LETTER

https://doi.org/10.1038/s41586-018-0872-x

Complex networks reveal global pattern of extreme-rainfall teleconnections

Niklas Boers[1,2]*, Bedartha Goswami[2,7], Aljoscha Rheinwalt[3,7], Bodo Bookhagen[3], Brian Hoskins[1,4] & Jürgen Kurths[2,5,6]

Climatic observables are often correlated across long spatial distances, and extreme events, such as heatwaves or floods, are typically assumed to be related to such teleconnections[1,2]. Revealing atmospheric teleconnection patterns and understanding their underlying mechanisms is of great importance for weather forecasting in general and extreme-event prediction in particular[3,4], especially considering that the characteristics of extreme events have been suggested to change under ongoing anthropogenic climate change[5–8]. Here we reveal the global coupling pattern of extreme-rainfall events by applying complex-network methodology to high-resolution satellite data and introducing a technique that corrects for multiple-comparison bias in functional networks. We find that the distance distribution of significant connections ($P < 0.005$) around the globe decays according to a power law up to distances of about 2,500 kilometres. For longer distances, the probability of significant connections is much higher than expected from the scaling of the power law. We attribute the shorter, power-law-distributed connections to regional weather systems. The longer, super-power-law-distributed connections form a global rainfall teleconnection pattern that is probably controlled by upper-level Rossby waves. We show that extreme-rainfall events in the monsoon systems of south-central Asia, east Asia and Africa are significantly synchronized. Moreover, we uncover concise links between south-central Asia and the European and North American extratropics, as well as the Southern Hemisphere extratropics. Analysis of the atmospheric conditions that lead to these teleconnections confirms Rossby waves as the physical mechanism underlying these global teleconnection patterns and emphasizes their crucial role in dynamical tropical–extratropical couplings. Our results provide insights into the function of Rossby waves in creating stable, global-scale dependencies of extreme-rainfall events, and into the potential predictability of associated natural hazards.

Using the framework of complex networks, we present an analysis of the spatial configuration of significant synchronizations between extreme-rainfall events (EREs) around the globe. Our analysis is based on the quasi-global, high-resolution satellite product Tropical Rainfall Measurement Mission (TRMM) 3B42 (see Methods section 'Data'). We define EREs locally as days with rainfall sums above the 95th percentile of wet days; events occurring on consecutive days are counted as single events and placed on the first day of occurrence (Fig. 1; see Methods section 'Data'). Typically, EREs do not occur as isolated events, but are associated with spatio-temporally organized weather patterns[9]. Depending on the relevant atmospheric mechanisms, this can lead to synchronizations between EREs at remote locations, as has been shown so far only on regional scales[10–12]. Revealing and explaining global ERE synchronization patterns, and in particular their spatial scales, is important for our understanding of the governing atmospheric processes. Advances in this regard have a key role in assessing the predictability of EREs and associated natural hazards[11], as well as for evaluating their representation in climate models[12]. In addition, better insights into the global coupling structure of extreme events are crucial for estimating how anthropogenic climate change may alter the dynamical characteristics of EREs[6,7].

We quantify the synchronicity of rainfall events using event synchronization (Methods section 'Event synchronization'). According to this measure, highly significant synchronizations (with a maximum delay of ten days) are represented as a network comprising 576,000 nodes, corresponding to the total number of spatial grid cells of the TRMM dataset. Network links are placed between two nodes if the corresponding synchronization values are significant, with $P < 0.005$ (see Methods sections 'Functional networks' and 'Significance of event synchronization and network construction').

A quantity of particular interest is the global distribution of the spatial distances across which significant synchronizations occur (Fig. 2). This distribution decays as the power law $p(d) \propto d^{-\alpha}$, with an exponent α very close to 1 for distances d below about 2,500 km, but exhibits super-power-law behaviour for longer distances. The scale-break at 2,500 km indicates that links representing significant synchronizations can be grouped into two distinct classes: first, links associated with regional weather systems with distances up to 2,500 km, which include mesoscale convective systems and tropical cyclones; second, links associated with global-scale teleconnections. Such teleconnections are generally understood to be caused either by direct signal transport due to atmospheric circulations or by propagating waves triggered by disturbances of these circulations[13]. From a statistical point of view, these global-scale teleconnections are remarkable because they are, as revealed here, far more likely than expected from the scaling regime of the shorter connections. Their existence can thus not be predicted from the scaling of the shorter links. Such instances have been referred to as 'dragon kings', and their occurrence is typically interpreted as a switch to a different physical regime[14,15]. They follow closely the distribution of all possible great-circle distances on Earth's surface, and their distances are hence restricted only by spatial embedding.

To elucidate the atmospheric processes leading to the global-scale rainfall teleconnections, we focus on the teleconnections associated with EREs in south-central Asia (SCA; Fig. 3), including parts of northern India, eastern Pakistan, Nepal and Tibet. This regional example is chosen because of its particular vulnerability to EREs and associated severe, record-breaking flood events, as occurred in each monsoon season from 2010 to 2014[3,4,8,16,17].

Before proceeding, however, a statistical problem needs to be solved, which arises in all data-driven interdependency analyses, and in particular in networks that are constructed on the basis of statistical similarities. Such approaches are generally biased because of multiple comparisons. In this case, we compare each timeseries with 575,999 other timeseries, which amounts to more than 10^{11} comparisons in total. Therefore, the network will contain links that—despite corresponding to statistically significant pairwise synchronization values—are present only because of random coincidences, and not because of physical mechanisms. Correspondingly, the spatial distribution of network links connected to SCA does not convey any meaningful information (Fig. 3a). We introduce here a technique to correct for the

[1]Grantham Institute for Climate Change, Imperial College, London, UK. [2]Potsdam Institute for Climate Impact Research, Potsdam, Germany. [3]Institute of Earth and Environmental Science, University of Potsdam, Potsdam, Germany. [4]Department of Meteorology, University of Reading, Reading, UK. [5]Department of Physics, Humboldt University, Berlin, Germany. [6]Saratov State University, Saratov, Russia. [7]These authors contributed equally: Bedartha Goswami, Aljoscha Rheinwalt. *e-mail: boers@pik-potsdam.de

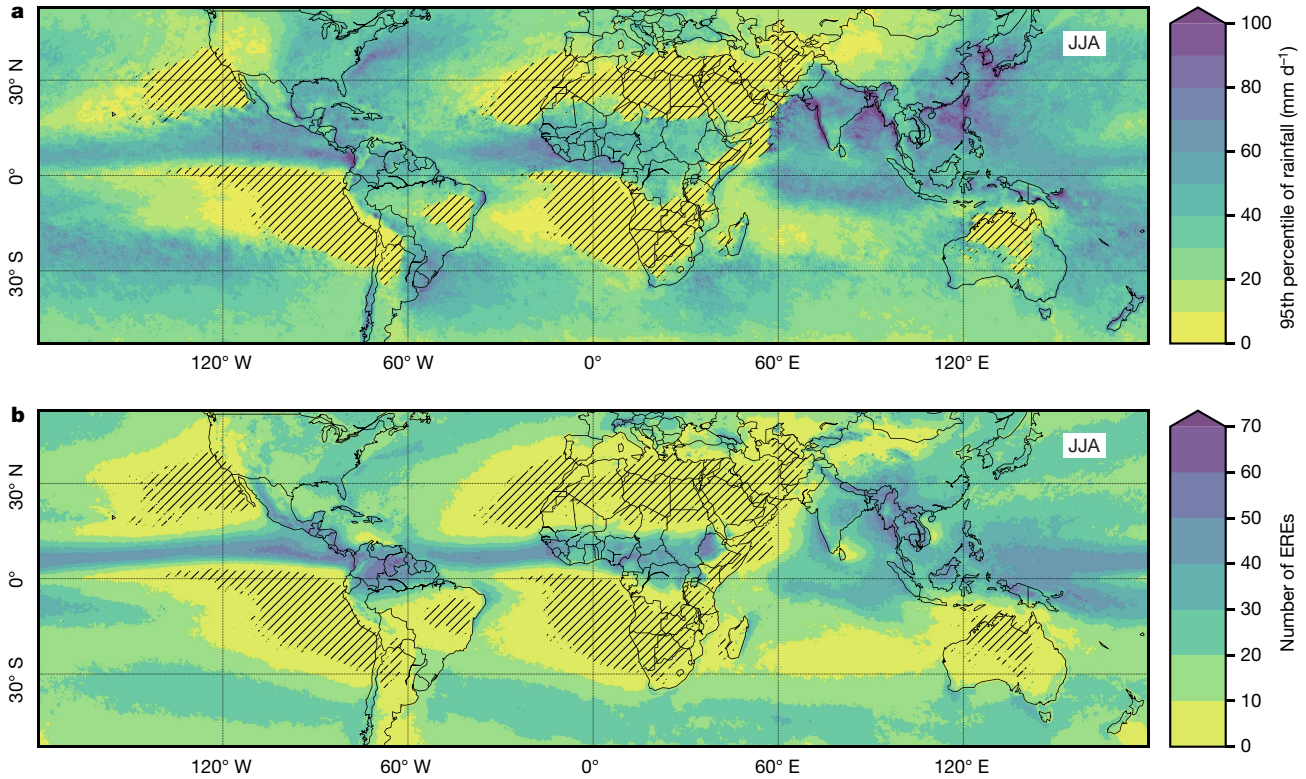


Fig. 1 | Percentiles and total numbers of extreme-rainfall events. **a**, Values of the 95th percentile of daily rainfall values for the June–July–August (JJA) season, computed individually for each of the 576,000 timeseries obtained from the 0.25° × 0.25° grid cells of the TRMM dataset. **b**, Total number of events above the threshold of the 95th percentile shown in **a**. Consecutive days with rainfall above the threshold are considered as single events. Hatched areas indicate regions with fewer than three events in total, which are excluded from the analysis.

multiple-comparison bias in spatially embedded functional networks, based on spatial density estimation to identify statistically significant ($P < 0.001$) spatial link bundles. Links that are not part of significant spatial bundles are then removed to extract the physically meaningful couplings (Fig. 3b, Extended Data Fig. 1; Methods section 'Significance of spatial patterns').

After applying our correction technique, a concise teleconnection pattern associated with the northern part of the South Asian monsoon is revealed: in addition to regional links covering most of the Indian subcontinent, we observe pronounced link bundles connecting SCA with eastern Asia, the African tropics, large parts of Europe and the eastern coast of North America, as well as the Southern Hemisphere extratropics. The break between regional and teleconnection scales (Fig. 2) is not affected by this correction (Extended Data Fig. 2).

Synchronizations between EREs in the tropical monsoon systems of Africa, India and eastern Asia can be understood, to a first approximation, in the context of the global monsoon, which affects all these regions and dominates during June, July and August[18]. We proceed with an analysis of the synchronizations between EREs in SCA and Europe. From a lead–lag correlation analysis of timeseries obtained by spatially averaging the numbers of EREs in both regions, we obtain an oscillatory pattern with a period of approximately 9 days: EREs in Europe lead EREs in SCA by 4–5 days, but EREs in Europe also lag EREs in SCA by 4–5 days, and therefore a one-sided causal relationship cannot easily be inferred (Fig. 4a). This oscillatory behaviour, with a period of about 9 days, and the phase shift of 4–5 days between Europe and SCA are consistent with a case study of the severe flood[4] in northern Pakistan in July 2010 (compare Fig. 4a with figure 7 in ref. [4]).

