Event synchrony measures for functional climate network analysis: A case study on South American rainfall dynamics

Frederik Wolf,[1,2] Jurek Bauer,[3] Niklas Boers,[1,4,5] and Reik V. Donner[1,6]

[1)] *Potsdam Institute for Climate Impact Research (PIK) – Member of the Leibniz Association, Telegrafenberg A56, 14473 Potsdam, Germany*

[2)] *Department of Physics, Humboldt University Berlin, Newtonstraße 15, 12489 Berlin, Germany*

[3)] *Department of Physics, Georg-August-University, Wilhelmsplatz 1, 37073 Göttingen, Germany*

[4)] *Department of Mathematics and Computer Science, Free University Berlin, Takustraße 9, 14195 Berlin, Germany*

[5)] *Global Systems Institute and Department of Mathematics, University of Exeter, Stocker Rd, Exeter EX4 4PY, United Kingdom*

[6)] *Department of Water, Environment, Construction and Safety, Magdeburg–Stendal University of Applied Sciences, Breitscheidstraße 2, 39114 Magdeburg, Germany*

(Dated: October 29, 2019)

Investigating climate extremes and their synchrony has recently attracted rising attention in the context of ongoing climate change. With enhanced computational capacity data driven methods such as functional climate networks have been proposed and have already contributed to significant advances in understanding and predicting extreme events, as well as identifying interrelations between the occurrences of various climatic phenomena. While the (in its basic setting) parameter free event synchronization (ES) has been widely applied to construct functional climate networks from extreme event series, its original definition has been realized to exhibit problems in handling events occurring at subsequent time steps, which need to be accounted for by a correction scheme. Along with the study of this conceptual limitation of the original ES approach, event coincidence analysis (ECA) has been suggested as an alternative approach, which incorporates an additional parameter for selecting certain timescales of event synchrony. In this work, we systematically compare functional climate network representations of South American heavy precipitation events obtained using ES and ECA without and with the correction for temporal event clustering. We find that both measures exhibit different types of biases, which are thoroughly explained based on the obtained network structures. By combining the complementary information captured by ES and ECA, we revisit the spatiotemporal organization of extreme events during the South American Monsoon season. While the corrected version of ES captures multiple timescales of heavy rainfall cascades at once, ECA allows to systematically disentangle them and thereby to trace the spatiotemporal propagation in greater detail.

The occurrence of extreme events and their dynamics is one central topic of Earth System Science. Such events are often not only a direct danger for people's lives, but can also cause financial damage. This holds particularly true for weather extremes, the intensity and frequency distributions of which can be heavily affected by climate change. In the last years, several studies have been conducted employing the concept of event synchronization (ES) to quantify the statistical association between the occurrences of rainfall extremes at different locations. This information allows constructing functional climate network representations, which can potentially reveal dynamical characteristics of the Earth's climate system that are hidden to more traditional techniques of statistical climatology. However, some recent studies have identified potential caveats of ES when being applied to temporally clustered events, which are a common situation in the context of climate extremes. While some works have introduced a simple correction to account for the resulting bias, the parallel development of event coincidence analysis (ECA) has provided a powerful alternative measure for quantifying event synchrony. In this context, it is important to better understand the conceptual benefits and limitations of both methods to put forward the appropriate interpretation of the results of event based functional climate network analyses. This work aims to provide the corresponding details, based on the reconsideration of the spatiotemporal organization of heavy rainfall over South America from the perspective of functional climate networks based upon either ES or ECA. Our results help in understanding the differences of the resulting network structures and, thus, potential issues with the climatological interpretation of previous studies.

I. INTRODUCTION

In the last two decades, computational power and data availability in many fields of science have grown by various orders of magnitude. Among the manifold approaches which have been proposed to cope with the sheer amount of data, complex networks have attracted increasing attention, providing a powerful tool to study different kinds of systems[1–4].

Simultaneously, the analysis of extreme events has gained importance in the scope of Earth System Science

due to their changing characteristics under ongoing climate change[5–7]. In recent years, several studies have effectively incorporated measures to quantify synchrony of such extreme events and revealed a multitude of related spatiotemporal patterns. Among the variety of synchrony measures, event synchronization (ES)[8,9] stands out as the most frequently applied method to construct functional climate networks based on the synchrony of extreme events[10–13] and has already successfully contributed to the prediction and classification of extreme event cascades[11,14,15].

Recent work has pointed out a conceptual limitation of ES to properly capture serially dependent events[16–18] and has suggested event coincidence analysis (ECA)[19,20] as an alternative method to quantify synchrony. Although the corresponding clustering bias of ES had already been mentioned in some earlier studies on event based functional climate networks and can be addressed using a simple algorithmic modification[11,15,21], several questions remain, especially regarding the differences between uncorrected and corrected ES, possible influences of the correction scheme on functional networks constructed utilizing ECA, and distinct information provided by functional climate networks based on the uncorrected and corrected versions of both, ES and ECA, respectively.

In this work, we systematically compare ES and ECA based functional climate networks without and with correction by reanalyzing two examples of spatiotemporal dynamics of heavy precipitation over South America along with the South American Monsoon System (SAMS)[21–25]. First, we investigate the spatial patterns of extreme event cascades over the South American continent. Here, we apply both, ES as well as ECA without and with the algorithmic correction to emphasize the differences among the resulting networks' degree fields, and provide explanations of the observed differences by drawing upon the recently introduced pairing coefficient[16,17]. Second, we specifically revisit extreme rainfall cascades propagating from South East South America (SESA) towards the Eastern Central Andes[11,14]. We demonstrate the possibility of assessing different time scales of the underlying climatological mechanisms by varying the intrinsic algorithmic parameter of ECA (the global coincidence window ΔT), while the corrected version of the parameter free ES provides a time scale integrated picture of the whole process.

The remainder of this paper is organized as follows: First, we review the contemporary definitions of ECA and ES along with the aforementioned correction for temporal event clustering. Second, we describe the data set used in this paper and elaborate on the different approaches to constructing functional climate networks. Third, we introduce the atmospheric setting of the SAMS and compare the resulting degree fields of functional climate networks after the application of the different event synchrony measures. Finally, we analyze rainfall cascades by measuring the cross degree between different regional subnetworks.

II. METHODOLOGY

In this section, we first introduce some basic concepts of network theory to the extent employed in this work. Then, we describe the concepts of ES and ECA and summarize their correct mathematical formulation as recently pointed out in refs.[16,17]. Subsequently, we discuss the motivation and implications of the correction of ES as employed in some previous studies[11,15].

A. Complex Networks

We investigate the topology of a complex network with N nodes and E edges by analyzing the symmetric, binary adjacency matrix $\mathbf{A}$. Its coefficients a_{ij} take values of 1 if node i is connected with node j, and 0 otherwise.

The *degree* k_i of node i characterizes the number of connections of node i to the rest of the network and is thus defined as

$$k_i = \sum_{j=1}^{N} a_{ij}. \tag{1}$$

In the context of functional climate networks, where nodes represent areas of the globe with potentially different sizes, network properties like the degree need to be properly corrected. For this purpose, we consider so-called nodesplitting invariant (n.s.i.) characteristics[26], which are obtained from the traditional ones by introducing proper node weights that reflect the areal share of each node. In the context of the degree, this leads to

$$k_i = \sum_{j=1}^{N} w_j a_{ij}, \tag{2}$$

where w_j represents the n.s.i. weight of node j (for a regular latitude-longitude grid as used in the present work, w_j corresponds to the cosine of the latitudinal position of j). In the remainder of this paper, we will exclusively consider the n.s.i. degree, but will omit this specification to simplify all discussions.

