Complex Network Analysis for Exploring Teleconnection Structures

From Global Patterns to Local Propagation Pathways

Dissertation
der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Felix Strnad, M. Sc.
aus Freiburg im Breisgau

Tübingen
2024

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	03. Dezember 2024
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Dr. Bedartha Goswami
2. Berichterstatter:	Prof. Dr. Ulrike von Luxburg

Für Jonah, Julia und Ompf

Disclaimer: This thesis uses Federico Marotta's `kaobook` template based on Ken Arroyo Ohori's doctoral thesis. I am grateful to Frank Schneider, Felix Dangel and Jakob Schlör who added many stylistic features.

Acknowledgments

This thesis would not have been possible without the support of many people who have accompanied me on my journey through the world of research at the intersection point of machine learning and climate science. I would like to express my deepest gratitude to all of them.

First and foremost, I would like to thank my first supervisor, Bedartha Goswami, for your continuous support and guidance throughout my doctoral studies. Your enthusiasm for statistical approaches to climate science has been a great inspiration to me. I am grateful for the freedom you gave me to explore my research interests and for keeping me motivated also in critical phases during this project.

I would like to thank the members of my Thesis Advisory Committee (TAC), Ulrike von Luxburg and Caterina de Bacco. Your expertise in complex network theory has been a great asset to my work. I am grateful for the many stimulating discussions we had and for the new perspectives, you brought to my research. Further, I would like to thank Jakob Macke for being part of the examination committee.

Thank you to Niklas Boers, for your valuable input and feedback on my research throughout my whole PhD project. I am grateful for the many inspiring discussions we had and for the collaborative projects we worked on together.

This work would not have been possible without the input obtained from great collaborators. Here, I want to thank especially my co-authors Ruth Geen and Kieran Hunt for their valuable contributions. Thank you for opening my eyes to the world of meteorology and teaching me how to interpret climate data. Thank you Ruth as well for allowing me to stay three months with you at the University of Birmingham, I learned so much during this time.

I was so lucky to have had the best colleagues one could image. Thank you Jakob Schlör, Jannik Thümmel, Jan Prosi and Moritz Haas of the Machine Learning in Climate Science group at the University of Tübingen for your day-to-day support, encouragement and the many discussions we had during lunch and various coffee breaks. We started as colleagues and ended up as friends.

I enjoyed supervising several students during my Ph.D. project. Thank you Merle Kammer, Julia Hellmig, Hendrik Krass and Dexter Früh. I hope you learned as much from me as I did from you.

I would like to thank Sebastian Schwenk, Elena Siziliani, and Shyamala Subramanian of the University of Tübingen, and Leila Masri, Heike Größl and Sara Sorce of the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for providing me support with all the administrative issues.

I thank all the people who have proofread this thesis, in no particular order Julia Brugger, Jakob Schlör, Jannik Thümmel, Moritz Haas, Alina Sage, Jakob Wedemeyer, Anna Brugger and Philipp Hess. All remaining errors are my own.

I want to thank my family, especially my parents, my brothers and my grandparents for their love, trust and support over decades. You have been always there whenever I needed you.

Finally, I'm grateful to Julia for so many things that I can't even list them all. Thank you for your love, your patience and for being such a wonderful person. I am indescribably happy about our son Jonah, who was born during this PhD time and who fills every day with so much joy.

Felix Strnad
Tübingen, September 15, 2024

Abstract

The term "teleconnections" is used to describe relationships between weather phenomena or climate patterns occurring at widely separated locations on Earth. The understanding of such connections enables the formulation of predictions with a lead time of several weeks to months which allows a better preparation for extreme events. A specific focus of this thesis is set on the investigation of teleconnection structures of extreme rainfall events (EREs) within the Indo-Pacific Monsoon domain. EREs are a major source of natural disasters in this region and are often associated with severe flooding, landslides, and loss of life and property. A better understanding of the teleconnection structures of EREs is crucial for the development of early warning systems in the region. However, the complexity of the climate system and the multitude of processes involved in the formation of teleconnections make it challenging to identify teleconnection structures and understand their underlying mechanisms.

This work uses so-called climate networks, derived from complex network science, to capture teleconnection structures. Climate networks are constructed by representing the climate system as a network of nodes representing grid points on Earth which are connected by links derived from statistical relationships between climate variables. The topology of the network allows conclusions about the underlying physical processes forming teleconnection patterns. In this work, two different network measures are introduced to quantify the strength and spatial organization of teleconnection patterns. The first measure to be drawn, the *curvature* approach, is based on the concept of detecting the structure of the most relevant bundles of links in the network. The second measure, the *community detection approach*, is based on the concept of detecting groups of densely connected nodes in the network. These two measures are applied in different studies:

i) The first study presents the curvature approach, which is an intuitive way to visualize differences between regional links and long-range teleconnection structures within climate networks. This study employs this to investigate differences in teleconnection patterns of the El Niño Southern Oscillation (ENSO). The analysis suggests that El Niño events with positive sea surface temperatures (SSTs) anomalies in the Eastern Pacific (EP) reveal teleconnection patterns primarily confined to the Tropics, whereas the teleconnection patterns of El Niño events with positive SST anomalies in the Central Pacific (CP) show teleconnections to the Subtropics and Midlatitudes as well.

ii) The second study investigates the teleconnection structures of EREs within the Indo-Pacific Monsoon domain. The community detection approach reveals that the spatial organization of EREs is dominated by the Boreal Summer Intraseasonal Oscillation (BSISO), a quasi-periodic north-eastward movement of convective precipitation from the Indian Ocean to the Western Pacific. This study reveals different local propagation pathways of EREs within the Indo-Pacific Monsoon domain and uncovers the El Niño Southern Oscillation (ENSO) as the key driver in modulating these pathways.

iii) The third study analyzes the synchronization between two of the most pronounced monsoon systems on Earth, the Indian and the West African Monsoon. This work identifies a teleconnection structure between these two monsoon systems that exists on intraseasonal timescale. The teleconnection is modulated by a variety of different atmospheric processes.

Taken together, the three studies included in this thesis present two different ways to intuitively visualize the most dominant teleconnections for the area under investigation. Using these insights our studies contribute to the understanding of teleconnection structures from local to global scales and provide a new perspective on the mechanisms that govern the yearly variability in the occurrences of EREs.

Zusammenfassung

Der Begriff “Teleconnections” wird verwendet, um Wechselwirkung zwischen Wetterphänomenen zu beschreiben, die an weit voneinander entfernten Orten auf der Erde auftreten. Das Verständnis solcher Zusammenhänge ermöglicht Wettervorhersagen in einem Zeitraum von mehreren Wochen bis Monaten, was eine bessere Vorbereitung auf Extremereignisse ermöglicht. Ein besonderer Schwerpunkt dieser Arbeit liegt auf Strukturen von Teleconnections von extremen Niederschlagsereignissen (EREs) im Bereich des indopazifischen Monsuns. EREs sind eine der Hauptursachen für Naturkatastrophen in der Region und gehen häufig mit schweren Überschwemmungen, Erdrutschen sowie dem Verlust von Menschenleben einher. Ein besseres Verständnis der Teleconnections von EREs ist für die Entwicklung von Frühwarnsystemen in der Region von hoher Bedeutung. Die Komplexität des Klimasystems und die Vielzahl der Prozesse, die an der Bildung von Teleconnections beteiligt sind, machen es jedoch schwierig, Teleconnections zu identifizieren und die ihnen zugrunde liegenden Mechanismen zu verstehen.

In dieser Arbeit werden sogenannte Klimanetzwerke verwendet, die aus dem Feld der komplexen Netzwerke hergeleitet werden, um Strukturen von Teleconnections besser zu erfassen. Klimanetzwerke werden konstruiert, indem das Klimasystem als ein Netzwerk von Knotenpunkten dargestellt wird, die Gitterpunkte auf der Erde repräsentieren und Verbindungen zueinander haben, die aus statistischen Beziehungen zwischen Klimavariablen abgeleitet werden. Die Topologie des Netzwerks erlaubt Rückschlüsse auf die zugrundeliegenden physikalischen Prozesse, die Muster aus Teleconnections bilden. Wir führen zwei verschiedene Netzwerkmaße ein, um die Stärke und räumliche Organisation von Teleconnections zu quantifizieren. Das erste Maß, der *Curvature*-Ansatz, basiert auf dem Konzept, die Struktur der relevantesten Bündel von Verbindungen im Netzwerk zu erkennen. Das zweite Maß, der *Community Detection approach*, basiert auf dem Konzept der Erkennung von Gruppen dicht verbundener Knoten im Netzwerk. Diese beiden Maße werden in unterschiedlichen Studien angewandt:

i) In der ersten Studie benutzt den Curvature-Ansatz, mit dem Unterschiede zwischen regionalen Verbindungen und weitreichenden Teleconections in Klimanetzwerken intuitiv sichtbar gemacht werden können. Die Analyse deutet darauf hin, dass El-Niño-Ereignisse mit anormal hohen Oberflächentemperaturen im östlichen Pazifik (EP) vorwiegend Teleconections innerhalb der Tropen aufweist, während die Teleconections von El-Niño-Ereignissen mit anormal hohen Oberflächentemperaturen im zentralen Pazifik (CP) auch Teleconections zu den Subtropen und den mittleren Breiten aufweisen.

ii) Die zweite Studie untersucht die Teleconections von EREs innerhalb des indo-pazifischen Monsungebietes. Der Community-Detection-Ansatz zeigt, dass die räumliche Organisation von EREs von der sogenannten Boreal Summer Intraseasonal Oscillation (BSISO) dominiert wird, einer annähernd periodisch wiederkehrenden Oszillation von Niederschläge, die sich vom Indischen Ozean in nordöstlicher Richtung bis in den westlichen Pazifik bewegt. Diese Studie zeigt verschiedene lokale Ausbreitungswege von EREs innerhalb des indo-pazifischen Monsuns auf und kann aufzeigen, dass die El-Niño-Südliche Oszillation (ENSO) die Hauptursache für die Veränderung dieser Wege ist.

iii) Die dritte Studie analysiert die Synchronisation zwischen zwei der bekanntesten Monsunsysteme der Erde: dem indischen und dem westafrikanischen Monsun. Diese Arbeit findet eine Verbindung zwischen diesen beiden Monsun-Systemen, die auf in einem Zeitraum innerhalb des Sommers existiert. Diese Teleconections wird durch eine Vielzahl unterschiedlicher atmosphärischer Prozesse moduliert.

Insgesamt zeigen die drei Studien in dieser Arbeit zwei verschiedene Möglichkeiten auf, die wichtigsten Teleconections für das untersuchte Gebiet intuitiv zu darzustellen. Dadurch tragen unsere Studien zum Verständnis der Teleconections von der lokalen bis zur globalen Ebene bei und bieten eine neue Perspektive auf die Mechanismen, die die jährliche Variabilität des Auftretens von Extremereignissen bestimmen.

List of publications

P1: **F. M. Strnad**, J. Schlör, C. Fröhlich, and B. Goswami (2022). 'Teleconnection Patterns of Different El Niño Types Revealed by Climate Network Curvature'. In: *Geophysical Research Letters* 49.17. DOI: `10.1029/2022GL098571`

P2: **F. M. Strnad**, J. Schlör, R. Geen, N. Boers, and B. Goswami (2023). 'Propagation pathways of Indo-Pacific rainfall extremes are modulated by Pacific sea surface temperatures'. In: *Nature Communications* 14.5708. DOI: `10.1038/s41467-023-41400-9`

P3: **F. M. Strnad**, K. M. R. Hunt, N. Boers, and B. Goswami (2024). 'Intraseasonal synchronization of extreme rainfalls between North India and the Sahel'. In: *in review at Quarterly Journal of the Royal Meteorological Society*. DOI: `10.48550/arXiv.2405.08492`. eprint: `2405.08492`

P4: P. Hess, M. Drüke, S. Petri, **F. M. Strnad**, and N. Boers (2022). 'Physically constrained generative adversarial networks for improving precipitation fields from Earth system models'. In: *Nature Machine Intelligence* 4. DOI: `10.1038/s42256-022-00540-1`

P5: J. Schlör, **F. M. Strnad**, A. Capotondi, and B. Goswami (2024). 'Contribution of El Niño Southern Oscillation (ENSO) Diversity to Low-Frequency Changes in ENSO Variance'. In: *Geophysical Research Letters* 51.14. DOI: `10.1029/2024GL109179`

P6: V. Lembo, S. Bordoni, E. Bevacqua, D. I. V. Domeisen, C. L. E. Franzke, V. M. Galfi, C. Garfinkel, C. I. Grams, A. Hochman, R. Jha, K. Kornhuber, F. Kwasniok, V. Lucarini, G. Messori, D. Pappert, I. Perez-Fernandez, J. Riboldi, E. Russo, T. A. Shaw, I. Strigunova, **F. M. Strnad**, P. Yiou, and N. Zagar (2024). 'Dynamics, statistics and predictability of Rossby waves, heatwaves and spatially compounded extreme events'. In: *Bull. Am. Meteorol. Soc.* -1.aop. DOI: `10.1175/BAMS-D-24-0145.1`

Contents

1.

Introduction

1.1. Motivation

Short-term variations in weather patterns around the globe that occur within a season impact the daily livelihood of people worldwide. They are often considered to be local phenomena, but are in fact due to large-scale circulation processes in the atmosphere that dynamically connect climate variables like temperature or precipitation across long spatial distances. These long-range links between remote regions are referred to as "*teleconnections*" (Trenberth, 2022) implying causal relationships between weather patterns in the tropics, subtropics and mid-latitudes (Di Capua et al., 2020). Understanding the mechanisms governing these teleconnections within the global circulation structure is crucial, as they are, for instance, responsible for the occurrence of extreme events (Lembo et al., 2024). Teleconnections are the result of the propagation of atmospheric or oceanic signals from one region to another region of the world as recurring and persistent patterns (Trenberth, 1997), which together give predictability in the climate system beyond the stochastic limit of numerical weather prediction of about 2 weeks (Zhang et al., 2019). In a next step, this can be used for the development of seasonal to sub-seasonal (S2S) forecasts and the prediction of extreme events (Vitart and Robertson, 2018).

In this work, we are especially interested in teleconnection structures that connect the occurrence of Extreme Rainfall Events (EREs), defined as days with precipitation exceeding a certain threshold of the daily precipitation distribution (Boers et al., 2019). EREs can cause devastating floods affecting infrastructure, agriculture, and ecosystems and their strong implications for the global economy are well documented (Kotz et al., 2022). The occurrences of EREs are globally related to teleconnection structures (Boers et al., 2019) but the current understanding of teleconnections of EREs is limited, as the occurrences of these events involve intrinsic stochasticity and the binary time series are more difficult to analyze from a practical point of view.

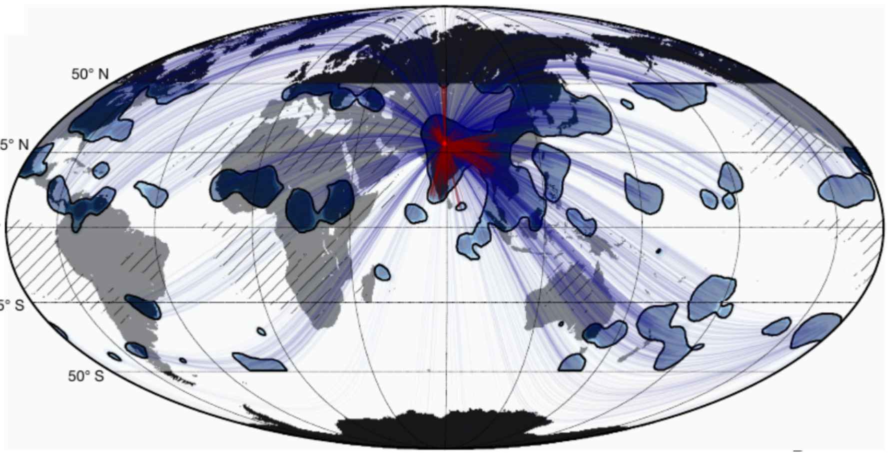


Figure 1.1.: Teleconnection pattern for North India for synchronous EREs Teleconnections of the North Indian region (denoted by the red dot) are shown by blue contours. Black contours visualizes statistically significant teleconnections. The statistical significance is accessed by comparing the local link density to a null model of randomly distributed links which has to be above the 99.9th percentile of the null model. The thin lines are for visualization purposes only, where blue visualizes long-range teleconnections and red regional links. The locations with only few wet days are indicated by the hatched contours and are excluded. The figure has been adapted from Boers et al., 2019.

The domain of the South Asian Summer Monsoon (SASM) of which the Indian Summer Monsoon (ISM) is part of, is particularly vulnerable to EREs, which often result in damage to both rural and urban facilities impacting agricultural production and the economy for many years (Gadgil and Gadgil, 2006). The vulnerability of society to EREs increases due to population density and land use development, necessitating the establishment of reliable early warning systems for EREs for the affected regions (Gallé and Katzenberger, 2024). A better understanding of the relevant teleconnection mechanisms can help here. Figure 1.1 visualizes teleconnection patterns of EREs for the North Indian Monsoon domain (Boers et al., 2019). In recent years, many of these patterns could be explained. For example the synchronization of North India with the Yellow River Basin in North China (Gupta et al., 2022), or to central Europe (Beverley et al., 2021) but many patterns remain poorly understood, for example the synchronization with the Maritime Continent (this is the group of islands in Southeast Asia, including the Philippines, Indonesia and Papua New Guinea) or with the Sahel region.

Teleconnections reflect changes in the atmospheric circulation, driven by large-scale modes of variability, such as the El Niño Southern Oscillation (ENSO), a coupled ocean-atmosphere phenomenon that originates in the tropical Pacific, or the Madden-Julian Oscillation (MJO) (Madden and Julian, 1971), a phenomenon that occurs most prominently in the Indo-Pacific region. These modulate the atmospheric dynamics by reinforcing or weakening the flow of air and the associated transport of moisture and energy through the atmosphere (Hoskins, 2013; Stott et al., 2016). Historically the field of climate science and meteorology mainly identified teleconnections via a Principal Component Analysis (PCA) (in climate science often referred to as Empirical Orthogonal Functions (EOF)-Analysis) where the most dominant modes reveal the spatial locations that show the same variability in time (Trenberth, 1997). Since PCA or related methods are linear time series analysis approaches, they often fail to identify teleconnection structures for less dominant teleconnection patterns as the Earth's climate system - the atmosphere, hydrosphere, cryosphere, lithosphere and biosphere - and their mutual interactions and feedback mechanisms form a highly non-linear complex system (Gulev et al., 2021). Teleconnections typically involve interactions between these components that potentially occur across multiple timescales which can obscure the atmospheric signals (Stan et al., 2017) and makes most teleconnection patterns to inherently non-linear processes.

Climate networks are one effective possibility to address this issue. This approach employs complex network theory (Amaral, 2022), a field of mathematics that investigates the characteristics of complex dynamical systems by representing them as a graph comprised of nodes and edges (Strogatz, 2001; Watts, 2004). In climate networks nodes represent different regions of the Earth's surface, and edges represent the strength of the statistical relationship between the time series of climate variables. The topology of these networks, i.e. the arrangement and interconnections of nodes and edges within the network graph, is assumed to reflect the driving physical mechanisms (Dijkstra et al., 2019). Exploring the topology provides a non-linear time series analysis tool, allowing conclusions to be drawn that may not be readily accessible using conventional approaches (Boers et al., 2021).

In numerous articles in recent years, it has been shown that the application of climate networks is well-suited to uncover hidden teleconnection structures (for example Boers et al., 2019; Tsonis and Swanson, 2008; Donges et al., 2009a; Malik et al., 2010; Ludescher et al., 2014; Boers et al., 2014; Stolbova et al., 2016; Gupta et al., 2021; Liu et al., 2023). However, most of the existing studies have focused on single patterns of the network than on the analysis of the overall network topology itself, and the physical interpretation of the network measures is often limited and the question of why certain patterns are formed remains unclear. Further, there are many climate network based studies that neglect that local teleconnection structures are often themselves influenced by global teleconnection structures. For example, a strong ENSO phase can enhance local teleconnection patterns of the Indian Summer Monsoon (Kumar et al., 2006). Or even ENSO teleconnections are known to be modified by the Pacific Decadal Oscillation on interannual timescales (Trenberth, 2022). These examples show that a global teleconnection which is particularly strong can amplify local teleconnection signals, thereby causing regional patterns to become more (or less) pronounced implying that local teleconnection patterns are not constant but can change over time depending on the strength of the global teleconnection patterns.

1.2. Scope and Content

The above discussion on teleconnections condenses into two key questions:

Key Questions

1. How can the spatial structures of teleconnections be uncovered?
2. How are local teleconnection patterns influenced by global teleconnection structures?

These questions have guided this PhD project over the last four years and required an intensive examination of the spatial structures of teleconnections. The study of teleconnections in the climate system is a field of active research in climate science, and one single thesis alone cannot claim to be able to answer the above questions conclusively. However, this work aims to provide a new perspective on these questions by combining methods from climate science, meteorology, network theory, and machine learning. A diverse array of analytical techniques is employed to integrate insights derived from network theory, with those derived from unsupervised machine learning, including latent space representations and clustering. The insights from these analyses are combined with approaches from meteorology and climate science. As a result, new ways to analyze and visualize teleconnection patterns in the climate system are presented. Using the specific region of the Indo-Pacific as a concrete example, this thesis approaches the question of how local and global teleconnection structures relate to each other.

1.3. Overview and Contributions

Part I provides the necessary background information, with Chapter 2 providing an overview of the different forms of teleconnection in the climate system and Chapter 3 introducing various methods to analyze climate data as employed in this thesis. Part II presents the main results of this thesis in three chapters. Chapter 4 and Chapter 5 are peer-reviewed publications and Chapter 6 is currently under review. The concluding Part III (Chapter 7) summarizes the answers to the key questions obtained in the publications and discusses their implications. This thesis is concluded by an outlook on potential directions of future research. Supporting Information are included in the Appendix in Part IV.

The author's individual contributions to the publications and a brief summary of each article are outlined below.

Publication 1: ***Teleconnection patterns of different El Niño types revealed by climate network curvature***

In Chapter 4, we adopt a network measure from complex network analysis called curvature, which we apply for the first time to climate networks to identify global teleconnection structures. This new approach allows an intuitive identification and visualization of teleconnection patterns in climate networks. This work uses climate networks that are constructed by using global surface air temperature data. We use the new curvature approach to reveal global teleconnection structures for two scenarios, in which the center of anomalous warming in the tropical Pacific differs: Eastern Pacific (EP) and Central Pacific (CP) El Niño events. We demonstrate that the curvature measure is highly effective in identifying long-range teleconnection structures, despite the absence of link length information in the climate network. We show that the curvature approach reveals a comprehensive picture of the spatial teleconnection structures of the full underlying data, in contrast to previous studies that have primarily focused on individual teleconnections. Our findings indicate that EP events exhibit teleconnections that are primarily occurring the tropics, whereas CP events have a more pronounced impact on the mid and higher latitudes. The methodological approach from this work has potential implications for the climate modeling community in evaluating whether teleconnection structures are well-represented in climate models or not.

Chapter 4 is based on the peer-reviewed journal publication with the following co-author contributions:

F. M. Strnad, J. Schlör, C. Fröhlich, and B. Goswami (2022). 'Teleconnection Patterns of Different El Niño Types Revealed by Climate Network Curvature'. In: *Geophysical Research Letters* 49.17. DOI: `10.1029/2022GL098571`

	Ideas	**Experiments**	**Analysis**	**Writing**
F. Strnad*	25 %	40 %	45 %	45 %
J. Schlör*	25 %	40 %	45 %	45 %
C. Fröhlich	25 %	20 %	0 %	0 %
B. Goswami	25 %	0 %	10 %	10 %

* Authors contributed equally.

Publication 2: ***Propagation Pathways of Indo-Pacific Rainfall Extremes are Modulated by Pacific Sea Surface Temperatures***

In Chapter 5 we shift our focus to regional-scale community structures using climate networks constructed from rainfall data in the Indo-Pacific region. In particular, we first identify distinct communities where EREs are likely to be synchronized. We then uncover the Boreal Summer Intraseasonal Oscillation (BSISO) as the responsible physical mechanism that modulates the form of the community structures. Next, we examine the propagation pathways of Indo-Pacific rainfall extremes between these communities. By using a clustering approach, we demonstrate that these pathways occur in three distinct modes. We reveal how these are modulated by Pacific sea surface temperatures and how this modulation is driven by an interaction of two of the primary drivers of boreal summer variability in the Indo-Pacific: ENSO and the BSISO. Lastly, we outline how these findings could improve the predictability of EREs on a time scale up to 4 weeks in advance which might help to develop early warning signals in the region. As the BSISO is one of the main drivers of the rainfall variability during Indian Summer Monsoon, the results of this work enable a better understanding of the active and break phases of the Indian Summer Monsoon and offer the potential for improved seasonal forecasts.

Chapter 5 is based on the peer-reviewed journal publication with the following co-author contributions:

F. M. Strnad, J. Schlör, R. Geen, N. Boers, and B. Goswami (2023). 'Propagation pathways of Indo-Pacific rainfall extremes are modulated by Pacific sea surface temperatures'. In: *Nature Communications* 14.5708. DOI: `10.1038/s41467-023-41400-9`

	Ideas	**Experiments**	**Analysis**	**Writing**
F. Strnad	70 %	90 %	65 %	70 %
J. Schlör	10 %	10 %	5 %	5 %
R. Geen	5 %	0 %	10 %	5 %
N. Boers	5 %	0 %	10 %	5 %
B. Goswami	10 %	0 %	10 %	15 %

Publication 3: *Intraseasonal synchronization of extreme rainfalls between North India and the Sahel*

In Chapter 6 we demonstrate that synchronization structures of EREs also exist at the global scale. We uncover a synchronization pattern of EREs between northern India and the Sahel. In contrast to previous studies, which typically treat monsoon systems as distinct entities, we find that rainfall in North India and the Sahel occurs synchronized on intraseasonal timescales. Furthermore, we identify the physical mechanisms that drive this synchronization. The BSISO and the Tropical Easterly Jet (TEJ) act in concert to initiate convection, which modulates the strength of the synchronization. The La Niña condition sets the background state that enables the synchronization. This work offers a new perspective on the teleconnection structures of EREs on continental scales and provides a better understanding of the rainfall variability in the Sahel.

Chapter 6 is based on the manuscript currently under review at Quarterly Journal of the Royal Meteorological Society (QJRMS) with the following co-author contributions:

F. M. Strnad, K. M. R. Hunt, N. Boers, and B. Goswami (2024). 'Intraseasonal synchronization of extreme rainfalls between North India and the Sahel'. In: *in review at Quarterly Journal of the Royal Meteorological Society*. DOI: `10.48550/arXiv.2405.08492`. eprint: `2405.08492`

	Ideas	**Experiments**	**Analysis**	**Writing**
F. Strnad	60 %	90 %	70 %	70 %
K.M.R. Hunt	10 %	0 %	10 %	10 %
N. Boers	20 %	0 %	10 %	10 %
B. Goswami	10 %	10 %	10 %	10 %

Part I.

BACKGROUND

Teleconnection of Extreme Rainfall Events

2.

2.1. Precipitation and Extreme Events

2.1.1. Generation of Rainfall

Precipitation can manifest as rain, snow, or hail, resulting from the condensation of rising warm air in the *atmosphere* when water droplets form around tiny particles, such as dust. Subsequently liquid droplets or ice crystals precipitates onto the Earth's surface (Strangeways, 2006). The lifting of air masses can be attributed to a number of factors, including convection, orographic lifting (e.g. when winds flow over a mountain chain), or the convergence of weather fronts (Trenberth et al., 2003). Convection most commonly occurs when the Earth's surface is heated by the sun, causing moist air to rise, called evapotranspiration, and leading to the formation of large cumulus clouds (Houze, 2014). In the tropics, most precipitation occurs as convective precipitation (Houze, 1997). Convective precipitation is characterized by intense, short-lived rainfall events that produce an anomalous amount of rainfall in a single day. We refer to these as *extreme rainfall events* (EREs).

Atmosphere: The atmosphere is a mixture of gases surrounding the Earth. It is composed primarily of nitrogen (78 %) and oxygen (21 %). The remaining 1 % include water vapor, carbon dioxide, and other gases. The atmosphere is divided into several layers, with the troposphere being the lowest layer where most weather dynamics occur. The troposphere extends from the Earth's surface to an altitude of about 8 – 15 km, depending on the latitude and season (Achatz, 2022). In this thesis, when we speak of the atmosphere, we refer to the troposphere.

2.1.2. Extreme Rainfall Events

There are different ways to define extreme events in climate science, but two approaches are the most common. The first approach is based on block maxima (BM), where the maximum values of finite sequences or blocks are considered as extremes (Serinaldi and Kilsby, 2014). One example is to define record-breaking precipitation sums per day as an extreme event (Lehmann et al., 2015). The second approach is called peak over threshold (POT). Here, values in a continuous record that exceed a certain threshold are considered extremes. In this approach, the threshold is often set based on percentiles of the observed distribution, such as the 90th (Gupta et al., 2022), 95th (Boers et al., 2019), or 99th percentile (Boers et al., 2014). The BM approach is more straightforward to interpret, but it requires a long time series to estimate the distribution of the maxima accurately. A potential bias with the POT approach can occur when there is a trend in the data, the occurrences of extremes may increase over time. This behavior has been observed for the Indian Summer Monsoon with an increase in moderate to extreme rainfall days (Katzenberger et al., 2022), potentially driven by global warming (Katzenberger et al., 2021). Therefore, one typically has to detrend the data which comes with its own assumptions like linearity of the trend which is debatable (Wu et al., 2007). The advantage of the POT approach is the pointwise definition of the threshold which allows easier comparison across different spatial locations. For instance, the total amount of rainfall over the sea is much higher than over land, mainly due to the much higher evaporation. As we are investigating the synchronization of EREs over large spatial distances, we will use the POT approach in this thesis which means we will transform the continuous precipitation time series into a binary event series of the same length, where each time step is either an extreme event (denoted as 1 in the event series) or not (denoted as 0).

EREs, regardless of the definition, are challenging to study mainly due to three reasons:

- **Quality of data** Ground-based observational precipitation data is often sparse, especially in remote regions where the number of rain gauges is limited. While satellites can provide broader coverage, they have limitations in terms of resolution, accuracy, and ability to measure different types of precipitation (e.g., snow versus rain). Interpolated data, which is often used to fill the gaps in the observational data, has strong assumptions (Dai and Bloecker, 2019), and downscaling methods (statistical or dynamical) introduce additional uncertainties (Gao, 2020; Cheng et al., 2024).
- **High temporal and spatial variability** Precipitation can vary substantially over short distances. For example, rainfall might be heavy in one location while an area only a few kilometers away remains dry. Also, the temporal variability is high: Precipitation can be short-lived (e.g., thunderstorms) or prolonged (e.g., steady rain) (Kendon et al., 2014). This involves the risk that actual extreme events are averaged out.
- **Nonlinearity** The relationship between atmospheric variables and the amount of rainfall is often nonlinear and includes local stochastic fluctuations which means that the amount of extreme events are over- or underestimated (O'Gorman, 2015).

When trying to understand the dynamics of EREs, it is crucial to consider the large-scale atmospheric circulation patterns. For this, precipitation data alone is often not sufficient, and one needs to consider additional variables that provide information on the atmospheric circulation:

- **Outgoing Longwave Radiation** (OLR) describes the outgoing radiation emitted from the Earth's surface. It occurs primarily in the infrared range (wavelength of $\approx 4-100\ \mu$m). OLR is related to cloud coverage, where more clouds produce lower values of OLR. This inverse relationship makes OLR a viable proxy for convection and precipitation in tropical and subtropical regions. The OLR data are directly available from satellite measurements, and can be assessed globally (in contrast to rain gauge measurements). It is a key variable to track the development and propagation of clouds. Since in the tropics, most precipitation occurs as convective precipitation, OLR is a good indicator for precipitation and its propagation pathways in these regions.
- **Wind Fields** provide information on the atmospheric circulation patterns. For example the formation of the monsoon rains can be related to the direction and strength of the winds that transport moist air from the oceans toward the land. Wind fields can also be used to identify the overturning circulation cells that drive convection, and accelerate or block convective systems (Schwendike et al., 2014, 2021).
- **Moist Static Energy** (MSE) is a thermodynamic variable. Examining the fluxes of MSE can provide insights into how energy is transported within the atmosphere (Sobel et al., 2008; Sobel and Maloney, 2012; Wang and Sobel, 2022), for example by indicating the initiation and propagation of convective processes (Kim et al., 2014). The column-integrated (over the vertical pressure levels) MSE variable is defined as the sum of enthalpy, latent heat, and potential energy as

$$MSE = \underbrace{C_p T}_{\text{enthalpy}} + \underbrace{L_v q}_{\text{latent heat}} + \underbrace{gz}_{\text{potential energy}} , \tag{2.1}$$

where T is temperature, C_p the specific heat of air at constant pressure, L_v latent heat of vaporization, q specific humidity, g the gravitational acceleration, and z geopotential height (Wallace and Hobbs, 2006). The moist static energy of an air parcel is a composite measure that includes the internal energy related to temperature, the latent energy associated with the water vapor content, and the potential energy due to the parcel's altitude.

2.2. Modulation of Atmospheric Teleconnections

The key to understanding the atmospheric modulation of teleconnections is the propagation of *Rossby Waves* in the atmosphere (Hoskins et al., 1977; Hoskins and Karoly, 1981). Rossby Waves are caused by the rotation of the Earth dependent on the variation of the Coriolis force with latitude which acts on moving particles in the ocean and the atmosphere (Achatz, 2022). It is a planetary-scale wave that is for example responsible for the characteristic meandering of jet streams (Achatz, 2022). Unlike water waves that show variations in the water's height, atmospheric Rossby Waves produce peaks and troughs of high and low air pressure. These resultant pressure patterns influence temperature, wind, and rainfall by induced convection (Boers et al., 2019; Beverley et al., 2021). Atmospheric Rossby Waves exist on timescales of days to weeks and are responsible for the propagation of weather systems across the globe (Vallis, 2006). Rossby Waves are likely to be triggered by large-scale modes of variability. These modes can be regarded as recurring patterns of atmospheric and oceanic circulation with substantial influence on weather and climate from regional and global scale by modulating atmospheric teleconnections. We briefly introduce the modes that were found to be most relevant to the studies in this thesis.

2.2.1. El Niño Southern Oscillation

The El Niño Southern Oscillation (ENSO) is the most dominant mode of sea surface temperature variability in the climate system characterized by the warming and cooling of the tropical Pacific Ocean in periods of around 2-7 years. The phase of ENSO characterized by anomalously warm sea surface temperatures (SSTs) is called El Niño (Figure 2.1 a), whereas the counterpart characterized by anomalously cool SSTs is called La Niña (Figure 2.1 b). ENSO is a coupled ocean-atmosphere phenomenon that influences the global atmospheric circulation through the alteration of the Walker circulation (Figure 2.2). This overturning circulation is characterized by the upward motion of air over the Maritime Continent, driven by the warm pool in the western tropical Pacific. The associated downward motions over the eastern Pacific together with the resulting low-level easterly winds and upper-level westerly winds complete the circulation cell (Bjerknes, 1969). ENSO

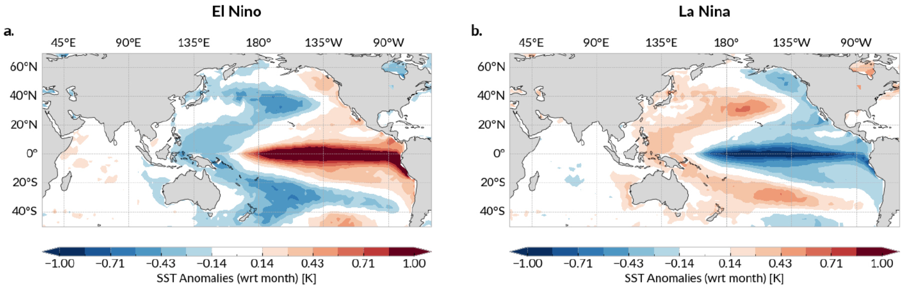


Figure 2.1.: Phases of the El Niño Southern Oscillation. El Niño (**a**) and La Niña (**b**) are shown by SST anomalies. The ENSO state is determined by using the Multivariate ENSO Index version 2 (MEIv2) for the time 1979-2022 (Wolter and Timlin, 2011).

events typically show similar development on seasonal timescales, and during this time they have global impacts on temperature variability, and also on precipitation and wind patterns around the globe (Capotondi et al., 2020).

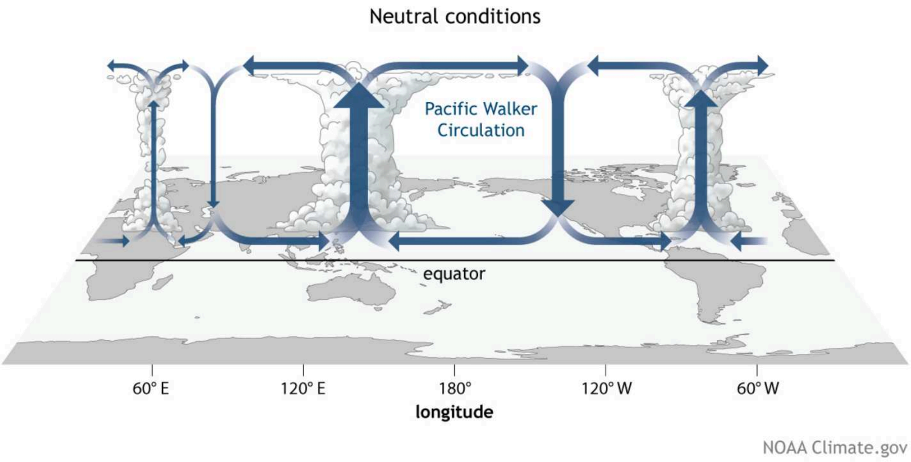


Figure 2.2.: Illustration of the Walker Circulation. The Walker Circulation during ENSO-neutral conditions is visualized. The areas of convection are associated with the ascending branches of the circulation. These occur strongest over the Maritime Continent and induce secondary cells over East Africa and South America. Figure adapted from `https://www.climate.gov/media/13542`

One of the most prominent teleconnections of ENSO is the modulation of the South Asian monsoon system (Kripalani and Kulkarni, 1997) through alternation of the Walker circulation: During La Niña events (Figure 2.1 b), the cooler SST temperatures in the eastern Pacific strengthen the low-level easterly winds, which increases the convection over the Maritime Continent and the South Asian main land, the circulation cell is intensified. This intensification leads consequently to anomalously high rainfall over the entire region. In contrast, during El Niño events warm waters from the equatorial western Pacific induce convection in the east Pacific which weakens the Walker circulation and the convection over the Western Pacific region, leading to reduced rainfall over the South Asian subcontinent (Kumar et al., 2006).

2.2.2. Madden-Julian Oscillation

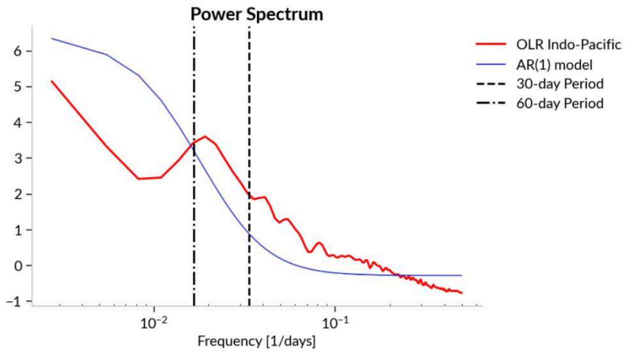


Figure 2.3.: Power spectrum of OLR field. The power spectrum of the mean Outgoing Longwave Radiation (OLR) time series in the tropical band between 15° S – 15° N shows the dominant frequency of the Madden-Julian Oscillation (MJO) around 30-60 days (dashed lines). For comparison, an AR1 model for the same time series is shown.

In 1971, Madden and Julian employed a spectral analysis of measures of OLR in the tropics and identified a dominant frequency of approximately 30-60 days (Figure 2.3), which was subsequently designated the Madden-Julian Oscillation (MJO) (Hand, 2015). This intraseasonal oscillation (ISO) is the leading mode of intraseasonal variability in the tropics characterized by periodically occurring fluctuations in atmospheric convection, rainfall, and wind patterns (Madden and Julian, 1971). Unlike ENSO, which is stationary, the MJO is a propagating convective system characterized by an eastward-moving band of clouds that circles the globe near the equator (Zhang, 2005; Zhang et al., 2020).

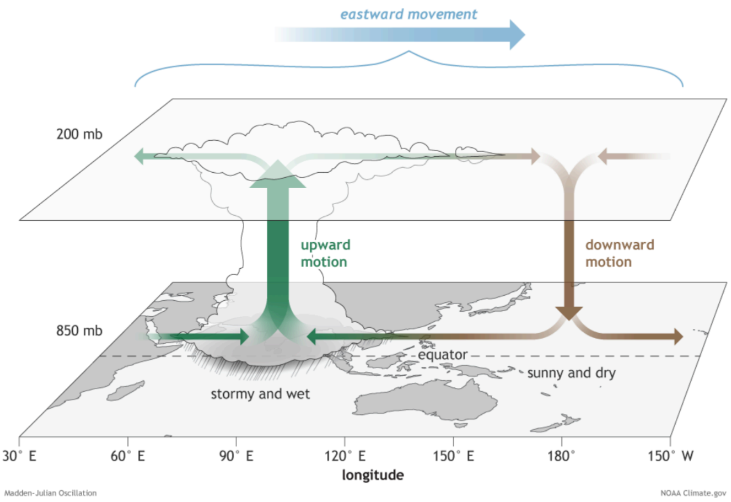


Figure 2.4.: Convective structure of the MJO. This schematic shows the vertical structure of a developing MJO with its enhanced convective phase (represented by the clouds) centered at the Indian Ocean. The associated suppressed convective phase is located over the equatorial West Pacific. The arrows are indicating the direction of the wind field. Here, green denotes an enhancement of the wind direction and brown a reduction. Together these winds describe the vertical circulation structure of the MJO. This whole system is moving eastward, denoted by the blue arrow. The figure has been adapted from `https://www.climate.gov/media/13488`

Figure 2.4 schematically visualizes the convective structure of the MJO. The MJO is composed of two parts: the enhanced (convective) rainfall phase and the suppressed rainfall phase forming a dipole pattern of wet and dry anomalies. In the region of enhanced convection, near-surface winds converge, causing air to rise (green arrows in Figure 2.4). This upward motion of air promotes condensation and leads to increased rainfall. As the air reaches the top of the atmosphere (at around 200 hPa), the winds diverge to the east and west. To the east of the convection center a convergence of winds at the upper levels of the atmosphere is established, causing the air to descend and subsequently diverge at the surface (brown arrows in Figure 2.4). As the air descends from higher altitudes, it warms and dries, leading to the suppression of rainfall. The mechanism driving the eastward movement of the convective system is still not fully understood, but it is believed to be due to the occurrences of large-scale pressure anomalies together with local heating in the lower equatorial troposphere (Rostami et al., 2022). This moving system is commonly categorized into phases from 1 to 8 (Wheeler and Hendon, 2004), depicted in Figure 2.5. The blue shaded areas denote regions of enhanced convection, and one can follow the eastward propagation of the MJO by following the blue shaded areas. Initiated in the Western Tropical Indian Ocean (Figure 2.5 a,b,c), the MJO propagates across the Maritime Continent (Figure 2.5 d,e,f) and toward the Western Pacific Oceans (Figure 2.5 g). Once the

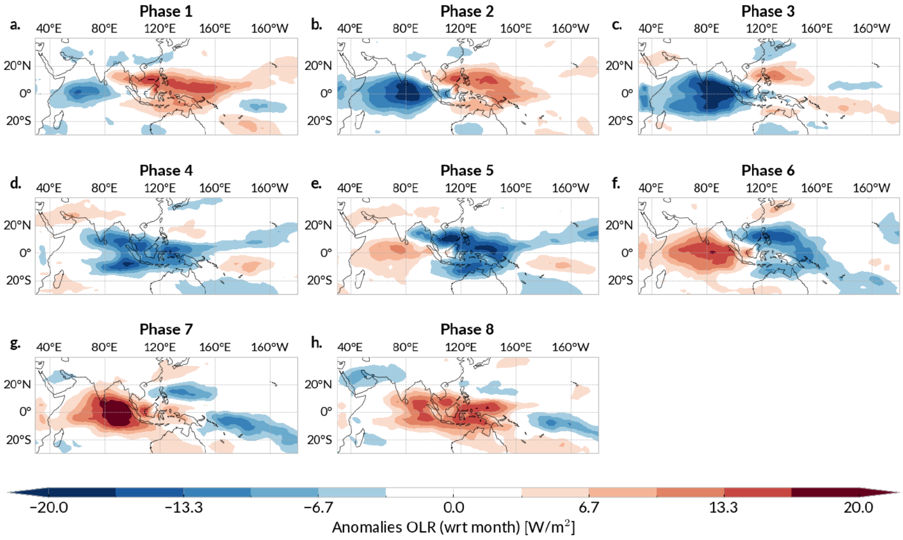


Figure 2.5.: Phases of the MJO. The different phases of the MJO are visualized by using outgoing longwave radiation (OLR) as a proxy for visualizing deep convection (see Subsection 2.1.1) for the years 1979-2022. Blue shaded areas denote regions of enhanced convection, and red denotes suppressed convection. The days for the MJO phases are taken from the RMM-index defined by Wheeler and Hendon, 2004. All days are considered as active MJO events according to the same definition as for the phases.

MJO reaches the Western Pacific, the cooler waters of the Pacific Ocean reduce the convective anomalies (Fig. Figure 2.5 h), but the wind circulations persist and continue their eastward propagation.

In the boreal summer the dynamics of the MJO changes slightly triggered by the change in circulation structure in the Northern hemisphere due to higher solar insolation (Sobel et al., 2008; Wang and Sobel, 2022; Wang and Xie, 1997). In addition to the eastward traveling direction the oscillation shows a further northward propagating component. This northeastward propagating convective system is called Boreal Summer Intraseasonal Oscillation (BSISO) (Kikuchi, 2021) and can be understood as the summer counterpart of the MJO (Wang and Sobel, 2022). The BSISO propagates over the Indian and South Asian mainland and modulates here the active-break cycles of the monsoon season, i.e. driving a strong alternation of EREs and dry phases (Kikuchi, 2021). Like for the MJO, an active BSISO is initiated in the Indian Ocean (Figure 2.6 a) and the whole system of enhanced/suppressed convection moves northeastward (Figure 2.6 b). When reaching the Maritime Continent, the convective system undergoes a weakening process, accompanied by the emission of moist Rossby waves that subsequently propagate in a northwesterly direction, ultimately reaching the Indian subcontinent (Figure 2.6 c). Here, the convective system expands into a northwest-southeast tilted band of clouds, ranging from southern Pakistan in the northwest over the Bay of Bengal to the Philippine Sea and the Maritime Content the southeast (Figure 2.6 d–f). This corridor of clouds is accompanied by heavy rainfalls. It slowly propagates northeastward and after approximately two weeks, it reaches the Western Pacific where it dissipates (Figure 2.6 g–h).

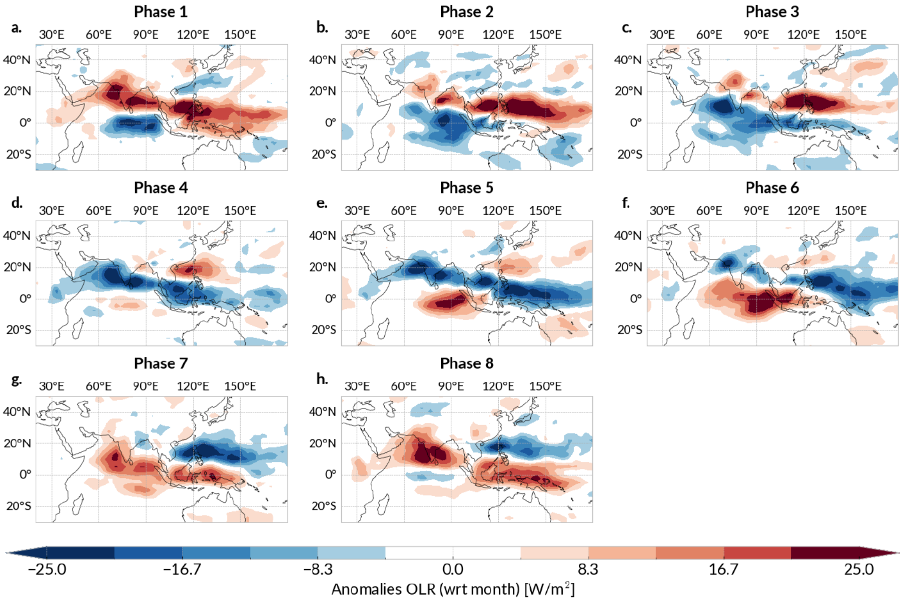


Figure 2.6.: Phases of the BSISO. The different phases of the Boreal Summer Intraseasonal Oscillation (BSISO) are visualized by using outgoing longwave radiation (OLR) as a proxy for visualizing deep convection for the years 1979-2022. As the BSISO is mainly occurring in boreal summer, only days from June to September are used. Blue shaded areas denote regions of enhanced convection, and red denotes suppressed convection. The days for the specific BSISO phases are taken from the index definition by Kikuchi et al., 2012. All days are considered as active BSISO events according to the same definition as for the phases.