To investigate the relationship between EREs in SCA and Europe in more detail, we compute composite anomalies of rainfall and meridional wind speeds (v) for the specific times during which EREs occur synchronously in Europe and SCA (Fig. 4; Methods section 'Identification of specific times with high synchronization'). The corresponding anomalies of further atmospheric variables are shown in Extended Data Fig. 3. The anomalies show that EREs in Europe are caused by Rossby waves originating from the mid-latitude Atlantic region. Subsequently, these waves amplify the Eurasian wave train and thereby enhance the upper-level anti-cyclone over the Tibetan plateau. This leads to enhanced EREs in SCA, but also strengthens the

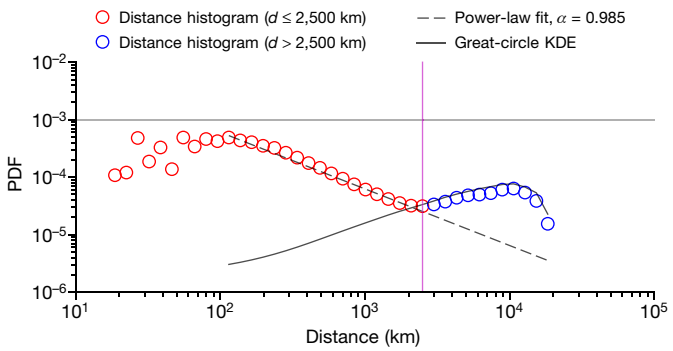


Fig. 2 | Distance distribution of extreme-event synchronizations. Probability density function (PDF) for the link distances (red and blue circles), the power-law fit for the range 100–2,500 km (dashed black line) and the kernel density estimate (KDE) of the distribution of all possible great-circle distances (solid black line) for events above the 95th percentile. The distributions are plotted on a log–log scale. The distribution of great-circle distances is inferred from a KDE of all possible great-circle distances between the 576,000 grid cells. The power-law exponent α implies that for distances d smaller than 2,500 km, their distribution is approximately proportional to d^{-1}. The vertical line at $d = 2{,}500$ km marks the regime shift from regional weather systems to large-scale teleconnections. For small distances (below about 100 km), the distribution is biased by the finite resolution of the 0.25° × 0.25° grid, with smaller geographical distances between grid cells closer to the poles. Results are shown for the JJA season, but are very similar in the December–January–February season (Extended Data Fig. 5).

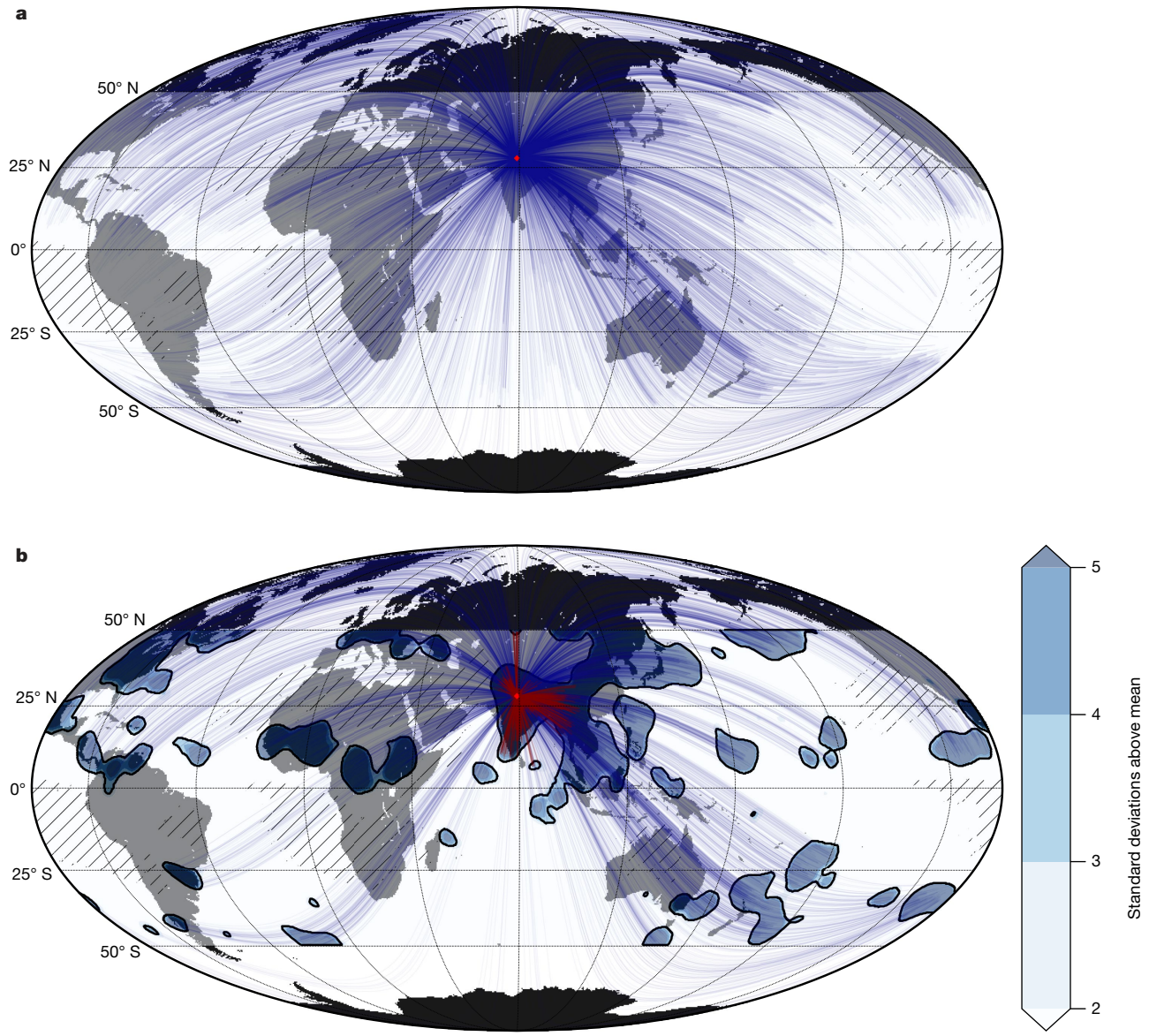


Fig. 3 | Teleconnection pattern for south-central Asia for events above the 95th percentile. a, Links attached to south-central Asia (SCA; 25° N to 32° N, 71° E to 88° E), before correcting for the multiple-comparison bias. **b**, Link bundles attached to SCA, after correcting for the multiple-comparison bias. Links shorter (longer) than 2,500 km are shown in red (blue). A spherical Gaussian KDE of the regional link density, in combination with a null model of randomly distributed links (Extended Data Fig. 1), is used to determine link bundles; links that are not part of significant bundles ($P < 0.001$) are omitted (Methods section 'Significance of spatial patterns'). Significant link bundles are shown by blue contours in units of standard deviations above the mean, as inferred from the null model. The black contour line delineates areas in which the regional link density is higher than the 99.9th percentile of the null-model distribution. The TRMM dataset used here ranges from 50° S to 50° N. The endpoints of all links lie within this latitudinal belt, even if the links appear to extend beyond it along their great-circle path, owing to the spherical projection.

upper-level high over eastern Asia, leading to synchronous EREs in the Yellow River basin (Figs. 3b, 4c).

Our results corroborate earlier studies on connections between the Eurasian wave train and the Indian summer monsoon[4,19] in terms of the 'Silk Road teleconnection'[20] and the 'circumglobal teleconnection' over the Northern Hemisphere[21], and demonstrate that statistically significant synchronizations between EREs across long spatial scales arise from these patterns. From the lead–lag correlation analysis we also infer a relationship between EREs in SCA and Europe in the opposite temporal direction. This is possibly related to the strengthening of the anti-cyclone over the Tibetan plateau, which, it has been argued, can enhance the amplitudes of the Rossby wave train upstream[22], thereby potentially triggering EREs in Europe.

The link bundles connecting SCA with the eastern coast of North America are probably the result of a chain of mechanisms: first, significant link bundles exist between SCA and tropical Africa. In the latter region, EREs are mainly controlled by the African easterly waves[23,24], which occasionally trigger storm systems and tropical cyclones that propagate across the tropical Atlantic Ocean before turning northwards, towards the Caribbean and the eastern coast of North America[25,26]. Furthermore, the link bundles extending to the Pacific Ocean are potentially associated with the Madden–Julian oscillation (MJO)[27,28], and in particular its phases 1 and 2 (Extended Data Fig. 4a). These two MJO phases have also been suggested to be related to the triggering of African easterly waves[29].

The fact that the global-scale teleconnections approximately follow the distance distribution of great circles (Fig. 2) suggests that Rossby waves are indeed the dominant mechanism behind these teleconnections, because these waves propagate—to first approximation—along great circles[30]. The super-power-law part of the distance distribution is strongly diminished if the analysis is restricted to the global tropics between 23.5° S and 23.5° N, whereas it remains substantial when