In addition to the full degree, the last part of this work will also consider the total *cross degree*[13,27] between different subnetworks characterizing distinct spatial regions, which is defined as

$$K^{m,n} = \sum_{x \in R_m} \sum_{y \in R_n} a_{xy} \tag{3}$$

with R_m and R_n specifying the regions between which the cross degree is computed. Since the differences in latitude are relatively small in the considered region, we will neglect the relevance of differently sized nodes for the calculation of the cross degree.

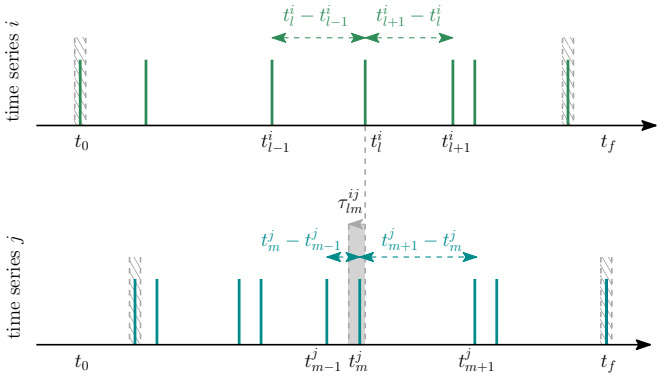

Fig. 1: Schematic sketch of the definition of event proximity used in the estimation of ES.

B. Event Synchronization

Event Synchronization (ES), first introduced by Quiroga et al.[8], is a parameter free method to quantify the synchrony of events, which has been originally developed for the analysis of spike train synchrony in EEG recordings, but has later also been applied for the construction of functional climate networks. Here, two events l and m in time series from observations made at node i and j at time t_l^i and t_m^j are considered *synchronized* if and only if they have occurred closer to each other than the *local (dynamical) coincidence interval*[18]

$$\tau_{lm}^{ij} = \frac{1}{2}\min\left\{t_{l+1}^i - t_l^i, t_l^i - t_{l-1}^i, t_{m+1}^j - t_m^j, t_m^j - t_{m-1}^j\right\}. \tag{4}$$

As we compare the temporal distances between the actual and the preceding and subsequent events, we need to exclude the last and the first event of each series in all further calculations[16,17]. Assuming that time series i and j have s_i and s_j events, respectively, we therefore set $l = 2, 3, ..., s_i - 1$ and $m = 2, 3, ..., s_j - 1$.

With these prerequisites, we next define the synchronization condition

$$\sigma_{lm}^{ij} = \begin{cases} 1, & \text{if } 0 < t_l^i - t_m^j \leq \tau_{lm}^{ij}, \\ 0, & \text{otherwise,} \end{cases} \tag{5}$$

depending on the temporal distance between the events and the local coincidence interval which, in principle, can be limited by an additional parameter τ_{max}, to avoid overly large coincidence intervals. The basic idea is depicted in Fig. 1. To compute the total number of "synchronized" events

$$c(i|j) = \sum_{l=2}^{s_i-1}\sum_{m=2}^{s_j-1} J_{lm}^{ij}, \tag{6}$$

implying that an event in time series j *precedes* an event in time series i, we sum over all coincidences captured by the following indicator function[16,17]:

$$J_{lm}^{ij} = \begin{cases} 1, & \text{if } \sigma_{lm}^{ij} = 1, \sigma_{m,l-1}^{ji} = 0 \text{ and } \sigma_{m+1,l}^{ji} = 0, \\ \frac{1}{2}, & \text{if either } t_l^i = t_m^j \\ & \text{or } \sigma_{lm}^{ij} = 1 \text{ and } (\sigma_{m,l-1}^{ji} = 1 \text{ or } \sigma_{m+1,l}^{ji} = 1), \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

Here, we have to use this rather unhandy version as we otherwise might end up double counting the same event pairs as synchronized in both, $c(i|j)$ and $c(j|i)$, where $c(j|i)$ denotes the total number of synchronized events, where an event in time series i precedes an event in time series j.

Finally, we calculate the *event synchronization strength* (hereafter shortly termed "event synchronization" or ES for brevity) as

$$Q_{ij}^{ES} = \frac{c\,(i|j) + c\,(j|i)}{\sqrt{(s_i - 2)\,(s_j - 2)}}. \tag{8}$$

The resulting symmetric matrix $\boldsymbol{Q}^{ES} = \left(Q_{ij}^{ES}\right)$ can be used to construct a functional climate network, as will be further detailed in Section III.

From a conceptual perspective, we note that in real world time series, events are often far from homogeneously distributed. In this regard, the use of a local coincidence interval τ_{lm}^{ij} in the computation of ES covers different time scales with one measure only. On one hand, this makes ES potentially effective in extracting meaningful patterns from climate data, where delays between events usually vary. On the other hand, mixing information from different time scales may blur the identifiability of certain climatological processes. Specifically, in cases where the waiting time distribution between subsequent events becomes very heterogeneous, depending on τ_{max}, the local coincidence intervals cover very different tolerable delays. Moreover, as we will further discuss in the course of this paper, temporal clustering of events may lead to a systematic bias of the estimated ES strength. This may impact the topological properties of functional network representations of spatiotemporal data obtained based on ES as a statistical association measure[16–18].

C. Event Coincidence Analysis

As a potential alternative to ES, ECA follows a generally similar rationale with the exception that the time interval used for distinguishing whether or not two events are synchronized is set as a global parameter – the *global (static) coincidence interval* ΔT – instead of being locally chosen in a fully data adaptive way.

Similar to ES, we again start by comparing events at times t_l^i and t_m^j. Two events are considered to be synchronized if they occur within the global coincidence interval, i.e., $0 < t_l^i - t_m^j < \Delta T$ (time symmetrized versions may be easily defined, but shall not be further discussed

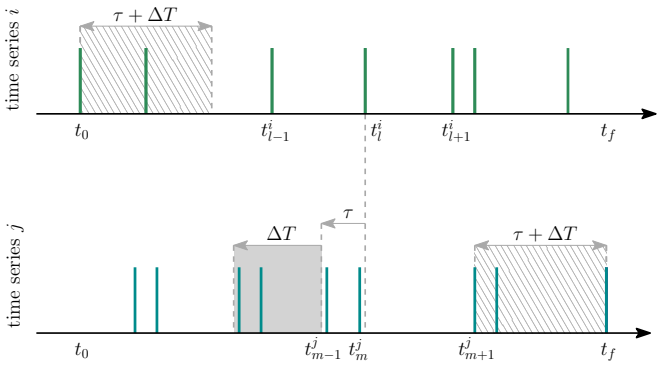

Fig. 2: Schematic sketch of the definition of event proximity used in ECA, including the extension to time delayed events.

here to maintain easy comparison with ES). In contrast to ES, we do not have to calculate local coincidence intervals for all pairs of events individually, which renders the computation somewhat faster and allows for analytical considerations[19,20]. The basic idea of the method is illustrated in Fig. 2.