2.3. The Global Monsoon System

A crucial factor in the Earth's hydrological cycle is the global monsoon system. It modulates many regional and global-scale teleconnections (Wang and Ding, 2008). Historically, the monsoons are regarded as large-scale sea breeze circulations. The temperature difference between the landmass and the surrounding oceans are assumed to create a low-pressure system that draws moist air from the ocean

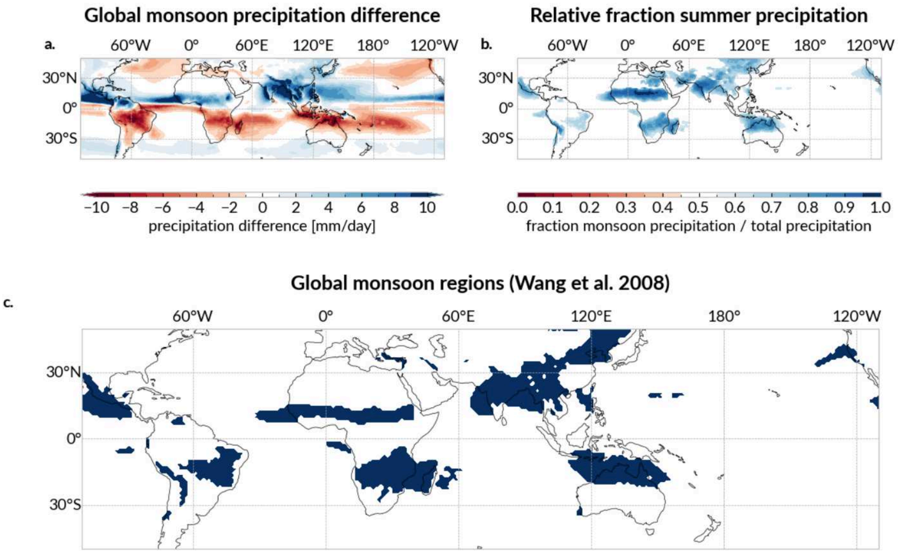


Figure 2.7.: Global Monsoon regions defined by Wang and Ding (2008). The definition has two criteria: (i) The regions are defined where pointwise the summer (JJAS) minus the winter (DJFM) precipitation exceeds a threshold of 2 mm/day (**a**) and (ii) where the summer accounts for at least 55% of the annual total annual precipitation (**b**). The intersection of both criteria defines the regions of the Global Monsoon (**c**) that are (from west to east): the North American Monsoon, the South American Monsoon, the West African Monsoon, the South African Monsoon, the South Asian Monsoon, and the North Australian Monsoon.

toward the land (Ramage, 1971). More recently, based on the observation that the monsoons in different regions of the world show strong similarity in their dynamics, a new perspective has emerged, which views the monsoons as part of a circulation system that spans the entire globe (Wang and Ding, 2008; Trenberth et al., 2000; Bordoni and Schneider, 2008; Geen et al., 2020). Wang and Ding (2008) introduced the concept of the "*Global Monsoon*" (GM)(Figure 2.7) characterized by rainy summers and dry winters with revised prevailing winds.

The perspective of a GM is based on the idea that monsoon circulations are cross-equatorial *Hadley circulations* (Bordoni and Schneider, 2008; Geen et al., 2020; Walker et al., 2015). This viewpoint differentiates the dynamics of the Hadley cell into two regimes that arise due to the seasonal movement of the tropical convergence zone driven by the seasonal variation in solar insolation: (i) The Intertropical Convergence Zone (ITCZ) regime with a convergence zone near the equator approximately co-located with the peak SST (Schneider et al., 2014) (Figure 2.8 a (ii) the monsoon regimes in the higher latitudes farther from the equator (around 15° N/S – 25° N/S) which are considered as the poleward extension of the Hadley circulation (Bordoni and Schneider, 2008; Geen et al., 2020) (Figure 2.8 b) where the convergence zone co-locates with MSE maximum.

Hadley Circulation: The Hadley circulation describes a meridional atmospheric overturning circulation cell that is driven by the temperature difference between the equator and the poles. Warm air near the equator rises due to intense solar heating and flows poleward near the tropopause (at around 100 – 250 hPa). Poleward, the air cools and descends in the subtropics (around 25° N/S). Finally, it returns equatorward near the surface, completing a cell labeled as the Hadley cell (Hoskins et al., 2020). This convergence zone moves seasonally, following position where the suns's rays are vertical and most intense.

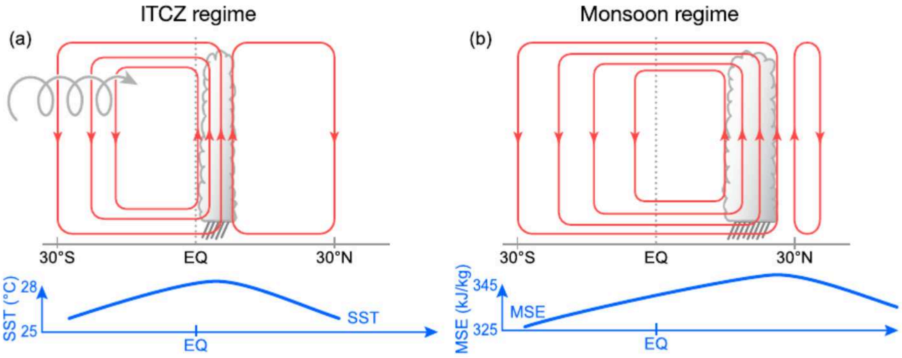


Figure 2.8.: Schematic of the Intertropical Convergence Zone (ITCZ) and the monsoon regime. ITCZ regime with convergence zone near the equator (EQ) (**a**) and monsoon regime **b** displayed. Red arrows denote the wind stream and the region of strongest convective precipitation is visualized by grey clouds. Figure adapted from (Geen et al., 2020).

The convergence zone of the Hadley cell is associated with the convection of moist air, leading to the formation of clouds and heavy precipitation. Further, the seasonal movement of the peak of solar insolation leads to the development of a low-pressure system over the continents and a high-pressure system over the oceans. The resulting pressure gradient causes the trade winds to shift, drawing moist air from the oceans towards the land and leading to the formation of the monsoon rains that taken together, account for approximately 31 % of total precipitation worldwide (Wang and Ding, 2008).

It is therefore logical to regard the monsoons as part of a connected global circulation system that spans the entire globe. Still, the framework of the GM is mainly applied to interannual over decadal to millennial timescales, for example, to study the development of the monsoon over different geological eras (Meijer et al., 2024). Most studies on specific monsoon regions on synoptic to interannual timescales focus on regional dynamics which means that they do not consider interactions between the different monsoon regions. In Chapter 6, we will challenge this viewpoint and show that two of the most prominent monsoon systems, namely the Indian and the West African Monsoon, interact on daily timescales.

3.

Spatio-Temporal Data Analysis

3.1. Spatio-Temporal Data

Climate data is typically stored as spatio-temporal datasets. These are four-dimensional arrays $X \in \mathbb{R}^{N \times T}$, where N represents the spatial locations and T is the number of time points. Commonly, the data is available on a regular grid, with N_{lat} latitude points and N_{lon} longitude points. For investigating atmospheric dynamics, one has to include the height by using different *isobaric surfaces* (constant pressure surfaces, see box), referred to as vertical pressure levels commonly from 50 hPa to 1000 hPa (surface pressure at sea level is at 1013 hPa). The number of vertical levels is denoted as N_{lev} and the number of spatial points is in total $N = N_{\text{lat}} \times N_{\text{lon}} \times N_{\text{lev}}$. There are various types of spatio-temporal datasets:

- **Observational data** are obtained from a variety of sources, such as ground-based weather stations, satellites, and weather balloons.
- **Climate model data** are generated using numerical models (so-called General Circulation Models or GCMs) of the Earth system. These models simulate the behavior of the atmosphere, oceans, land surface, cryosphere and their interactions based on numerical solving of physical differential equations. The complexity of different models varies, ranging from simple energy balance models to complex GCMs and is mainly determined by the resolution and the number of processes resolved (Jeevanjee et al., 2017). The most important GCMs are included in the Coupled Model Intercomparison Project (CMIP) which found the scientific basis for the IPCC reports (Eyring et al., 2016).
- **Reanalysis data** combines (potentially sparse) observational data from direct in situ measurements (temperature, pressure, wind, humidity and precipitation) with numerical weather models to create a consistent and continuous dataset of past meteorological variables form the Earth's surface to the top of the atmosphere.
- **Multi-source ensemble data** combine multiple sources, such as datasets from observations, models, and various reanalysis products, to create a more robust and reliable dataset. These datasets can help to reduce uncertainties and biases in individual datasets (Beck et al., 2019b). One example for precipitation data is the Multi-Source Weighted-Ensemble Precipitation (MSWEP) dataset (Beck et al., 2019b) which particularly provides more accurate and reliable information on EREs (Beck et al., 2019a).

Why vertical pressure levels instead of altitude: In the horizontal direction, winds move along isobaric levels. The pressure levels are defined by the hydrostatic equation

$$\frac{\partial p}{\partial z} = -\rho g \, ,$$

which relates the pressure p to the altitude z, the density ρ of the air and the gravitational acceleration g. In good approximation, the altitude can be obtained by dividing geopotential height by the gravitational acceleration assuming it to be constant with altitude $g \approx g_0 = 9.81$ (Achatz, 2022).

Geopotential: Geopotential Φ is the potential energy per unit mass in the Earth's gravitational field. It is a function of the height z and the gravitational acceleration g, depending on the latitude ϕ:

$$\Phi(z) = \int_0^z g(\phi, z) \, . \qquad (3.1)$$

The geopotential height is a measure of the height of an atmospheric pressure level. It is used for example to map the movement of low and high pressure systems and the associated wind fields.

3.2. Traditional Climate Data Analysis

3.2.1. Dimensionality Reduction

The high dimensionality of Earth system data presents a substantial challenge for data analysis as already visualizing high-dimensional climate data is difficult. Dimensionality reduction techniques are employed to reduce the complexity of the data while retaining the most pertinent information. This allows for a more straightforward interpretation of the data and can help to identify the most pronounced patterns. Below, we introduce some common dimensionality reduction techniques used in climate science.

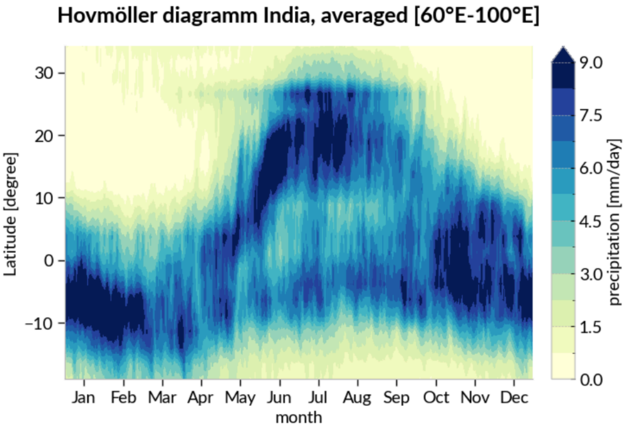


Figure 3.1.: Hovmöller diagram of precipitation in India. The x axis represents the day of the year, and the y axis represents the latitude band. For each latitude value one averages about the longitude range from 60° E – 100° E (i.e. the Indian monsoon domain). The color scale represents the precipitation amount. The diagram shows the daily precipitation over time and latitude for the monsoon region of India. The sinusoidal pattern indicates the seasonal migration of the monsoon rainband.

Hovmöller Diagrams Hovmöller diagrams (Hovmöller, 1949) are a tool for visualizing the behavior of time-varying geospatial data. They are two-dimensional plots that show the evolution of a variable like precipitation over time and latitude *or* longitude where the color scale represents the variable of interest. Commonly one averages over the other spatial dimension (longitude or latitude). In Figure 3.1, we show an example of a Hovmöller diagram composed of precipitation data for the region of the Indian monsoon. It visualizes intuitively the seasonal migration of the monsoon rainband over the Indian subcontinent.

Principal Component Analysis Empirical Orthogonal Function (EOF) analysis, also referred to as Principal Component Analysis (PCA) (Lorenz, 1956) reduces the dimensionality of spatiotemporal data while trying to retain the largest possible fraction of its variance. A spatiotemporal field is denoted as $\vec{x} \in \mathbf{X}$, where $\vec{x}(t)$ is the flattened vector field at time $t \in [1, \dots, T]$ of dimension $N = N_{\text{lon}} \cdot N_{\text{lat}}$. The principal components are defined as the vector $\vec{z}_r(t) \in \mathbb{R}^P$ with $r \in [1, \dots, P]$ being the rth component at time t is given by the linear combination $z_r(t) = \sum_s^N x_s(t) \cdot u_{r,s}$. The space spanned by the eigenvectors $\vec{z}_r$ is called the latent space and is constructed from the rth eigenvectors $\vec{u}_r$ of the covariance matrix $\langle \mathbf{X}\mathbf{X}^T \rangle$. By sorting the corresponding eigenvalues λ_r in descending order, the PCs are ordered by the amount of variance they explain. Using the $P \ll N$ most dominant PCs, one substantially reduces the dimensionality. For example,

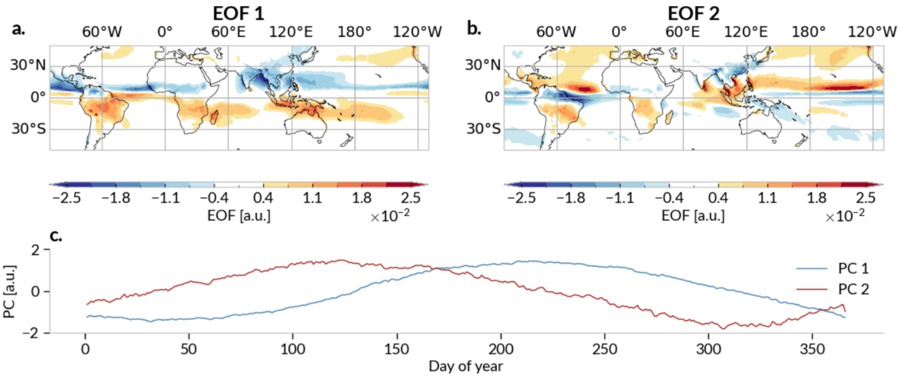


Figure 3.2.: EOF analysis of global precipitation field. In **a,b** the spatial patterns of the first two EOF modes of the climatological daily mean precipitation for the period of 1979 – 2022 are shown. **a** shows the first EOF mode, the solstitial mode, and **b** shows the second EOF mode, the equinoctial asymmetric mode. Both reflect the global monsoon patterns (Wang and Ding, 2008). In **c** the corresponding time series of the first two PCs are shown.

the concept of the Global Monsoon (Section 2.3) is formally introduced by using the first two EOFs of the global precipitation field (Figure 3.2 a,b). The first mode denotes the solstitial mode, and the second mode, called equinoctial asymmetric mode, reflects the spring-fall asymmetry of the ITCZ. Figure 3.2 illustrates that the primary climatology modes in tropical precipitation can be characterized by the solstitial mode and the spring-fall asymmetric mode. These are driven by the annual cycle of insolation and are called the global monsoon patterns (Wang and Ding, 2008; Geen et al., 2020). For this reason, the monsoon season is commonly defined by the peak of the highest precipitation for PC1 from June to September (Figure 3.2 c) in the Northern Hemisphere and from December to March in the Southern Hemisphere.

3.2.2. Time Series Analysis

Anomaly Time Series The climate system underlies various external forcings, of which for our purposes the most important are the daily and seasonal cycles as well as anthropogenic climate change. These forcings can bias for example correlation-based analyses. To remove these biases and to focus on the variability of the climate system one often uses anomaly time series. An anomaly time series is obtained by calculating deviations from the climatological mean (the so-called climatology), allowing to focus on the internal variability of the data rather than on the absolute values. Typically, anomaly time series are based on deviations from the day of the year, the month or the seasonal mean.

Climatology: The climatology is the average of a variable over a specific period, for example, the average temperature per month, averaged over all months in the dataset or the average precipitation in the June through September (JJAS) season.

Climate Indices Climate indices are time series that capture the behavior of specific meteorological phenomena over time. They are often derived from spatiotemporal fields, such as sea surface temperature or precipitation, using PCA or averaging over a specific region that is known to be most strongly influenced by the meteorological phenomenon. Due to their intuitive form, climate indices are widely used in climate science to monitor large-scale modes of atmospheric variability (Section 2.2). These are commonly represented by a climate index and their influence is frequently assessed by computing the correlation between the climate index and spatiotemporal fields.

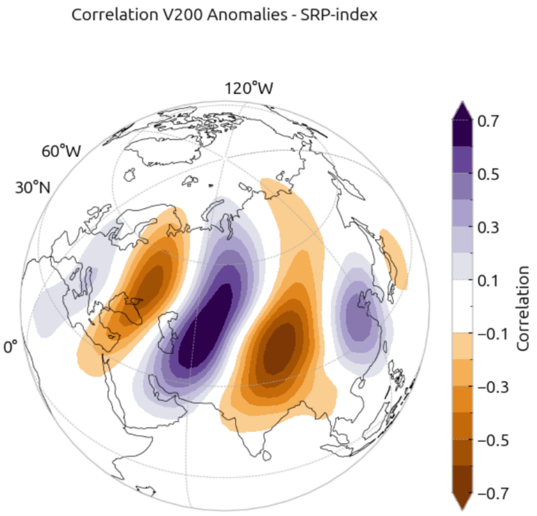


Figure 3.3.: Correlation of the CGT index with the meridional wind field at 200 hPa. The map shows spatial locations where the CGT index is significantly correlated (or anti-correlated) with the meridional V wind field. The alternating positive and negative correlation values indicate the wave-like structure of Rossby waves in the subtropical jet stream.

A famous example is the Euroasian wave train pattern, that is part of the Circumglobal Teleconnection Pattern (CGT) (Ding and Wang, 2005). It is a global scale atmospheric Rossby wave pattern occurring in the subtropical jet at around 200 hPa. Most importantly it modulates the influence of Asian to European weather (positive CGT pattern) and vice versa (negative CGT) (Beverley et al., 2019). The CGT index is defined as the average over the region 35° N – 40° N and 60° E – 70° E of the 200 hPa geopotential height anomaly field (Ding and Wang, 2005). The spatial pattern over the Euroasian continent, called the Silk Road Pattern (SRP), can be visualized by using the correlation of the CGT index with the meridional (v) wind field at 200 hPa (Figure 3.3). The correlation map shows the regions where the CGT index is significantly correlated with the meridional v-wind field. The alternating positive and negative correlation values indicate the wave-like structure of Rossby waves in the subtropical jet stream which transports moisture between Europe and the Indian monsoon system (Beverley et al., 2021).

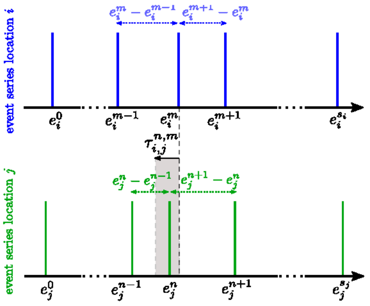


Figure 3.4.: Scheme of time-delayed event synchronization. The blue (green) bars denote events that happen at the locations in i (j) at the time points e_i^m (e_j^n). m (n) denotes the event number at location i (j). The arrows indicate the time differences between neighboring events at the respective locations, i.e. $d_{i,i}^{m,m-1}$, $d_{i,i}^{m,m+1}$, $d_{j,j}^{n,n-1}$ and $d_{j,j}^{n,n+1}$. In this exemplary sketch, the point in time e_i^m is identified as synchronous to e_j^n since the time difference between e_i^m and e_j^n is within the allowed range given by $\tau_{ij}^{m,n}$ according to the definition expressed in equation 3.3.

Event Synchronization The Event Synchronization (ES) algorithm (Quian Quiroga et al., 2002) quantifies the degree of synchronization between pairs of time series. For a pair of event sequences $\{e_i^m\}_{m=1}^{s_i}$ and $\{e_j^n\}_{n=1}^{s_j}$, where s_i (s_j) represents the total number of events at location i (j), and e_i^m (e_j^n) denotes the point in time of the mth (nth) event in that location one counts the number of synchronous events. In Fig. 3.4, we show an exemplary scheme of two events in two event series that are considered synchronous. The time difference d between an event e_i^m in location i and an event e_j^n in location j is denoted as $d_{ij}^{m,n} = e_i^m - e_j^n$. The set $D_{i,j}^{m,n}$ contains the time differences to the neighboring events of e_i^m and e_j^n respectively:

$$D_{i,j}^{m,n} = \left\{ d_{i,i}^{m,m-1}, d_{i,i}^{m,m+1}, d_{j,j}^{n,n-1}, d_{j,j}^{n,n+1}, 2\tau_{\max} \right\}, \tag{3.2}$$

The dynamical delay $\tau_{i,j}^{m,n}$ is calculated as half of the minimum time contained in D, limited to a maximum value $\tau_{\max}$:

$$\tau_{ij}^{m,n} = \frac{1}{2} \min_{\forall d \in D_{ij}^{m,n}} d\,. \tag{3.3}$$

The event synchronization strength $R_{i,j}$ sums up the number of synchronous events between the two event series i and j:

$$R_{i,j} = \sum_{m=1}^{s_i} \sum_{n=1}^{s_j} S_{i,j}^{m,n}\,, \quad \text{where} \quad S_{i,j}^{m,n} = \begin{cases} 1 & 0 < d_{ij}^{m,n} < \tau_{i,j}^{m,n}\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{3.4}$$

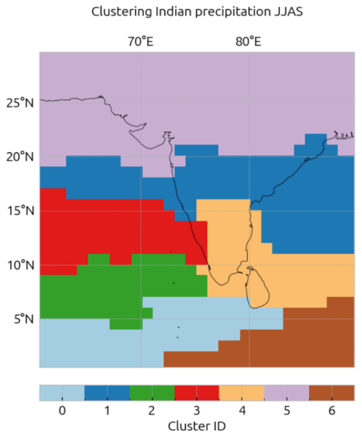


Figure 3.5.: Clustering of Indian Summer Monsoon precipitation. The correlation matrix of the precipitation time series of different regions for the summer months (JJAS) for the period 1979–2022 is used as a distance matrix to cluster the regions. The clustered regions show geographically separated regions with different rainfall patterns.

Clustering Analysis Clustering is a technique used to group data points to identify patterns and relationships in the data. This method assigns a discrete membership $1, \ldots, k \in \mathbf{N}$ to each data point based on the similarity of the data points. The similarity is often calculated using a distance metric, such as the Euclidean distance but there are numerous alternative ways to define a distance. For example, as done in Rheinwalt et al. (2015) in the context of the Indian Summer monsoon, one can cluster the rainfall time series of different spatial locations and define the distance between points based on their inverse correlation value (Figure 3.5). The clustered locations show geographically separated regions with different rainfall patterns as for example the Western Ghats, the Western Coast, and North Indian including the Himalayan mountain chain. For climate data, the spatial dimensionality is often high compared to the number of time points, which can make clustering challenging. Therefore, sometimes a combination of clustering and with additional dimensionality reduction techniques is used to identify patterns in high-dimensional spatio-temporal data (Schlör et al., 2024). The approach is a two-step process: First, one uses PCA to reduce the dimensionality of the data and then applies a clustering algorithm in the latent space spanned by the most dominant EOFs (Figure 3.6). The inverse transformation from a representative point of a cluster in the latent space allows for a reconstruction of a representation of the cluster in real space.

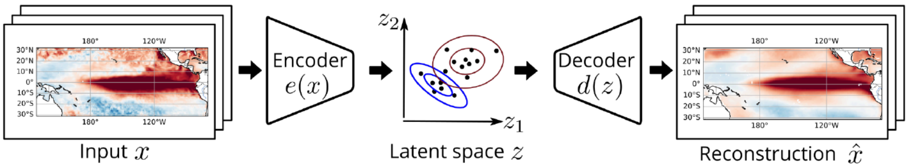


Figure 3.6.: Clustering in a latent space The real space, represented as a SST field with $N_{\mathrm{lon}} \cdot N_{\mathrm{lat}}$ dimensions is transformed to a 2-dimensional latent space z via the encoder function $e(z)$. The distribution of points in the latent space is then clustered. The back-transformation by the decoder function $d(z)$ gives the representation of the clusters in real space. The schematic is adapted from Schlör et al. (2024).

In Chapter 5 and Chapter 6, we use a similar approach to cluster propagation pathways of anomalous precipitation events in the Indian Summer Monsoon by either first applying a Hovmöller dimensionality reduction (Chapter 5) or projecting the data into a latent space by a PCA (Chapter 6).

3.3. Identifying Teleconnection Patterns

Teleconnections are commonly detected through statistical correlations between variables, such as temperature, precipitation, or atmospheric pressure, in geographically separated regions. An arguably intuitive approach is to use climate networks (Subsection 3.3.1) to identify the connections between different regions based on the similarity of their climate variability. Here, edges between nodes of the network (i.e., spatial locations) can be directly regarded as teleconnections.

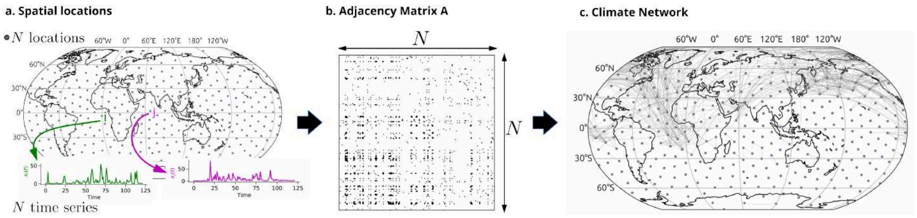
3.3.1. Climate Network Construction

A spatiotemporal dataset of time series (for example temperature or precipitation) is given as $\mathbf{X} \in \mathbb{R}^{N \times T}$, N denoting the number of spatial locations and T the number of time points (Figure 3.7 a). The climate network is mathematically represented by the graph $\mathcal{G} = (V, E)$. Each geographical location given by its index $i \in 1, \dots, N$ translates to a node $n_i \in V$, where V is the set of nodes. Each node is associated with the time series $\vec{x}_i = (x_{i,0}, \dots, x_{i,T}) \in \mathbf{X}$ of the spatial location i. An edge $e_{ij} \in E$ between two nodes n_i and n_j encodes a statistically significant dependence between the time series $\vec{x}_i$ and $\vec{x}_j(t)$, where E is the set of all edges. The matrix $\mathbf{A}$, defined as $\mathbf{A} \in \{0, 1\}^{N \times N}$ is called the adjacency. $A_{ij} = 1$ denotes that there is an edge between node n_i and n_j and $A_{ij} = 0$ otherwise (Figure 3.7 b). The network can be visualized by plotting the nodes at their geographical locations and the edges as curved lines (Figure 3.7 c).

There are multiple ways to define statistical dependence, depending on the type of time series:

- **Correlation-based climate networks** For continuous time series, one often uses correlation-based methods. Typically, the underlying variables are slowly varying in time like temperature or surface pressure. The network

Figure 3.7.: Climate network construction. The statistical dependence between all pairs of time series $x_i(t)$ and $x_j(t)$ are calculated (**a**). This results in the adjacency matrix **A** (**b**). The nodes are plotted at their geographical locations as dots together with their edges visualized by curved lines (**c**).



is constructed by calculating the correlation matrix **W** using Pearson correlation (Tsonis and Swanson, 2008), Spearman correlation (Donges et al., 2009b) or mutual information (Donges et al., 2009a) and thresholding the correlation matrix **W** at a certain level θ to obtain the adjacency matrix **A**:

$$A_{ij} = \begin{cases} 1 & \text{if } |W_{ij}| > \theta \ , \\ 0 & \text{otherwise} \ . \end{cases} \tag{3.5}$$

The threshold θ is chosen to retain only the most significant correlations. Still, one has to account for random correlations due to the high number of comparisons (Haas et al., 2023). For example, one can use multiple testing corrections like the False Discovery Rate (FDR) (Benjamini and Hochberg, 1995) or the Holm and Bonferroni correction (Holm, 1979; Hochberg and Benjamini, 1990).

- **Event-based climate networks** For binary time series one often uses the event synchronization algorithm (for example Boers et al., 2019; Malik et al., 2010; Boers et al., 2014; Stolbova et al., 2016; Wolf et al., 2021) based on the pair-wise event-to-event comparisons (see Subsection 3.2.2, Figure 3.4). The statistical significance of the synchronization is determined by comparing the number of synchronous events to a null model of randomly distributed events. When a significant synchronization between i and j is found, an edge from node n_i to n_j is established, and the adjacency matrix **A** is set to $A_{i,j} = 1$.

3.3.2. Uncovering Spatial Teleconnection Patterns

Climate networks are useful for detecting single teleconnections but also to uncover interaction and relationships between multiple processes in the climate system. Applying the tools of complex network science to analyze the topology of the network allow making conclusions on the underlying physical processes that are responsible for generating the specific topology. One intuitive example is the (weighted) node degree k_i ($\tilde{k}_i$), a frequently applied in studies using climate networks (Gupta et al., 2022; Donges et al., 2009a; Boers et al., 2013; Donges et al., 2015). For example, when the node degree is plotted for a network of surface air temperature anomalies, it reveals a pronounced region in the Eastern Tropical Pacific region (Figure 3.8). This is the key region of ENSO's influence on the global SAT variability (Donges et al., 2009a).

Node Degree: The node degree k_i ($\tilde{k}_i$) is defined as the (weighted) number of nodes adjacent to node n_i:

$$k_i = \sum_j^N \mathbf{A}_{ij} \ , \quad \tilde{k}_i = \sum_j^N \mathbf{W}_{ij} \ ,$$

where N denotes the number of nodes $n_i \in V$, **A** the adjacency matrix, and **W** the correlation matrix.

Investigating the network's topology is possible in different ways and depends on the specific context. In this work, we focus on two approaches that allows an all-encompassing macroscopic view on the structure of the network. In particular, this means to analyze the network with respect to its community structure (i.e. groups of locally densely connected nodes) and their links between these communities. We adopt two recent applications from complex network science, namely network curvature (Chapter 4) and community detection (Chapter 5, Chapter 6).

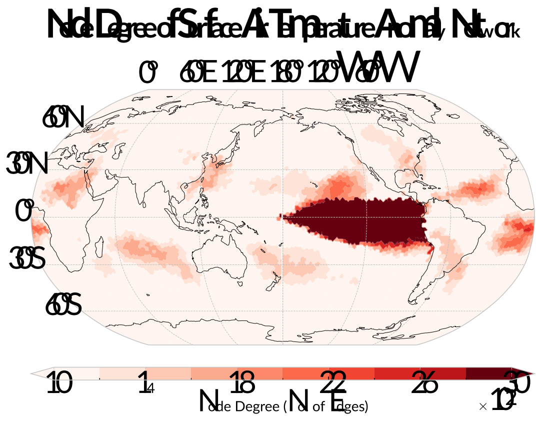


Figure 3.8.: Node degree for surface air temperature network. The node degree is visualized for a network of surface air temperature anomalies. The network is calculated with a correlation threshold of $\theta = 0.01$ and based on a Fekete grid.

Curvature The curvature method is a novel approach to analyze the network structure by assigning a curvature value to each edge of the network. It allows identifying important links in the network that connect different regions which we interpret as the most prominent teleconnections in the network (see Chapter 4). To provide the reader with some intuition on the network curvature, we use a toy dataset with known covariance structure (see box for details). Figure 3.9 visualizes the intuition behind the curvature approach: The curvature measure (called Forman-curvature) allows to distinguish the between-community links (blue edges, negative Forman- curvature) from the within-community links (red edges, positive Forman-curvature) by the sign of the curvature. Similar network measures like the edge betweenness centrality which is often used in climate network studies to identify edges that connect communities (for example Stolbova et al., 2016; Donges et al., 2009b; Stolbova et al., 2014) do not reveal this intuitive separation of network links (Figure 3.9 b). The rationale behind this approach is that when all negative links (i.e. between-community links) are plotted on a spatial map in real space, one obtains a global representation of the most prominent teleconnections all at once. We apply this approach to climate networks of surface air temperature anomalies to estimate global differences in teleconnection structures of different El Niño events (Chapter 4).

Toy Dataset: $X_{\text{toy}} \in \mathbb{R}^{NxT}$ is generated from a random graph with four communities of equal size with the following edge probabilities:

$$P = \begin{bmatrix} 0.7 & 0.02 & 0.02 & 0.03 \\ 0.02 & 0.6 & 0.02 & 0.01 \\ 0.02 & 0.02 & 0.6 & 0.02 \\ 0.03 & 0.01 & 0.02 & 0.7 \end{bmatrix}.$$

Time series with $T = 1000$ time points are sampled from a multivariate Gaussian, $\mathbf{N}\left(x, 0, A_{\text{toy}}\right)$, N denoting the number of nodes (i.e. data points), with the adjacency matrix $A_{\text{toy}} \in \mathbf{R}^N$ of the graph used as the covariance matrix.

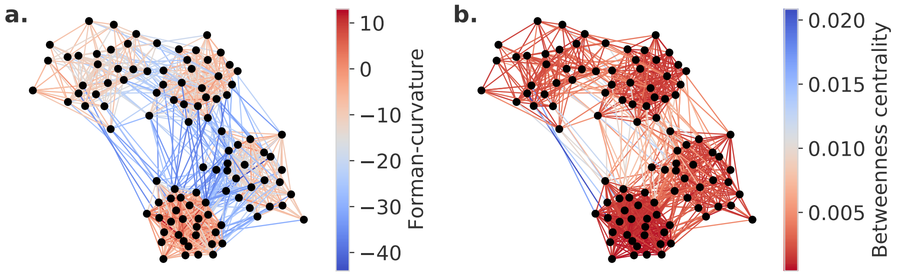


Figure 3.9.: In the toy example network, **a** shows the edges colored according to their Forman-Curvature value. b shows the same network, but each edge colored according to its betweenness centrality value. The network is constructed as outlined in the box above.

Community Detection Community detection describes the process of identifying densely connected groups of nodes with similar contributions to the network structure. Nodes within a community are more densely connected within the group than between groups. The identification of communities within a network can provide insights into its macro-scale structure, highlighting important patterns and pointing to the underlying mechanisms that contribute to its formation, for example through physical processes. Therefore, it is beneficial to use inferential methods to identify the community structure of a network (Peixoto, 2023). Inferential methods are based on the assumption that the network is generated by a generative model. The process of identifying communities translates to reordering the rows and columns of the adjacency matrix **A** such that a block-diagonal structure emerges (Fortunato and Hric, 2016). The Stochastic Block Model (SBM) generates a network graph by using as input a given "block-structure" model of the adjacency matrix in which the edges attached to any given node are determined by two probabilities: a within-block edge probability, and an across-block probability. For example, this model has been used to generate the toy example above (see Figure 3.9). The SBM is a generative model for network graphs where k blocks would be described by $\frac{k(k+1)}{2}$ probabilities. The probabilities can then be calculated using Bayesian inference (Peixoto, 2019). Our idea behind the application of the community detection algorithm is that it helps to identify regions with synchronous occurrences of EREs and assume implicitly their shape allows conclusions on macro-scale physical processes behind (see Chapter 5 and Chapter 6).

Part II.

Investigating Teleconnection Structures by Complex Networks

4. Teleconnection Patterns of Different El Niño types Revealed by Climate Network Curvature

This chapter shows the following article:

F. M. Strnad, J. Schlör, C. Fröhlich, and B. Goswami (2022). ‘Teleconnection Patterns of Different El Niño Types Revealed by Climate Network Curvature’. In: *Geophysical Research Letters* 49.17. DOI: `10.1029/2022GL098571`.

Geophysical Research Letters

RESEARCH LETTER
10.1029/2022GL098571

Teleconnection Patterns of Different El Niño Types Revealed by Climate Network Curvature

Felix M. Strnad[1], Jakob Schlör[1], Christian Fröhlich[1], and Bedartha Goswami[1]

[1]Cluster of Excellence Machine Learning: New Perspectives for Science, Universität Tübingen, Tübingen, Germany

Felix M. Strnad and Jakob Schlör contributed equally to this work.

Key Points:

- Ricci curvature of boreal winter climate networks reveals long-range teleconnection structure
- Eastern Pacific (EP) El Niños show primarily teleconnections in tropical while Central Pacific El Niños teleconnections on all latitudes
- The EP contains robust teleconnections for both El Niño types

Supporting Information:

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

Correspondence to:

F. M. Strnad and J. Schlör,
felix.strnad@uni-tuebingen.de;
jakob.schloer@uni-tuebingen.de

Citation:

Strnad, F. M., Schlör, J., Fröhlich, C., & Goswami, B. (2022). Teleconnection patterns of different El Niño types revealed by climate network curvature. *Geophysical Research Letters*, *49*, e2022GL098571. https://doi.org/10.1029/2022GL098571

Received 4 MAR 2022
Accepted 24 AUG 2022

Author Contributions:

Conceptualization: Felix M. Strnad, Jakob Schlör, Bedartha Goswami
Formal analysis: Felix M. Strnad, Jakob Schlör, Bedartha Goswami
Investigation: Felix M. Strnad, Jakob Schlör, Christian Fröhlich, Bedartha Goswami
Methodology: Felix M. Strnad, Jakob Schlör, Christian Fröhlich
Software: Felix M. Strnad, Jakob Schlör, Bedartha Goswami
Supervision: Bedartha Goswami
Validation: Felix M. Strnad, Jakob Schlör, Bedartha Goswami

Abstract The diversity of El Niño events is commonly described by two distinct flavors, the Eastern Pacific (EP) and Central Pacific (CP) type. While the remote impacts, that is, teleconnections, of EP and CP events have been studied for different regions individually, a global picture of their structure is still lacking. Here, we use Forman-Ricci curvature applied on climate networks constructed from surface air temperature data to distinguish regional links from teleconnections. Our results confirm that both El Niño types influence the teleconnection patterns, however, with different spatial manifestations. Our analysis suggests that EP El Niños alter the general circulation which changes the teleconnection structure to primarily tropical teleconnections. In contrast, the teleconnection pattern of CP El Niños show only subtle changes to normal conditions. Moreover, this work identifies the dynamics of the Eastern Pacific as a proxy for the remote impact of both El Niño types.

Plain Language Summary El Niño events, characterized by anomalous sea surface temperatures (SSTs) in the Tropical Pacific, come in two flavors; Eastern Pacific (EP) and Central Pacific (CP) types, depending on the longitudinal location of the strongest SST anomalies. Their remote impacts, known as teleconnections, differ. Although there are many studies investigating teleconnections of EP and CP events for individual target regions, a global analysis of the spatial distribution of their teleconnections is still lacking. In this study, we use the theory of complex networks to study EP and CP El Niño teleconnections. We construct “climate networks” from global surface air temperature data and use the notion of “curvature” of a network link to uncover their spatial organization. We show that the most negatively curved links highlight important teleconnection patterns that differ depending on the El Niño type. EP events change the teleconnection structure to the tropics while CP and Normal year conditions reveal teleconnections to all latitudes. Interestingly, the Central Pacific does not show many teleconnections, even during CP El Niño events which we attribute to the varying location of warm water anomalies in the Central Pacific. The Eastern Pacific changes more consistently allowing identifying remote impacts of both El Niños types.

1. Introduction

The El Niño-Southern Oscillation (ENSO) is the most dominant interannual variation in the global climate system. It is a dynamical atmospheric and oceanic phenomenon characterized by anomalously warm (El Niño) or cold (La Niña) phases of sea surface temperatures (SSTs) in the Equatorial Pacific. Both phases are known to impact Earth's climate significantly on large spatial scales, typically referred to as teleconnections (Trenberth, 1997) and thus have been investigated in many studies over the past two decades (Capotondi et al., 2015, 2020; Timmermann et al., 2018).

Yet, significant differences in the downstream impacts of El Niño events are reported (Shi et al., 2019), depending on the amplitude and spatial position of SST anomalies. These differences can be partly related to the type of El Niño. The diversity of El Niño events is typically characterized by two modes: The “canonical” or Eastern Pacific (EP) El Niño (Rasmusson & Carpenter, 1982) with peak SST anomalies in the eastern equatorial Pacific, and the “El Niño Modoki” (Ashok et al., 2007) or Central Pacific (CP) El Niño with peak SST anomalies in the Central Equatorial Pacific (Kao & Yu, 2009). Although the effect of both El Niño types on different locations of the Earth—such as the Indian Ocean (IO) (e.g., Klein et al., 1999), maritime continent (e.g., G. Wang and Hendon 2007), tropical Atlantic (e.g., Huang 2004), and Northern America (e.g., Yu et al., 2012)—has been studied thoroughly (see Okumura 2019 and Taschetto et al., 2020 for an overview), previous work has mainly focused on single teleconnections of the El Niño types.

In comparison, little is known about differences in the spatial extent of global teleconnection patterns between EP and CP events. In this study, we address this issue by introducing a novel machine learning approach that

Visualization: Felix M. Strnad, Jakob Schlör
Writing – original draft: Felix M. Strnad, Jakob Schlör
Writing – review & editing: Felix M. Strnad, Jakob Schlör, Christian Fröhlich, Bedartha Goswami

employs climate networks combined with Ricci-curvature, an abstract tool from the theory of complex networks. Climate networks (Dijkstra et al., 2019) have gained increasing interest in the analysis of spatial dependencies of climatic variables through their ability to reduce data to relevant climatic patterns, and therefore, have been widely used in the analysis of ENSO. Tsonis and Swanson (2008) investigated the topology of El Niño and La Niña networks of surface air temperature. The global impact of El Niño on various geographical zones (Yamasaki et al., 2008), on geographical long-range teleconnections (Donges et al., 2009; Zhou et al., 2015), and its diversity (Kittel et al., 2021; Radebach et al., 2013) have been analyzed by evolving climate network analyses. Wiedermann et al. (2016) employ them to find a robust way to distinguish between different types of El Niños flavors and Lu et al. (2020) estimate the expected El Niño impacts by using climate networks. Ricci-curvature of complex networks is a recent approach to visualize the structure of a network intuitively by highlighting whether an edge of the network connects nodes within a community (i.e., a group of densely connected nodes) or bridges two communities (Ollivier, 2009; Sreejith et al., 2016). It has been proven useful, for example, in the analysis of financial markets (R. S. Sandhu et al., 2016), gene expressions (Pouryahya et al., 2018; R. Sandhu et al., 2015), brain connectivity (Farooq et al., 2019), urban transportation (Gao et al., 2019), power grids (Jonckheere & Grippo, 2019), and epidemiology (de Souza et al., 2021). We show that our approach is an intuitive yet informative tool to analyze the spatial organization of teleconnections that outlines structural differences between EP and CP El Niño impacts.

2. Data and Methods

2.1. Data

We use daily surface air temperature (SAT, 2-m air temperature) data for the years 1959–2020 from the ERA5 Global Reanalysis database (Hersbach et al., 2020). We first detrend each time series, then subtract the daily climatology of the whole time period resulting in anomaly time series with respect to the day of the year. We use next-neighbor interpolation to map the data to a grid of spatially approximately uniformly distributed points using the Fekete algorithm (Bendito et al., 2007) to avoid spurious correlation patterns close to the poles (Ebert-Uphoff & Deng, 2012) (see Section 4.1 and Figure S5 in Supporting Information S1). The distance between grid points in the Fekete grid corresponds to the distance between two points at the equator of a 2.5° Gaussian grid, resulting in a total of ≈6,000 grid points.

2.2. Classification of EP and CP El Niño Conditions

We use December–February daily SAT anomalies and select EP El Niño, CP El Niño, and “Normal” winters based on the average DJF SST anomalies in the Niño 3 (N3) and Niño 4 (N4) region (Trenberth & Stepaniak, 2001). We classify a winter as EP (CP) event if N3 is greater (less) than N4 and N3 (N4) larger than 0.5 (Capotondi et al., 2020). Winters with N3 and N4 between −0.5 and 0.5 are labeled as “Normal”. The SST anomalies are calculated using a centered sliding-window 30-year base period, successively updated in 5-year steps (see Section 1 and Table S1 in Supporting Information S1). Although the onset and duration of El Niño events show high variability, we restrict our analysis to the December–February period where El Niño events show in general the highest intensities to avoid seasonality effects in our analysis.

2.3. Ricci Curvature of Correlation-Based Climate Networks

We demonstrate the process of computing the curvature of a climate network using a toy data set. Figure 1a denotes a set of time series distributed over a sphere. Correlations between all pairs of time series are *predefined* (in contrast to our actual SAT data set) with a priori fixed covariance structure created from a stochastic block model (see Section 2 in Supporting Information S1). In order to calculate the Ricci curvature, we first need to create a climate network, which signifies the most related grid points over the entire globe. The network is determined by considering the 2% strongest ($\rho_{0.98}$) statistically significant correlations ρ_{ij} between pairs of time series $x_i(t)$, $x_j(t)$ which is described by the weighted adjacency matrix $\mathbf{W}_{ij}$ (Figure 1b):

$$\mathbf{W}_{ij} = \begin{cases} |\rho_{ij}|, & |\rho_{ij}| > \rho_{0.98}, \\ 0, & \text{otherwise}. \end{cases} \tag{1}$$

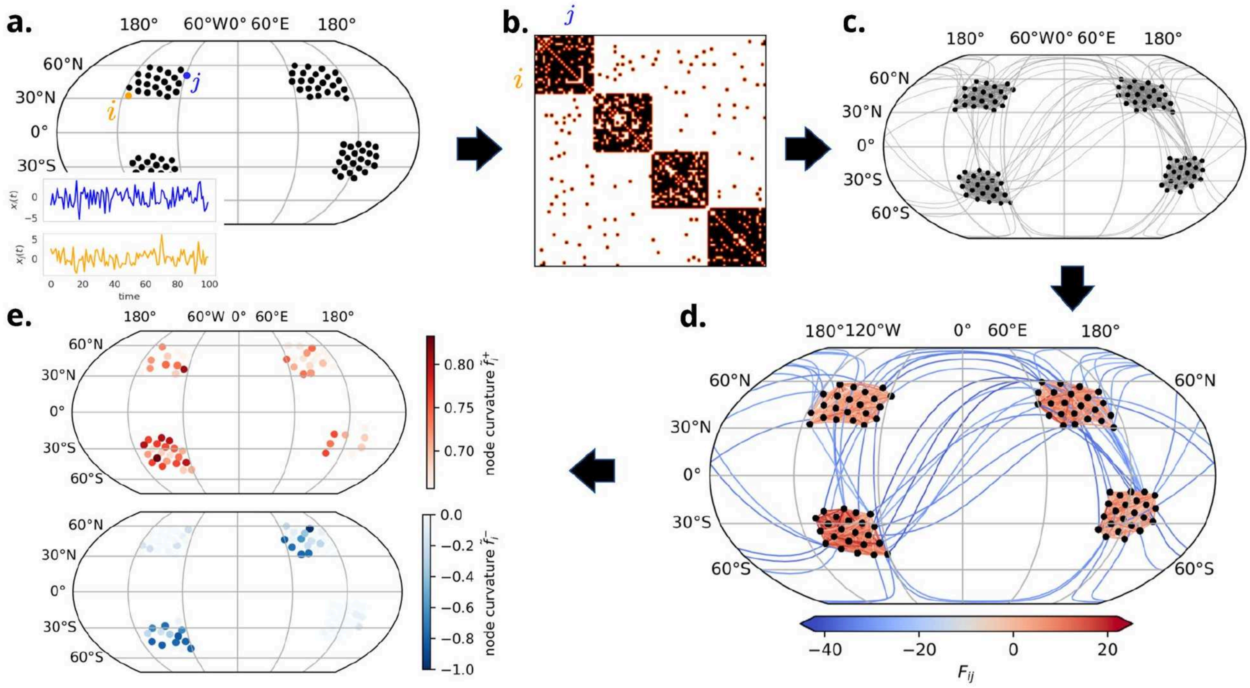


Figure 1. Construction of a climate network and its curvature measure for a toy data set. By computing correlations of time series between all pairs of locations (a), we obtain the adjacency matrix $\mathbf{W}_{ij}$ (Equation 1), orange dots denote points of significant correlations between time series $x_i(t)$, $x_j(t)$ (b), visualized as a network graph (c). The Forman curvature F_{ij} of each edge (d) reveals the property of an edge as within-community (positive curvature) or bridging communities (negative curvature). Positive (negative) node curvature hotspots f_{ij}^+ $\left(f_{ij}^-\right)$ (e) are obtained by aggregating the 10% most positively and 10% most negatively curved edges on the nodes.