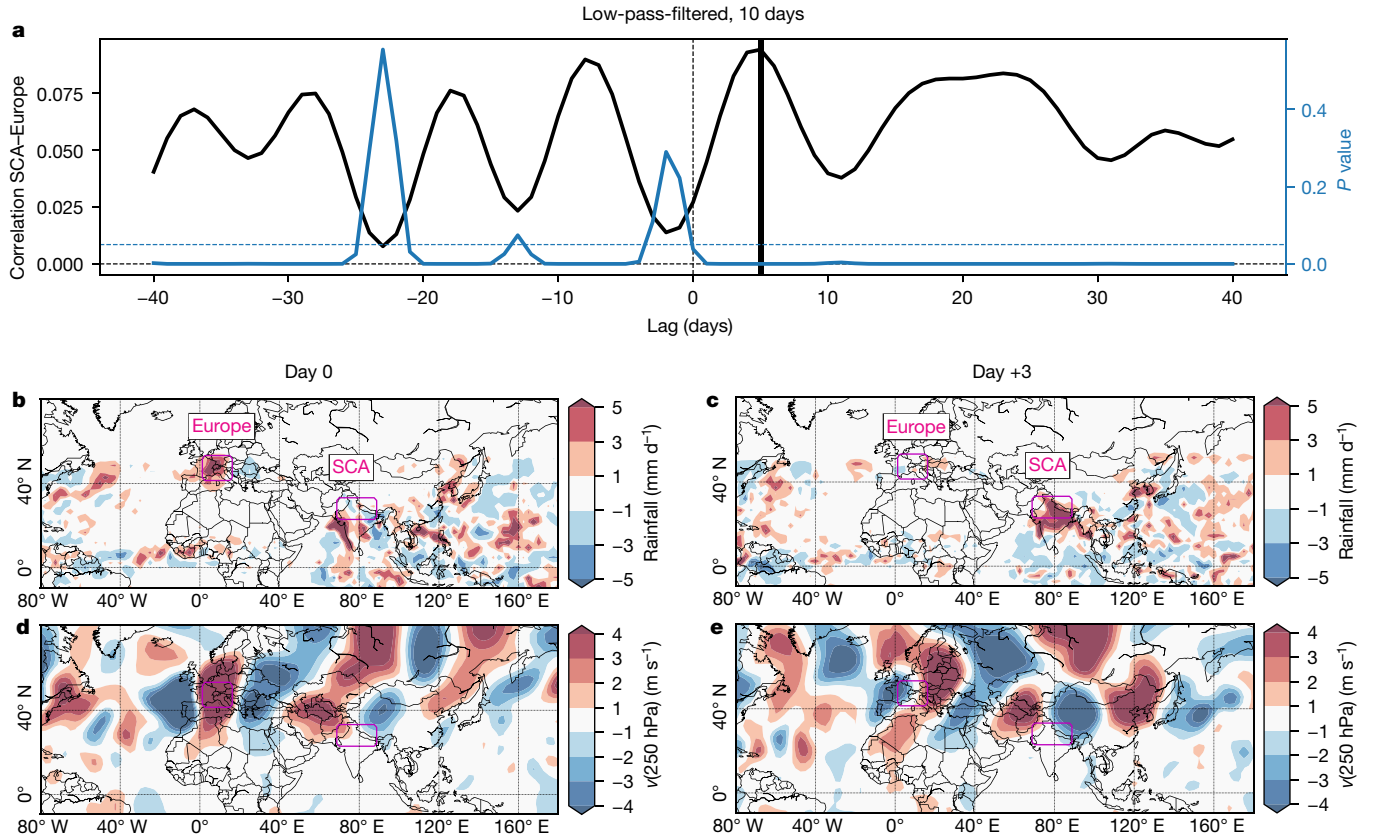


Fig. 4 | Atmospheric conditions for the teleconnection pattern between Europe and SCA. a, Lead–lag correlations (solid black line) of timeseries, obtained from spatially averaging the daily numbers of events above the 95th percentile in Europe (42° N to 50° N, 3° E to 15° E) and SCA (magenta boxes in **b**–**e**). The timeseries are low-pass-filtered with a cutoff period of 10 days (Methods section 'Identification of specific times with high synchronization'). The blue line shows the *P* values of the correlations at each lead/lag. **b**, Composite anomalies of TRMM rainfall with respect to the JJA climatology, for days with high numbers of EREs in Europe that are followed by associated EREs in SCA. **c**, Same as **b**, but 3 days later. **d**, Composite anomalies of the meridional wind component *v* at 250 hPa, with respect to the JJA climatology, for the same time steps as in **b**. **e**, Same as **d**, but 3 days later. The 'Silk Road' Rossby wave pattern[20] is clearly visible in terms of the east–west-oriented oscillation of positive and negative anomalies in **d** and **e**. The wave train strengthens towards the east within 3 days after the initial ERE occurrence in Europe. The dominant wavenumber associated with this Rossby wave pattern is $k = 6$, determined from the spatial power spectral density of the latitude belt from 37.5° N to 47.5° N.

including the links between the tropics and the extratropics of both hemispheres (Extended Data Fig. 5). This suggests a dominant role of processes connecting extratropical and tropical atmospheric dynamics—such as the Rossby wave train embedded on the mid-latitude westerly jet stream, which causes synchronous EREs in Europe and SCA—in shaping the global teleconnection pattern uncovered here.

The approach presented here provides the framework for further investigations of rainfall teleconnections in arbitrary regions of interest. Whereas we focus on events above the 95th percentile, the global distance distribution shown in Fig. 2 remains similar across event thresholds ranging from the 80th to 99th percentile, with slight increases in the tails of the distributions for the highest percentiles. Moreover, the spatial teleconnection pattern shown in Fig. 3b is robust across this range of thresholds (Extended Data Fig. 6). Further robustness tests with respect to the chosen threshold (Extended Data Fig. 7), the maximum delay τ_{max} (Extended Data Fig. 8), the cutoff frequencies (Extended Data Fig. 9) and the employed dataset (Extended Data Fig. 10) are described in Methods section 'Robustness tests'.

In the above example, we revealed the complete rainfall teleconnection pattern associated with the monsoon region of SCA. Some of these teleconnections can be explained with reference to previous studies, such as those investigating connections to Europe[20,21] and to northern China via the Silk Road pattern. Further investigations of the atmospheric dynamics will be needed to determine whether the above, rather speculative, explanation of the connections to tropical Africa and eastern North America via African easterly waves, as well as their connection to the Atlantic tropical cyclones, is valid. Although the teleconnections revealed here are consistent with the results of earlier studies on African easterly waves and their role in triggering tropical cyclones, additional investigations will be necessary to prove that these processes are indeed responsible. Furthermore, the mechanisms causing significant teleconnections between SCA and the Southern Hemisphere extratropics, as well as the suggested relationship with the MJO, will be addressed in future research.

We hope that our results will inspire further work on the predictability of EREs arising from these large-scale teleconnection patterns and on their representation in weather and climate models. Many studies have recently raised the concern that the characteristics of extreme events will change under ongoing climate change. A particular challenge in this regard is the discrimination between natural variability and anthropogenic influences[6,7]. With the increasing temporal lengths of global, high-resolution rainfall datasets, investigating changes of the global rainfall teleconnection structure along the lines of this study should become possible in the near future.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-018-0872-x.

Received: 7 August 2016; Accepted: 4 December 2018;
Published online 30 January 2019.

1. Hong, C. C., Hsu, H. H., Lin, N. H. & Chiu, H. Roles of European blocking and tropical–extratropical interaction in the 2010 Pakistan flooding. *Geophys. Res. Lett.* **38**, L13806 (2011).
2. Lau, W. K. M. & Kim, K.-M. The 2010 Pakistan flood and Russian heat wave: teleconnection of hydrometeorological extremes. *J. Hydrometeorol.* **13**, 392–403 (2012).
3. Webster, P. J., Toma, V. E. & Kim, H. M. Were the 2010 Pakistan floods predictable? *Geophys. Res. Lett.* **38**, L04806 (2011).
4. Hoskins, B. The potential for skill across the range of the seamless weather-climate prediction problem: a stimulus for our science. *Q. J. R. Meteorol. Soc.* **139**, 573–584 (2013).
5. Trenberth, K. E. & Fasullo, J. T. Climate extremes and climate change: the Russian heat wave and other climate extremes of 2010. *J. Geophys. Res. D* **117**, D17103 (2012).
6. Shepherd, T. G. Atmospheric circulation as a source of uncertainty in climate change projections. *Nat. Geosci.* **7**, 703–708 (2014).
7. Trenberth, K. E., Fasullo, J. T. & Shepherd, T. G. Attribution of climate extreme events. *Nat. Clim. Change* **5**, 725–730 (2015).
8. Cho, C., Li, R., Wang, S. Y., Yoon, J. H. & Gillies, R. R. Anthropogenic footprint of climate change in the June 2013 northern India flood. *Clim. Dyn.* **46**, 797–805 (2016).
9. Casanueva, A., Rodríguez-Puebla, C., Frías, M. D. & González-Reviriego, N. Variability of extreme precipitation over Europe and its relationships with teleconnection patterns. *Hydrol. Earth Syst. Sci.* **18**, 709–725 (2014).
10. Boers, N. et al. The South American rainfall dipole: a complex network analysis of extreme events. *Geophys. Res. Lett.* **41**, 7397–7405 (2014).
11. Boers, N. et al. Prediction of extreme floods in the eastern Central Andes based on a complex network approach. *Nat. Commun.* **5**, 5199 (2014).
12. Boers, N. et al. Extreme rainfall of the South American monsoon system: a dataset comparison using complex networks. *J. Clim.* **28**, 1031–1056 (2015).
13. Chase, T. N., Pielke, R. A. Sr & Avissar, R. Teleconnections in the Earth system. In *Encyclopedia of Hydrological Sciences* (eds Anderson M. G. & McDonnell J. J. (Wiley Online Library, 2006); https://doi.org/10.1002/0470848944.hsa190.
14. Sornette, D. Dragon-kings, black swans, and the prediction of crises. *Int. J. Terrasp. Sci. Eng.* **2**, 1–18 (2009).
15. Sachs, M. K., Yoder, M. R., Turcotte, D. L., Rundle, J. B. & Malamud, B. D. Black swans, power laws, and dragon-kings: earthquakes, volcanic eruptions, landslides, wildfires, floods, and SOC models. *Eur. Phys. J. Spec. Top.* **205**, 167–182 (2012).
16. Rasmussen, K. L. et al. Multiscale analysis of three consecutive years of anomalous flooding in Pakistan. *Q. J. R. Meteorol. Soc.* **141**, 1259–1276 (2015).
17. Shaevitz, D. A., Nie, J. & Sobel, A. H. The 2010 and 2014 floods in India and Pakistan: dynamical influences on vertical motion and precipitation. Preprint at http://arxiv.org/abs/1603.01317 (2016).
18. Wang, B. & Ding, Q. Global monsoon: dominant mode of annual variation in the tropics. *Dyn. Atmos. Oceans* **44**, 165–183 (2008).
19. Ding, Q. & Wang, B. Intraseasonal teleconnection between the summer Eurasian wave train and the Indian Monsoon. *J. Clim.* **20**, 3751–3767 (2007).
20. Enomoto, T., Hoskins, B. J. & Matsuda, Y. The formation mechanism of the Bonin high in August. *Q. J. R. Meteorol. Soc.* **129**, 157–178 (2003).
21. Ding, Q. & Wang, B. Circumglobal teleconnection in the Northern Hemisphere summer. *J. Clim.* **18**, 3483–3505 (2005).
22. Lin, H. & Wu, Z. Indian summer monsoon influence on the climate in the North Atlantic–European region. *Clim. Dyn.* **39**, 303–311 (2012).
23. Burpee, R. W. The origin and structure of easterly waves in the lower troposphere of North Africa. *J. Atmos. Sci.* **29**, 77–90 (1972).
24. Thorncroft, C. D. & Hoskins, B. J. An idealized study of African easterly waves. I: a linear view. *Q. J. R. Meteorol. Soc.* **120**, 953–982 (1994).
25. Thorncroft, C. & Hodges, K. African easterly wave variability and its relationship to Atlantic tropical cyclone activity. *J. Clim.* **14**, 1166–1179 (2001).
26. Chen, T. C., Wang, S. Y. & Clark, A. J. North Atlantic hurricanes contributed by African easterly waves north and south of the African easterly jet. *J. Clim.* **21**, 6767–6776 (2008).
27. Zhang, C. et al. Cracking the MJO nut. *Geophys. Res. Lett.* **40**, 1223–1230 (2013).
28. Mishra, S. K., Sahany, S. & Salunke, P. Linkages between MJO and summer monsoon rainfall over India and surrounding region. *Meteorol. Atmos. Phys.* **129**, 283–296 (2017).
29. Alaka, G. J. & Maloney, E. D. The influence of the MJO on upstream precursors to African easterly waves. *J. Clim.* **25**, 3219–3236 (2012).
30. Hoskins, B. J. & Ambrizzi, T. Rossby wave propagation on a realistic longitudinally varying flow. *J. Atmos. Sci.* **50**, 1661–1671 (1993).

Acknowledgements This study was conducted within the scope of IRTG 1740/TRP 2015/50122-0, funded by the German Science Foundation (DFG)/FAPESP. N.B. acknowledges further funding by the Alexander von Humboldt Foundation, the German Federal Ministry for Education and Research and DFG (reference BO 4455/1-1). B.G. is funded by the DFG project IUCLiD (project number DFG MA4759/8) and has received additional partial funding from the European Union's Horizon 2020 Research and Innovation programme under the Marie Skłodowska-Curie grant agreement number 691037 (project QUEST) and MWFK Brandenburg.

