AGU ADVANCING EARTH AND SPACE SCIENCES

Geophysical Research Letters

RESEARCH LETTER
10.1029/2024GL111513

Clustering the Diurnal Cycle of Precipitation Using Global Satellite Data

Penghan Chen[1], Aifang Chen[2], Shuiqing Yin[3], Yuxin Li[4], and Junguo Liu[5,6]

[1]School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China, [2]School of Environment and Civil Engineering, Dongguan University of Technology, Dongguan, China, [3]State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science, Beijing Normal University, Beijing, China, [4]Department of Geography, The University of Hong Kong, Hong Kong, China, [5]Yellow River Research Institute, North China University of Water Resources and Electric Power, Zhengzhou, China, [6]Henan Provincial Key Laboratory of Hydrosphere and Watershed Water Security, North China University of Water Resources and Electric Power, Zhengzhou, China

Penghan Chen and Aifang Chen contributed equally to this work.

Key Points:

- Hierarchical clustering and high-resolution data reveal intricate global patterns of the diurnal cycle of precipitation (DCP)
- Diurnal cycles of precipitation amount and frequency are stronger than that of intensity, and exhibit significant seasonal differences
- We provide detailed, user-friendly DCP maps that include amount, frequency, intensity, and seasonality

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

J. Liu,
junguo.liu@gmail.com

Citation:

Chen, P., Chen, A., Yin, S., Li, Y., & Liu, J. (2024). Clustering the diurnal cycle of precipitation using global satellite data. *Geophysical Research Letters*, *51*, e2024GL111513. https://doi.org/10.1029/2024GL111513

Received 20 JUL 2024
Accepted 21 NOV 2024

Abstract The diurnal cycle of precipitation (DCP) is a fundamental component of the precipitation cycles across various temporal scales. However, the absence of a comprehensive and intuitive clustering description of the DCP, including precipitation amount, frequency, intensity, and seasonality on the global scale, has impeded understanding of climate impacts on society. This study investigates the diurnal cycle and its seasonal differences in global precipitation using hierarchical clustering and high-resolution satellite precipitation data. Results indicate that regions with coastal and topographic complexity exhibit intricate clustering patterns globally. Significant inter-cluster differences suggest that the precipitation amount and frequency have stronger diurnal cycles than that of the precipitation intensity. Seasonal cluster transitions indicate that the global DCP exhibits significant seasonal differences. This study provides multiple sets of information-rich and user-friendly maps of the DCP, enhancing research on sub-daily precipitation mechanisms and their climatic impacts.

Plain Language Summary The diurnal precipitation cycle significantly impacts human society and ecosystems, with spatial variations across the globe. Some regions experience concentrated precipitation at specific times of the day, while others display a more even distribution. However, there is a need for a clearer understanding of these spatial variations. This study employs a hierarchical approach to categorize the global diurnal precipitation cycle using high-resolution satellite data. Our findings highlight that the diurnal cycle of precipitation amount and frequency is most pronounced in summer. These maps provide richer details than conventional methods and are easier to interpret when comparing seasons. Similar to established climate classifications, these global diurnal precipitation cycle maps offer valuable insights for advancing research into the diurnal cycle of precipitation.

1. Introduction

As a fundamental mode of atmospheric variability, the diurnal cycle of precipitation (DCP) is prevalent worldwide (Janowiak et al., 2005). The DCP varies spatially as it is influenced by the local, regional and large-scale dynamic and thermal processes (Yu et al., 2014). Various mechanisms influence the DCP, such as the underlying complex topography, land-sea breeze and anthropogenic influences (Cook & Vizy, 2024; Doan et al., 2021; Qiu et al., 2023; Wang et al., 2023; Wu et al., 2018; Yin et al., 2011). As the most fundamental component of multi-temporal scale precipitation cycles, the DCP profoundly impacts both nature and society. For example, the DCP can affect local evapotranspiration and gross primary productivity, thereby impacting local agriculture economic development (Kincer, 1916; Yang et al., 2023). Therefore, an accurate understanding of the global DCP would support diagnosing the impact of climate on society.

Previous research has revealed significant spatial disparities in the regional DCP. Land precipitation tends to peak in the afternoon, particularly in coastal regions such as the southeast coast of China, the Indochinese Peninsula, and maritime continents (Jiang et al., 2017; Lu et al., 2021; Takahashi et al., 2010). In regions characterized by complex terrain like basins and mountain ranges, the DCP exhibits diverse characteristics, as seen in the surrounding areas of the Amazon Basin, Qinghai-Tibet Plateau and Sichuan Basin (Giles et al., 2020; Lu et al., 2024; Zhang et al., 2019). Coastal regions with intricate topography like the Cordillera mountain range, are influenced by both coastal-land distinctions and terrain, resulting in more intricate DCP attributes (Ruiz-Hernández

et al., 2021). In oceans, precipitation tends to peak in the morning and demonstrates a spatial-temporal pattern of offshore propagation (Aoki & Shige, 2024; Coppin & Bellon, 2019; Fang & Du, 2022; Li & Carbone, 2015). Mid-low latitude regions experience higher solar radiation than high-latitude regions, leading to more pronounced spatial heterogeneity in the DCP (Jeong et al., 2011).

Many metrics and methods have been developed to describe the spatial heterogeneity of the DCP. As the DCP is essentially a description of distribution, using hourly discrete data directly for this purpose may not effectively highlight the key aspects of interest. One common approach involves using the first-order amplitude and phase obtained from frequency domain transformations to characterize DCP (Dai, 2024; Dai et al., 2007; Hayden & Liu, 2021; Jeong et al., 2011; Mao & Wu, 2012). However, this method simplifies the DCP to a sinusoidal model and overlooks high-frequency information. Another method, known as empirical orthogonal functions (EOF), has also been utilized to capture spatial variations in DCP (Kikuchi & Wang, 2008; Lu et al., 2021; Wilks, 2011). However, this method has limitations, as each spatial location must exhibit complete opposites from other locations, potentially limiting the detailed information provided (Covey et al., 2016). Moreover, interpreting the raw DCP often requires a process of reconstruction. Additionally, parameters such as peak time, kurtosis, skewness, and indicators based on precipitation differences over specific time periods are also used to describe DCP (Deng et al., 2022; Jiang et al., 2017; Johnson, 2011; Lu et al., 2024; Yu et al., 2007).

Apart from the aforementioned description methods, clustering analysis has been used to extract the main patterns of regional DCP, for example, Fuzzy C-Means and K-Means methods (Chen et al., 2009; Liu et al., 2021; Mu et al., 2021; Wu et al., 2018; Yaqub et al., 2011). Although these methods can provide a more intuitive display of regional DCP while retaining richer information, their clustering process is often challenging to interpret. Besides, these methods have poor scalability because finer clustering cannot obtain any valuable information from coarser clustering results. Hierarchical clustering differs from traditional clustering as it does not require a predefined number of clusters; instead, it progressively merges or splits clusters based on the similarity between data points. This clustering method naturally reflects the hierarchical structure of the data, facilitating an intuitive understanding of relationships between data points. The clustering tree generated by hierarchical clustering clearly displays the hierarchical relationships between data, making the clustering results more intuitive and interpretable (Murtagh & Contreras, 2012).