This is equivalent to a network G whose nodes $V = \{v_i: i = 1, .., N\}$ correspond to spatial locations $i = 1, \ldots, N$ with edge e_{ij} connecting the nodes $(v_i, v_j) \in V$ if $|\rho_{ij}| > \rho_{0.98}$ (Figure 1c) weighted by ρ_{ij} (for details see Section 3.2 in Supporting Information S1).

Ricci curvature provides a continuous measure over network links (i.e., independent from the spatial grid) using the connectivity between points to describe their relationship in the network. Figure 1d shows the network of our toy example with their edges e_{ij} colored by their curvature F_{ij}.

We see that positive curvature indicates an edge being embedded within a community of nodes. The neighborhood of a positively curved edge is more densely connected than a regular graph which is a network with each node containing the same number of neighboring nodes. Positive curvature thus indicates a community of nodes with similar dynamics. Negative curvature corresponds to an edge connecting two communities, with its own neighborhood containing fewer connections than a regular graph. It thus can indicate teleconnections as shown in Section 3.1.

Since multiple edges are difficult to visualize ($\approx 10^6$ links for the SAT network), we define node curvature f_i of the node v_i as the summation of the edge-based measure F_{ij} of all edges e_{ij} connected to v_i. The node curvature is easier to visualize on a map and allows to identify geographical locations connected to strongly negatively or positively curved links. More precisely, we define the upper 90th (lower 10th) percentile of all edges as F_{ij}^+ $\left(F_{ij}^-\right)$ and analogously, and the aggregation of the upper 90th (lower 10th) percentile of the curvature values as f_i^+ $\left(f_i^-\right)$, which we denote as "hotspots" (Figure 1e). As the value ranges differ between networks, we normalize them by using the min-max transformation to $(-1, 1)$, denoted by $\tilde{F}_{ij}$ and analogously, $\tilde{f}_i$ (Equation 9 in Supporting Information S1). The intensity of the positive node-curvature hotspots $\tilde{f}_i^+$ reveals the communities with the highest edge probabilities in the stochastic block model outlining a community of nodes that behave similarly in time. Negative node-curvature hotspots $\tilde{f}_i^-$ coincide with the underlying probabilities used for constructing the data set indicating locations that are highly teleconnected on the globe.

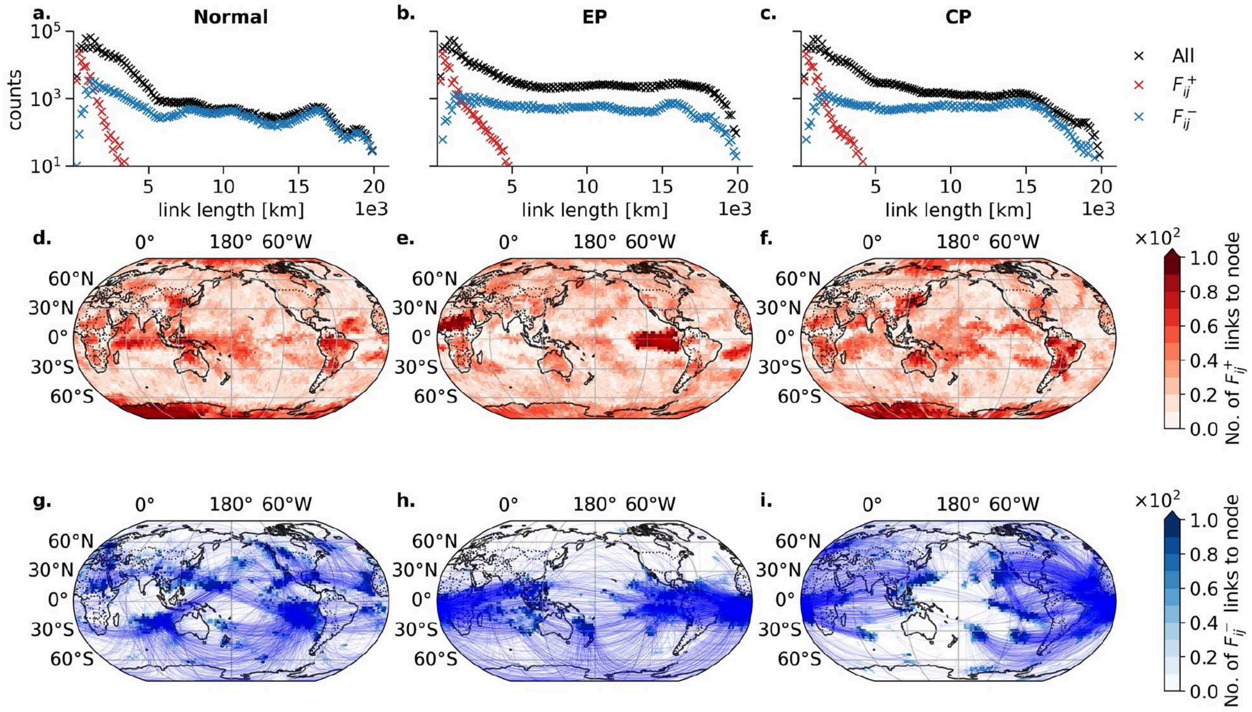


Figure 2. Forman curvature of Normal, Eastern Pacific (EP), and Central Pacific (CP) El Niño climate network links. The networks are computed from surface air temperature (SAT) anomalies for Normal (left column), EP (middle column), and CP (right column) conditions. The first row (a–c) depicts the spatial link length distribution for all (black), most positively F_{ij}^+ (red) and most negatively curved F_{ij}^- (blue) edges. The second row shows F_{ij}^+ (d–f), the third row F_{ij}^- (g–i). Colorbars indicate the number of incoming edges to a node. For visual reasons only every 20th edge is plotted in (d–i).

Of the two numerical approximations of Ricci curvature on networks—Forman-Ricci curvature (Forman, 2003; Sreejith et al., 2016) and Ollivier-Ricci curvature (Ollivier, 2010)—we use Forman-Ricci curvature (henceforth simply Forman curvature) as it is computationally cheaper. Both definitions are highly correlated, barring slight differences in extreme values (Samal et al., 2018). For details on the mathematical description of the curvature method, we refer to Section 4 in Supporting Information S1.

3. Results and Discussion

3.1. Spatial Organization of Teleconnections Depends on El Niño Type

We compute correlation-based climate networks and their Forman curvature using global SAT data for EP, CP, and Normal conditions. Figures 2a–2c show the distribution of spatial link lengths as the shortest distance between two connected points on the globe, the so-called great-circle lengths. Shown are all links of the network (black markers), the most positively curved F_{ij}^+ (red) and most negatively curved F_{ij}^- (blue) network links for Normal, EP, and CP conditions.

The spatial distribution of F_{ij}^+ (Figures 2d–2f) shows that positive curvature occurs at only regional scales $\leq 5 \cdot 10^3$ km, resulting in the enforcement of local community structures. While F_{ij}^- links occur at all spatial scales, for lengths $\geq 10^4$ km, the F_{ij}^- and "All" link distributions almost overlaps for CP and Normal conditions. Long-range teleconnections are here thus modulated predominantly via negatively curved links. As the curvature estimation of climate network links does not include any information on the spatial length, this suggests that the relationship between negative climate network link curvature and long spatial scales is an intrinsic topological property of the SAT dynamics. Therefore, negative curvature results in link bundles that are related to well-known teleconnection patterns which are discussed in Section 3.3.

In contrast to CP and Normal conditions, for EP El Niño conditions the distribution of F_{ij}^{-} does not describe all spatially long-range links (Figure 2b) and the most negative network links undergo a drastic spatial reorganization (compare Figure 2h to Figures 2g and 2i). In particular, we observe that the connection between the Tropical Pacific and the southern Atlantic is strengthened during EP conditions (Figure 2h), while during CP El Niños the Tropical Pacific shows an enhanced connection to the extratropical Pacific as well as to the mid-latitude North Atlantic region (Figure 2i, Section 3.3).

Changes in most positively curved links F_{ij}^{+} are more subtle between Normal (Figures 2d and 2f), CP (Figure 2f) and EP conditions (Figure 2e), such as the weakening of regional correlation structures in the tropical Atlantic (CP conditions) and the strengthening of correlation in the Eastern Pacific and the West African monsoon belt (EP conditions).

Note, we mainly discover teleconnections within the global oceans because correlations are generally higher over oceans than over land due to slower oceanic SAT variability (Lambert et al., 2011).

3.2. EP El Niño Conditions Lead to Enforcement of Teleconnections to the Tropics

The global teleconnection pattern is under EP conditions substantially different when compared to Normal conditions (Figures 2 and 3). We suggest that the differences can be attributed to the intensity of the events. EP events tend to have higher temperature anomalies (Figures S2 and S3 in Supporting Information S1) and more stable warming locations in the tropics than CP events (Figure S4 in Supporting Information S1) leading to a strong chan the global atmospheric circulation which results in the confinement of hotspots around the tropics for both regional links and teleconnections. In contrast, CP conditions show unstable warming locations (Figures S2 and S4 in Supporting Information S1) and a weaker amplitude and thus cannot alter the general circulation. Therefore, Normal and CP networks reveal various similarities resulting in positive and negative curvature hotspots over all latitudes (Figure 3). This is further confirmed by the zonal medians (Figures 3g–3i). We repeat our analysis for strong and moderate El Niño events separately (see Section 1 in Supporting Information S1) and find that the EP curvature results (Figure 3) are mainly dominated by the strong EP El Niños (Figures S16 and S17 in Supporting Information S1). We thus conclude that the differences between EP conditions to CP and Normal conditions are driven by strong EP El Niño events. This corroborates previous works by Lu et al. (2020), Wiedermann et al. (2016) who report a strong localization of climate network links during EP conditions as well.

For EP conditions, positive curvature hotspots $\tilde{f}_i^{+}$ reveal the well-known ENSO tongue (Figure 3b) typically observed in empirical orthogonal function analysis of SST data (Johnson, 2013). We also find pronounced regions of $\tilde{f}_i^{+}$ in the IO and the tropical Atlantic which are known to be affected by strong EP El Niño events (Klein et al., 1999; Rodrigues et al., 2015; X. Wang & Wang, 2014; W. Zhang et al., 2015). Under CP conditions, $\tilde{f}_i^{+}$ is spread over all latitudes and over different regions of the globe (Figure 3c). For instance, we observe a hotspot in the Tropical Pacific similar to the El Niño tongue, which is however shifted toward the dateline and also extended southwards.

Negative node-curvature hotspots, $\tilde{f}_i^{-}$, for EP conditions (Figure 3e) show enhanced teleconnections in the South China Sea, tropical IO, Eastern Tropical Pacific, and the tropical Atlantic. This coincides with a decrease in teleconnections in the extratropical Pacific, southern IO, North Atlantic (near Greenland), and the Southern Ocean. Under CP conditions (Figure 3f), we find negative node curvature hotspots in the extratropics and mid-latitudes similar to $\tilde{f}_i^{+}$. While some similarities between CP and Normal conditions are found (Figures 3d and 3f), for example, at the US-West Coast (Capotondi et al., 2019), in the Northern Tropical Pacific, and in the tropical Atlantic, the major difference can be found over the European continent, the Northern Atlantic, and the Labrador Sea (see as well Figures 2g and 2i).

We repeat our analysis using 1,000 years of pre-industrial run of the UKESM1-0-LL model from the CMIP6 project (Eyring et al., 2016) which shows a comparable variation of SSTA in the Tropical Pacific to observational data (Dieppois et al., 2021). We obtain a localization of positive and negative curvature hotspots to the Tropics under EP conditions and teleconnections to the mid-latitudes under CP conditions (see Figure S14 in Supporting Information S1), similar to our analysis of observational data (Figure 3). The longer time period of

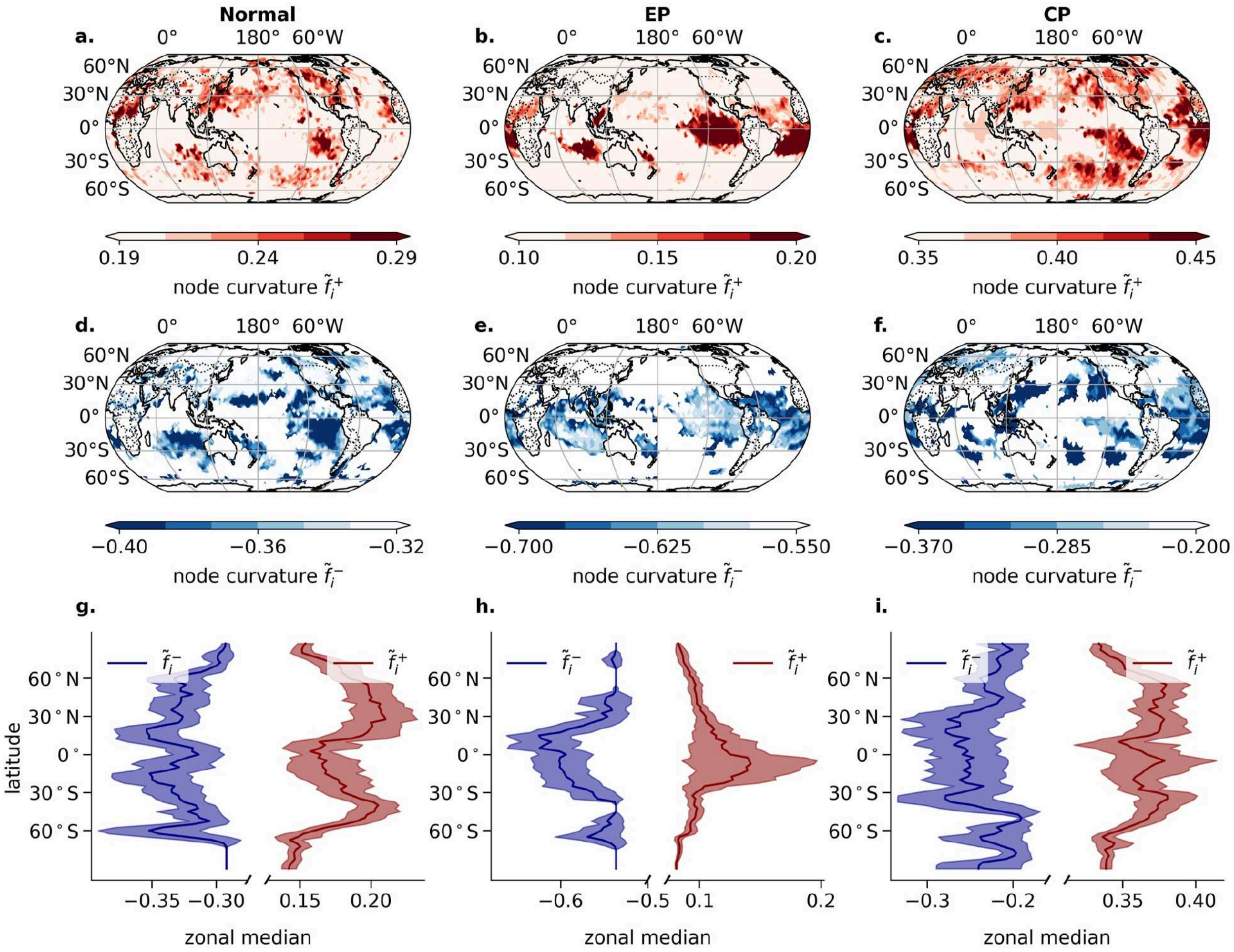


Figure 3. Node curvature of Normal, Eastern Pacific (EP) El Niño, and Central Pacific (CP) El Niño climate networks. Node curvature hotspots for Normal (left column), EP (middle column) CP (right column) hotspots are shown summing the most positively curved $\tilde{f}_i^+$ (a–c), most negatively curved $\tilde{f}_i^-$ (d–f) links, as well as the aggregated zonal medians of $\tilde{f}_i^+$ (red) and $\tilde{f}_i^-$ (blue) (g–i) are shown.

the pre-industrial run allows us to control for the small number of EP and CP events in the reanalysis data set (see Section 7 in Supporting Information S1). We find that the curvature analysis of the reanalysis data with only 7 EP/CP events tend to overestimate the number of negatively curved links. However, qualitatively the spatial patterns of positive and negative curved links are robust in comparison to the 297 EP events obtained in the whole pre-industrial time period (see Figure S15 in Supporting Information S1).

3.3. EP and CP El Niño Teleconnection Patterns of Eastern and Central Pacific Ocean, Indian Ocean, and the Labrador Sea

During EP conditions, teleconnections link the N3 region (Figure 4a) to the tropical Atlantic, supporting earlier work showing that strong El Niños can lead to warming in the tropical Atlantic mediated by the tropospheric temperature mechanism (Chang et al., 2006) and the atmospheric bridge via the Pacific North American (PNA) pattern (Alexander et al., 2002; Rodrigues et al., 2011). The links connecting N3 to the IO are likely apparent because of the influence of Eastern Tropical Pacific SSTs on the IO during and after El Niño events attributed to net heat flux anomalies due to changes in the atmospheric circulation of ENSO (Klein et al., 1999). During CP conditions, we observe links from the N3 box as well, primarily connecting to the extratropical Pacific. These are

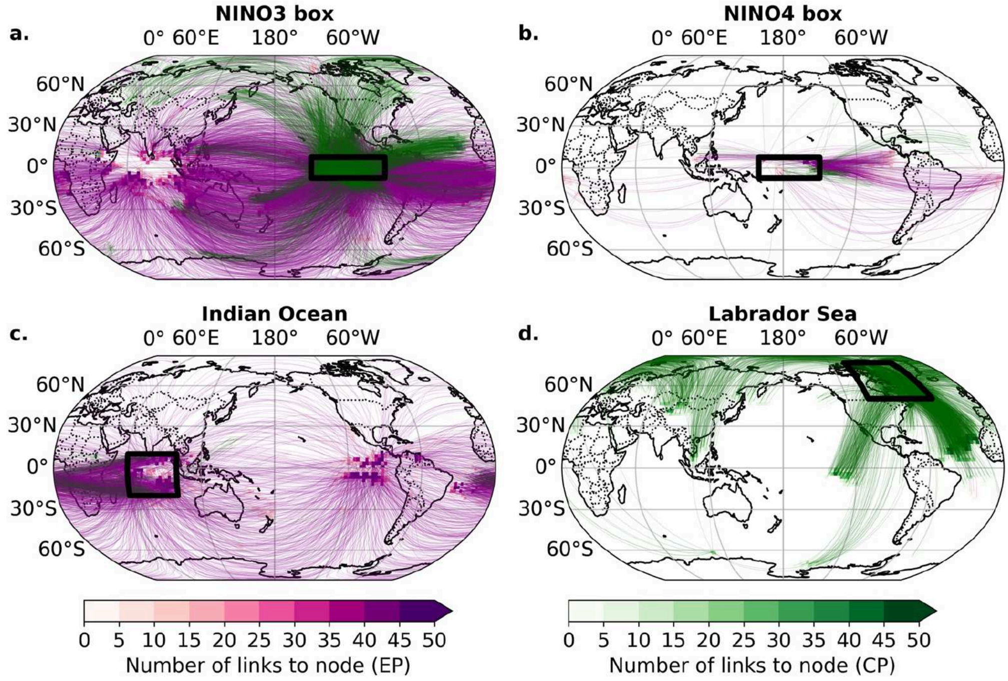


Figure 4. Teleconnections from Eastern and Central Pacific Ocean, Indian Ocean (IO), and the Labrador Sea. F_{ij}^{-}, that is, teleconnections, for EP (purple) and CP (green) El Niño events connected to the four selected regions (black rectangles): Niño 3 (a), Niño 4 (b), IO (c), and Labrador Sea (d). For visual clarity, only a third of all links are shown.

likely due to the North and South Pacific Meridional Mode (NPMM and SPMM), as atmospheric and oceanic anomalies in the extratropics associated with SPMM and NPMM affect the intensity and flavor of El Niño (You & Furtado, 2018). Capotondi and Ricciardulli (2021) showed that the dominant patterns of wind variability in the Northern and Southern Hemispheres, which represent the atmospheric expressions of the NPMM and SPMM, have a stronger relationship with CP events than EP events (Amaya, 2019; Capotondi et al., 2019; Capotondi & Ricciardulli, 2021), consistent with the extratropical links observed during CP events (Figure 4a).

The teleconnections of the N4 region (Figure 4b) are far fewer compared to the N3 region (Figure 4a). It is rather surprising that the N4 region is not well connected even during CP conditions, although a CP El Niño is primarily characterized by higher SST anomalies in the N4 region (Figures S1–S3 in Supporting Information S1). These findings are in apparent disagreement with Barsugli and Sardeshmukh (2002) and Shi et al. (2019) who showed that the PNA and the Northern Pacific are most sensitive to changes of SST in the Central Tropical Pacific using General Circulation Models (GCMs). GCMs are known to produce SST anomalies too far in the west (Capotondi et al., 2006; Ham & Kug, 2012; Kug et al., 2010), however, this can only partially explain the missing teleconnections in the Central Pacific. We find that although anomalies in the Central Pacific are stronger during CP conditions (Figure S1 in Supporting Information S1), the Eastern Pacific shows significantly higher local correlations (Figure S13 in Supporting Information S1). Strong anomalies do not necessarily imply high correlations (compare Section 1 and Figures S2–S4 in Supporting Information S1), therefore, we attribute the weak correlations in the Central Pacific under CP conditions to the higher spatial variability of anomalously warm temperatures between CP El Niño events than between EP El Niño events (Figures S4 and S5 in Supporting Information S1) (Chen et al., 2021; C. Wang & Wang, 2013). During CP conditions, anomalies in the tropical Eastern Pacific are weaker but behave more consistently in time than in the Central Pacific (see Figure S16 in Supporting Information S1). We thus suggest that the teleconnections from the N3 region are artifacts of the correlation but functions as a proxy for the stimulation by the Central Pacific. The results from the curvature analysis using the model output from the UKESM1-0-LL GCM further corroborate our findings. Here as well, N3 teleconnections to the Northern Pacific and west American coast are found in the CP network while the Central Pacific does not exhibit strong teleconnections (Figure S14 in Supporting Information S1). Further analysis using model runs with nudged SSTA could therefore allow disentangling the role of the central and Eastern Pacific on teleconnections to the Northern Pacific and North America.

The IO has a large number of most negatively curved links during EP conditions but not for CP (Figures 3c and 3d). The EP event teleconnections link the IO to the tropical Atlantic and Pacific basins (Figure 4c). The links to the tropical Atlantic could be either attributed to indirect links mediated by the impact of the Ninõ 3 region on the tropical Atlantic or might resemble direct impact between the oceans as recently described by L. Zhang and Han (2021). CP conditions do not result in teleconnections in the IO as CP events are generally weaker (W. Zhang et al., 2015).

The Labrador Sea (Figure 4d) is another pronounced area of most negative node-curvature in the CP network not present in the EP network (compare Figure 3c and 3d. Edges with the most negative curvature adjacent to the Labrador Sea connect to the extra topical Atlantic and, in contrast to Normal year conditions (Figure S13 in Supporting Information S1), to the Eastern Tropical Pacific. This pattern may be attributed to the North Atlantic Oscillation (NAO), which refers to sea level pressure changes in the Arctic and subtropical Atlantic (Jiménez-Esteve & Domeisen, 2018). El Niño is known to cause a negative NAO pattern driven by the PNA pattern where a negative NAO phase is attributable to CP El Niño events via the subtropical bridge (Domeisen et al., 2019; Graf & Zanchettin, 2012).

Note that we can identify the teleconnection patterns primarily because we have the results from the curvature-based climate network analysis to identify differences in the teleconnection structure and to guide our interpretations. Without the curvature analysis, for instance using classical complex network measures like betweenness centrality (Figure S7 in Supporting Information S1), it is not trivial to figure out the most important regions for each El Niño flavor.

4. Conclusion

We presented a new approach to estimate global teleconnection patterns of surface air temperature and used it to investigate the teleconnections of Eastern Pacific and Central Pacific El Niño events. Our approach involves the combination of correlation-based climate networks with a recently established network measure based on Ricci-curvature. In particular, we used Forman-Ricci curvature to distinguish links related to small-scale regional structures (positively curved links) from long-range teleconnections connecting regions from different parts of the globe (negatively curved links).

We showed that El Niño diversity drastically impacts the spatial organization of teleconnections. We identified teleconnection hotspots for both EP and CP conditions and showed that EP conditions strongly alternate the teleconnection structure of the climate network to be mainly confined to the tropics, whereas CP network teleconnections were found in northern and southern mid-latitudes as well. A comparison of the results from the application of our approach on SAT from the UKESM-LL2 Model from the CMIP6 project further corroborates our findings.

We further investigated the impact of ENSO diversity on the teleconnection patterns of four specific regions: the Niño 3 region, the Niño 4 region, the northern IO, and the Labrador Sea. We found that the Niño 3 region in the Eastern Pacific has a large number of teleconnections irrespective of whether we consider EP or CP conditions. The Niño 4 region in the Central Pacific has, in comparison, very few teleconnections under both EP and CP conditions and our analysis suggests that this is likely due to the higher spatio-temporal variability of anomalous SSTs in the Central Pacific. We conclude that the Eastern Pacific is the primary mediator of El Niño impacts irrespective of the El Niño type, and acknowledge that further work on the role of the Eastern Pacific during CP El Niños is needed.

We found that the northern IO and the Labrador Sea show teleconnections almost only under EP and CP conditions respectively. While the teleconnections of the IO region to the Niño 3 region during EP conditions are well-known, the links between the Labrador Sea to the Eastern Tropical Pacific and northern tropical Atlantic are not fully understood.

Data Availability Statement

Data sets for this research are available from Copernicus Climate Change Service. The data from 1959 till date was taken from Hersbach et al. (2018). The code for generating and analyzing the networks is made publicly available under Strnad and Schlör (2022). The code for reproducing the analysis of the network curvature described in this paper is publicly available under Schlör and Strnad (2022).

Acknowledgments

The authors would like to thank the reviewers for their helpful comments and suggestions. The authors acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC number 2064/1—Project number 390727645. The authors thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting Jakob Schlör and Felix Strnad. Open Access funding enabled and organized by Projekt DEAL.

References

Alexander, M. A., Bladé, I., Newman, M., Lanzante, J. R., Lau, N.-C., & Scott, J. D. (2002). The atmospheric bridge: The influence of ENSO teleconnections on air–sea interaction over the Global Oceans. *Journal of Climate*, *15*(16), 2205–2231. https://doi.org/10.1175/1520-0442(2002)015<2205:TABTIO>2.0.CO;2

Amaya, D. J. (2019). The Pacific Meridional Mode and ENSO: A review. *Current Climate Change Reports*, *5*(4), 296–307. https://doi.org/10.1007/s40641-019-00142-x

Ashok, K., Behera, S. K., Rao, S. A., Weng, H., & Yamagata, T. (2007). El Niño Modoki and its possible teleconnection. *Journal of Geophysical Research*, *112*(C11), C11007. https://doi.org/10.1029/2006JC003798

Barsugli, J. J., & Sardeshmukh, P. D. (2002). Global atmospheric sensitivity to tropical SST anomalies throughout the Indo-Pacific basin. *Journal of Climate*, *15*(23), 3427–3442. https://doi.org/10.1175/1520-0442(2002)015<3427:GASTTS>2.0.CO;2

Bendito, E., Carmona, A., Encinas, A. M., & Gesto, J. M. (2007). Estimation of Fekete points. *Journal of Computational Physics*, *225*(2), 2354–2376. https://doi.org/10.1016/j.jcp.2007.03.017

Capotondi, A., & Ricciardulli, L. (2021). The influence of pacific winds on ENSO diversity. *Scientific Reports*, *11*(18672), 1–11. https://doi.org/10.1038/s41598-021-97963-4

Capotondi, A., Sardeshmukh, P. D., Di Lorenzo, E., Subramanian, A. C., & Miller, A. J. (2019). Predictability of US West Coast ocean temperatures is not solely due to ENSO. *Scientific Reports*, *9*(10993), 1–10. https://doi.org/10.1038/s41598-019-47400-4

Capotondi, A., Wittenberg, A., & Masina, S. (2006). Spatial and temporal structure of Tropical Pacific interannual variability in 20th century coupled simulations. *Ocean Modelling*, *15*(3), 274–298. https://doi.org/10.1016/j.ocemod.2006.02.004

Capotondi, A., Wittenberg, A. T., Kug, J.-S., Takahashi, K., & McPhaden, M. J. (2020). ENSO diversity. In *El Niño Southern Oscillation in a changing climate* (pp. 65–86). American Geophysical Union (AGU). https://doi.org/10.1002/9781119548164.ch4

Capotondi, A., Wittenberg, A. T., Newman, M., Di Lorenzo, E., Yu, J.-Y., Braconnot, P., et al. (2015). Understanding ENSO diversity. *Bulletin of the American Meteorological Society*, *96*(6), 921–938. https://doi.org/10.1175/BAMS-D-13-00117.1

Chang, P., Fang, Y., Saravanan, R., Ji, L., & Seidel, H. (2006). The cause of the fragile relationship between the Pacific El Niño and the Atlantic Niño. *Nature*, *443*(7109), 324–328. https://doi.org/10.1038/nature05053

Chen, M., Yu, J.-Y., Wang, X., & Chen, S. (2021). Distinct onset mechanisms of two subtypes of CP El Niño and their changes in future warming. *Geophysical Research Letters*, *48*(14), e2021GL093707. https://doi.org/10.1029/2021GL093707

de Souza, D. B., da Cunha, J. T. S., dos Santos, E. F., Correia, J. B., da Silva, H. P., de Lima Filho, J. L., et al. (2021). Using discrete Ricci curvatures to infer COVID-19 epidemic network fragility and systemic risk. *Journal of Statistical Mechanics: Theory and Experiment*, *2021*(5), 053501. https://doi.org/10.1088/1742-5468/abed4e

Dieppois, B., Capotondi, A., Pohl, B., Chun, K. P., Monerie, P.-A., & Eden, J. (2021). ENSO diversity shows robust decadal variations that must be captured for accurate future projections. *Communications Earth & Environment*, *2*(212), 1–13. https://doi.org/10.1038/s43247-021-00285-6

Dijkstra, H. A., Hernández-García, E., Masoller, C., & Barreiro, M. (2019). Networks in climate. https://doi.org/10.1017/9781316275757

Domeisen, D. I. V., Garfinkel, C. I., & Butler, A. H. (2019). The teleconnection of El Niño Southern Oscillation to the stratosphere. *Review of Geophysics*, *57*(1), 5–47. https://doi.org/10.1029/2018RG000596

Donges, J. F., Zou, Y., Marwan, N., & Kurths, J. (2009). Complex networks in climate dynamics. *The European Physical Journal—Special Topics*, *174*(1), 157–179. https://doi.org/10.1140/epjst/e2009-01098-2

Ebert-Uphoff, I., & Deng, Y. (2012). A new type of climate network based on probabilistic graphical models: Results of boreal winter versus summer. *Geophysical Research Letters*, *39*(18), 1–7. https://doi.org/10.1029/2012GL053269

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Farooq, H., Chen, Y., Georgiou, T. T., Tannenbaum, A., & Lenglet, C. (2019). Network curvature as a hallmark of brain structural connectivity. *Nature Communications*, *10*(4937), 1–11. https://doi.org/10.1038/s41467-019-12915-x

Forman, R. (2003). Bochner's method for cell complexes and combinatorial Ricci curvature. *Discrete & Computational Geometry*, *29*(3), 323–374. https://doi.org/10.1007/s00454-002-0743-x

Gao, L., Liu, X., Liu, Y., Wang, P., Deng, M., Zhu, Q., & Li, H. (2019). Measuring road network topology vulnerability by Ricci curvature. *Physica A*, *527*, 121071. https://doi.org/10.1016/j.physa.2019.121071

Graf, H.-F., & Zanchettin, D. (2012). Central Pacific El Niño, the "subtropical bridge," and Eurasian climate. *Journal of Geophysical Research*, *117*(D1), 1102. https://doi.org/10.1029/2011JD016493

Ham, Y.-G., & Kug, J.-S. (2012). How well do current climate models simulate two types of El Nino? *Climate Dynamics*, *39*(1), 383–398. https://doi.org/10.1007/s00382-011-1157-3

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2018). Era5 hourly data on single levels from 1979 to present. Copernicus climate change Service (C3S) climate data Store (CDS). https://doi.org/10.24381/cds.adbb2d47

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

Huang, B. (2004). Remotely forced variability in the Tropical Atlantic Ocean. *Climate Dynamics*, *23*(2), 133–152. https://doi.org/10.1007/s00382-004-0443-8

Jiménez-Esteve, B., & Domeisen, D. I. V. (2018). The tropospheric pathway of the ENSO–North Atlantic teleconnection. *Journal of Climate*, *31*(11), 4563–4584. https://doi.org/10.1175/JCLI-D-17-0716.1

Johnson, N. C. (2013). How many ENSO flavors can we distinguish? *Journal of Climate*, *26*(13), 4816–4827. https://doi.org/10.1175/JCLI-D-12-00649.1

Jonckheere, E., & Grippo, E. (2019). Ollivier-Ricci curvature approach to cost-effective power grid congestion management. In *2019 Chinese control and decision conference (CCDC)* (pp. 2118–2123). IEEE. https://doi.org/10.1109/CCDC.2019.8832819

Kao, H.-Y., & Yu, J.-Y. (2009). Contrasting eastern-Pacific and Central-Pacific types of ENSO. *Journal of Climate*, *22*(3), 615–632. https://doi.org/10.1175/2008JCLI2309.1

Kittel, T., Ciemer, C., Lotfi, N., Peron, T., Rodrigues, F., Kurths, J., & Donner, R. V. (2021). Evolving climate network perspectives on global surface air temperature effects of ENSO and strong volcanic eruptions. *The European Physical Journal—Special Topics*, *230*(14), 3075–3100. https://doi.org/10.1140/epjs/s11734-021-00269-9

Klein, S. A., Soden, B. J., & Lau, N.-C. (1999). Remote sea surface temperature variations during ENSO: Evidence for a tropical atmospheric bridge. *Journal of Climate*, *12*(4), 917–932. https://doi.org/10.1175/1520-0442(1999)012<0917:RSSTVD>2.0.CO;2

Kug, J.-S., Choi, J., An, S.-I., Jin, F.-F., & Wittenberg, A. T. (2010). Warm pool and cold tongue El Niño events as simulated by the GFDL 2.1 coupled GCM. *Journal of Climate*, *23*(5), 1226–1239. https://doi.org/10.1175/2009JCLI3293.1

Lambert, F. H., Webb, M. J., & Joshi, M. M. (2011). The relationship between land–ocean surface temperature contrast and radiative forcing. *Journal of Climate*, *24*(13), 3239–3256. https://doi.org/10.1175/2011JCLI3893.1

Lu, Z., Yuan, N., Chen, L., & Gong, Z. (2020). On the impacts of El Niño events: A new monitoring approach using complex network analysis. *Geophysical Research Letters*, *47*(6), e2019GL086533. https://doi.org/10.1029/2019GL086533

Okumura, Y. M. (2019). ENSO diversity from an atmospheric perspective. *Current Climate Change Reports*, *5*(3), 245–257. https://doi.org/10.1007/s40641-019-00138-7

Ollivier, Y. (2009). Ricci curvature of Markov chains on metric spaces. *Journal of Functional Analysis*, *256*(3), 810–864. https://doi.org/10.1016/j.jfa.2008.11.001

Ollivier, Y. (2010). A survey of Ricci curvature for metric spaces and Markov chains. In *Probabilistic approach to geometry* (pp. 343–381). Mathematical Society of Japan.

Pouryahya, M., Oh, J. H., Mathews, J. C., Deasy, J. O., & Tannenbaum, A. R. (2018). Characterizing cancer drug response and biological correlates: A geometric network approach. *Scientific Reports*, *8*(6402), 1–12. https://doi.org/10.1038/s41598-018-24679-3

Radebach, A., Donner, R. V., Runge, J., Donges, J. F., & Kurths, J. (2013). Disentangling different types of El Niño episodes by evolving climate network analysis. *Physical Review E—Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *88*(5), 052807. https://doi.org/10.1103/PhysRevE.88.052807

Rasmusson, E. M., & Carpenter, T. H. (1982). Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. *Monthly Weather Review*, *110*(5), 354–384. https://doi.org/10.1175/1520-0493(1982)110<0354:vitsst>2.0.co;2

Rodrigues, R. R., Campos, E. J. D., & Haarsma, R. (2015). The impact of ENSO on the South Atlantic subtropical dipole mode. *Journal of Climate*, *28*(7), 2691–2705. https://doi.org/10.1175/JCLI-D-14-00483.1

Rodrigues, R. R., Haarsma, R. J., Campos, E. J. D., & Ambrizzi, T. (2011). The impacts of Inter–El Niño variability on the Tropical Atlantic and Northeast Brazil climate. *Journal of Climate*, *24*(13), 3402–3422. https://doi.org/10.1175/2011JCLI3983.1

Samal, A., Sreejith, R. P., Gu, J., Liu, S., Saucan, E., & Jost, J. (2018). Comparative analysis of two discretizations of Ricci curvature for complex networks. *Scientific Reports*, *8*(8650), 1–16. https://doi.org/10.1038/s41598-018-27001-3

Sandhu, R., Georgiou, T., Reznik, E., Zhu, L., Kolesov, I., Senbabaoglu, Y., & Tannenbaum, A. (2015). Graph curvature for differentiating cancer networks. *Scientific Reports*, *5*(12323), 1–13. https://doi.org/10.1038/srep12323

Sandhu, R. S., Georgiou, T. T., & Tannenbaum, A. R. (2016). Ricci curvature: An economic indicator for market fragility and systemic risk. *Science Advances*, *2*(5). https://doi.org/10.1126/sciadv.1501495%23F2

Schlör, J., & Strnad, F. (2022). Netcurvature. *Zenodo*. https://doi.org/10.5281/zenodo.6325299

Shi, J., Fedorov, A. V., & Hu, S. (2019). North Pacific temperature and precipitation response to El Niño-like equatorial heating: Sensitivity to forcing location. *Climate Dynamics*, *53*(5), 2731–2741. https://doi.org/10.1007/s00382-019-04655-x

Sreejith, R., Mohanraj, K., Jost, J., Saucan, E., & Samal, A. (2016). Forman curvature for complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2016*(6), 063206. https://doi.org/10.1088/1742-5468/2016/06/063206

Strnad, F., & Schlör, J. (2022). Climnet. *Zenodo*. https://doi.org/10.5281/zenodo.6325661

Taschetto, A. S., Ummenhofer, C. C., Stuecker, M. F., Dommenget, D., Ashok, K., Rodrigues, R. R., & Yeh, S.-W. (2020). ENSO atmospheric teleconnections. In *El Niño Southern Oscillation in a changing climate* (pp. 309–335). American Geophysical Union (AGU). https://doi.org/10.1002/9781119548164.ch14

Timmermann, A., An, S.-I., Kug, J.-S., Jin, F.-F., Cai, W., Capotondi, A., et al. (2018). El Niño–Southern Oscillation complexity. *Nature*, *559*(7715), 535–545. https://doi.org/10.1038/s41586-018-0252-6

Trenberth, K. E. (1997). The definition of El Niño. *Bulletin of the American Meteorological Society*, *78*(12), 2771–2778. https://doi.org/10.1175/1520-0477(1997)078<2771:tdoeno>2.0.co;2

Trenberth, K. E., & Stepaniak, D. P. (2001). Indices of El Niño Evolution. *Journal of Climate*, *14*(8), 1697–1701. https://doi.org/10.1175/1520-0442(2001)014<1697:lioeno>2.0.co;2

Tsonis, A. A., & Swanson, K. L. (2008). Topology and predictability of El Niño and La Niña networks. *Physical Review Letters*, *100*(22), 228502. https://doi.org/10.1103/PhysRevLett.100.228502

Wang, C., & Wang, X. (2013). Classifying El Niño Modoki I and II by different impacts on rainfall in Southern China and typhoon tracks. *Journal of Climate*, *26*(4), 1322–1338. https://doi.org/10.1175/JCLI-D-12-00107.1

Wang, G., & Hendon, H. H. (2007). Sensitivity of Australian rainfall to Inter–El Niño variations. *Journal of Climate*, *20*(16), 4211–4226. https://doi.org/10.1175/JCLI4228.1

Wang, X., & Wang, C. (2014). Different impacts of various El Niño events on the Indian Ocean dipole. *Climate Dynamics*, *42*(3), 991–1005. https://doi.org/10.1007/s00382-013-1711-2

Wiedermann, M., Radebach, A., Donges, J. F., Kurths, J., & Donner, R. V. (2016). A climate network-based index to discriminate different types of El Niño and La Niña. *Geophysical Research Letters*, *43*(13), 7176–7185. https://doi.org/10.1002/2016GL069119

Yamasaki, K., Gozolchiani, A., & Havlin, S. (2008). Climate networks around the globe are significantly affected by El Niño. *Physical Review Letters*, *100*(22), 228501. https://doi.org/10.1103/PhysRevLett.100.228501

You, Y., & Furtado, J. C. (2018). The South Pacific Meridional Mode and its role in tropical Pacific climate variability. *Journal of Climate*, *31*(24), 10141–10163. https://doi.org/10.1175/JCLI-D-17-0860.1

Yu, J.-Y., Zou, Y., Kim, S. T., & Lee, T. (2012). The changing impact of El Niño on US winter temperatures. *Geophysical Research Letters*, *39*(15). https://doi.org/10.1029/2012GL052483

Zhang, L., & Han, W. (2021). Indian Ocean dipole leads to Atlantic Niño. *Nature Communications*, *12*(5952), 1–9. https://doi.org/10.1038/s41467-021-26223-w

Zhang, W., Wang, Y., Jin, F.-F., Stuecker, M. F., & Turner, A. G. (2015). Impact of different El Niño types on the El Niño/IOD relationship. *Geophysical Research Letters*, *42*(20), 8570–8576. https://doi.org/10.1002/2015GL065703

Zhou, D., Gozolchiani, A., Ashkenazy, Y., & Havlin, S. (2015). Teleconnection paths via climate network direct link detection. *Physical Review Letters*, *115*(26), 268501. https://doi.org/10.1103/PhysRevLett.115.268501

Propagation Pathways of Indo-Pacific Rainfall Extremes are Modulated by Pacific Sea Surface Temperatures

5.

This chapter shows the following article:

F. M. Strnad, J. Schlör, R. Geen, N. Boers, and B. Goswami (2023). 'Propagation pathways of Indo-Pacific rainfall extremes are modulated by Pacific sea surface temperatures'. In: *Nature Communications* 14.5708. DOI: `10.1038/s41467-023-41400-9`.

nature communications

Article https://doi.org/10.1038/s41467-023-41400-9

Propagation pathways of Indo-Pacific rainfall extremes are modulated by Pacific sea surface temperatures

Received: 1 March 2023

Accepted: 1 September 2023

Felix M. Strnad[1], **Jakob Schlör**[1], **Ruth Geen**[2], **Niklas Boers**[3,4,5] & **Bedartha Goswami**[1]

Intraseasonal variation of rainfall extremes within boreal summer in the Indo-Pacific region is driven by the Boreal Summer Intraseasonal Oscillation (BSISO), a quasi-periodic north-eastward movement of convective precipitation from the Indian Ocean to the Western Pacific. Predicting the spatio-temporal location of the BSISO is essential for subseasonal prediction of rainfall extremes but still remains a major challenge due to insufficient understanding of its propagation pathway. Here, using unsupervised machine learning, we characterize how rainfall extremes travel within the region and reveal three distinct propagation modes: north-eastward, eastward-blocked, and quasi-stationary. We show that Pacific sea surface temperatures modulate BSISO propagation – with El Niño-like (La Niña-like) conditions favoring quasi-stationary (eastward-blocked) modes—by changing the background moist static energy via local overturning circulations. Finally, we demonstrate the potential for early warning of rainfall extremes in the region up to four weeks in advance.

The Indo-Asia Pacific region and its population of around 2.5 billion people receive most of its annual rainfall during the monsoon season from June through September (JJAS)[1]. A defining feature of the region is the intraseasonal variation of heavy precipitation and convergent wind circulation[2], which occurs periodically on time scales of around 40 days during boreal summer[3]. Precipitation peaks and troughs are known as "active" and "break" periods[4], the active phase being often marked by widespread extreme rainfall events (EREs)[5]. The intraseasonal timings of active and break periods can leave long-lasting impacts on crop yields and harvest, and suddenly occurring EREs often wreak havoc on rural and urban infrastructures[6].

The Boreal Summer Intraseasonal Oscillation (BSISO) exerts a substantial influence on precipitation dynamics over the oceans and land masses of the Indo-Pacific domain and constitutes a major source of rainfall variability and the occurrence of EREs on intraseasonal time scales[3]. Active phases of the BSISO are initiated in the Indian Ocean triggering a forced Kelvin wave response to the east of the convective anomaly. As the eastward propagating convective system reaches the Maritime Continent, the convection weakens, and moist Rossby waves are emanated, which then move north-westwards toward India. This results in a northwest-southeast tilted band of heavy rainfall that ranges from southern Pakistan in the northwest end to the Philippine Sea and Guam in the southeast. This rainfall band then slowly propagates northward and eastward, finally dissipating over the western Pacific around two weeks later.

Various theories have been put forth to elucidate the BSISO's propagation mechanism. The traditional view suggests that the initiation produces easterly wind flows associated with a slowly eastward propagating convective Kelvin wave[7,8]. The northward propagation is explained via the vertical shear mechanism, which posits that

---

[1]Machine Learning in Climate Science, University of Tübingen, Tübingen, Germany. [2]School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK. [3]School of Engineering & Design, Earth System Modelling, Technical University Munich, Munich, Germany. [4]Potsdam Institute for Climate Impact Research, Potsdam, Germany. [5]Department of Mathematics and Global Systems Institute, University of Exeter, Exeter, UK.
e-mail: felix.strnad@uni-tuebingen.de

the background seasonal mean vertical wind shear interacts with upward moving air parcels in the BSISO convective center and, due to the meridional gradient of their vertical velocities, generates cyclonic vorticity and boundary layer convergence to the north of the BSISO cloud band[7,9]. More recently, however, observed fluctuations in sea surface temperatures (SSTs) that appear coherently with BSISO convection[10] have received enhanced attention, emphasizing the role of air-sea interaction. The idea is that the energy source for the propagation is provided by wind-induced surface heat exchange[11] through the feedback processes of the uprising air associated with convection[12]. Based on this mechanism, another compelling explanation has been put forth to account for the northeastward propagation, known as the moisture mode theory[13–15]. It extends the eastward propagation concept of the Madden-Julian Oscillation (MJO)[16], which is mainly observed during boreal winter, to the boreal summer season. Following the initiation of anomalous convection over the tropical Indian Ocean, moistening to the east and drying to the west of the anomalous convection, a typical characteristic of both MJO and BSISO, drive its eastward propagation. It has been explained as a result of the advection of the MJO moisture anomalies by background zonal winds[14], the advection of background moisture by the anomalous flow[17], or the combination of both components[15]. Further studies emphasized the role of the meridional advection caused by the anomalous zonal advection[18] and the coupling between MJO convection and the mean monsoon flow[19] to explain the northward propagation. The theory is supported by empirical evidence derived from observational data[17,20] and has been recently bolstered by advancements in theoretical understanding which incorporate both MJO and BSISO characteristics into a unified framework[21].

Most observational studies concur with the theories and taken together, the combined movement of the eastward and northward propagation characterizes the “canonical” BSISO propagation[2]: A dominant low-frequency mode (30–60 days)[22] in the form of a deep convection zone carrying heavy rainfall emerges in the equatorial Indian Ocean and moves simultaneously eastward and northward, forming a northwest-southeast tilted convection band which, after transgressing the Maritime Continent barrier, progresses further to the Pacific Ocean[2,3,23,24] (exemplified in Figs. S11 and S12). However, not every anomalous convective activity in the Indian Ocean that is associated with the BSISO follows the canonical propagation pathway. Several studies have reported anomalous convective activity that fails to propagate north-eastward and remains stationary in the equatorial Indian Ocean[25,26]. One possible factor modulating the BSISO propagation could be the sea surface temperature (SST) variability associated with the El Niño Southern Oscillation (ENSO), as it is known to affect the rainfall dynamics during the South Asian summer monsoon season, mainly through inducing changes in the Walker circulation[27,28]. But empirical evidence to show clearly that the ENSO influences BSISO propagation is still lacking and the interactions of the Pacific SSTs with the north-eastward propagating convective BSISO system remain poorly understood to date.

Here, we investigate the spatial patterns associated with the propagation of Indo-Asia Pacific rainfall extremes and show that they are clearly linked to different phases of the BSISO. We further address the influence of the SST background state by analyzing the occurrence of synchronous EREs over large spatial areas in the Indo-Pacific domain. We use the fact that the BSISO is a large-scale convective system; thus, BSISO-driven EREs are likely to emerge as spatiotemporally organized weather systems connected via long-range teleconnections[29]. We develop a simple heuristic to identify regions of synchronous BSISO-driven EREs by using the framework of climate networks derived from observational rainfall event data[30]. Our method identifies statistically robust geographical regions that tend to have similar active and break phase timings. BSISO propagation can thus be investigated as the progression of EREs from one region to the next. Based on these propagation pathways, we cluster them by using an unsupervised spatial clustering method and discover three distinct propagation modes of the BSISO. We further find that the background state in the tropical Pacific does affect BSISO propagation but not its initiation in the equatorial Indian Ocean.