To quantify the overall synchrony between events in the two time series, we consider *event coincidence rates*

$$r\left(i|j;\Delta T\right) = \frac{1}{s_j - s'_j} \sum_{m=1}^{s_j - s'_j} \Theta\left\{ \sum_{l=1}^{s_i} 1_{[0,\Delta T]}\left(t^i_l - t^j_m\right) \right\} \tag{9}$$

with the indicator function

$$1_I\left(x\right) = \begin{cases} 1, & \text{if } x \in I, \\ 0, & \text{otherwise,} \end{cases} \tag{10}$$

and the left-continuous Heaviside step function

$$\Theta\left(x\right) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{otherwise,} \end{cases} \tag{11}$$

which prevents double counting of events. Hence, this event coincidence rate gives the fraction of events in i that are preceded (within ΔT) by at least one event in j. For a correct normalization of the event coincidence rate[16,17], we consider the number of events occurring between $t_f - \Delta T$ and t_f,

$$s'_j = \sum_{m=1}^{s_j} 1_{[t_f - \Delta T, t_f]}\left(t^i_l\right), \tag{12}$$

and remove the latter from the total number of events at node j.

Similar as for ES, we may be interested in a symmetric matrix of pairwise statistical association coefficients. For the directional coincidence rates $r\left(i|j;\Delta T\right)$ and $r\left(j|i;\Delta T\right)$, this is typically done in either of the two following ways: highlighting strong unidirectional associations by calculating the maximum of the two pairwise event coincidence rates

$$Q^{ECA,max}_{ij} = \max\left(r\left(i|j;\Delta T\right), r\left(j|i;\Delta T\right)\right), \tag{13}$$

or the bidirectionality of undirected connections by calculating the mean of the event coincidence rates

$$Q^{ECA,mean}_{ij} = \frac{r\left(i|j;\Delta T\right) + r\left(j|i;\Delta T\right)}{2}. \tag{14}$$

Notably, a similar duality could be incorporated in the study of ES as well by implementing a symmetrization as in Eq. (13) into Eq. (8); however, this has not been commonly done in previous works, so that we refrain here from studying both variants also for ES.

Finally, for a multivariate (spatiotemporal) climate dataset, we obtain a functional network representation by utilizing the similarity matrix $\boldsymbol{Q}^{ECA} = \left(Q^{ECA}_{ij}\right)$, which is further described in Section III.

Along with our use of ECA, it should be noted that we have ignored the distinction between precursor and trigger coincidence rates raised in previous works[20], while only considering here the latter version (corresponding to what has been described above). In our real world climate example discussed below, we have found that the difference between the results obtained based on both versions is rather negligible (not shown).

D. Time Delayed Versions of ES and ECA

In addition to the global coincidence window ΔT, ECA commonly features a second parameter, the time lag τ[20] (cf. Figure 2). Despite being widely considered an algorithmic parameter of ECA but not ES, in the context of our conceptual comparison with ES, we emphasize that this time lag is in fact not a unique feature of ECA, but shall be considered in the same fashion like in lagged cross correlations and, hence, could be equally implemented in the ES to search for time delayed event synchronization. In the course of this work, we will, however, consider only time lagged versions of ECA when discussing the application to studying the spatiotemporal propagation of rainfall extremes over South America, while the corresponding analysis could be easily extended to ES. Specifically, everything that changes here with respect to the equations provided in the previous subsections is a shift of one of the event timing sequences by a constant value τ.

E. Correction scheme for Clustered Events

The common version of ES as described above inherits one possible caveat related to the treatment of temporally clustered events[16,18]. Specifically, if two events in one series occur at subsequent time steps (or very close in time),

the local coincidence window will collapse to $\frac{1}{2}$ timestep (see Fig. 1), so that there is an increased likelihood that two events actually occurring in close succession are *not* identified as synchronized. Notably, due to the consideration of a global coincidence window, this problem is not shared by ECA. In a parallel study[17], a detailed analysis of data sets with serially dependent events has been carried out, which demonstrates that the characteristics of the single time series tend to dominate the values of ES in some scenarios. This is compatible with recent numerical results of modeled spreading phenomena on different types of networks[18].

The issue mentioned above has already been addressed in previous studies[11,15] where a modification of the ES computation has been suggested: to avoid temporal clustering and the resulting bias of ES, each time series has to be analyzed individually first to eliminate all but the first event of each event cluster.

Depending on the density of events in time and the employed significance criterion for defining an edge, the application of this correction step can have substantial effects on the functional network structure inferred from the matrix of pairwise ES strengths. Beyond the effect of providing a tool for handling the issue of subsequent events, we can further motivate the proposed correction scheme as a method to interpret a cluster of multiple events as a single persistent event. In the context of climate extremes, this persistence can be caused by both, temporal persistence and spatial extent of a weather system feeding heavy rainfall in a large area. From this perspective, it appears reasonable to study the consequences of applying the same correction scheme to ECA as well. However, we stress that the application of the declustering step prior to applying ES and ECA is motivated by different rationales.

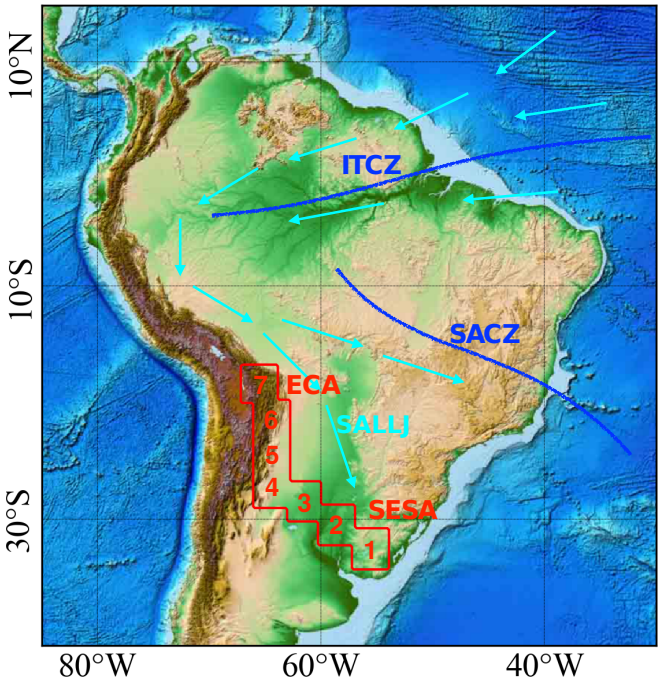

Fig. 3: Topography of South America and key features of the South American Monsoon System (SAMS), including typical wind directions (light blue arrows) and the South American Low-Level Jet (SALLJ). The climatological positions of the Intertropical Convergence Zone (ITCZ) and the South Atlantic Convergence Zone (SACZ) are shown by dashed dark blue lines. The red boxes illustrate the parcellation of the study area into 7 boxes to track the propagation of extreme precipitation events (see text).

III. FUNCTIONAL NETWORK REPRESENTATION OF SOUTH AMERICAN PRECIPITATION EXTREMES

As a particular application of both ES and ECA, we focus on spatiotemporal patterns of extreme precipitation associated with the South American Monsoon System (SAMS), which is schematically shown in Figure 3. The monsoon season in South America, which lasts from December to February (DJF), is strongly affected by moisture influx from the tropical Atlantic Ocean associated with trade winds converging at the Intertropical Convergence Zone (ITCZ)[28]. Driven by low level winds, moisture is recycled over and transported across the Amazon Basin towards the northern part of the Andes mountain range, where the low level winds get blocked and reflected, and moisture transport is channeled southwards[29,30]. Depending on the Rossby wave train phase[31], either the South American Low Level Jet (SALLJ) distributes moisture towards South Eastern South America (SESA)[32–35], or low level winds transport moisture along the South American Convergence Zone (SACZ) towards South East Brazil (SEBRA)[33,36]. This behavior results in the most prominent rainfall variability pattern in South America, the South American Rainfall Dipole[23,31,32,37].