Although clustering is an excellent method for describing DCP, it has not yet been applied at the global scale, let alone hierarchical clustering. It is worth noting that clustering is also well-suited for describing seasonal differences in DCP. This is because after clustering annual DCP, we obtain a well-trained model (classifier) that can cluster summer and winter DCP. By comparing the differences in clustering results, we can intuitively understand the seasonal differences in DCP. In addition, the DCP involves three essential aspects, precipitation amount, frequency, and intensity, which are commonly used to provide a more comprehensive description of precipitation (Wu et al., 2018). Most existing studies primarily focus on the amount; however, they lack a complete description of DCP. Besides, the seasonal differences in the DCP also lack intuitive global descriptions. Summer precipitation receives greater attention than winter precipitation in studies of the DCP (Chen et al., 2018; Giles et al., 2020; Mu et al., 2021; Song & Wei, 2021) and investigations into seasonal differences in the DCP often rely on the one-sided metrics (Dai, 2024; Deng et al., 2022; Hirose & Nakamura, 2005; Olaguera et al., 2024; Yin et al., 2009). Hence, a comprehensive and intuitive clustering description of the global characteristics of DCP, including amount, frequency, intensity and seasonality, is necessary for a better understanding of precipitation mechanisms.

The study of DCP requires the use of sub-daily data, specifically hourly data. While instrumental observation data remains the gold standard for accuracy, data sets like Integrated Surface Data set (ISD), Hadley Center Integrated Surface Data set (HadISD) and Global Sub-Daily Rainfall (GSDR) (Dunn et al., 2012; Lewis et al., 2019; Smith et al., 2011), which provide hourly instrument observation precipitation data, are predominantly available in the United States, Europe, and Japan. However, a notable disparity exists in the time spans covered by different stations, posing a challenge for global-scale studies. The Integrated Multi-satellitE Retrievals for GPM (IMERG) and the European Center for Medium-Range Weather Forecasts Reanalysis v5 (ERA5) (Hersbach et al., 2020; Huffman et al., 2015) have been proven to have good capabilities in representing the DCP (Qin et al., 2021; Tan et al., 2019; Zhang et al., 2022) and have garnered widespread use in research related to hourly precipitation, including regional DCP and tropical cyclone precipitation (Fang & Du, 2022; Maulana et al., 2023; Tu et al., 2022; Wang et al., 2023).

Therefore, this study uses IMERG as the primary data source, combined with hierarchical clustering, to analyze the diurnal cycle of precipitation amount, frequency, and intensity globally. Some results from the ERA5 data, used for auxiliary validation, can be found in Supporting Information S1. Additionally, clustering is developed to demonstrate the seasonal differences of the DCP. Finally, leveraging the interpretability of the hierarchical approach, we delved into the decision-making process underlying the global DCP clustering. Similar to climate classification zones, our results of the information-rich and intuitive maps of the DCP can provide values for sub-daily precipitation-related research.

2. Materials and Methods

2.1. Data

The IMERG precipitation data utilized in this study spans from 2001 to 2020, featuring a spatial resolution of 0.1° and a temporal resolution of 0.5 hr, it has been aggregated to a 1-hr temporal resolution for this analysis. The ERA5 precipitation data utilized in this study covers the period from 1979 to 2020, with a spatial resolution of 0.25° and a temporal resolution of 1 hr. To facilitate analysis, the data is converted from Coordinated Universal Time (UTC) to Local Solar Time (LST). It's worth noting that due to reduced solar radiation in high-latitude regions, the DCP is less pronounced, and satellites encounter challenges in accurately estimating precipitation in these areas. As a result, this study primarily focuses on the global mid-low latitudes (60°N-60°S). The elevation and land-sea boundary data are sourced from the Copernicus digital elevation model (Guth & Geoffroy, 2021).

2.2. Precipitation Amount, Frequency and Intensity

The diurnal cycle of precipitation (DCP) refers to the variations or cycles of precipitation within a natural day. Amount, frequency, and intensity are three crucial aspects for describing precipitation behavior (Dai et al., 2007). In this study, precipitation amount (PA), frequency (PF), and intensity (PI) used to describe the DCP are defined as follows (Wu et al., 2018). For PA, we first calculate the hourly precipitation for each year, then average the hourly precipitation over multiple years, and finally normalize it within a day to reflect diurnal variations. For PF, we calculate the number of hours in local time where precipitation exceeding 0.1 mm/hr occurs and then normalize it within a day. For PI, we normalize the quotient of PA and PF. The relevant formulas are as follows:

$$A_t = \mathrm{Norm}_t\left(\sum_{i=Y_s}^{Y_e}\sum_{j=0}^{N_i} P_{i,j,t}\right)$$

$$F_t = \mathrm{Norm}_t\left(\sum_{i=Y_s}^{Y_e}\sum_{j=0}^{N_i} \Theta(P_{i,j,t} - 0.1)\right)$$

$$I_t = \mathrm{Norm}_t\left(\frac{A_t}{F_t}\right)$$

In the equations, A_t, F_t, and I_t represent the PA, PF and PI at local time t, respectively. Norm_t is normalized over 24 hr of local time (divided by the average), Y_s and Y_e are the start and end years, and N_i is the total number of days in the i-th year. i, j, and t denote the year, day, and local hour, respectively. $P_{i,j,t}$ represents the precipitation amount at the t-th local hour of the j-th day of the i-th year. $\Theta(x)$ is the step function, taking the value 1 when x is greater than or equal to 0, and 0 otherwise.

2.3. Hierarchical Clustering of Global Diurnal Precipitation Cycle

In this study, we employed the Bisecting K-Means algorithm, a hierarchical variant of the K-Means algorithm. Our validation showed that the Bisecting K-Means algorithm performed comparably to K-Means, maintaining a similar sum of squared errors (SSE). We opted for 8 clusters, as further increases yielded only marginal improvements in clustering performance, with a diminishing decrease in SSE (Figure S1 in Supporting Information S1). This choice strikes an optimal balance between visualization clarity and clustering performance.

When partitioning samples, the Bisecting K-Means algorithm selects the cluster with the largest SSE for further division. After obtaining the final clustering results, we obtain a clustering binary tree that describes the clustering

process (Di & Gou, 2018). When clustering global DCP, we can clearly understand the results of each binary split of the clustering model, thereby gaining a clearer insight into the spatial differences in global DCP. Additionally, after performing hierarchical clustering, we obtain a trained clustering model that can be applied to cluster other samples. The Bisecting K-Means clustering is implemented using Python's Scikit-Learn library (Pedregosa et al., 2011).

In this study, we conduct separate clustering analyses of the global annual PA, PF, and PI derived from IMERG data sets. Subsequently, we use the clustering models trained on the annual DCP to the summer (AMJJAS in the Northern Hemisphere and ONDJFM in the Southern Hemisphere) and winter (ONDJFM in the Northern Hemisphere and AMJJAS in the Southern Hemisphere) to determine seasonal differences. Finally, leveraging the interpretability of the hierarchical method, we discuss the clustering process of the global diurnal cycle of precipitation amount (DCPA) and its underlying basis.