While BSISO’s impact on annual monsoonal rainfall has been analyzed thoroughly[3,31–33], the propagation pathways of rainfall extremes linked to the BSISO and the potential influence of the SST background state have received less attention. Previous studies have shown that propagation patterns of convective anomalies during the May-June period exhibit distinct variations compared to those observed from August to October[12] and that ENSO can affect BSISO intensity[34] and propagation over the Maritime Continent. In particular, it was found that the premoistening in the Western Pacific, primarily modulated by ENSO, is influencing the eastward propagation[35]. El Niño-like (La Niña-like) conditions suppress (enhance) the propagation over the Maritime Continent[36]. BSISO activity in the East Asian-western North Pacific region was shown to be influenced by ENSO[37] and the variability in the northward propagation of the BSISO has been related to different cloud hydrometeors[38]. Also, the east-, north- and north-eastward propagation of BSISO-related convection has been investigated, based on predefined propagation directions[18] or on the basis of convective anomalies in the equatorial Indian Ocean[39]. However, these studies do not report any influence of the background SST state on the propagation, and the causes for varying propagation pathways remain unresolved. A mechanistic understanding of the propagation diversity is still lacking, limiting the forecast skill of the BSISO[40] and the ability of numerical models to describe correctly the north-eastward propagation over the Indo-Pacific domain[3,41]. Our work offers a new perspective on BSISO diversity with implications for improving climate model simulations and shows the potential to develop early-warning signals of EREs along the propagation pathway on subseasonal time scales in a prediction period of more than four weeks in advance.

Results

Fingerprint of the BSISO on the spatial organization of EREs

In order to explore the BSISO propagation pathway in boreal summer from June through September (JJAS), we first detect its signature in regions with similar active and break phase timings. We thus identify geographical regions where EREs (defined locally as days with rainfall sums above the 90th percentile of wet days) occur synchronously (within up to 10 days) on average over the boreal summer JJAS data period, and whose average ERE timings are distinct from the rest of the study area. The regions identified correspond to “communities” of a climate network constructed by estimating event synchronization[29,42–44] from extreme rainfall event data of the Indo-Asia Pacific domain (illustrated schematically in Fig. 1 and explained in detail in “Methods”). As our community detection model is inherently probabilistic, we repeat the community detection step multiple times and use the distribution of different community detection outputs (Fig. 1c, d) to quantify the membership likelihood of spatial locations of belonging to a particular community (Fig. 1e). The low variances in the shape of the communities (Fig. S4) confirm that the communities are stable manifestations of spatial patterns associated with synchronous EREs.

The community detection reveals six geographical regions, labeled here as the equatorial Indian Ocean (EIO), Bay of Bengal (BoB), Maritime Continent (MC), South Asia (SA), West Pacific (WP), and North India-China (NIC) (Fig. 2a). EIO consists solely of the equatorial and northern Indian Ocean, whereas the BoB region connects India with the Maritime Continent via the Bay of Bengal. The V-shaped form of the BoB region has been reported in modeling studies of the BSISO[7]. SA is a northwest-southeast tilted region connecting the South East Asian Monsoon domain with central India also reported in previous BSISO studies[8,24,39]. WP is located in the North-Western Pacific north of the Maritime Continent and reveals a long south-north shape along the

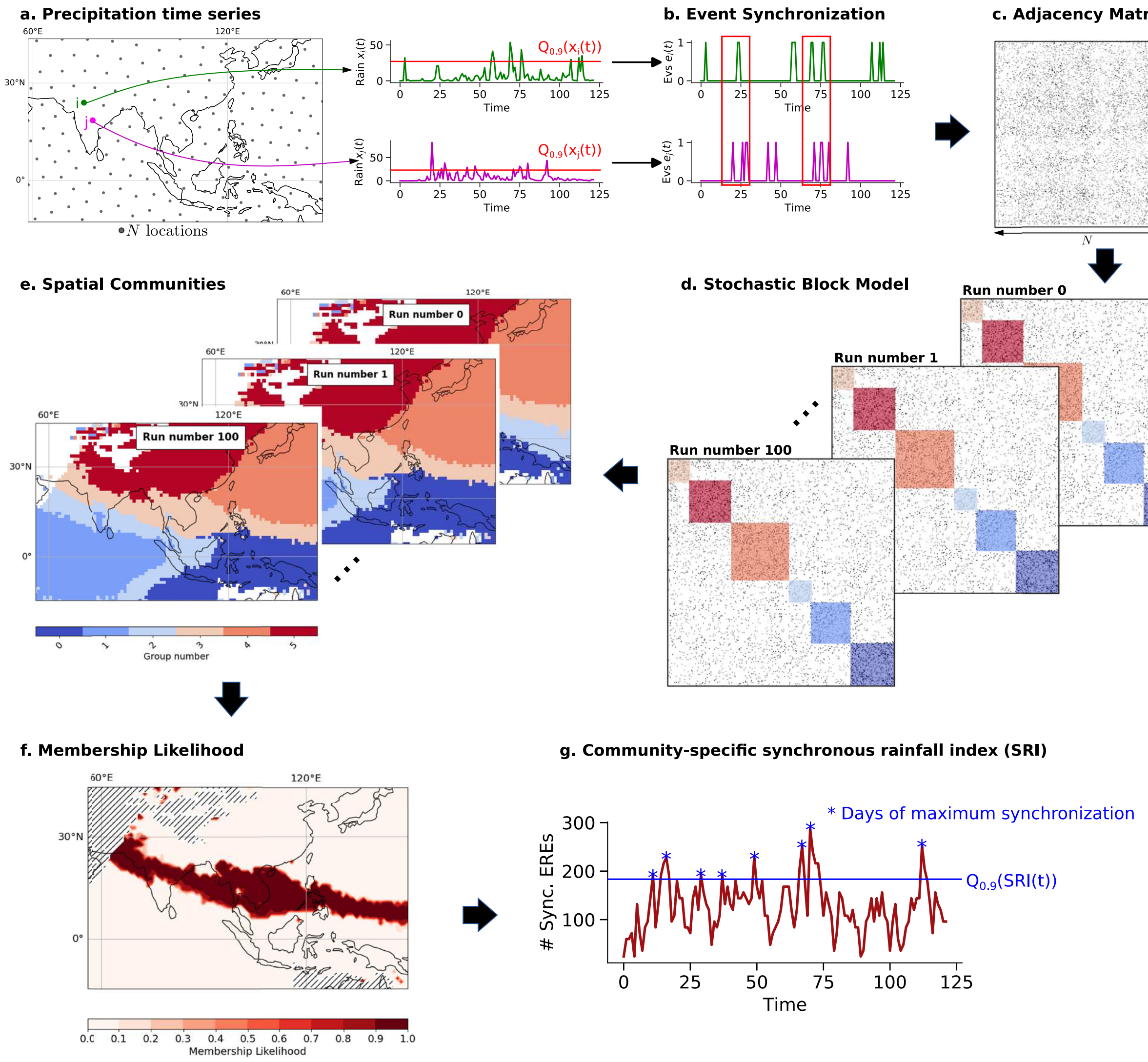


Fig. 1 | Identification of communities of synchronous Extreme Rainfall Events. **a** The dataset comprises a collection of *N* spatially distributed (depicted as gray circles) rainfall time series (only every 5th location shown). **b** The precipitation time series are transformed into extreme event series, estimated locally by days that exceed the 90th percentile of all wet days. The network is constructed by point-to-point comparison using the event synchronization technique[29,42–44], which quantifies the degree of synchronization between all pairs of single event series (e_i, e_j) (see "Methods", Eq. (2)). **c** The network is represented by its $N\times N$ adjacency matrix **A**, where $A_{ij}=1$ if the synchronization between locations *i* and *j* is statistically significant (denoted by black dots in **A**). **d** Communities are identified as blocks in **A** by reordering rows and columns using a probabilistic community detection algorithm based on the Stochastic Block Model (SBM, see "Methods"). **e** The network nodes within these blocks correspond to the spatial locations. **f** The membership likelihood describes the probability of a point belonging to a respective community estimated by the overlap of the 100 independent runs. One exemplary community is displayed (other communities: Fig. S4). Hatched regions indicate areas without a sufficiently large number of extreme rainfall events (EREs), and, thus, are excluded from the analysis. **g** The community-specific synchronous rainfall index ($SRI(t)$) counts the number of synchronously occurring EREs per day within a community (see "Methods", Eq. (4)). The time series shown is for illustrative purposes only. The blue stars indicate the local maxima above the 90th percentile of the $SRI(t)$ index, referred to as days of maximum synchronization.

coastline of East Asia including the island of Japan. NIC is solely over land, including the Himalayan Mountains, the Tibetan Plateau, the Ganges Delta, and the Chinese mainland. The regions of synchronization show long spatial extension, e.g., the SA spans over approximately 9000 km from South Pakistan in the west to the West equatorial Pacific in the east. The community structure indicates stable synchronization patterns of EREs in both west-east as well as south-north directions.

Propagation of EREs. During boreal summer, EREs propagate along the sequence of the regions EIO → BoB → MC → SA → WP, i.e. from southwest to north-east, in approximately 25 days (Fig. 2b). Such a propagation pathway has also been reported in previous studies[12]. We find that EREs which occur in EIO are particularly likely to take place in BoB +4 days later (Fig. 2c). In the same way, we see that EREs in BoB are likely to arrive at MC at +4 days later (Fig. 2d), from MC to SA with around +6 days (Fig. 2e) and from SA to WP at approximately +6–11 days later (Fig. 2f). EREs within NIC do not show any significant lagged correlation to the other regions, most likely because it is primarily over land, unlike the other regions. The propagation sequence of EREs along the identified regions is uncovered by using "days of maximum synchronization" within EIO (using the community-specific ERE index, explained in "Methods"), denoted as day 0, and counting the number of EREs in all communities during these and the following 30 days. The maximum time delay between EREs of different communities is estimated by lead-lag correlation analysis. This propagation

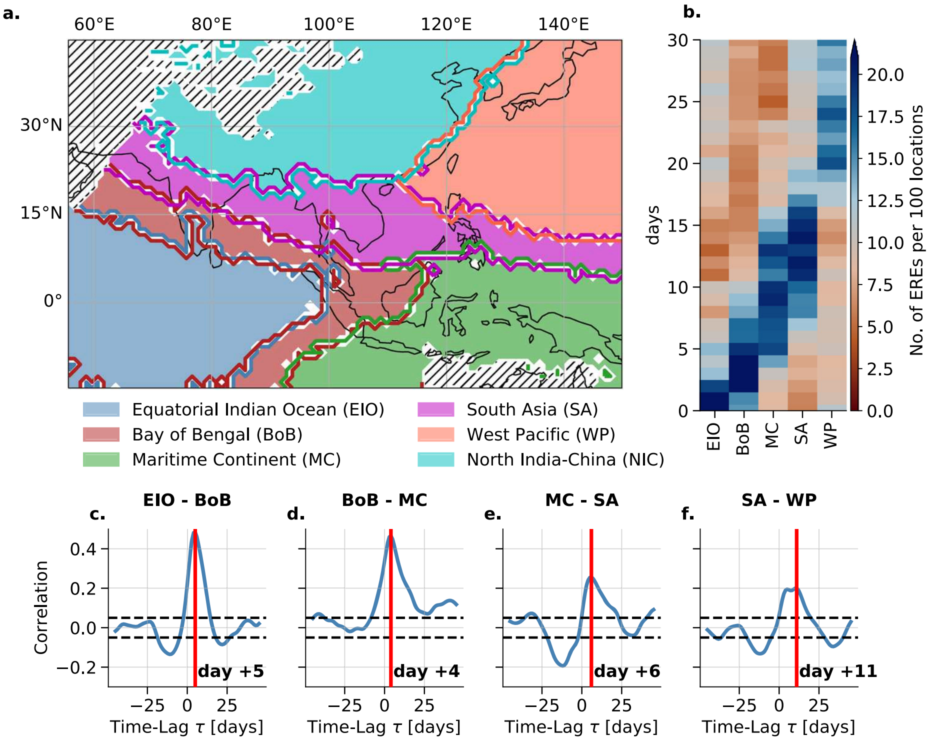


Fig. 2 | Communities of synchronous extreme rainfall events in the Indo-Pacific domain and propagation characteristics. The communities are determined using a probabilistic community detection algorithm and overlaps of 100 independent runs. **a** Six regions of synchronous extreme rainfall events (EREs). These are labeled according to their spatial mean position, i.e. equatorial Indian Ocean (EIO), Bay of Bengal (BoB), Maritime Continent (MC), South Asia (SA), Western Pacific (WP), and North India-China (NIC). Hatched areas indicate regions with too little precipitation, which are excluded from the analysis. **b** Temporal evolution of EREs using the most synchronous days within EIO, normalized by the number of grid points per community as day 0. **c**–**f** Lead-lag correlation analysis between pairs of the synchronous EREs belonging to the communities shown in (a). The time shifts of maximum correlation are denoted by vertical red lines and the respective days are displayed.

scheme can also be observed for specific events (see Supplementary Note 5, Fig. S11, S12).

BSISO modulates the ERE occurrences. We find that, except for the NIC, in all regions synchronous EREs are significantly more likely to occur in some particular phases of the BSISO (Fig. 3). We estimate the dependencies between phases of the BSISO (as defined in[23]) and the regions of synchronous EREs (Fig. 2a) using a conditional dependence test interpreted as the conditional probability of synchronous rainfall events subject to (i) phase (ii) active (inactive) BSISO days. As we define days of high synchronization within a community as the top 10% of the community-specific synchronous ERE index, by definition at most 10% of days in JJAS are days of maximum synchronization (dashed lines in Fig. 3). The corresponding null model for a day being a day of high synchronization (P(EREs = 1)) is therefore by construction 10%. For the NIC region, the likelihood of ERES for certain BSISO phases is not considerably different from the null model (Fig. 3f). Therefore, hereafter, we ignore the NIC region and focus on the other five regions (see Supplementary Note 14 for a discussion on the NIC). While an active BSISO increases the likelihood in the communities substantially, inactive days are distributed close to the null model (Fig. 3a–e). The influence of the propagating BSISO on the BoB region (Fig. 3b)also helps to explain the occurrence of the asymmetric V-shaped form (Fig. 2a) resulting from the decreasing zonal wind speed north- and southward of the equator. We observe that the regions that are predominantly oceanic (Fig. 3a–e) show a substantially increased likelihood for certain BSISO phases. The effect of the Maritime Continent barrier is reflected in the comparably lower likelihood for the MC region (Fig. 3c) but it still shows an increased likelihood for BSISO phase 4 and 5. The dependency of the occurrence of EREs and the BSISO can also be shown by a correlation analysis of the community-specific synchronous ERE indices to the BSISO index (Fig. S18), as well as by a linear regression test (see Supplementary Note 7). We also observe the BSISO characteristic 30-60 day oscillation[2,3,22] in all the community-specific synchronous ERE indices (Fig. S7). The communities MC and WP resemble the 10–20 day oscillation (Fig. S7c, e) which has been reported in that region[45].

Modes of BSISO propagation determined by Pacific SST background state

The preceding subsection has demonstrated that the BSISO plays a crucial role in shaping the spatial distribution of extreme rainfalls and provides insights into potential propagation pathways (Fig. 2b). However, the signal from the BSISO in the Equatorial Indian Ocean towards the Western Pacific weakens over time (Fig. 2c–f). Therefore, we investigate potential drivers of this diversity in propagation. Since BSISO propagation is clearly accompanied by the ERE progression and the occurrence of regions of highly synchronous rainfalls (Fig. 2), we consider the days of maximum synchronization in the EIO region to be potential BSISO initiation time points. We create for each individual time point two Hovmöller diagrams of outgoing longwave radiation (OLR) anomalies in the zonal and meridional direction to capture both the east- and northward propagation characteristics of the BSISO. The composited Hovmöller diagrams for the initiation time points show discernible patterns of eastward propagation through the Maritime Continent and simultaneous northward propagation, but nevertheless, the clarity of the propagation pattern diminishes approximately 5–10 days post initiation, indicating a degree of variability within the

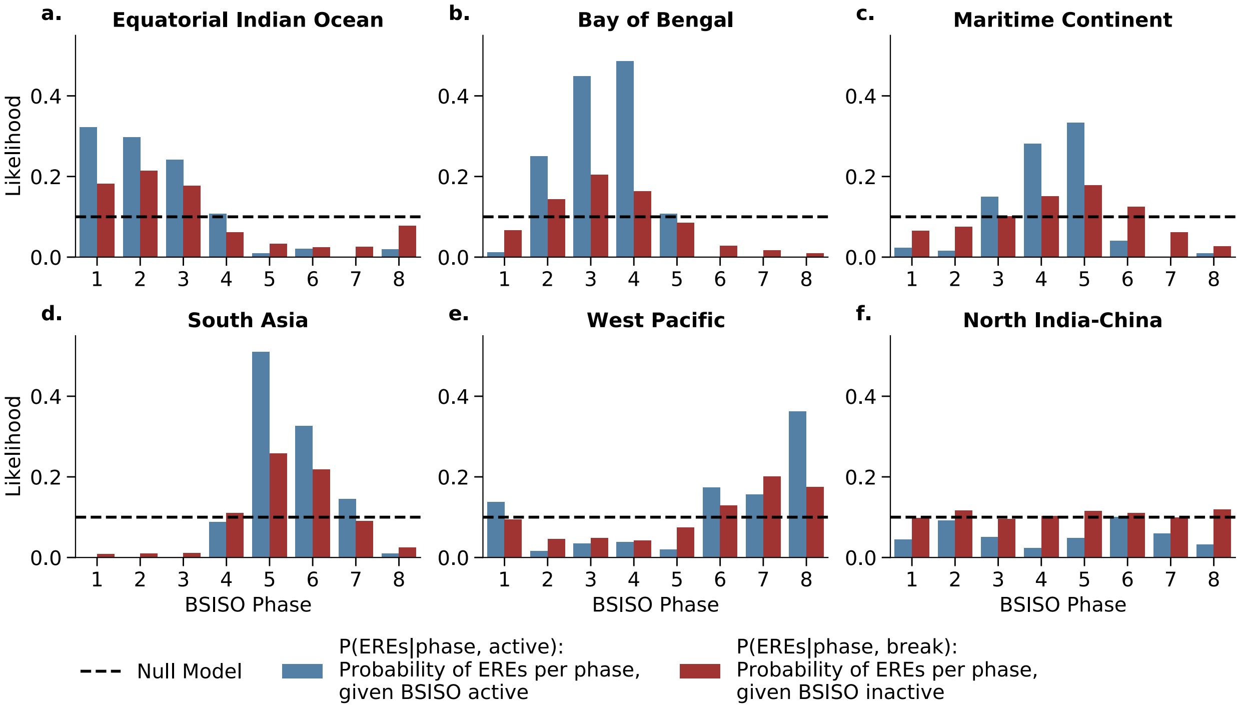


Fig. 3 | Likelihood of synchronous events for active/inactive Boreal Summer Intraseasonal Oscillation phases. The likelihood of the occurrence of synchronous events (*P*(*EREs*)) is analyzed for active (blue) and inactive (red) Boreal Summer Intraseasonal Oscillation (BSISO) phases (as defined here[23]) in the regions of Fig. 2a: **a** equatorial Indian Ocean, **b** Bay of Bengal, **c** Maritime Continent, **d** South Asia, **e** Western Pacific, and **f** North India-China. The dashed line illustrates the likelihood of synchronous events estimated from a null model for synchronous events being randomly selected (i.e. by construction 0.1 as the top 10% of all days are defined as days of maximum synchronization using the community-specific synchronous rainfall extremes index, see "Methods").

propagation pathways (Fig. S8). We thus use the single diagrams for each initiation point in time as input samples to a K-means clustering algorithm (see "Methods"). We obtain three clusters with different propagation features, even though all events initiated in EIO show similar enhanced convection at day 0:

- **Canonical propagation** This propagation cluster consists of 49 distinct Hovmöller diagrams. The convective system reveals a propagation speed in eastward direction of approximately 5.4 m/s (Fig. 4a), and in the northward direction of approx. 1.7 m/s (Fig. 4b). The phase speed in eastward direction is similar to the speed of the related Madden-Julian Oscillation (MJO) propagation[26]. The rainfall anomalies travel from the Indian Ocean towards the Western Pacific, passing the Maritime Continent barrier (Fig. 4a). The transition over the Maritime continent coincides with the temporary decrease of the convection anomalies at around day 10. The east-and northward propagation pattern resembles characteristics described in previous studies[46]. The anomalous wet phase is followed by an anomalously dry phase. The Canonical propagation mode occurs approximately twice as frequently as each non-canonical propagation mode[18]. We confirm that the bipolar pattern between enhanced convection in EIO and dry anomalies at the South Asian mainland and the Western Pacific (Fig. 4a) is characteristic in the eastward propagation[25].
- **Eastward Blocked propagation** This cluster of 28 samples shows a similarly fast northward propagation of 1.4 m/s and a suppressed eastward propagation with a speed of 4 m/s (Fig. 4c) that does not transgress the Maritime Continent barrier (dashed lines in Fig. 4c). Thus, its progression is "eastward blocked". Similar to the Canonical cluster, it progresses to latitudes north of 20° N (Fig. 4b), however, its wet phase is not directly followed by an anomalously dry phase.
- **Quasi-stationary propagation** This mode, consisting of 27 samples, shows a slow propagation that is, however, confined to the Indian Ocean from 50° E to 90° E (Fig. 4e), and a slightly poleward propagation (Fig. 4f) from day −5 to day 10. Further, it is even characterized by anomalously dry conditions in zonal and latitudinal directions besides the enhanced rainfalls in EIO (Fig. 4e, f). Note that keeping the classical definition of the BSISO in mind, these quasi-stationary events should not be called BSISO events as the BSISO is defined by the north-eastward propagation. However, there is literature discussing cases of non-propagating EREs that are still associative with the large-scale modes of variability[25].

Modulation by SST background state. We explore how the Pacific SST background is connected to the BSISO propagation. Figure. 5a, c, e shows the background SST anomalies that are associated with the three BSISO modes. The Canonical propagation mode corresponds to conditions without anomalous SSTs in the Pacific (Fig. 5a). A La Niña-like condition, expressed by anomalous cooling in the central Pacific (Fig. 5c), is likely to occur together with the Eastward Blocked Propagation. An El Niño-like condition represented by anomalously warm SSTs in the Pacific Ocean (Fig. 5e) likely occurs together with the Quasi-stationary Propagation. We confirm this connection to ENSO using a conditional dependence test (Fig. S32) showing that the likelihood of a specific BSISO propagation pathway is substantially increased by the respective ENSO state, i.e. Normal conditions favor the Canonical propagation, whereas La Niña- (El Niño-) like conditions favor the Eastward Blocked (Quasi-stationary) mode.

We observe that the column-integrated moist static energy (MSE) is in phase or even leading the spatial distribution of BSISO rainfall anomalies (see Fig. S37) which aligns with the moisture mode theory[17],

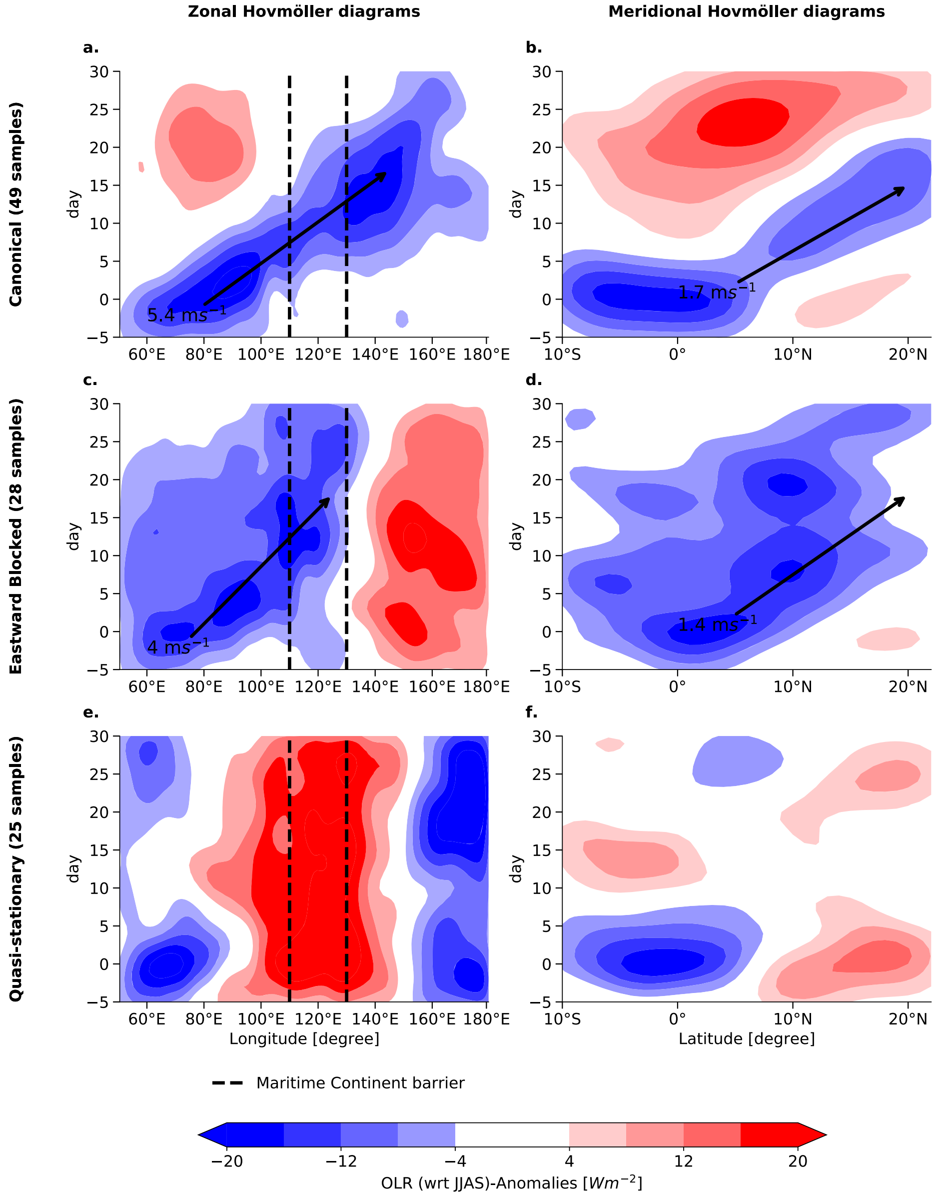


Fig. 4 | Clustering results of the the different Boreal Summer Intraseasonal Oscillation propagation pathways. Single Outgoing Longwave Radiation (OLR) Hovmöller diagrams for propagation pathways initiated in the Equatorial Indian Ocean (EIO) community are clustered. The first row shows the propagation for the Canonical mode, the second row the Eastward Blocked mode, and the third row the Quasi-stationary mode. Three different clusters are detected, labeled as Canonical, Eastward Blocked and Quasi-stationary. The first column (**a**, **c**, **e**) shows the composited Hovmöller diagrams in the zonal direction (averaged between [5°S, 5°N]), and the second column (**b**, **d**, **f**) in the meridional direction (averaged between [70°E, 80°E]). All anomalies are computed with respect to the JJAS seasonality. Day 0 describes the days of maximum synchronization within the EIO community (Fig. 2a). The dashed lines mark the area of the Maritime Continent barrier roughly estimated to be from 110° E to 130° E and the black arrows visualize the estimated propagation direction and velocity of the convective system for the propagating modes.

which asserts that meridional and zonal gradients in background MSE enable the propagation of a convective mode such as the BSISO (see Introduction). The Canonical mode (Fig. 5b) exhibits an eastward MSE gradient between the IO and the Western Pacific, crossing the Maritime Continent along with a poleward gradient between the east IO and South East Asia enabling the north-eastward propagation[21]. We observe changes in the background MSE gradient for the Eastward Blocked and the Quasi-stationary mode The Eastward Blocked mode displays a different pattern, characterized by the sign reversal of the zonal gradient around 140° E between the Maritime Continent and the Western Pacific (Fig. 5d), while the poleward gradient persists. Thus for this condition, only a northward propagation is encountered after the propagation reaches the Maritime Continent. The Quasi-stationary mode, however, exhibits neither a poleward gradient nor an eastward gradient (Fig. 5f) and thus remains trapped in the Indian Ocean. This observation is consistent with observations on the MJO propagation diversity as well[25].

Interaction of ENSO with BSISO via changes in the overturning circulation. Changes in the local zonal (i.e. the Walker) and meridional (i.e. the Hadley) overturning circulation (see "Methods") help to understand the changes in background moisture driving the interaction of the BSISO with ENSO. The Pacific Ocean does not show a substantial anomalous zonal wind flow for the Canonical mode (Fig. 6a).

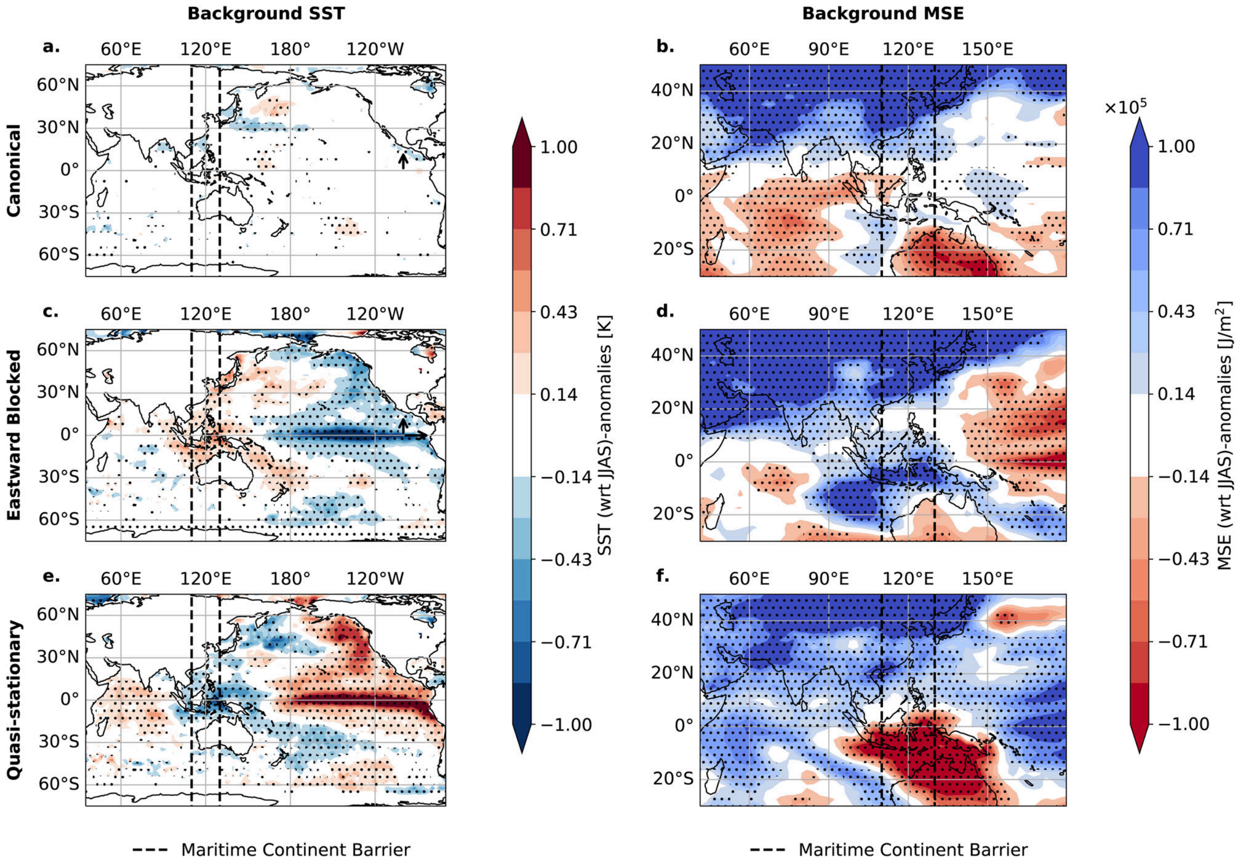


Fig. 5 | Sea surface temperature and moist static energy background state of the three different Boreal Summer Intraseasonal Oscillation propagation modes. Background states for sea surface temperature (SST) (**a**, **c**, **e**) and moist static energy (MSE) (**b**, **d**, **f**) for the Canonical (1st row), Eastward Blocked (2nd row), and Quasi-stationary propagation modes (3rd row). Composited anomalies of background column-integrated MSE and SST conditions are shown for days averaged 25–30 days before day 0 (as defined by the days of maximum synchronization), relative to the JJAS climatology. Stipples indicate significance at the 95% confidence level using Student's *t*-test. The dashed lines indicate the region of the Maritime Continent barrier (110° E to 130° E, compare Fig. 4c, e).

The enhanced local zonal overturning circulation in the equatorial Indian Ocean (Fig. 6b) is a consequence of the enhanced convection through the BSISO.

The Eastward Blocked mode reveals an enhanced zonal overturning circulation with the ascending (descending) branch over the Maritime Continent (Fig. 6e, f). The induced anomalous Walker cell over the Pacific Ocean with clockwise circulating air masses connects the circulation in the Indian Ocean with the circulation in the Pacific Ocean. The observed circulation pattern for the Eastward Blocked mode (Fig. 6e) deviates from the conventional La Niña conditions in JJAS (Fig. S34i), featuring a westward displacement of the Walker cell with the updraft of air masses dominated by the BSISO-associated convection (similar to the Canonical mode conditions).

We find two opposing zonal circulations for the Quasi-stationary mode (Fig. 6i). The vertical profile near the Maritime Continent exhibits a descending branch (Fig. 6j) that serves as a barrier, separating the anomalous clockwise zonal circulation in the Indian Ocean from the counterclockwise circulation which is also usually prevalent during El Niño events in the Pacific (Fig. S34j). These descending air masses, in opposition to the convective ascending air masses associated with the BSISO, conceivably suppress the BSISO-associated convection explaining the absence of a zonal MSE gradient (Fig. 5f).

We also observe a strongly enhanced Hadley circulation in the Central Indian Ocean for all three propagation modes (Fig. 6c, g, j) that are the result of the convective anomalies in the EIO. This anomalous overturning circulation pattern in the Indian Ocean is primarily driven by the arising convective anomalies of the BSISO[47]. However, the meridional circulation exhibits strong spatial variations between the different propagation modes (Fig. 6d, h, i), over the Maritime Continent and the Western Pacific. The Eastward Blocked propagation mode shows an elongated convergence corridor over the Western Pacific until the dateline (Fig. 6h) with an opposing circulation to the circulation in the Indian Ocean. The Quasi-stationary mode reveals a bipolar pattern in the region of the Maritime Continent and anomalously ascending air around the equator (Fig. 6l).

Propagation mechanism. Taking together the above results, we propose the following mechanism for the three BSISO propagation modes, schematically visualized in Fig. 7. Differences in the ENSO state in the tropical Pacific induce changes in the background MSE conditions over the Maritime Continent via modulation of the local Walker circulation. In all three modes intensified convection is observed in the equatorial Indian Ocean. For the Canonical mode (Fig. 7a) the MSE background condition has a zonal gradient over the Maritime Continent (Fig. 5a) which is induced by the Walker circulation. For the Eastward Blocked mode (Fig. 7b) the La Niña-like conditions trigger enhanced convection over the Maritime Continent (Fig. S35b) and anomalously wet conditions. The ascending air at the Maritime Continent provides an explanation for the eastward blocking (Fig. 4d) at around 120° N since the incoming winds from the Pacific Ocean oppose the eastward propagation of the BSISO convective system. We suggest that this potentially contributes to the observed two opposing MSE gradients preventing the propagation in zonal direction (Fig. 5d). The northward MSE gradient component remains

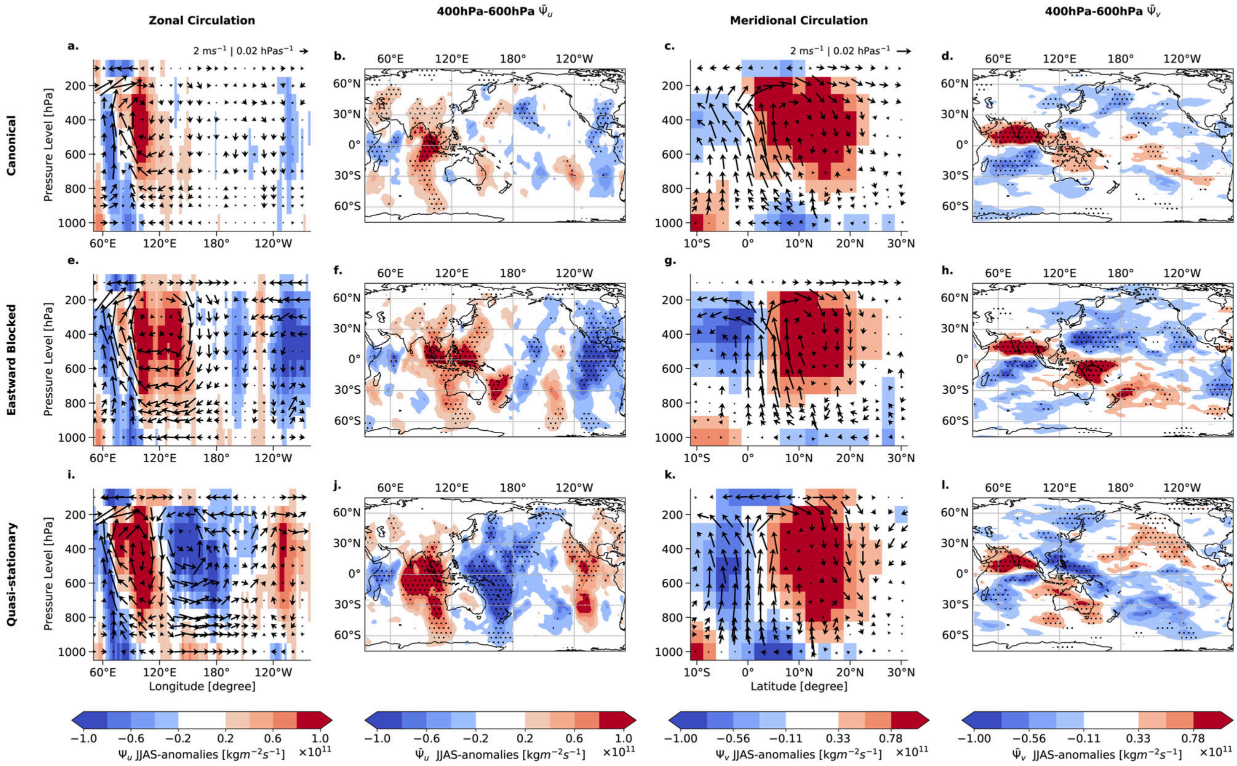


Fig. 6 | Mass stream function anomalies of overturning circulation for Boreal Summer Intraseasonal Oscillation propagation modes. For the days 0–5 after initiation in the equatorial Indian Ocean (EIO), the overturning circulation structure is analyzed. The first column (**a**, **e**, **i**) shows the composited zonal circulation, meridionally averaged from 0° S to 10° N, the second column (**b**, **f**, **j**) the zonally dependent circulation $\bar{\Psi}_u$ averaged between 400 hPa and 600 hPa. The color shading denotes the mass stream function in zonal direction. Red (blue) indicates irrotational lower-level easterlies (westerlies) and upper-level westerlies (easterlies). The third column (**c**, **g**, **k**) displays the zonally averaged (between 70°*E*-80°*E*) meridional circulation in the Central Indian Ocean. The fourth column (**d**, **h**, **l**) depicts the meridionally dependent circulation $\bar{\Psi}_v$ averaged between 400 hPa and 600 hPa. The color shading denotes the mass stream function in the meridional direction. Here, red (blue) indicates irrotational lower-level northerlies (southerlies) and upper-level southerlies (northerlies). The wind fields in the zonal (meridional) circulation plots are estimated using the meridionally (zonally) averaged u (v) components, measured in m/s, and the vertical velocity ω in the horizontal direction, measured in hPa/s. For visual clarity, only every 3rd wind arrow is plotted. Stipples denote anomalies that are significant at a 95% confidence level using Student's *t*-test.

unaffected[48]. Contrary, for the Quasi-stationary mode (Fig. 7c) the El Niño-like conditions induce a suppression of convection through the reduced low-level winds over the tropical Pacific leading to anomalously dry conditions over the Maritime Continent. The zonal MSE gradient is therefore already greatly reduced east of the Maritime Continent and the convective system remains over the initiation region in the Equatorial Indian Ocean[49,50].

Characteristics of the BSISO propagation diversity

Differences in the propagation modes are reflected in differences in terms of the Kelvin- and Rossby wave responses to the anomalous convection in the Equatorial Indian Ocean.

Eastward propagation. The Canonical and Eastward Blocked modes resemble a characteristic Kelvin wave signature[15,26,51]. The observed strong easterlies over the Bay of Bengal to the convection center in EIO, expressed by enhanced negative OLR anomalies (Fig. 8a), are previously reported as Kelvin wave responses[51,52] enabling the eastward propagation of the convective cell. Also, the characteristic moistening preceding the eastward-moving convection center is encountered (Fig. S37)[51]. The patterns of the wind fields for the Canonical mode and the Eastward Blocked mode (Fig. 8a, c) are very similar to each other, explaining the similar eastward propagation velocity. The Quasi-stationary propagation mode does not show the characteristic Kelvin wave signature (Fig. 8e) explaining the quasi-stationary propagation characteristics. We suggest that the descending dry air at the Maritime Continent leads to winds that are opposite to the convective uprising moist air (Fig. 8e). Hence, the Kelvin wave response to the anomalous convection in the EIO fails, so that the deep convection center remains stationary in EIO and vanishes after some days (Fig. 4e).

Poleward propagation. We observe the characteristic zonally Rossby wave pattern in the low-level winds[7,8,15,17,19] initiated at the Maritime Continent at ~120° E most clearly for the Canonical mode in Fig. 8b. The meridional wind anomalies reveal a westward drift in northeast-southwest tilted bands. The westward oriented waves occur 10–20 days after initiation consistent with the propagation of EREs (compare Fig. 2b). According to previous literature[7,15,19], the anomalous BSISO circulation near the equator exhibits modified Gill-type responses[53] with a stronger amplitude to the north than to the south of the equator. These are associated with the slanted northwest-southeast BSISO rainfall anomalies near the equator and suppressed convection north of 10^{circ}N that we also observe in Fig. 8a, c. Therefore, easterly anomalous winds due to the Rossby wave response (Fig. 8b, e) to BSISO-associated convection generate a meridional dipole-like MSE tendency pattern that drives the northward propagation over the Indian subcontinent at ~70°*E* (Fig. 8a) in a way as it was already shown in[7]. The Rossby wave pattern for the Eastward Blocked mode (Fig. 8d) is less clearly visible compared to the Canonical mode but starts

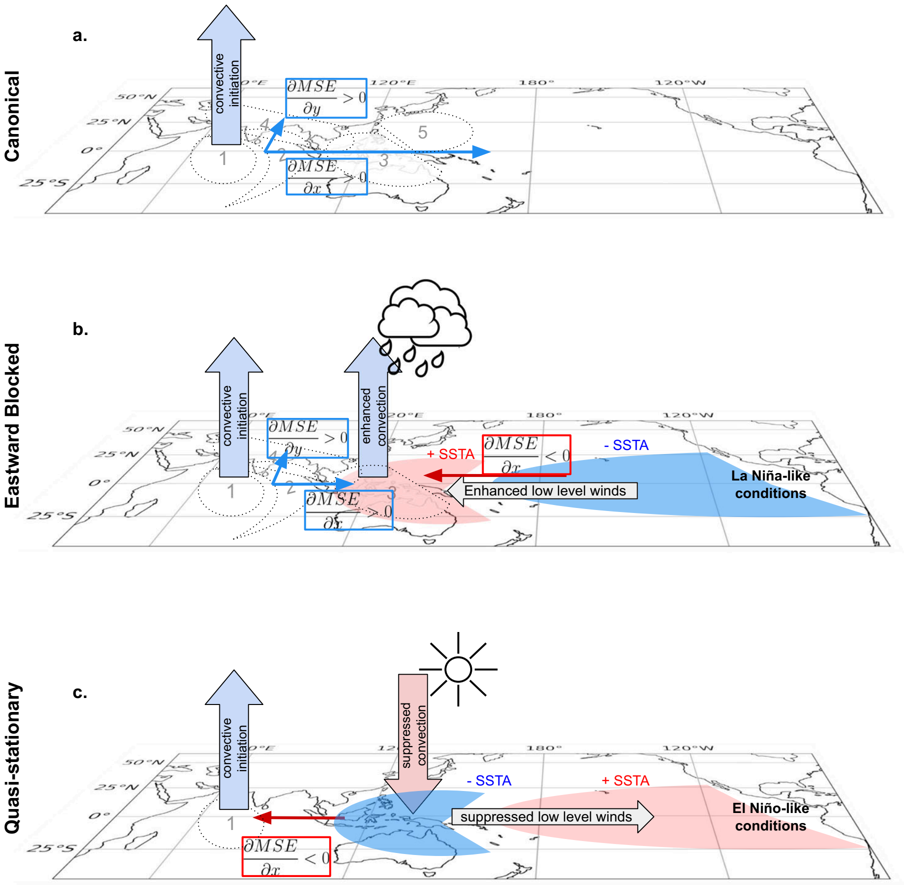


Fig. 7 | Proposed mechanism for El Niño Southern Oscillation modulation of Boreal Summer Intraseasonal Oscillation propagation modes. Our proposed mechanism for the three Boreal Summer Intraseasonal Oscillation (BSISO) propagation modes is represented schematically, showing the Canonical propagation (**a**), the Eastward Blocked mode (**b**) and the Quasi-stationary mode (**c**). The increasing (decreasing) gradient of the background moist static energy (MSE) state is visualized in blue (red) boxes in zonal and meridional direction. The respective ENSO condition and its influence on the MSE background state is visualized in the tropical Pacific. The dashed ovals indicate the communities identified (Fig. 2a) and the number denote their propagation in time. Small arrows denote the direction of the background MSE gradients. Big arrows in blue and red indicate the enhanced (suppressed) convection. The rainy clouds (sun) represent anomalously wet (dry) conditions over the Maritime Continent.

occurring at around 10–15 days after initiation in the EIO. The characteristic Rossby wave pattern does not exist in the Quasi-stationary mode (Fig. 8f) which can be explained by the absence of the eastward traveling Kelvin wave and a meridional background MSE gradient.

Potential for early-warning signals for EREs

The different BSISO propagation modes have a direct consequence on whether or not a given location in the Indo-Pacific region domain experiences an ERE on a given day once a convective anomaly has been initiated in the EIO. For normal conditions without substantial SST anomalies in the tropical Pacific, convective anomalies in the EIO are likely to follow the canonical north-eastward propagation (Fig. 9a and Fig. S13).

For La Niña-like conditions the propagation is trapped in the region of the Maritime continent, and the heavy rainfall remains in India and the South Asian subcontinent for a longer time span (Fig. 9b, and Fig. S14). We further observe that the convective anomalies are more likely to propagate towards higher latitudes in northern India upon the Himalayan foothills (Fig. 9b). If El Niño-like SST conditions coincide with a BSISO initiation in the EIO, our results show that India, the Maritime Continent, and the South Asian mainland are likely to only experience very few convective anomalies (Fig. 9c and Fig. S15). Most of the anomalous convection happens in the central and Western Pacific likely induced by the shifted Walker circulation.

These three propagation modes also translate to propagation of EREs via the identified communities (Fig. 2a), presented in Fig. S16. Using these different propagation modes, it is therefore justified to explore the possibility of early-warning signals (EWS) for EREs that are driven by the BSISO at a time horizon of multiple weeks. We sketch the potential for EWS for the Canonical mode: We estimate the days of maximum synchronization in all regions and calculate the fraction of events that are subsequent to days of maximum synchronization in EIO within a range of three days. In BoB 66% of the days of maximum synchronization occur 4–6 days after days of maximum synchronization in EIO. Subsequently, in the MC region 59% of the days of maximum synchronization are observed 6-9 days later than in the EIO. Similarly, 61% of the events in the SA community events (within 15-18 days and 39% of the events in WP (within 21–24 days) are subsequent to days of maximum synchronization in EIO. Even this simple approach shows the potential for large-scale spatially resolved ERE predictions up to 25 days in advance which is also the target range of the subseasonal to seasonal (S2S) prediction project[54].

