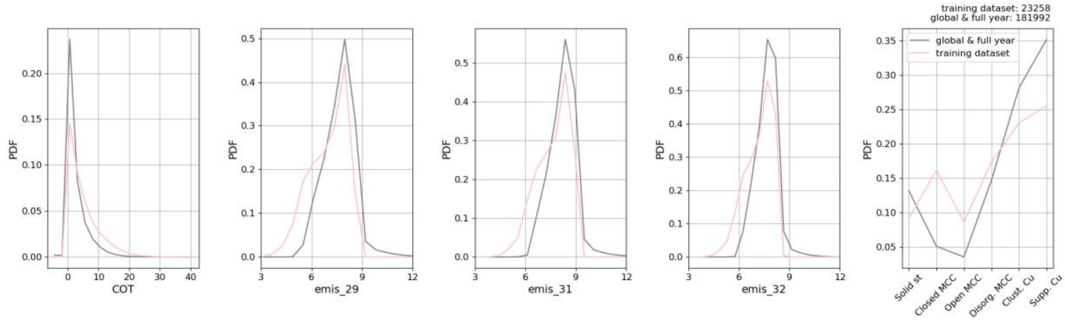
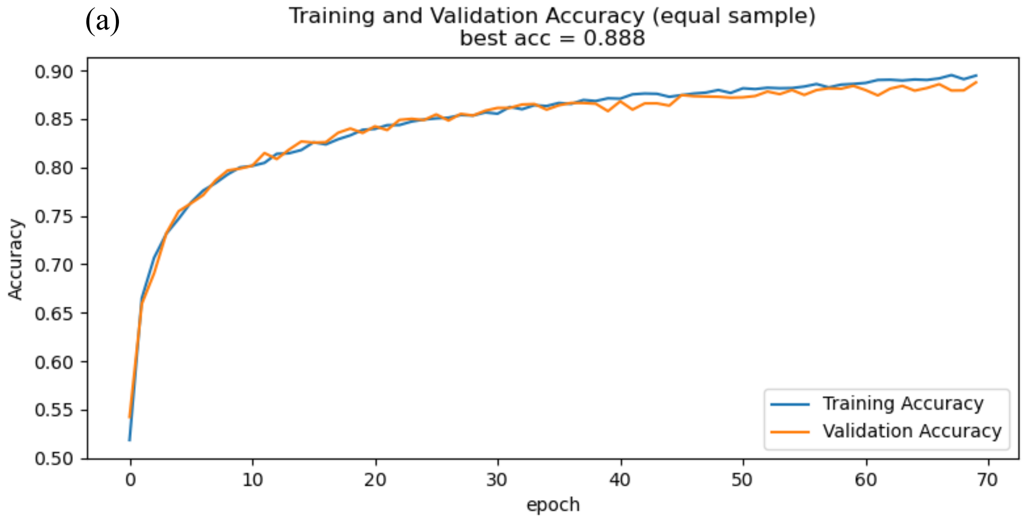


Figure S1: The comparison of probability density functions (PDFs) between our training dataset and a global full-year dataset. (a) PDFs of cloud optical thickness (COT); PDFs of radiance data from infrared channels: (b) 29, (c) 31, (d) 32; (e) PDFs of cloud morphology.

(a)



(b)