3. Results and Discussion

3.1. Global Diurnal Cycle of Precipitation Amount, Frequency and Intensity

From the clustering results of the global diurnal cycle of precipitation amount (DCPA, Figure 1a), we can gain a more intuitive understanding of information such as the peak hour, peak values, and uniformity of global DCPA. The findings underscore that regions characterized by coastal and topographical complexity exhibit intricate patterns in global DCPA. Specifically, significant clustering transitions can be observed from coastlines to adjacent areas and from high-altitude regions to their surroundings. These areas encompass Central America, Central Africa, the Qinghai-Tibet Plateau and its environs, maritime continents, and their coastal waters. In the western mountains of the United States, the highest elevations correspond to cluster eight on the clustering map. As the elevation decreases eastward, the clusters transition progressively to clusters 7 and 6. A similar phenomenon is observed on the Tibetan Plateau. In certain mid- and low-latitude marine regions, clear transitions between clusters are evident on either side of the coastline as the distance varies. The comparison of clustering results and specific terrestrial and marine topographic features (land characterized by elevation, sea by distance from the coastline) in regions with complex DCP can be referenced in the supplementary information (Figure S2 in Supporting Information S1). This complexity arises due to the abundant solar radiation and significant thermal capacity gradients in these regions (Ruiz-Hernández et al., 2021; Wang et al., 2023).

Taking into account the specific diurnal pattern for each cluster, in low-latitude coastal land, precipitation mainly occurs during the afternoon and evening, with a notable cluster transition in DCPA from coastal to inland regions (from cluster 8 to 7 to 6 in Figure 1a). This transition is marked by a gradual decline in peak values and a shift in the peak hour toward later times. High-altitude areas exhibit an extreme afternoon precipitation peak similar to that of coastal land. As the transition to lower altitudes occurs, precipitation shows a later peak time and greater uniformity. Marine areas exhibit a more uniform DCPA pattern, characterized by distinct nighttime peaks along coastlines and transitioning toward offshore zones (from cluster 4 to 3 to 1 in Figure 1a). The variation in the proportion of different clusters with distance from the coastline can be found in the supplementary information (Figure S3 in Supporting Information S1). These observations on cluster transition align with previous research, illustrating the behavior of precipitation on both sides of the coastline, with onshore propagation on the land side and offshore propagation on the ocean side (Coppin & Bellon, 2019; Du & Rotunno, 2018; Fang & Du, 2022; Li & Carbone, 2015). Clustering maps provide a clearer understanding of the global regions exhibiting this propagation pattern.

While most clusters exhibit a single peak, multiple peaks are also observed. In cluster 4 of the DCPA, a primary peak at 1 a.m. and a secondary peak at 9 a.m. are evident, primarily in nearshore waters and certain basins. Additionally, our statistics indicate that multiple peaks are not predominant, accounting for only 6.7% of the total samples. Therefore, the global-scale clustering cannot directly capture all the details. The spatial distribution of different types of multiple peaks (Figure S4 in Supporting Information S1) indicates that the highest number of bimodal samples is found along the west coasts of Africa and South America, where precipitation is extremely limited due to the influence of the Benguela Current and the Humboldt Current (Garreaud et al., 2010; Lima et al., 2019). This phenomenon may be related to local meteorological conditions, ocean-atmosphere interactions, limitations of satellite precipitation or other factors that are not yet fully understood. It is also worth noting that the

10.1029/2024GL111513

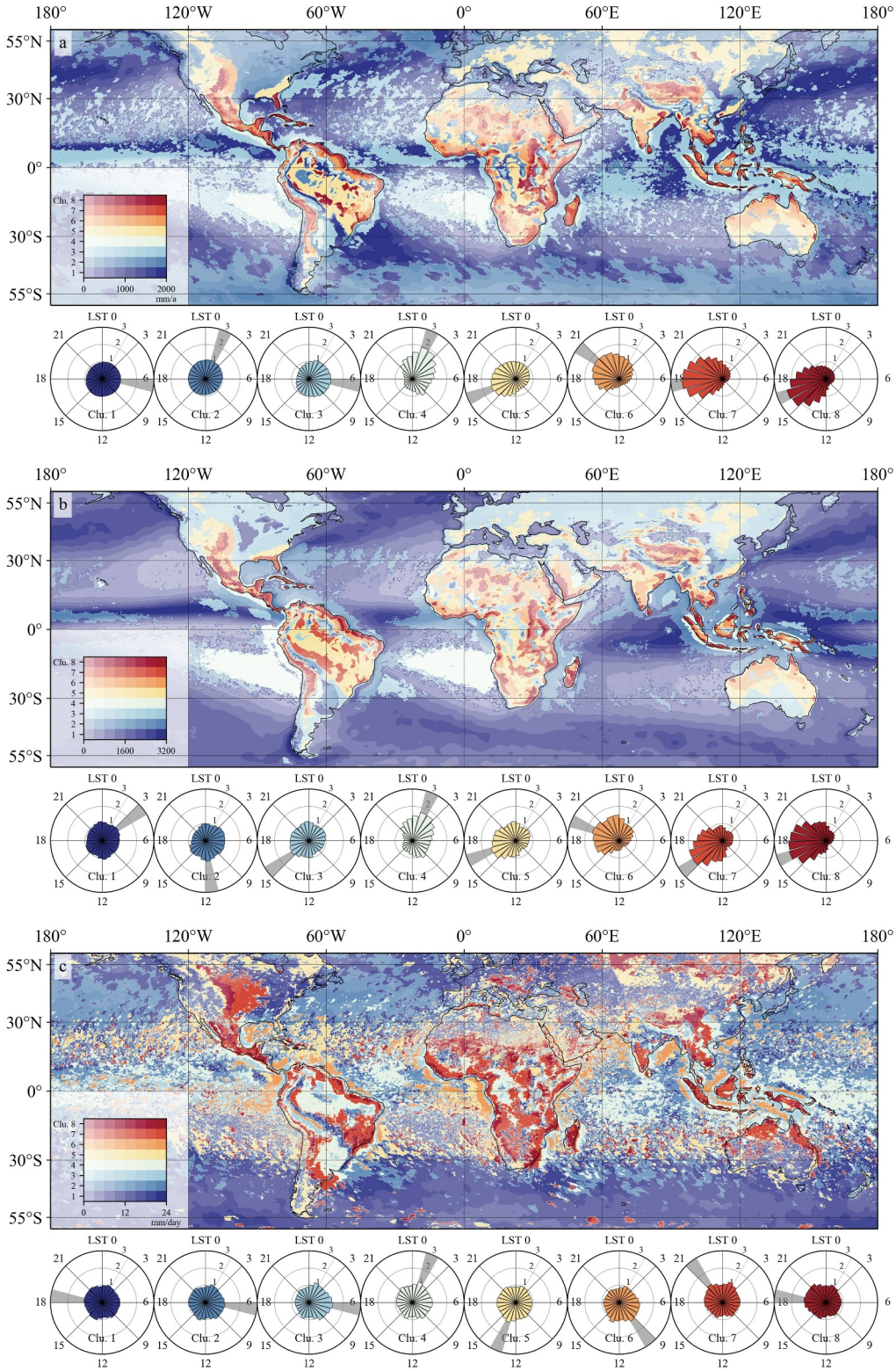

Figure 1.

rise before the peak of DCPA is more intense compared to the decline after the peak especially in clusters 5, 6 and 7, which is a characteristic that the sinusoidal model associated with the amplitude-phase method cannot capture. This phenomenon aligns with previously observed asymmetries in precipitation processes, likely due to atmospheric instability and the duration of precipitation events (Yu et al., 2013).