Discussion

The aim of this article has been to reveal the specific BSISO propagation pathways and to improve the mechanistic understanding of the BSISO so that a future potential early-warning system for EREs during boreal summer may be established. We uncovered three dynamically

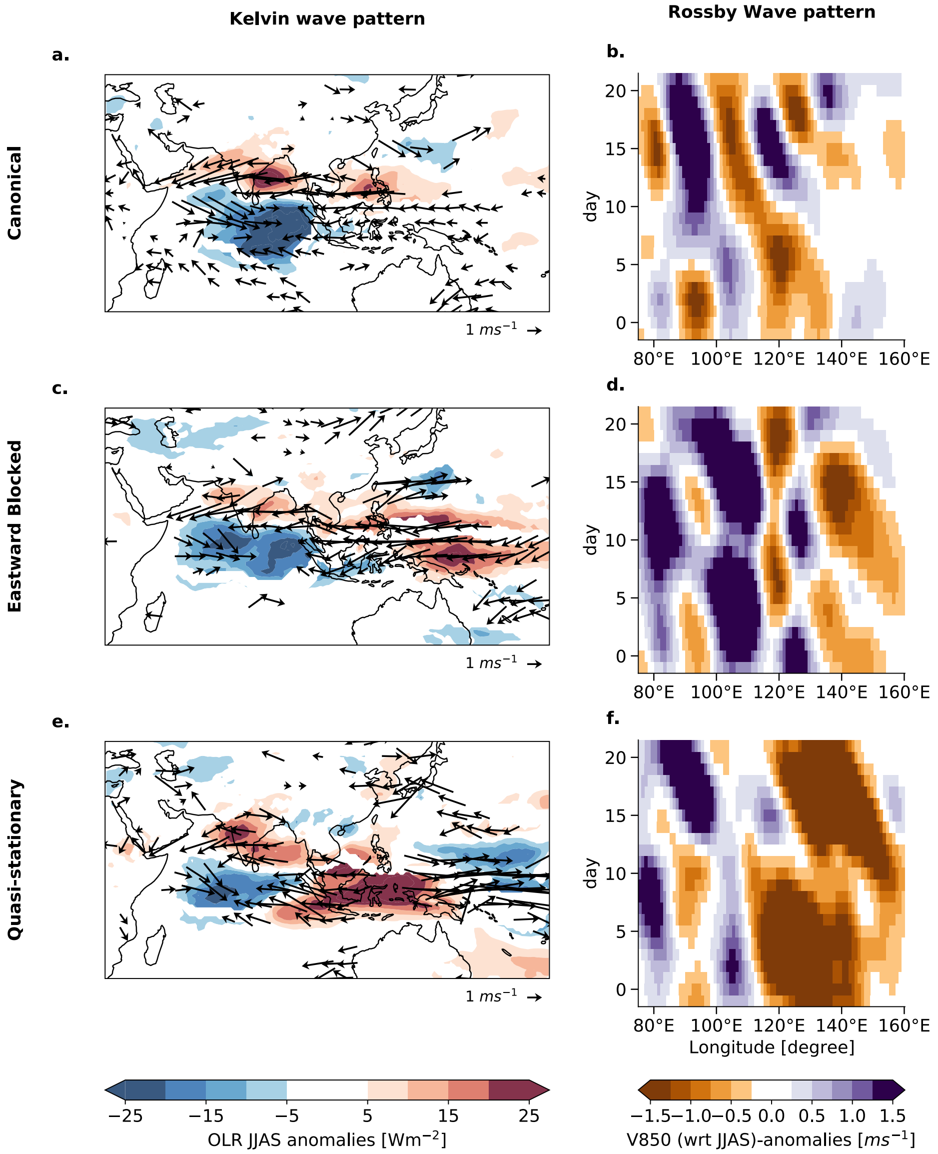


Fig. 8 | Kelvin and Rossby wave patterns after initiation in the equatorial Indian Ocean. Different atmospheric conditions as a response to initiation in the equatorial Indian Ocean (EIO) (day 0) are shown for the Canonical (1st row), Eastward Blocked (2nd row), and Quasi-stationary case (3rd row). The first column (**a**, **c**, **e**) shows the Kelvin wave pattern by contour lines of outgoing longwave radiation (OLR) and composited wind fields at 850 hPa plotted as arrows averaged for days 2–5 after initiation. Only the statistically significant contour lines and wind field anomaly arrows at 95% confidence level and for visual clarity only every third arrow are shown. The second column (**b**, **d**, **f**) shows the corresponding Rossby wave pattern by Hovmöller diagrams of low-level meridional V-winds at 850 hPa. The meridional average is applied for the range of [11°N, 16°N]. Day 0 denotes the initiation in the EIO.

resolved propagation modes of the BSISO that show different spatial manifestations and are strongly influenced by the SST background state of the tropical Pacific Ocean. We argued that understanding these three modes has important implications for the predictability of the occurrence of EREs in the Indo-Pacific region.

We demonstrated that the BSISO is a dominant driver of the spatiotemporal organization of EREs during the boreal summer in the Indo-Pacific region. To uncover the BSISO propagation, we introduced a new approach that combines climate networks based on a non-linear event synchronization measure with a probabilistic network community detection algorithm. Our approach identified macroscale structures of spatially coherent patterns of EREs, involving long-range teleconnections between regions from different parts of the Indo-Pacific domain. The results of our community detection approach are independent of the chosen dataset (see Supplementary Note 8.1) and remain robust also when using alternative community detection implementations, for instance using the Parallel Louvain Method (PLM) implemented in the NetworkIT packages[55] (see Supplementary Note 8.2). Using a posterior likelihood estimation conditioned on the BSISO index, our representation of spatial BSISO locations revealed a skewed distribution over multiple BSISO phases from the classical definition[23,24]. This confirms the relationship between active and break cycles of monsoon precipitation and the particular phases of the BSISO[3]. Our analysis also provided for the first time a detailed understanding of the spatiotemporal organization of the BSISO-driven rainfall extremes that emerge directly from the data. In this sense, our results present an alternative, impact-focused definition of the BSISO based on ERE data which remains still consistent with the classical definition in terms of atmospheric anomalies. The BSISO is characterized by two modes (10–20 days and 30–60 days) (e.g.[2]). Follow-

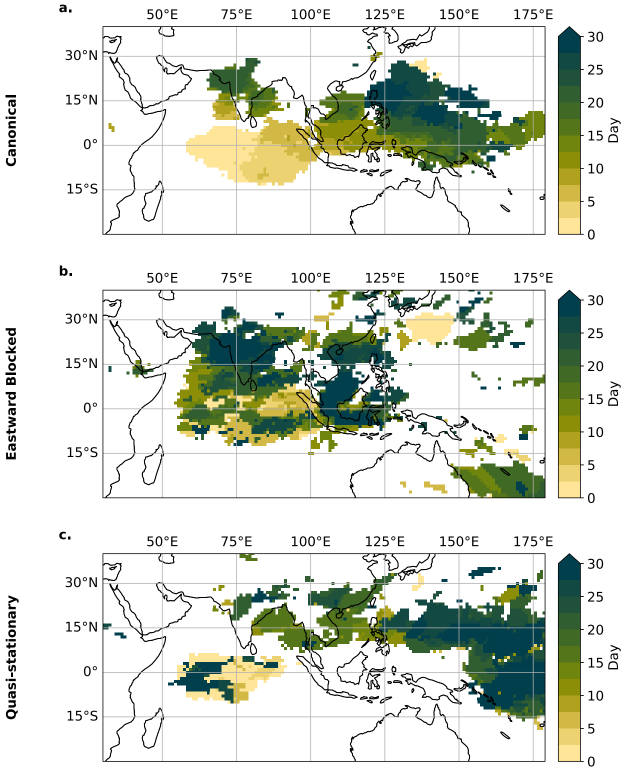


Fig. 9 | Evolution of anomalous rainfall for the three propagation modes. Regions with the highest rainfall intensity for days after maximum synchronization in the equatorial Indian Ocean (EIO) community are shown for the Canonical mode (**a**), the Eastward Blocked mode (**b**), and the Quasi-stationary mode (**c**). Outgoing longwave radiation (OLR) anomalies are computed according to the JJAS climatology. For every day, the statistically significant mean anomalous OLR above the 95th percentile is shown in its respective color of the day. Day 0 denotes the days of maximum synchronization in the EIO community.

up studies could investigate the influence of these modes on the occurrence and propagation of rainfall extremes.

We identified three distinct propagation modes of BSISO-driven EREs: the canonical northeastward propagating mode; the Eastward-blocked mode with continuous propagation only in the northward direction while eastward propagation ceases at the Maritime Continent; and the Quasi-stationary mode. Although the small observational sample size limits the level of statistical significance, our results are robust to randomly chosen initial configurations of the K-means algorithm and subsets of the input data. Note that a recent study reported three different modes of BSISO propagation as well[39], including the “canonical BSISO” propagation mode, and an “Eastward Expansion” mode, which shares some similarities to the Eastward Blocked mode identified in our work. However, the study did not report any modulation of the propagation modes by the background SST state. We contend that the discrepancy may be due to differences in the input to their clustering algorithm, which considered data from May–October and used local minima of an Indian Ocean box-averaged intraseasonal OLR time series to create pentad mean maps of OLR anomalies. Our study focused on the core monsoon season (JJAS) and hence looked at a situation in which the Walker circulation is more shifted to the Western Pacific[56]. In addition, we identified the initialization of BSISO events via regions of highly synchronous EREs ensuring spatially and temporally coherent patterns; and moreover accounted explicitly for zonal and meridional propagation characteristics by using separate Hovmöller diagrams in both these directions.

We reported a plausible mechanism that determines the different propagation modes. Our results provide a new perspective on how ENSO background state interacts with the BSISO on intraseasonal time scales. While the BSISO is known to be marginally correlated with ENSO[3], we showed that the ENSO state strongly influences the BSISO propagation, leading to local variability in rainfall[36,57]. We demonstrated how the coupling of ENSO to the BSISO-driven ERE propagation is mediated by the anomalously dry (moist) background moisture and the cooling (warming) near the Maritime Continent during El Niño (La Niña)-like conditions. We further showed how the interplay between the ENSO-related overturning circulation and the BSISO propagation inhibits the east- and northward propagation of EREs under El Niño-like conditions. This is in agreement with ISM failures observed during El Niño events[28,58]. Conversely, La Niña-like conditions in the tropical Pacific are found to favor the Eastward Blocked mode, such that EREs move only northward and bring extended periods of strong rainfall to the South Asian and Indian mainland as well as to the Maritime Continent. This result further ties in with the longer active spells observed in La Niña years over Indian and South Asian mainland[58,59]. For El Niño conditions the convection is suppressed, whereas for La Niña conditions the anomalous warming at the Maritime Continent leads to opposing winds that prevent the propagation beyond the Maritime Continent. Consequently, the Maritime Continent experiences a weakening (enhancement) of the rainfalls for El Niño (La Niña) conditions corroborating results of ref. 36. Our results thus imply a pronounced role of the Maritime Continent region for the BSISO propagation: The overturning circulation conditions at the Maritime continent are the key to deciding on the most probable propagation type. Our findings, therefore, offer a new perspective on its barrier effect upon existing theories based on the complex topography[60], the high land-sea thermal contrast[61], and the diurnal convection over land[62].

The proposed mechanism of BSISO diversity may provide a framework for understanding why models fail to simulate the BSISO propagation over the Indo-Pacific domain realistically[3,40,41], and thus could offer a new validation scheme for GCM development. In addition, we outlined the potential for medium-range forecasts of EREs and for developing a risk assessment for floods in the South Asian Monsoon domain and along the coast of South East Asia[63,64], on the scale of more than 4 weeks in advance. In comparison, current forecast models of the European Centre for Medium-Range Weather Forecasts (ECMWF) are predicting at a forecast lead time of around 14 days[65]. Although the discussed Early-Warning Signals approach is simplistic, the prediction skill could likely be substantially increased and a location-specific early-warning indicator could be developed by using tools from pattern recognition tasks in machine learning in a similar way as it has been proposed in a recent perspectives paper[66]. However, it should be noted that even though our study identifies the SST variability as a main driving force of the diverse propagation patterns, there are further factors determining the propagation patterns, for example, seasonal differences in the air-sea interaction[12], that need to be considered for an early-warning system.

Methods

Data

We use daily precipitation sums for the time period of 1979–2020 from the Multi-Source Weighted-Ensemble Precipitation (MSWEP) V2.2 dataset that merges gauge, satellite, and reanalysis data provided in a resolution of 0. 1° × 0.1°[30]. We used this dataset as it covers a longer time range than other available multi-satellite precipitation products to assure statistical robustness and it has been shown to represent high rainfall quantiles well on both global scales[67] and locally over India[68,69] and South East Asia[70,71]. We restrict our analysis to the tropical Indian Ocean and the South and East Asian Monsoon domain including the Western Pacific (55° E–140° E, 20° S–50° N). We use next-neighbor interpolation to map the data to a grid of spatially approximately uniformly distributed points employing the Fekete algorithm[72]. The distance between two points corresponds to the spatial distance between two points at the equator of a Gaussian 1° grid, resulting in a total of ≈4700 grid points. We linearly detrend the precipitation time series. The event time series is constructed from 'wet days' only, defined as days with rainfall of at least 1 mm/day. ERE days for a single location are defined as those days where the daily precipitation sum exceeds the 90th percentile of all wet days at that location.

To assess the robustness of our analysis to the choice of dataset, we conducted a comparative investigation using the Tropical Rainfall Measuring Mission (TRMM) dataset[73]. The TRMM dataset was applied to the identical spatial region, interpolated to the equivalent Fekete grid resolution employed for the MSWEP dataset. Our analysis revealed both qualitative and quantitative similarities in the patterns obtained from the TRMM dataset, corroborating our findings derived from the MSWEP dataset (see Figs. S20 and S21).

Further observational datasets used in this study are daily top net thermal radiation, translating to Outgoing Longwave Radiation (OLR), latent heat flux, sea surface temperature, and multi-pressure level variables on 50–1000 hPa of (u, v)-wind fields, vertical velocity w, and specific humidity q and temperature T taken from the ERA5 Global Reanalysis dataset[74]. The datasets are interpolated to 2.5° × 2.5° grid. The NINO3.4 index[75] was estimated using SST anomaly fields from the ERA5 dataset.

The column-integrated moist static energy (MSE) is a crucial factor in investigating the north-eastward propagation of the BSISO with respect to the moisture mode theory[3,14,21]. It is defined as the sum of sensible heat, latent heat, and potential energy as $\mathrm{MSE} = C_p T + L_v q + gz$, where T is temperature, z geopotential height, q specific humidity, C_p the specific heat of air at constant pressure, g the gravitational acceleration, and L_v latent heat of vaporization.

The daily resolved BSISO index by ref. 23 is taken from https://iprc.soest.hawaii.edu/users/kazuyosh/Bimodal_ISO.html(Last Accessed: 10th May 2023). The index is calculated by Principal Component Analysis of OLR for days from May until October. The first two leading principal components (PCs) are used to define the state of the BSISO by the amplitude $A = \sqrt{\mathrm{PC}_1^2 + \mathrm{PC}_2^2}$, where $A \geq 1$ $(A < 1)$ is regarded as active (inactive)[76]. The two-dimensional space spanned by PC_1 and PC_2 is subdivided into eight equally sized sections that denote the phase of the BSISO. The chosen BSISO index[23] was shown to capture both north- and eastward propagation in a coherent way and is well suited for tracking the BSISO-associated convection[77]. There are further BSISO indices available (e.g.[24,78,79]) but differences in our results remain minor (see Fig. S17 for a comparison).

Event synchronization-based climate networks

We apply a climate network approach, derived from complex network science. It offers a valuable extension to traditional methods by utilizing the toolset of complex network analysis[80] such as node degree analysis[81] betweenness centrality[82], and network curvature[83], to provide insights into climatic patterns and their interconnections. These network-based metrics provide a non-linear time series analysis of the underlying climate system, allowing conclusions to be drawn that may not be readily accessible through conventional approaches.

Assume a spatiotemporal dataset $X \in \mathbf{R}^{N \times T}$, where N denotes the number of datapoints and T is the number of points in time. The climate network $\mathcal{G}$ is defined as $\mathcal{G} = (V, E)$ where each geographical position of the dataset $x_i(t) \in X$ corresponds to a node $n \in V$ and E is the set of edges. Network edges $e_{ij} \in E$ encode strong statistical dependencies between pairs of time series $x_i(t)$ and $x_j(t)$.

To assess the degree of synchronization between pairs of time series, we use the Event Synchronization algorithm[42]. The number of temporally coinciding events is counted between pairs of event sequences $\{e_i^m\}_{m=1}^{s_i}$ and $\{e_j^n\}_{n=1}^{s_j}$, where s_i (s_j) are the total number of events at location i (j), and e_i^m (e_j^n) describes the timing of an event in i (j). The delay between an event e_i^m in i and an event e_j^n in j is denoted as $d_{ij}^{m,n} = e_i^m - e_j^n$. Defining the set $D_{ij}(e_i^m, e_j^n)$ as the set that contains all four neighboring events of e_i^m, e_j^n,

$$D_{i,j}^{m,n} = \left\{ d_{i,i}^{m,m-1}, d_{i,i}^{m,m+1}, d_{j,j}^{n,n-1}, d_{j,j}^{n,n+1}, 2\tau_{\max} \right\}, \tag{1}$$

the dynamical delay, $\tau_{i,j}^{m,n}$, is defined as half of the minimal waiting time of subsequent events in both time series around event e_i^m and e_j^n and not larger than a predefined maximal value $\tau_{\max}$ (Fig. 1b),

$$\tau_{ij}^{m,n} = \frac{1}{2} \min_{\forall d \in D_{ij}^{m,n}} d. \tag{2}$$

It encodes a small deviation between the occurrences, allowing for a time delay between two events. The parameter $\tau_{\max}$ separates time scales of ERE synchronization and is set to a maximum delay of $\tau_{\max} = 10$ days to ensure both the high- and low-frequency modes of the intraseasonal oscillations are captured. The event synchronization strength $R_{i,j}$ between locations i and j is the sum of all synchronous time points between all pairs of event sequences $\{e_i^m\}_{m=1}^{s_i}$ and $\{e_j^n\}_{n=1}^{s_j}$,

$$R_{i,j} = \sum_{m=1}^{s_i} \sum_{n=1}^{s_j} S_{i,j}^{m,n} \quad \text{where} \quad S_{i,j}^{m,n} = \begin{cases} 1 & 0 < d_{ij}^{m,n} < \tau_{i,j}^{m,n}, \\ 0 & \text{otherwise}. \end{cases} \tag{3}$$

Blocks of consecutive events are counted as one event, placed on the point in time of the first event to avoid the dynamical delay $\tau_{i,j}^{m,n}$ resulting in a value of 1/2, leading to a case where two sequentially occurring events would not be recognized as synchronous.

The adjacency $\mathbf{A}$ (Fig. 1c) of a network characterizes the interconnections and linkages between nodes, delineating the network's underlying topology. It is a mathematical representation of these connections and captures the presence or absence of links between nodes expressed as a $N \times N$ matrix, where $A_{i,j} = 1$ indicates that events at location i are statistically significantly followed by events at location j. We estimate the statistical significance using a null-model test. Our null hypothesis is that an observed $R_{i,j}$ value occurs from a pair of purely random event sequences with the same number of events s_i, s_j as in the observed sequences. To encode the null hypothesis, we construct surrogate event sequences e_i', e_j' with s_i, s_j randomly uniformly distributed events. Event series e_i is considered to be significantly synchronous to e_j if their corresponding $R_{i,j}$ value is higher than the 95 percentile of $R_{i',j'}$ values obtained using 2000 pairs of surrogate event sequences e_i', e_j'. Significant $R_{i,j}$ values imply that we place an edge from node n_i to n_j and set $A_{i,j} = 1$.

As the number of comparisons becomes very high even for moderately large datasets (in our case 10^8 comparisons), there is a non-negligible chance to consider singular pairs of time series as statistically significant, even though their significance is just by coincidence[84]. To avoid such spurious links, we assume that synchronous time series are supposed to be caused by physical mechanisms and thus show spatially coherent patterns[29]. For each spatial location, we rewire its network links randomly 2000 times. We use a Gaussian kernel density estimator (KDE) with the bandwidth selected according to Scott's Rule of Thumb to compute the spatial link distribution of every random sample. A link is only found significant if its regional link distribution (also obtained by a Gaussian KDE) is above the 99.9 percentile.

Estimating communities within climate networks

Determining macroscale regions of synchronously occurring EREs translates to identifying communities within the network, i.e. groups of nodes in the network that have a much higher number of edges within themselves than to nodes of the rest of the network. To identify communities, we thus need to reorder the rows of the adjacency matrix **A** such that a clear block structure is obtained (Fig. 1d) which is, in principle a *NP*-hard problem. The problem of identifying communities has been extensively studied in complex network science and many possible solutions have been proposed[85,86]. In our work, we use the Stochastic Block Model (SBM), which essentially formulates a 'block-structure' model of the adjacency matrix in which the edges attached to any given node are determined by two probabilities: a within-block edge probability, and an across-block probability. Thus, a SBM with k blocks would be described by $\frac{k(k+1)}{2}$ probabilities. Typically, the probabilities are then estimated using Bayesian inference techniques. In particular, we use the SBM implementation in the network package `graph_tool`[87,88], as it offers several advantages in terms of speed and efficient data handling. We refer the readers to[87] and[88] for a more detailed explanation of the SBM implementation, but would like to note here that the optimization is guided by the Minimum Description Length (MDL) principle, which favors simpler network structures with equal explanatory power, i.e. it uses the principle of Occam's Razor at its core. The SBM implementation in[87] can thus also fit a hierarchy of SBM's to the observed data and provide a most likely 'optimal' number of communities based on parsimony.

The SBM implementation uses Markov Chain Monte Carlo (MCMC) methods to estimate the posterior likelihoods of model parameters and hence, the inference algorithm is stochastic, meaning that it may produce (slightly) different results at each run[87]. We thus use multiple runs of the model to estimate the uncertainties associated with in assigning nodes to specific communities. In particular, we use a simple heuristic to estimate the posterior likelihood that a geographical location belongs to a particular climate network community: the percent of total runs that a node belongs to a given community is its 'community membership.'

Here, we run the SBM algorithm 100 times on the event synchronization-based climate network, with the constraint that it can have at most 10 communities (as we are interested in large spatial scales). Most of the runs identify 6 communities (with very few exceptions of 5 and 7 communities) with similar spatial shapes (Fig. 1d). Subsequently, by examining the overlap among all 100 SBM runs, we estimate the community membership (Fig. S4).

Community-specific synchronous ERE index

We introduce the community-specific synchronous ERE index, denoted as $SRI(t)$, to assess for a community its degree of synchronization over time. The index is computed for a particular set of locations A, describing one community within the network. For each time step t, we count the number of EREs that occur in all event series e_k:

$$\mathrm{SRI}_A(t) = \sum_{k\in A} e_k(t). \tag{4}$$

This counting process enables us to quantify the frequency of synchronously occurring EREs within this community A per day. To further pinpoint the points in time of exceptionally strong synchronization, we identify the local maxima in the time series SRI_A that are above the 90th percentile and define these as the "days of maximum synchronization" (blue stars in Fig. 1g).

Estimation of conditional probabilities

The probability for the occurrence of synchronous rainfall days within a cluster (denoted as $s=1$) under a condition a is calculated as follows: Assume the set of days with exceptionally synchronous events being S and the set of days that fulfill the condition a being A. Then $P(s=1|a) = \frac{P(s=1,a)}{P(a)} = \frac{||S\cap A||}{||A||}$ describes the conditional probability for synchronous events under condition a. Here, $||\cdot||$ denotes set cardinality and $S\cap A$ the intersection of S and A. Accordingly, the conditional probability for a second condition b with a set of days B is computed as:

$$P(s=1|a,b) = \frac{P(s=1,a)}{P(a,b)} = \frac{||S\cap A\cap B||}{||A\cap B||}. \tag{5}$$

A corresponding null model is estimated by counting the days of maximum synchronization $||S||$ divided by the total number of days (i.e. in our case ≤0.1). Hence, the upper limit for the null model is $P_{\mathrm{nullmodel}}(s=1) = 0.1$.

Clustering of propagation times

The BSISO propagation events we define as a day of maximum synchronization within the region EIO (Fig. 2a and see "Methods") and are denoted as day 0. Consecutive dates by less than 20 days are removed. In total 110 events are considered. We choose the propagation time range to be 5 days before and 30 days after the initiation day 0. The propagation of different synchronous extreme rainfall events is investigated by a K-means cluster analysis[89]. To account for the eastward as well as the northward propagation, propagation patterns are analyzed by Hovmöller diagrams of the OLR anomalies along a zonal band averaged between 0°*S* and 10°*N* and a meridional band averaged between 70° E and 80° E. OLR is like precipitation an indicator for deep convective activity, but, compared to precipitation, OLR exhibits a considerably smoother pattern, and it offers the advantage of direct measurement without requiring an inverse algorithm (which can introduce errors) as in the case of precipitation[90]. To ignore daily variations and to make the macroscale propagation patterns better distinguishable, we apply a 2D-smoothing Gaussian filter on the Hovmöller diagrams with 5 Pixels as the width of the filter. We use the silhouette coefficient method to determine the optimal number of groups and find that the samples can be best fitted into 3 distinct clusters. The silhouette coefficient indicates how similar a member is to its own cluster. We use it properly to remove outliers that have a silhouette coefficient lower than 0.05 from the cluster analysis. This further reduces the number of input samples by 13 events to 97 events in total.

Local overturning circulation analysis

In order to assess the relative contributions of the mass fluxes in the troposphere to the pair of meridional and zonal overturning circulations we use the method by[56,91–93]. A Helmholtz decomposition is applied to the global wind field $\mathbf{V} = (u, v)$, to separate the divergent component from the rotational component as $\mathbf{V} = \mathbf{V}_{\mathrm{div}} + \mathbf{V}_{\mathrm{rot}}$. The meridional (zonal) component of the divergent wind $\mathbf{V}_{\mathrm{div}} = (u_{\mathrm{div}}, v_{\mathrm{div}})$ is associated with the north-south (east-west) oriented circulations, commonly known as the Hadley (Walker) cell. The longitudinally dependent meridional circulation is calculated as the mass stream function Ψ_v as a vertical integration over the pressure levels:

$$\Psi_v(\lambda,\phi,p,t) = \frac{2\pi R}{g} cos(\phi) \int_0^p \mathrm{d}p' v_{\mathrm{div}}(\lambda,\phi,p',t), \tag{6}$$

where R denotes Earth's radius, g the gravitational constant, ϕ latitude, λ longitude, p pressure level and t time. The zonal mass stream function Ψ_u is analogously computed as:

$$\Psi_u(\lambda,\phi,p,t) = \frac{2\pi R}{g} \int_0^p \mathrm{d}p' u_{\mathrm{div}}(\lambda,\phi,p',t), \tag{7}$$

In our analysis, we use a simplified representation of Ψ_u and Ψ_v by averaging between 400 and 600 hPa.

Data availability

All data needed to evaluate the conclusions in the paper are present in the paper or the Supplementary Materials. Precipitation data were taken from the MSWEP dataset (https://www.gloh2o.org/mswep)[30]. Datasets for the composite analysis from 1979 till date were taken from Copernicus Climate Change Service (C3S) (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels?tab=overview)[94]. Plots were generated using the Cartopy library[95].

Code availability

The code for generating and analyzing the networks is made publicly available under ref. 96. The code for reproducing the analysis of the network communities and the spatial clustering described in this paper is publicly available under ref. 97.

References

1. Webster, P. *Dynamics of The Tropical Atmosphere and Oceans* (John Wiley & Sons, Ltd, 2020).
2. Wang, B. Intraseasonal Modulation of the Indian Summer Monsoon. In *Oxford Research Encyclopedia of Climate Science* (Oxford University Press, 2018).
3. Kikuchi, K. The Boreal Summer Intraseasonal Oscillation (BSISO): a review. *J. Meteorol. Soc. Jpn. Ser. II*. https://doi.org/10.2151/jmsj.2021-045 (2021).
4. Rajeevan, M., Gadgil, S. & Bhate, J. Active and break spells of the Indian summer monsoon. *J. Earth Syst. Sci.* **119**, 229–247 (2010).
5. Ding, Q. & Wang, B. Predicting extreme phases of the Indian summer monsoon. *J. Clim.* **22**, 346–363 (2009).
6. Gadgil, S. & Gadgil, S. The Indian monsoon, GDP and agriculture. *Econ. Polit. Wkly* **41**, 4887–4895 (2006).
7. Wang, B. & Xie, X. A Model for the Boreal Summer Intraseasonal Oscillation. *J. Atmos. Sci.* **54**, 72–86 (1997).
8. Wang, B., Webster, P. J. & Teng, H. Antecedents and self-induction of active-break south Asian monsoon unraveled by satellites. *Geophys. Res. Lett.* **32**. https://doi.org/10.1029/2004GL020996 (2005).
9. Hoskins, B. & Wang, B. Large-scale atmospheric dynamics. In *The Asian Monsoon*, 357–415, https://doi.org/10.1007/3-540-37722-0_9 (Springer, Berlin, Germany, 2006).
10. Sengupta, D., Goswami, B. N. & Senan, R. Coherent intraseasonal oscillations of ocean and atmosphere during the Asian Summer Monsoon. *Geophys. Res. Lett.* **28**, 4127–4130 (2001).
11. Sobel, A. H., Maloney, E. D., Bellon, G. & Frierson, D. M. The role of surface heat fluxes in tropical intraseasonal oscillations. *Nat. Geosci.* **1**, 653–657 (2008).
12. Kemball-Cook, S. & Wang, B. Equatorial waves and air–sea interaction in the boreal summer intraseasonal oscillation. *J. Clim.* **14**, 2923–2942 (2001).
13. Raymond, D. J. & Fuchs, Ž. Moisture modes and the Madden–Julian oscillation. *J. Clim.* **22**, 3031–3046 (2009).
14. Sobel, A. & Maloney, E. An idealized semi-empirical framework for modeling the Madden–Julian oscillation. *J. Atmos. Sci.* **69**, 1691–1705 (2012).
15. Adames, Á. F. & Kim, D. The MJO as a dispersive, convectively coupled moisture wave: theory and observations. *J. Atmos. Sci.* **73**, 913–941 (2016).
16. Madden, R. A. & Julian, P. R. Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.* **28**, 702–708 (1971).
17. Jiang, X., Adames, Á. F., Zhao, M., Waliser, D. & Maloney, E. A unified moisture mode framework for seasonality of the Madden–Julian Oscillation. *J. Clim.* **31**, 4215–4224 (2018).
18. Pillai, P. A. & Sahai, A. K. Moisture dynamics of the northward and eastward propagating boreal summer intraseasonal oscillations: possible role of tropical Indo-west Pacific SST and circulation. *Clim. Dyn.* **47**, 1335–1350 (2016).
19. Wang, T. & Li, T. Diagnosing the column-integrated moist static energy budget associated with the northward-propagating boreal summer intraseasonal oscillation. *Clim. Dyn.* **54**, 4711–4732 (2020).
20. Wang, S. et al. Propagating mechanisms of the 2016 summer BSISO event: air-sea coupling, vorticity, and moisture. *J. Geophys. Res. Atmos.* **126**, e2020JD033284 (2021).
21. Wang, S. & Sobel, A. H. A unified moisture mode theory for the Madden–Julian oscillation and the boreal summer intraseasonal oscillation. *J. Clim.* **35**, 1267–1291 (2022).
22. Lin, H. Monitoring and predicting the intraseasonal variability of the East Asian–Western North Pacific summer monsoon. *Mon. Weather Rev.* **141**, 1124–1138 (2013).
23. Kikuchi, K., Wang, B. & Kajikawa, Y. Bimodal representation of the tropical intraseasonal oscillation. *Clim. Dyn.* **38**, 1989–2000 (2012).
24. Lee, J.-Y. et al. Real-time multivariate indices for the boreal summer intraseasonal oscillation over the Asian summer monsoon region. *Clim. Dyn.* **40**, 493–509 (2013).
25. Kim, D., Kug, J.-S. & Sobel, A. H. Propagating versus nonpropagating Madden–Julian oscillation events. *J. Clim.* **27**, 111–125 (2014).
26. Wang, S., Sobel, A. H., Tippett, M. K. & Vitart, F. Prediction and predictability of tropical intraseasonal convection: seasonal dependence and the Maritime Continent prediction barrier. *Clim. Dyn.* **52**, 6015–6031 (2019).
27. Krishnamurthy, V. & Goswami, B. N. Indian monsoon–ENSO relationship on interdecadal timescale. *J. Clim.* **13**, 579–595 (2000).
28. Kumar, K. K., Rajagopalan, B., Hoerling, M., Bates, G. & Cane, M. Unraveling the mystery of Indian monsoon failure during El Niño. *Science* **314**, 115–119 (2006).
29. Boers, N. et al. Complex networks reveal global pattern of extreme-rainfall teleconnections. *Nature* **566**, 373–377 (2019).
30. Beck, H. E. et al. MSWEP V2 Global 3-Hourly 0.1° precipitation: methodology and quantitative assessment. *Bull. Am. Meteorol. Soc.* **100**, 473–500 (2019).
31. Di Capua, G. et al. Tropical and mid-latitude teleconnections interacting with the Indian summer monsoon rainfall: a theory-guided causal effect network approach. *Earth Syst. Dyn.* **11**, 17–34 (2020).
32. Karmakar, N., Boos, W. R. & Misra, V. Influence of intraseasonal variability on the development of monsoon depressions. *Geophys. Res. Lett.* **48**, e2020GL090425 (2021).
33. Hunt, K. M. R. & Turner, A. G. Nonlinear intensification of monsoon low pressure systems by the BSISO. *Weather Clim. Dyn. Discuss.* 1–28. https://doi.org/10.5194/wcd-2022-31 (2022).
34. Lin, A. & Li, T. Energy spectrum characteristics of boreal summer intraseasonal oscillations: climatology and variations during the ENSO developing and decaying phases. *J. Clim.* **21**, 6304–6320 (2008).
35. Wei, Y. et al. Western pacific premoistening for eastward-propagating BSISO and Its ENSO modulation. *J. Clim.* **35**, 4979–4996 (2022).
36. Liu, F., Li, T., Wang, H., Deng, L. & Zhang, Y. Modulation of boreal summer intraseasonal oscillations over the Western North Pacific by ENSO. *J. Clim.* **29**, 7189–7201 (2016).
37. Lin, H. Long-lead ENSO control of the boreal summer intraseasonal oscillation in the East Asian-western North Pacific region. *npj Clim. Atmos. Sci.* **2**, 1–6 (2019).
38. Abhik, S., Halder, M., Mukhopadhyay, P., Jiang, X. & Goswami, B. N. A possible new mechanism for northward propagation of boreal summer intraseasonal oscillations based on TRMM and MERRA reanalysis. *Clim. Dyn.* **40**, 1611–1624 (2013).
39. Chen, G. & Wang, B. Diversity of the boreal summer intraseasonal oscillation. *J. Geophys. Res. Atmos.* **126**, e2020JD034137 (2021).
40. He, Z., Hsu, P., Liu, X., Wu, T. & Gao, Y. Factors limiting the forecast skill of the boreal summer intraseasonal oscillation in a subseasonal-to-seasonal model. *Adv. Atmos. Sci.* **36**, 104–118 (2019).
41. Nakano, M. & Kikuchi, K. Seasonality of intraseasonal variability in global climate models. *Geophys. Res. Lett.* **46**, 4441–4449 (2019).

42. 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).
43. Malik, N., Marwan, N. & Kurths, J. Spatial structures and directionalities in Monsoonal precipitation over South Asia. *Nonlinear Processes Geophys.* **17**, 371–381 (2010).
44. Boers, N. et al. Prediction of extreme floods in the eastern Central Andes based on a complex networks approach. *Nat. Commun.* **5**, 1–7 (2014).
45. Gao, J., Lin, H., You, L. & Chen, S. Monitoring early-flood season intraseasonal oscillations and persistent heavy rainfall in South China. *Clim. Dyn.* **47**, 3845–3861 (2016).
46. Waliser, D. et al. Mjo simulation diagnostics. *J. Climate* **22**, 3006–3030 (2009).
47. Schwendike, J., Berry, G. J., Fodor, K. & Reeder, M. J. On the relationship between the Madden-Julian oscillation and the Hadley and Walker circulations. *J. Geophys. Res. Atmos.* **126**, e2019JD032117 (2021).
48. Zhu, W., Li, T., Fu, X. & Luo, J.-J. Influence of the maritime continent on the boreal summer intraseasonal oscillation. *J. Meteorol. Soc. Jpn. Ser. II* **88**, 395–407 (2010).
49. Ahn, M.-S., Kim, D., Ham, Y.-G. & Park, S. Role of maritime continent land convection on the mean STATe and MJO propagation. *J. Clim.* **33**, 1659–1675 (2020).
50. Liu, J., Da, Y., Li, T. & Hu, F. Impact of ENSO on MJO pattern evolution over the maritime continent. *J. Meteorol. Res.* **34**, 1151–1166 (2020).
51. Vallis, G. K. Distilling the mechanism for the Madden–Julian Oscillation into a simple translating structure. *Q. J. R. Meteorol. Soc.* **147**, 3032–3047 (2021).
52. Chen, G. & Wang, B. Dynamic moisture mode versus moisture mode in MJO dynamics: importance of the wave feedback and boundary layer convergence feedback. *Clim. Dyn.* **52**, 5127–5143 (2019).
53. Gill, A. E. Some simple solutions for heat-induced tropical circulation. *Q. J. R. Meteorol. Soc.* **106**, 447–462 (1980).
54. Vitart, F. et al. The subseasonal to seasonal (S2S) prediction project database. *Bull. Am. Meteorol. Soc.* **98**, 163–173 (2017).
55. Staudt, C. L., Sazonovs, A. & Meyerhenke, H. NetworKit: A tool suite for large-scale complex network analysis. *Netw. Sci.* **4**, 508–530 (2016).
56. Schwendike, J. et al. Local partitioning of the overturning circulation in the tropics and the connection to the Hadley and Walker circulations. *J. Geophys. Res. Atmos.* **119**, 1322–1339 (2014).
57. Li, J. & Mao, J. Factors controlling the interannual variation of 30–60-day boreal summer intraseasonal oscillation over the Asian summer monsoon region. *Clim. Dyn.* **52**, 1651–1672 (2019).
58. Joseph, S., Sahai, A. K., Chattopadhyay, R. & Goswami, B. N. Can El Niño–Southern oscillation (ENSO) events modulate intraseasonal oscillations of Indian summer monsoon?*J. Geophys. Res. Atmos.* **116**. https://doi.org/10.1029/2010JD015510 (2011).
59. Dwivedi, S., Goswami, B. N. & Kucharski, F. Unraveling the missing link of ENSO control over the Indian monsoon rainfall. *Geophys. Res. Lett.* **42**, 8201–8207 (2015).
60. Kim, D., Kim, H. & Lee, M.-I. Why does the MJO detour the Maritime Continent during austral summer? *Geophys. Res. Lett.* **44**, 2579–2587 (2017).
61. Zhang, C. & Ling, J. Barrier effect of the Indo-Pacific maritime continent on the MJO: perspectives from tracking MJO precipitation. *J. Clim.* **30**, 3439–3459 (2017).
62. Ling, J., Zhang, C., Joyce, R., Xie, P.-p & Chen, G. Possible role of the diurnal cycle in land convection in the barrier effect on the MJO by the maritime continent. *Geophys. Res. Lett.* **46**, 3001–3011 (2019).
63. Zhu, C., Nakazawa, T., Li, J. & Chen, L. The 30–60 day intraseasonal oscillation over the western North Pacific Ocean and its impacts on summer flooding in China during 1998. *Geophys. Res. Lett.* **30**. https://doi.org/10.1029/2003GL017817 (2003).
64. Li, J., Mao, J. & Wu, G. A case study of the impact of boreal summer intraseasonal oscillations on Yangtze rainfall. *Clim. Dyn.* **44**, 2683–2702 (2015).
65. Wu, J., Li, J., Zhu, Z. & Hsu, P.-C. Factors determining the subseasonal prediction skill of summer extreme rainfall over southern China. *Clim. Dyn.* 1–18. https://doi.org/10.1007/s00382-022-06326-w (2022).
66. Mariotti, A. et al. Windows of Opportunity for Skillful Forecasts Subseasonal to Seasonal and Beyond. *Bull. Am. Meteorol. Soc.* **101**, E608–E625 (2020).
67. Beck, H. E. et al. Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. *Hydrol. Earth Syst. Sci.* **23**, 207–224 (2019).
68. Ali, H. & Mishra, V. Increase in subdaily precipitation extremes in India under 1.5 and 2.0 °C warming worlds. *Geophys. Res. Lett.* **45**, 6972–6982 (2018).
69. Prakash, S. Performance assessment of CHIRPS, MSWEP, SM2RAIN-CCI, and TMPA precipitation products across India. *J. Hydrol.* **571**, 50–59 (2019).
70. Wu, X. & Zhao, N. Evaluation and comparison of six high-resolution daily precipitation products in Mainland China. *Remote Sens.* **15**, 223 (2022).
71. Du, Y., Wang, D., Zhu, J., Lin, Z. & Zhong, Y. Intercomparison of multiple high-resolution precipitation products over China: climatology and extremes. *Atmos. Res.* **278**, 106342 (2022).
72. Bendito, E., Carmona, A., Encinas, A. M. & Gesto, J. M. Estimation of Fekete points. *J. Comput. Phys.* **225**, 2354–2376 (2007).
73. Huffman, G. J. et al. The TRMM Multisatellite Precipitation Analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* **8**, 38–55 (2007).
74. Hersbach, H. et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
75. Trenberth, K. E. The definition of El Niño. *Bull. Am. Meteorol. Soc.* **78**, 2771–2778 (1997).
76. Wheeler, M. C. & Hendon, H. H. An all-season real-time multivariate MJO Index: development of an index for monitoring and prediction. *Mon. Weather Rev.* **132**, 1917–1932 (2004).
77. Wang, S., Ma, D., Sobel, A. H. & Tippett, M. K. Propagation characteristics of BSISO indices. *Geophys. Res. Lett.* **45**, 9934–9943 (2018).
78. Kiladis, G. N. et al. A comparison of OLR and circulation-based indices for tracking the MJO. *Mon. Weather Rev.* **142**, 1697–1715 (2014).
79. Wang, S. A precipitation-based index for tropical intraseasonal oscillations. *J. Clim.* **33**, 805–823 (2019).
80. Dijkstra, H. A., Hernández-García, E., Masoller, C. & Barreiro, M. *Networks in Climate* (Cambridge University Press, Cambridge, England, UK, 2019).
81. Donges, J. F., Zou, Y., Marwan, N. & Kurths, J. The backbone of the climate network. *EPL (Europhys. Lett.)* **87**, 48007 (2009).
82. Stolbova, V., Surovyatkina, E., Bookhagen, B. & Kurths, J. Tipping elements of the Indian monsoon: Prediction of onset and withdrawal. *Geophys. Res. Lett.* **43**, 3982–3990 (2016).
83. Strnad, F. M., Schlör, J., Fröhlich, C. & Goswami, B. Teleconnection patterns of different El Niño types revealed by climate network curvature. *Geophys. Res. Lett.* **49**, e2022GL098571 (2022).
84. Haas, M., Goswami, B. & von Luxburg, U. Pitfalls of climate network construction—a statistical perspective. *J. Clim.* **36**, 3321–3342 (2023).
85. Fortunato, S. & Hric, D. Community detection in networks: a user guide. *Phys. Rep.* **659**, 1–44 (2016).
86. Peixoto, T. P. Descriptive vs. Inferential Community Detection in Networks: Pitfalls, Myths and Half-Truths (Elements in the Structure and Dynamics of Complex Networks). (Cambridge: Cambridge University Press, 2023). https://doi.org/10.1017/9781009118897.

87. Peixoto, T. P. Hierarchical block structures and high-resolution model selection in large networks. *Phys. Rev. X* **4**, 1–18 (2014).
88. Peixoto, T. P. Bayesian stochastic blockmodeling. In *Advances in Network Clustering and Blockmodeling*, 289–332, (John Wiley & Sons, Ltd, Chichester, England, UK, 2019).
89. Lloyd, S. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **28**, 129–137 (1982).
90. Waliser, D. E., Graham, N. E. & Gautier, C. Comparison of the highly reflective cloud and outgoing longwave radiation datasets for use in estimating tropical deep convection. *J. Clim.* **6**, 331–353 (1993).
91. Hu, S., Cheng, J. & Chou, J. Novel three-pattern decomposition of global atmospheric circulation: generalization of traditional two-dimensional decomposition. *Clim. Dyn.* **49**, 3573–3586 (2017).
92. Raiter, D., Galanti, E. & Kaspi, Y. The tropical atmospheric conveyor belt: a coupled Eulerian-Lagrangian analysis of the large-scale tropical circulation. *Geophys. Res. Lett.* **47**, e2019GL086437 (2020).
93. Galanti, E., Raiter, D., Kaspi, Y. & Tziperman, E. Spatial patterns of the tropical meridional circulation: drivers and teleconnections. *J. Geophys. Res. Atmos.* **127**, e2021JD035531 (2022).
94. Hersbach, H. et al. Era5 hourly data on single levels from 1979 to present (accessed 02 March 2022) https://doi.org/10.24381/cds.adbb2d47 (2018).
95. Met Office. *Cartopy: a Cartographic Python Library with a Matplotlib interface* (Met Office, Exeter, Devon, 2010–2015).
96. Strnad, F. & Schlör, J. climnet v.2.2.0. https://doi.org/10.5281/zenodo.8233634 (2023).
97. Strnad, F. Netcommunities v.2.0.0. https://doi.org/10.5281/zenodo.8233579 (2023).

Acknowledgements

F.S., J.S., and B.G. acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC number 2064/1 - Project number 390727645. F.S. and J.S. thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting their PhD program. N.B. acknowledges funding by the Volkswagen foundation, the European Union's Horizon 2020 research and innovation program under grant agreement No. 820970 and under the Marie Sklodowska-Curie grant agreement No. 956170, as well as from the Federal Ministry of Education and Research under grant No. 01LS2001A. We acknowledge support from the Open Access Publication Fund of the University of Tübingen.

Author contributions

F.S. and B.G. conceived and designed the study. F.S. conducted the analysis. F.S. and B.G. prepared the manuscript. F.S., J.S., R.G., N.B. and B.G. discussed the results and edited the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-023-41400-9.

Correspondence and requests for materials should be addressed to Felix M. Strnad.

Peer review information *Nature Communications* thanks Kevin Cheung, Ding Ma and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Intraseasonal Synchronization of Extreme Rainfalls between North India and the Sahel

6.

This chapter shows the following manuscript (under review):

F. M. Strnad, K. M. R. Hunt, N. Boers, and B. Goswami (2024). ‘Intraseasonal synchronization of extreme rainfalls between North India and the Sahel’. In: *in review at Quarterly Journal of the Royal Meteorological Society*. DOI: `10.48550/arXiv.2405.08492`. eprint: `2405.08492`.

INTRASEASONAL SYNCHRONIZATION OF EXTREME RAINFALLS BETWEEN NORTH INDIA AND THE SAHEL

A PREPRINT

Felix M. Strnad
Machine Learning in Climate Science, University of Tübingen, Germany

Kieran M. R. Hunt
Department of Meteorology, University of Reading, UK
National Centre for Atmospheric Science, University of Reading, UK

Niklas Boers
Earth System Modelling, School of Engineering & Design, Technical University of Munich, Germany
Potsdam Institute for Climate Impact Research, Potsdam, Germany
Department of Mathematics and Global Systems Institute, University of Exeter, Exeter, UK

Bedartha Goswami
Machine Learning in Climate Science, University of Tübingen, Germany

September 16, 2024

ABSTRACT

The Indian Summer Monsoon (ISM) and the West African Monsoon (WAM) are dominant drivers of boreal summer precipitation variability in tropical and subtropical regions. Although the regional precipitation dynamics in these two regions have been extensively studied, the intraseasonal interactions between the ISM and WAM remain poorly understood. Here, we employ a climate network approach based on extreme rainfall events to uncover synchronously occurring extreme rainfall patterns across the two monsoon systems. We reveal strong synchronization of extreme rainfall events during the peak monsoon period in July and August, linking heavy rainfall over North India to that over the Sahel with a lag of around 12 days. We find that La Niña-like conditions in combination with the Boreal Summer Intraseasonal Oscillation and an enhanced Tropical Easterly Jet (TEJ) foster the synchronization between the ISM and the WAM. Convective clouds are transported by an intensified TEJ from southwestern Asia toward North Africa, supporting anomalous deep convection over the Sahel region.

Key points:

- Climate networks reveal synchronous occurrence of extreme rainfall events in North India and the Sahel.
- Propagation pathways leading to the synchronization are identified using multivariate latent space clustering.
- The synchronization is driven by a combination of active BSISO, La Niña-like conditions, and a strong Tropical Easterly Jet.

