Chaos

The climatic interdependence of extreme-rainfall events around the globe ⒻF

ⒾD **Zhen Su,** ⒾD **Henning Meyerhenke and** ⒾD **Jürgen Kurths**

COLLECTIONS

Paper published as part of the special topic on Theory-informed and Data-driven Approaches to Advance Climate Sciences

ⒻF   This paper was selected as Featured

openaccess   View Online   Export Citation   CrossMark

ARTICLES YOU MAY BE INTERESTED IN

Network approach reveals the spatiotemporal influence of traffic on air pollution under COVID-19
Chaos: An Interdisciplinary Journal of Nonlinear Science **32**, 041106 (2022); https://doi.org/10.1063/5.0087844

A metapopulation approach to identify targets for Wolbachia-based dengue control
Chaos: An Interdisciplinary Journal of Nonlinear Science **32**, 041105 (2022); https://doi.org/10.1063/5.0087435

Phase and frequency linear response theory for hyperbolic chaotic oscillators
Chaos: An Interdisciplinary Journal of Nonlinear Science **32**, 043124 (2022); https://doi.org/10.1063/5.0064519

AIP Publishing

The climatic interdependence of extreme-rainfall events around the globe Ⓕ

View Online   Export Citation   CrossMark

Zhen Su,[1,2,a)] Ⓘ Henning Meyerhenke,[2] Ⓘ and Jürgen Kurths[1,3] Ⓘ

AFFILIATIONS

[1] Potsdam Institute for Climate Impact Research, 14473 Potsdam, Germany
[2] Department of Computer Science, Humboldt-Universität zu Berlin, 12489 Berlin, Germany
[3] Department of Physics, Humboldt-Universität zu Berlin, 12489 Berlin, Germany

Note: This article is part of the Focus Issue, Theory-informed and Data-driven Approaches to Advance Climate Sciences.
[a)] **Author to whom correspondence should be addressed:** zhen.su@pik-potsdam.de

ABSTRACT

The identification of regions of similar climatological behavior can be utilized for the discovery of spatial relationships over long-range scales, including teleconnections. Additionally, it provides insights for the improvement of corresponding interaction processes in general circulation models. In this regard, the global picture of the interdependence patterns of extreme-rainfall events (EREs) still needs to be further explored. To this end, we propose a top-down complex-network-based clustering workflow, with the combination of consensus clustering and mutual correspondences. Consensus clustering provides a reliable community structure under each dataset, while mutual correspondences build a matching relationship between different community structures obtained from different datasets. This approach ensures the robustness of the identified structures when multiple datasets are available. By applying it simultaneously to two satellite-derived precipitation datasets, we identify consistent synchronized structures of EREs around the globe, during boreal summer. Two of them show independent spatiotemporal characteristics, uncovering the primary compositions of different monsoon systems. They explicitly manifest the primary intraseasonal variability in the context of the global monsoon, in particular, the "monsoon jump" over both East Asia and West Africa and the mid-summer drought over Central America and southern Mexico. Through a case study related to the Asian summer monsoon, we verify that the intraseasonal changes of upper-level atmospheric conditions are preserved by significant connections within the global synchronization structure. Our work advances network-based clustering methodology for (i) decoding the spatiotemporal configuration of interdependence patterns of natural variability and for (ii) the intercomparison of these patterns, especially regarding their spatial distributions over different datasets.

Precipitation variability of monsoons affects over two-thirds of the world's population, and regional monsoons have their own characteristics due to specific land–ocean and topographic conditions.[1] In spite of being distributed in different continental regions, they are essentially driven and synchronized by the annual cycle of solar radiation. The connections between them are via the global divergent circulation characterized by global-scale persistent overturning of the atmosphere varying with time.[2,3] An integration of these monsoons forms the concept of the global monsoon in terms of similar dynamics and behaviors. The synchronization in the context of the global monsoon is not limited to the tropical regions, but also extends to the subtropics, with the East Asian monsoon being a typical example. Therefore, identifying the synchronization structure on a global scale helps to understand the interaction with mid-latitude regions. A clustering workflow with higher robustness, by combining consensus clustering and mutual correspondences, is proposed for this purpose.

I. INTRODUCTION

Regions of similar climatological properties are somewhat coherent in time and space. Therefore, regional weather systems, formed within these areas, can exhibit synchronization behavior. Synchronization is also observed in the climate system, which spans thousands of kilometers. This may be an indication of a

teleconnection.[4,5] Advances in the identification of these synchronized regions help one to better understand the self-organized structure of the climate system and to further improve prediction skills.[6,7]

A complex-network-based clustering approach has proved to be useful in this respect due to the climatological interpretation of the identified communities.[8] It takes as input complex networks reconstructed from climate data. Generally speaking, the reconstruction process treats grid points as nodes and establishes connections between nodes with a high correlation according to the corresponding timeseries. Clustering such networks has been used for the discovery of spatial relationships in climate variables,[8–10] for the extraction of climate indices in a data-driven fashion,[6,7,11] and for the intercomparison of performance of general circulation models (GCMs).[12] However, cluster analysis remains a challenging problem in this domain because of the availability of multiple datasets on each single climate variable. For example, each dataset of precipitation produces a network, which further yields a community structure. How to extract representative and consistent communities when multiple community structures are obtained from different datasets on the same variable is a key question.

Regarding global-scale analysis,[8–10,13] previous studies based on network clustering put little focus on precipitation data, let alone EREs. This is essentially due to the bias of precipitation estimation in reanalysis products.[8,14] However, it has recently been pointed out in the Sixth Assessment of the Intergovernmental Panel on Climate Change (IPCC) that precipitation extremes will be very likely to become more frequent in most locations.[15] Indeed, the investigation of synchronization behavior of extreme rainfall is of societal relevance due to the occurrences of natural hazards, such as floods or landslides.[14,16] A particular complex-network-based study on EREs[17] starts from the utilization of event synchronization [ES, a method originally introduced to study electroencephalogram (EEG) signals][18] as a nonlinear similarity measure for the reconstruction of the functional climate networks. Using observational data, some studies emphasize how the network-based clustering decodes the spatial relationship on a regional scale.[16,19–21] A recent investigation reveals the global structure of synchronization of EREs based on high-resolution satellite data.[22] However, it relies on a given region of interest, and therefore, it provides a partial view. The focus of this study is, therefore, to reveal a comprehensive view of this global structure.

We address the above-mentioned issues through a systematic network-based clustering workflow. Our aim is to unravel the synchronization of EREs around the globe, especially in terms of intraseasonal variability over different monsoon systems. Specifically, the representation of climate data in the form of networks is implemented based on ES.[14,18,22] The reconstructed functional climate networks are then viewed as input for a downstream cluster analysis. Through this proposed clustering workflow, we identify two primary monsoon-related structures of distinct spatiotemporal characteristics during boreal summer. Our work is the first combined application of consensus clustering[23] and mutual correspondences[24] for climate extreme analysis. This combination provides the identification of reliable interdependence patterns not only within each dataset, but also over different ones.

II. METHODOLOGY

We develop here a generic workflow for the identification of the global synchronization structure of EREs, as presented in Fig. 1. The network reconstruction of functional climate networks, i.e., Steps (1) and (2), is based on Ref. 22. For the downstream analysis, our emphasis is on the proposed combined clustering workflow, in particular, Steps (3.1) and (3.2). At the same time, we also include the correction for the multiple-comparison bias in Step (3.3) by using the technique introduced in Ref. 22.

A. The reconstruction of functional climate networks

1. Definition of EREs based on thresholding

For rainfall data, we adopt the $p = 90$th percentile of wet days (daily sum rainfall above 1 mm), by the definition of extreme weather events from IPCC,[25] as the threshold to indicate EREs. Suppose that a pair of grid points i and j are randomly chosen from the entire set of grid cells. For the two timeseries at i and j, days with rainfall values higher than the given threshold are kept to indicate the occurrences of EREs. For consecutive days with EREs, only the first day is preserved to represent an independent occurrence of events.[22] We denote the set of grid points with at least three events as V. The sets that incorporate the final event series, after data preprocessing at i and j, are then defined as $e_i^p := \{t_{i,\alpha}\}$ and $e_j^p := \{t_{j,\beta}\}$ with $\alpha = 1, \ldots, n_i, \beta = 1, \ldots, n_j$, respectively. For i, $t_{i,\alpha}$ denotes the time when the αth event occurs and n_i is the total number of events. Similarly, $t_{j,\beta}$ and n_j are the corresponding variables for j.
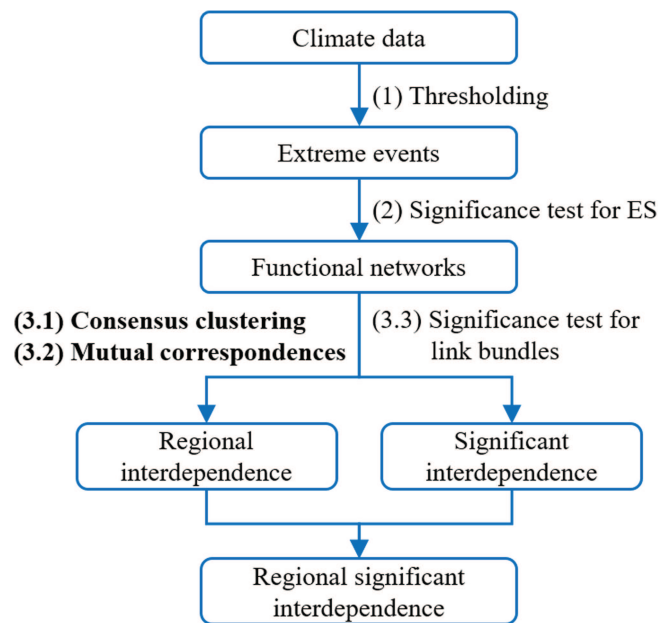

FIG. 1. Schematic overview of the complex-network-based clustering workflow. Steps (1), (2), and (3.3) of the framework have been provided in Ref. 22. The new steps (3.1) and (3.2) are proposed in this paper.

2. ES and the corresponding significance test

To quantify and determine the interactions between i and j, ES considers all possible pairs of events $t_{i,\alpha}$ and $t_{j,\beta}$ and measures the closeness of each pair by imposing the condition that the absolute value of temporal delay $t_{i,j}^{\alpha,\beta} := |t_{i,\alpha} - t_{j,\beta}|$ must be smaller than a dynamical delay $\tau_{i,j}^{\alpha,\beta} := \frac{1}{2}\min\{t_{i,i}^{\alpha,\alpha-1}, t_{i,i}^{\alpha,\alpha+1}, t_{j,j}^{\beta,\beta-1}, t_{j,j}^{\beta,\beta+1}\}$. Together with a maximum delay $\tau_{max} = 10$ days, which is used to exclude the occurrences of unreasonably long delays, the $ES_{i,j}$ is, therefore, defined between e_i^p and e_j^p as[22]

$$ES_{i,j} := \left| \left\{ (\alpha, \beta) : t_{i,j}^{\alpha,\beta} < \tau_{i,j}^{\alpha,\beta} \wedge t_{i,j}^{\alpha,\beta} \leq \tau_{max} \right\} \right|, \quad (1)$$

where $|\cdot|$ denotes the cardinality of a given set, i.e., the number of synchronized event pairs.

The placement of network links between each pair of grid points is determined statistically. Taking i and j as an example, by randomly distributing the same number of events 2000 times within the June–July–August (JJA) season and computing the corresponding ES values for the 2000 pairs of surrogate event series, a null-model distribution is obtained.[22] We take the 99.5th percentile of the distribution as the threshold. A significant connection with $P < 0.005$ is determined if $ES_{i,j}$ is above the threshold. Since the significance thresholds are only associated with the number of events, we manage this process independently for all possible pairs of event numbers.

In many scenarios, links of a network can be attached with features, such as weight and direction. For the reconstructed functional climate networks, each significant link has an ES value as a weight to indicate the link strength. This is an important factor in Step (3.1) (see Sec. II B) to determine the affinities of different nodes. Regarding the link direction, it can also be incorporated into the calculation of ES and the corresponding test of significance. For example, the synchronization from i to j is defined as

$$ES_{i\to j} := \left| \left\{ (\alpha, \beta) : t_{i\to j}^{\alpha,\beta} < \tau_{i,j}^{\alpha,\beta} \wedge t_{i\to j}^{\alpha,\beta} \leq \tau_{max} \right\} \right|, \quad (2)$$

where $t_{i\to j}^{\alpha,\beta} := t_{j,\beta} - t_{i,\alpha}$ and $t_{i\to j}^{\alpha,\beta} > 0$ indicates the occurrence of an event at i with a subsequent one at j. The significance test for the directed case follows the same procedure as the one for the undirected one. We take the undirected weighted network as the default setting over the following main text. The directed approach is only used when identifying days of high synchronization for atmospheric condition analysis in Sec. II B.

B. The network-based clustering workflow

1. Consensus clustering

The significant link bundles given by Step (3.3)[22] have limited capability to capture regions following similar climatological behavior around the globe. This is essentially due to the limited link distribution based on a selected region (see Fig. 6 in Appendix A as an example). Therefore, we adopt here complex-network-based clustering, namely, community detection, to fill this gap, by taking into account the link distribution of a whole reconstructed functional network. However, most community detection methods are not deterministic, indicating that partitions delivered by them have

certain fluctuations, e.g., due to randomness. One way to manage this problem is to use consensus clustering.[23] Using the consensus of a combination of partitions provides a stable result out of a set of candidates. In the context of climate science, to the best of our knowledge, the use of consensus clustering represents a novelty; using a stable partition derived this way yields an algorithmically reliable and climatologically interpretable community structure for a given dataset.

Given the reconstructed functional climate network G with $|V|$ nodes and a community detection method A, below is the detailed procedure of consensus clustering:

(A1) Apply A on G for 1000 times, yielding 1000 partitions.
(A2) Rank the 1000 partitions based on modularity[26] and select the first n_p partitions of the highest modularity.
(A3) Compute the consensus matrix D: D_{ij} is the number of partitions where nodes i and j belong to the same community, divided by n_p.
(A4) All entries of D below a chosen threshold θ are set to zero.
(A5) Apply A on D for n_p times, yielding n_p partitions.
(A6) If the partitions in (A5) are all equal, that is, D is block-diagonal, then stop; otherwise, go back to (A3).