Similarly, the diurnal cycle of precipitation frequency (DCPF, Figure 1b) mirrors DCPA. However, compared to DCPA and DCPF, the diurnal cycle of precipitation intensity (DCPI, Figure 1c) shows minimal discrepancies between different clusters and presents a relatively uniform diurnal distribution. This indicates that DCPI maintains uniformity across most global regions, exhibiting low spatial variability. Like satellite precipitation, reanalysis precipitation demonstrates broadly similar characteristics in DCPA and DCPF (Figure S5 in Supporting Information S1). Overall, the findings indicate intricate clustering patterns globally, and the precipitation amount and frequency have stronger diurnal cycles than that of the precipitation intensity.

3.2. Seasonal Difference of Diurnal Precipitation Cycle

To better capture seasonal differences, the clustering in this section employs the model trained on clustering the annual DCP, ensuring consistency with the clusters depicted in Figure 1a. Upon validation, the clustering model trained on annual DCP exhibits acceptable performance on seasonal DCP samples across different latitude intervals (Figures S6, S7, S8 in Supporting Information S1 for annual, summer and winter, respectively).

Figure 2 illustrates the global spatial distribution of DCPA during the summer (a) and winter (b), calculated using satellite data. Compared to index-based descriptions, the DCP clustering maps we provide for different seasons offer a more intuitive way to understand the seasonal characteristics and differences in DCP at any given location, while also allowing for a comprehensive examination of these seasonal differences on a global scale.

The results indicate that global DCPA exhibits significant seasonal differences, with 68.0% of land areas and 51.7% of sea areas undergoing seasonal cluster transitions (area-weighted average). Considering the varying degrees of transition across different clusters, for instance, the significant shift from cluster 1 to 8 compared to the modest transition from cluster 1 to 2 in DCPA clustering, the supplementary information (Figure S9 in Supporting Information S1) provides the regions that experienced transitions and the Euclidean distances between the summer and winter clusters, offering a more comprehensive perspective on the seasonal differences in DCPA. In certain areas, DCPA shows notable seasonal differences. For example, the southeastern coast of China, the Florida Peninsula, and the western mountains of the United States transition from clusters 5, 6, 7, and 8 (with relatively distinct peaks) in summer to clusters 1, 2, and 3 (relatively uniform) in winter. The northern part of the Eurasian continent transitions from predominantly cluster 5 to predominantly cluster 2. Overall, from summer to winter, the proportions of DCPA clusters with pronounced diurnal variation characteristics (concentrated precipitation) decrease, and this difference is more evident in regions above 15° latitude. This seasonal difference exhibited by the clustering supports previous research findings (Bedoya-Soto et al., 2019; Chen et al., 2009; Deng et al., 2022; Olaguera et al., 2024; Xiao et al., 2018; Zhao et al., 2022). This is because lower solar radiation in winter makes it difficult to drive local wind and diurnal water vapor transport (Barman et al., 2021; Shen et al., 2021). Additionally, at lower latitudes, the transition rate is relatively lower compared to other latitudes due to the indistinct seasonal boundaries in these regions. For an overview of seasonal clustering transitions across different latitude intervals, the Sankey diagram in the supplementary information (Figure S10 in Supporting Information S1) can be consulted. Global DCPF shows similar seasonal differences compared to DCPA (Figure S11 in Supporting Information S1). Given the minimal diurnal variation, seasonal differences in DCPI are not particularly meaningful to study. In addition, reanalysis precipitation demonstrates similar characteristics of global DCPA and DCPF in both summer and winter, including seasonal differences in land areas with latitudes higher than approximately 15° (Figures S12 and S13 in Supporting Information S1).

Figure 1. (a) Global diurnal cycle of precipitation amount. Different colors represent different clusters, while different levels of transparency represent different annual precipitation amount (sum of 24 hr). The rose diagrams below show the specific DCP patterns for each cluster. The radius represents the average of normalized hourly precipitation amount, while the angle corresponds to different local solar times. The shaded bar highlights the peak hour (same for later figures). (b) Global diurnal cycle of precipitation frequency. Different colors represent different clusters, while different levels of transparency represent different annual precipitation frequency. (c) Global diurnal cycle of precipitation intensity. Different colors represent different clusters, while different levels of transparency represent different daily precipitation intensity.

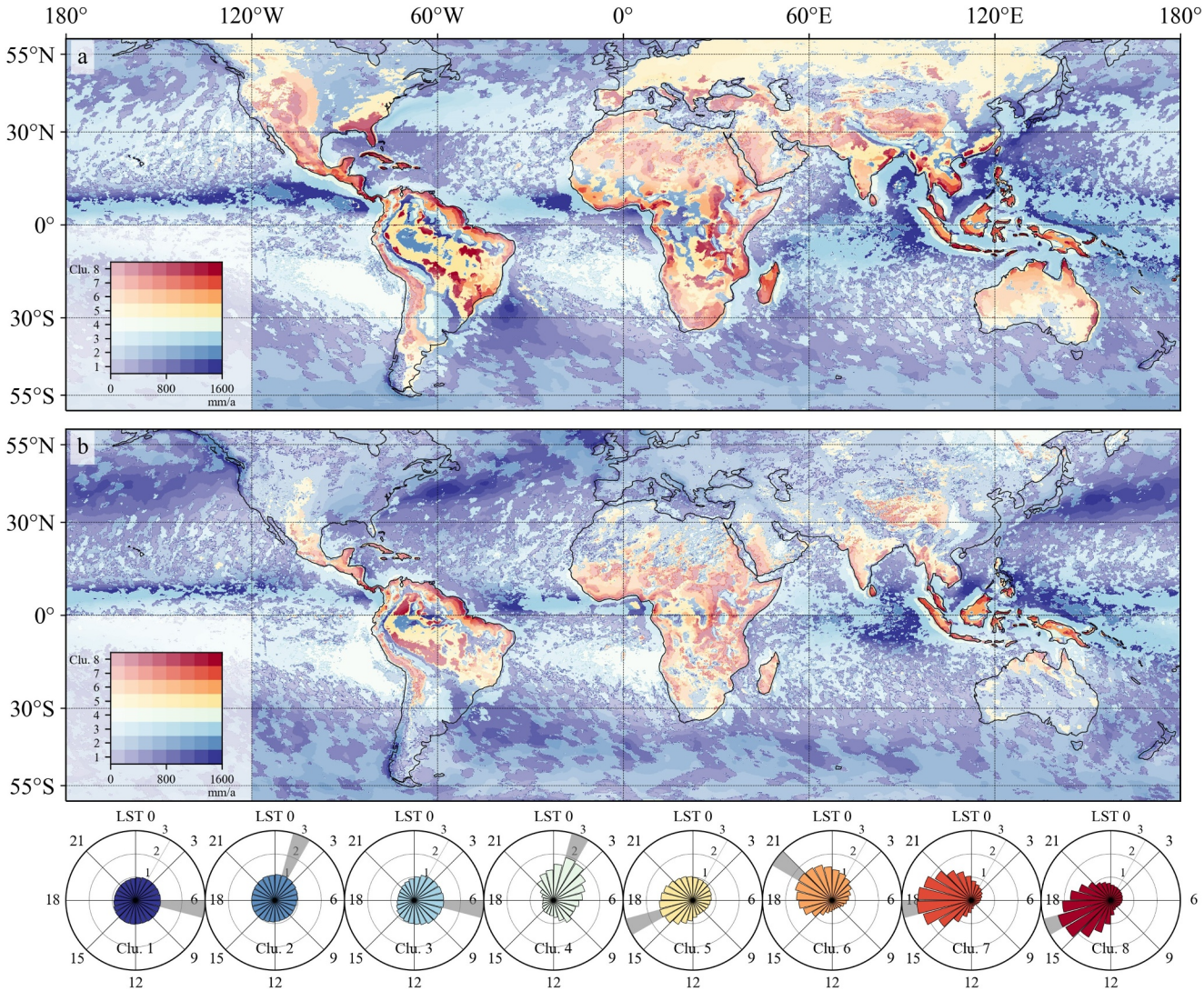

Figure 2. Global diurnal cycle of precipitation amount, (a) summer, (b) winter. Different colors represent different clusters, while different levels of transparency represent different annual precipitation amount.