Keywords Communities of rainfall extremes · Long-range rainfall teleconnection · intraseasonal rainfall dynamics

1 Introduction

The Indian Summer Monsoon (ISM) and the West African Monsoon (WAM) are primary drivers of boreal summer precipitation variability in tropical and subtropical regions. The ISM is characterized by a strong meridional overturning circulation, with a low-level monsoon flow from the Indian Ocean to the Indian subcontinent and a return flow at upper levels as the northward migration of the Intertropical Convergence Zone (ITCZ) (Bordoni and Schneider, 2008). Similarly, the WAM is characterized by a low-level inflow of moist air from the Atlantic Ocean, which converges with the dry, hot air from the Sahara desert. Both the North Indian and Sahel regions exhibit substantial variations in monsoon rainfall known as "active" and "break" phases. The active phase is often associated with extreme rainfall events, which can have severe socioeconomic impacts (Kotz et al., 2022).

The intraseasonal variability comes about through a complex interplay of tropical and subtropical large-scale modes of variability (Priya et al., 2015; Sooraj et al., 2020; Nikumbh et al., 2023). For example, the rainfall dynamics in both North India and the Sahel are known to be correlated to the intensity and location of the Tropical Easterly Jet (TEJ) (Huang et al., 2019). The TEJ is a strong zonal wind maximum in the upper troposphere that is located over the tropical Indian Ocean and is mainly present during boreal summer months (usually June–September). It is established due to the meridional temperature difference between the Equatorial Indian Ocean and the Tibetan Plateau and extends from the Pacific Ocean to the West Coast of the Sahel Zone (SZ) (Koteswaram, 1958). Previous studies have found a strong connection between the strength of the TEJ and ISM rainfall (Pattanaik and Satyan, 2000; Madhu, 2014; Huang et al., 2019, 2021). Through latent heat release, the TEJ becomes stronger (weaker) during wet (dry) years of the ISM (Sathiyamoorthy et al., 2007; Rao and Srinivasan, 2016). Conversely, a stronger upper-level jet increases the vertical wind shear, which in turn leads to more convective organization and also enhances the northward migration of the ITCZ to the Indian subcontinent (Jiang et al., 2004; Bickle et al., 2021). This consequently increases the rainfall in India. The physical mechanism underlying the link between the TEJ and Sahel rainfall is less well understood, even though the strong statistical correlation between the TEJ and Sahel rainfall is well documented and consistent over several decades (Grist and Nicholson, 2001; Sathiyamoorthy, 2005; Nicholson, 2008). It is assumed that the TEJ represents an upper boundary for a region that is bounded downwards by the African Easterly Jet (AEJ). Between these two bounds African Easterly Wave (AEW) activity are causing around $90\ \%$ of the annual rainfall in the Sahel region. These mesoscale convective systems are embedded in the tracks of these waves (Nicholson, 2013) and potentially play a role in the occurrence of EREs in the western Sahel (Nicholson et al., 2008). Several studies have shown that wet years in the Sahel are often characterized by a regionally stronger TEJ even on interannual timescales (Lemburg et al., 2019). A strong (weak) TEJ is conducive to wet (dry) conditions and recent work has also shown that there is only a weak relationship between TEJ variability and daily variations in convection over the Sahel (Nicholson and Klotter, 2021). The authors conclude that the intensity of the TEJ is not influenced by rainfall variations over North West and Central Africa. Bickle et al. (2021) show that increased vertical shear in the Sahel region also modulates the occurrence of rainfall. Further, Whittleston et al. (2017) showed that the coupling of the different zonal jets (TEJ, African Easterly Jet and African easterly wave activity) over West Africa and the Sahel together modulate the rainfall but also the majority of climate models fail to capture the observed relationship between the jets and Sahel rainfall (Whittleston et al., 2017).

While the TEJ is influenced by strong convection over India, the occurrence of extreme rainfall events in the Indo-Pacific region itself is substantially influenced by another large-scale mode of variability: the Boreal Summer Intraseasonal Oscillation (BSISO) (Kikuchi, 2021), sometimes also referred to as Monsoon Intraseasonal Oscillation (MISO) (Webster, 2020; Nikumbh et al., 2021). The BSISO is a mode of large-scale convective variability that originates in the Indian Ocean and propagates northeastward across the Maritime Continent and the Western Pacific (Kikuchi et al., 2012; Lee et al., 2013; Kiladis et al., 2014). This oscillation plays a crucial role in the generation of deep convection and interacts with the local circulation patterns over the Indian subcontinent, e.g. by modulating the frequency and behavior of low-pressure systems (Hunt and Turner, 2022) modulating the occurrences of EREs (Nikumbh et al., 2021). The ISM rainfall dynamic is further modulated, on interannual timescales, by the El Niño Southern Oscillation (ENSO). Indian rainfall is enhanced (reduced) during La Niña (El Niño) years, typically explained by the modulation of the Walker circulation (Kumar et al., 2006). The BSISO itself is modulated by ENSO with La Niña-like conditions leading to an intensified northward propagation and more EREs in North India (Strnad et al., 2023). Similarly to the Indian rainfall, a teleconnection between ENSO and Sahel rainfall has been noted although the connection is more controversial (Joly et al., 2007). The modulation of the TEJ by ENSO has been suggested to explain the teleconnection (Chen and van Loon, 1987). Due to constructive (destructive) inference from the Walker circulation in La Niña (El Niño) years, the TEJ intensifies (weakens), and expands both horizontally and vertically (contracts spatially) (Nithya et al., 2017) inducing more (less) rainfall over the Sahel. This mechanism is used for instance to explain why El Niño events reduce rainfall and foster the occurrence of droughts in Ethiopia (Gleixner et al., 2017) and is also present in some global circulation models (Bader and Latif, 2003; Vashisht et al., 2021).

Summarized, the two monsoon systems of the ISM and WAM are both influenced by the same global large-scale modes of variability like ENSO (Shaman and Tziperman, 2007). Further a direct modulation of the ISM by the tropical Atlantic via by atmospheric Rossby waves in the Tropics has been reported (Kucharski et al., 2009). However, the occurrence of rainfall in the two monsoon systems of the ISM and WAM is commonly still mainly investigated on interannual to decadal timescales in the framework of the ‘Global Monsoon’ (Wang and Ding, 2008; Geen et al., 2020). Possible interactions between both monsoon systems on intraseasonal timescales, however, have not been systematically investigated. In their study, Boers et al. (2019) reveal a synchronization pattern of extreme rainfall events (EREs) between North India and the Sahel on daily timescales. While some of the other teleconnections of the ISM shown in Boers et al. (2019) have been explained in recent years; for example, the dynamical mechanism linking the ISM to the circumglobal teleconnection (Beverley et al., 2019, 2021) or to the Yellow River Basin (Gupta et al., 2022), the reasons for the synchronization between the ISM and WAM remain unclear. Our study thus aims to identify the atmospheric processes that drive the synchronization between the ISM and the WAM on the timescale of multiple days.

Previous studies on rainfall variability during boreal summer monsoon have traditionally used methods such as linear regression, empirical orthogonal functions (EOFs), and composite analyses (Ding and Wang, 2005, 2007; Wang and Ding, 2008; Vellore et al., 2014; Walker et al., 2015). These are not suitable to capture the spatial characteristics of EREs. Additionally, the identification of specific and possibly varying time lags associated with particular interaction patterns of different atmospheric processes is challenging. Therefore, in this study, we examine the spatial patterns associated with EREs that occur concurrently in the tropics and subtropics by employing a climate network approach (Tsonis and Swanson, 2008; Malik et al., 2010; Boers et al., 2014, 2019; Strnad et al., 2022, 2023). Climate networks are networks where links reflect a strong statistical correlation between corresponding nodes, typically representing time series from different spatial locations. Here, the correlation is accessed by quantifying the synchronicity of EREs using the event synchronization algorithm (Quian Quiroga et al., 2002). We adopt the methodology introduced in Strnad et al. (2023) to identify groups of densely connected nodes in the network, referred to as communities. These can be understood as spatial regions where EREs occur significantly synchronously.

Our method uncovers a community of synchronous EREs comprising North India, China, and the Sahel region. We then investigate the mechanisms responsible for the synchronization by building upon the method from Schlör et al. (2024), clustering the propagation of EREs in a multivariate latent space. Our approach reveals a robust time-lagged connection between the ISM and WAM monsoon regions, prompting a subsequent exploration using traditional meteorological analysis tools. We can show that the aforementioned large-scale modes of variability, namely the BSISO, the TEJ, and ENSO interact to foster the synchronization between the ISM and the WAM.

2 Data and Methods

2.1 Data

This study focuses on the analysis of EREs occurring within the monsoon season, set as June through September (JJAS). Consequently, the analysis is limited to this specific time frame.

Precipitation data We use the $0.25°$ daily resolved precipitation data from the Multi-Source Weighted-Ensemble Precipitation (MSWEP) dataset (Beck et al., 2019b) for the period of 1980–2022. The MSWEP dataset is chosen for its longer time range compared to other available multi-satellite precipitation products, improving statistical robustness. It has been shown to represent high rainfall quantiles well on global scales (Beck et al., 2019a). We restrict our analysis to the tropics and the subtropics of the Northern Hemisphere (180° W – 180° E, 30° S – 70° N, see Fig. 1 a).

Extreme Rainfall Events We generate binary event series (Evs) from the precipitation dataset. We only consider “wet days”. These are days with a rainfall sum of at least 1 mm. Events are point-wise defined: We determine one ERE day as a day when the daily precipitation sum exceeds the 0.9th quantile $(Q_{0.9}(\cdot))$ of all wet days at that specific location (Fig. 1 b). This results in a binary time series for each grid point, where 1 indicates an ERE and 0 indicates no ERE.

Reanalysis data To analyze large-scale patterns associated with the synchronization we use the following variables from the ERA5 Global Reanalysis dataset (Hersbach et al., 2020): Daily outgoing longwave radiation (OLR), sea surface temperature (SST), and horizontal (u, v)-wind fields at pressure levels from 50–1000 hPa in steps of 50 Pa, vertical velocity w, specific humidity q and temperature T. The datasets used for the multivariate PCA analysis are interpolated onto a $1° \times 1°$ grid, and all remaining datasets are interpolated onto a $2.5° \times 2.5°$ grid. To ensure the robustness of our analysis and to avoid confounding effects from long-term climate change, we apply a linear detrending to all reanalysis datasets.

ENSO index The ENSO state is obtained using the Multivariate ENSO index version 2 (MEIv2) (Wolter and Timlin, 2011). The MEIv2 is a multivariate index that combines six observed variables over the tropical Pacific Ocean to provide a comprehensive measure of the ENSO state. The MEIv2 is available from the NOAA Physical Sciences Laboratory (PSL) at `https://psl.noaa.gov/enso/mei/` (Last Accessed: 10th April 2024).

BSISO index The daily resolved BSISO index by Kikuchi et al. (2012) is taken from `https://iprc.soest.hawaii.edu/users/kazuyosh/Bimodal_ISO.html` (Last Accessed: 10th May 2023). The index consists of two time series BSISO1 and BSISO2. The amplitude is given by $A = \sqrt{BSISO1^2 + BSISO2^2}$, where $A \geq 1$ ($A < 1$) is regarded as active (inactive) (Wheeler and Hendon, 2004) BSISO. The two-dimensional phase space spanned by BSISO1 and BSISO2 is subdivided into eight equally sized sections that denote the phase of the BSISO.

Tropical Easterly Jet index We employ the index definition of Huang et al. (2019) to describe the characteristics in terms of amplitude and variability of the TEJ for the period from June through September. This index, denoted as Tropical Easterly Jet Index (TEJI), is mathematically formalized as the mean anomaly time series of the zonal winds (u) at the 200 hPa pressure level, of locations in the box within the interval [0° E, 70° E], [0° N, 15° N] (magenta boxes in SI Fig. S4). The metric of TEJI is computed by spatially averaging the u-wind anomalies enclosed within these boxes. We define positive and negative phases of the TEJI as events exceeding one standard deviation from the mean (below for negative, above for positive; see SI Fig. S4).

All India Rainfall Index The All India Rainfall Index (AIRI) by Parthasarathy et al. (1994) available from the India Meteorological Department (IMD) at `https://www.data.gov.in/catalog/rainfall-india` (Last Accessed: 29th August 2024) is used to represent the ISM rainfall variability in JJAS. We consider days above (below) one standard deviation from the mean as active (break) ISM days.

2.2 Communities of Synchronous Extreme Rainfall Events

We use climate networks to determine regions of significantly synchronously occurring EREs. This involves multiple analysis steps, schematically visualized in Fig. 1.

Mapping to a grid of spatially uniformly distributed points On a regular (rectangular) grid on a sphere, grid points are spatially closer together the closer they are to the poles and therefore also more likely correlated. This could cause biases in the community detection algorithm later (see Sec. 2.2). To avoid these confounding correlation effects, we first map the data to a grid of spatially approximately uniformly distributed points using nearest-neighbor interpolation (Fig. 1 a) by employing the Fekete algorithm (Bendito et al., 2007). For computational reasons, the distance between two points of the new grid is set to around 111 km, which corresponds to the spatial distance between two points at the equator of a regular Gaussian 1° grid, resulting in a total of approximately 9600 grid points.

Climate Network Definition A spatiotemporal dataset of rainfall time series is denoted as $\mathbf{X} \in \mathbb{R}^{N \times T}$, where N represents the spatial locations and T is the number of time points (i.e. all days in JJAS for the period 1980-2022, i.e. in total $T = 5002$). The climate network is then defined as $\mathcal{G} = (V, E)$, where each geographical location $i \in 1, \ldots, N$ corresponds to a node $n_i \in V$, and is associated with the rainfall time series $\vec{x}_i = (x_{i,0}, \ldots, x_{i,T}) \in \mathbf{X}$. E represents the set of edges within the network. An edge between two nodes n_i and n_j encodes a robust statistical dependence between the time series $\vec{x}_i(t)$ and $\vec{x}_j(t)$ labeled as edge $e_{ij} \in E$.

Event Synchronization First, all rainfall time series $\vec{x}_i(t)$ are transformed into binary event series $\vec{e}_i(t)$ by thresholding at the 0.9th quantile of all wet days (Fig. 1 b; see above). The Event Synchronization algorithm (Quian Quiroga et al., 2002) is employed to evaluate the statistical dependencies between pairs of event series, serving as a metric for quantifying the synchronization between all pairs of time series within the network (Fig. 1 c, Fig. S1). This method involves counting concurrently occurring events within event sequences from different locations, allowing a temporal gap – the so-called dynamical delay τ – between events in these sequences to account for time deviations of at most 10 days (see SI Sec. SI 1 or Malik et al. (2010)). The strength of the synchronization between locations is assessed by summing up synchronous time points across all event pairs. The statistical significance of synchronization between a pair of event series is estimated using a null-model test with 2000 random surrogates for each pair of event series separately (see SI 1, statistical significance). For details, we refer the reader to Sec. SI 1 or to the detailed descriptions in Boers et al. (2014, 2019); Strnad et al. (2023). The adjacency matrix $\mathbf{A}$ of a network characterizes the linkages between nodes, delineating the network's underlying topology which can be visualized in the form of a graph. It is a mathematical representation of these connections and captures the presence or absence of links between nodes expressed as a $N \times N$ matrix, where $A_{i,j} = 1$ indicates that events at location i are statistically significantly followed by events at location j, and is set to 0 otherwise (Fig. 1 d).

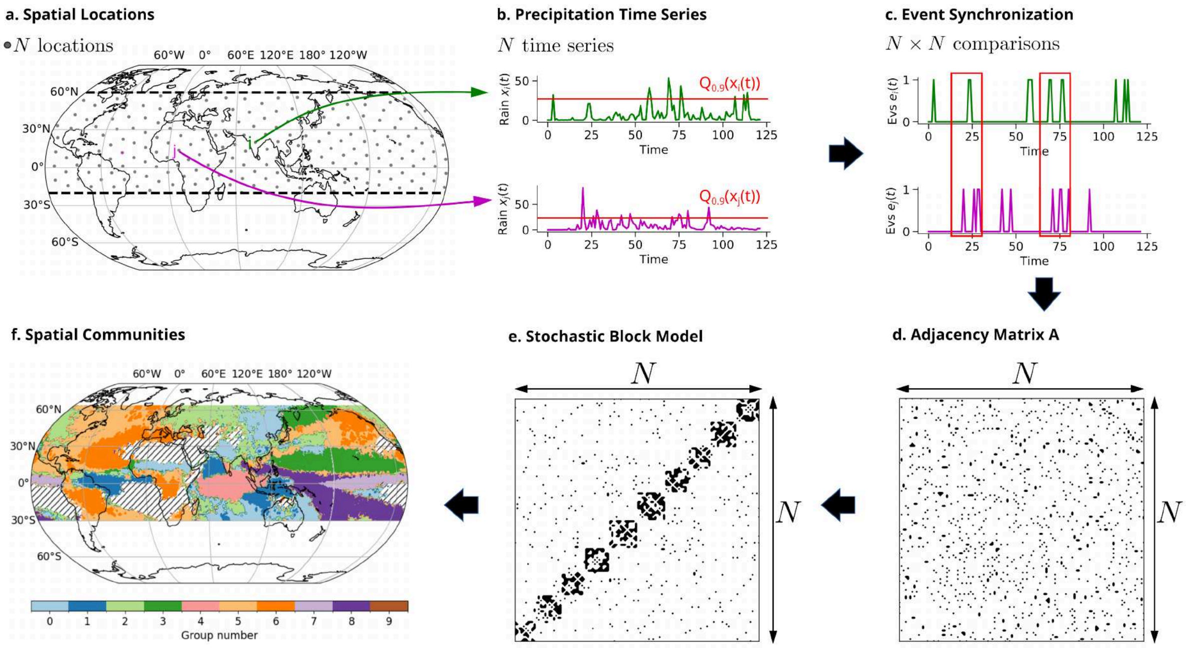


Figure 1: **Schematic of the Community Detection approach.** **a** The climate network is constructed by mapping the data to a spatial grid of N approximately uniformly distributed points using the Fekete algorithm (Bendito et al., 2007). For visualization purposes, only every 10th grid point is plotted. Dashed lines indicate the latitudinal range of the analysis $[20°\,\mathrm{S}, 60°\,\mathrm{N}]$. **b** Every grid point is associated with a precipitation time series. By thresholding locally at the 90th percentile ($Q_{0.9}$), we obtain binary event series (Evs). **c** The Event Synchronization algorithm (Quian Quiroga et al., 2002) is employed to evaluate the statistical dependencies between all pairs of time series (i.e. a total of $N \times N$ comparisons). The time series and event series in **b** and **c** are just for illustrative purposes. **d** The adjacency matrix $\mathbf{A}$ of the network characterizes the linkages between nodes, delineating the network's underlying topology. Each black dot represents a statistically significant synchronization between two nodes. Again, this matrix is for illustrative purposes. **e** Communities are detected by identifying blocks in the adjacency matrix $\mathbf{A}$ via reordering of rows and columns using a probabilistic community detection algorithm. Communities within the climate network represent groups of nodes characterized by stronger internal connectivity compared to their connections with nodes outside the community. **f** The block structure is projected back to the spatial map. Nodes that are assigned to the same community are colored in the same color. The hatched areas indicate regions with only few wet days, which are excluded from the analysis.

Communities of Synchronous Extreme Rainfall Events Communities within the climate network represent groups of nodes characterized by a higher internal density of links compared to their link density with groups of nodes outside the community. Keeping in mind that the network is constructed based on pairs of event series exhibiting statistically significant synchronicity in the occurrence of EREs, these communities correspond to spatial regions where EREs are more likely to occur synchronously. These communities plotted on a map, often correspond to spatially connected regions possibly over large spatial distances (Fig. 1 f). The assumption is that points within the same community are attributable to one (or even multiple) shared underlying physical mechanism(s) (Strnad et al., 2023). In practice, communities are determined by reordering rows and columns of the adjacency matrix $\mathbf{A}$ to obtain a block structure (Fig. 1 e). Very common is an approach that uses a stochastic block model (SBM). The SBM serves as a generative model for random graphs, generating community structures - specifically, subsets of nodes referred to as "blocks". These exhibit distinct connectivity patterns, being connected with one another at particular densities. The model implementation used in this work is the network analysis package *graph_tool* (Peixoto, 2014b) favored for its computational efficiency and its probabilistic output. This implementation can identify the optimal number of communities in the data based on the principle of parsimony, i.e. simple structures with a low number of communities are preferred to more complex ones with a higher number of communities. In our case, the optimal number of communities for the whole network was determined to be 10. Further, the Bayesian approach of the algorithm, makes it possible to quantify the posterior probability of each node belonging to a specific community by sampling from the posterior distribution. We call this the membership likelihood of a node to a community. For more in-depth details, we refer the reader to Peixoto (2014a, 2019).

2.3 Lagged Synchronous Rainfall Index

We introduce an index, denoted as $SRI(t)$, to assess for two (potentially distinct) sets of nodes A and B of the network their time-dependent level of (lagged) synchronicity in the occurrence of EREs. For each time point t, we count the number of EREs that occur at a lag τ_{lag} in all event series e_n^A in set A and are followed by events in all event series e_m^B in set B, where n (m) describe the nth (mth) location in A (B) with N (M) total number of points:

$$SRI_{\tau_{\text{lag}}}(t) = \sum_{n=1}^{N} \sum_{m=1}^{M} e_n^A(t) \cdot e_m^B(t - \tau_{\text{lag}}) \; . \tag{1}$$

We further observe that the lag with the highest synchronization between synchronous events in A and B can vary (Fig. S3 a). Thus, we also count all synchronizations between A and B within a certain range $[\tau_{\text{min}}, \tau_{\text{max}}]$, by using eq. 1. The synchronous rainfall index SRI is estimated as:

$$SRI(t) = \sum_{\tau=\tau_{\text{min}}}^{\tau_{\text{max}}} SRI_{\tau}(t) \; . \tag{2}$$

This counting process enables us to quantify the frequency of EREs within a set of locations A that are followed by events in a set of locations B within a range of possible lags. To further pinpoint the points in time of exceptionally strong (lagged) synchronization, we identify the local maxima in the time series SRI that are above the 90th percentile and define these as the *most synchronous days* (MSDs).

2.4 Latent Space Clustering of Propagation Pathways

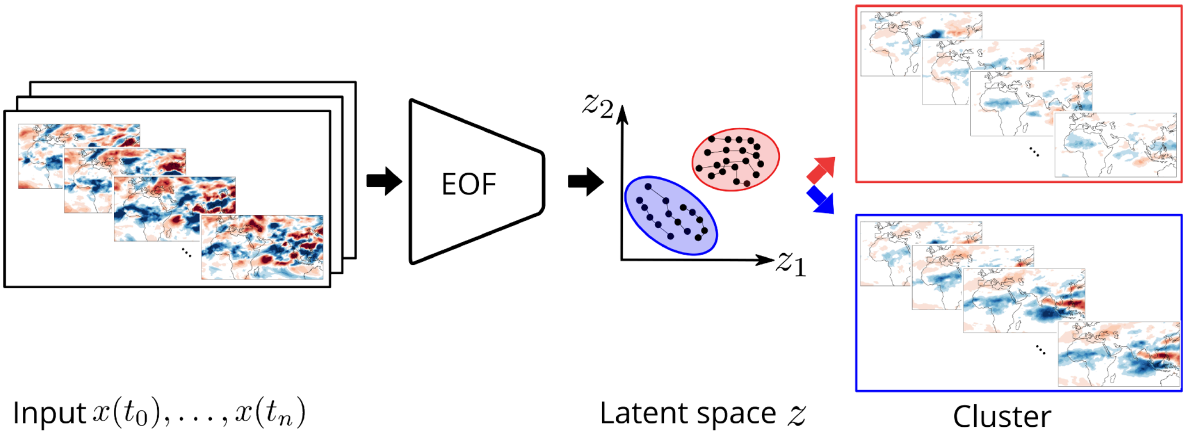


Figure 2: **Sketch of the latent space clustering.** The propagation pathways are clustered via traces in the latent space z. The high-dimensional input field is transformed to a low-dimensional latent space which is denoted by the encoder. In our case, the encoder function is based on a Multivariate Principal Component Analysis (MV-PCA) on 200-hPa u and v, 500-hPa vertical velocity ω and 400-hPa relative humidity. Input fields, $x(t_1), \ldots, x(t_n)$, for a propagation sample of n consecutive time steps are encoded to the latent space z (displayed as black dots). These build a trace in the latent space (visualized by connected lines). These traces are clustered by a clustering algorithm, the clusters are visualized as red and blue regions in the latent space. These correspond to the composite anomaly maps of the input field in the original space.

To obtain meaningful clusters of the propagation pathways, we propose a variation of latent space clustering, schematically visualized in Fig. 2, as it was applied in, e.g., (Schlör et al., 2024). We initially employ empirical orthogonal function (EOF) analysis, also known as principal component analysis (PCA). This well-established technique is used for reducing the dimensionality of spatiotemporal data, while retaining the largest possible fraction of its original variance.

A spatiotemporal field $x \in \mathbf{X}$ can consist of multiple variables. At time t a transformation, $z_t = e(x_t)$, is applied to map the high-dimensional space ,$\mathbb{R}^N$, to a lower-dimensional space,$\mathbb{R}^M$, where $M \ll N$.

To cluster the dynamics of the propagation up to a lag of $+15$ days, we compute the traces in the latent space z, where each sample s concatenates the steps in the latent space z from day 0 to day $+15$: $s = \{z_0, z_1, \ldots, z_{15}\}$. Therefore, each sample t is a vector of length $15 \times n_{EOFs}$, where n_{EOFs} is the number of EOFs used for the clustering. To select the EOFs that best represent the propagation pathways, we consider the EOFs whose PCs have the highest correlation coefficient with the lagged synchronous index $SRI(t)$ (eq. 2).

We employed the k-means clustering algorithm to cluster the traces in the latent space z. To determine the optimal number of clusters, we utilized the silhouette score (Rousseeuw, 1987). It measures the similarity of an object to its own cluster compared to other clusters, ranging from -1 to 1. A high value indicates a strong match to its cluster and a poor match to neighboring clusters. A high average silhouette score across all samples indicates an appropriate clustering configuration. We performed a grid search over the number of clusters and EOFs, ranging from 2 to 10 clusters and 1 to 10 EOFs, in order to identify the optimal clustering configuration. Here, the best clustering is obtained for 2 clusters and a single EOF.

2.5 Estimation of Conditional Probabilities

The probability that a day is a day that belongs to the MSDs (i.e. days where we observe a high number of synchronizations between North India and the Sahel, see Sec. 2.3) under a condition, a (for example the condition that SST state in the tropical Pacific fulfills an La Niña condition), is calculated as follows. Be s denoted as the condition that a day belongs to the set of MSDs S and a a certain condition, where the set of days that fulfill the condition a is labeled as A. Then

$$P(s|a) = \frac{P(s,a)}{P(a)} = \frac{||S \cap A||}{||A||}$$

describes the conditional probability for synchronous events under condition a, where $P(s,a)$ is the joint probability of s and a and $P(a)$ the probability of a. Here, $||\cdot||$ denotes the set cardinality and $S \cap A$ the intersection of days in S and A, i.e. all days that are MSDs and fulfill condition a. Accordingly, the conditional probability for a further, second condition b with a set of days B is computed as:

$$P(s|a,b) = \frac{P(s,a,b)}{P(a,b)} = \frac{||S \cap A \cap B||}{||A \cap B||} \,. \tag{3}$$

This procedure is analogously for further conditions c, d, etc. In our case, the conditional probabilities are estimated for MSDs conditioned on the ENSO state, the BSISO phase and activity, and the TEJ state. A corresponding null model is estimated by assuming that the MSDs are randomly distributed over the time. Then the null model probability that a day is a MSDs is computed by dividing $||S||$ by the total number of days weighted by the relative occurrence of a condition a over time.

3 Results

3.1 Lagged extreme rainfall synchronization between North India and the Sahel

Pattern of the synchronization The community detection algorithm (see Sec. 2.2) applied on the on the global climate network reveals 10 regions of synchronous EREs (Fig. 1 e). These regions are part of a “community” within the climate network constructed by estimating patterns of simultaneously occurring extreme rainfall events of a domain restricted to the tropics and boreal subtropics. Each of these communities could be investigated on its own. This study focuses on the community that encompasses North India including the Indo-Gangetic Plain, the Himalayan foothills (NI, red contour in Fig. 3 a), East Asia (EA, pink contour in Fig. 3 a) delineated by the Himalayan mountain chain, and the Sahel Zone (SZ, green contour in Fig. 3 a) bounded between the tropical rainforest and the Sahara desert. This community is characterized by a high membership likelihood of the nodes in the respective regions to the community. The high spatial coherence indicates that the community is a stable manifestation of the synchronization pattern associated with synchronous EREs (Fig. S2). The synchronization between NI and EA has been analyzed in previous studies and can be attributed to the Eurasian Wave Train associated with the Silk Road pattern (Gupta et al., 2022). The time lag between NI and EA of 3 days (Fig. 3 b, pink line) is consistent with the propagation speed of the Eurasian Wave Train and the results in (Gupta et al., 2022). In the following, we focuses on exploring why the EREs manifest synchronously (within a certain time range) over the large continental-scale distance between North India (red contour in Fig. 3 a) and the Sahel Zone (green contour in Fig. 3a).

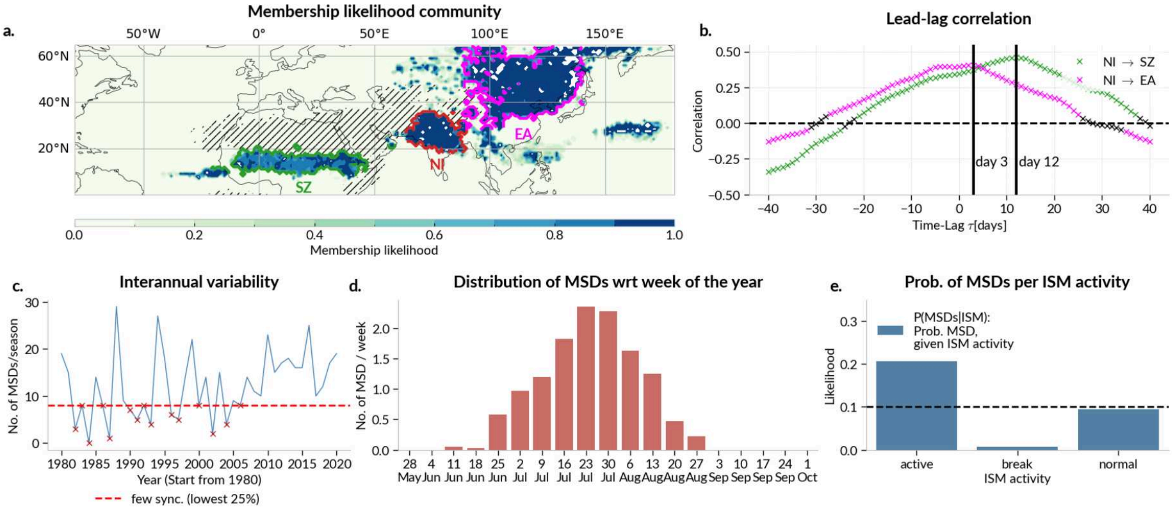


Figure 3: **Synchronization pattern of extreme rainfalls between North India and the Sahel.** **a** Regions of statistically significant synchronization between spatial locations are plotted according to their membership likelihood of being in the community comprising NI (red contour) and SZ (green contour). The communities are determined using a probabilistic hierarchical community detection algorithm based on the Stochastic Block Model and overlaps of 100 independent runs. Hatched areas indicate regions with few wet days, which are excluded from the analysis. We find a community that comprises the regions of North India (NI), East Asia (EA), and the Sahel Zone (SZ). The connection between NI (red contours), SZ (green contours), and EA (magenta contours) is investigated. **b** Lead-lag correlations in steps of days between time series obtained from counting the number of extreme events of locations in the NI and the SZ (green line) and NI and EA (magenta line). Only correlations with $p < 0.001$ are considered significant and shown by colors, while black crosses denote non-significant correlation. The maximum correlation is marked by a vertical solid line. **c** Interannual variability of the synchronization between NI and SZ is displayed by counting the number of MSDs per JJAS season. **d** For the MSDs, we estimate the distribution over the week of the year in the JJAS observation period. **e** The occurrence of MSDs is conditioned on the ISM activity. The dashed line indicates the null model of randomly distributed MSDs per season.

Time lags of the synchronization Synchronizations were observed to peak in strength at a time lag of around 12 days, on average, throughout boreal summer (June, July, August, September) (Fig. 3 b). This characteristic time delay between EREs within the NI region and the SZ locations is estimated by lead-lag correlation analysis. The number of occurrences of synchronization patterns between NI and SZ is not constant but varies substantially between different years (Fig. 3 c). We obtain this yearly fluctuation by counting the number of MSDs per year. For later reference, we mark the lowest 25 % as the least synchronous years (red markers in Fig. 3 c). We also observe that the synchronization occurs mainly during the core monsoon season from July through mid-August (Fig. 3 d). The distribution over the JJAS period is estimated by the week of the year for all identified MSDs. Therefore, hereafter we focus on synchronization events in July and August (JA), which we consider to be the background state under which most synchronization happens. While the occurrence of synchronous events is high mainly during July and August, coinciding with the period of highest daily rainfall sums in India, there is still strong interannual variability within these months.

3.2 Distinct synchronization patterns with varying strength

The synchronization between North India and the Sahel Zone is not instantaneous but lags by almost two weeks (Fig. 3 b). Further, when investigating individual years, we observe that the lag is not constant but varies across years (SI Fig. S3 a) with varying SST background state in the tropical Pacific (SI Fig. S3 b,c,d). We thus investigate potential variations in the diversity of the synchronization patterns. To obtain a meaningful classification of the synchronization events, we apply a clustering approach. As the number of dimensions in the real space is high, we first reduce the dimensionality of the fields by a multivariate PCA (MV-PCA) to identify the principal components (PCs) and cluster in the latent space based on the components that are best correlated with the lagged synchronous index. We expect that the synchronization events are somehow driven by the convective activity, the subsequent moistening of the upper

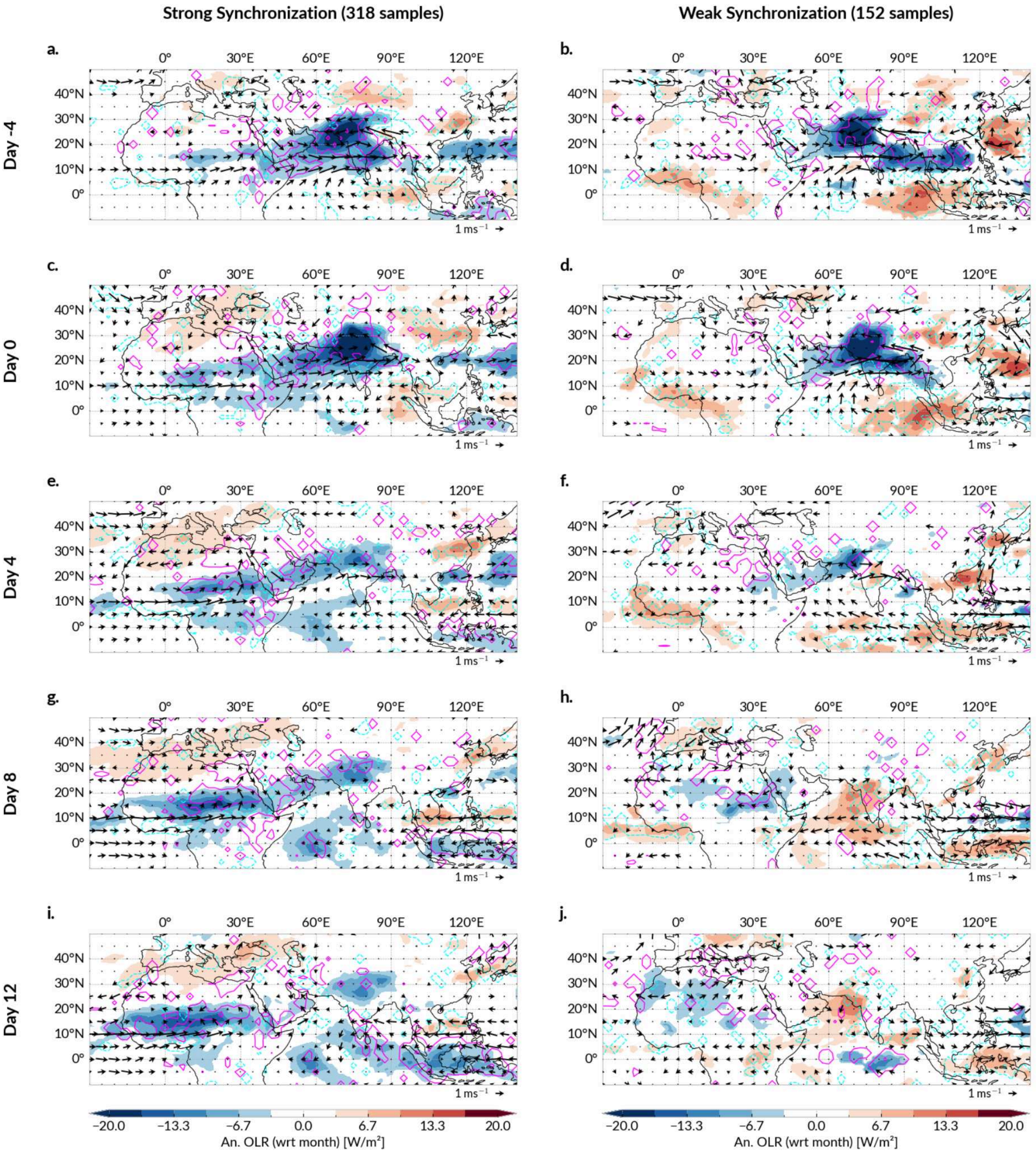


Figure 4: **Propagation for OLR and vertical velocity.** We cluster the propagation pathways based on the most synchronous days using the lagged synchronous index (compare Fig. 3 b). The first column (panels **a,c,e,f**) shows the first cluster of propagation pathways using OLR. The second column (panels **b,d,f,h**) shows the second cluster. We create composite anomaly maps (estimated with respect to the month of the year) for days -4 to +12 OLR in shading, overlapped by vertical velocity ω at 600 hPa (measured in Pa/s) in magenta line contours denoting anomalously rising (sinking) air with ω anomalies larger (smaller) than 1 Pa/s. Colored areas, contours and wind vectors imply statistical significance at 95 % confidence level.

atmosphere and anomalous processes in the upper troposphere (around 200 hPa). This motivates the choice of the following variables for a multivariate PCA (MV-PCA). We use vertical velocity ω at 500 hPa as a proxy for convective activity, relative humidity at 400 hPa to account for the moistening of the mid and upper troposphere and horizontal (u) and horizontal (v) winds at 200 hPa for the propagation of the signal from NI towards SZ. Where clusters of consecutive days occur in the set of MSDs, we remove all but the first to avoid including the same event more than once in our composite. The MV-PCA helps to substantially reduce the dimensionality of the fields and to identify the PCs of the OLR field that are associated with the synchronization (see Sec. 2.4 for details).

We create a trace in the latent space of 15 consecutive days for each time point that is identified as an MSD. We use these traces as input features for a k-means clustering algorithm. The clustering with the highest silhouette score is obtained for $k = 2$ clusters and 1 EOF (Fig. 4). We thus find two distinct clusters, and consistently, the years in which these clusters occur are largely distinct from each other in terms of timing (Fig. S8 d). We observe that the propagation of the negative OLR anomalies is accompanied by a strong rising motion of the air in the mid-troposphere (Fig. 4 a-h). Composite anomaly maps are computed for each cluster and each day in steps of $+3$ days (Fig. 4) revealing two different propagation clusters:

- **Strong Synchronization** The first column in Fig. 4, demonstrates the propagation of anomalous OLR in steps of 4 days, from day 0 to day 12. Note that negative OLR anomalies often coincide with intense rainfall. We see a region of negative OLR anomalies propagating from North India toward the Sahel. These are largely confined to North India and re-emerge later in the Sahel. Strongly anomalous OLR values over the full range of the Sahel are observed from around day $+8$ on (Fig. 4 e). The associated rainfall in the Sahel lags the rainfall in North India by around $10 - 12$ days.
- **Weak Synchronization** The second column of Fig. 4 does not show a clear propagation pattern and the anomalies over SZ are of reduced intensity. Substantially different from the first cluster is not only the intensity but also the speed of the propagation. This is reflected by strongly anomalous OLR values which are observed from around day $+10$ on (Fig. 4 g). The rainfall over the Sahel lags the rainfall in North India by around $+12 - 15$ days.

The clusters also exhibit different background conditions in the Indian Ocean and the Western Pacific. While the first cluster shows persistent convective anomalies over NI and parts of the Maritime Continent, the second cluster reveals a developing pattern of reduced convection at NI, parts of the Arabian Sea and around the equatorial Western Pacific (Fig. 4 b,d,f). However, as the first cluster shows a more intense propagation and the number of samples in cluster 0 is twice the number of samples in cluster 1, we suggest that the first cluster is the main manifestation of the mechanism resulting in the synchronization between NI and SZ, while the second cluster is a weaker variation of the main mechanism. Further, given the smaller cluster size, we cannot rule out that some reasonable fraction of events are not causally connected but arise by chance. Therefore, in the following, we will focus on the first cluster.

3.3 Background conditions favoring the ISM–WAM synchronization

The question arises as to why the synchronization between North India and the Sahel occurs in some years but not in others (Fig. 3 c). We first note that the synchronization is linked to the ISM variability, as the likelihood of the occurrence of the MSD is more likely in active ISM years while there are hardly any synchronizations during break ISM events and normal years correspond more or less to the null mode (Fig. 3 e). To get a clearer understanding of the initiation mechanism, we thus investigate different conditions influencing the variability of the ISM activity.

SST background state enhancing synchronization likelihood We find that the background SST state shows distinct patterns for the two clusters. The Strong Synchronization cluster shows an anomalous cooling in the central Pacific (Fig. 5 a) – a La Niña like pattern – and a band of anomalously warm SSTs in the northwest Pacific at around 40° N, which is known as the Kuroshio-Oyashio extension (Di Lorenzo et al., 2023; Joh et al., 2023). Together, these patterns in the Pacific Ocean resemble the patterns of a negative phase of the Pacific Decadal Oscillation (PDO) in JJAS (see Fig. S7). We find that starting from around 1986, the occurrences of samples that are clustered as Strong Synchronization (Weak Synchronization) are well aligned with the negative (positive) phase of the PDO (Fig. S8a-d) even though the match is not perfect. The negative phase of the PDO favors the occurrence of La Niña events (Di Lorenzo et al., 2023). La Niña-like conditions lead to an anomalously wet monsoon season in North India through modulating the Walker circulation (Xavier et al., 2007) and by assisting the northward propagation of the BSISO towards the north of India (Strnad et al., 2023). We additionally observe associated cool SSTs over the Arabian Sea and the Bay of Bengal, which are probably due to the enhanced convection over the region, i.e. cooling the surface through increased insolation and precipitation. This pattern might also be associated with a negative Indian Ocean Dipole which is often dynamically linked to La Niña events (Meyers et al., 2007; Cherchi and Navarra, 2013). The second cluster does not show a strong SST pattern in the tropical Pacific Ocean (Fig. 5 b). We identify a weak warming pattern in

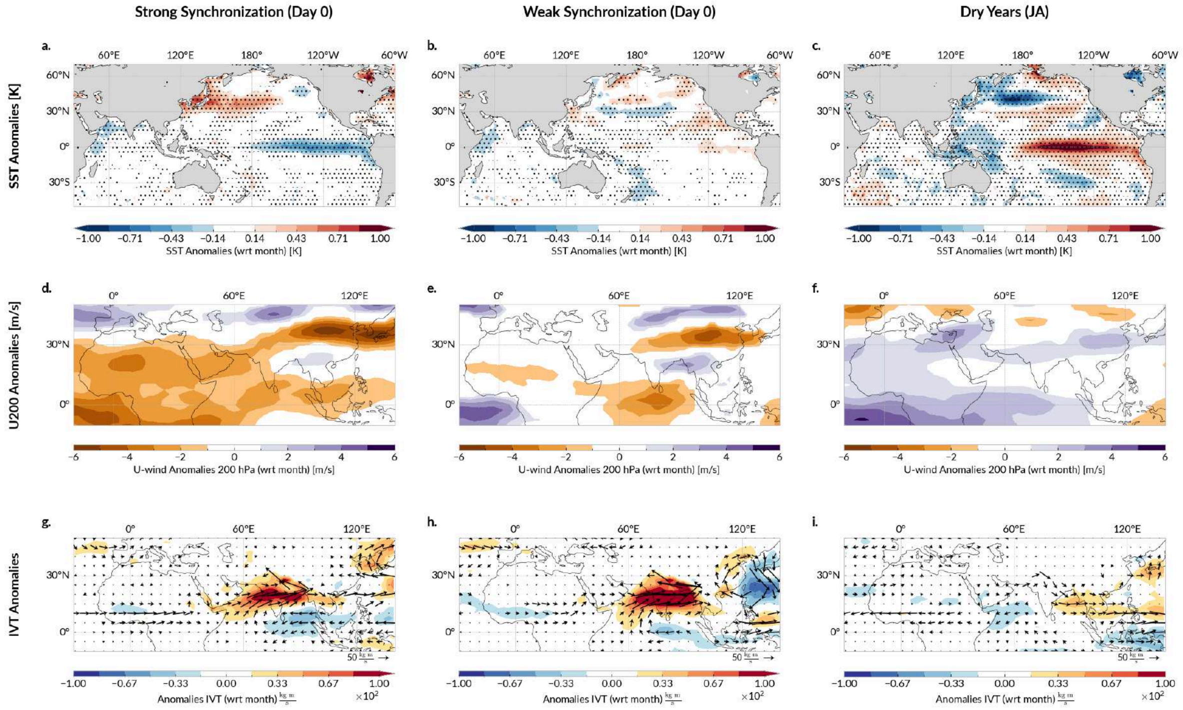


Figure 5: **Background states for propagation clusters.** We visualize different conditions for the samples of the days classified as most synchronous days that were clustered (Fig. 4) into Strong Synchronization (first column) and Weak Synchronization (second column). For comparison, conditions in July-August (JA) for years with few/no synchronizations (see Fig. 3 c) are also shown (third column). The first row (**a–c**) The second row (**g–i**) shows the composited zonal U-wind anomalies at 200 hPa. The third row (**j–l**) shows the Integrated Vapour Transport (IVT) anomalies, arrows denote the corresponding east- and northward components. All composited anomalies are computed with respect to the month of the year. In all panels, colored areas imply statistical significance at 95 $\%$ confidence level using the Student's t-test and are shown. The significance for the wind arrows is computed separately for the U and V component and a wind arrow is only shown if both components are statistically significant.

the Eastern Pacific represented by anomalously warm SSTs in the Pacific Ocean (Fig. 5 b). However, this pattern is less pronounced than the La Niña-like one, suggesting that the La Niña background state provides an important intensification of the synchronization, even if the synchronization can also occur without a La Niña event. For years with only a few synchronizations, which we call dry years, we identify a more pronounced El Niño-like pattern represented by anomalously warm SSTs in the central equatorial Pacific Ocean (Fig. 5 c). This is consistent with the observation that central Pacific El Niño events are more often associated with a weakening of the Indian monsoon (Kumar et al., 2006; Fan et al., 2017) and thus also weakening the synchronization between North India and the Sahel.

Synchronization is guided by the Tropical Easterly Jet The composite wind field at 200 hPa for the Strong Synchronization cluster shows an established corridor that zonally connects the area around the Yellow River Basin with North India and SZ (Fig. 5 a). The flow is guided by the orography of the Tibetan Plateau and the pattern is especially prominent to the south of the Tibetan Plateau. These patterns resemble the characteristics of an enhanced TEJ (Fig. S4). The enhancement is likely a result of the La Niña-like conditions (Fig. 5 a) and the associated Walker circulation response through upper-level disturbances which has been first described by (Shaman and Tziperman, 2007). Induced by La Niña, there is anomalous heating over the Tibetan Plateau (Duan et al., 2012). The subsequent anomalous meridional temperature gradient leads to a strengthening of the TEJ core position over the Indian Ocean (Fig 5 a; also Nithya et al., 2017). Further evidence is provided by the spatial correlation patterns obtained by correlating the 200-hPa zonal wind, 200-hPa meridional wind, and 500-hPa vertical velocity respectively with the lagged synchronous index (see eq.1). The correlation pattern resembles the structure of the tropical easterly jet (Nicholson and Klotter, 2021) (see also SI Fig. S4). The pattern of an enhanced TEJ is not visible for the Weak Synchronization cluster (Fig. 5 b).

Here, the zonal winds are only moderately enhanced over the equatorial Indian Ocean. In dry years, the zonal wind field is not enhanced over the Indian Ocean and the Bay of Bengal (Fig. 5 c). On the contrary, the pattern resembles a weakened TEJ (Fig. S4 b), which can be explained by the El Niño-like SST pattern (Fig. 5 c).