Note that A serves as a basic method for the entire procedure, and we use modularity to estimate the quality/strength of a partition due to the lack of climate-related benchmarks. Modularity compares the fraction of intra-community links with its expected number in a null model, which is a network with the same degree sequence but links placed at random.[26] Intuitively speaking, networks with high modularity have dense intra-community connections but are sparsely connected between different communities.

In (A1), we use the parallel Louvain method (PLM)[27] as the basic method as algorithm A. We choose PLM because (a) it is flexible enough to control the granularity/size of communities with only one parameter γ, (b) its speed allows one to handle large-scale networks, and (c) it is based on the well-known Louvain method.[28] The Louvain method is a greedy local search algorithm with a bottom-up multilevel approach. Specifically, the local search phase moves each node to the neighbor community where this move yields the largest increase in modularity (if any). After this process has converged, each community is contracted to a super-node. The super-nodes are linked together by weighted edges (and self-loops) whose weights are set according to the inter-community (intra-community) edge weights in the larger graph. The smaller network derived this way is treated as new input for the next iteration of local search and coarsening (until no further community changes are observed). Among others, the main change of PLM compared to plain Louvain is to use parallelism. PLM is part of the network analysis toolkit NetworKit,[29] which allows the integration of PLM into more complex workflows. We verify in our work (not shown) that the default setting $\gamma = 1$ yields an appropriate community resolution with the highest modularity.[27,30] The 1000 partitions are particularly generated in search of a broad range of possible solutions since the inherent nondeterminism due to thread parallelism can easily change solutions, especially for those algorithms relying on stochastic search strategies (see Fig. 7 in Appendix A as an example).[23] In (A2), we choose $n_p \in \{25, 50, 100\}$ candidates of the highest modularity as input for reaching a consensus. In (A4), the optimal θ is situation-dependent.[31]

Because this threshold parameter also implies the probability of two nodes being settled into the same community, we vary it in our work and let $\theta \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$. In (A6), considering multiple partitions to be obtained under different values of n_p and θ, the final decision is made based on the modularity score again.

2. Mutual correspondences

Comparing different communities based on single similarity values does neither provide information on how they differ nor any spatial information on the robustness of the identified interdependence patterns of EREs. That is why we employ mutual correspondences,[24] which provide insights on how different partitions agree with each other on a meta-community level (i.e., a group of communities). Suppose that there are two stable partitions $C = \{c_1, \ldots, c_{|C|}\}$ and $C' = \{c'_1, \ldots, c'_{|C'|}\}$ obtained, for example, from two datasets. Given $\varnothing \neq S \subsetneq C$ and $\varnothing \neq S' \subsetneq C'$, a correspondence (S, S') is mutual when the following conditions are met: (a) $|\cup_{i=1}^{|S|} S_i \cap s'| > \frac{1}{2}|s'|$ for any $s' \in S'$, (b) $|\cup_{i=1}^{|S|} S_i \cap s'| \leq \frac{1}{2}|s'|$ for any $s' \in C' \setminus S'$, (c) $|\cup_{i=1}^{|S'|} S'_i \cap s| > \frac{1}{2}|s|$ for any $s \in S$, and (d) $|\cup_{i=1}^{|S'|} S'_i \cap s| \leq \frac{1}{2}|s|$ for any $s \in C \setminus S$. The four conditions above guarantee that any community s from S preserves more than 50% similarity to the counterpart s' in S' and vice versa. We explain below the detailed procedure in finding such a mutual correspondence:

(B1) Initialize S with S_t and $\varnothing \neq S_t \subsetneq C$. Only for the first iteration, S_t is chosen based on domain knowledge, such as well-known teleconnection patterns and is preserved by $S_{init} = S_t$ for the final check in (B5); otherwise, S equals S_t, which is found in (B3).

(B2) Each element $c' \in C'$ satisfying $|\cup_{i=1}^{|S|} S_i \cap c'| > \frac{1}{2}|c'|$, where $S = S_t$ as in (B1), is placed into S'_t, with the correspondence direction of $S_t \to S'_t$.

(B3) Each element $c \in C$ satisfying $|\cup_{i=1}^{|S'|} S'_i \cap c| > \frac{1}{2}|c|$, where S' equals S'_t found in (B2), is placed into S_t, with the reversed correspondence direction of $S'_t \to S_t$.

(B4) If $(S_t \cup S'_t)_{S_t \to S'_t}$ given in (B2) equals $(S'_t \cup S_t)_{S'_t \to S_t}$ given in (B3), then stop; otherwise, go back to (B1) for the next iteration.

(B5) If the updated S_t and S'_t still satisfy $\varnothing \neq S_t \subsetneq C$, $\varnothing \neq S'_t \subsetneq C'$ and $S_{init} \subseteq S_t$ when the iteration process stops in (B4), then a mutual correspondence $(S = S_t, S' = S'_t)$ is established.

It is also possible to start by initializing S' in (B1). The resulting correspondence relationship is $(S' = S'_t, S = S_t)$, which is equal to $(S = S_t, S' = S'_t)$ obtained by initializing S in (B1).

Together with consensus clustering, mutual correspondences lead to a robust top-down clustering workflow. Both combined allow a reliable identification of a synchronization structure in the context of global monsoon. A mutual correspondence can be one-to-many or many-to-one to indicate the split or combination of rainfall patterns over different datasets. For a many-to-many correspondence, it is reasonable to only consider neighborhood communities due to spatial coherence. Finally, we estimate the meta-community level

similarity based on the relative correspondence rate, defined as $RCR(S, S') = \frac{|S \cap S'|}{|S \cup S'|}$.

a. Identification of synchronized days for meta-communities. After identifying meta-communities of good mutual correspondence, we further determine synchronized days for them. Assume that $(t_{i,\alpha}, t_{j,\beta})$ is one pair of synchronized events when $ES_{i,j}$ is significant. It is, therefore, plausible to take $\min\{t_{i,\alpha}, t_{j,\beta}\}$ as the day to indicate the occurrence of synchronization. By repeating this process for all synchronized pairs between i and j and for all other significant links within a meta-community, we obtain the distribution of these days over the JJA season. Such a distribution indicates the temporal characteristic of a given meta-community.

3. Significance test for link bundles

During the (re)construction process of functional climate networks, each link is determined statistically by multiple comparisons in Step (2).[22] A large number of comparisons can lead to some links being preserved by chance. Therefore, we need to correct this problem afterward and to complement the proposed clustering workflow. The basic idea is to identify some regions of higher link density formed by physical mechanisms. Specifically, given a region and all links coming from this region, a spherical Gaussian kernel density estimate (KDE) provides the link density estimation for each grid point regarding the distribution of the original regional link configuration. By randomly distributing those links 1000 times to grid points with at least three events, a corresponding null-model distribution based on the same KDE is obtained for each grid point. The thresholds are then taken in the same way as in Step (2). We also use the 99.9th percentile to determine statistically significant link bundles ($P < 0.001$).

The combination of Steps (3.1), (3.2), and (3.3) gives the regional significant synchronization structure of EREs. Specifically, the regional interdependence by Steps (3.1) and (3.2) takes the form of meta-communities. For a specified region "A" within a meta-community S (S'), we determine all other regions that are significantly synchronized with "A" based on Step (3.3). Some of these regions fall into S (S'), forming the final significantly synchronized regional structure associated with "A." Such a structure could indicate rainfall teleconnections, which can further be verified by atmospheric conditions.

C. Atmospheric condition analysis

This part is the verification for the identified synchronization structure. Therefore, it is not indicated in the workflow (see Fig. 1). Specifically, to investigate composite anomalies, we first identify days of high synchronization. Suppose that we need to find out the specific days when two regions "A" and "B" are highly synchronized. We first determine significant links between them, with direction from "A" ("B") to "B" ("A"), under a significance level of $P < 0.005$. Then, for each time step, we identify the number of associated EREs occurring in "B" ("A") within $\tau_{max} = 10$ days, when there are events in "A" ("B"), based on these links. All these event numbers form a timeseries from "A" ("B") to "B" ("A"), with the same time steps as

datasets. They are then both filtered by a Butterworth low-pass filter with a cutoff frequency of 10 days. We take those time steps as days of high synchronization, where there are local maxima above the 90th percentile of JJA season in the filtered timeseries. With these days, we calculate the composite anomalies regarding the JJA climatology.

III. RESULTS

A. Data

For the reconstruction of functional climate networks regarding EREs, we choose observational data due to the bias of precipitation estimation in reanalysis products.[8] Two satellite-derived datasets, i.e., TRMM 3B42 V7[32] and GPM IMERG[33] from NASA, are analyzed here. We utilize the daily rainfall sums within the JJA season from 2000 to 2019 with the spatial resolution of $1° × 1°$. For a comparative analysis to TRMM, the GPM data are cropped with a boundary from 50°S to 50°N, named as B-GPM.

For the understanding of climatological behavior behind the obtained rainfall patterns, we choose horizontal and vertical wind fields from ECMWF (ERA5)[34] to obtain the corresponding atmospheric conditions. The spatial (i.e., $1° × 1°$) and temporal (i.e., daily estimations ranging from 2000 to 2019) resolutions are consistent with TRMM.

B. Regional interdependence of EREs

The global monsoon represents the integration of all regional monsoons. Different regional monsoons are synchronized since they are fundamentally driven by the annual variation in solar radiation.[3]

In the context of the global monsoon, the following two points need to be further explored: (i) The synchronization structure with higher latitudes and (ii) the spatiotemporal characteristics of intraseasonal variability. We apply the proposed clustering workflow (see Fig. 1) to the global precipitation datasets and give appropriate interpretations of the results.

Two main monsoon-related meta-communities are identified during the JJA season, as shown in Figs. 2 and 3. An overview of the entire community structure is presented in Fig. 8 in Appendix B. The meta-communities describe a global view of synchronized EREs, with a synchronization structure extending to mid-latitude regions [to point (i)] and distinct intraseasonal changes on the frequency of synchronization [to point (ii)]. Temporal distributions indicate two recognizable periods of synchronization: from early June to mid-July [see Figs. 2(b) and 2(d)] and from mid-July to late-August [see Figs. 3(b) and 3(d)]. The former peaks in early June with a sharply decreasing trend thereafter, while the latter shows a gradual increase with a peak in mid-August. Since the temporal distributions are obtained based on the timing of all synchronous EREs within the identified meta-communities, it can be concluded that those teleconnection patterns carried by significant connections within meta-communities are also experiencing a similar temporal variability. We verify this with the Asian summer monsoon (ASM)-related case study in Sec. III C. Regarding the corresponding spatial distributions, they decode the compositions of regional EREs, including different monsoon systems. They are robust under different choices of maximum delays (see Figs. 15–18 in Appendix E) and significance levels (see Figs. 21–24 in Appendix F). Based on regional
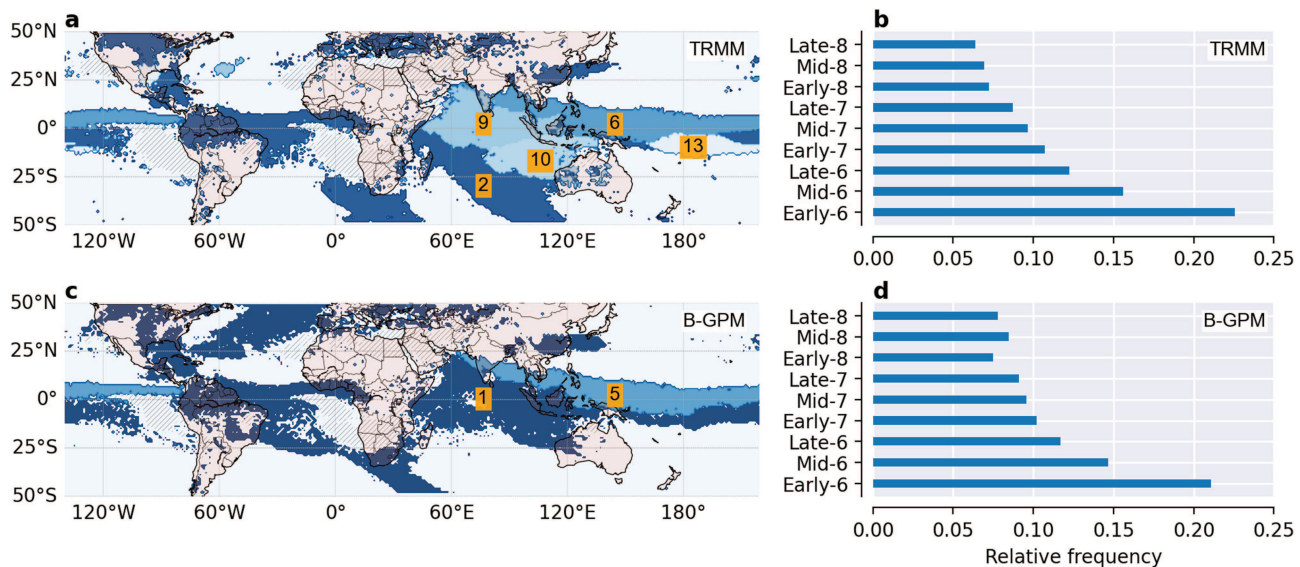