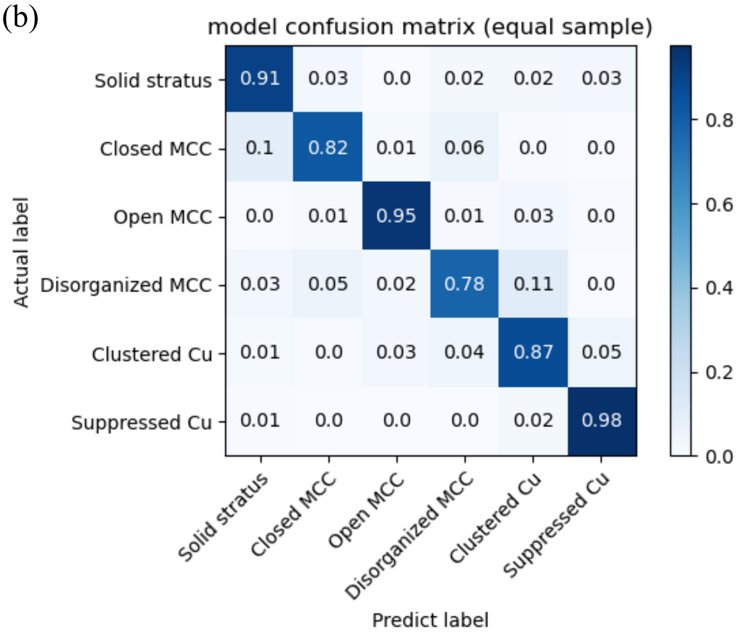
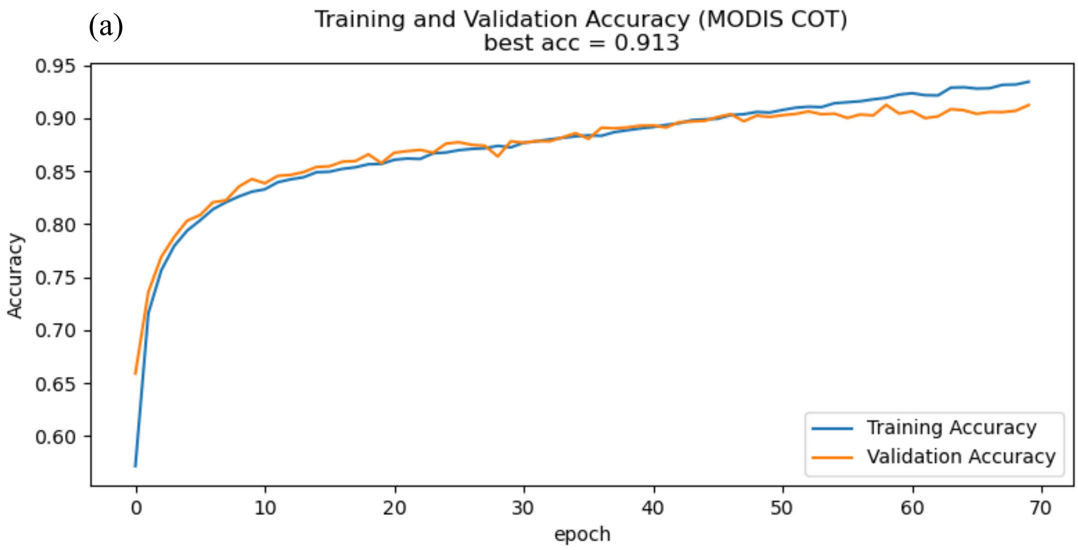


Figure S2: Results of the model trained on a balanced sample dataset (2,000 samples per category). (a) the model's accuracy on the training and validation datasets. (b) the confusion matrix of the model.

(a)



(b)

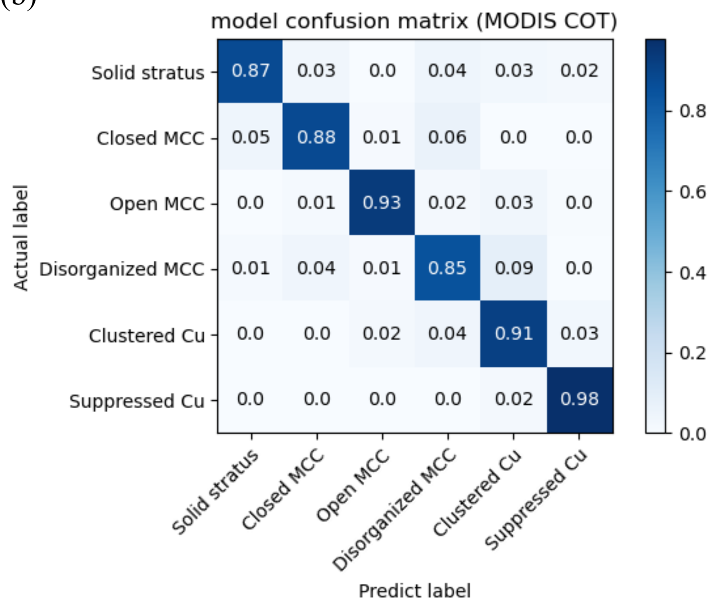


Figure S3: Model training results based on MODIS COT. (a) the model's accuracy on the training and validation datasets. (b) the confusion matrix of the model.