For the identification of heavy precipitation events, we utilize rainfall estimates from the Tropical Rainfall Monsoon Mission (TRMM, version 3B42 V7)[38], which are available on a $0.25^\circ \times 0.25^\circ$ spatial grid at a temporal resolution of 3 hours between 1998 and 2015. As we are specifically interested in the rainfall events during the monsoon season, we employ 3 hourly and daily precipitation sums for DJF. For each grid point, we consider an event (i.e., precipitation value) to be *extreme* if the rainfall estimate exceeds the empirical 98th percentile of the raw 3-hourly data or the 90th percentile of the daily sums at each given site (node), respectively. Thereby, we obtain the same number of extreme events in most time series and ignore dry spots with an insufficient number of wet days during the study period. Thereafter, we will investigate the features of the resulting extreme event series by applying ES and ECA in different settings for constructing functional climate networks (see Section IV).

The widely used way to construct such networks (af-

ter identifying the events of interest) consists of applying either ES or ECA and thresholding the resulting similarity matrix $\mathbf{Q}$ at some value to obtain an adjacency matrix with a desired link density. For the uncorrected versions of both measures and therefore the same number of extreme events in all time series, this is a reasonable strategy[39,40], which we will also adopt in the following (the chosen link density will be specified along with the respective discussion). In the corrected versions of both methods, we apply the declustering scheme as discussed above, which modifies the characteristics of the event sequences. It should be kept in mind, that although the correction scheme has the advantage of accounting for the clustering bias, it also reduces the number of events in some time series. At this point, there are several ways to tackle this issue.

In recent studies by Boers et al.[11,15,21], each link has been defined based on significance testing utilizing surrogates which account for different numbers of extreme events in each pair of time series. Accordingly, network construction has been consistently adapted to the described correction scheme. Although this strategy does not conserve the serial temporal dependency structure of extreme events in each time series, it is the most often applied strategy to address the problem of temporally clustered events. A possible alternative way to cope with the clustering bias would be iteratively considering new events of lower magnitude as additional events until the desired number has been reached, thereby successively reducing the empirical percentile that has been effectively used for defining an event in each time series. While this would preserve the number of events, it also renders the analysis computationally costly.

As our study focuses on the different synchrony concepts and is not primarily meant to reveal novel climatological information, we will continue with the most naive approach and neglect the influence of different numbers of extreme events in the time series. This choice appears justified in the present setup, where the number of events does not vary drastically after the application of the correction scheme in most of the study area (see Fig. 4). Exceptions include the Pacific ocean region off the Chilean coast, the Atacama desert and the Orinoco Basin, along with some regions in the Atlantic ocean to be further discussed below. In the following, we, therefore, obtain the adjacency matrices for the different functional climate networks (as previously described) by thresholding the respective similarity matrix at a certain value to obtain a specific link density.

For more in-depth future studies, we recommend obtaining links in functional climate networks through one of the described strategies and not by varying link densities until certain features become visible. Varying link densities go in hand with substantial changes in network structure[41,42] and should be therefore handled with care. In our analysis, we on purpose chose a well-studied region to highlight the differences between the different approaches while being confident about the interpretation of the described climatic features.

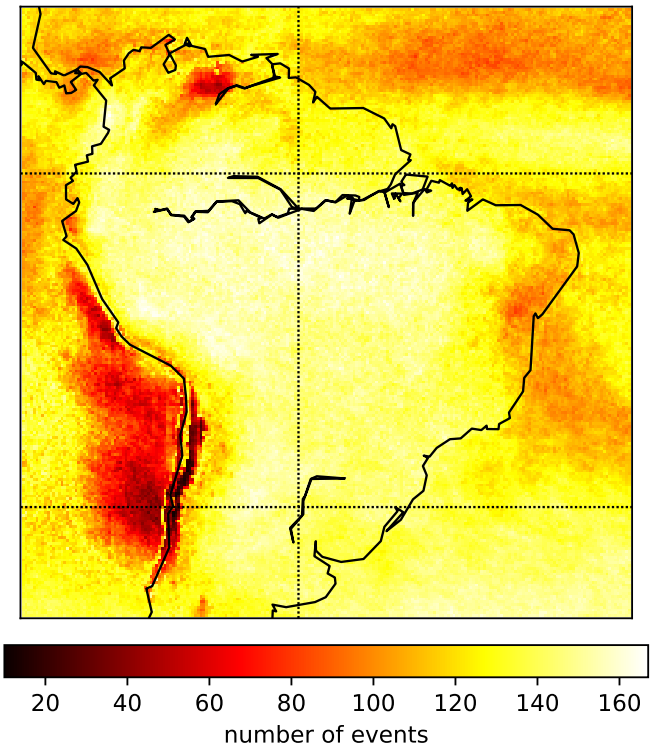

Fig. 4: Number of remaining events after declustering the obtained event sequences with initially a constant number of events for daily precipitation data (see text for details).

IV. RESULTS

A. Method Intercomparison

The first aim of this paper is to highlight the different features of functional climate networks from precipitation data when inferred by using different event based statistical association measures. To tie in with previous studies, our analysis copies a previously employed setup[22] for studying daily precipitation estimates from the TRMM satellite mission, setting an upper bound to the local coincidence interval of ES as $\tau_{lm}^{ij} \leq \tau_{max} = 3$ days (note that this should potentially have a similar effect like the global coincidence interval of ECA, avoiding synchronized events to be too heavily mutually lagged). We set the link density of the resulting functional climate network to $\rho = 0.02$, which is a common choice for climate networks with high spatial resolution[13,41].

In line with ref.[22], we initially apply ES in the original version without the algorithmic correction and show the resulting node degree pattern of the described setup in Figure 5a. We identify a low degree channel in the area of the ITCZ as well as elevated node degrees at the position of the SALLJ driven moisture pathways along the Andes,

while the moisture exit zone in SEBRA is not clearly recognizable.

To emphasize the effect of temporal declustering on the results of ES based functional climate networks (i.e., the network was constructed after we deleted clustered events in the way described above), the resulting degree field is depicted in Fig. 5b. Here, we observe more pronounced features in comparison with the degree field of the network obtained by the application of the uncorrected ES version. While the degree field in Figure 5a appears blurry especially in the southern Amazon basin, we can differentiate regions of high degree much better from such with lower degree in Figure 5b. This effect is caused directly by the correction scheme, which prevents a systematic underrepresentation of regions with clustered events (i.e., nodes exhibiting a negative bias of the ES strength[16–18]). Among the visible differences, we firstly observe a narrow and distinct low degree band associated with the ITCZ. Secondly, the northern part of the Amazon basin, where heavy precipitation events and moisture recycling take place, is characterized by elevated degree values. Thirdly, we observe relatively high degree values at the Eastern flank of the Andes mountain range, where the SALLJ transports moisture towards the southern parts of South America. Finally, the strongest variability pattern associated with the SAMS, the South American Rainfall Dipole with maximal variance in SESA and SEBRA, is also well visible in the degree field.

To further compare the networks obtained based on ES with those inferred by utilizing ECA, Figure 5c shows the degree pattern of the network constructed by employing ECA with $\Delta T = 3$ and using the mean of the two pairwise event coincidence rates as the proxy for statistical association. This choice ensures that we capture similar time scales as in the ES based networks. Note that in the ECA, the global coincidence interval ΔT defines the maximal time lag between synchronized extreme events, which agrees with the limitation of the local coincidence interval of ES to $\tau^{ij}_{lm} \leq 3$ days used above. In comparison with previous degree fields based on ES, the pattern derived by the application of the ECA exhibits major differences. The area north of the ITCZ is much stronger represented in the ECA based network and is characterized by much higher degree values than any other region in the ES based networks. In particular, we again clearly recognize the ITCZ as a key feature. As we have employed the same fixed link density, the other features previously observed in the ES based networks necessarily appear less pronounced. While the rainfall dipole areas in SESA and SEBRA are partly visible and the northern part of the Amazon basin is represented by elevated degree, there is no coherent pattern at the Eastern flank of the Andes mountain range.