3.3. Hierarchical Clustering Process of Diurnal Precipitation Cycle

Figure 3 illustrates the clustering process of global DCPA. When the number of clusters is set to 8, the hierarchical clustering method naturally progresses through different clustering levels from 1 to 8 because of its binary nature. As a result, the attribution in hierarchical clustering can be inherited across various levels. In contrast, non-hierarchical clustering methods require independent specification of cluster numbers from 1 to 8, with no relationship between different clustering levels (Figure S14 in Supporting Information S1). Due to sensitivity to initial conditions, the instability of non-hierarchical methods is also observed, evidenced by an anomalous red thread at the junction of two blue regions over the ocean, which does not correspond to the actual conditions (Figure S15 in Supporting Information S1).

The hierarchical clustering process helps us understand the principles of model partitioning. By combining similar external features within each cluster (such as terrain, land-sea relationships, etc.), we can explore the multiple factors that influence the DCPA. The logic here is that DCP determines clustering, and the external environment determines DCP. Therefore, the external environment determines clustering. Given the clustering results at different levels, we can infer the significance of the external environment in reverse.

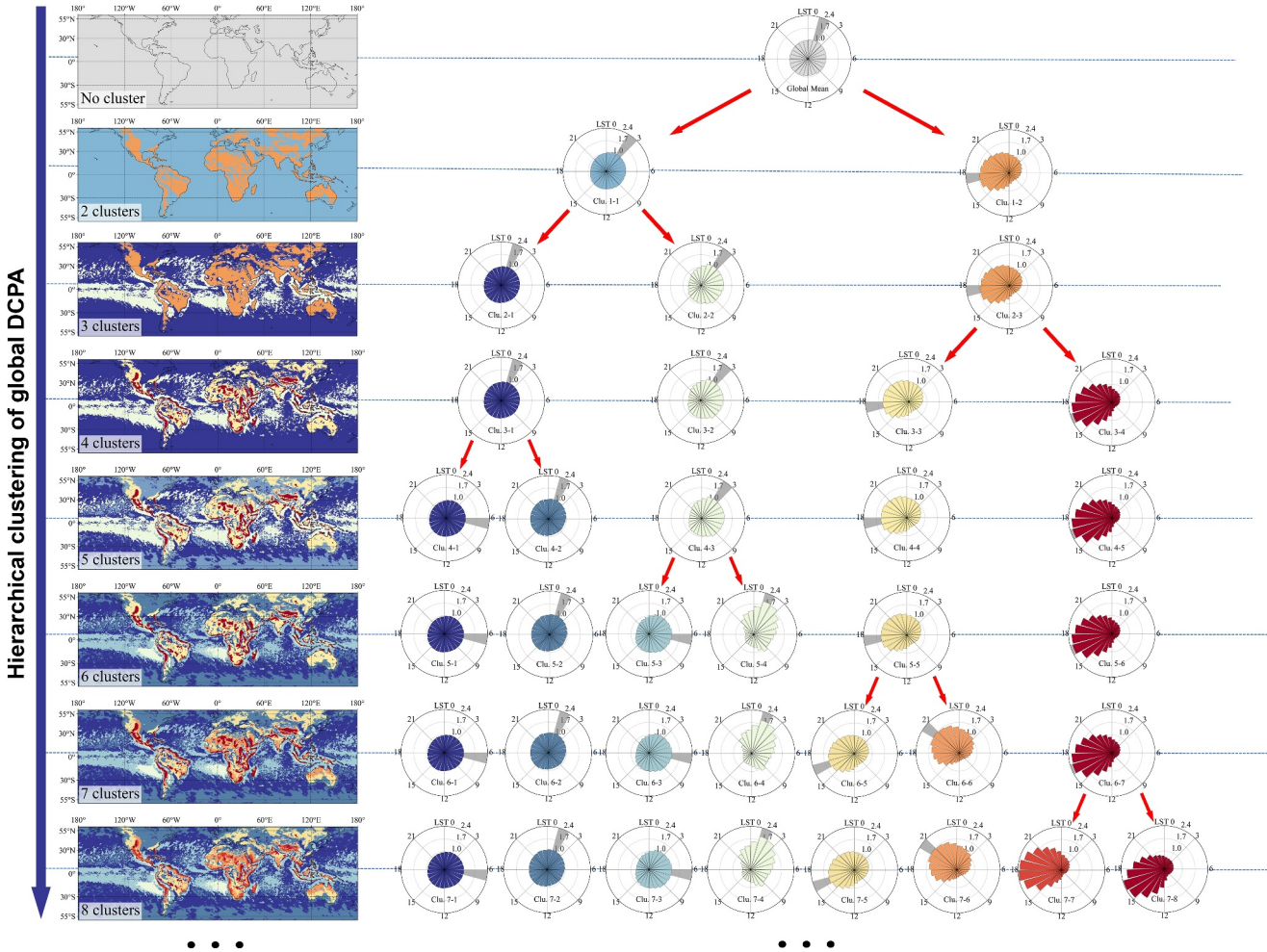

Figure 3. The clustering process of the global diurnal cycle of precipitation amount. Different colors represent different clusters. The rose diagrams show the average for each cluster.

During the first split, the model minimizes SSE by selecting the two most distinct clusters. In addition to the hourly precipitation itself, samples within the same cluster exhibit other common characteristics. For instance, cluster 2-2 shares similar spatial features, predominantly distributed over land and local maximum altitudes, whereas cluster 2–1 is primarily found over oceans or other terrestrial regions. Given the extensive impact of topography and land-sea relationships on precipitation, combined with these clustering results, we can infer that the primary factors influencing the spatial heterogeneity of global DCPA are the extreme thermal contrasts between land and sea, as well as between mountains and valleys, which is also supported by previous EOF-based studies (Lu et al., 2021; Pritchard & Somerville, 2009; Teo et al., 2011).

Further subdivision reveals the primary factors influencing DCP variation within each respective cluster, which are also significant factors affecting global DCPA. For the majority of oceans represented by cluster 2–1, the distance from the coastline is a crucial factor influencing their DCP. Similarly, for the majority of land areas represented by cluster 2-2, both the distance from the coastline and elevation contrast may be key factors affecting their DCP. These phenomena may be linked to the gravity waves and spatial scale of sea-land breeze (Huang & Wang, 2014; Rani et al., 2010; Short et al., 2019; Vincent & Lane, 2016; Yamanaka et al., 2018; Zhu et al., 2022). In summary, the hierarchical method offers good interpretability for DCP clustering, facilitating the presentation and comprehension of its clustering process, making it highly valuable in understanding high-resolution spatial differences and their causes of DCP.