Reviewer information *Nature* thanks M. Barreiro, A. Gozolchiani and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions N.B. conceived the study, carried out the analysis and prepared the manuscript. N.B., B.G. and A.R. developed the methodology. All authors interpreted the results and edited the manuscript.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0872-x.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to N.B.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Data. We employ the gauge-calibrated, satellite-derived rainfall dataset of TRMM 3B42 V7[31], with daily temporal resolution, provided on a 0.25° × 0.25° spatial grid ranging from 50° N to 50° S (Fig. 1) for the time period 1998–2016. Rainfall events of increasing strength are identified in this dataset as days with rainfall sums above the 80th, 81st, …, 99th percentiles of wet days (that is, days with rainfall exceeding 1 mm) at each grid cell, for the JJA season. Consecutive days with rainfall above the threshold are considered as single events and placed on the first day of occurrence. The 576,000 grid cells and related timeseries of the TRMM dataset are associated with the nodes of the functional network.

To test whether the results of our study are independent of the specific dataset (see section 'Robustness tests'), we also employ daily data from the Global Precipitation Climatology Project[32] (GPCP 1DD V1.2), available at a global grid with 1° × 1° resolution. For the composite analysis of atmospheric variables, we employ the NCEP/NCAR Reanalysis 1 data[33].

Synchronization of extreme events. *Event synchronization.* Even for fixed spatial distances, the delay between synchronous rainfall events at different locations may vary in time because of varying scales of the driving atmospheric processes, for example, changing group or phase velocities of atmospheric waves. A suitable similarity measure that is capable of dealing with the technical challenges of event-like timeseries and the additional complication of varying temporal delays between events is event synchronization (ES)[34], which has been successfully employed to analyse the spatial synchronization structure of EREs on regional scales[10–12,35–37].

ES is defined as follows: let the set of events above a given percentile p at grid cell i be denoted by $\{{}^{p}e_i^{\mu}\}_{\mu=1,\ldots,l_i}$, where l_i denotes the total number of events at grid cell i. Events occurring on consecutive days are counted as single events. For each pair of grid cells (i, j), ES counts the number of pairs of uniquely associable events, where uniqueness is ensured by imposing the condition that the absolute value of the temporal delay between any two synchronous events μ and ν, $t_{i,j}^{\mu,\nu} := {}^{p}e_i^{\mu} - {}^{p}e_j^{\nu}$, must be smaller than $\tau_{i,j}^{\mu,\nu} := \frac{\min\{t_{i,i}^{\mu,\mu-1}, t_{i,i}^{\mu+1,\mu}, t_{j,j}^{\nu,\nu-1}, t_{j,j}^{\nu+1,\nu}\}}{2}$. As a maximum temporal delay between events at different locations we impose $\tau_{\max} = 10$ days. We hence define ES for the pair of locations (i, j) as the number of event pairs (μ, ν) meeting these conditions:

$$\mathrm{ES}_{ij} := \left\| \{(\mu,\nu) : |t_{i,j}^{\mu,\nu}| < \tau_{i,j}^{\mu,\nu} \wedge |t_{i,j}^{\mu,\nu}| \leq \tau_{\max}\} \right\| \quad (1)$$

where $|a|$ denotes the absolute value of a scalar quantity a, and $\|M\|$ the cardinality of a set M. We note that the original paper[34] suggests normalizing ES by multiplying with $(l_i l_j)^{-1/2}$ to make the measure independent of the event rates in the timeseries at i and j. It can be shown, however, that this normalization does not entirely remove this bias, because for higher event rates the probability of random synchronizations increases[37,38]. We therefore omit this normalization and instead focus on the statistical significance of each empirical value ES_{ij}, as defined in the above equation, on the basis of a null model that, for each pair (i, j), incorporates the same number of events as the original timeseries i and j.

Identification of specific times with high synchronization. A particular advantage of the similarity measure ES is that it allows for a dynamical delay $\tau_{i,j}^{\mu,\nu}$ within the range $[0, \tau_{\max}]$, in contrast to the static delay inferred from ordinary lead–lag correlation analyses. Furthermore, a modification of this measure allows us to determine the specific times during which extreme-event synchronization between two regions of interest is high, while keeping track of the temporal order. For two sets of timeseries, A and B, associated with two different regions of interest, we define:

$$\mathrm{ES}_{A\to B}^{\mu} := \left\| \{(i,j) \in A \times B : -\tau_{i,j}^{\mu,\nu} < t_{i,j}^{\mu,\nu} \leq 0 \wedge |t_{i,j}^{\mu,\nu}| \leq \tau_{\max}\} \right\| \quad (2)$$

and

$$\mathrm{ES}_{B\to A}^{\nu} := \left\| \{(i,j) \in A \times B : 0 \leq t_{i,j}^{\mu,\nu} < \tau_{i,j}^{\mu,\nu} \wedge |t_{i,j}^{\mu,\nu}| \leq \tau_{\max}\} \right\| \quad (3)$$

where $A \times B$ denotes the Cartesian product of sets A and B, that is, all possible combinations (i, j) such that $i \in A$ and $j \in B$.

Hence, $\mathrm{ES}_{A\to B}^{\mu}(\mathrm{ES}_{B\to A}^{\nu})$ can be viewed as a timeseries that gives for each time step the number of events in region A (B) that have a subsequent, uniquely associable event in region B (A). For the example of synchronized extreme events in SCA and Europe that is described in the main text, we first apply a low-pass filter with a cutoff frequency of 10 days to the two corresponding timeseries, and then determine days of strong synchronization by identifying the local maxima of the timeseries that are above the 90th percentile of the entire timeseries. These time points, which occur mainly during July and August and much less frequently during June and September (Extended Data Fig. 4b), are then used to compute the composite anomalies of rainfall and meridional wind speeds shown in Fig. 4, as well as the composite anomalies of precipitable water, geopotential height and streamfunction shown in Extended Data Fig. 3.

Functional networks. Functional networks are defined as networks for which each link is placed in accordance with statistically similar behaviour of the two corresponding nodes. Usually, the nodes of a functional network are identified with timeseries, and links are placed between the nodes if the two corresponding timeseries are strongly correlated. The connectivity structure of a dataset is thereby represented by the topology of the network and thus made mathematically accessible. In the recent past, functional-network techniques have been applied to analyse various properties of Earth's climate, such as global surface-temperature anomalies related to the El Niño Southern Oscillation[39–42], atmospheric teleconnections and waves[43,44], regional patterns of simultaneous rainfall occurrences in association with frontal systems[10,37], and prediction of EREs in the South American Andes[11,38,45].

Significance tests. *Significance of event synchronization and network construction.* The statistical significance for each observed ES_{ij} value is estimated as follows. First, a null-model distribution is numerically obtained by computing ES for 2,000 pairs of surrogate event series with l_i and l_j uniformly and randomly distributed events. For each pair (l_i, l_j) of event numbers, the 99.5th percentile of the corresponding distribution is determined as the significance threshold. Finally, a network link is placed between i and j if ES_{ij} is above this threshold—that is, if it is significant (according to this null-model) at a significance level of 0.005. Equivalent approaches to construct null models for ERE synchronizations have been employed in previous studies[10,11,37]. Of course, one could also construct a null model that preserves the observed distribution of inter-arrival times, but this would in fact lead to a less restrictive statistical test[11,37] because such a null model would preserve potential serial correlations: for each event e_i^{μ} at location i, ES looks for an event e_j^{ν} at location j such that the lead or lag between e_j^{ν} and e_i^{μ} is smaller than the absolute value of the adaptive delay $\tau_{i,j}^{\mu,\nu}$. But because stronger serial correlations lead to stronger clustering of events in each event timeseries, this would also imply a smaller $\tau_{i,j}^{\mu,\nu}$ on average. Consequently, the probability of finding an event e_j^{ν} within the range $\tau_{i,j}^{\mu,\nu}$ around e_i^{μ} will decrease with increasing temporal clustering, leading to lower values of ES in the null model. Hence, the null model is more conservative for less-clustered events. We further note that by placing network links in accordance with the statistical significance of ES_{ij}, potential biases due to different event rates are excluded.

Significance of spatial patterns. As noted in the main text, results derived from data-driven interdependency analyses, and in particular the link configuration of functional networks, are generally biased owing to multiple comparisons. This problem is independent of the employed similarity measure and other technical details, and even persists when conditional correlations are considered, such as in Bayesian networks[46] or causal inference studies[47].

There exist, of course, various approaches to adjust the significance level of the pairwise significance tests, with Bonferroni's and Sidak's corrections being the most prominent ones[48–50]. However, these techniques do not provide a suitable solution to the specific multiple-comparison problem presented here: on the one hand, lowering the significance threshold will also exclude synchronization values that may be caused by physical coupling mechanisms. On the other hand, after such corrections, there will still remain links that are caused by random coincidences. In fact, when analysing large datasets, such as the TRMM satellite product with 576,000 timeseries, it is not clear whether the fraction of links that are due to random coincidences would decrease when adjusting the significance level.

We suggest here a different approach to correct for biases caused by multiple comparisons, which exploits the fact that the system under study is spatially embedded. The key idea is that links that are caused by physical mechanisms should exhibit spatially coherent patterns, in contrast to links caused by random coincidences. A Gaussian KDE of the spatial link distribution shown in Fig. 3a, using Scott's rule for bandwidth selection[51] and the Haversine metric to account for spherical embedding, reveals that there indeed exist regions with substantially higher link density than that found in other parts of the globe (Fig. 3b).

To construct a statistical null model for this regional link density, we first distribute the same number of links randomly across those grid cells that lie within the latitudinal belt between 50° S and 50° N and have at least three events above the respective percentile threshold during the study period (the non-hatched areas in Fig. 1 for events above the 95th percentile). We then estimate the regional link density of the randomly distributed links in the same way as for the original link configuration, and repeat this procedure 1,000 times to obtain null-model distributions for each grid cell, which can be used to determine regions in which links form significant bundles. The means and standard deviations of this null-model distribution are shown in Extended Data Fig. 1. Specifically, we consider all grid cells with regional link density above the 99.9th percentile (that is, $P < 0.001$) of the null-model distribution to be part of a significant link bundle (Fig. 3b).

A comparison of the link-distance distribution of all links connected to SCA with the distance distribution of links that remain after correction reveals that the overall shape of the distribution, and in particular the scale break distinguishing regional weather systems from teleconnections, is not affected by our correction technique (red versus blue circles in Extended Data Fig. 2).

Robustness tests. For our statistical analysis, several parameter values need to be determined. In this section we show that our results are robust to variations of the choices presented in the main text. First, we show that we obtain very similar results when defining EREs at the 94th or 96th percentile, instead of the 95th percentile (Extended Data Fig. 7). Furthermore, our results depend only weakly on different choices of maximum delay τ_{max} between synchronous events, namely, $\tau_{max} = 3$ or $\tau_{max} = 30$ instead of $\tau_{max} = 10$ (Extended Data Fig. 8).

To compute the lead–lag correlations in Fig. 4a and determine times of strong synchronizations between SCA and Europe, a Chebycheff type I low-pass filter of order 8 with a cutoff period of 10 days is used to eliminate low-frequency noise. The results remain very similar when changing the cutoff period within a range of 8 to 12 days (Extended Data Fig. 9).