Finally, to investigate the effect of event declustering on the ECA based networks, Figure 5d shows the node degree based on the ECA in combination with the correction scheme. While there are certain differences between the uncorrected ECA and the corrected version, the main features are similar to those of the networks based on corrected ES (Figure 5b) and the corrected ECA and only differ in terms of a weaker effect in the Eastern Central Andes along the SALLJ and a slightly intensified differentiation of the ITCZ and the South American Rainfall Dipole in the ECA based network.

B. Network Patterns and Event Clustering

To further understand the differences between the ES and ECA based networks, we investigate the pairing coefficient P_i of each underlying time series[16,17]

$$P_i = \frac{1}{s_i - 1} \sum_{l=1}^{s_i - 1} \delta[\left(t^i_{l+1} - t^i_l\right) - 1]. \qquad (15)$$

The pairing coefficient quantifies the temporal clustering of events in a time series and is normalized such that $P_i = 0$ if there are no pairs of events at subsequent time steps present in the event time series, while $P_i = 1$ corresponds to a situation where all events occur at subsequent time steps. Note that the pairing coefficient does not measure any association between two time series, but a feature of each single time series at each grid point.

Figure 6 reveals the corresponding intrinsic features of the event time series. Notably, there is no simple direct correspondence between the pairing coefficient and the number of events after declustering (Fig. 4), since both are affected differently by different size distributions of clustered event sequences (not shown). Interestingly, the prominent region of high pairing coefficients coincides with the region north of the ITCZ where the ECA based network exhibits relatively high degree values. As the uncorrected ES does not properly address series of subsequent events due to the definition of a possibly overly narrow local coincidence interval, regions in which events occur frequently at subsequent time steps get systematically underrepresented by the uncorrected ES. In turn, the uncorrected ECA counts every single event as an isolated event even though some occur at subsequent time steps, which automatically leads to higher possible event coincidence rates. As weather systems normally have a spatial extent that covers multiple grid points, this effect gets amplified, because these systems can cause extreme events at subsequent time steps among a multitude of geographically nearby nodes. As a consequence, we observe higher degree values in the uncorrected ECA based network in comparison with the uncorrected ES based network in regions with high pairing coefficient.

In the corrected version of ES and ECA, we reduce all event clusters before applying the synchrony measures. Therefore, as the absolute number of events in such time series decreases, the possible number of synchronous events decreases as well. This can lead to a reduction of the estimated value of the considered association measure and, thus, a lower node degree in the

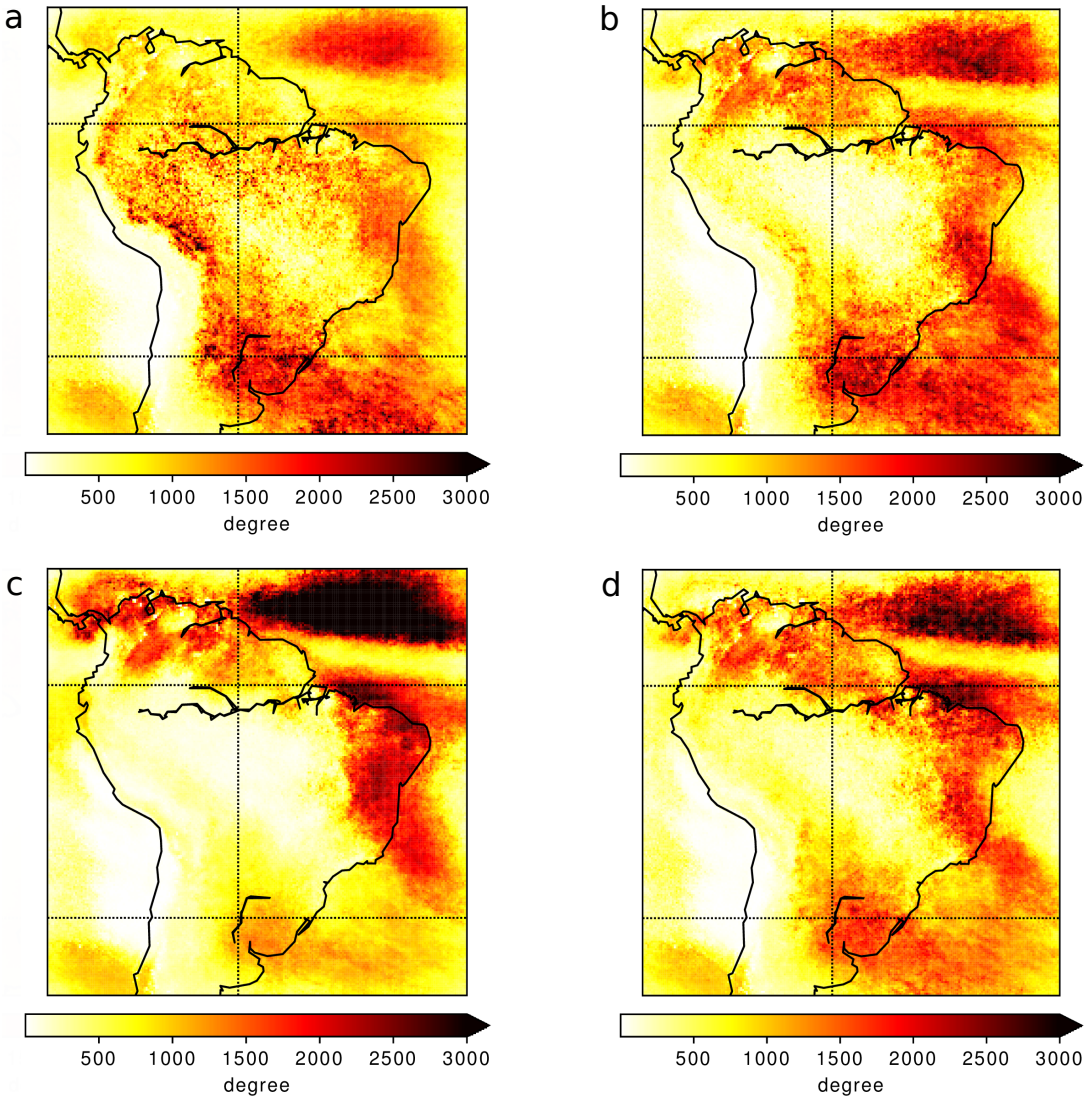

Fig. 5: Degree of the functional climate network representations of heavy rainfall events based on (a,b) ES and (c,d) ECA without (a,c) and with (b,d) the utilization of the correction scheme for temporally clustered events. For the two ES based networks, we set $\tau_{lm}^{ij} \leq 3$ days, while for ECA, $\Delta T = 3$ days. All networks exhibit a link density of 0.02.

affected regions of the resulting network. As ES and ECA exhibit certain structural differences, we correct here for different effects by applying the correction scheme prior to the utilization of the synchrony measures. On the one hand, ES underrepresents regions with high pairing coefficient due to sometimes overly narrow local coincidence intervals. On the other hand, ECA potentially overrepresents regions with high pairing coefficient, because subsequent rainfall events may belong to the same weather system and therefore should be considered as one event that should be counted once instead of multiple times. However, we obtain rather similar results with both measures when we restrict the considered time scales to the same (short) length and apply the same declustering procedure.