4. Conclusions

Based on satellite data of hourly precipitation, we use a hierarchical method to objectively cluster the global diurnal cycle of precipitation amount, frequency, and intensity, respectively, throughout the entire year, as well as separately for summer and winter. The results suggest that areas characterized by complex coastal and topographical features display more intricate clustering patterns in the diurnal precipitation cycle. Clusters with features such as peak value, peak time, and uniformity undergo significant transitions from coastlines to adjacent areas and from high-altitude regions to their surroundings, particularly noticeable in low-latitude regions. The significant inter-cluster differences indicate that the precipitation amount and frequency have stronger diurnal cycles than that of the precipitation intensity globally. The seasonal transitions of clusters suggest that the global diurnal cycle of precipitation exhibits significant seasonal differences. The hierarchical clustering process of the global diurnal cycle of precipitation demonstrates better interpretability. Compared to existing studies, our results fill gaps in the research landscape concerning diurnal precipitation cycle clustering, encompassing precipitation amount, frequency, intensity, and seasonality on a global scale. Our information-rich global maps of the diurnal precipitation cycle can provide valuable insights into the mechanisms underlying the diurnal cycle of precipitation, sub-daily simulation by climate models and the impacts of the diurnal precipitation cycle on nature and society.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

IMERG v06 precipitation data are available at the NASA precipitation processing system (Huffman et al., 2015). ERA5 precipitation data were downloaded from the ECMWF Climate Data Store (Hersbach et al., 2020). Copernicus digital elevation model is available from Open Topography (Copernicus, 2024). The clustering algorithm in this article is implemented using Scikit-learn (Pedregosa et al., 2011). Figures and maps in the article are plotted using Matplotlib (Hunter, 2007).

Acknowledgments
This research was supported by the National Natural Science Foundation of China (Grant 42361144001; 42101041) and Guangdong Provincial Key Laboratory of Intelligent Disaster Prevention and Emergency Technologies for Urban Lifeline Engineering (2022, Grant 2022B1212010016). It was also partly supported by the Henan Provincial Key Laboratory of Hydrosphere and Watershed Water Security. Computation in this study was supported by the Centre for Computational Science and Engineering at the Southern University of Science and Technology.

References

Aoki, S., & Shige, S. (2024). Control of low-level wind on the diurnal cycle of tropical coastal precipitation. *Journal of Climate*, *37*(1), 229–247. https://doi.org/10.1175/JCLI-D-23-0180.1

Barman, N., Borgohain, A., Kundu, S. S., & Kumar, N. K. (2021). Seasonal variation of mountain-valley wind circulation and surface layer parameters over the mountainous terrain of the northeastern region of India. *Theoretical and Applied Climatology*, *143*(3–4), 1501–1512. https://doi.org/10.1007/s00704-020-03491-y

Bedoya-Soto, J. M., Aristizábal, E., Carmona, A. M., & Poveda, G. (2019). Seasonal shift of the diurnal cycle of rainfall over Medellin's Valley, Central Andes of Colombia (1998–2005). *Frontiers in Earth Science*, *7*, 92. https://doi.org/10.3389/feart.2019.00092

Chen, G., Lan, R., Zeng, W., Pan, H., & Li, W. (2018). Diurnal variations of rainfall in surface and satellite observations at the monsoon coast (South China). *Journal of Climate*, *31*(5), 1703–1724. https://doi.org/10.1175/JCLI-D-17-0373.1

Chen, G., Sha, W., & Iwasaki, T. (2009). Diurnal variation of precipitation over southeastern China: Spatial distribution and its seasonality. *Journal of Geophysical Research*, *114*(D13). https://doi.org/10.1029/2008JD011103

Cook, K. H., & Vizy, E. K. (2024). Understanding the regionality and diurnal cycles of precipitation in the Lake Victoria Basin during Boreal fall. *Climate Dynamics*, *62*(2), 1359–1378. https://doi.org/10.1007/s00382-023-06950-0

Copernicus. (2024). Copernicus global digital elevation model [Dataset]. *Open Topography*. https://doi.org/10.5069/G9028PQB

Coppin, D., & Bellon, G. (2019). Physical mechanisms controlling the offshore propagation of convection in the tropics: 1. Flat island. *Journal of Advances in Modeling Earth Systems*, *11*(9), 3042–3056. https://doi.org/10.1029/2019MS001793

Covey, C., Gleckler, P. J., Doutriaux, C., Williams, D. N., Dai, A., Fasullo, J., et al. (2016). Metrics for the diurnal cycle of precipitation: Toward routine benchmarks for climate models. *Journal of Climate*, *29*(12), 4461–4471. https://doi.org/10.1175/JCLI-D-15-0664.1

Dai, A. (2024). The diurnal cycle from observations and ERA5 in precipitation, clouds, boundary layer height, buoyancy, and surface fluxes. *Climate Dynamics*, 1–30. https://doi.org/10.1007/s00382-024-07182-6

Dai, A., Lin, X., & Hsu, K.-L. (2007). The frequency, intensity, and diurnal cycle of precipitation in surface and satellite observations over low- and mid-latitudes. *Climate Dynamics*, *29*(7–8), 727–744. https://doi.org/10.1007/s00382-007-0260-y

Deng, H., Pepin, N., Chen, Y., Guo, B., Zhang, S., Zhang, Y., et al. (2022). Dynamics of diurnal precipitation differences and their spatial variations in China. *Journal of Applied Meteorology and Climatology*, *61*(8), 1015–1027. https://doi.org/10.1175/JAMC-D-21-0232.1

Di, J., & Gou, X. (2018). Bisecting K-means algorithm based on K-valued selfdetermining and clustering center optimization. *Journal of Computers*, *13*(6), 588–595. https://doi.org/10.17706/jcp.13.6

Doan, Q. V., Dipankar, A., Simón-Moral, A., Sanchez, C., Prasanna, V., Roth, M., & Huang, X. Y. (2021). Urban-induced modifications to the diurnal cycle of rainfall over a tropical city. *Quarterly Journal of the Royal Meteorological Society*, *147*(735), 1189–1201. https://doi.org/10.1002/qj.3966

Du, Y., & Rotunno, R. (2018). Diurnal cycle of rainfall and winds near the south coast of China. *Journal of the Atmospheric Sciences*, *75*(6), 2065–2082. https://doi.org/10.1175/JAS-D-17-0397.1

Dunn, R. J., Willett, K. M., Thorne, P. W., Woolley, E. V., Durre, I., Dai, A., et al. (2012). HadISD: A quality-controlled global synoptic report database for selected variables at long-term stations from 1973–2011. *Climate of the Past*, *8*(5), 1649–1679. https://doi.org/10.5194/cp-8-1649-2012

Fang, J., & Du, Y. (2022). A global survey of diurnal offshore propagation of rainfall. *Nature Communications*, *13*(1), 7437. https://doi.org/10.1038/s41467-022-34842-0

Garreaud, R. D., Molina, A., & Farias, M. (2010). Andean uplift, ocean cooling and atacama hyperaridity: A climate modeling perspective. *Earth and Planetary Science Letters*, *292*(1–2), 39–50. https://doi.org/10.1016/j.epsl.2010.01.017

Giles, J. A., Ruscica, R. C., & Menéndez, C. G. (2020). The diurnal cycle of precipitation over South America represented by five gridded datasets. *International Journal of Climatology*, *40*(2), 668–686. https://doi.org/10.1002/joc.6229

Guth, P. L., & Geoffroy, T. M. (2021). LiDAR point cloud and ICESat-2 evaluation of 1 second global digital elevation models: Copernicus wins. *Transactions in GIS*, *25*(5), 2245–2261. https://doi.org/10.1111/tgis.12825