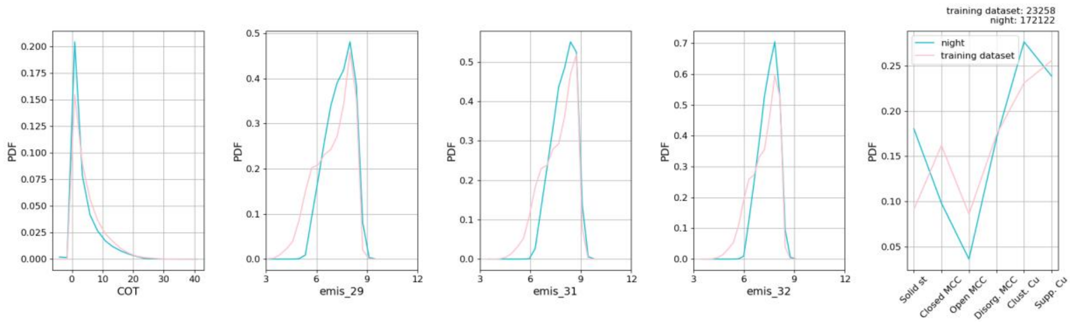


Figure S4: The comparison of probability density functions (PDFs) between our training dataset and a nighttime dataset. (a) PDFs of cloud optical thickness (COT); PDFs of radiance data from infrared channels: (b) 29, (c) 31, (d) 32; (e) PDFs of cloud morphology.

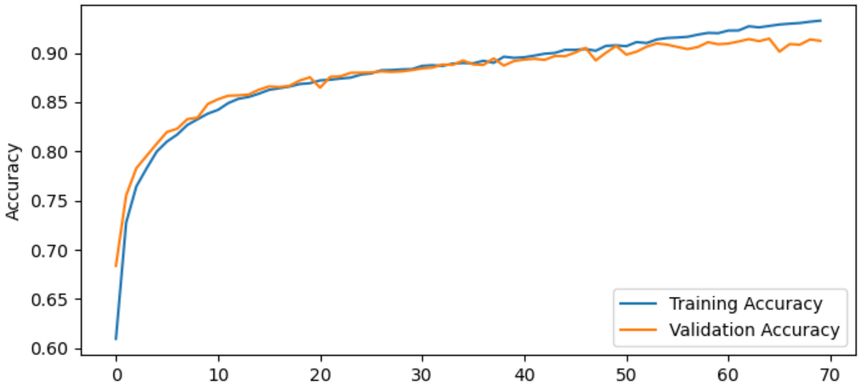


Figure S5: The training and validation accuracy curves. Validation accuracy reaches its maximum around epoch 65.