To further highlight the differences between ES and ECA based networks, we show scatter plots of the obtained degrees for all nodes in Fig. 7, which underline the limitations of, and the differences between the different measures. To confirm that the characteristics of the time series are responsible for the observed differences, we color-code all displayed pairs of values according to the pairing coefficient of the associated node. We observe that nodes with high and low pairing coefficients are separated in the scatter plots based on the uncorrected measures, see Figure 7a. While only nodes with a high pairing coefficient can exhibit a high degree in the ECA based networks, such nodes tend to exhibit low to intermediate degree in the ES based network[16,17]. In turn, nodes with a low pairing coefficient can have both,

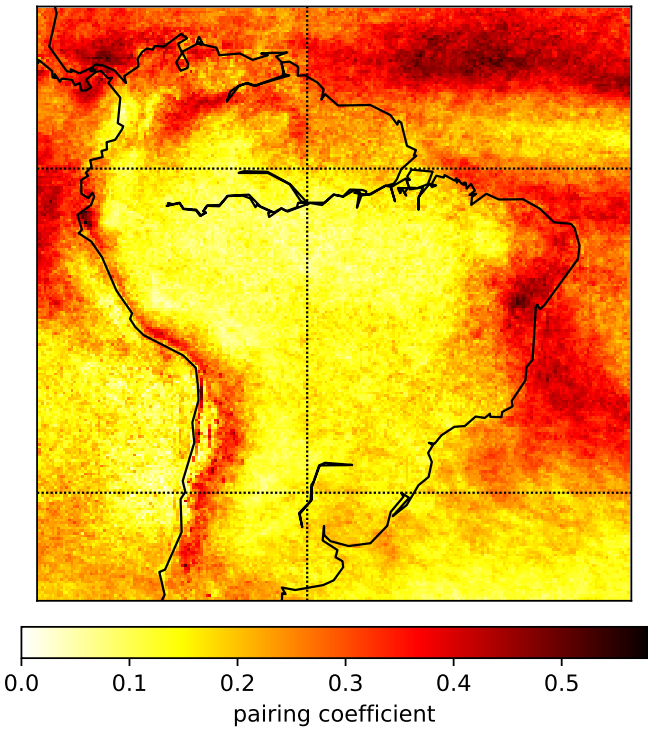

Fig. 6: Pairing coefficient of heavy daily precipitation events.

low or high degrees in the ES based network, but display mostly low degrees in the ECA based network. These findings are consistent with the expected behaviors discussed in the previous paragraph.

In Figure 7b, we show the degrees of each node in the networks constructed from the corrected versions of both ES and ECA. Here, the points appear more closely aligned along the line of identity, indicating that the degree of the nodes is commonly comparable in both networks. In addition, we do not observe a marked separation between nodes with high and low pairing coefficient. This is in line with Figure 5, where we had shown the spatial patterns of the node degree and observed a strong similarity between the ES and the ECA in their corrected versions, which is a direct consequence of the limitation of $\tau_{max} = 3$ for ES (and setting $\Delta T = 3$ for ECA accordingly). Residual differences hence originate from the existence of local coincidence windows of ES that are smaller than the global one of ECA (yet still larger than zero), thereby focusing on different delay time scales.

C. South American Moisture Pathways Across Scales

To highlight the relevance of the different aspects of event synchrony captured by ES and ECA, we finally apply both measures to extreme 3-hourly rainfall sums over South America, to track associated rainfall patterns traversing from SESA to the Eastern Central Andes[11,14]. Specifically, we are interested in the capabilities of both approaches to reveal distinct features of the climate dynamics in comparison with the results of a previous ES based study[11]. In the latter reference, the authors presented a comprehensive framework employing ES along with the declustering of events to unveil the sources, pathways, and time scales of extreme precipitation events in the region of interest. For this purpose, several regions of interest have been identified within the study area (cf. red boxes in Fig. 3), and the direction of extreme precipitation event propagation between those regions has been unveiled by utilizing a directed network approach, focusing on network divergence as the difference between the numbers of incoming and outgoing links at each node. Specifically, these previous works analyzed the time scales of subsequent rainfall events which first occur in box 1 followed by heavy precipitation in the other boxes. Along with the utilization of the ES concept, they quantified the number of extremes in the consecutively numbered regions that occurred in close succession to those in region 1.

Unlike in those earlier works, we here aim to identify the characteristic time scales of those subsequent rainfall events by applying ECA and using the resulting functional climate network properties only. For this purpose, we consider a set of functional climate networks based on the same data, which are distinguished by different parameter settings of the ECA. In particular, we set the link density for all ECA configurations to $\rho = 0.05$ and also fix $\Delta T = 1$ time step (3 hours), while systematically varying the delay $\tau \in [0, 9]$ time steps. As we choose $\Delta T = 1$, we do not correct for temporally clustered events in this scenario. In each network configuration, we measure the total cross degree between region 1 and each of the other regions, $K^{1,n}$, and subtract the corresponding value $K^{1,r;e}$ of the same measure that one would expect when considering the link distance distribution of the entire functional climate network and the fraction of area covered by the respective region (i.e., the number of nodes per region):

$$K^{1,n*} = K^{1,n} - K^{1,n;e} = (1 - F_n)K^{1,n} \tag{16}$$

with

$$F_n = \frac{|E^{1,n}|}{\sum_{e \in E} \left[\Theta\left(d(e) - d^{1,n}_{min}\right) - \Theta\left(d(e) - d^{1,n}_{max}\right)\right]} \tag{17}$$

being the ratio between the total number of *possible* edges between the source region 1 and the target region n and the *actual* edges in the network in the range of allowed link distances between both regions. $\Theta(\bullet)$ again denotes the Heaviside step-function and $d(e)$ the distance covered by an edge $e \in E$. To compare the results obtained from the ECA based networks to those of the networks based on the parameter free version of the corrected ES (no time lag, no upper bound to the local coincidence interval), we additionally perform the same analysis for the

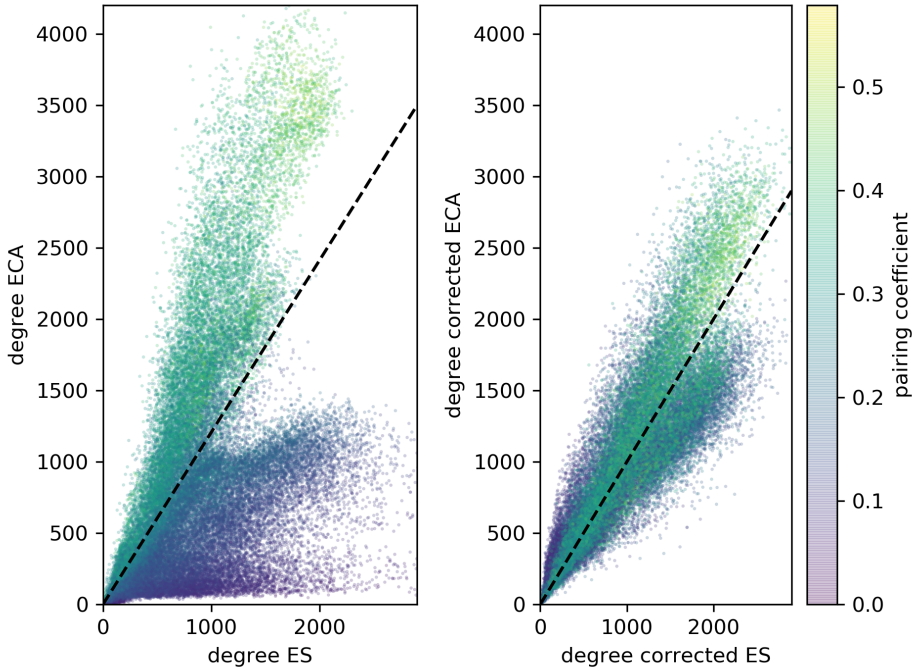

Fig. 7: Degree–degree scatter plot for the uncorrected (left) and corrected (right) versions of ES and ECA. The x axes show the degree of a node in the ES based network, while the y axes show the degree in the ECA based network. The color code indicates the respective pairing coefficient, the dashed line the line of identity.