Hayden, L., & Liu, C. (2021). Differences in the diurnal variation of precipitation estimated by spaceborne radar, passive microwave radiometer, and IMERG. *Journal of Geophysical Research: Atmospheres*, *126*(9), e2020JD033020. https://doi.org/10.1029/2020JD033020

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis [Dataset]. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hirose, M., & Nakamura, K. (2005). Spatial and diurnal variation of precipitation systems over Asia observed by the TRMM Precipitation Radar. *Journal of Geophysical Research*, *110*(D5). https://doi.org/10.1029/2004JD004815

Huang, W.-R., & Wang, S.-Y. (2014). Impact of land–sea breezes at different scales on the diurnal rainfall in Taiwan. *Climate Dynamics*, *43*(7–8), 1951–1963. https://doi.org/10.1007/s00382-013-2018-z

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Xie, P., & Yoo, S.-H. (2015). NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG) [Dataset]. *Algorithm Theoretical Basis Document (ATBD) Version*, *4*(26). https://doi.org/10.1007/978-3-030-24568-9_19

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment [Software]. *Computing in Science & Engineering*, *9*(03), 90–95. https://doi.org/10.1109/MCSE.2007.55

Janowiak, J. E., Kousky, V. E., & Joyce, R. J. (2005). Diurnal cycle of precipitation determined from the CMORPH high spatial and temporal resolution global precipitation analyses. *Journal of Geophysical Research*, *110*(D23). https://doi.org/10.1029/2005JD006156

Jeong, J.-H., Walther, A., Nikulin, G., Chen, D., & Jones, C. (2011). Diurnal cycle of precipitation amount and frequency in Sweden: Observation versus model simulation. *Tellus A: Dynamic Meteorology and Oceanography*, *63*(4), 664–674. https://doi.org/10.1111/j.1600-0870.2011.00517.x

Jiang, Z., Zhang, D.-L., Xia, R., & Qian, T. (2017). Diurnal variations of presummer rainfall over southern China. *Journal of Climate*, *30*(2), 755–773. https://doi.org/10.1175/JCLI-D-15-0666.1

Johnson, R. H. (2011). Diurnal cycle of monsoon convection. In *The global monsoon system: Research and forecast* (pp. 257–276). World Scientific. https://doi.org/10.1142/9789814343411_0015

Kikuchi, K., & Wang, B. (2008). Diurnal precipitation regimes in the global tropics. *Journal of Climate*, *21*(11), 2680–2696. https://doi.org/10.1175/2007JCLI2051.1

Kincer, J. B. (1916). Daytime and nighttime precipitation and their economic significance. *Monthly Weather Review*, *44*(11), 628–633. https://doi.org/10.1175/1520-0493(1916)44<628:DANPAT>2.0.CO;2

Lewis, E., Fowler, H., Alexander, L., Dunn, R., McClean, F., Barbero, R., et al. (2019). Gsdr: A global sub-daily rainfall dataset. *Journal of Climate*, *32*(15), 4715–4729. https://doi.org/10.1175/JCLI-D-18-0143.1

Li, Y., & Carbone, R. (2015). Offshore propagation of coastal precipitation. *Journal of the Atmospheric Sciences*, *72*(12), 4553–4568. https://doi.org/10.1175/JAS-D-15-0104.1

Lima, D. C., Soares, P. M., Semedo, A., Cardoso, R. M., Cabos, W., & Sein, D. V. (2019). A climatological analysis of the Benguela coastal low-level jet. *Journal of Geophysical Research: Atmospheres*, *124*(7), 3960–3978. https://doi.org/10.1029/2018JD028944

Liu, J., Yang, L., Jiang, J., Yuan, W., & Duan, Z. (2021). Mapping diurnal cycles of precipitation over China through clustering. *Journal of Hydrology*, *592*, 125804. https://doi.org/10.1016/j.jhydrol.2020.125804

Lu, H.-L., Qiu, J., Li, M.-J., Zuo, H.-M., Li, J.-L., Hu, B. X., & Li, F.-F. (2024). Temporal and spatial variations in the sub-daily precipitation structure over the Qinghai–Tibet Plateau (QTP). *Science of the Total Environment*, *915*, 170153. https://doi.org/10.1016/j.scitotenv.2024.170153

Lu, J., Li, T., & Wang, L. (2021). Precipitation diurnal cycle over the maritime continent modulated by the climatological annual cycle. *Journal of Climate*, *34*(4), 1387–1402. https://doi.org/10.1175/JCLI-D-20-0130.1

Mao, J., & Wu, G. (2012). Diurnal variations of summer precipitation over the Asian monsoon region as revealed by TRMM satellite data. *Science China Earth Sciences*, *55*(4), 554–566. https://doi.org/10.1007/s11430-011-4315-x

Maulana, M. T., Yamazaki, T., Iwasaki, T., & Abdillah, M. R. (2023). Regional variation of the influence of cross-equatorial northerly surge towards diurnal cycle of rainfall over Java Island. *Geoscience Letters*, *10*(1), 40. https://doi.org/10.1186/s40562-023-00293-8

Mu, X., Huang, A., Wu, Y., Xu, Q., Zheng, Y., Lin, H., et al. (2021). Characteristics of the precipitation diurnal variation and underlying mechanisms over Jiangsu, eastern China, during warm season. *Frontiers in Earth Science*, *9*, 703071. https://doi.org/10.3389/feart.2021.703071

Murtagh, F., & Contreras, P. (2012). Algorithms for hierarchical clustering: An overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *2*(1), 86–97. https://doi.org/10.1002/widm.53

Olaguera, L. M. P., Llorin, A. G. A., Magnaye, A. M. T., Cruz, F. A. T., Villarin, J. R. T., & Topacio, X. G. V. M. (2024). Observed characteristics and seasonality of the diurnal patterns of precipitation over Metro Manila, Philippines. *Theoretical and Applied Climatology*, *155*(2), 1123–1138. https://doi.org/10.1007/s00704-023-04684-x

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python [Software]. *Journal of Machine Learning Research*, *12*, 2825–2830. https://doi.org/10.5555/1953048.2078195

Pritchard, M. S., & Somerville, R. C. (2009). Empirical orthogonal function analysis of the diurnal cycle of precipitation in a multi-scale climate model. *Geophysical Research Letters*, *36*(5). https://doi.org/10.1029/2008GL036964

Qin, S., Wang, K., Wu, G., & Ma, Z. (2021). Variability of hourly precipitation during the warm season over eastern China using gauge observations and ERA5. *Atmospheric Research*, *264*, 105872. https://doi.org/10.1016/j.atmosres.2021.105872

Qiu, C., Dong, X. g., Shen, L. c., & Wen, J. h. (2023). Diurnal cycle of precipitation and its possible causes: A case study in Shandong Province, China. *International Journal of Climatology*, *43*(13), 5966–5983. https://doi.org/10.1002/joc.8183

Rani, S. I., Ramachandran, R., Subrahamanyam, D. B., Alappattu, D. P., & Kunhikrishnan, P. (2010). Characterization of sea/land breeze circulation along the west coast of Indian sub-continent during pre-monsoon season. *Atmospheric Research*, *95*(4), 367–378. https://doi.org/10.1016/j.atmosres.2009.10.009