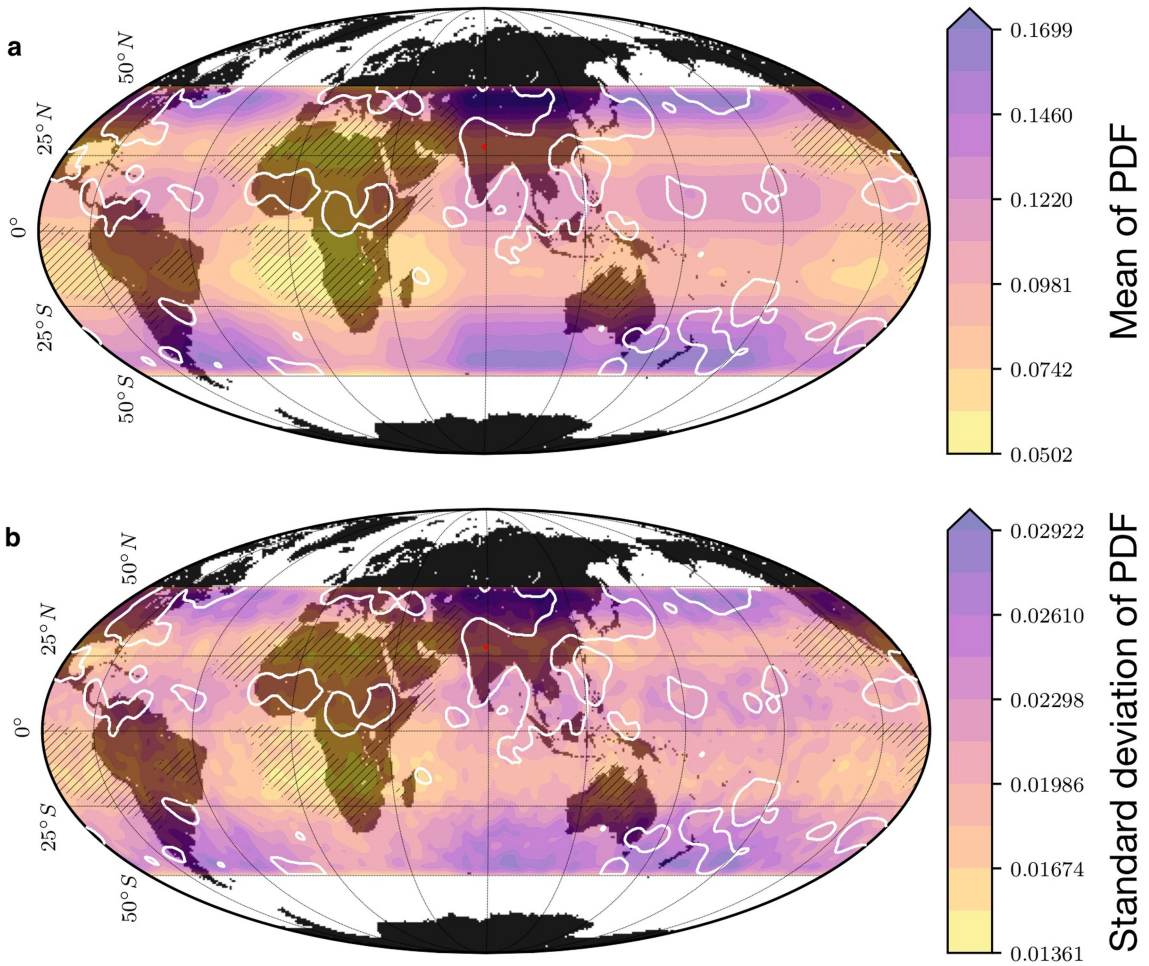
Finally, we test whether our results are independent of the TRMM dataset used here. When repeating the analysis with rainfall data from GPCP[32], we obtain—despite the substantially coarser spatial resolution of 1° instead of 0.25°—a link-distance distribution that closely resembles the one obtained for the TRMM data (compare Fig. 2 and Extended Data Fig. 10), including the scale break at 2,500 km and the super-power-law behaviour for longer distances.

Code availability. The Python code used for the analysis is available on GitHub (https://github.com/niklasboers/rainfall-teleconnections.git).

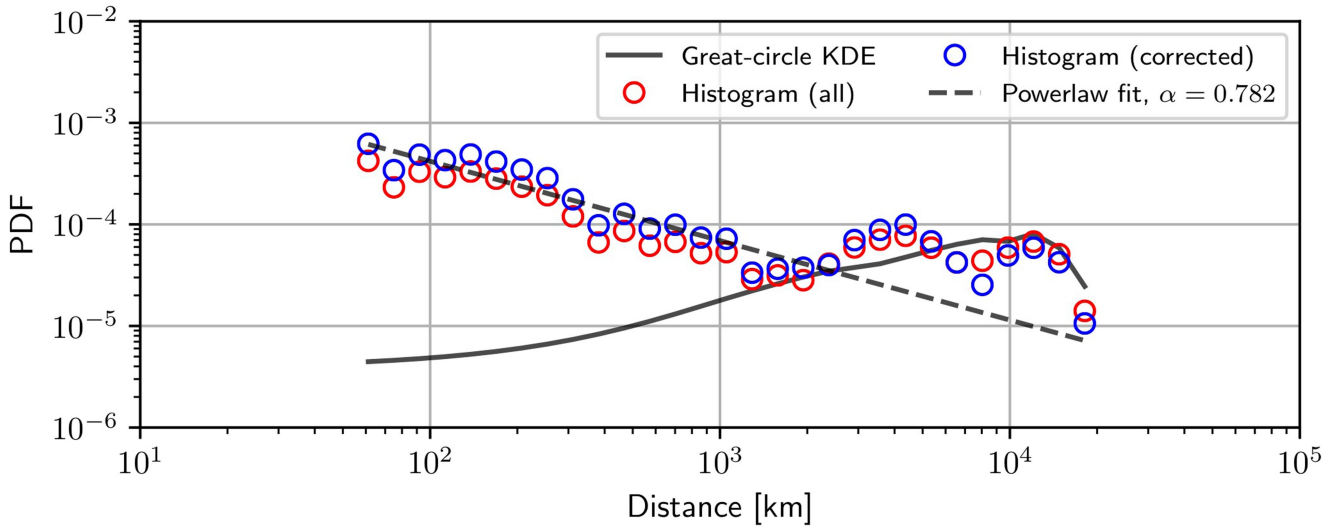
Data availability

The data/reanalysis that support the findings of this study are publicly available online: TRMM 3B42 V7[31], https://pmm.nasa.gov/data-access/downloads/trmm; GPCP 1DD V1.2[32], https://precip.gsfc.nasa.gov/; NCEP/NCAR Reanalysis 1[33], https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html.

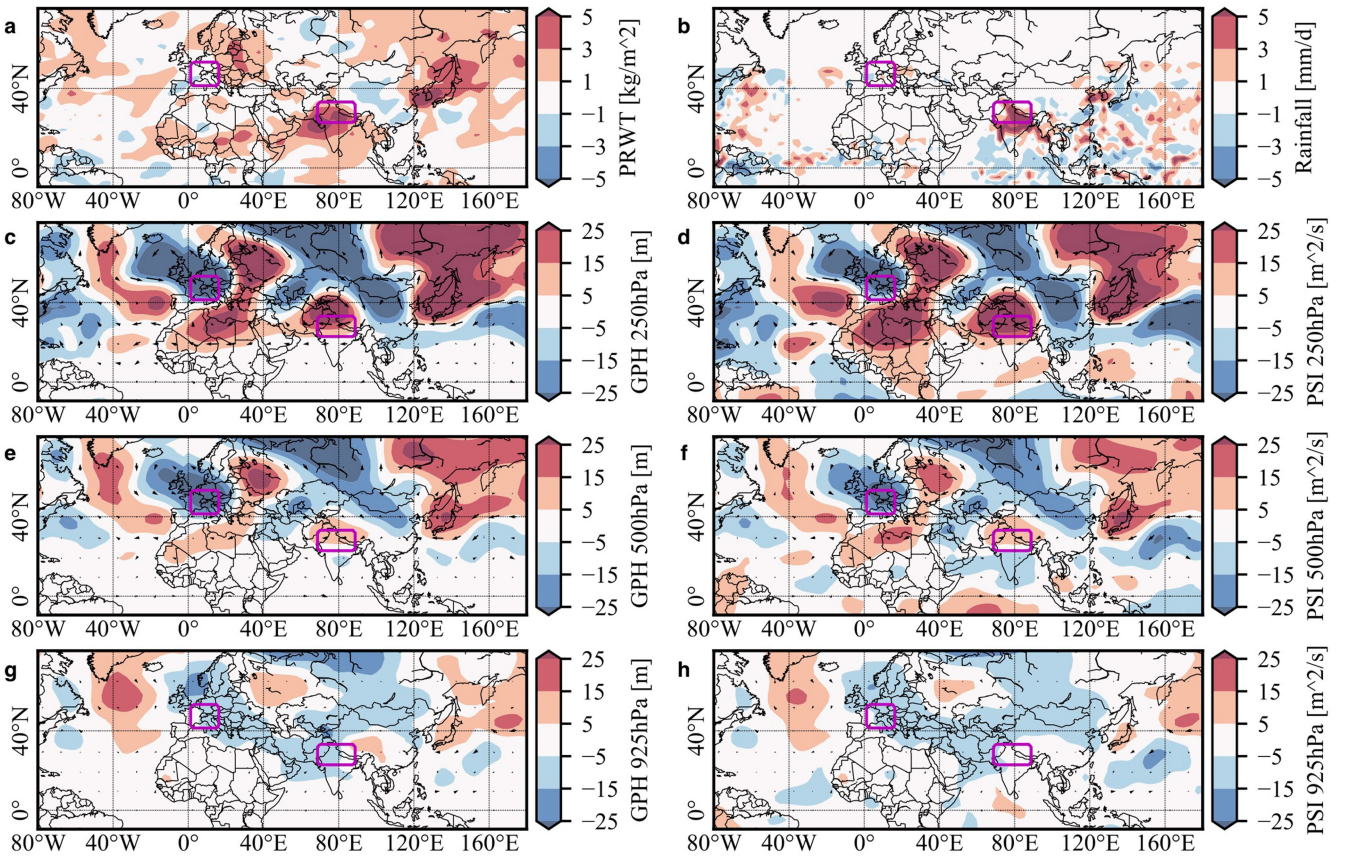
31. Huffman, G. et al. The TRMM Multisatellite Precipitation Analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.* **8**, 38–55 (2007).
32. Huffman, G. J. et al. Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeorol.* **2**, 36–50 (2001).
33. Kalnay, E. et al. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**, 437–471 (1996).
34. Quian Quiroga, R., Kreuz, T. & Grassberger, P. Event synchronization: a simple and fast method to measure synchronicity and time delay patterns. *Phys. Rev. E* **66**, 041904 (2002).
35. Malik, N., Bookhagen, B., Marwan, N. & Kurths, J. Analysis of spatial and temporal extreme monsoonal rainfall over South Asia using complex networks. *Clim. Dyn.* **39**, 971–987 (2012).
36. Boers, N., Bookhagen, B., Marwan, N., Kurths, J. & Marengo, J. Complex networks identify spatial patterns of extreme rainfall events of the South American Monsoon System. *Geophys. Res. Lett.* **40**, 4386–4392 (2013).
37. Rheinwalt, A. et al. Non-linear time series analysis of precipitation events using regional climate networks for the region of Germany. *Clim. Dyn.* **46**, 1065–1074 (2016).
38. Boers, N., Bookhagen, B., Marwan, N. & Kurths, J. Spatiotemporal characteristics and synchronization of extreme rainfall in South America with focus on the Andes Mountain range. *Clim. Dyn.* **46**, 601–617 (2016).
39. Tsonis, A. & Swanson, K. Topology and predictability of El Niño and La Niña networks. *Phys. Rev. Lett.* **100**, 228502 (2008).
40. Yamasaki, K., Gozolchiani, A. & Havlin, S. Climate networks around the globe are significantly affected by El Niño. *Phys. Rev. Lett.* **100**, 228501 (2008).
41. Gozolchiani, A., Havlin, S. & Yamasaki, K. Emergence of El Niño as an autonomous component in the climate network. *Phys. Rev. Lett.* **107**, 148501 (2011).
42. Ludescher, J. et al. Improved El Niño forecasting by cooperativity detection. *Proc. Natl Acad. Sci. USA* **110**, 11742–11745 (2013); correction **110**, 19172–19173 (2013).
43. Wang, Y. et al. Dominant imprint of Rossby waves in the climate network. *Phys. Rev. Lett.* **111**, 138501 (2013).
44. Zhou, D., Gozolchiani, A., Ashkenazy, Y. & Havlin, S. Teleconnection paths via climate network direct link detection. *Phys. Rev. Lett.* **115**, 268501 (2015).
45. Boers, N. et al. Propagation of strong rainfall events from southeastern South America to the central Andes. *J. Clim.* **28**, 7641–7658 (2015).
46. Ebert-Uphoff, I. & Deng, Y. A new type of climate network based on probabilistic graphical models: results of boreal winter versus summer. *Geophys. Res. Lett.* **39**, L19701 (2012).
47. Runge, J. et al. Identifying causal gateways and mediators in complex spatio-temporal systems. *Nat. Commun.* **6**, 8502 (2015).
48. Bonferroni, C. E. Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni del R. Istituto Superiore di Scienze Economiche e Commerciali di Firenze* **8**, 3–62 (1936).
49. Westfall, P. H., Johnson, W. O. & Utts, J. M. A Bayesian perspective on the Bonferroni adjustment. *Biometrika* **84**, 419–427 (1997).
50. Abdi, H. in *Encyclopedia of Measurement and Statistics* (ed. Salkind, N. J.) 103–107 (Sage, Thousand Oaks, 2007).
51. Scott, D. W. *Multivariate Density Estimation: Theory, Practice, and Visualization* (John Wiley & Sons, Hoboken, 1992).