FIG. 2. Communities of mutual correspondence with a similarity of $RCR(TRMM = \{2, 6, 9, 10, 13\}, B − GPM = \{1, 5\}) = 0.54$. (a) Communities 2, 6, 9, 10, and 13 from Fig. 8(a) in Appendix B. (c) Communities 1 and 5 from Fig. 8(b) in Appendix B. (b) and (d) Frequency of synchronized days over JJA, within meta-communities in (a) and (c), respectively (see Sec. II B 2 "Identification of synchronized days for meta-communities"). Grid points with less than three events are excluded when extracting EREs and visualized as hatched areas here.
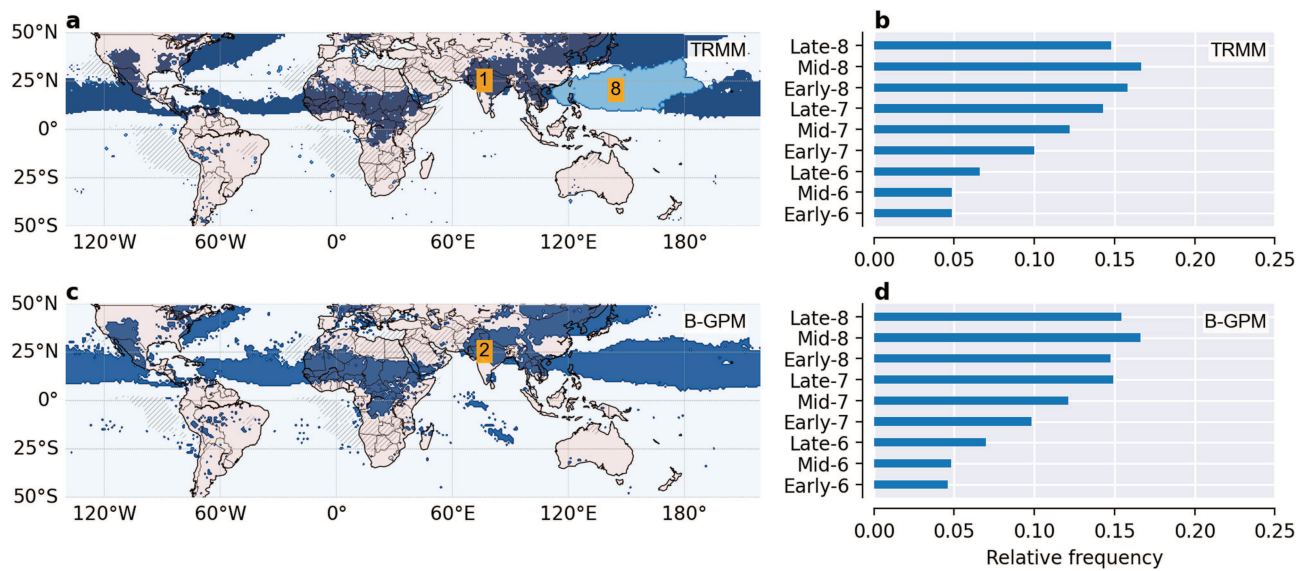
FIG. 3. Same as Fig. 2, but for a different mutual correspondence with a similarity of $RCR(TRMM = \{1, 8\}, B - GPM = \{2\}) = 0.63$ from Fig. 8 in Appendix B.

monsoons over Asia, West Africa, and Southwest United States in boreal summer, we now provide detailed evidence corroborating the spatiotemporal characteristics of intraseasonal variability found in Figs. 2 and 3.

Our results are in line with the known teleconnection pattern inside of ASM. The Meiyu/Baiu rainband,[35] as in Figs. 2(a) and 2(c), is automatically identified by using the proposed clustering workflow. The coexistence of this noticeable region, the Bay of Bengal, and the western coast of India in the same meta-community supports the synchronized initiation (i.e., the "south" teleconnection) of the Indian rainy season and Meiyu/Baiu.[35,36] Figures 3(a) and 3(c) preserve the "north" teleconnection between India and North China, which is built-up after a rapid northward jump of the rain belt.[37] For adjacent oceanic areas, the connection between Indonesian rainfall with Indian and Pacific Oceans[38] is partly revealed by the coexistence of these regions in Fig. 2. In Fig. 3, the northwestward propagation of cyclones in a high phase of the Pacific-Japan pattern establishes the connection of EREs between the western North Pacific tropics and the northern South China Sea; during a low phase, cyclones follow a recurved propagation route to a higher latitude, close to the east of Japan.[39] These cyclones form the synchronization structure relating to the global monsoon, especially over oceanic areas.

Figures 2 and 3 capture the abrupt intraseasonal change in West Africa, i.e., from the oceanic regime [5° North of equator, see Figs. 2(a) and 2(c)] to the continental regime [10° North of equator, see Figs. 3(a) and 3(c)].[40] Between Indian and African monsoons, there exists a connection provided by westward-propagating Rossby waves from the Pacific warm pool in boreal summer.[41] They help modulate the easterly wave activity and moisture transport in West Africa.[42] Meta-communities in Figs. 2(a) and 2(c), covering the Gulf of Guinea with extension to tropical South America, Caribbean Sea, and the eastern tropical Pacific, manifest the

underlying connection via easterly wave disturbances over the equatorial Atlantic basin.[43,44] These disturbances contribute to a large proportion of rainfall in these tropical regions. In Figs. 3(a) and 3(c), meta-communities over West Africa experience a northward shift, which is also known as "monsoon jump."[45,46] Along with this, easterly waves emanating from West Africa are enhanced in August over the tropical North Atlantic.[47] They are the primary precursors of tropical cyclones, which propagate toward the Caribbean Sea and North America.[22,48–50] It has also been proposed that the easterly waves from West Africa have a weak correlation with tropical cyclogenesis in the eastern North Pacific.[51] Such a correlation provides a possible explanation for the synchronized EREs between them [see Figs. 3(a) and 3(c)].

For surrounding regions related to the Southwest United States monsoon, our results catch the mid-summer drought from July to August, especially over Central America, southern Mexico, and part of the Caribbean Sea [see Figs. 3(a) and 3(c)].[52,53] In June, the intensified Caribbean and Great Plains low-level jets together contribute to the moisture supply to rainfall in parts of North America.[54] This leads to the synchronized EREs along the way from the Caribbean Sea to the central United States [see Figs. 2(a) and 2(c)].

Apart from the above-mentioned synchronization related to monsoon, it is also observed that the meta-communities in Figs. 2 and 3 span the mid-latitude belt. The connections between them are mainly via upper-level teleconnection patterns, such as the "Silk Road teleconnection"[55] and the "circumglobal teleconnection"[56,57] in the Northern Hemisphere. For example, the summer rainfall variation of North America stems from East Asian subtropical monsoon heating by the Asia–North America teleconnection,[58] as meta-communities covering both Meiyu/Baiu region and parts of North America in Fig. 2; the positive phase of circumglobal wave train is associated with significantly enhanced precipitation over western

Europe and northwestern India,[56,59] as in Fig. 3. A detailed case study on this is provided in Sec. III C.

C. Regional significant interdependence of EREs: A case study related to ASM

To further verify the intraseasonal change given by the identified meta-communities, as in Figs. 2 and 3, we here present an ASM-related case study on the upper-level atmospheric circulation. Two pairs of relationships are chosen based on significant connections within meta-communities (see Figs. 9 and 10 in Appendix C). A particular reason for the first pair in Figs. 9(e) and 9(f) is that, in the early summer of 2018, rainfall extremes occurred in South-East Europe and South-West Japan nearly simultaneously.[57] Instead of southwestern Japan, we select a large area covering parts of southern China and the Yangtze River Valley since they are all within the Meiyu/Baiu rainband. Meanwhile, it has also been suggested that a strong Indian summer monsoon is often accompanied by significant above normal rainfall over part of western and central Europe.[22,60] We select the second pair based on this as indicated in Figs. 10(e) and 10(f). The composite anomalies, obtained according to the two pairs of relationships, are given in Figs. 4 and 5.

In Figs. 4(b) and 4(d), more synchronizations appear in June followed by a decreasing trend, while Figs. 5(b) and 5(d) present a reversed tendency. There is a recognizable replacement of an eastward upper-level zonal wind movement, as in Figs. 4(a) and 4(c), with a westward direction, as in Figs. 5(a) and 5(c) over Japan, for example. This is consistent with the end of the Meiyu, indicating

the disappearance of the high-altitude westerly jet streams and the appearance of the easterlies over the same region.[35] Additionally, the upper-level meridional wind movement, as shown in Figs. 11 and 12 in Appendix D, confirms the function of Rossby waves in creating synchronized EREs for remote mid-latitude regions.[22,55,56] The identified memberships of two pairs of significant connections to different meta-communities are in agreement with the temporal order and upper-level atmospheric conditions between them. This case study also suggests that the global synchronization structures in Figs. 2 and 3 resemble the distribution of rainfall anomalies (above normal) since only the occurrences of EREs are considered.

IV. DISCUSSION OF ROBUSTNESS

The here proposed clustering workflow enables us to identify regions following similar dynamical properties in an automatic and reliable way. In Sec. III, we obtain two primary global synchronization structures characterizing the intraseasonal changes by using both TRMM and B-GPM data. Therefore, this section is to show whether these two structures are robust concerning the choice of the parameters. The parameters are divided into two groups based on their usages: (i) the network reconstruction step needs an event threshold ($p = 90$th by IPCC), the significance level, and the maximum delay (τ_{max}) and (ii) the clustering workflow needs γ, n_p, and θ.

The parameters in group (i) determine the structure of the constructed networks, which serve as input for the clustering analysis.
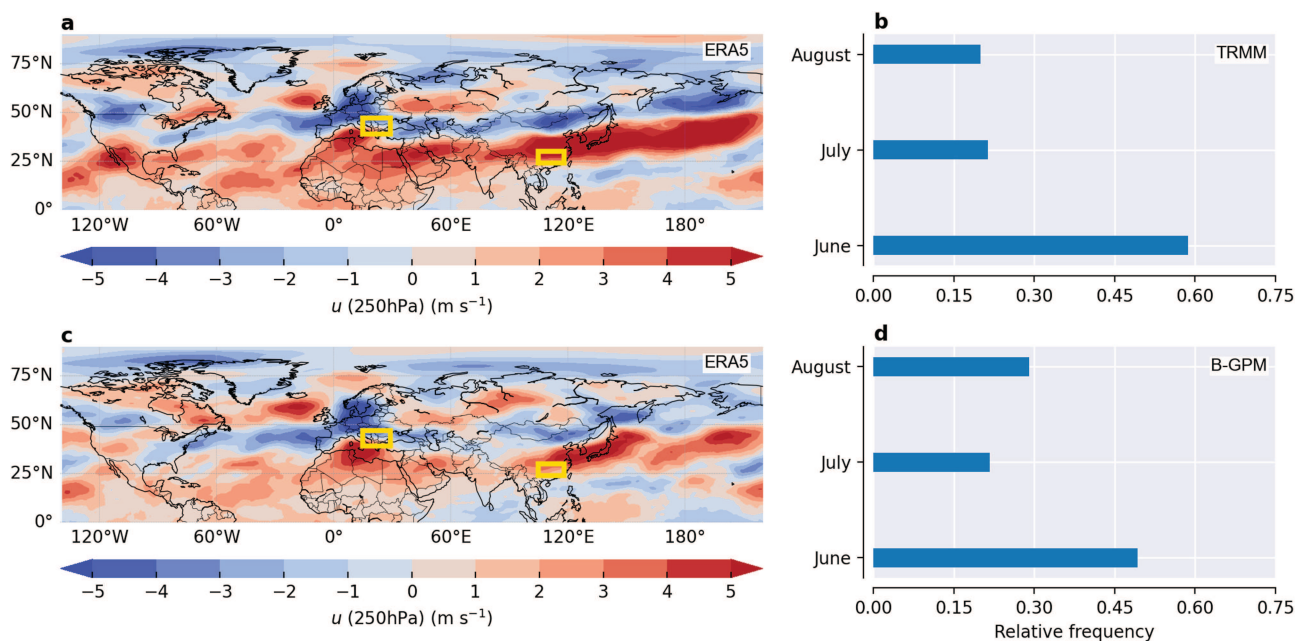

FIG. 4. Atmospheric conditions for the teleconnection pattern between South-East Europe and South China, marked with yellow boxes. (a) and (c) Composite anomalies of 250 hPa zonal wind component u, with respect to the JJA climatology, based on days when the two regions are highly synchronized (see Sec. II C). For (a) and (c), the days are obtained based on TRMM and B-GPM, respectively. (b) and (d) Frequency of these days over JJA.
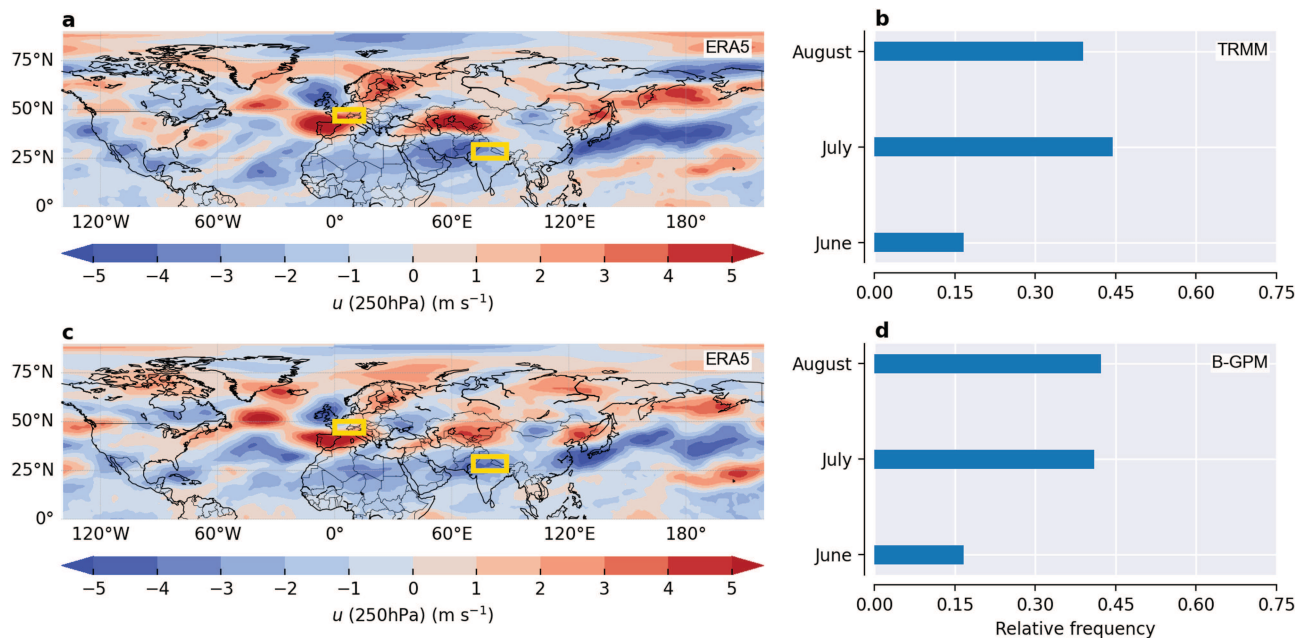
FIG. 5. Same as Fig. 4, but between western and central Europe and northern India.