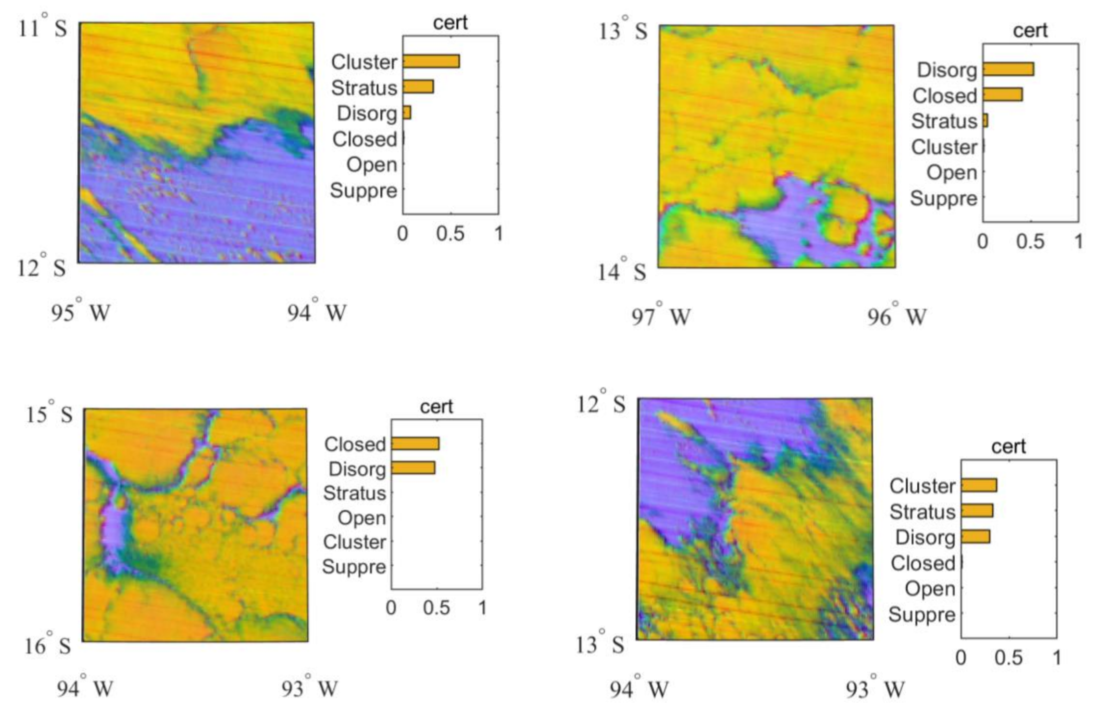


Figure S6: Examples of some mixed and transition scenes. Here, 'cert' represents the confidence level of each category. The sum of the confidence levels for the six categories equals to 1.

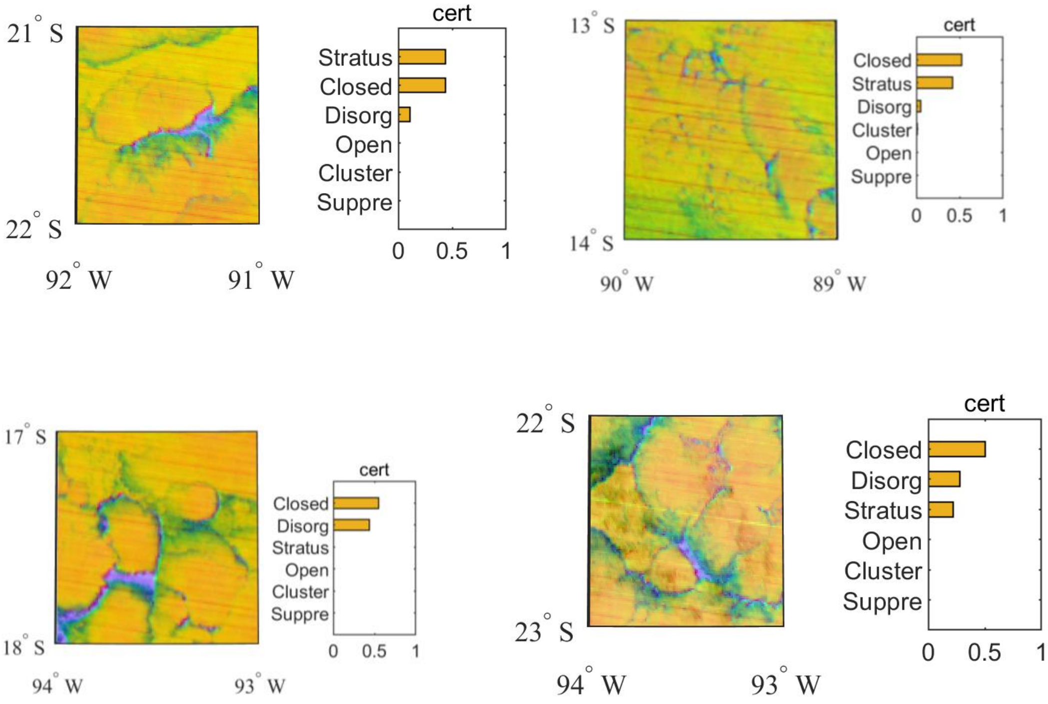
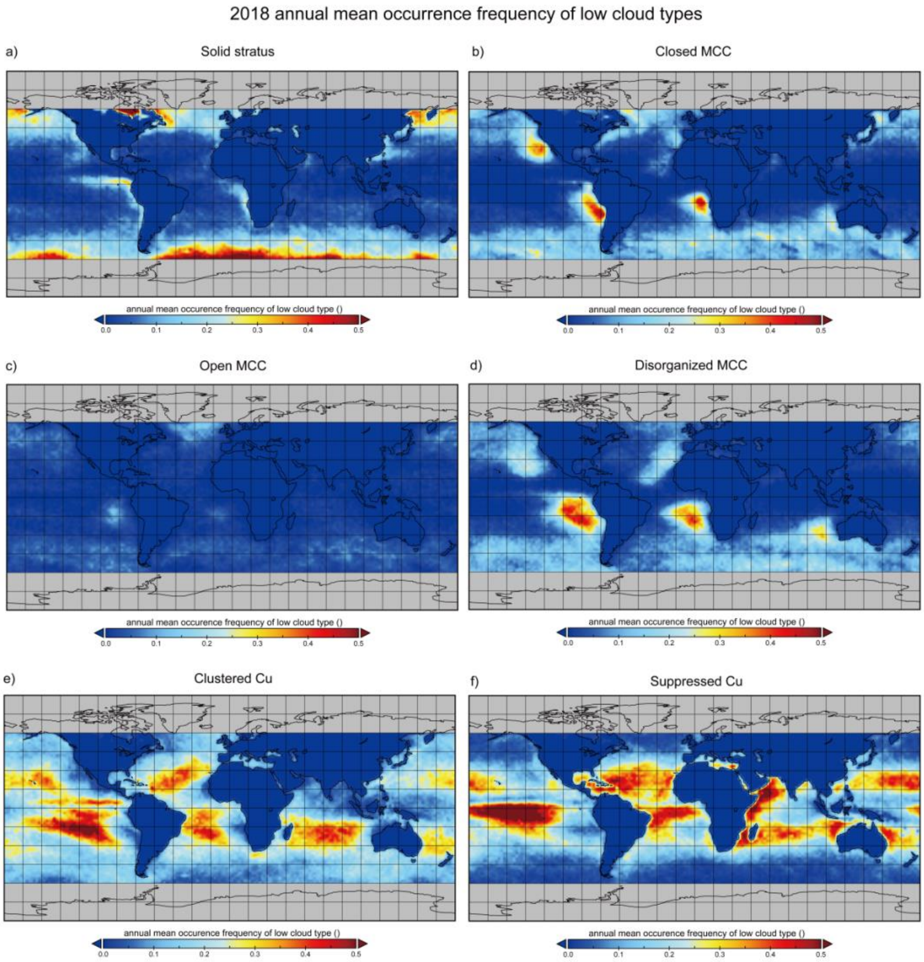