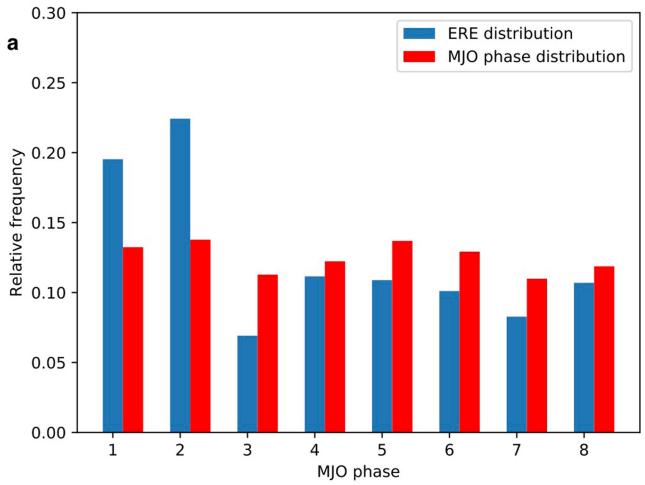
Extended Data Fig. 1 | Means and standard deviations of the null-model distribution for the regional link density. a, Mean of the null-model distribution for the regional link density, obtained by randomly redistributing the links as described in the Methods section 'Significance of spatial patterns'. **b**, Standard deviation of the same null-model distribution as in **a**. The white contour lines indicate regions in which the regional link density of the observations is higher than the 99.9th percentile of the null model. Hatched areas indicate regions with fewer than three events in total, which are excluded from the analysis.

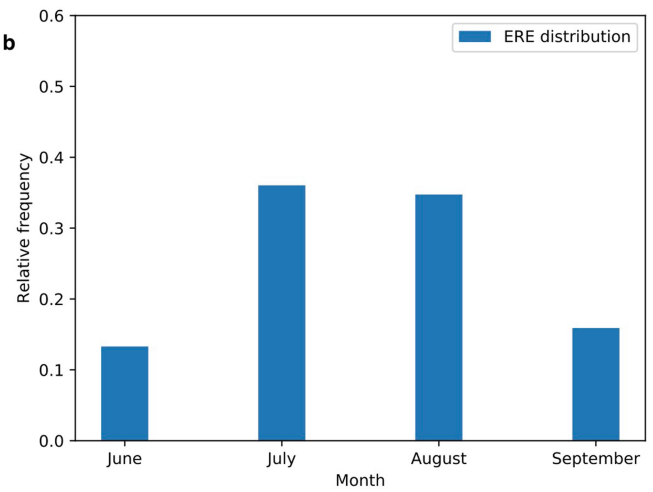


Extended Data Fig. 2 | Distance distribution of extreme-event synchronizations, restricted to SCA. Plot of link distances (red and blue circles) restricted to links attached to SCA (red circles), the power-law fit for the range 100–2,500 km (dashed black line) and the KDE of the distribution of all possible great-circle distances (solid black line). The distance distribution of links that remain after correcting for the multiple-comparison bias (blue circles; see Methods section 'Significance of spatial patterns') resembles the original distribution closely.

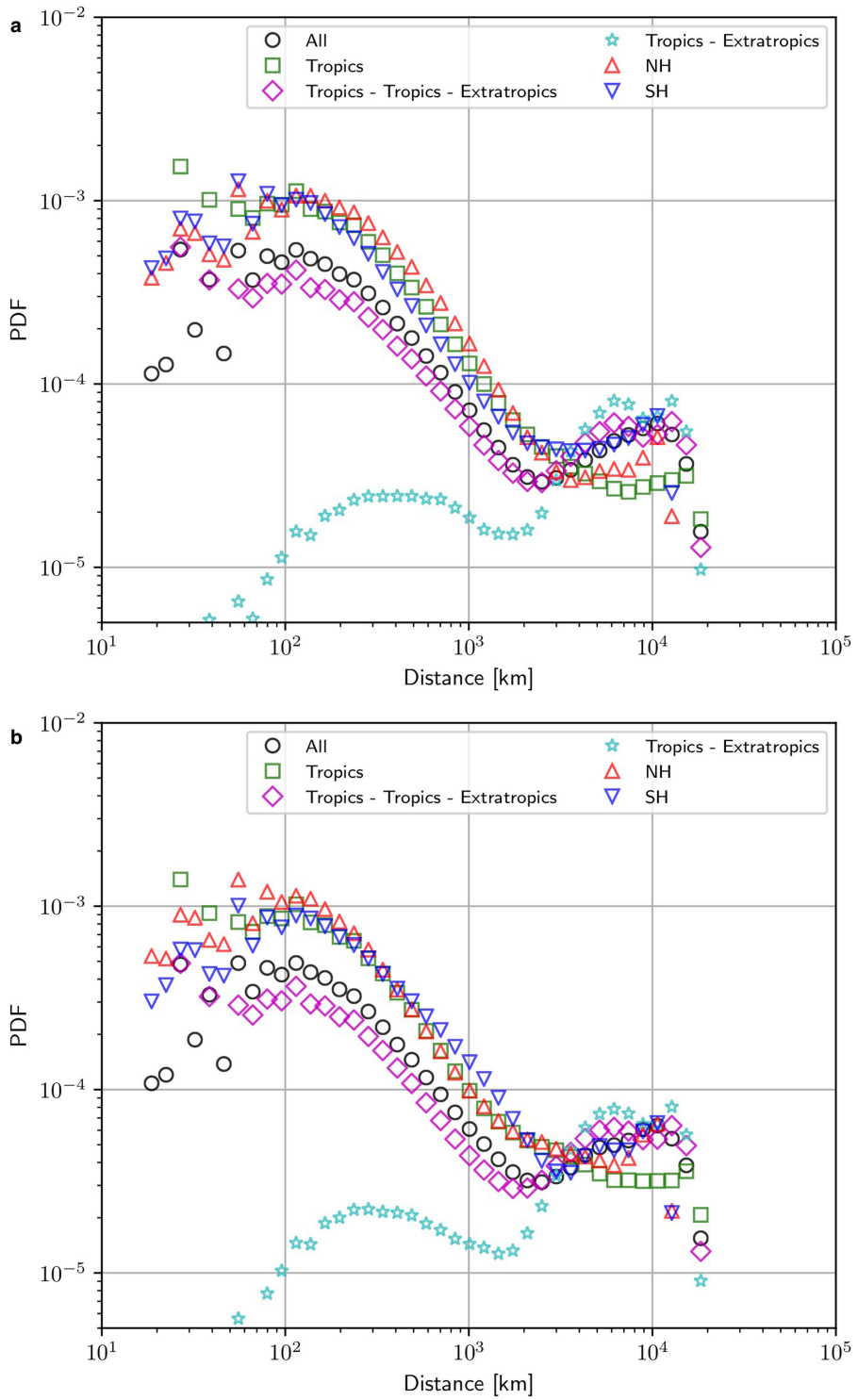


Extended Data Fig. 3 | Extended atmospheric conditions for the teleconnection pattern between Europe and SCA. a–h, Composite anomalies of precipitable water (PRWT; **a**), TRMM rainfall (**b**), geopotential height (GPH; **c**, **e**, **g**) and streamfunction (PSI; **d**, **f**, **h**) at high (**c**, **d**), middle (**e**, **f**) and low (**g**, **h**) atmospheric levels. The anomalies are centred at day 3 after the maximum ERE occurrence in Europe, with significant subsequent counterparts in SCA. The two regions are indicated by magenta boxes.