latter, using the same link density of $\rho = 0.05$. The corresponding results are shown in Fig. 8 and feature multiple characteristics that indicate the specific capabilities of the different methods:

Firstly, we obtain rather distinct results for the various parameter settings of the ECA. For small time delays τ, we find high rescaled total cross degrees for the areas 2 and 3 which are close to the source region 1, as opposed to low values for those areas at larger distance to the source region. With increasing delay, the maxima of the rescaled total cross degree continuously shift towards the more distant regions.

Secondly, we highlight the differences between the two previously suggested symmetrization methods for event coincidence rates that can be used to obtain the adjacency matrix for the ECA based networks. Specifically, we find that utilizing the maximum of the two pairwise event coincidence rates (Eq. 13) allows a better tracking of the event propagation than when employing their mean value (Eq. 14). This is a natural consequence of the physical process of directed propagation of extreme events, which is much better represented by keeping the unidirectional information and leads to the clear shift of the maximum total cross degrees from regions close to the source region to the more distant regions best visible for the maximum symmetrization. Notably, this approach also reproduces the time scale of the event propagation from SESA to the Eastern Central Andes of $18\text{h} - 24\text{h}$, which was found earlier by Boers et al.[11]. By contrast, the average symmetrization results in positive total cross degree values for region 2 for all considered values of the delay parameters. This is likely a consequence of regionally confined (as opposed to northwestward traveling) weather systems associated with heavy rainfall in SESA.

Finally, we also find that the corrected version of the ES can capture the full cascade of extreme events in all regions at once without selecting any parameters, which can be seen in terms of consistently positive rescaled total cross degrees between the source region 1 and all other regions. However, all these values are relatively low in magnitude in comparison to the maximum rescaled total cross degrees values in the ECA based networks, which in turn commonly do not feature the complete cascade individually. Moreover, there is no clear order of the rescaled total cross degree in the ES based networks without delay, thereby not allowing to infer the detailed pathways of event propagation.

V. CONCLUSION AND OUTLOOK

We have studied synchrony measures for constructing functional climate networks from event time series. Specifically, we have investigated event coincidence analysis (ECA) and event synchronization (ES) in their most appropriate definitions, including a correction scheme for the application of both synchrony measures to temporally

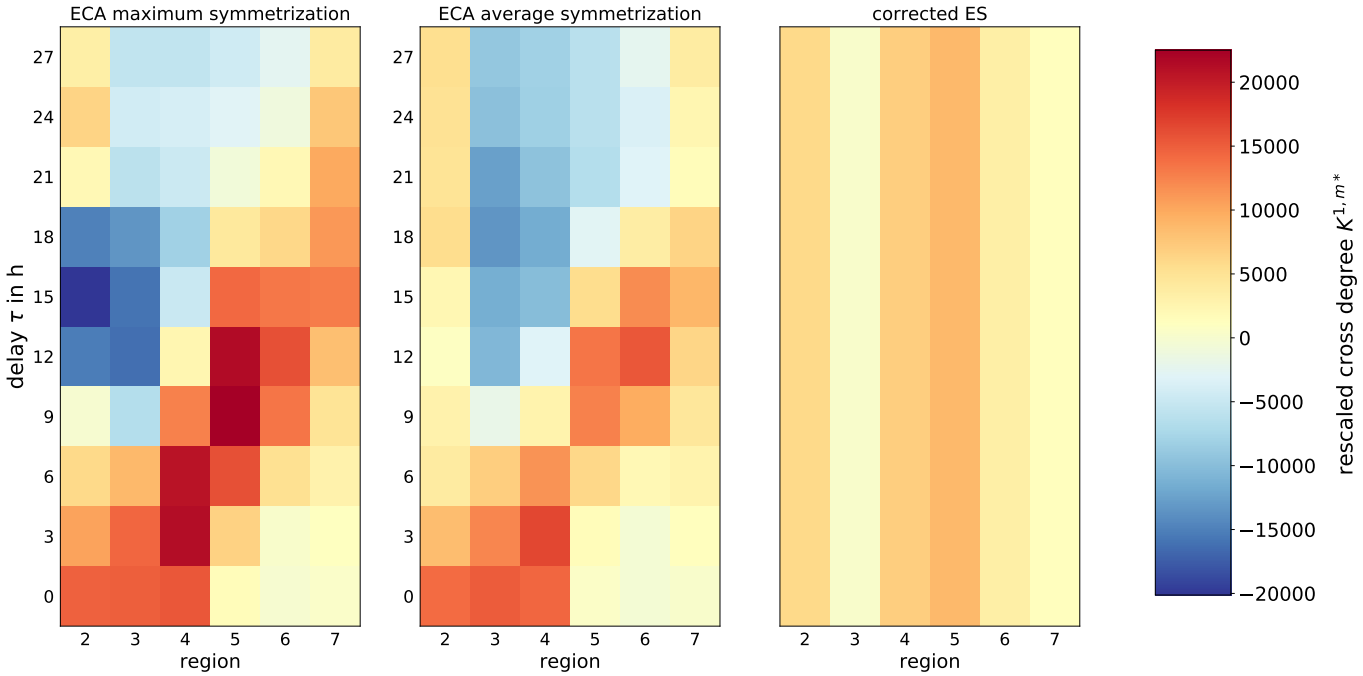

Fig. 8: Rescaled total cross degree $K^{1,m*}$ between region 1 and the other boxes in Figure 3 for (left) ECA using the maximum of the pairwise event coincidence rates, (center) ECA using the mean of the pairwise rates, and (right) corrected ES, allowing for a varying delay in the ECA based networks.

clustered events.

To highlight the influence of the said correction scheme as well as the differences between ES and ECA, we have investigated the spatiotemporal dynamics of heavy precipitation in the context of the South American Monsoon System as a well studied regional climate phenomenon in terms of different network analyses. We have illustrated the effect of the correction scheme on both, ES and ECA, by comparing the resulting node degree fields constructed using both approaches without and with the correction scheme. The difference is remarkable and indicates that studies using event synchrony measures should not only consider the structural bias of the original formulation of ES due to temporal event clustering but also precisely elaborate on the event definition and take possible biases resulting from temporal clustering of events into account when applying the ECA as well.

Finally, we have analyzed the propagation of heavy rainfall events in South America with the parameter free version of ES in combination with the correction scheme as well as ECA in various parameter settings and with both common choices of symmetrizations of pairwise (directional) event coincidence rates. We have shown that the ES in combination with the correction scheme can capture the presence of propagating events without any further adaptation or filtering of the data by estimating the total cross degree between different regions within a functional climate network. Utilizing ECA along with the implementation and systematic variation of possible time delays, we have demonstrated that the latter approach effectively highlights the different time scales of the extreme precipitation cascade traveling through different regions.