We have no prior knowledge of the underlying network structure. Therefore, different choices for the maximum delay (τ_{max}) and the significance level should be first considered. The parameters in group (ii), although the underlying community structure is also unknown, we can still use modularity as the measure to quantify the quality of the community structure. That is, one chooses the community structure with the highest modularity.[26] Therefore, all parameters in group (ii) are chosen according to the highest modularity solution; thus, we only discuss τ_{max} and the significance level in the remainder of this section. It should also be noted that the robustness analysis in this section is different from the significance test, which is already considered in the reconstruction step of the functional climate networks in Sec. II A.

A. Robustness over τ_{max}

Different choices of maximum delays with $\tau_{max} \in \{3, 30\}$ days are considered, as in Figs. 13 and 14 in Appendix E. When τ_{max} decreases (increases) to 3 (30) days, the obtained monsoon-related spatial distribution shrinks (extends) to relatively local (wide) areas (see Figs. 15–18 in Appendix E). Indeed, a shorter delay indicates that only those closely synchronized event pairs in time are counted. These pairs are more likely to be found in neighborhood regions. The assumption of a short delay might be too strong to capture a long-distance relationship established by the physical atmospheric process. However, it is also probable that such a relationship appears as the result of a chain of mechanisms with indirect transitivity. This can be revealed from the comparisons between Figs. 15 and 17, or Figs. 16 and 18 since the spatial pattern under the delay of 3 days

is preserved under that of 30 days. That is, the identified intraseasonal synchronization structures are stable inside the interval $\tau_{max} \in [3, 30]$.

B. Robustness over a significance level

The significance level determines the topological structure of the reconstructed functional climate network. A particularly low threshold may eliminate some important synchronization relationships,[22] while a higher value amplifies the multiple-comparison bias. We, therefore, consider the conservative choices of $P < 0.006$ and $P < 0.004$ (see Figs. 19 and 20 in Appendix F). The monsoon-related spatial distributions, shown in Figs. 21–24 in Appendix F, remain consistent with Figs. 2 and 3. By means of the comparison between Figs. 8 and 19, or Figs. 20 and 19, lowering the significance threshold yields a slight increase in modularity. Such a signal implies the community structure as an intrinsic attribute of the underlying self-organized climate system.

In spite of the robustness of the identified two synchronization structures, it is still worth noting that several regions remain dissimilar to a certain degree. For example, the meta-community in Fig. 2(c) covers a larger area of the North Atlantic than that in Fig. 2(a). We attribute this to the uncertainties from rainfall data since fluctuations caused by the network-based clustering approach are corrected in our proposed workflow. Ideally, when two precipitation datasets are exactly the same, the reconstructed network structures are identical and the resulting spatial distributions can be perfectly matched. However, satellite observations contain non-negligible random errors and biases during the generation of datasets due to factors, such as inadequate sampling and deficiencies

in algorithms.[61] These uncertainties are further retained, leading to discrepancies in the identified synchronization structures. However, the discrepancy information can also be important to measure the extent to which different datasets reproduce the same climate phenomenon.

V. CONCLUSION

In summary, our work starts from the identification of regions of similar climatological behavior and focuses on revealing the global view of the interdependence patterns for EREs in the context of the global monsoon. For this, we provide a network-based clustering workflow that combines consensus clustering and mutual correspondences. By means of this clustering workflow, we identify two main monsoon-related synchronization structures. They have independent temporal and spatial characteristics. The first structure, primarily from early June to mid-July, shows early summer features over different monsoon systems, including the synchronized initiation of the Indian rainy season and Meiyu/Baiu over ASM, the oceanic regime over West Africa, and early-rainy season over the Caribbean Sea toward the central United States. The second structure, from mid-July to late-August, manifests the states after the "monsoon jump," including the establishment of the "north" teleconnection inside the ASM between Indian summer monsoon and the rainfall in North China and the continental regime over West Africa. Meanwhile, the mid-summer drought is dominating Central America and southern Mexico. Also, we show that the significant connections inside the identified meta-communities capture the synchronized behavior via the "Silk Road teleconnection"[55] and the "circumglobal teleconnection"[56,57] for the mid-latitude belt.

The successful identification of intraseasonal spatiotemporal variability essentially stems from the combination of ES and complex-network theory to represent the underlying climate system as networks. As an extension of this research, there are several potentially relevant directions. First, how interaction processes between different monsoons form at different levels is crucial for the physical understanding of the identified global monsoon structure. Second, this clustering workflow can be applied to other seasonal analyses, such as the boreal winter, which brings marked monsoonal characteristics to the Southern Hemisphere. Third, the application on higher spatial resolution scenarios (e.g, $0.5° \times 0.5°$) is a computational problem worthy of further consideration. Fourth, the combined application of event synchronization and community detection to other types of data, in particular, in neuroscience, such as EEG signals and magnetic resonance imaging (MRI), is another promising direction.

AUTHOR DECLARATIONS
Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

The data that support the findings of this study are available within the article.

APPENDIX A: THE MOTIVATION FOR THE NETWORK-BASED WORKFLOW

Illustration of the limitation of the significant test for link bundles under TRMM (Fig. 6). Illustration of fluctuations of clustering approaches, such as PLM, under (a) TRMM and (b) B-GPM, respectively (Fig. 7).
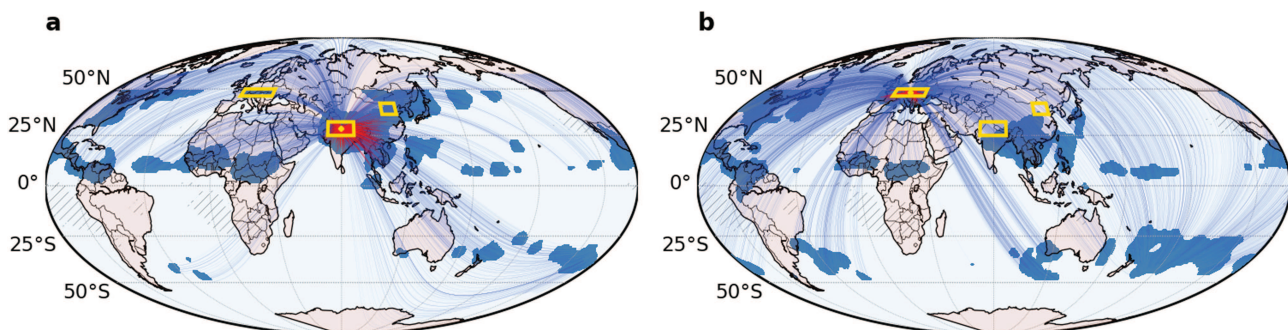

FIG. 6. Illustration of the limitation of the significant test for link bundles under TRMM. Three yellow boxes are positioned over Europe, northern India, and northern China, respectively. (a) When the box in northern India is selected as the central region, associated significant link bundles connect to the entire boxes in both Europe and northern China. (b) When the central region is changed to the box in Europe, only a few link bundles connect to the boxes in northern India and northern China. This indicates that the significant link bundles are limited by the link distribution of a given region and, therefore, can only reveal a partial view of regions that follow similar climatological behavior.
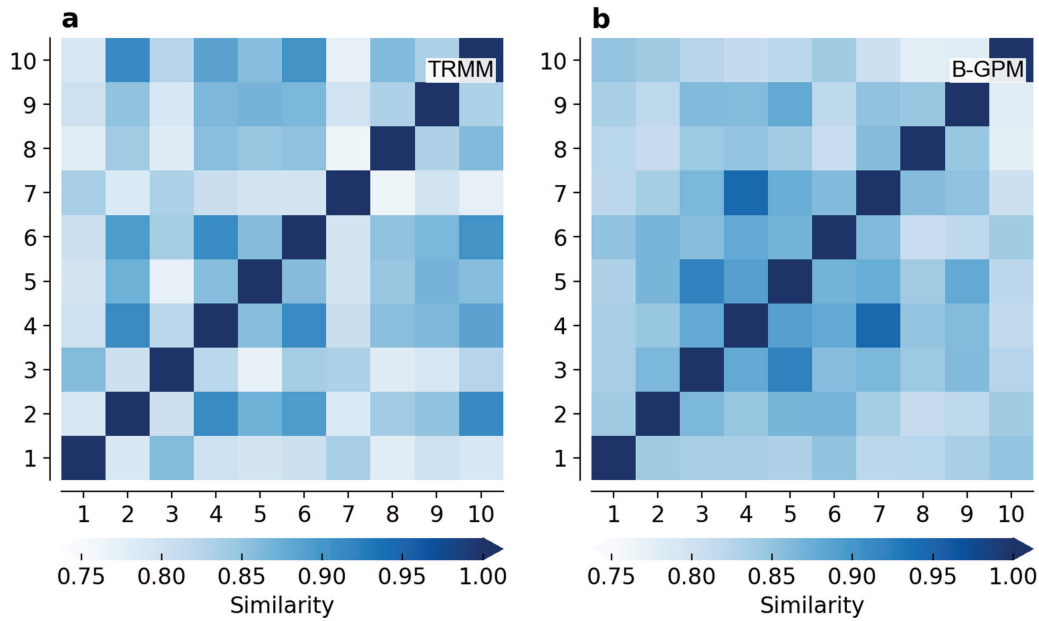
FIG. 7. Illustration of fluctuations of clustering approaches, such as PLM, under (a) TRMM and (b) B-GPM, respectively. Taking TRMM as an example, the first 10 partitions of the highest modularity are selected from 1000 randomly generated partitions. Any pair between them should show a similarity of 1, based on the normalized mutual information score, if the clustering process was deterministic. However, (a) and (b) show the average deviations of 0.15 and 0.14, respectively, indicating a certain degree of differences between solutions.
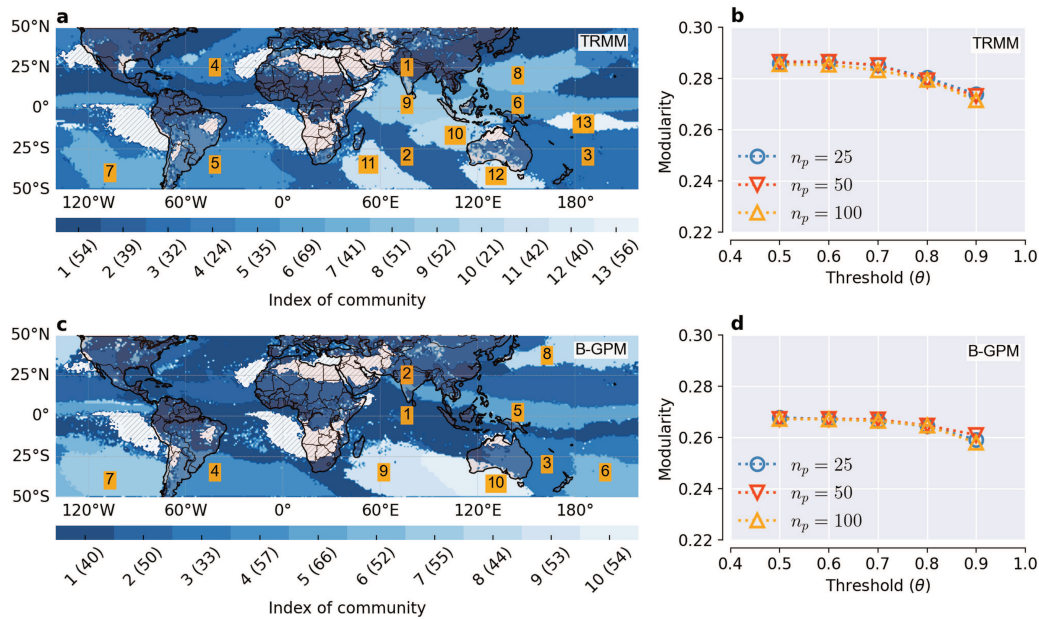
FIG. 8. Community structures for TRMM and B-GPM, respectively. (b) and (d) Modularity as a function of n_p and τ for the selection of community structure for TRMM and B-GPM. (a) and (c) share a similarity of 0.58 based on the normalized mutual information score. They are chosen according to (b) with $n_p = 25$ and $\theta = 0.6$ and (d) with $n_p = 25$ and $\theta = 0.5$, respectively, to indicate the maximum modularity after consensus clustering. Both are sorted by the size of each community, namely, the number of grid points inside a community. Each index is attached with the average number of EREs in the corresponding community. Communities smaller than $0.01*|V|$ are excluded from the visualization. Grid points with less than three events are also excluded when extracting EREs and visualized as hatched areas.
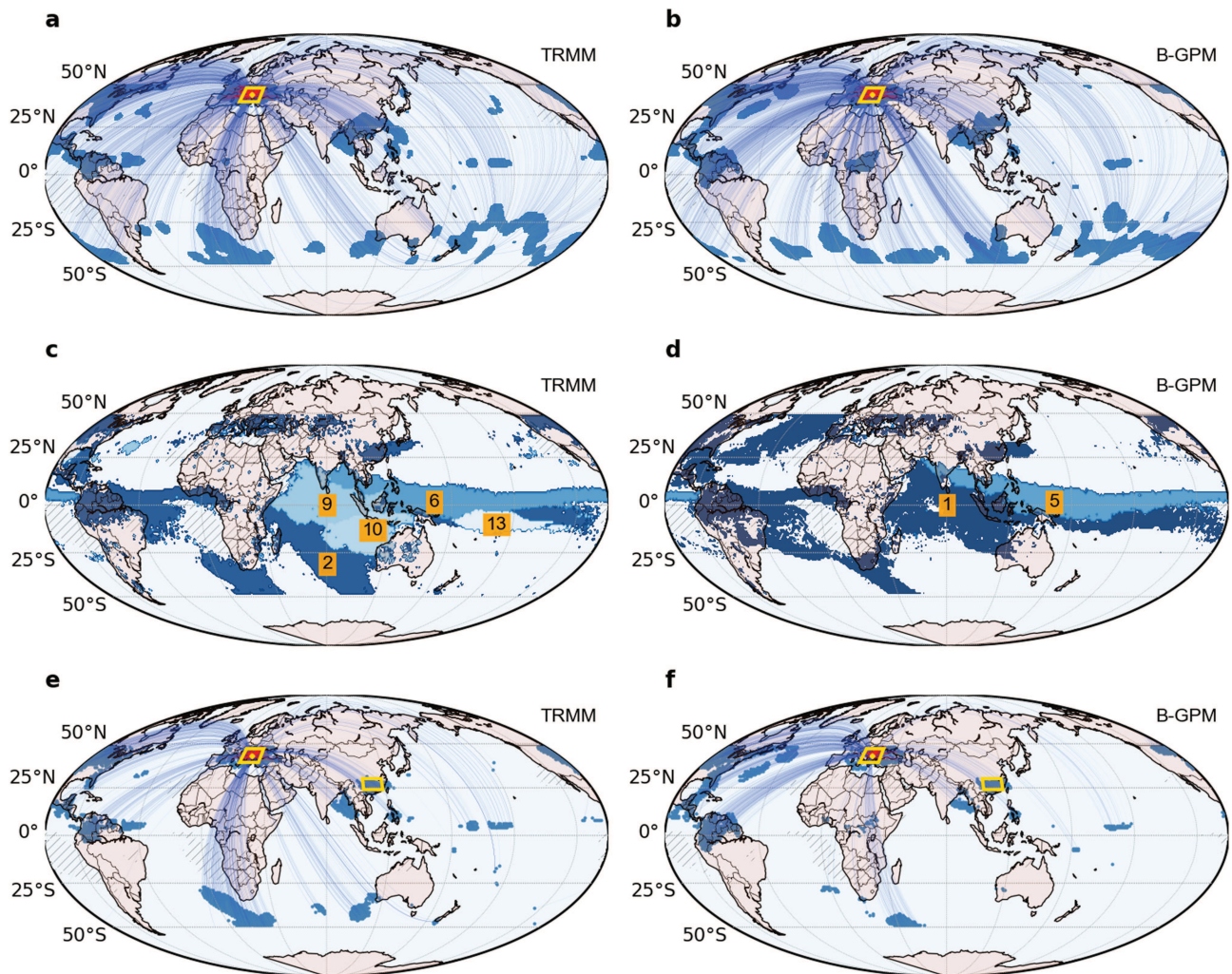
FIG. 9. Teleconnection pattern given for South-East Europe for TRMM [(a), (c), and (e)] and B-GPM [(b), (d), and (f)]. (a) and (b) Significant link bundles as given in Ref. 22 but for South-East Europe (see Sec. II B 3). (c) and (d) Meta-communities extracted from Fig. 8 and the same as Fig. 2 (see Sec. II B 1 and II B 2). (e) and (f) Regional significant interdependence based on the intersections of (a) and (c) and (b) and (d), respectively. Yellow boxes positioned in (e) and (f) represent the presence of significant connections between South-East Europe (39°N–47°N, 15°E–29°E) and South China (24°N–30°N, 105°E–118°E) to be considered in Sec. III C.