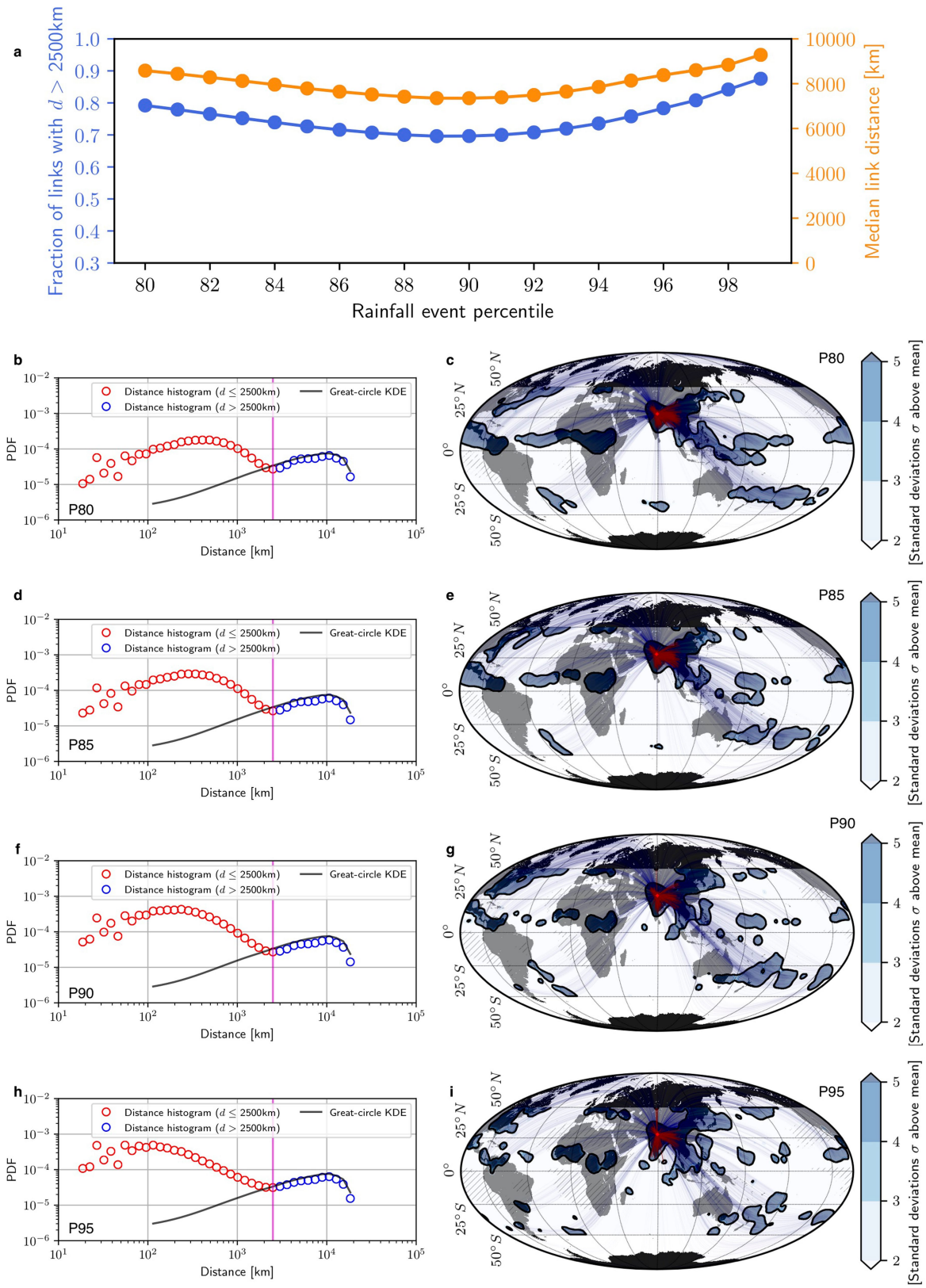




Extended Data Fig. 4 | Dependence of ERE occurrence in SCA on the phase of the MJO and the time of the season. a, Frequency of EREs in SCA over the eight phases of the MJO (blue) and corresponding distribution for all days during the JJA season over these phases (red). We note that EREs in SCA occur predominantly during phases 1 and 2. **b**, Frequency of EREs in Europe with synchronous subsequent counterparts in SCA over the months of June, July, August and September.

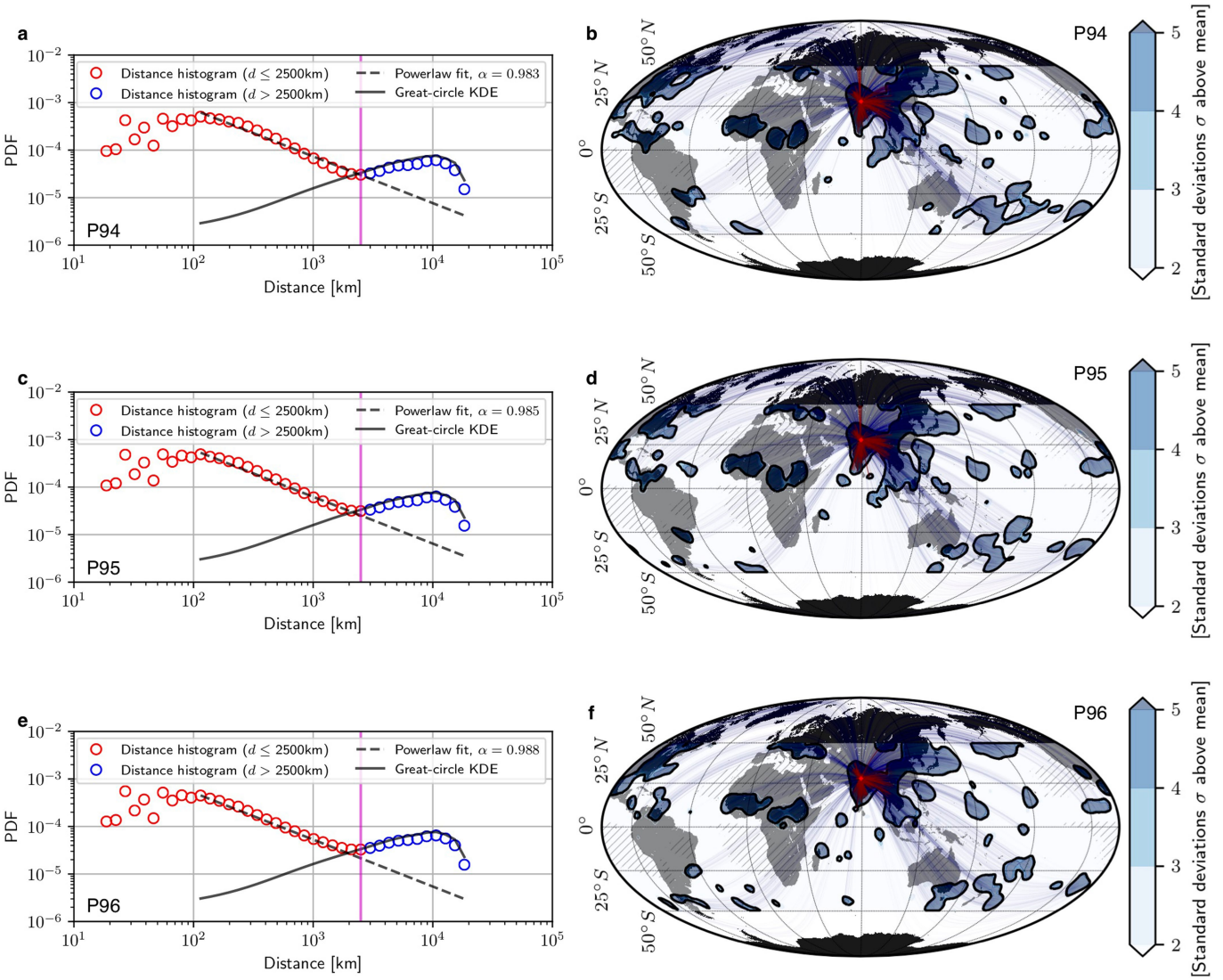


Extended Data Fig. 5 | Distance distributions of extreme-event synchronizations in the tropics and extratropics. a, Link distance distributions for the December–January–February season are shown for the following cases: considering only links within the tropics (green squares), links within the tropics and to the extratropics of both hemispheres (magenta diamonds), links connecting the tropics and the extratropics of both hemispheres (cyan stars), links within the Northern Hemisphere (NH; upward red triangles) and links within the Southern Hemisphere (SH; downward blue triangles). **b**, Same as **a**, but for the JJA season. The distance distributions for all links taken together (as in Fig. 2) are indicated by black circles in both panels. We note that the super-power-law part of the distribution (that is, the part of the distribution for distances longer than 2,500 km) is substantially suppressed if the analysis is restricted to the global tropics (green squares), whereas it remains strong if links to the extratropics are included (magenta diamonds). In particular, the super-power-law part is much more pronounced if only the links connecting the tropics with the extratropics are considered, compared with the distribution of all links (black circles). We also note that the distributions are very similar for both seasons.

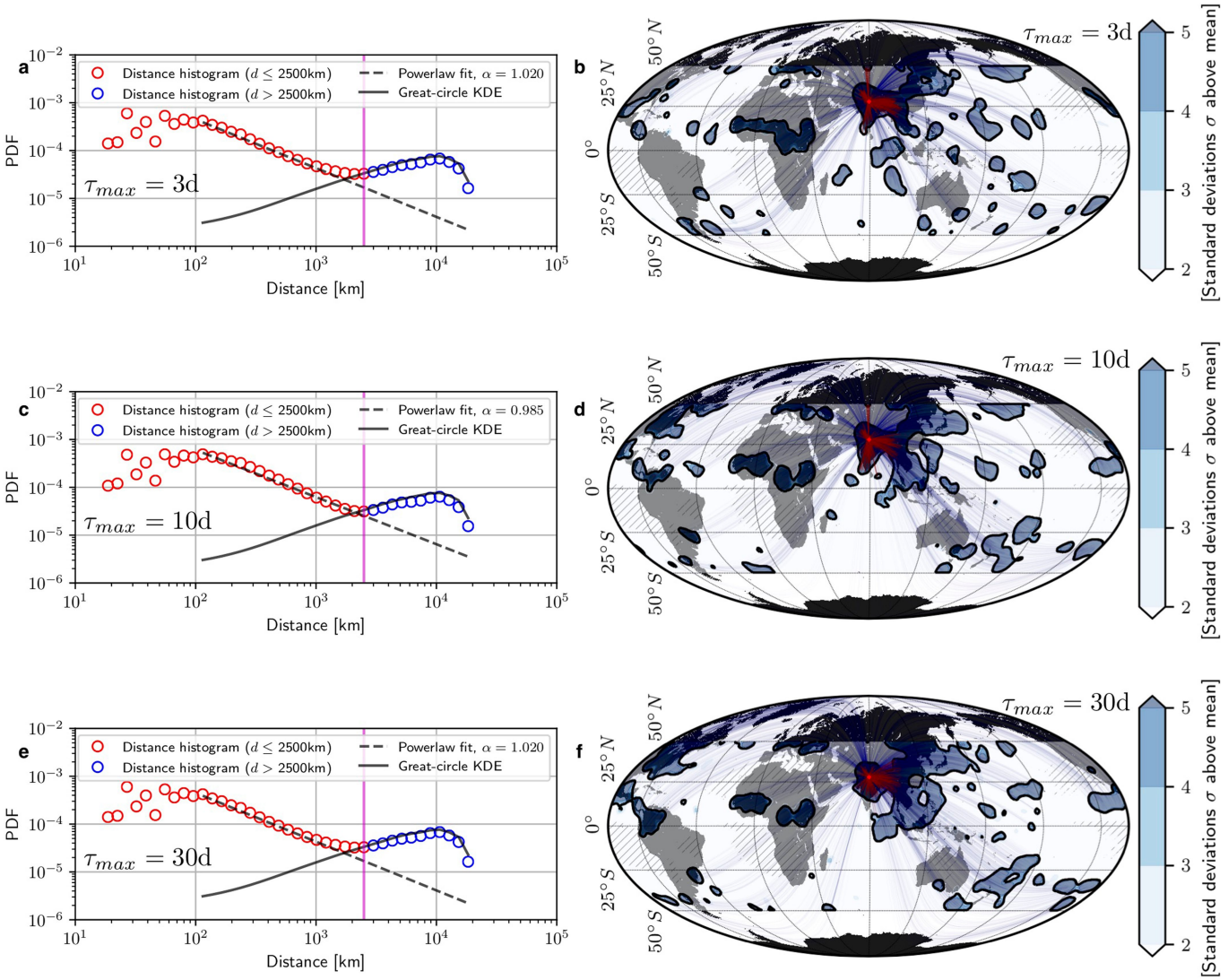


Extended Data Fig. 6 | See next page for caption.