Figure S7: Same as Fig. S2. Examples of some misclassified scenes because of model's limited capacity to distinguish between stratiform structures and convective cells.



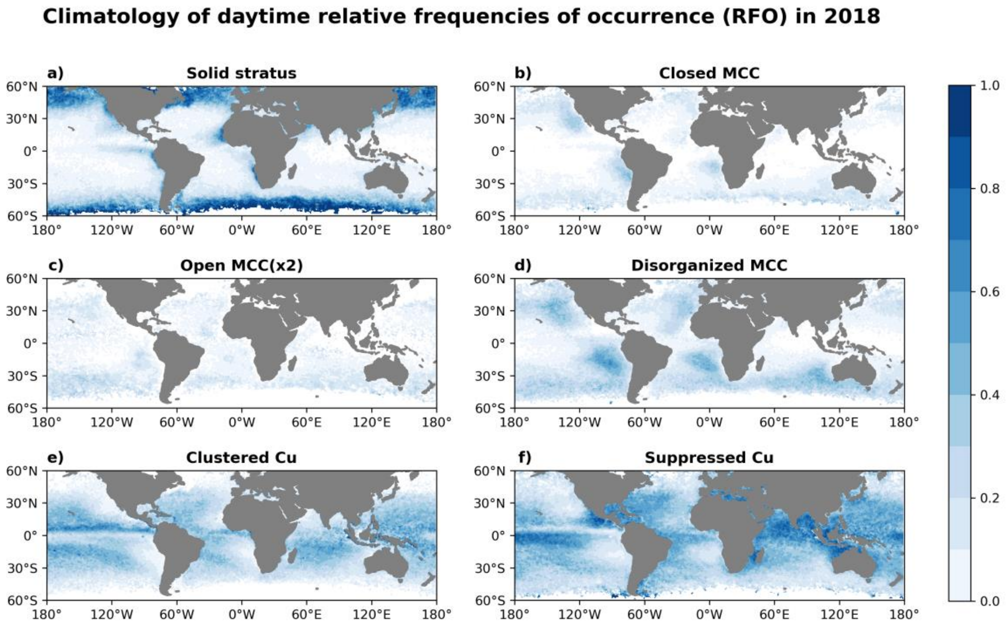


Figure S8: 2018 annual mean occurrence frequency of low cloud types from Yuan et al. (2020) and from our model.