APPENDIX B: COMMUNITY STRUCTURE UNDER THE DELAY OF 10 DAYS

Community structures for TRMM and B-GPM, respectively (Fig. 8).

APPENDIX C: THE PROCESS TO DETERMINE REGIONAL SIGNIFICANT INTERDEPENDENCE FOR THE ASM-RELATED CASE STUDY

Teleconnection pattern given for South-East Europe for TRMM [(a), (c), and (e)] and B-GPM [(b), (d), and (f)] (Fig. 9).

Teleconnection pattern given for northern India for TRMM [(a), (c), and (e)] and B-GPM [(b), (d), and (f)] (Fig. 10).

APPENDIX D: THE MERIDIONAL COMPONENT FOR THE ASM-RELATED CASE STUDY

Atmospheric conditions for the teleconnection pattern between South-East Europe and South China, marked with yellow boxes (Fig. 11). Same as Fig. 11, but between western and central Europe and northern India (Fig. 12).
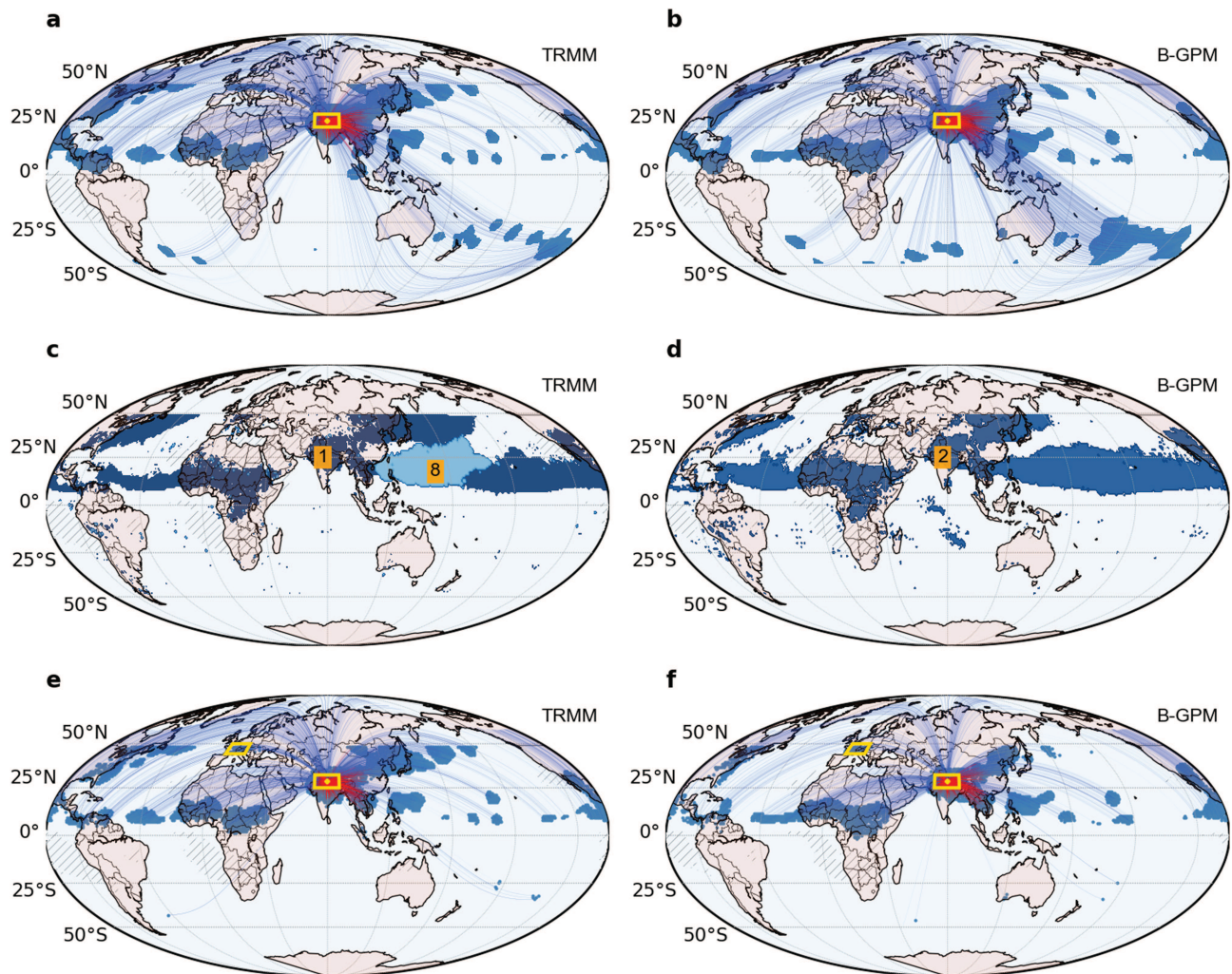
FIG. 10. Teleconnection pattern given for northern India for TRMM [(a), (c), and (e)] and B-GPM [(b), (d), and (f)]. (a) and (b) Significant link bundles as given in Ref. 22 (see Sec. II B 3). (c) and (d) Meta-communities extracted from Fig. 8 and the same as Fig. 3 (see Secs. II B 1 and II B 2). (e) and (f) Regional significant interdependence based on the intersections of (a) and (c) and (b) and (d), respectively. Yellow boxes positioned in (e) and (f) represent the presence of significant connections between northern India (25°N–32°N, 71°E–88°E) and part of western and central Europe (44°N–50°N, 0°E–15°E), to be considered in Sec. III C.

APPENDIX E: THE ROBUSTNESS TEST ON DIFFERENT DELAYS

Same as Fig. 8, but for a different $\tau_{max} = 3$ days (Fig. 13). Same as Fig. 8, but for a different $\tau_{max} = 30$ days (Fig. 14). Same as Fig. 2, but for a different $\tau_{max} = 3$ days (Fig. 15). Same as Fig. 3, but for a different $\tau_{max} = 3$ days (Fig. 16). Same as Fig. 2, but for a different $\tau_{max} = 30$ days (Fig. 17). Same as Fig. 3, but for a different $\tau_{max} = 30$ days (Fig. 18).

APPENDIX F: THE ROBUSTNESS TEST ON DIFFERENT SIGNIFICANCE LEVELS

Same as Fig. 8, but for a different significance level $P < 0.006$ (Fig. 19). Same as Fig. 8, but for a different significance level $P < 0.004$ (Fig. 20). Same as Fig. 2, but for a different significance level $P < 0.006$ (Fig. 21). Same as Fig. 3, but for a different significance level $P < 0.006$ (Fig. 22). Same as Fig. 2, but for a different significance level $P < 0.004$ (Fig. 23).
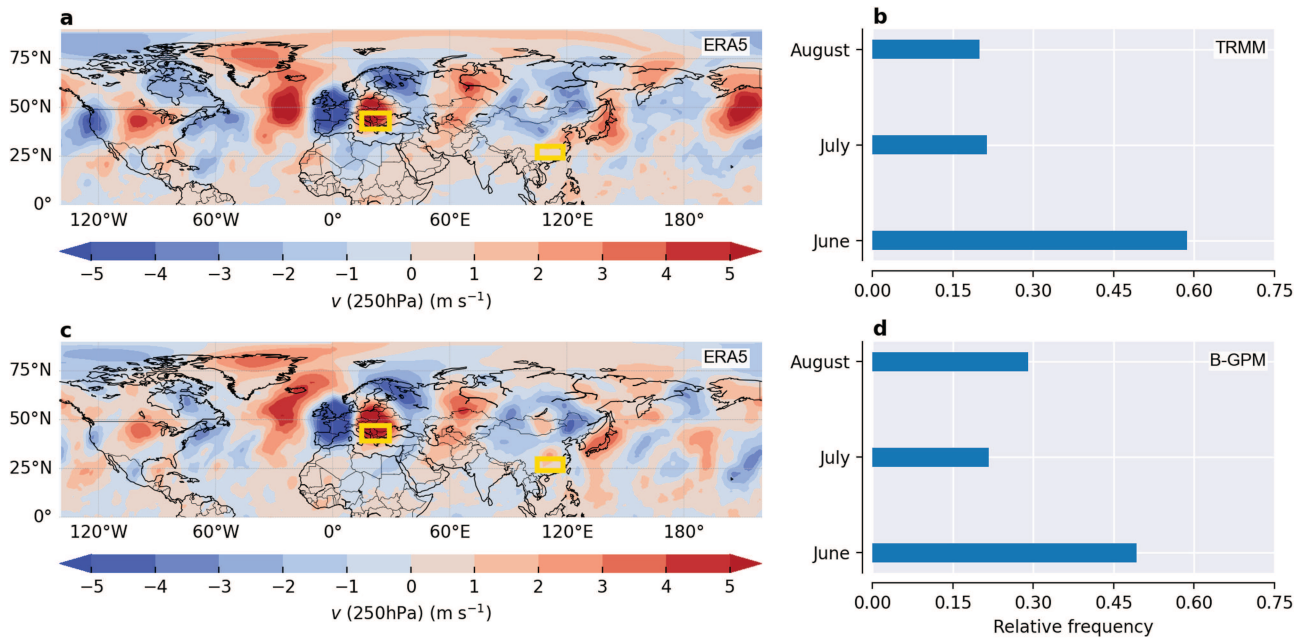
FIG. 11. Atmospheric conditions for the teleconnection pattern between South-East Europe and South China, marked with yellow boxes. (a) and (c) Composite anomalies of 250 hPa meridional wind component v, with respect to the JJA climatology, based on days when the two regions are highly synchronized (see Sec. II C). For (a) and (c), the days are obtained based on TRMM and B-GPM, respectively. (b) and (d) Frequency of these days over JJA.
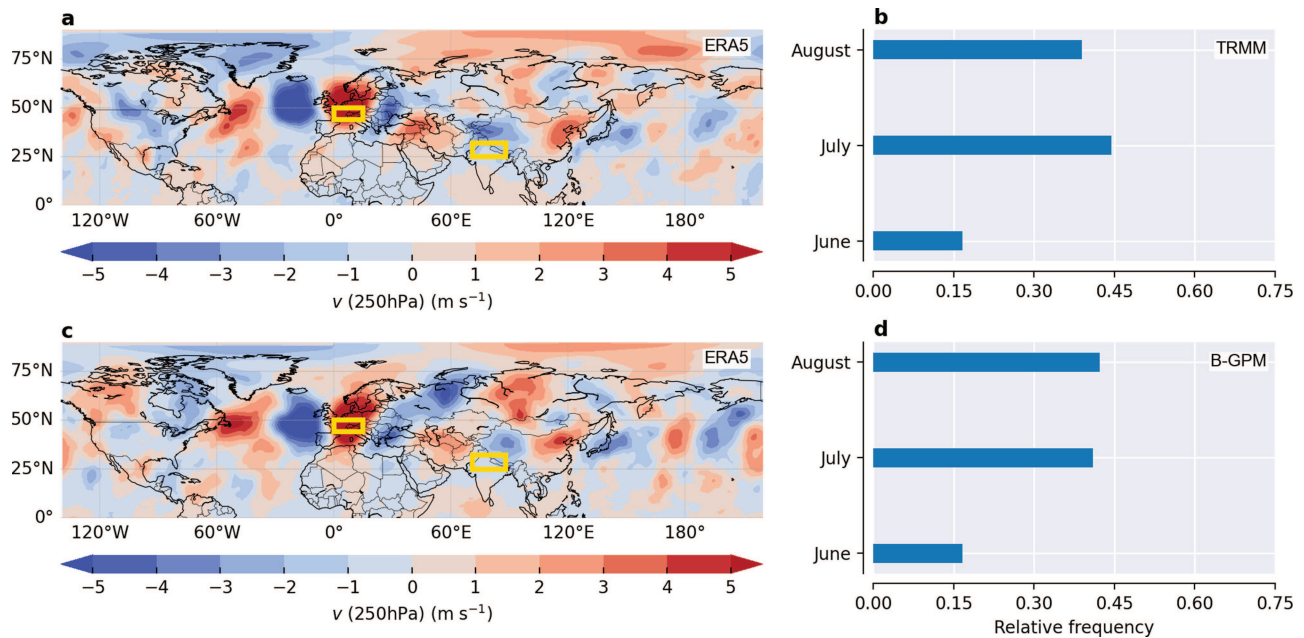