As the main focus of this study was to compare the well established ES and the more recently introduced ECA in the context of functional climate networks, we have not attempted to provide detailed climatic interpretations of the obtained network structures, particularly not the features like the region north of the ITCZ and in the eastern Pacific characterized by elevated degree, which were not found in previous ES based studies but highlighted by ECA. In addition, we have not further investigated the differences between the different strategies for defining links in functional climate networks, which could further clarify differences between common approaches to construct such networks. To this end, we outline corresponding further investigations and developments as subjects of future research.

ACKNOWLEDGMENTS

This work has been financially supported by the IRTG 1740/TRP 2011/50151-0 (funded by DFG and FAPESP) and by the German Federal Ministry for Education and Research (BMBF) via the Young Investigators Group CoSy-CC^2: Complex Systems Approaches to Understanding Causes and Consequences of Past, Present and Future Climate Change (grant no. 01LN1306A) and the Belmont Forum/JPI Climate project GOTHAM

(Globally Observed Teleconnections and their Representation in Hierarchies of Atmospheric Models, grant no. 01LP16MA). NB acknowledges funding by the Humboldt foundation, the Volkswagen Foundation, and the European Union's Horizon 2020 research and innovation programme under grant agreement No 820970.

REFERENCES

[1]S. H. Strogatz, Nature **410**, 268 (2001).
[2]R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[3]M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[4]S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, Phys. Rep. **424**, 175 (2006).
[5]R. W. Katz and B. Brown, Barbara, Clim. Chang. **21**, 289 (1992).
[6]D. R. Easterling, J. L. Evans, P. Y. Groisman, T. R. Karl, K. E. Kunkel, and P. Ambenje, Bull. Am. Meteorol. Soc. **81**, 417 (1999).
[7]C. Rosenzweig, A. Iglesias, X. B. Yang, P. R. Epstein, and E. Chivian, Glob. Chang. Hum. Heal. **2**, 90 (2001).
[8]R. Q. Quiroga, T. Kreuz, and P. Grassberger, Phys. Rev. E **66**, 041904 (2002).
[9]N. Malik, N. Marwan, and J. Kurths, Nonlinear Process. Geophys. **17**, 371 (2010).
[10]N. Malik, B. Bookhagen, N. Marwan, and J. Kurths, Clim. Dyn. **39**, 971 (2012).
[11]N. Boers, B. Bookhagen, H. M. J. Barbosa, N. Marwan, J. Kurths, and J. A. Marengo, Nat. Commun. **5**, 5199 (2014).
[12]V. Stolbova, P. Martin, B. Bookhagen, N. Marwan, and J. Kurths, Nonlinear Process. Geophys. **21**, 901 (2014).
[13]R. V. Donner, M. Wiedermann, and J. F. Donges, *Nonlinear and Stochastic Climate Dynamics*, 1st ed., edited by C. Franzke and T. O'Kane (Cambridge University Press, Cambridge, 2017) pp. 159–183.
[14]N. Boers, H. M. J. Barbosa, B. Bookhagen, J. A. Marengo, N. Marwan, and J. Kurths, J. Clim. **28**, 7641 (2015).
[15]N. Boers, B. Goswami, A. Rheinwalt, B. Bookhagen, B. Hoskins, and J. Kurths, Nature **566**, 373 (2019).
[16]A. Odenweller, *Coupled Complex Network Analysis of Extreme Precipitation in the Asian Monsoon Region*, Bsc thesis, University of Cologne (2017).
[17]A. Odenweller and R. V. Donner, "Disentangling synchrony from serial dependency in paired event time series," (2019), arXiv:1910.12343 [physics.data-an].
[18]F. Hassanibesheli and R. V. Donner, Chaos **29**, 083125 (2019).
[19]J. F. Donges, R. V. Donner, M. H. Trauth, N. Marwan, H.-J. Schellnhuber, and J. Kurths, Proceedings of the National Academy of Sciences **108**, 20422 (2011).
[20]J. F. Donges, C. F. Schleussner, J. F. Siegmund, and R. V. Donner, Eur. Phys. J. Spec. Top. **225**, 471 (2016).
[21]N. Boers, B. Bookhagen, N. Marwan, and J. Kurths, Clim. Dyn. **46**, 601 (2015).
[22]N. Boers, B. Bookhagen, N. Marwan, J. Kurths, and J. A. Marengo, Geophys. Res. Lett. **40**, 4386 (2013).
[23]N. Boers, A. Rheinwalt, B. Bookhagen, H. M. J. Barbosa, N. Marwan, J. A. Marengo, and J. Kurths, Geophys. Res. Lett. **41**, 7397 (2014).
[24]N. Boers, A. Rheinwalt, B. Bookhagen, H. M. J. Barbosa, N. Marwan, J. A. Marengo, and J. Kurths, Geophys. Res. Lett. **41**, 7397 (2014).
[25]N. Boers, B. Bookhagen, J. A. Marengo, N. Marwan, J.-S. von Storch, and J. Kurths, J. Clim. **28**, 1031 (2015).
[26]J. Heitzig, J. F. Donges, Y. Zou, N. Marwan, and J. Kurths, Eur. Phys. J. B **85**, 38 (2012).
[27]J. F. Donges, H. C. Schultz, N. Marwan, Y. Zou, and J. Kurths, Eur. Phys. J. B **84**, 635 (2011), arXiv:1102.3067.
[28]J. Zhou and K. M. Lau, J. Clim. **11**, 1020 (1998).
[29]B. Bookhagen and M. R. Strecker, Geophys. Res. Lett. **35**, 1 (2008).
[30]D. C. Zemp, C. Schleussner, H. M. J. Barbosa, R. J. van der Ent, J. F. Donges, J. Heinke, G. Sampaio, and A. Rammig, Atmos. Chem. Phys. **14**, 13337 (2014).
[31]M. Gelbrecht, N. Boers, and J. Kurths, Sci. Adv. **8**, eaau3191 (2018).
[32]J. A. Marengo and W. R. Soares, J. Clim. **17**, 2261 (2004).
[33]B. Liebmann, G. N. Kiladis, C. S. Vera, A. C. Saulo, and L. M. V. Carvalho, J. Clim. **17**, 3829 (2004).
[34]P. Salio, M. Nicolini, and E. J. Zipser, Mon. Weather Rev. **135**, 1290 (2007).
[35]J. D. Durkee, T. L. Mote, and J. M. Shepherd, J. Clim. **22**, 4590 (2009).
[36]L. M. V. Carvalho, C. Jones, and B. Liebmann, J. Clim. **15**, 2377 (2002).
[37]C. Vera, W. Higgins, J. Amador, T. Ambrizzi, R. Garreaud, D. Gochis, D. Gutzler, D. Lettenmaier, J. A. Marengo, C. R. Mechoso, J. Nogues-Paegle, P. L. S. Dias, and C. Zhang, J. Clim. - Spec. Sect. **19**, 4977 (2006).
[38]Goddard Earth Sciences Data and Information Services Center (GES DISC): TRMM (TMPA) Precipitation L3 1 day 0.25 degree x 0.25 degree V7 **35** (2016).
[39]A. Agarwal, N. Marwan, M. Rathinasamy, B. Merz, and J. Kurths, Nonlinear Process. Geophys. **24**, 599 (2017).
[40]U. Ozturk, N. Malik, K. Cheung, N. Marwan, and J. Kurths, Clim. Dyn. **53**, 521 (2019).
[41]A. Radebach, R. V. Donner, J. Runge, J. F. Donges, and J. Kurths, Phys. Rev. E **88**, 052807 (2013).
[42]M. Wiedermann, J. F. Donges, J. Kurths, and R. V. Donner, Phys. Rev. E **96**, 042304 (2017).