The 200-hPa meridional wind anomalies for composited time points of the Strong Synchronization cluster (Fig. 5 d) show a large-scale wave train pattern (Fig. 5 g) ranging from Japan to North India that resembles the known silk road pattern (Enomoto et al., 2003) (Fig. S6). The ridge in Fig. 5 d is guided by an anticyclone-cyclone-anticyclone (A-C-A) circulation pattern. We observe these as a subtropical anticyclone near Japan (Fig. 5 g) occurring together with an anomalous high known as the Bonin high, a cyclone-like structure over the South of China and anticyclone over the Himalayans, suggesting a relationship with the Silk Road Pattern (Enomoto et al., 2003). The positive CGT drives southwesterlies into the northwest and north of India (Fig. 5 g), which brings additional moisture from the Arabian Sea. At the local-synoptic level, this is similar to the Western disturbances (WD)-tropical storm interactions (Hunt et al., 2021). This wave train pattern reduces to weak northerly anomalies in the second cluster (Fig. 5 e, h). For dry years (Fig. 5 f) it vanishes altogether, leading to a case where the anomalies flow in the opposite direction to the TEJ stream (Fig. 5 i).

Enhanced moisture transport from the Arabian Sea Further, during the MSDs for both clusters, we observe an increased inflow of moist air from the Arabian Sea via the cross-equatorial Somali Jet (Fig. 5 g, h). The Somali Jet is a major source of moisture during the ISM (Rai et al., 2018). It is therefore consistent that the synchronization pattern is mainly established during the core monsoon season (Fig. 3 c) when the Somali Jet is most active. Further evidence provides the observed cold SST pattern in the Arabian Sean (Fig. 5 a, b) which is associated with an increased evaporation and upwelling under the wind stress forcing due to a strengthened Somali jet. When this moisture flux hits the western Himalayan mountains, it leads to forced deep convection (Fig. 6 a, b, d). The convection that initiates the synchronization is thus – at least partially – forced by the orography. For dry years (Fig. 3 c), the synchronization is less pronounced, and the enhanced Somali-Jet inflow vanishes (Fig. 5 i), indicating that the moisture flux from the Arabian Sea is crucial for the initiation of the synchronization.

Anomalous convection in the northwest of India The synchronization is initiated by anomalously strong convection in northwest India (Fig. 6 a). We observe two characteristics. Firstly, the anomalously strong convection in the northwest also coincides with the locations where the strongest rainfall occurs (Fig. 6 b,d) as is usually expected in the tropics. Secondly, we also note a deepening of the monsoon trough for the band of enhanced negative OLR anomalies at around 15° N (Fig. 6 a). We further observe a strong anomalous moistening of the upper atmosphere in Fig. 6 b and d, as well as anomalous easterly zonal winds in the upper atmosphere around 200 hPa.

The anomalously strong convection in northwest India already suggests the presence of the BSISO. An analysis based on conditional probabilities (see Sec. 2.5) supports this impression and shows a substantially increased likelihood for MSDs when the BSISO is in phases 6 and 7 (Fig. 6 c, SI Fig. S12 e). The convective band is thus likely to be associated with the active phase of the BSISO, which is characterized by heavy rainfall (see e.g., Kikuchi (2021); Strnad et al. (2023)).

Combination of processes initiating the synchronization Taken together, the initiation of the synchronization mechanism can be attributed to a threefold process: Firstly, a negative PDO phase favors La Niña conditions (Fig.5a), promoting deep convection over peninsular India, facilitated by the anomalous Walker circulation response (Fig.6b). Secondly, this response supports the progression of the BSISO towards northwest India (Fig. 6 a), substantially enhancing the frequency of deep convection there. Thirdly, La Niña-like conditions also enhance the TEJ (Fig. 5 d, SI Fig. S5). We can provide strong statistical evidence for the importance of these three factors coming together. By estimating conditional probabilities (eq. 3) for the occurrence of MSDs, conditioned on the three impacting factors BSISO, ENSO and the TEJ, we observe that the conditional probabilities for a day in JJAS being a MSD conditioned on La Niña, TEJ, as well as an active BSISO phase is substantially enhanced compared to a respective null model (Fig. 7 e). An enhanced TEJ seems to be a necessary condition for the synchronization to occur, as the likelihood of MSDs is increased when the TEJ is enhanced (Fig. 7 a, c, e) also for El Niño or neutral conditions. For a reduced TEJ we observe hardly any synchronizations, even if the BSISO is active and La Niña conditions are present (Fig. 7 b,d,f) emphasizing the necessary role the TEJ plays in the propagation.

The combination of these three factors also becomes evident in a visual comparison: The pattern of anomalously negative OLR in Fig. 6 a occurs further north than on average for BSISO phases 6 and 7 (SI Fig. S12). However, when compositing all time points that occur in La Niña years, show an enhanced TEJ and are in BSISO phases 6 and 7 (Fig. S13 e,f) the composited map resembles the composited MSD map (Fig. 6 a). The enhanced BSISO pushes moisture further north and northwest towards Pakistan. In this region, the dry static stability is very low, and anomalous

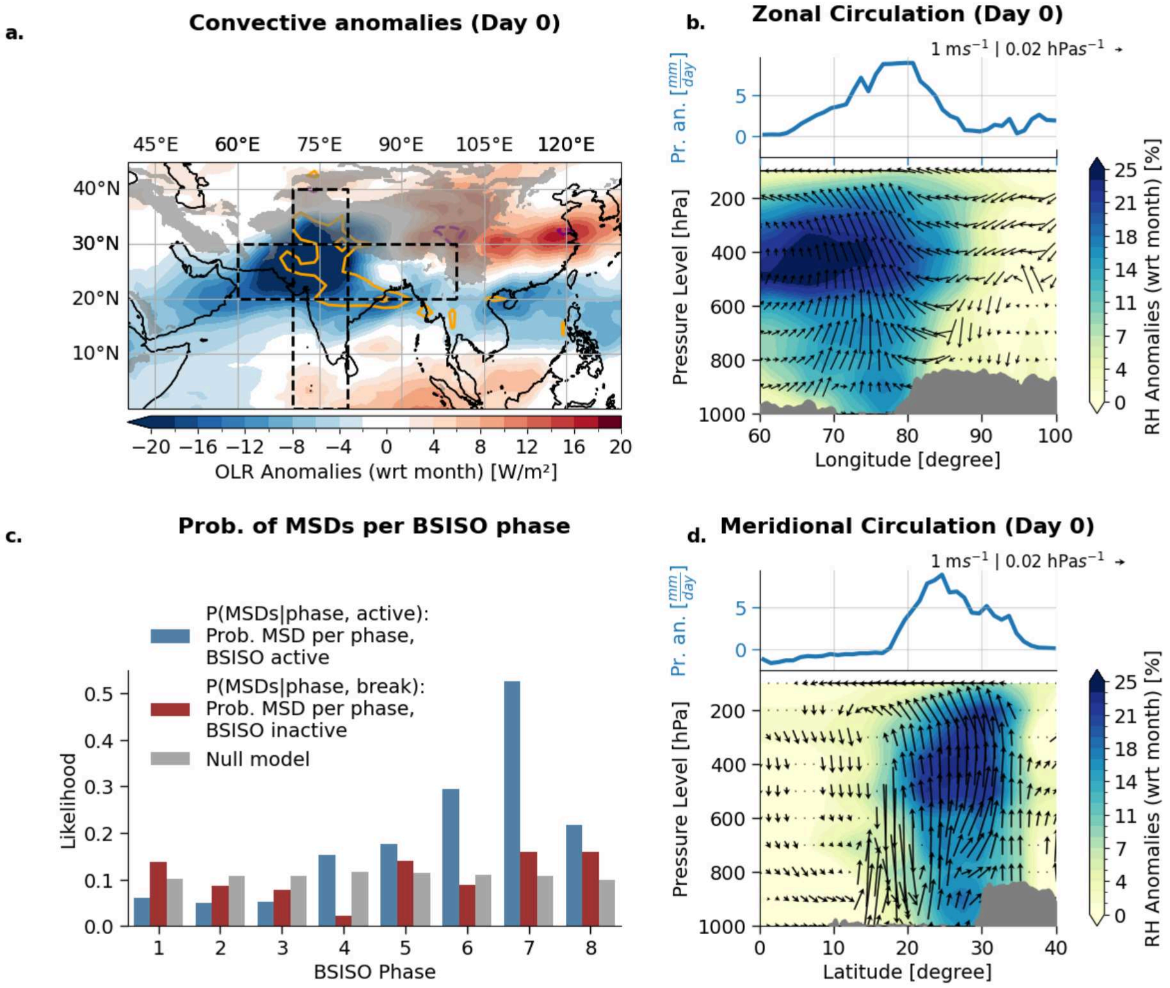


Figure 6: **Convective anomalies in North India.** Panel **a** shows composites of the MSDs for the Outgoing Longwave Radiation (OLR) superimposed by the vertical velocities ω at 400 hPa (measured in Pa/s) by colored contours. Orange solid (purple dashed) contours denote anomalously rising conditions for regions with anomalies larger (smaller) than 3 Pa/s. Grey shading represents the Himalayan mountains. Panel **b** (**d**) visualizes the zonal (meridional) circulation averaged between $20°$ N $-$ $30°$ N ($70°$ E $-$ $80°$ E), visualized by dashed rectangles in **a**. All composites are computed for the conditions on Day 0. Panels **b**,**d** are split into two parts: on top the meridionally (zonally) averaged precipitation anomalies (with respect to the month) are displayed; on the bottom, the vertical circulation is shown by composites of relative humidity anomalies, computed with respect to the month (shading) and wind field anomalies (arrows). In **a**, **b**, **d** colored areas imply statistical significance at a 95 % confidence level (otherwise white) using a two-sided t-test. Grey contours show the highest 10% of the orography, visualizing the Himalayan mountain chain. The wind fields in the zonal (meridional) circulation plots are estimated using the meridionally (zonally) averaged u (v) anomalies, measured in m s^{-1}, and the vertical velocity ω in the horizontal direction, measured in hPa s^{-1}. Only statistically significant arrows at a 95% confidence level using a two-sided t-test are shown. Panel **c** shows the condition probabilities (Sec. 2.5) of the most synchronous days (MSDs) conditioned on active (inactive) BSISO phases marked by blue (red) bars. The gray bars denote the likelihood of a respective null model that randomly distributes the MSDs over the BSISO phases with respect to the relative occurrence of each phase.

moisture readily triggers convection. In consequence, we observe a significantly increased moistening of the middle to upper atmosphere from 500 hPa to 200 hPa (Fig. 6 b).

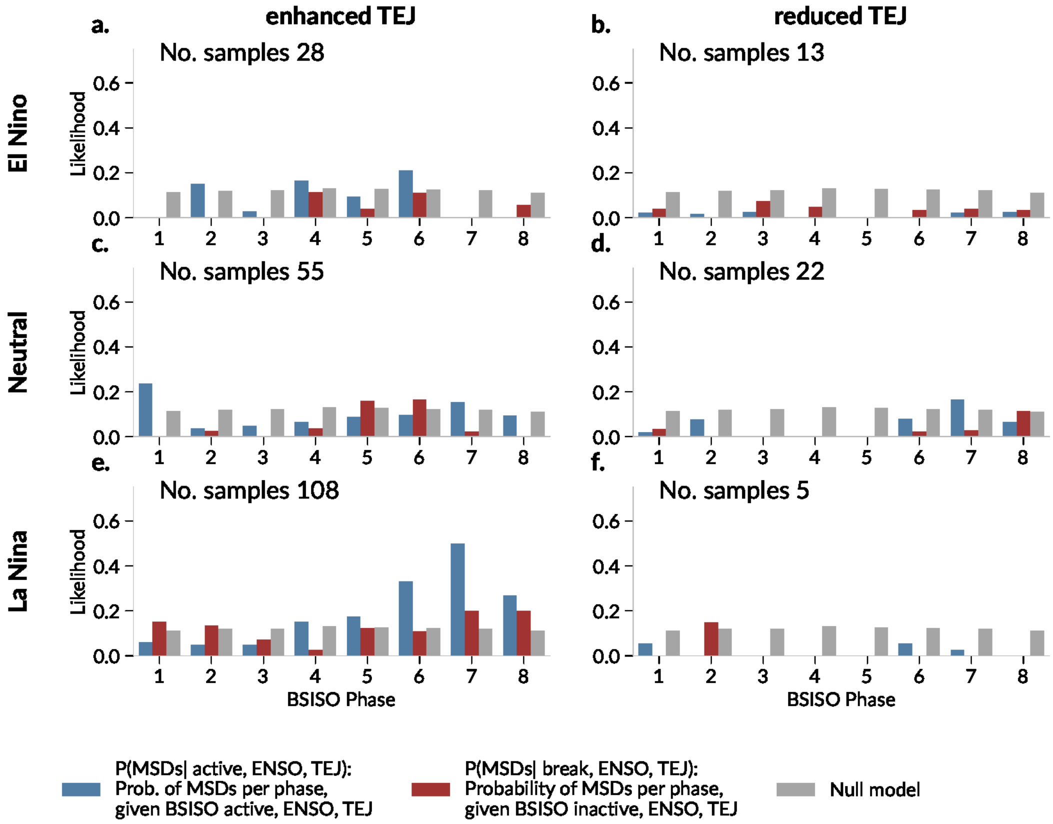


Figure 7: **Probabilities for the occurrence of synchronization conditioned on BSISO phases, ENSO and TEJ. a-f** Most synchronous days (MSDs) are defined using the 0.9 quantile of the lagged synchronous index (see Sec. 2.5 d). Their probability of occurrence is conditioned on the state of the Boreal Summer Intraseasonal Oscillation (BSISO), the Tropical Easterly Jet (TEJ), and the respective phase of the El Niño Southern Oscillation (ENSO) for the respective time points computed as in eq: 3. Blue (red) bars indicate an active (inactive) BSISO state conditioned on the state of ENSO and the TEJ. Grey bars denote the respective null model, i.e. MSDs are assigned randomly to days in the JJAS season and assigned to the distribution of BSISO phases.

3.4 Mechanism driving precipitation in the Sahel Zone

Our results indicate that the anomalous deep convection over NI (Fig. 6 b,c) leads to anomalous moistening of the mid- and upper-troposphere, centered at around 400 hPa, that is then transported towards SZ via the TEJ. The TEJ is also enhanced further southwest of NI such that the propagation over the Arabian Sea and the Sahel Zone is guided by an anomalously strong flow of easterly winds (Fig. 5 d,j). We applied a Lagrangian trajectory analysis (Sprenger and Wernli, 2015) to investigate the number of days it takes to transport the moist anomalies from NI to SZ (see SI Sec. SI 9, Fig. S14). The analysis confirmed that the moist anomaly takes around 6–9 days to arrive over SZ, consistent with the propagation pathways in Fig. 4 a-e.

For timepoints around day 0, we observe a weak orographic forcing over the Ethiopian Highlands and the western coast (Fig. 8). We also see an enhancement of the TEJ over the Sahel, but there is no substantial increase in moisture at around 400 hPa (Fig. 8 a,b). These patterns change around 1–2 days after the arrival of the anomalous moisture in the upper troposphere. We now observe anomalous heavy rainfall over the Sahel (Fig. 8 c,d). The rainfall extends over the full zonal and meridional range of the Sahel from $-5°$ E – 40° E; 5° N – 15° N with two peaks in intensity, one is over the Ethiopian Highlands, one over the Northwest African coast (Fig. 8 c). Composites of Strong Synchronization MSDs show only marginally enhanced intensity at the region of the upper-level TEJ at 200 hPa (compare the zonal circulation at day 0, Fig. 8 a, with day +12, Fig. 8 c). The upper-level westerlies do not vary substantially with respect to the climatology, consistent with (Lemburg et al., 2019). Instead, anomalous moistening happens from near the surface up

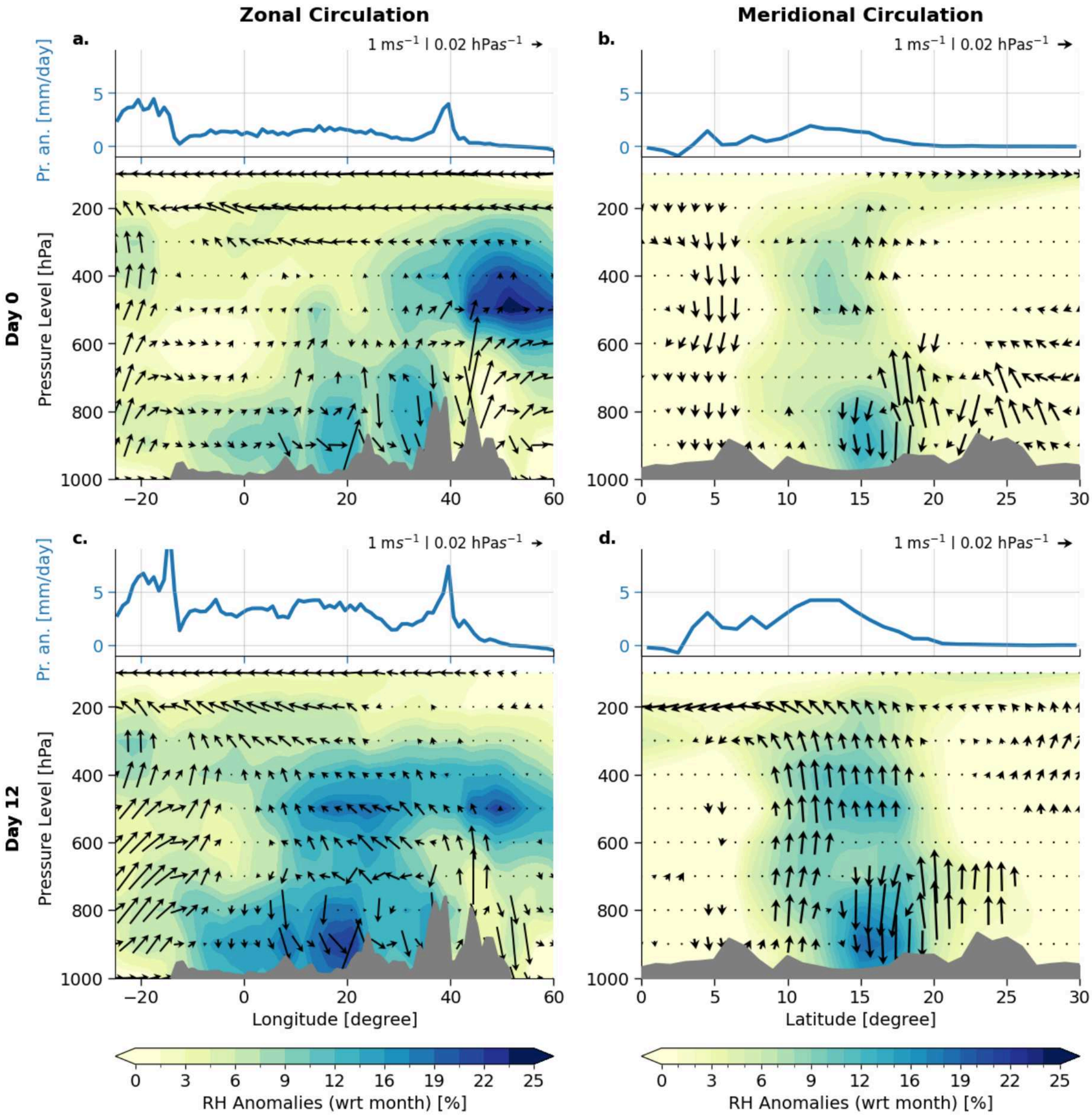


Figure 8: **Composite thermodynamic structure for synchronized EREs over the Sahel.** Conditions for the MSDs of the Strong Synchronization cluster are shown. The first column shows the zonal circulation averaged over 10° N − 18° N, while the second column shows the meridional circulation, averaged over 10° E − 20° E. The first row shows the conditions on Day 0 and the second row shows the average for days 10–12, which are the days of the strongest synchronizations (see Fig. S2 b). Each subplot consists of two panels. The top panel shows the meridionally (zonally) averaged precipitation anomalies (with respect to the month). The bottom panel shows the vertical structure of the circulation in arrows and the filled contours show relative humidity anomalies (computed with respect to the month). Grey contours visualize the orography. The wind fields in the zonal (meridional) circulation plots are estimated using the meridionally (zonally) averaged u (v) anomalies, measured in m s^{-1}, and the vertical velocity ω in the horizontal direction, measured in Pa s^{-1}. Only statistically significant wind arrows at a 95 % confidence level using a two-sided t-test are shown.

to about 400 hPa (Fig. 8 c,d). Hence, it is not the local variability of the TEJ on weekly to subseasonal timescales that modulates the Sahel rainfall; rather its variability on interannual timescales. The strength of the teleconnection also, therefore, does not depend on the local strength but the positioning of the TEJ.

Already at day 0, we find weak anomalous orographic forcing over the Ethiopian Highlands but further moisture incursion is restricted by the Simien Mountains (around 40° E in Fig. 8 a). However, as we see on day +12, the upper-tropospheric moisture anomalies can readily pass over this mountain chain barrier, carried by the TEJ (Fig. 8 c). The anomalous moisture in the east ($>$ 45° E) is transported towards the west and the north of the Sahel Zone (Fig. 8 c,d). At this stage the center of the composite deep convection is around 20° E and 15° N (Fig. 4 e).

The meridional cross-section on day 12 (Fig. 8 d) shows an overturning circulation, where convection occurs slightly off the equator at around 10° N and is extremely deep, extending to the tropopause. This circulation overturns and leads to subsidence at both 5° N and 15° N. The anomalous moisture brought to the mid- and upper-troposphere above the SZ by the TEJ leads to deep instability in a region already characterized by widespread convection during this time of year. Further, the anomalous upper-level moisture supports sustained deep convection by reducing dry entrainment, ultimately leading to heavier widespread precipitation. There is also positive feedback with the overturning circulation facilitating convergence in the lower levels, leading to the observed increase in near-surface humidity. In particular, we observe enhanced integrated vapor transport (IVT) influx driven by the increased convective activity (Fig. S9), which brings additional moisture from the Atlantic Ocean into the continent at levels around 900 hPa (Fig. 8 c, Fig. S10).

4 Discussion

We have identified and explained a teleconnection pattern linking extreme rainfall events (EREs) on a continental scale between North India as part of the ISM and the Sahel Zone as part of the WAM. We first demonstrated that the synchronization pattern of EREs between North India and the Sahel Zone is a robust feature of the boreal summer monsoon system. The synchronization pattern is not a statistical artifact but is rather driven by anomalous large-scale atmospheric circulation, supported by the background state of the climate system. Our work provides a physical explanation for the teleconnection between the Asian and the African Monsoon domains.

We find that the synchronization is initiated by anomalously strong convection in the northwest of India, which is associated with an enhanced active phase of the BSISO. During La Niña conditions there is typically a stronger South Asian monsoon (via the Walker circulation) and, driven by the BSISO, it is more likely for EREs to arise in northwest India (Strnad et al., 2023). Importantly, these EREs substantially moisten the upper troposphere over NI via detrainment associated with deep convection. Subsequently, the anomalous upper-level moisture is advected over the Arabian Sea and Sahel by the TEJ – itself strengthened by La Niña conditions. Moist air aloft over the Sahel provides conditions that are then favorable for supporting and sustaining deep convection (and hence EREs) there. Hence, the strength of this teleconnection depends on the positioning of the TEJ and not on its regional intensity over SZ.

While prior investigations have provided detailed insights into intraseasonal precipitation variability at a regional level in North India (Malik et al., 2010; Stolbova et al., 2014; Boers et al., 2019; Hunt et al., 2021; Hunt and Turner, 2022) and parts of the Sahel (Gleixner et al., 2017; Lemburg et al., 2019; Vashisht et al., 2021), our study integrates and consolidates these findings into a broader context. This synthesis enhances our comprehension of the underlying physical mechanisms, thereby holding the potential to enhance seasonal and sub-seasonal forecasts using windows of opportunities (Mariotti et al., 2020) during boreal summer in the tropical monsoon domain.

Still, some open questions remain. Firstly, the propagation pathway shown in this study was uncovered by using OLR as a proxy for intense rainfall. OLR has the advantage of being directly measurable and therefore a reliable variable. However, it is not a direct measure of precipitation. Secondly, the local causes of the enhanced rainfall in the Sahel Zone remain unclear. Deep convective precipitation over the SZ relies on the interaction of several complex processes, including interactions of mesoscale convective systems with the East African Jet at around 600 hPa and the TEJ. Thirdly, the role of the CGT in the synchronization between North India and the Sahel Zone is not fully understood. We find some patterns indicating a partial modulation by the CGT, but its effect seems to be minor. Further, future research could better investigate the Weak Synchronization propagation cluster. As the previous analysis suggests that the Strong Synchronization is the dominant process, the Weak Synchronization could be a combination of multiple and presumably more local interactions.

There are multiple ways to extend this work. One first step would be to test the occurrence of the synchronization pattern in current general circulation models (GCMs). On the one hand, an analysis of the teleconnection pattern in GCMs could provide insights into the robustness of the synchronization pattern. On the other hand, as precipitation dynamics and teleconnections are not always well reproduced in the current GCMs (Boyle and Klein, 2010; Hess et al., 2022; Douville et al., 2023), the comparison could help to identify potential biases in the models. Another next step could be to integrate the knowledge about the uncovered teleconnection in current operational forecast systems. Of particular interest will be the representation of these synchronizations in deep learning global forecasts (Lam et al., 2023; Bi et al., 2023; Lessig et al., 2023), which are better capable of capturing complex non-linear relationships in the data.

Data Availability

Precipitation data was taken from the MSWEP dataset (https://www.gloh2o.org/mswep)(Beck et al., 2019b). Datasets for the composite analysis from 1979 till date were taken from Copernicus Climate Change Service (C3S) (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels?tab=overview) (Hersbach et al., 2018).

Code Availability

The code for generating and analyzing the networks is made publicly available under Strnad and Schlör (2023). The analysis uses the geoutils package (Strnad, 2023). The code for reproducing the analysis of the network communities and the data analysis in this paper is publicly available under Strnad (2024). Plots were generated using the Cartopy (Met Office, 2010 - 2015) and the Metpy v1.6 libraries (May et al., 2022).

Acknowledgements

F.S. and B.G. acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC number 2064/1 - Project number 390727645. F.S. thanks the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting his PhD program. N.B. acknowledges funding by the Volkswagen Foundation, the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 956170, as well as from the European Union's Horizon Europe research and innovation program under grant agreement No. 101137601. K.M.R.H. is supported by a NERC Independent Research Fellowship (MITRE; NE/W007924/1).

Author's contribution

F.S. and B.G. conceived and designed the study. F.S. conducted the analysis. F.S. and B.G. prepared the manuscript. F.S., K.M.R.H., N.B., and B.G. discussed the results and edited the manuscript.

Competing Interests

The authors declare that they have no competing interests.

References

J. Bader and M. Latif. The impact of decadal-scale Indian Ocean sea surface temperature anomalies on Sahelian rainfall and the North Atlantic Oscillation. *Geophys. Res. Lett.*, 30(22), Nov. 2003. ISSN 0094-8276. doi:10.1029/2003GL018426.

H. E. Beck, M. Pan, T. Roy, G. P. Weedon, F. Pappenberger, A. I. J. M. van Dijk, G. J. Huffman, R. F. Adler, and E. F. Wood. Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. *Hydrol. Earth Syst. Sci.*, 23(1):207–224, Jan. 2019a. ISSN 1027-5606. doi:10.5194/hess-23-207-2019.

H. E. Beck, E. F. Wood, M. Pan, C. K. Fisher, D. G. Miralles, A. I. J. M. van Dijk, T. R. McVicar, and R. F. Adler. MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bull. Am. Meteorol. Soc.*, 100(3):473–500, Mar 2019b. ISSN 0003-0007. doi:10.1175/BAMS-D-17-0138.1.

E. Bendito, A. Carmona, A. M. Encinas, and J. M. Gesto. Estimation of Fekete points. *J. Comput. Phys.*, 225(2): 2354–2376, Aug 2007. ISSN 0021-9991. doi:10.1016/j.jcp.2007.03.017.

J. D. Beverley, S. J. Woolnough, L. H. Baker, S. J. Johnson, and A. Weisheimer. The northern hemisphere circumglobal teleconnection in a seasonal forecast model and its relationship to European summer forecast skill. *Clim. Dyn.*, 52(5): 3759–3771, Mar. 2019. ISSN 1432-0894. doi:10.1007/s00382-018-4371-4.

J. D. Beverley, S. J. Woolnough, L. H. Baker, S. J. Johnson, A. Weisheimer, and C. H. O'Reilly. Dynamical mechanisms linking Indian monsoon precipitation and the circumglobal teleconnection. *Clim. Dyn.*, 57(9):2615–2636, Nov. 2021. ISSN 1432-0894. doi:10.1007/s00382-021-05825-6.

K. Bi, L. Xie, H. Zhang, X. Chen, X. Gu, and Q. Tian. Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, 619:533–538, July 2023. ISSN 1476-4687. doi:10.1038/s41586-023-06185-3.

M. E. Bickle, J. H. Marsham, A. N. Ross, D. P. Rowell, D. J. Parker, and C. M. Taylor. Understanding mechanisms for trends in Sahelian squall lines: Roles of thermodynamics and shear. *Q. J. R. Meteorolog. Soc.*, 147(735):983–1006, Jan. 2021. ISSN 0035-9009. doi:10.1002/qj.3955.

N. Boers, B. Bookhagen, H. M. J. Barbosa, N. Marwan, J. Kurths, and J. A. Marengo. Prediction of extreme floods in the eastern Central Andes based on a complex networks approach. *Nat. Commun.*, 5(5199):1–7, Oct 2014. ISSN 2041-1723. doi:10.1038/ncomms6199.

N. Boers, B. Goswami, A. Rheinwalt, B. Bookhagen, B. Hoskins, and J. Kurths. Complex networks reveal global pattern of extreme-rainfall teleconnections. *Nature*, 566(7744):373–377, feb 2019. ISSN 0028-0836. doi:10.1038/s41586-018-0872-x. URL `http://www.nature.com/articles/s41586-018-0872-x`.

S. Bordoni and T. Schneider. Monsoons as eddy-mediated regime transitions of the tropical overturning circulation. *Nat. Geosci.*, 1:515–519, Aug. 2008. ISSN 1752-0908. doi:10.1038/ngeo248.

J. Boyle and S. A. Klein. Impact of horizontal resolution on climate model forecasts of tropical precipitation and diabatic heating for the TWP-ICE period. *J. Geophys. Res. Atmos.*, 115(D23), Dec. 2010. ISSN 0148-0227. doi:10.1029/2010JD014262.

T. Chen and H. van Loon. Interannual variation of the tropical easterly jet. *Monthly Weather Review*, 115(8):1739–1759, 1987.

A. Cherchi and A. Navarra. Influence of ENSO and of the Indian Ocean Dipole on the Indian summer monsoon variability. *Clim. Dyn.*, 41(1):81–103, July 2013. ISSN 1432-0894. doi:10.1007/s00382-012-1602-y.

E. Di Lorenzo, T. Xu, Y. Zhao, M. Newman, A. Capotondi, S. Stevenson, D. J. Amaya, B. T. Anderson, R. Ding, J. C. Furtado, Y. Joh, G. Liguori, J. Lou, A. J. Miller, G. Navarra, N. Schneider, D. J. Vimont, S. Wu, and H. Zhang. Modes and Mechanisms of Pacific Decadal-Scale Variability. *Annu. Rev. Mar. Sci.*, (Volume 15, 2023):249–275, Jan. 2023. doi:10.1146/annurev-marine-040422-084555.

Q. Ding and B. Wang. Circumglobal Teleconnection in the Northern Hemisphere Summer. *J. Clim.*, 18(17):3483–3505, Sept. 2005. ISSN 0894-8755. doi:10.1175/JCLI3473.1.

Q. Ding and B. Wang. Intraseasonal Teleconnection between the Summer Eurasian Wave Train and the Indian Monsoon. *J. Clim.*, 20(15):3751–3767, Aug. 2007. ISSN 0894-8755. doi:10.1175/JCLI4221.1.

H. Douville, K. Raghavan, J. Renwick, R. Allan, P. Arias, M. Barlow, R. Cerezo-Mota, A. Cherchi, T. Gan, J. Gergis, D. Jiang, A. Khan, W. P. Mba, D. Rosenfeld, J. Tierney, , O. Zolina, and . W. C. Changes. *Water Cycle Changes*, pages 1055–1210. Cambridge University Press, 2023. doi:10.1017/9781009157896.010.

A. Duan, G. Wu, Y. Liu, Y. Ma, and P. Zhao. Weather and climate effects of the Tibetan Plateau. *Adv. Atmos. Sci.*, 29 (5):978–992, Sept. 2012. ISSN 1861-9533. doi:10.1007/s00376-012-1220-y.

T. Enomoto, B. J. Hoskins, and Y. Matsuda. The formation mechanism of the Bonin high in August. *Q. J. R. Meteorolog. Soc.*, 129(587):157–178, Jan. 2003. ISSN 0035-9009. doi:10.1256/qj.01.211.

F. Fan, X. Dong, X. Fang, F. Xue, F. Zheng, and J. Zhu. Revisiting the relationship between the South Asian summer monsoon drought and El Niño warming pattern. *Atmos. Sci. Lett.*, 18(4):175–182, Apr. 2017. ISSN 1530-261X. doi:10.1002/asl.740.

J. Findlater. A major low-level air current near the Indian Ocean during the northern summer. *Q. J. R. Meteorolog. Soc.*, 95(404):362–380, Apr. 1969. ISSN 0035-9009. doi:10.1002/qj.49709540409.

R. Geen, S. Bordoni, D. S. Battisti, and K. Hui. Monsoons, ITCZs, and the Concept of the Global Monsoon. *Rev. Geophys.*, 58(4):e2020RG000700, Dec. 2020. ISSN 8755-1209. doi:10.1029/2020RG000700.

S. Gleixner, N. Keenlyside, E. Viste, and D. Korecha. The El Niño effect on Ethiopian summer rainfall. *Clim. Dyn.*, 49 (5):1865–1883, Sept. 2017. ISSN 1432-0894. doi:10.1007/s00382-016-3421-z.

J. P. Grist and S. E. Nicholson. A Study of the Dynamic Factors Influencing the Rainfall Variability in the West African Sahel. *J. Clim.*, 14(7):1337–1359, Apr. 2001. ISSN 0894-8755. doi:10.1175/1520-0442(2001)014<1337:ASOTDF>2.0.CO;2.

S. Gupta, Z. Su, N. Boers, J. Kurths, N. Marwan, and F. Pappenberger. Interconnection between the Indian and the East Asian Summer Monsoon: spatial synchronization patterns of extreme rainfall events. *Int. J. Climatol.*, n/a(n/a), Sept. 2022. ISSN 0899-8418. doi:10.1002/joc.7861.

D. Halpern and P. M. Woiceshyn. Somali Jet in the Arabian Sea, El Niño, and India Rainfall. *J. Clim.*, 14(3):434–441, Feb. 2001. ISSN 0894-8755. doi:10.1175/1520-0442(2001)014<0434:SJITAS>2.0.CO;2.

H. Hersbach, B. Bell, P. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, J. Nicolas, C. Peubey, R. Radu, I. Rozum, D. Schepers, A. Simmons, C. Soci, D. Dee, and J.-N. Thépaut. Era5 hourly data on single levels from 1979 to present, 2018. Accessed on 02-03-2022.

H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. De Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume, and J. N. Thépaut. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020. ISSN 1477870X. doi:10.1002/qj.3803.

P. Hess, M. Drüke, S. Petri, F. M. Strnad, and N. Boers. Physically constrained generative adversarial networks for improving precipitation fields from Earth system models. *Nat. Mach. Intell.*, 4:828–839, Oct. 2022. ISSN 2522-5839. doi:10.1038/s42256-022-00540-1.

S. Huang, Z. Wen, Z. Chen, X. Li, R. Chen, and Y. Guo. Interdecadal change in the relationship between the tropical easterly jet and tropical sea surface temperature anomalies in boreal summer. *Clim. Dyn.*, 53(3):2119–2131, Aug. 2019. ISSN 1432-0894. doi:10.1007/s00382-019-04801-5.

S. Huang, B. Wang, Z. Wen, and Z. Chen. Enhanced Tropical Eastern Indian Ocean Rainfall Breaks Down the Tropical Easterly Jet–Indian Rainfall Relationship. *J. Clim.*, 34(8):3039–3048, Apr. 2021. ISSN 0894-8755. doi:10.1175/JCLI-D-20-0631.1.

K. M. R. Hunt and A. G. Turner. Nonlinear intensification of monsoon low pressure systems by the BSISO. *Weather and Climate Dynamics Discussions*, pages 1–28, June 2022. doi:10.5194/wcd-2022-31.

K. M. R. Hunt, A. G. Turner, and R. K. H. Schiemann. How Interactions between Tropical Depressions and Western Disturbances Affect Heavy Precipitation in South Asia. *Mon. Weather Rev.*, 149(6):1801–1825, June 2021. ISSN 1520-0493. doi:10.1175/MWR-D-20-0373.1.

X. Jiang, T. Li, and B. Wang. Structures and Mechanisms of the Northward Propagating Boreal Summer Intraseasonal Oscillation. *J. Clim.*, 17(5):1022–1039, Mar. 2004. ISSN 0894-8755. doi:10.1175/1520-0442(2004)017<1022:SAMOTN>2.0.CO;2.

Y. Joh, T. L. Delworth, A. T. Wittenberg, X. Yang, A. Rosati, N. C. Johnson, and L. Jia. The role of upper-ocean variations of the Kuroshio-Oyashio Extension in seasonal-to-decadal air-sea heat flux variability. *npj Clim. Atmos. Sci.*, 6(123):1–11, Aug. 2023. ISSN 2397-3722. doi:10.1038/s41612-023-00453-9.

M. Joly, A. Voldoire, H. Douville, P. Terray, and J.-F. Royer. African monsoon teleconnections with tropical SSTs: validation and evolution in a set of IPCC4 simulations. *Clim. Dyn.*, 29(1):1–20, July 2007. ISSN 1432-0894. doi:10.1007/s00382-006-0215-8.

K. Kikuchi. The Boreal Summer Intraseasonal Oscillation (BSISO): A Review. *Journal of the Meteorological Society of Japan. Ser. II*, pages 2021–045, Mar. 2021. ISSN 0026-1165. doi:10.2151/jmsj.2021-045.

K. Kikuchi, B. Wang, and Y. Kajikawa. Bimodal representation of the tropical intraseasonal oscillation. *Clim. Dyn.*, 38 (9):1989–2000, May 2012. ISSN 1432-0894. doi:10.1007/s00382-011-1159-1.

G. N. Kiladis, J. Dias, K. H. Straub, M. C. Wheeler, S. N. Tulich, K. Kikuchi, K. M. Weickmann, and M. J. Ventrice. A Comparison of OLR and Circulation-Based Indices for Tracking the MJO. *Mon. Weather Rev.*, 142(5):1697–1715, May 2014. ISSN 1520-0493. doi:10.1175/MWR-D-13-00301.1.

P. Koteswaram. The Easterly Jet Stream in the Tropics∗. *Tellus*, Jan. 1958. ISSN 1180-5694. URL `https://www.tandfonline.com/doi/abs/10.3402/tellusa.v10i1.9220`.

M. Kotz, A. Levermann, and L. Wenz. The effect of rainfall changes on economic production. *Nature*, 601:223–227, Jan. 2022. ISSN 1476-4687. doi:10.1038/s41586-021-04283-8.

F. Kucharski, A. Bracco, J. H. Yoo, A. M. Tompkins, L. Feudale, P. Ruti, and A. Dell'Aquila. A Gill–Matsuno-type mechanism explains the tropical Atlantic influence on African and Indian monsoon rainfall. *Q. J. R. Meteorolog. Soc.*, 135(640):569–579, Apr. 2009. ISSN 0035-9009. doi:10.1002/qj.406.

K. K. Kumar, B. Rajagopalan, M. Hoerling, G. Bates, and M. Cane. Unraveling the Mystery of Indian Monsoon Failure During El Niño. *Science*, 314(5796):115–119, Oct. 2006. ISSN 0036-8075. doi:10.1126/science.1131152.

R. Lam, A. Sanchez-Gonzalez, M. Willson, P. Wirnsberger, M. Fortunato, F. Alet, S. Ravuri, T. Ewalds, Z. Eaton-Rosen, W. Hu, A. Merose, S. Hoyer, G. Holland, O. Vinyals, J. Stott, A. Pritzel, S. Mohamed, and P. Battaglia. Learning skillful medium-range global weather forecasting. *Science*, 382(6677):1416–1421, Dec. 2023. ISSN 0036-8075. doi:10.1126/science.adi2336.

J.-Y. Lee, B. Wang, M. C. Wheeler, X. Fu, D. E. Waliser, and I.-S. Kang. Real-time multivariate indices for the boreal summer intraseasonal oscillation over the Asian summer monsoon region. *Clim. Dyn.*, 40(1):493–509, Jan. 2013. ISSN 1432-0894. doi:10.1007/s00382-012-1544-4.

A. Lemburg, J. Bader, and M. Claussen. Sahel Rainfall–Tropical Easterly Jet Relationship on Synoptic to Intraseasonal Time Scales. *Mon. Weather Rev.*, 147(5):1733–1752, May 2019. ISSN 1520-0493. doi:10.1175/MWR-D-18-0254.1.

C. Lessig, I. Luise, B. Gong, M. Langguth, S. Stadtler, and M. Schultz. AtmoRep: A stochastic model of atmosphere dynamics using large scale representation learning. *arXiv*, Aug. 2023. doi:10.48550/arXiv.2308.13280.

V. Madhu. Variation of Zonal Winds in the Upper Troposphere and Lower Stratosphere in Association with Deficient and Excess Indian Summer Monsoon Scenario. *Atmospheric and Climate Sciences*, 4(4):685–695, Aug. 2014. doi:10.4236/acs.2014.44062.

N. Malik, N. Marwan, and J. Kurths. Spatial structures and directionalities in Monsoonal precipitation over South Asia. *Nonlinear Processes Geophys.*, 17(5):371–381, Sep 2010. ISSN 1023-5809. doi:10.5194/npg-17-371-2010.

A. Mariotti, C. Baggett, E. A. Barnes, E. Becker, A. Butler, D. C. Collins, P. A. Dirmeyer, L. Ferranti, N. C. Johnson, J. Jones, B. P. Kirtman, A. L. Lang, A. Molod, M. Newman, A. W. Robertson, S. Schubert, D. E. Waliser, and J. Albers. Windows of Opportunity for Skillful Forecasts Subseasonal to Seasonal and Beyond. *Bull. Am. Meteorol. Soc.*, 101(5):E608–E625, May 2020. ISSN 0003-0007. doi:10.1175/BAMS-D-18-0326.1.

R. Masiwal, V. Dixit, and A. K. Seshadri. Explaining Dynamics and Rapid Onset of the Somali Jet through Its Kinetic Energy Budget. *J. Atmos. Sci.*, 80(3):833–847, Feb. 2023. ISSN 0022-4928. doi:10.1175/JAS-D-22-0160.1.

R. M. May, K. H. Goebbert, J. E. Thielen, J. R. Leeman, M. D. Camron, Z. Bruick, E. C. Bruning, R. P. Manser, S. C. Arms, and P. T. Marsh. Metpy: A meteorological python library for data analysis and visualization. *Bulletin of the American Meteorological Society*, 103(10):E2273 – E2284, 2022. doi:10.1175/BAMS-D-21-0125.1. URL `https://journals.ametsoc.org/view/journals/bams/103/10/BAMS-D-21-0125.1.xml`.

Met Office. *Cartopy: a cartographic python library with a Matplotlib interface*. Exeter, Devon, 2010 - 2015. URL `http://scitools.org.uk/cartopy`.

G. Meyers, P. McIntosh, L. Pigot, and M. Pook. The Years of El Niño, La Niña, and Interactions with the Tropical Indian Ocean. *J. Clim.*, 20(13):2872–2880, July 2007. ISSN 0894-8755. doi:10.1175/JCLI4152.1.

S. E. Nicholson. The intensity, location and structure of the tropical rainbelt over west Africa as factors in interannual variability. *Int. J. Climatol.*, 28(13):1775–1785, Nov. 2008. ISSN 0899-8418. doi:10.1002/joc.1507.

S. E. Nicholson. The West African Sahel: A Review of Recent Studies on the Rainfall Regime and Its Interannual Variability. *Intl. Scholarly Res. Notices*, 2013(1):453521, Jan. 2013. ISSN 2356-7872. doi:10.1155/2013/453521.

S. E. Nicholson and D. Klotter. The Tropical Easterly Jet over Africa, its representation in six reanalysis products, and its association with Sahel rainfall. *Int. J. Climatol.*, 41(1):328–347, Jan. 2021. ISSN 0899-8418. doi:10.1002/joc.6623.

S. E. Nicholson, A. I. Barcilon, and M. Challa. An Analysis of West African Dynamics Using a Linearized GCM. *J. Atmos. Sci.*, 65(4):1182–1203, Apr. 2008. ISSN 0022-4928. doi:10.1175/2007JAS2194.1.

A. C. Nikumbh, A. Chakraborty, G. S. Bhat, and D. M. W. Frierson. Multiscale Interactions between Monsoon Intraseasonal Oscillations and Low Pressure Systems That Produce Heavy Rainfall Events of Different Spatial Extents. *J. Clim.*, 34(23):9523–9534, Dec. 2021. ISSN 0894-8755. doi:10.1175/JCLI-D-21-0231.1.

A. C. Nikumbh, A. B. S. Thakur, A. Chakraborty, G. S. Bhat, and J. Sukhatme. The Role of the North Atlantic Blocking High during Large-Scale Heavy Rainfall Events over Central India. *J. Atmos. Sci.*, 80(7):1815–1827, July 2023. ISSN 0022-4928. doi:10.1175/JAS-D-22-0185.1.

K. Nithya, M. G. Manoj, and K. Mohankumar. Effect of El Niño/La Niña on tropical easterly jet stream during Asian summer monsoon season. *Int. J. Climatol.*, 37(15):4994–5004, Dec. 2017. ISSN 0899-8418. doi:10.1002/joc.5137.

B. Parthasarathy, A. A. Munot, and D. R. Kothawale. All-India monthly and seasonal rainfall series: 1871–1993. *Theor. Appl. Climatol.*, 49(4):217–224, Dec. 1994. ISSN 1434-4483. doi:10.1007/BF00867461.

D. R. Pattanaik and V. Satyan. Fluctuations of Tropical Easterly Jet during contrasting monsoons over India: A GCM study. *Meteorol. Atmos. Phys.*, 75(1):51–60, Nov. 2000. ISSN 1436-5065. doi:10.1007/s007030070015.

T. P. Peixoto. Hierarchical block structures and high-resolution model selection in large networks. *Physical Review X*, 4 (1):1–18, 2014a. ISSN 21603308. doi:10.1103/PhysRevX.4.011047.

T. P. Peixoto. The graph-tool python library. *figshare*, 2014b. doi:10.6084/m9.figshare.1164194. URL `http://figshare.com/articles/graph_tool/1164194`.

T. P. Peixoto. Bayesian Stochastic Blockmodeling. In *Advances in Network Clustering and Blockmodeling*, pages 289–332. John Wiley & Sons, Ltd, Chichester, England, UK, Nov. 2019. ISBN 978-1-11948329-8. doi:10.1002/9781119483298.ch11.

P. Priya, M. Mujumdar, T. P. Sabin, P. Terray, and R. Krishnan. Impacts of Indo-Pacific Sea Surface Temperature Anomalies on the Summer Monsoon Circulation and Heavy Precipitation over Northwest India–Pakistan Region during 2010. *J. Clim.*, 28(9):3714–3730, May 2015. ISSN 0894-8755. doi:10.1175/JCLI-D-14-00595.1.

R. Quian Quiroga, T. Kreuz, and P. Grassberger. Event synchronization: A simple and fast method to measure synchronicity and time delay patterns. *Phys. Rev. E*, 66(4):041904, Oct. 2002. ISSN 2470-0053. doi:10.1103/PhysRevE.66.041904.

P. Rai, M. Joshi, A. P. Dimri, and A. G. Turner. The role of potential vorticity anomalies in the Somali Jet on Indian Summer Monsoon Intraseasonal Variability. *Clim. Dyn.*, 50(11):4149–4169, June 2018. ISSN 1432-0894. doi:10.1007/s00382-017-3865-9.

S. Rao and J. Srinivasan. The impact of latent heating on the location and strength of the tropical easterly jet. *Meteorol. Atmos. Phys.*, 128(2):247–261, Apr. 2016. ISSN 1436-5065. doi:10.1007/s00703-015-0407-z.

P. J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.*, 20:53–65, Nov. 1987. ISSN 0377-0427. doi:10.1016/0377-0427(87)90125-7.

V. Sathiyamoorthy. Large scale reduction in the size of the Tropical Easterly Jet. *Geophys. Res. Lett.*, 32(14), July 2005. ISSN 0094-8276. doi:10.1029/2005GL022956.

V. Sathiyamoorthy, P. K. Pal, and P. C. Joshi. Intraseasonal variability of the Tropical Easterly Jet. *Meteorol. Atmos. Phys.*, 96(3):305–316, June 2007. ISSN 1436-5065. doi:10.1007/s00703-006-0214-7.