FIG. 12. Same as Fig. 11, but between western and central Europe and northern India.
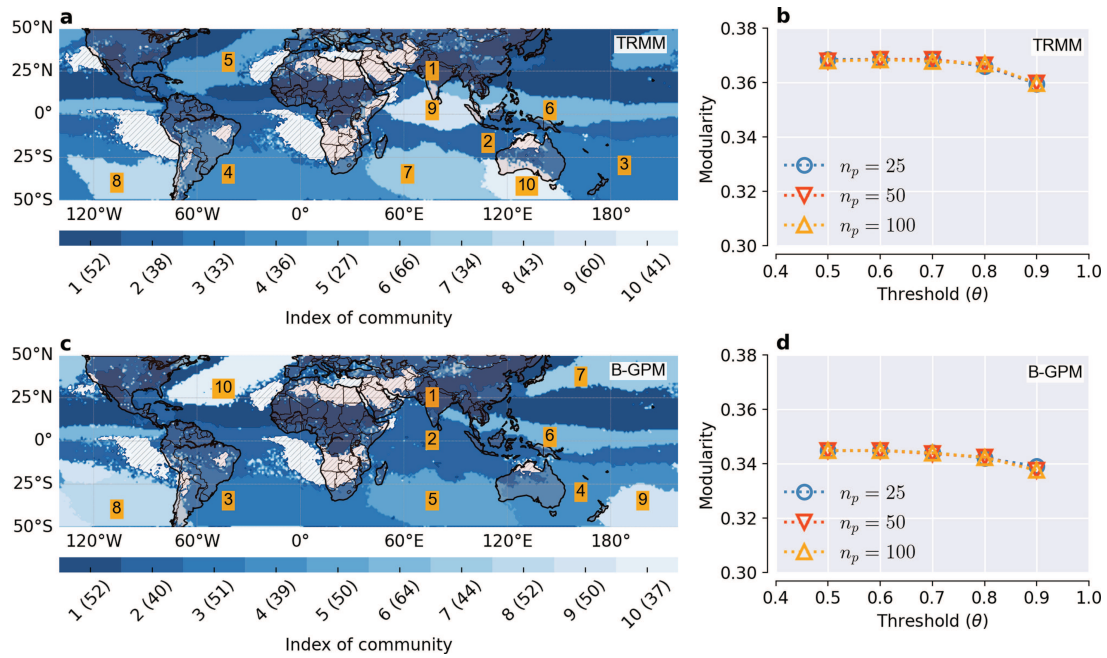
FIG. 13. Same as Fig. 8, but for a different $\tau_{max} = 3$ days. In (a) and (c), community structures for TRMM and B-GPM are chosen according to (b) with $n_p = 25$ and $\theta = 0.5$ and (d) with $n_p = 50$ and $\theta = 0.6$, respectively, to indicate the highest modularity value. They share a similarity of 0.66 based on the normalized mutual information score.
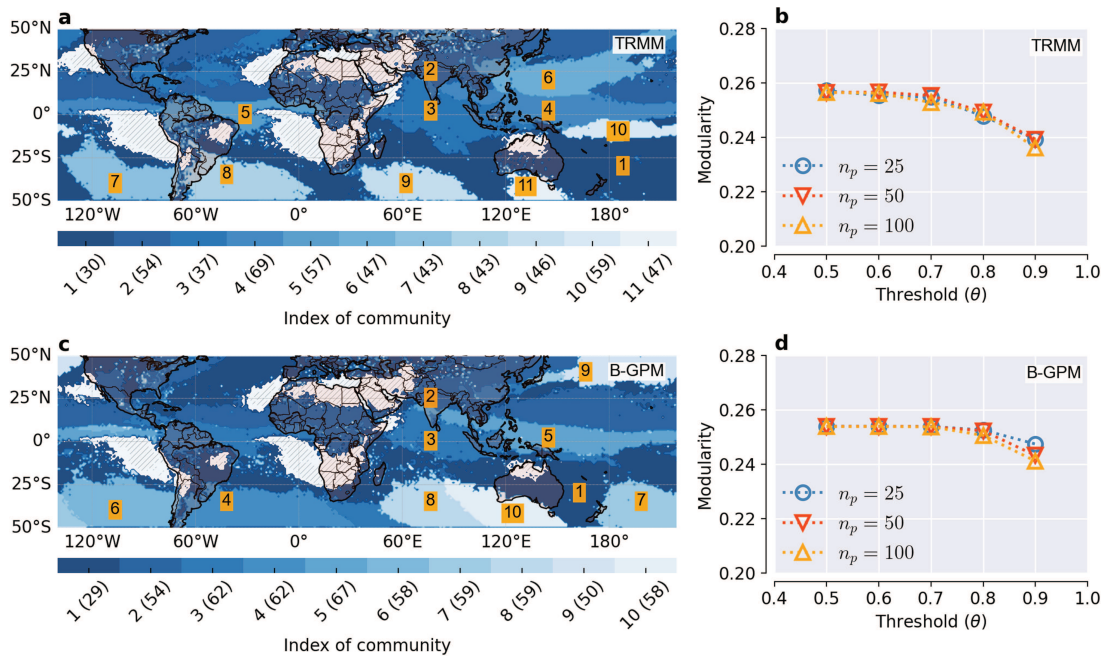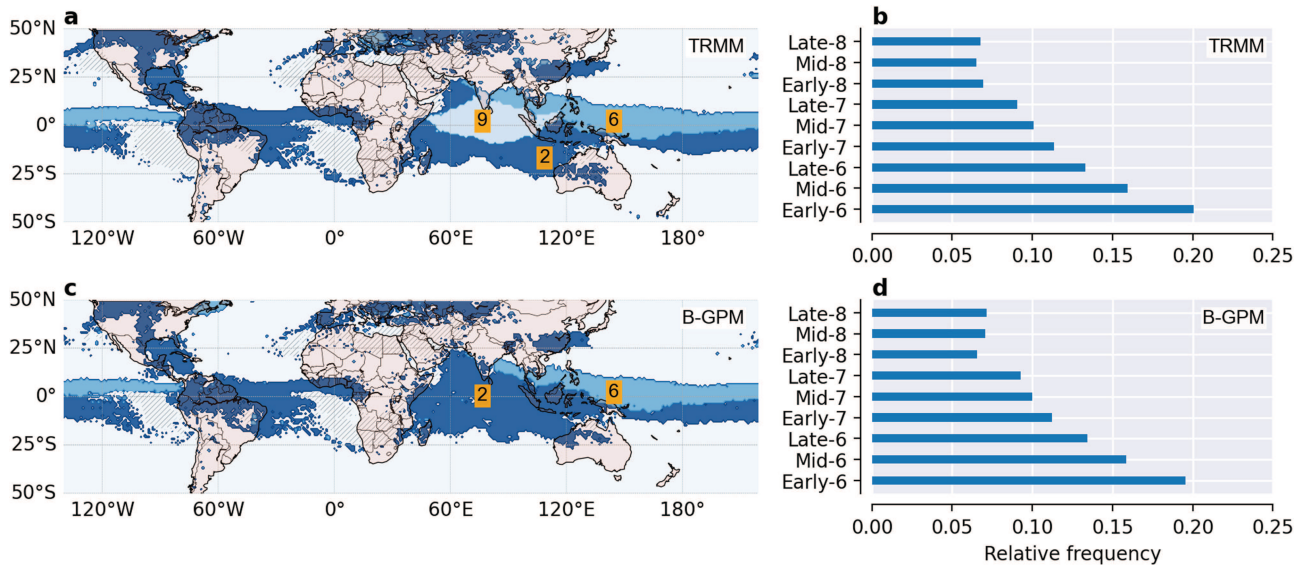

FIG. 14. Same as Fig. 8, but for a different $\tau_{max} = 30$ days. Community structures for (a) TRMM and (c) B-GPM are chosen according to (b) with $n_p = 25$ and $\theta = 0.5$ and (d) with $n_p = 25$ and $\theta = 0.5$, respectively, sharing a similarity of 0.60.

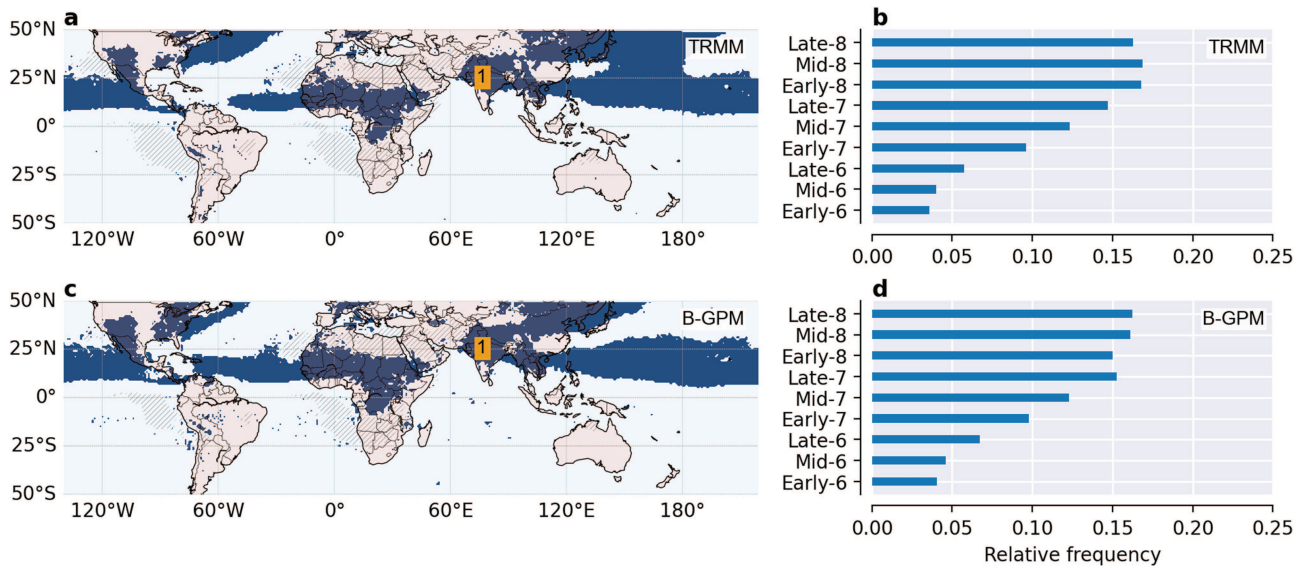**FIG. 15.** Same as Fig. 2, but for a different $\tau_{max} = 3$ days. Communities of mutual correspondence with a similarity of $RCR(TRMM = \{2, 6, 9\}, B - GPM = \{2, 6\}) = 0.72$. (a) Communities 2, 6, and 9 from Fig. 13(a). (c) Communities 2 and 6 from Fig. 13(b). (b) and (d) Frequency of synchronized days within meta-communities in (a) and (c), respectively (see Sec. II B 2 "Identification of synchronized days for meta-communities"). Grid points with less than three events are excluded when extracting EREs and visualized as hatched areas here.

FIG. 16. Same as Fig. 3, but for a different $\tau_{max} = 3$ days. Communities of mutual correspondence with a similarity of $RCR(TRMM = \{1\}, B - GPM = \{1\}) = 0.65$ from Fig. 13.
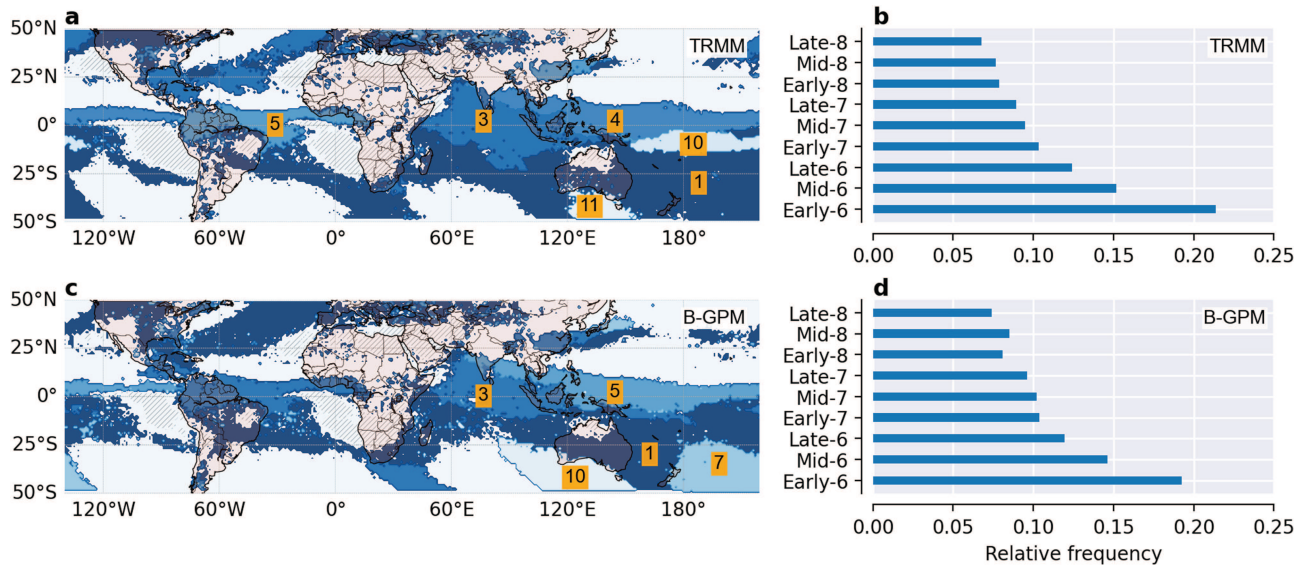
FIG. 17. Same as Fig. 2, but for a different $\tau_{max} = 30$ days. Communities of mutual correspondence with a similarity of $RCR(TRMM = \{1, 3, 4, 5, 10, 11\}$, $B - GPM = \{1, 3, 5, 7, 10\}) = 0.72$. (a) Communities 1, 3, 4, 5, 10, and 11 from Fig. 14(a). (c) Communities 1, 3, 5, 7, and 10 from Fig. 14(b). (b) and (d) Frequency of synchronized days within meta-communities in (a) and (c), respectively (see Sec. II B 2 "Identification of synchronized days for meta-communities"). Grid points with less than three events are excluded when extracting EREs and visualized as hatched areas here.
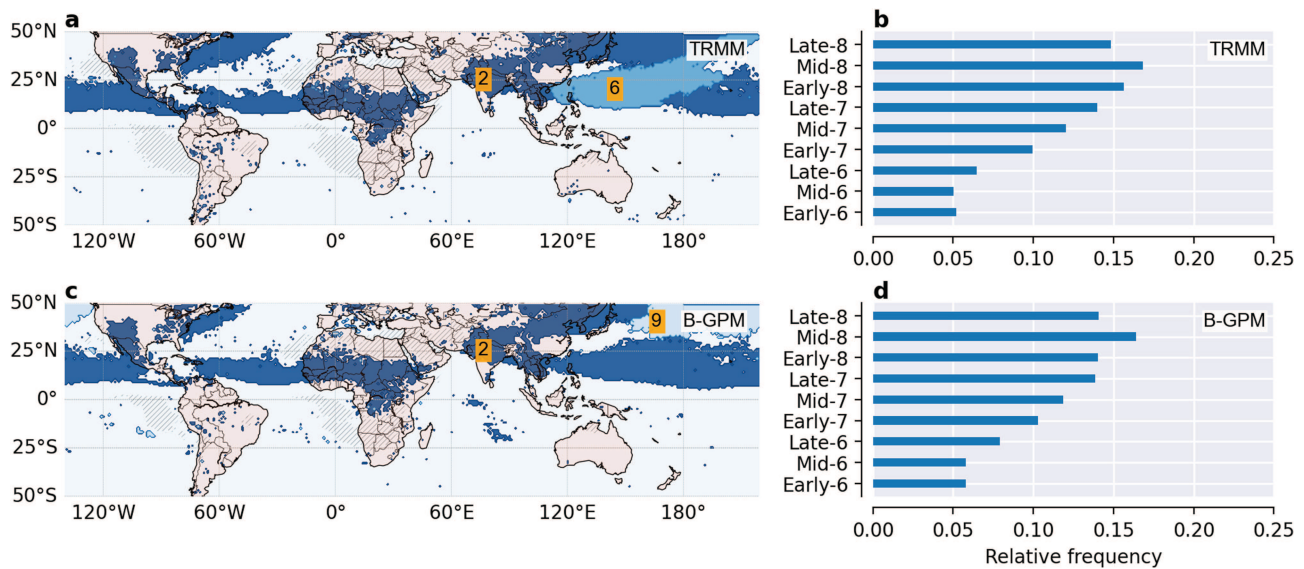