Ruiz-Hernández, J.-C., Condom, T., Ribstein, P., Le Moine, N., Espinoza, J.-C., Junquas, C., et al. (2021). Spatial variability of diurnal to seasonal cycles of precipitation from a high-altitude equatorial Andean valley to the Amazon Basin. *Journal of Hydrology: Regional Studies*, *38*, 100924. https://doi.org/10.1016/j.ejrh.2021.100924

Shen, L., Zhao, C., & Yang, X. (2021). Insight into the seasonal variations of the sea-land breeze in Los Angeles with respect to the effects of solar radiation and climate type. *Journal of Geophysical Research: Atmospheres*, *126*(6), e2020JD033197. https://doi.org/10.1029/2020JD033197

Short, E., Vincent, C. L., & Lane, T. P. (2019). Diurnal cycle of surface winds in the Maritime Continent observed through satellite scatterometry. *Monthly Weather Review*, *147*(6), 2023–2044. https://doi.org/10.1175/MWR-D-18-0433.1

Smith, A., Lott, N., & Vose, R. (2011). The integrated surface database: Recent developments and partnerships. *Bulletin of the American Meteorological Society*, *92*(6), 704–708. https://doi.org/10.1175/2011BAMS3015.1

Song, Y., & Wei, J. (2021). Diurnal cycle of summer precipitation over the North China Plain and associated land–atmosphere interactions: Evaluation of ERA5 and MERRA-2. *International Journal of Climatology*, *41*(13), 6031–6046. https://doi.org/10.1002/joc.7166

Takahashi, H., Fujinami, H., Yasunari, T., & Matsumoto, J. (2010). Diurnal rainfall pattern observed by tropical rainfall measuring mission precipitation radar (TRMM-PR) around the Indochina peninsula. *Journal of Geophysical Research*, *115*(D7). https://doi.org/10.1029/2009JD012155

Tan, J., Huffman, G. J., Bolvin, D. T., & Nelkin, E. J. (2019). Diurnal cycle of IMERG V06 precipitation. *Geophysical Research Letters*, *46*(22), 13584–13592. https://doi.org/10.1029/2019GL085395

Teo, C.-K., Koh, T.-Y., Chun-Fung Lo, J., & Chandra Bhatt, B. (2011). Principal component analysis of observed and modeled diurnal rainfall in the Maritime Continent. *Journal of Climate*, *24*(17), 4662–4675. https://doi.org/10.1175/2011JCLI4047.1

Tu, S., Chan, J. C., Xu, J., Zhong, Q., Zhou, W., & Zhang, Y. (2022). Increase in tropical cyclone rain rate with translation speed. *Nature Communications*, *13*(1), 7325. https://doi.org/10.1038/s41467-022-35113-8

Vincent, C. L., & Lane, T. P. (2016). Evolution of the diurnal precipitation cycle with the passage of a Madden–Julian oscillation event through the Maritime Continent. *Monthly Weather Review*, *144*(5), 1983–2005. https://doi.org/10.1175/MWR-D-15-0326.1

Wang, J., Yuan, H., Wang, X., Cui, C., & Wang, X. (2023). Impact of thermally forced circulations on the diurnal cycle of summer precipitation over the southeastern Tibetan Plateau. *Geophysical Research Letters*, *50*(6), e2022GL100951. https://doi.org/10.1029/2022GL100951

Wilks, D. S. (2011). Principal component (EOF) analysis. *International Geophysics*, *100*, 519–562. https://doi.org/10.1016/B978-0-12-385022-5.00012-9

Wu, Y., Huang, A., Huang, D., Chen, F., Yang, B., Zhou, Y., et al. (2018). Diurnal variations of summer precipitation over the regions east to Tibetan Plateau. *Climate Dynamics*, *51*(11), 4287–4307. https://doi.org/10.1007/s00382-017-4042-x

Xiao, C., Yuan, W., & Yu, R. (2018). Diurnal cycle of rainfall in amount, frequency, intensity, duration, and the seasonality over the UK. *International Journal of Climatology*, *38*(13), 4967–4978. https://doi.org/10.1002/joc.5790

Yamanaka, M. D., Ogino, S.-Y., Wu, P.-M., Jun-Ichi, H., Mori, S., Matsumoto, J., & Syamsudin, F. (2018). Maritime continent coastlines controlling Earth's climate. *Progress in Earth and Planetary Science*, *5*(1), 1–28. https://doi.org/10.1186/s40645-018-0174-9

Yang, M., Liu, J., Wang, Y., Chen, J. M., Cui, Z., Zhang, Z., et al. (2023). Prominent impact of diurnal rainfall variations on evapotranspiration and gross primary productivity in forests over low latitudes. *Agricultural and Forest Meteorology*, *342*, 109740. https://doi.org/10.1016/j.agrformet.2023.109740

Yaqub, A., Seibert, P., & Formayer, H. (2011). Diurnal precipitation cycle in Austria. *Theoretical and Applied Climatology*, *103*(1–2), 109–118. https://doi.org/10.1007/s00704-010-0281-z

Yin, S., Chen, D., & Xie, Y. (2009). Diurnal variations of precipitation during the warm season over China. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *29*(8), 1154–1170. https://doi.org/10.1002/joc.1758

Yin, S., Li, W., Chen, D., Jeong, J.-H., & Guo, W. (2011). Diurnal variations of summer precipitation in the Beijing area and the possible effect of topography and urbanization. *Advances in Atmospheric Sciences*, *28*(4), 725–734. https://doi.org/10.1007/s00376-010-9240-y

Yu, R., Li, J., Chen, H., & Yuan, W. (2014). Progress in studies of the precipitation diurnal variation over contiguous China. *Journal of Meteorological Research*, *28*(5), 877–902. https://doi.org/10.1007/s13351-014-3272-7

Yu, R., Yuan, W., & Li, J. (2013). The asymmetry of rainfall process. *Chinese Science Bulletin*, *58*(16), 1850–1856. https://doi.org/10.1007/s11434-012-5653-6

Yu, R., Zhou, T., Xiong, A., Zhu, Y., & Li, J. (2007). Diurnal variations of summer precipitation over contiguous China. *Geophysical Research Letters*, *34*(1). https://doi.org/10.1029/2006GL028129

Zhang, L., Chen, X., Lai, R., & Zhu, Z. (2022). Performance of satellite-based and reanalysis precipitation products under multi-temporal scales and extreme weather in mainland China. *Journal of Hydrology*, *605*, 127389. https://doi.org/10.1016/j.jhydrol.2021.127389

Zhang, Y., Xue, M., Zhu, K., & Zhou, B. (2019). What is the main cause of diurnal variation and nocturnal peak of summer precipitation in Sichuan Basin, China? The key role of boundary layer low-level jet inertial oscillations. *Journal of Geophysical Research: Atmospheres*, *124*(5), 2643–2664. https://doi.org/10.1029/2018JD029834

Zhao, D., Dong, W., Lin, Y., Hu, Y., & Cao, D. (2022). Diurnal variation of precipitation over the high mountain Asia: Spatial distribution and its seasonality. *Journal of Hydrometeorology*, *23*(12), 1945–1959. https://doi.org/10.1175/JHM-D-21-0243.1

Zhu, B., Du, Y., & Gao, Z. (2022). Influences of MJO on the diurnal variation and associated offshore propagation of rainfall near western coast of sumatra. *Atmosphere*, *13*(2), 330. https://doi.org/10.3390/atmos13020330