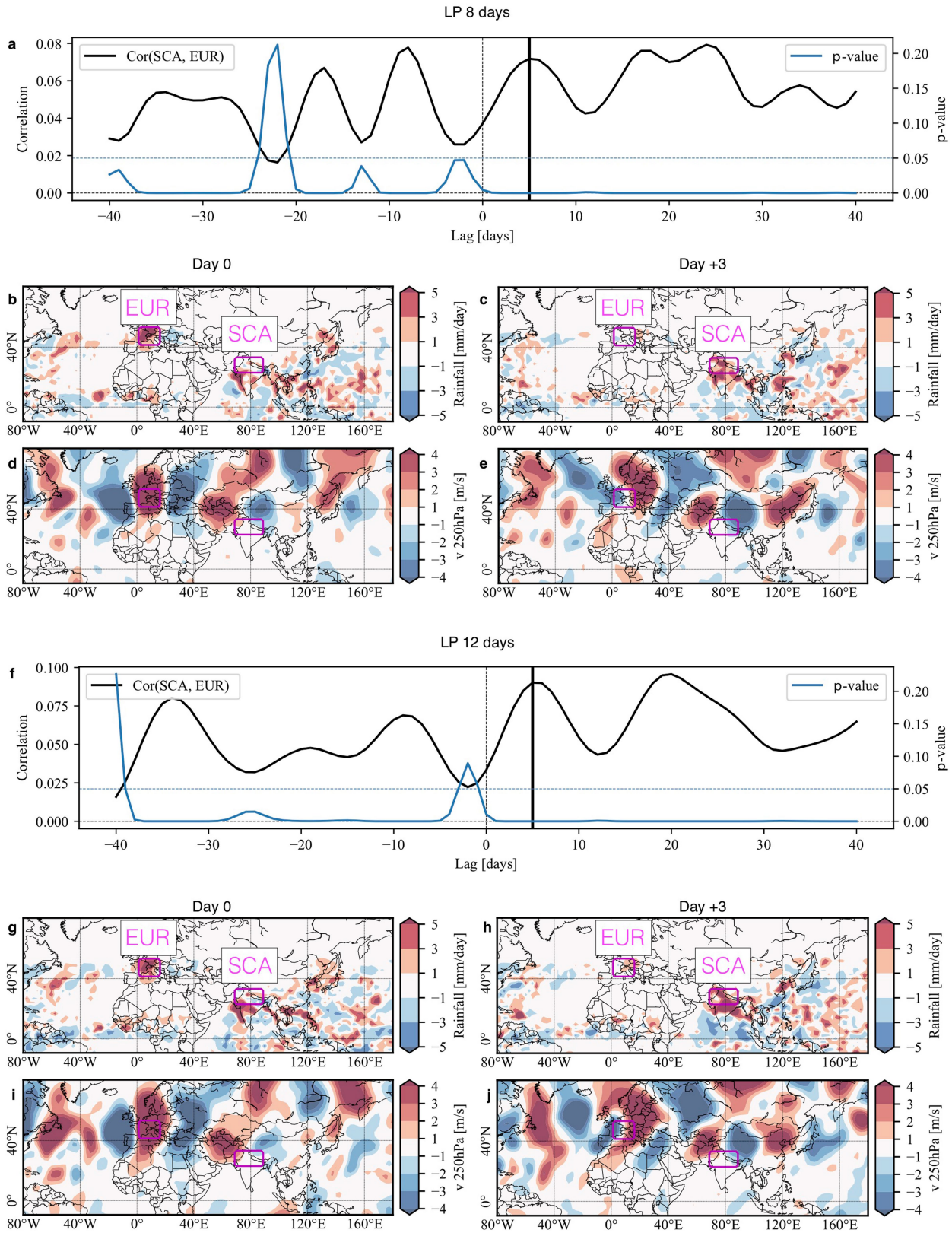
Extended Data Fig. 6 | Comparison of the global distance distribution and teleconnection patterns in SCA for different event thresholds. **a**, Fraction of links longer than 2,500 km (blue) and median of the link-distance distribution (orange) for different event percentile thresholds. Both quantities remain similar over the range of thresholds, with slight increases towards the strongest events. **b**, **d**, **f**, **h**, The corresponding distance distributions are shown for the 80th (**b**), 85th (**d**), 90th (**f**) and 95th (**h**) percentiles for comparison. **c**, **e**, **g**, **i**, Significant link bundles attached to SCA are also shown for events above the 80th (**c**), 85th (**e**), 90th (**g**) and 95th (**i**) percentiles. Links shorter (longer) than 2,500 km are shown in red (blue). A spherical Gaussian KDE of the regional link density, in combination with a null model of randomly distributed links, is used to determine link bundles; links that are not part of significant bundles are omitted (Methods section 'Significance of spatial patterns'). Significant link bundles are shown by blue contours in units of standard deviations above the mean. The mean and the standard deviation are inferred from the null model of the regional link density. The black contour line delineates areas in which the regional link density is higher than the 99.9th percentile of the null-model distribution.



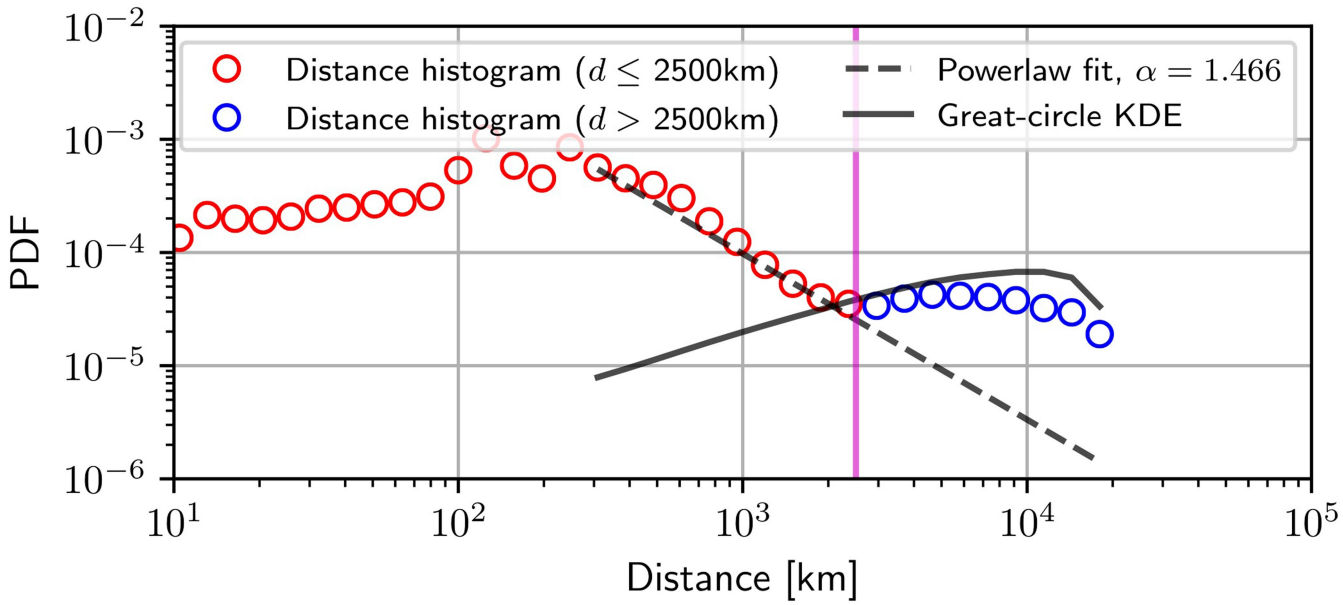
Extended Data Fig. 7 | Distance distribution and teleconnection pattern in SCA for different extreme-event thresholds. a, **c**, **e**, Plots of link distances (red and blue circles), power-law fits for the range 100–2,500 km (dashed black lines), and KDEs of the distribution of all possible great-circle distances (solid black lines) for EREs above the 94th (**a**), 95th (**c**) and 96th (**e**) percentile. The vertical line at $d = 2{,}500$ km marks the regime shift from regional weather systems to large-scale teleconnections. We note that the power-law exponent remains very similar over this range, indicating that the $1/d$ decay of the distance distribution is robust. **b**, **d**, **f**, Link bundles attached to SCA are shown for EREs above the 94th (P94; **b**), 95th (P95; **d**) and 96th (P96; **f**) percentile, after correcting for the multiple-comparison bias. Links shorter (longer) than 2,500 km are shown in red (blue). A spherical Gaussian KDE of the regional link density, in combination with a null model of randomly distributed links, is used to determine link bundles; links that are not part of significant bundles are omitted (Methods section 'Significance of spatial patterns'). Significant link bundles are shown by blue contours in units of standard deviations above the mean. The mean and the standard deviation are inferred from the null model of the regional link density. The black contour lines delineate areas in which the regional link density is higher than the 99.9th percentile of the null-model distribution.



Extended Data Fig. 8 | Distance distribution of extreme-event synchronizations and teleconnection pattern in SCA for different values of τ_{max}. a, **c**, **e**, Plots of link distances (red and blue circles), power-law fits for the range 100–2,500 km (dashed black lines) and KDEs of the distribution of all possible great-circle distances (solid black lines) for $\tau_{max} = 3$ days (**a**), $\tau_{max} = 10$ days (**b**) and $\tau_{max} = 30$ days (**c**). The vertical line at $d = 2{,}500$ km marks the regime shift from regional weather systems to large-scale teleconnections. We note that the distribution of significant link distances below 2,500 km (red circles) decays slightly faster for $\tau_{max} = 3$ days than for $\tau_{max} = 10$ days or $\tau_{max} = 30$ days, implying that 3 days are not sufficient to capture the entire global-scale teleconnection pattern. **b**, **d**, **f**, Link bundles attached to SCA are shown for $\tau_{max} = 3$ days (**b**), $\tau_{max} = 10$ days (**d**) and $\tau_{max} = 30$ days (**f**), after correcting for the multiple-comparison bias. Links shorter (longer) than 2,500 km are shown in red (blue). A spherical Gaussian KDE of the regional link density, in combination with a null model of randomly distributed links, is used to determine link bundles; links that are not part of significant bundles are omitted (Methods section 'Significance of spatial patterns'). Significant link bundles are shown by blue contours in units of standard deviations above the mean. The mean and the standard deviation are inferred from the null model of the regional link density. The black contour lines delineate areas in which the regional link density is higher than the 99.9th percentile of the null-model distribution.



Extended Data Fig. 9 | Atmospheric conditions for the teleconnection pattern between Europe and SCA for different cutoff values of the low-pass filter. a, Lead–lag correlations (solid black line) of timeseries obtained from spatially averaging the daily numbers of EREs in boxes in Europe (EUR; 42° N to 50° N, 3° E to 15° E) and SCA. The timeseries are low-pass-filtered (LP) at a cutoff period of 8 days (Methods section 'Identification of specific times with high synchronization'). **b**, Composite anomalies of TRMM rainfall for days with high numbers of EREs in Europe that are followed by associated EREs in SCA. **c**, Same as **b**, but 3 days later. **d**, Composite anomalies of the meridional wind component *v* at 250 hPa for the same time steps as in **b**. **e**, Same as **d**, but 3 days later. **f**–**j**, Same as **a**–**e**, but for a cutoff of 12 days.



Extended Data Fig. 10 | Distance distribution of extreme-event synchronizations based on the GPCP instead of the TRMM dataset. Plot of link distances (red and blue circles), power-law fit for the range 100 km–2,500 km (dashed black line) and KDE of the distribution of all possible great-circle distances (solid black line) for EREs above the 95th percentile, derived from the GPCP instead of the TRMM dataset. The vertical line at $d = 2{,}500$ km marks the regime shift from regional weather systems to large-scale teleconnections. We note that in contrast to the TRMM data, the GPCP data extend to latitudes λ beyond 50°. The spatial distance that corresponds to a resolution of 1° scales with (111 km) × cos(λ), and therefore the distance distribution includes distances below 100 km. The part of the distribution that is relevant for comparison with that obtained from the TRMM data (Fig. 2) is for distances above 100 km. For smaller distances, a bias exists owing to the very small distances between grid cells near the poles.