FIG. 18. Same as Fig. 3, but for a different $\tau_{max} = 30$ days. Communities of mutual correspondence with a similarity of $RCR(TRMM=\{2, 6\}$, B-GPM=\{2, 9\}) = 0.67$ from Fig. 14.
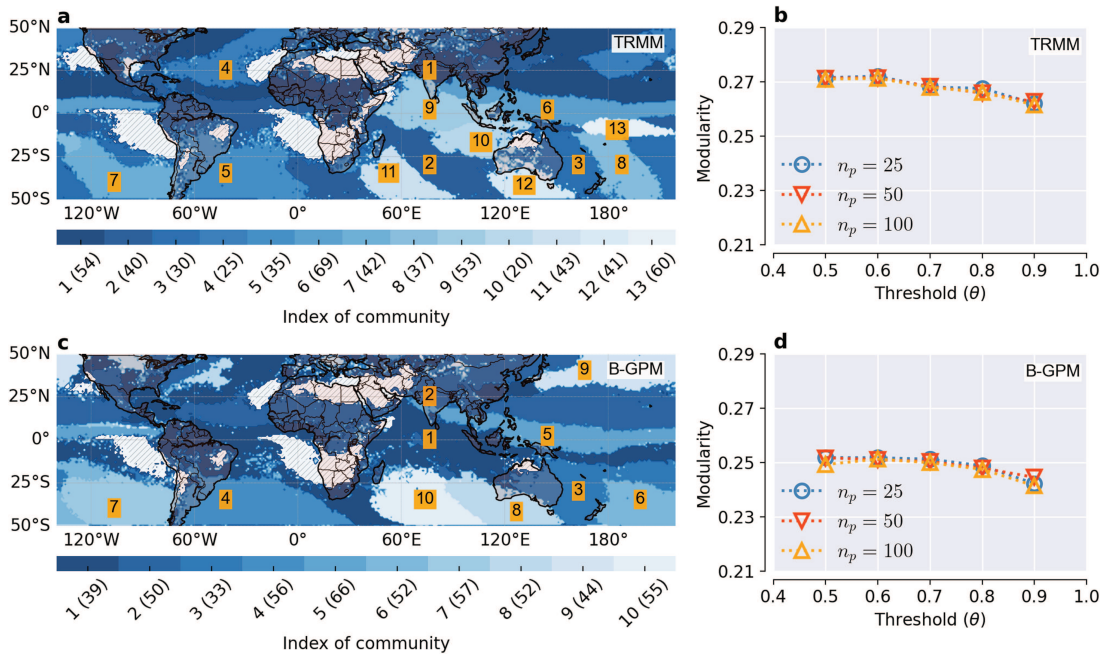
FIG. 19. Same as Fig. 8, but for a different significance level $P < 0.006$. In (a) and (c), community structures for TRMM and B-GPM are chosen according to (b) with $n_p = 25$ and $\theta = 0.6$ and (d) with $n_p = 25$ and $\theta = 0.5$, respectively, to indicate the largest modularity value. They share a similarity of 0.57 based on the normalized mutual information score.
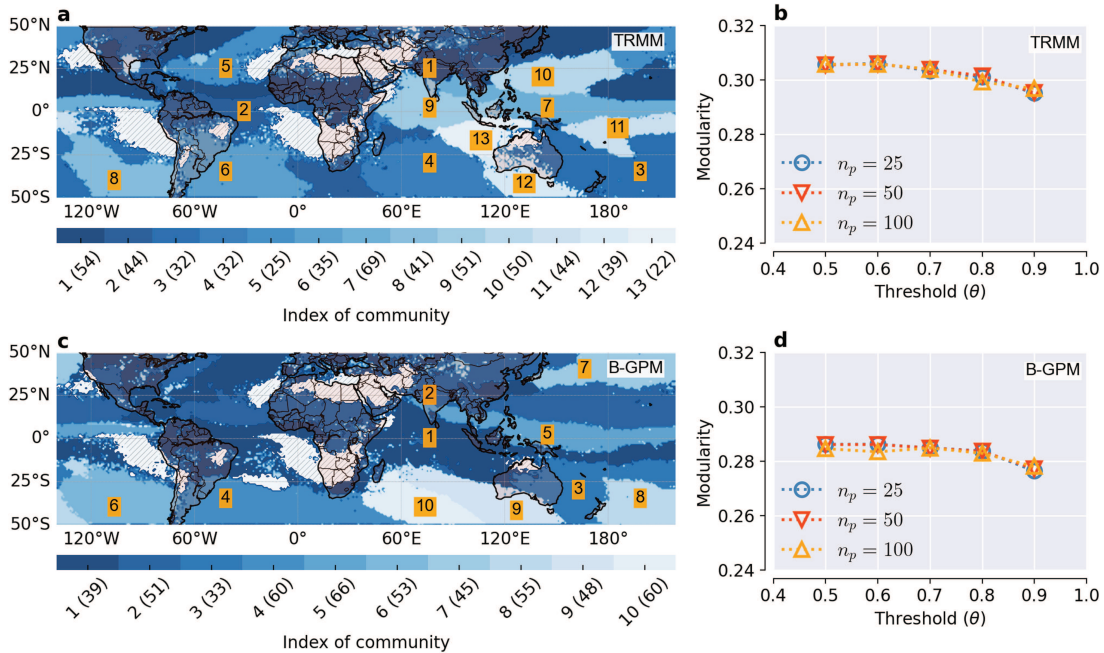

FIG. 20. Same as Fig. 8, but for a different significance level $P < 0.004$. Community structures for (a) TRMM and (c) B-GPM are chosen according to (b) with $n_p = 25$ and $\theta = 0.6$ and (d) with $n_p = 25$ and $\theta = 0.6$, respectively, sharing a similarity of 0.59.
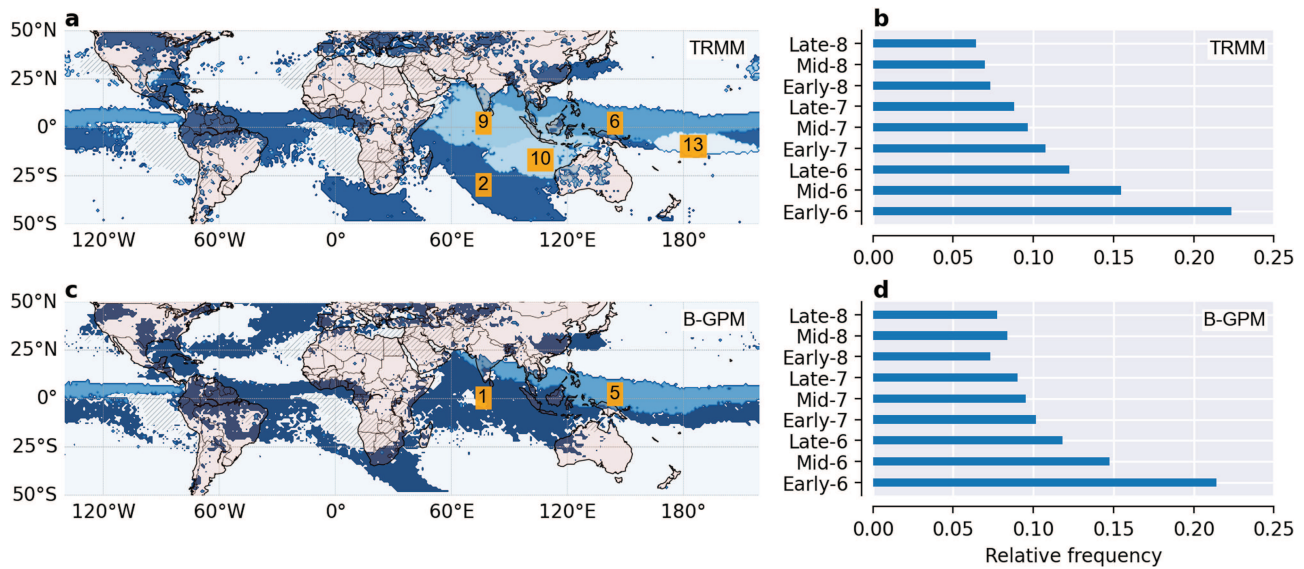
FIG. 21. Same as Fig. 2, but for a different significance level $P < 0.006$. Communities of mutual correspondence with a similarity of $RCR(TRMM = \{2, 6, 9, 10, 13\}, B - GPM = \{1, 5\}) = 0.53$. (a) Communities 2, 6, 9, 10, and 13 from Fig. 19(a). (c) Communities 1 and 5 from Fig. 19(b). (b) and (d) Frequency of synchronized days within meta-communities in (a) and (c), respectively (see Sec. II B 2 "Identification of synchronized days for meta-communities"). Grid points with less than three events are excluded when extracting EREs and visualized as hatched areas here.
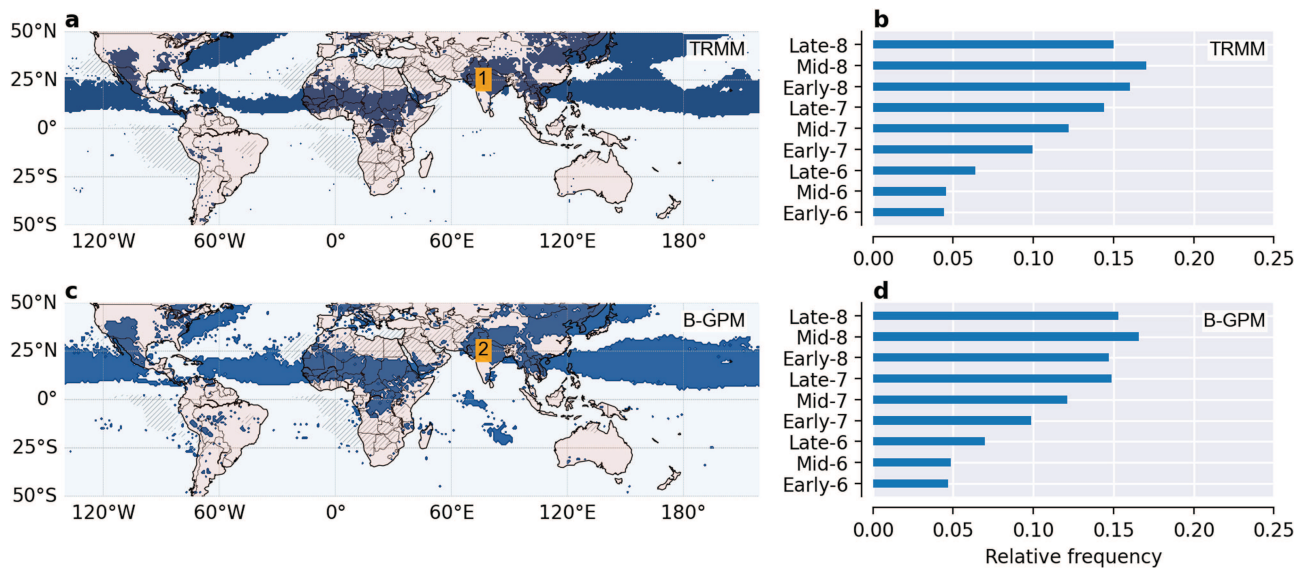

FIG. 22. Same as Fig. 3, but for a different significance level $P < 0.006$. Communities of mutual correspondence with a similarity of $RCR(TRMM = \{1\}, B - GPM = \{2\}) = 0.61$ from Fig. 19.
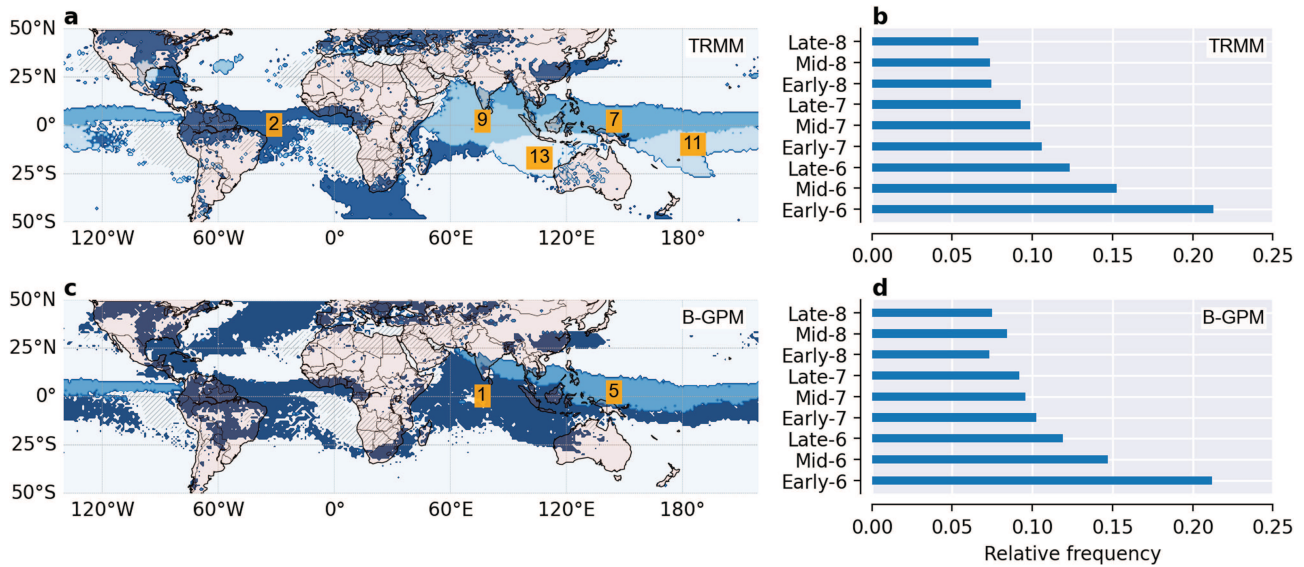
FIG. 23. Same as Fig. 2, but for a different significance level $P < 0.004$. Communities of mutual correspondence with a similarity of $RCR(TRMM = \{2, 7, 9, 11, 13\}$, $B - GPM = \{1, 5\}) = 0.57$. (a) Communities 2, 7, 9, 11, and 13 from Fig. 20(a). (c) Communities 1 and 5 from Fig. 20(b). (b) and (d) Frequency of synchronized days within meta-communities in (a) and (c), respectively (see Sec. II B 2 "Identification of synchronized days for meta-communities"). Grid points with less than three events are excluded when extracting EREs and visualized as hatched areas here.
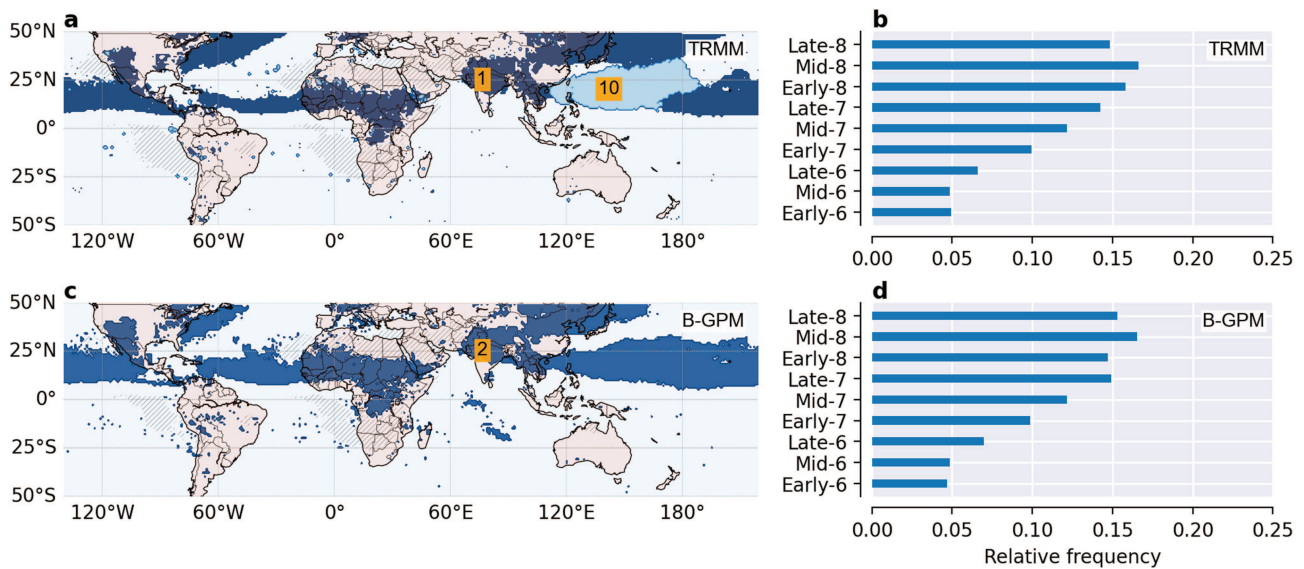

FIG. 24. Same as Fig. 3, but for a different significance level $P < 0.004$. Communities of mutual correspondence with a similarity of $RCR(TRMM = \{1, 10\}$, $B - GPM = \{2\}) = 0.64$ from Fig. 20.

Same as Fig. 3, but for a different significance level $P < 0.004$ (Fig. 24).

REFERENCES

[1] B. Wang, J. Liu, H.-J. Kim, P. J. Webster, and S.-Y. Yim, Clim. Dyn. **39**, 1123 (2012).

[2] B. Wang and Q. Ding, Dyn. Atmos. Oceans **44**, 165 (2008).

[3] A. Zhisheng, W. Guoxiong, L. Jianping, S. Youbin, L. Yimin, Z. Weijian, C. Yanjun, D. Anmin, L. Li, M. Jiangyu *et al.*, Annu. Rev. Earth Planet. Sci. **43**, 29 (2015).

[4] A. A. Tsonis and P. J. Roebber, Physica A **333**, 497 (2004).

[5] A. A. Tsonis, K. L. Swanson, and P. J. Roebber, Bull. Am. Meteorol. Soc. **87**, 585 (2006).

[6] G. A. Bello, M. Angus, N. Pedemane, J. K. Harlalka, F. H. Semazzi, V. Kumar, and N. F. Samatova, in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (Springer, 2015), pp. 736–751.

[7] M. Saha and P. Mitra, Meteorol. Atmos. Phys. **131**, 1301 (2019).

[8] K. Steinhaeuser, N. V. Chawla, and A. R. Ganguly, in *Proceedings of the Third International Workshop on Knowledge Discovery from Sensor Data* (Association for Computing Machinery, 2009), pp. 23–31.

[9] A. A. Tsonis, G. Wang, K. L. Swanson, F. A. Rodrigues, and L. da Fontura Costa, Clim. Dyn. **37**, 933 (2011).

[10] W. Hendrix, I. K. Tetteh, A. Agrawal, F. Semazzi, W.-K. Liao, and A. Choudhary, in *2011 IEEE 11th International Conference on Data Mining Workshops* (IEEE, 2011), pp. 9–14.

[11] K. Steinhaeuser, N. V. Chawla, and A. R. Ganguly, Stat. Anal. Data Min.: ASA Data Sci. J. **4**, 497 (2011).

[12] K. Steinhaeuser and A. A. Tsonis, Clim. Dyn. **42**, 1665 (2014).

[13] K. Steinhaeuser, A. R. Ganguly, and N. V. Chawla, Clim. Dyn. **39**, 889 (2012).

[14] N. Boers, B. Bookhagen, H. M. Barbosa, N. Marwan, J. Kurths, and J. Marengo, Nat. Commun. **5**, 5199 (2014).

[15] See https://www.ipcc.ch/assessment-report/ar6/ for "Sixth Assessment Report."

[16] U. Ozturk, N. Malik, K. Cheung, N. Marwan, and J. Kurths, Clim. Dyn. **53**, 521 (2019).

[17] N. Malik, B. Bookhagen, N. Marwan, and J. Kurths, Clim. Dyn. **39**, 971 (2012).

[18] R. Q. Quiroga, T. Kreuz, and P. Grassberger, Phys. Rev. E **66**, 041904 (2002).

[19] M. Laib, F. Guignard, M. Kanevski, and L. Telesca, Chaos **29**, 043107 (2019).

[20] J. Kurths, A. Agarwal, R. Shukla, N. Marwan, M. Rathinasamy, L. Caesar, R. Krishnan, and B. Merz, Nonlinear Process. Geophys. **26**, 251 (2019).

[21] K. K. Cheung and U. Ozturk, Chaos **30**, 063117 (2020).

[22] N. Boers, B. Goswami, A. Rheinwalt, B. Bookhagen, B. Hoskins, and J. Kurths, Nature **566**, 373 (2019).

[23] A. Lancichinetti and S. Fortunato, Sci. Rep. **2**, 336 (2012).

[24] R. Glantz and H. Meyerhenke, in *Proceedings of the 2018 SIAM International Conference on Data Mining* (SIAM, 2018), pp. 1–9.

[25] See https://www.ipcc.ch/report/ar5/wg1/ for "Climate Change 2013: The Physical Science Basis."

[26] M. E. Newman and M. Girvan, Phys. Rev. E **69**, 026113 (2004).

[27] C. L. Staudt and H. Meyerhenke, IEEE Trans. Parallel Distrib. Syst. **27**, 171 (2016).

[28] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, J. Stat. Mech.: Theory Exp. **2008**, P10008.

[29] C. L. Staudt, A. Sazonovs, and H. Meyerhenke, Netw. Sci. **4**, 508 (2016).

[30] J. Reichardt and S. Bornholdt, Phys. Rev. Lett. **93**, 218701 (2004).

[31] L. G. Jeub, O. Sporns, and S. Fortunato, Sci. Rep. **8**, 3259 (2018).

[32] G. J. Huffman, D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker, J. Hydrometeorol. **8**, 38 (2007).

[33] G. J. Huffman, D. T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, P. Xie, and S.-H. Yoo, Algorithm Theoretical Basis Document (ATBD) Version 4, 26 (2015).

[34] H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers *et al.*, Q. J. R. Meteorol. Soc. **146**, 1999 (2020).

[35] D. Yihui and J. C. Chan, Meteorol. Atmos. Phys. **89**, 117 (2005).

[36] Y. Liu and Y. Ding, Sci. China Ser. D: Earth Sci. **51**, 1021 (2008).

[37] Y. Y. Liu and Y. H. Ding, Acta Meteor. Sin. **66**, 789 (2008).

[38] A. Kurniadi, E. Weller, S.-K. Min, and M.-G. Seong, Int. J. Climatol. **41**, 3640 (2021).

[39] K.-S. Choi, C.-C. Wu, and E.-J. Cha, J. Geophys. Res.: Atmos. **115**, D19114, https://doi.org/10.1029/2010JD013866 (2010).

[40] T. Lebel, A. Diedhiou, and H. Laurent, J. Geophys. Res.: Atmos. **108**, 77, https://doi.org/10.1029/2001JD001580 (2003).

[41] S. Janicot, F. Mounier, N. M. Hall, S. Leroux, B. Sultan, and G. N. Kiladis, J. Clim. **22**, 1541 (2009).

[42] S. Janicot, F. Mounier, S. Gervois, B. Sultan, and G. N. Kiladis, J. Clim. **23**, 4005 (2010).

[43] H. B. Gomes, T. Ambrizzi, B. F. P. da Silva, K. Hodges, P. L. S. Dias, D. L. Herdies, M. C. L. Silva, and H. B. Gomes, Clim. Dyn. **53**, 1393 (2019).

[44] C. Dominguez, J. M. Done, and C. L. Bruyère, Clim. Dyn. **54**, 191 (2020).

[45] B. Sultan and S. Janicot, Geophys. Res. Lett. **27**, 3353, https://doi.org/10.1029/1999GL011285 (2000).

[46] S. E. Nicholson, Int. Sch. Res. Not. **2013**, 453521 (2013).

[47] A. Diedhiou, S. Janicot, A. Viltard, P. De Felice, and H. Laurent, Clim. Dyn. **15**, 795 (1999).

[48] C. W. Landsea, Mon. Weather Rev. **121**, 1703 (1993).

[49] C. Wang, D. B. Enfield, S.-K. Lee, and C. W. Landsea, J. Clim. **19**, 3011 (2006).

[50] T.-C. Chen, S.-Y. Wang, and A. J. Clark, J. Clim. **21**, 6767 (2008).

[51] L. A. Avila and R. J. Pasch, Mon. Weather Rev. **120**, 2688 (1992).

[52] V. Magaña, J. A. Amador, and S. Medina, J. Clim. **12**, 1577 (1999).

[53] C. Martinez, L. Goddard, Y. Kushnir, and M. Ting, Clim. Dyn. **53**, 825 (2019).

[54] C. Wang, S.-K. Lee, and D. B. Enfield, J. Clim. **20**, 5021 (2007).

[55] T. Enomoto, B. J. Hoskins, and Y. Matsuda, Q. J. R. Meteorol. Soc. **129**, 157 (2003).

[56] Q. Ding and B. Wang, J. Clim. **18**, 3483 (2005).

[57] K. Kornhuber, S. Osprey, D. Coumou, S. Petri, V. Petoukhov, S. Rahmstorf, and L. Gray, Environ. Res. Lett. **14**, 054002 (2019).

[58] Z. Zhu and T. Li, Clim. Dyn. **50**, 3523 (2018).

[59] S. Saeed, N. Van Lipzig, W. A. Müller, F. Saeed, and D. Zanchettin, Clim. Dyn. **43**, 503 (2014).

[60] H. Lin and Z. Wu, Clim. Dyn. **39**, 303 (2012).

[61] Q. Sun, C. Miao, Q. Duan, H. Ashouri, S. Sorooshian, and K.-L. Hsu, Rev. Geophys. **56**, 79, https://doi.org/10.1002/2017RG000574 (2018).