J. Schlör, F. Strnad, A. Capotondi, and B. Goswami. Contribution of El Niño Southern Oscillation (ENSO) Diversity to Low-Frequency Changes in ENSO Variance. *Geophys. Res. Lett.*, 51(14):e2024GL109179, July 2024. ISSN 0094-8276. doi:10.1029/2024GL109179.

J. Shaman and E. Tziperman. Summertime ENSO–North African–Asian Jet teleconnection and implications for the Indian monsoons. *Geophys. Res. Lett.*, 34(11), June 2007. ISSN 0094-8276. doi:10.1029/2006GL029143.

K. P. Sooraj, P. Terray, A. Shilin, and M. Mujumdar. Dynamics of rainfall extremes over India: A new perspective. *Int. J. Climatol.*, 40(12):5223–5245, Oct. 2020. ISSN 0899-8418. doi:10.1002/joc.6516.

M. Sprenger and H. Wernli. The LAGRANTO Lagrangian analysis tool – version 2.0. *Geosci. Model Dev.*, 8(8): 2569–2586, Aug. 2015. ISSN 1991-959X. doi:10.5194/gmd-8-2569-2015.

V. Stolbova, P. Martin, B. Bookhagen, N. Marwan, and J. Kurths. Topology and seasonal evolution of the network of extreme precipitation over the Indian subcontinent and Sri Lanka. *Nonlinear Processes Geophys.*, 21(4):901–917, Aug 2014. ISSN 1023-5809. doi:10.5194/npg-21-901-2014.

F. Strnad. fstrnad/geoutils: geoutils, Sept. 2023. URL `https://doi.org/10.5281/zenodo.8354788`.

F. Strnad. monsoon synchronization, 2024. URL `https://github.com/fstrnad/monsoon_synchronization`.

F. Strnad and J. Schlör. climnet v.2.2.0, 2023.

F. M. Strnad, J. Schlör, C. Fröhlich, and B. Goswami. Teleconnection Patterns of Different El Niño Types Revealed by Climate Network Curvature. *Geophys. Res. Lett.*, 49(17):e2022GL098571, Sept. 2022. ISSN 0094-8276. doi:10.1029/2022GL098571.

F. M. Strnad, J. Schlör, R. Geen, N. Boers, and B. Goswami. Propagation pathways of Indo-Pacific rainfall extremes are modulated by Pacific sea surface temperatures. *Nat. Commun.*, 14(5708):1–16, Sept. 2023. ISSN 2041-1723. doi:10.1038/s41467-023-41400-9.

A. A. Tsonis and K. L. Swanson. Topology and Predictability of El Ni\~no and La Ni\~na Networks. *Phys. Rev. Lett.*, 100(22):228502, Jun 2008. ISSN 1079-7114. doi:10.1103/PhysRevLett.100.228502.

A. Vashisht, B. Zaitchik, and A. Gnanadesikan. ENSO Teleconnection to Eastern African Summer Rainfall in Global Climate Models: Role of the Tropical Easterly Jet. *J. Clim.*, 34(1):293–312, Jan. 2021. ISSN 0894-8755. doi:10.1175/JCLI-D-20-0222.1.

R. K. Vellore, R. Krishnan, J. Pendharkar, A. D. Choudhury, and T. P. Sabin. On the anomalous precipitation enhancement over the Himalayan foothills during monsoon breaks. *Clim. Dyn.*, 43(7):2009–2031, Oct. 2014. ISSN 1432-0894. doi:10.1007/s00382-013-2024-1.

J. M. Walker, S. Bordoni, and T. Schneider. Interannual Variability in the Large-Scale Dynamics of the South Asian Summer Monsoon. *J. Clim.*, 28(9):3731–3750, May 2015. ISSN 0894-8755. doi:10.1175/JCLI-D-14-00612.1.

B. Wang and Q. Ding. Global monsoon: Dominant mode of annual variation in the tropics. *Dyn. Atmos. Oceans*, 44(3): 165–183, Mar. 2008. ISSN 0377-0265. doi:10.1016/j.dynatmoce.2007.05.002.

P. Webster. *Dynamics of The Tropical Atmosphere and Oceans*. John Wiley & Sons, Ltd, Mar. 2020. ISBN 978-0-47066256-4. doi:10.1002/9781118648469.

M. C. Wheeler and H. H. Hendon. An All-Season Real-Time Multivariate MJO Index: Development of an Index for Monitoring and Prediction. *Mon. Weather Rev.*, 132(8):1917–1932, Aug 2004. ISSN 1520-0493. doi:10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.

D. Whittleston, S. E. Nicholson, A. Schlosser, and D. Entekhabi. Climate Models Lack Jet–Rainfall Coupling over West Africa. *J. Clim.*, 30(12):4625–4632, June 2017. ISSN 0894-8755. doi:10.1175/JCLI-D-16-0579.1.

K. Wolter and M. S. Timlin. El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI.ext). *Int. J. Climatol.*, 31(7):1074–1087, June 2011. ISSN 0899-8418. doi:10.1002/joc.2336.

P. K. Xavier, C. Marzin, and B. N. Goswami. An objective definition of the Indian summer monsoon season and a new perspective on the ENSO–monsoon relationship. *Q. J. R. Meteorolog. Soc.*, 133(624):749–764, Apr. 2007. ISSN 0035-9009. doi:10.1002/qj.45.

Part III.

DISCUSSION AND CONCLUSION

Discussion & Conclusion 7.

Exploring teleconnection structures in the climate system is crucial for understanding the global large-scale atmospheric dynamics and its impact on local weather patterns. This thesis has investigated the teleconnection patterns from regional to global scale with a focus on extreme rainfall events in the Indo-Pacific Monsoon domain. Three studies were presented that have used different methods to uncover teleconnection structures in the climate system. This final chapter explores the answers given in the manuscripts in Part II to the key questions posed in Chapter 1 and tries to condense and unify the results of the articles. This thesis is concluded by providing an outlook on potential future research directions and try to place the results into a larger context in the field of data-driven Earth system analysis.

7.1. Approaches to Uncover Teleconnection Structures

The first key question raised in Chapter 1 was *how the spatial structures of teleconnections can be uncovered*. The articles included in this thesis interpreted the term "teleconnection structure" in two different ways. It can be understood either as the structure in single teleconnections between various remote regions or as the division of the entire area under investigation into distinct regions that share some characteristics. This thesis has presented approaches to uncover teleconnection structures which each deal with one of these different perspectives on teleconnections: the network curvature approach in Chapter 4 and the community detection approach in Chapter 5 and Chapter 6. The network curvature in Chapter 6 presents an intuitive tool to visualize the most dominant teleconnections for the area under investigation. In Chapter 5 and Chapter 6 teleconnection patterns are interpreted as groups of densely connected nodes in a climate network. The community detection approach in complex networks is used to identify these groups and showed that plotted on a map they correspond to large, spatially extended regions where EREs are significantly synchronized.

The essential difference between the network curvature approach in Chapter 4 and the community-detection approach in Chapter 5 and Chapter 6 is that the former assigns a value to each edge in the network, thereby focusing on single connections. In contrast, the latter focuses on the density of groups of nodes that are connected more densely among each other than to the rest of the network and focuses less on individual links. Consequently, the former approach is more suited to the identification of single teleconnections between two (or more) regions, whereas the latter approach is more appropriate for finding a meaningful subdivision of the entire area under investigation. Both approaches can, in principle, be extended to different large-scale climate phenomena and observable.

Throughout the process of this thesis and the studies presented in Part II, several limitations and drawbacks of the climate network approach to uncover teleconnection structures have been encountered. From the experience collected through the work on the three articles, these can be broadly classified into two categories: fundamental limitations and more practical considerations.

Fundamental limitations

- **Intrinsic noise in network construction** : Using climate networks involves many pitfalls that occur directly in their construction. These pitfalls are discussed in detail in Haas et al. (2023). Due to the covariance-based construction of the network and a high number of comparisons, it is unavoidable that networks contain strong bias through anisotropic autocorrelation that to date can only be overcorrected. In particular, the authors identify various patterns of spurious behavior. For the studies of this thesis, the most relevant are: (i) Unsuitable similarity measure, i.e. if the underlying data is heavy-tailed (like precipitation data used in Chapter 5 and Chapter 6), the correlation coefficient might induce many (bundled) false links. This made it unavoidable to use binary event time series. (ii) Spurious link bundles occur due to local correlation structures. To overcome these issues Haas et al. (2023) propose to construct ensembles of networks by bootstrapping the data. This was not feasible in the studies due to computational constraints for either computing the curvature measure (Chapter 4) or even more the event synchronization algorithm used in Chapter 5 and Chapter 6 which makes it likely that the networks contain a high number of false links. One can still assume that the macroscopic patterns are reliable as both the curvature and community detection approach are analyzing the overall topology of the network.
- **Causal mechanism of Teleconnection structures** Assessing causality in teleconnection structures through climate networks is difficult, as the networks used in the studies are based on correlations without clarifying the underlying causal mechanisms. Classical meteorology refers to the term "causality" with respect to the physical mechanisms that govern teleconnection structures (Kretschmer et al., 2021) by using the concept of counterfactuals (Hannart et al., 2016). Here, one uses climate models and changes artificially the state of the system and observes the response (Mengel et al., 2021). For the studies of this thesis, this was not feasible, and further analysis steps were needed to understand the physical mechanisms, e.g. the Hovmöller diagrams in Chapter 5 and the latent space clustering in Chapter 6. Exploring causal relationships through statistical data analysis is in general challenging, as it requires distinguishing between direct and indirect causal effects, which is often difficult to achieve and a field of active ongoing research (Schölkopf, 2022; Kaddour et al., 2022). One way to obtain causal graphs would be to apply statistical causal discovery algorithms to the network structures, for example by using the PCMCI algorithm based on Granger-causality (Runge et al., 2019). However, the notion of "causality" is controversial in the context of Granger causality which is based on the quantification of how well one time series can be forecasted by another. Some argue it should be better described as "presedence" (Leamer, 1985). Also from a practical point of view, the high computational costs of causal discovery algorithms makes it infeasible to generate a causal graph even for moderately large climate networks with around 1000 nodes as used in this thesis.

Practical limitations

- **Flexibility of the network topology** Once constructed the topology of the network is more or less fixed. This makes it in particular difficult to capture changing spatial patterns over time. Some studies use a moving window approach to capture these changes (for example Ludescher et al., 2014; Gupta et al., 2021; Wiedermann et al., 2016; Fan et al., 2017), but the lower number of datapoints per time series can affect a robust statistics and carries the risk of false link detection due to the intrinsic noise for shorter time scales (Haas et al., 2023).

- **Implementation and Reproducibility** The implementation of constructing climate networks is not straightforward, necessitating complex data preprocessing like the regridding to the Fekete grid (Bendito et al., 2007), intricate statistical analysis (like the event synchronization algorithm in Chapter 5 and Chapter 6), and substantial computational resources, which can pose substantial technical challenges and can make it hard for other researchers to re-implement.

These limitations and drawbacks of the climate network approach are unavoidable. Identified structures should therefore be considered with caution when interpreting the results. In all the studies presented in Part II, additional analysis steps were applied that were to some degree independent of the climate network approach like the Hovmöller diagram approach in Chapter 5 or the latent space clustering in Chapter 6. For all teleconnection patterns that were discussed in this thesis, plausible physical mechanisms were found that could explain the patterns. These mechanisms were not only derived from data used for the network analysis but also from additional data sources like reanalysis data from ERA5.

Overall, the methods used in this thesis demonstrate that the climate network approach can be a powerful tool for the first research step in explaining an observed pattern where a teleconnection mechanism is suspected to be involved: It tells the researcher *where* to look and *at which time points* to look for teleconnection patterns. The question of the *why* often cannot be answered directly and requires additional steps of analysis.

7.2. Modulation of Local Teleconnection Patterns by Global Teleconnections

The second key question of this thesis aimed to answer *how global teleconnections modulate local teleconnections*. Chapter 5 showed how the local ERE teleconnection structure, driven by the BSISO, is modulated by ENSO and could provide a physical explanation how this interaction between ENSO and the BSISO works. This could, for the first time, unify various studies that observed a relation between ENSO and BSISO for individual regions (for example Lin, 2019; Ahn et al., 2020; Liu et al., 2016; Lee et al., 2022). This work allows to place these studies into the broader context of the entire atmospheric dynamics of the whole Indo-Pacific domain. Furthermore, previous work often lacked an impact-based perspective. This study, in contrast, provided a direct link to EREs by design. Similarly, Chapter 6 showed how the local teleconnection pattern between the Indian and Sahel region is modulated by ENSO and its impact on the tropical easterly jet (TEJ) and the BSISO. Even though the term "local" teleconnection pattern might be misleading as the regions are geographically far apart, the term is used here to emphasize that this is still a locally clearly defined and enclosed teleconnection pattern. For the teleconnection between the Indian and Sahel region, it was found that ENSO plays a crucial role that is key to uncover the physical mechanism that modulates the synchronization of EREs. The influence of ENSO was demonstrated to operate in multiple ways. First, La Niña conditions intensify the TEJ what generates an unstable atmosphere above the Sahel. Second, and here the results from the previous study in Chapter 5 could be nicely picked up, La Niña conditions modulate the propagation of the BSISO in the northward direction which brings additional deep convection to the Indian Summer Monsoon domain.

The studies in Chapter 5 and Chapter 6 can be regarded as a proposal how local teleconnection structures can be analyzed with respect to their global modulation. There are some studies that have examined how local weather patterns are modulated by large-scale climate phenomena (for example Cardil et al., 2023; Dalelane et al., 2023; Chen, 2021; Steptoe et al., 2018). The studies of this thesis analyze local teleconnection pattern of precipitation. This set them apart from the majority of studies that are mainly based on slowly varying observable like sea surface temperature (Trenberth et al., 1998), geopotential height (Dalelane et al., 2023) or one-dimensional climate indices (Cardil et al., 2023). Precipitation data is more difficult to analyze due to its high variability and the need for a suitable similarity measure (see Subsection 2.1.1). This is presumably the reason why precipitation teleconnections are comparably less investigated.

7.3. Outlook

7.3.1. Data-Driven Subseasonal-to-Seasonal (S2S) Forecast System

In the past, weather and climate predictions and research therein have been quite strictly separated: On the one hand, weather prediction focuses on forecasting daily weather patterns up to approximately two weeks ahead, constrained by the chaotic behavior of the atmosphere (Zhang et al., 2019). On the other hand, climate forecasting averages fluctuations over a season and beyond to make predictions on a timescale of seasons or longer (Vitart and Robertson, 2019). The development of a data-driven sub-seasonal to seasonal (S2S) forecast system tries to bridge the gap between weather and climate prediction (Hoskins, 2013). As outlined by Mariotti et al. (2020) phenomena like the MJO, BSISO and ENSO modulate teleconnections which implies potential predictive power for the atmospheric processes on the respective time scale of these phenomena. Especially the MJO and its summer counterpart the BSISO is considered a leading source of potential predictability on the sub-seasonal time scale (i.e. multiple weeks ahead). Even though not all mechanisms are fully understood yet, this would allow S2S forecasts not only in the tropics but also to higher latitudes on the characteristic time scale of the MJO of around $30-60$ days (Woolnough, 2019). Of particular interest will be the integration of these teleconnection patterns in deep learning global forecast models (Lam et al., 2023; Bi et al., 2023; Lessig et al., 2023). Deep learning algorithms have shown great potential in capturing complex patterns and relationships in data, and they could be used to identify the patterns and precursors associated with EREs. It is a current field of active research to explore how these deep learning models perform also on the S2S scale.

In this sense, leveraging the insights from this thesis could improve the accuracy and reliability of a data-driven S2S forecast system to enable predictions in the Indo-Pacific Summer Monsoon. Chapter 5 suggests that including information about the SST state of the tropical Pacific Ocean should improve the prediction of EREs in the Indo-Pacific region. A way to test this hypothesis would be to train a statistical model to predict the OLR field in the Indo-Pacific region given the current field of OLR, upper and lower-level wind fields (to capture the circulation structure of the BSISO) and the state of the tropical Pacific Ocean and compare the results to a model that does not include the information on the Pacific SSTs. Here, the former should have a substantially higher skill in predicting the OLR field. In the outlook section in Chapter 5, it was shown that the BSISO-driven propagation pathways of EREs enable predictability beyond the 14-day limit of medium-range weather forecasts which is a promising result as precipitation and especially EREs are difficult to predict even on short timescales (Mariotti et al., 2020). An early warning system for EREs would probably not provide a precise prediction of the amount of rainfall but rather provide a probability map that indicates for a certain forecast horizon (e.g. 4 weeks) the likelihood of an ERE to occur. This could be a valuable tool for disaster management and agriculture in the region.

7.3.2. Future changes of teleconnection structures

The climate system is undergoing rapid changes as a result of anthropogenic climate change (Gulev et al., 2021), and these changes are expected to have significant impacts on teleconnection structures (Rezaei et al., 2023). First, warmer global temperatures directly affect ENSO teleconnections (Geng et al., 2024). Second, following the Clausius-Clapyron relation, it is expected that the frequency and intensity of extreme weather events will increase due to the increased amount of water vapor in the Earth system (Ummenhofer and Meehl, 2017; Mo et al., 2019; Rodell and Li, 2023). The impact of this distribution shift on the patterns of teleconnections is currently widely discussed (Liu et al., 2023; Haszpra et al., 2020). It is an unexplored research question of the studies presented in this thesis to understand how the identified teleconnection patterns are changing and to assess the potential impact of these changes on the predictability of extreme weather events. The most straightforward approach to achieve this would be to utilize the model data from the CMIP6 ensemble (Eyring et al., 2016) under different future emission scenarios (the so-called SSP scenarios, Gulev et al., 2021) and proceed in a three-step analysis: The initial step is to ascertain whether the observed teleconnections are present in the historical runs, evaluating the models' general capacity to reproduce the teleconnections. Secondly, an investigation should be conducted using climate networks with a sliding window of, for example, 30 years, in order to ascertain how the teleconnection structures are changing over time under different SSP scenarios. Thirdly, the objective is to identify the underlying physical mechanisms that may be responsible for any observed changes in the teleconnection structures. However, this approach would involve certain practical challenges. It would be necessary to eliminate potential spatial biases in the CMIP6 data such that the distribution of rainfall in the model data aligns with the observed distribution. This could be achieved, for example, by using generative neural networks (Hess et al., 2022, 2023). Further, the low resolution of the majority of CMIP6 models necessitates additional downscaling of the data to capture the regional patterns of teleconnections. For the downscaling of global precipitation data, recent research has demonstrated that Generative Foundation Models represent the current state of the art (Hess et al., 2024).

Nevertheless, based on the findings of previous studies, an attempt is made to hypothesize how the teleconnection structures might change. With ongoing anthropogenic warming, the Walker Circulation over the Pacific is expected to weaken, promoting the increased occurrences of both strong El Niño and La Niña events (Cai et al., 2015) making the associated teleconnections more stable and pronounced in the future (Alizadeh, 2024). Thus, the identified differences of global teleconnection structures of EP and CP El Niño events in Chapter 4 will potentially become even more pronounced. Most CMIP models concur that the amount of rainfall in the Indo-Pacific region will increase (Turner and Annamalai, 2012; Papalexiou and Montanari, 2019). Additional studies indicate that the tropical oscillations (MJO, BSISO) are likely to intensify in terms of precipitation (Maloney et al., 2019) while the eastward propagation accelerates (Cui and Li, 2019) and more regular oscillation patterns occur (Du et al., 2024). Taken together, this would imply that the community structures identified in Chapter 5 will become even more stable. The enhanced intensity of ENSO events may facilitate the interaction between the BSISO and ENSO, resulting in a more pronounced modulation of the propagation pathways of EREs. Therefore, it may be expected that the predictability of the propagation of rainfall extremes will also increase, as a consequence of the enhanced stability of the teleconnection structures. Recent studies have indicated an enhanced predictability of the region over the past decade (Du et al., 2024). Changes in the Tropical Easterly Jet (TEJ) are less frequently discussed in the literature. Studies on historical changes in the TEJ have shown that the TEJ is stronger (weaker) during La Niña (El Niño) years (Huang et al., 2019). One recent study has found a moderate strengthening of the TEJ in the last two decades (Sharma et al., 2024). This would imply that the propagation pathway identified in Chapter 6 will also become more pronounced and the impact on EREs over the Sahel would increase.

7.3.3. Future Perspectives

On a more general level, this work has underlined the potential and importance of integrating new methods from the field of statistical machine learning in the traditional field of climate science. The approach followed in all our studies, namely starting from a macroscopic perspective, followed by reducing the data to the most relevant patterns and finally investigating the underlying physical mechanisms, has proven to be useful and successful.

In the future, the increasing availability of high-resolution climate data and the advancements in deep learning will allow a more comprehensive feature extraction from the data. Nevertheless, finding the respective physical explanation will still remain the challenging but also fascinating part. I am convinced that the combination of deep learning and physical understanding of the processes in the climate system will be key to future climate research, and I am looking forward to seeing the future development in the field.

Part IV.

APPENDIX

Supplementary Material for Chapter 4

A.

GEOPHYSICAL RESEARCH LETTERS

Supporting Information for “Teleconnection patterns of different El Niño types revealed by climate network curvature”

Felix M. Strnad[1,*], Jakob Schlör[1,*], Christian Fröhlich[1], Bedartha Goswami[1]

[1]Cluster of Excellence Machine Learning: New Perspectives for Science, Universität Tübingen, 72072 Tübingen, Germany

[*]These two authors contributed equally to the paper

Contents of this file

1. Text S1 to S8
2. Figures S1 to S16
3. Table S1

Introduction

In this Supplementary Material to our article, we describe in detail the definition of different El Niño flavors (Text S1, Table S1). The average climatic conditions for the respective conditions are visualized in Fig. S1. The individual seasonal averages for the El Niño EP and CP conditions are shown in Fig. S2 and Fig. S3 respectively.

We further describe the variability of differently strong El Niños (Text S2, Fig. S2, S3, S4). The implementation details of the toy dataset are provided in Text S3. Further, we give details of the estimation of climate networks (Text S4, Fig.S5, S6), and provide some information on the commonly used network measure of betweenness centrality as a comparison to our curvature method (Fig. S7). A mathematically precise description of the Forman-Ricci curvature method is provided (Text S5, Fig. S8). We also present some further sensitivity analysis checks to different network densities (Text S6, Fig. S11, S10) and to differently strong El Niño events (Fig. S12, S13) We compare our observational data results to results obtained from taking the model output from the UKESM1-0-LL model that is part of the CMIP6 project (Text S7, Fig.S14) and use the pre-industrial control run output to estimate the robustness of our method to small sample sizes of events (Fig. S15). We also present additional figures for the discussion of the results including a local correlation analysis (Text S8, Fig.7).

Text S1. Additional information on identifying El Niño types

To separate El Niño events into Eastern Pacific (EP) and Central Pacific (CP) types, we follow the definition of Capotondi, Wittenberg, Kug, Takahashi, and McPhaden (2020): The DJF Niño-3 (N3) is larger than 0.5 and larger than the DJF Niño-4 (N4) for EP events, while the DJF Niño-4 is larger than 0.5 and larger than the DJF Niño-3 for CP events. We use the December to March period of the identified EP and CP events to construct the networks. The classification of El Niño events and normal conditions are shown in table S1. The monthly SST indices are obtained by averaging the monthly SST anomalies from ERA5 dataset (Hersbach et al., 2018). Anomalies are computed with respect to multiple centered 30-year base periods, which are successively updated in 5-year steps. For the end (beginning) of the time period, we continue to use the last (first) available 30-year base period. We further distinguish El Niño types between “moderate” and “strong”. A moderate EP El Niño is identified as N3 (N4) being $0.5\,\mathrm{K} < N3(N4) < 1\,\mathrm{K}$, a strong El Niño as $N3(N4) \geq 1\,\mathrm{K}$.

Text S2. The spatial extension and variability of El Niño flavors

We are interested in the center of the anomalies during the specific El Niños. The contours for every year are computed based on the quantiles Q of the seasonal spatial maps for each individual El Niño event (Fig. S2, S3). The aggregation of the yearly $Q(0.85)$-DJF contours yields Fig. S4 suggesting a more extensive spatial diversity of the 8 CP events compared to the 7 EP events.

Text S3. Additional information on toy dataset

We use a toy dataset to provide an intuition for the network construction and network curvature. We use a stochastic block model (SBM) to generate a random graph with four equal-sized communities. We set their edge probabilities to

$$P = \begin{bmatrix} 0.7 & 0.02 & 0.02 & 0.03 \\ 0.02 & 0.6 & 0.02 & 0.01 \\ 0.02 & 0.02 & 0.6 & 0.02 \\ 0.03 & 0.01 & 0.02 & 0.7 \end{bmatrix}.$$

The adjacency matrix $A_{\text{toy}} \in \mathcal{R}^N$ of the SBM graph are used to specify the covariance structure of a multivariate Gaussian, $\mathcal{N}\,(x, 0, A_{\text{toy}})$, with N is the number of nodes/data points. We draw $T = 1000$ samples from the Gaussian which forms the N time series of length T of our toy dataset $X_{\text{toy}} \in \mathcal{R}^{NxT}$ with known covariance structure.

Text S4. Additional information on climate network estimation

0.1. Grid

We account for the fact that on a rectangular Gaussian grid the data points are spatially much closer to each other at the poles than at the equator resulting in artificially higher correlation values at the poles which are solely due to the shorter geographical distances not because of physical mechanism (Rheinwalt et al., 2012). We use next-neighbor interpolation to map the data to a grid of spatially approximately uniformly distributed points using the Fekete algorithm (Bendito et al., 2007) to avoid these spurious correlation patterns close to the poles (Ebert-Uphoff & Deng, 2012). In Figure S5 we visualized the effect of using the Fekete interpolation. Note that the reason why the points at the Fekte grid do not seem to be uniformly distributed as well is due to the distortion bias of the chosen projection (i.e. 3-dimensional sphere cannot be projected onto a 2-dimensional space without a bias).

0.2. Climate network construction

Assume complex network graph G, defined by its set of nodes V, connected by its set of edges E, where e_{ij} denotes an edge connecting node v_i to node v_j. This type of graph is described by its adjacency matrix $\mathbf{A}$:

$$\mathbf{A}_{ij} = \begin{cases} 1, & e_{ij} \in E \\ 0, & \text{otherwise} \end{cases},$$

If the edges e_{ij} are weighted by weights w_{ij}, the weighted graph is described by

$$\mathbf{W}_{ij} = \begin{cases} w_{ij}, & e_{ij} \in E \\ 0, & \text{otherwise} \end{cases},$$

We estimate the weighted adjacency matrix $\mathbf{W}$ of the climate network by placing links between pairs of locations which have a correlation value among the 2% strongest absolute correlations. Thus, for the correlation threshold $\rho_{0.98} = Q_{|\rho|}(0.98)$, where $Q_X(\cdot)$ denotes the quantile function for X, we define

$$\mathbf{W}_{ij} = \begin{cases} |\rho_{ij}|, & |\rho_{ij}| > \rho_{0.98} \\ 0, & \text{otherwise} \end{cases}, \tag{1}$$

where $i, j \in \{1, \ldots, N\}$ index spatial locations and ρ_{ij} is the Dec–Feb correlation between locations i and j. $\mathbf{W}_{ij}$ thus defines a network G by a set of edges (or links) $e_{ij} \in E$, where e_{ij} connects the nodes $(v_i, v_j) \in V$ with link weight $w_{ij} = |\rho_{ij}|$. We compute instantaneous correlations using Spearman's rank-order correlation to capture nonlinear behavior between all pairs of time series. For a detailed introduction to climate networks, we refer the reader to (Dijkstra et al., 2019).

Accounting for autocorrelation in the data, a two-sided test for non-random correlations at a confidence level of 1% yields a threshold $|\hat{\rho}| = 0.1$. While our threshold is far higher, due to the high number of hypotheses tests (3.6×10^7), we nevertheless expect a non-negligible number of network links to be false positives. We thus additionally use a spatial null model which assumes that correlations caused by physical mechanisms are likely to occur in bundles of links (Boers et al., 2019).

For each spatial location, we randomly rewire its corresponding network links 2000 times and use a Gaussian kernel density estimator (KDE) to get the likelihood of a link to the chosen location. An observed network link to the chosen location is considered statistically significant if the spatial likelihood of the link (also obtained via Gaussian KDE) is above the 99.9-th percentile of the local null model link distribution. For our SAT climate networks, we found that 2 % of all links were identified as spurious links. In Fig S6 we visualized an example of the correction method by the Gaussian KDE for the location of the Indian subcontinent. In this example, spurious links in the Equatorial and Northern Atlantic Ocean and the Southern Indian Ocean are removed.

0.3. Complex Network Measures

Betweenness Centrality: Betweenness Centrality is defined for nodes v and edges e as,

$$BC_v(v_i) = \sum_{s,t}^{N} \frac{\sigma(v_s, v_t | v_i)}{\sigma(v_s, v_t)} \;, \tag{2}$$

$$BC_e(e_{ij}) = \sum_{s,t}^{N} \frac{\sigma(v_s, v_t | e_{ij})}{\sigma(v_s, v_t)} \;, \tag{3}$$

where $\sigma(v_s, v_t)$ denotes the number of shortest paths between nodes v_s and v_t and $\sigma(v_s, v_t|v_i) \leq \sigma(v_s, v_t)$ the number of all shortest paths that include node v_i. Similarly for edge betweenness centrality $\sigma(v_s, v_t|e_{ij}) \leq \sigma(v_s, v_t)$ yields the number of all shortest paths that include edge e_{ij}. For instance, betweenness centrality measures were used in (Freeman, 1977; Donges et al., 2009; Boers et al., 2013; Ciemer et al., 2018). BC_n is often referred to as the pathway of a variable through the network and therefore taken as an indicator for the flow of the variable of interest. Figure S7 yields the node betweeness centrality for Normal (S7 Fig. **a**), EP (Fig. S7 **b**) and CP (Fig. S7 **c**) conditions.

Text S5. Additional information on estimation of Forman-Ricci curvature

The Ricci-curvature of a network link describes how the connectivity of its network neighborhood differs from the connectivity of a regular grid. Forman curvature of edge e_{ij} in an undirected network with weight $w_{ij} \in \mathbf{W}$ is estimated as,

$$F_{ij} = w_{ij} \left(|\mathcal{T}_{ij}| \cdot w_{ij} + \frac{2}{w_{ij}} - \sum_{\substack{k=1 \\ w_{ik}w_{kj}=0}}^{N} \sum_{\substack{l \in \{i,j\} \\ w_{lk}>0}} \frac{1}{\sqrt{w_{ij} \cdot w_{lk}}} \right), \qquad (4)$$

where $\mathcal{T}_{ij} := \{v_k : w_{ik}w_{jk} > 0\}$ denotes the set of nodes in the neighborhood of v_i and v_j which form triangles containing edge e_{ij} (green edges in Fig. S8a) and $|\cdot|$ denotes set cardinality. The last term in Eq. 4 counts the number of edges adjacent to node v_i and v_j which do not form triangles with edge e_{ij} (black edges in Fig. S8 a). Equation 4 approximates the "augmented" Forman curvature (cf. Samal et al. (2018), Eq. 9) by considering only triangles, no node weights, and no cycle weights. Since curvature is in general not symmetrically distributed around zero, we use the 10% most positive curvatures,

$$F^{+}_{ij} := \{F_{ij} : F_{ij} > Q_F(0.9)\}, \qquad (5)$$

and the 10% most negative curvatures F_{ij}^- (defined similarly as values less than $Q_F(0.1)$. We also define the Forman-curvature of a network node,

$$f_i := \frac{\sum_j F_{ij}}{k_i} \tag{6}$$

to identify geographical locations, or 'hotspots,' connected to strongly negatively or positively curved links. Since curvature is distributed around zero but not necessarily symmetrically, i.e the value ranges of curvature differ for different networks, we use the min-max transform to normalize it to $(-1, 1)$, denoted by $\tilde{F}_{ij}$. Corresponding to $\tilde{F}_{ij}^+$, we use Eq. 6 to define $\tilde{f}_i^+$ as positive node-curvature hotspots and similarly, $\tilde{f}_i^-$ for negative node-curvature hotspots. Consider a random network with four communities (Fig. S8 b, c) generated using a stochastic block model from NetworkX (Hagberg et al., 2008). Forman-curvature separates the between-community links from the within-community links. Within-community links are typically part of triangles, indicating local convergence of shortest paths, i.e. positive curvature. Conversely, links connecting nodes with a high degree that are not part of triangles indicate local divergence of shortest paths and negative curvature. Forman-curvature provides a continuous measure over network links that indicates if an edge is inside a community or if it straddles two communities. By comparison, the edge betweenness centrality (see sec. 0.3)— typically used to identify edges that connect communities — fails to identify many between-community links (Fig. S8 c). This is likely due to the binary notion of the shortest path - either a path is the shortest or it is not, implying that "almost-shortest" paths are not considered (Eq. 3).

Text S6. Sensitivity analysis of climate network estimation

The density of the climate networks used for the analysis was chosen following the literature on climate networks ((Donges et al., 2009; Wiedermann et al., 2016; Ciemer et al., 2020). The density is supposed to be a good trade-off between significant correlation and sufficiently sparse networks. Here we show, that the results of our analysis are however robust to moderate modifications. We analyze networks of densities 2 % and 5 %. As shown in Figure S10 and Figure S11 the qualitative patterns for positive F_{ij}^{+} and negative F_{ij}^{-} node curvature remain the same. Especially, the preference of negative links for EP conditions to the tropics (Figure S11) and to the extratropics for CP conditions (Figure S10) is still clearly visible.

We further examine the difference in the curvature patterns between climate networks that are composed of only strong (moderate) El Niños. As shown in Fig **??**, it is the strong El Niños that dominate the tropically concentrated teleconnection pattern for the EP El Niños.

Text S7. Additional information on comparison to CMIP6 model data

In order to estimate the error due to the finite amount of EP and CP El Niño events available in the reanalysis data, we repeat our analysis on data from General Circulation Models (GCMs). Model data have the advantage that they can run for a longer period of time and thereby in principle produce an infinite amount of data to compare to. Here, we use the pre-industrial control run of the UKESM1-0-LL model from the CMIP6 project which covers 1000 years. We use UKESM1-0-LL because it is known to reproduce similar sea surface temperature variability in the tropical Pacific than found in observations (Dieppois et al., 2021). We again use surface air temperatures on a daily resolution interpolated on an equidistant Fekete grid. Using sea surface temperature data from the same model run, we identify 296 EP El Niño, 168 CP El Niño, and 653 Normal winters in the pre-industrial run.

We compute the correlation-based networks and Forman curvature on the pre-industrial model run. The most positive and most negative curvature hotspots are shown in Fig. S14. The UKESM1-0-LL model resembles similar teleconnection patterns found in the reanalysis data. Teleconnections are more constrained to the tropical ocean basins under EP conditions while under CP conditions teleconnections are found in the mid-latitudes.

With 296 EP events from the pre-industrial run at hand, we can estimate the error due to the finite amount of events in the observational period. We randomly sample 7 from the 296 available EP events and calculate the curvature. We repeat the sampling procedure 10 times. We compare the obtained distribution of curvature values of the 7 events to the distribution obtained with all 296 EP events (see Fig. S15 b). The curvature distribution of two samplings are shown in Fig. S15 c. We find that the distributions cover the same curvature range and show a peak close to zero which is also the case for the curvature distribution obtained from ERA5 (shown in Fig. S15 a.). This suggests that the UKESM1-0-LL model represents the topology of the SAT reanalysis data well.

However, the curvature distributions obtained from 7 events show higher densities at negative curvatures than the distribution obtained from all EP events. The overestimation of negatively curved links in the network due to the small number of events is also found in the analysis of ERA5 data (see Fig. S15 a c). We hypothesize that a shorter time period leads to spurious pair-wise correlations which form links over long spatial distances and are thus negatively curved.

This hypothesis is supported by comparing the spatial extent of the curvature obtained by using all 296 EP events from the pre-industrial run (Fig. S15 d-e) and the average over curvatures obtained from repeated sampling with 7 EP events each (Fig. S15 f-g). While the most spatial patterns of positively curved edges are similar, we obtain more negatively curved edges to the extra-tropics in the networks constructed from only 7 events. However, pronounced regions of negative curvature are consistent between the analysis using 296 EP events and 7 randomly sampled events. We also find an overestimation of negative curvature using only 7 randomly sampled CP events out of the 168 events in the pre-industrial run (not shown).

Text S8. Additional information on the mathematical definition of local correlation

We introduce the average local correlation R_i to better understand the regional dynamics and variability of spatially close points. We define the local correlation as the average correlation of the time series $x_i(t)$ associated with the node v_i to every time series of the node v_b in the set of nodes $\mathcal{B}_i$, where $\mathcal{B}_i$ describes the set of the $B = 200$ spatially closest next-neighbor nodes to v_i:

$$R_i = \frac{1}{B} \sum_{v_b \in \mathcal{B}_i} \rho(x_i(t), x_b(t)) \ . \tag{7}$$

Here, ρ describes the Spearman's rank order correlation function between the time series $x_i(t)$ and $x_b(t)$. Note, as we have interpolated the grid to a Fekete Grid, the nodes are approximately uniformly distributed, and therefore no north-south bias due to higher latitudes is included (Heitzig et al., 2012). The results are depicted in Fig. S16.

Open Research

Datasets for this research are available from Copernicus Climate Change Service. The data from 1979 till date was taken from Hersbach et al. (2018) and the data from 1959 to 1979 from Bell et al. (2020). The code for generating and analyzing the networks is made publicly available under Strnad and Schlör (2022). The code for analyzing the network curvature and reproducing the analysis described in this Supporting Information file is publicly available under Schlör and Strnad (2022).

References

Bell, B., Hersbach, H., Berrisford, P., Dahlgren, P., Horányi, A., Muñoz Sabater, J., … Thépaut, J.-N. (2020). *Era5 hourly data on single levels from 1950 to 1978 (preliminary version).* Copernicus Climate Change Service (C3S) Climate Data Store (CDS). Retrieved from `https://cds.climate.copernicus-climate.eu/cdsapp#!/dataset/reanalysis-era5-single-levels-preliminary-back-extension` (Accessed on 02-03-2022)

Bendito, E., Carmona, A., Encinas, A. M., & Gesto, J. M. (2007, Aug). Estimation of Fekete points. *J. Comput. Phys.*, *225*(2), 2354–2376. doi: 10.1016/j.jcp.2007.03.017

Boers, N., Bookhagen, B., Marwan, N., Kurths, J., & Marengo, J. (2013, Aug). Complex networks identify spatial patterns of extreme rainfall events of the South American Monsoon System. *Geophys. Res. Lett.*, *40*(16), 4386–4392. doi: 10.1002/grl.50681

Boers, N., Goswami, B., Rheinwalt, A., Bookhagen, B., Hoskins, B., & Kurths, J.

(2019). Complex networks reveal global pattern of extreme-rainfall teleconnections. *Nature*, *566*(7744), 373–377. Retrieved from `http://dx.doi.org/10.1038/s41586-018-0872-x` doi: 10.1038/s41586-018-0872-x

Capotondi, A., Wittenberg, A. T., Kug, J.-S., Takahashi, K., & McPhaden, M. J. (2020, Nov). ENSO Diversity. In *El Niño Southern Oscillation in a Changing Climate* (pp. 65–86). American Geophysical Union (AGU). doi: 10.1002/9781119548164.ch4

Ciemer, C., Boers, N., Barbosa, H. M. J., Kurths, J., & Rammig, A. (2018, Jul). Temporal evolution of the spatial covariability of rainfall in South America. *Clim. Dyn.*, *51*(1), 371–382. doi: 10.1007/s00382-017-3929-x

Ciemer, C., Rehm, L., Kurths, J., Donner, R. V., Winkelmann, R., & Boers, N. (2020, Sep). An early-warning indicator for Amazon droughts exclusively based on tropical Atlantic sea surface temperatures. *Environ. Res. Lett.*, *15*(9), 094087. doi: 10.1088/1748-9326/ab9cff

Dieppois, B., Capotondi, A., Pohl, B., Chun, K. P., Monerie, P.-A., & Eden, J. (2021, October). ENSO diversity shows robust decadal variations that must be captured for accurate future projections. *Commun. Earth Environ.*, *2*(212), 1–13. doi: 10.1038/s43247-021-00285-6

Dijkstra, H. A., Hernández-García, E., Masoller, C., & Barreiro, M. (2019). *Networks in climate.* doi: 10.1017/9781316275757

Donges, J. F., Zou, Y., Marwan, N., & Kurths, J. (2009, sep). The backbone of the climate network. *EPL (Europhysics Letters)*, *87*(4), 48007. doi: 10.1209/0295-5075/87/48007

Ebert-Uphoff, I., & Deng, Y. (2012). A new type of climate network based on probabilistic graphical models: Results of boreal winter versus summer. *Geophysical Research Letters*, *39*(18), 1–7. doi: 10.1029/2012GL053269

Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, *40*(1), 35–41. Retrieved from `https://www.jstor.org/stable/3033543`

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function using networkx. In G. Varoquaux, T. Vaught, & J. Millman (Eds.), *Proceedings of the 7th python in science conference* (p. 11 - 15). Pasadena, CA USA. Retrieved from `http://conference.scipy.org/proceedings/SciPy2008/paper_2`

Heitzig, J., Donges, J. F., Zou, Y., Marwan, N., & Kurths, J. (2012, Jan). Node-weighted measures for complex networks with spatially embedded, sampled, or differently sized nodes. *Eur. Phys. J. B*, *85*(1), 1–22. doi: 10.1140/epjb/e2011-20678-7

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., ... Thépaut, J.-N. (2018). *Era5 hourly data on single levels from 1979 to present.* Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 02-03-2022) doi: 10.24381/cds.adbb2d47

Rheinwalt, A., Marwan, N., Kurths, J., Werner, P., & Gerstengarbe, F.-W. (2012, October). Boundary effects in network measures of spatially embedded networks. *Europhys. Lett.*, *100*(2), 28002. doi: 10.1209/0295-5075/100/28002

Samal, A., Sreejith, R. P., Gu, J., Liu, S., Saucan, E., & Jost, J. (2018, Jun). Comparative analysis of two discretizations of Ricci curvature for complex networks - Scientific

Reports. *Sci. Rep.*, *8*(8650), 1–16. doi: 10.1038/s41598-018-27001-3

Schlör, J., & Strnad, F. (2022). *netcurvature.* Zenodo. doi: https://doi.org/10.5281/zenodo.6325299

Strnad, F., & Schlör, J. (2022). *climnet.* Zenodo. doi: https://doi.org/10.5281/zenodo.6325661

Wiedermann, M., Radebach, A., Donges, J. F., Kurths, J., & Donner, R. V. (2016, Jul). A climate network-based index to discriminate different types of El Niño and La Niña. *Geophys. Res. Lett.*, *43*(13), 7176–7185. doi: 10.1002/2016GL069119

Table S1. Classification of EP El Niño, CP El Niño and ‘normal’ conditions based-on Niño-3 – Ninño-4 approach. The DJF Niño-3 is larger than 0.5 and larger than the DJF Niño-4 for EP events, while the DJF Niño-4 is larger than 0.5 and larger than the DJF Niño-3 for CP events. Boreal winters where the average Niño-3 and Niño-4 index is between -0.5 and 0.5 are labeled as ‘Normal’

Year	Condition	Year	Condition
1959 – 1960	Normal	1990 – 1991	CP
1960 – 1961	Normal	1991 – 1992	EP
1961 – 1962	Normal	1992 – 1993	Normal
1963 – 1964	CP	1993 – 1994	Normal
1966 – 1967	Normal	1997 – 1998	EP
1968 – 1969	CP	2001 – 2002	Normal
1969 – 1970	Normal	2003 – 2004	Normal
1972 – 1973	EP	2004 – 2005	CP
1976 – 1977	Normal	2006 – 2007	EP
1977 – 1978	CP	2012 – 2013	Normal
1978 – 1979	Normal	2013 – 2014	Normal
1979 – 1980	CP	2014 – 2015	CP
1981 – 1982	Normal	2015 – 2016	EP
1982 – 1983	EP	2016 – 2017	Normal
1986 – 1987	EP	2019 – 2020	CP
1989 – 1990	Normal		

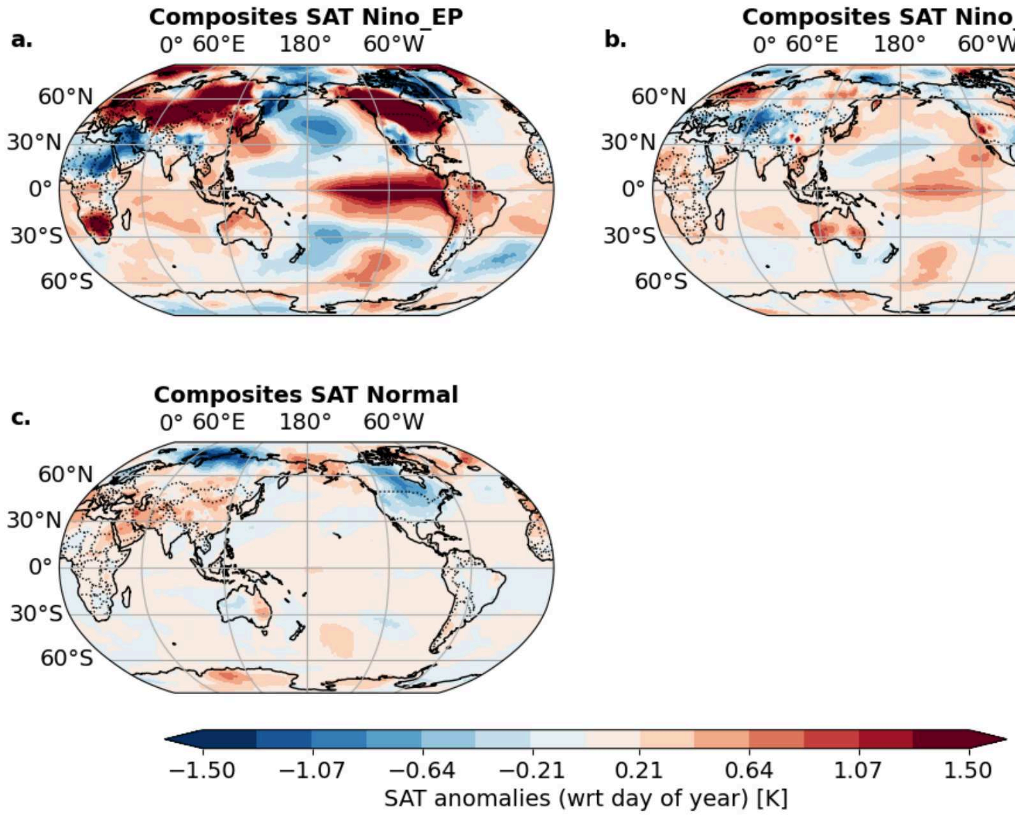


Figure S1. Averages of all DJF seasons identified as EP, CP and Normal condition. We compute composite anomalies for DJF time series for (a) EP El Niño, (b) CP El Niño, and (c) Normal conditions.

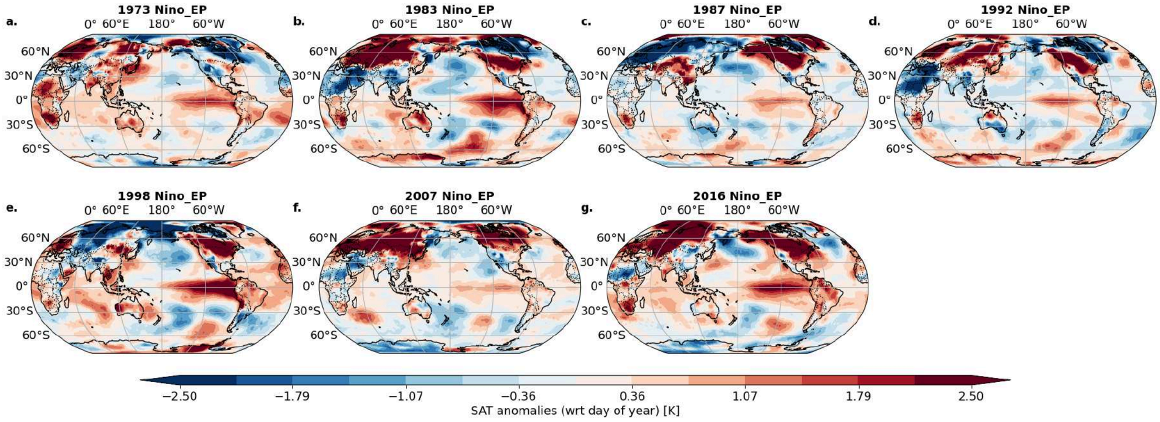


Figure S2. Seasonal DJF averages for all years identified as EP El Niño. We compute seasonal average anomalies for DJF time series separately for each year between 1959 and 2020. The title of each subplot denotes the year of the end of the season in which the event occurred. In total 7 events have been identified as EP El Niño.

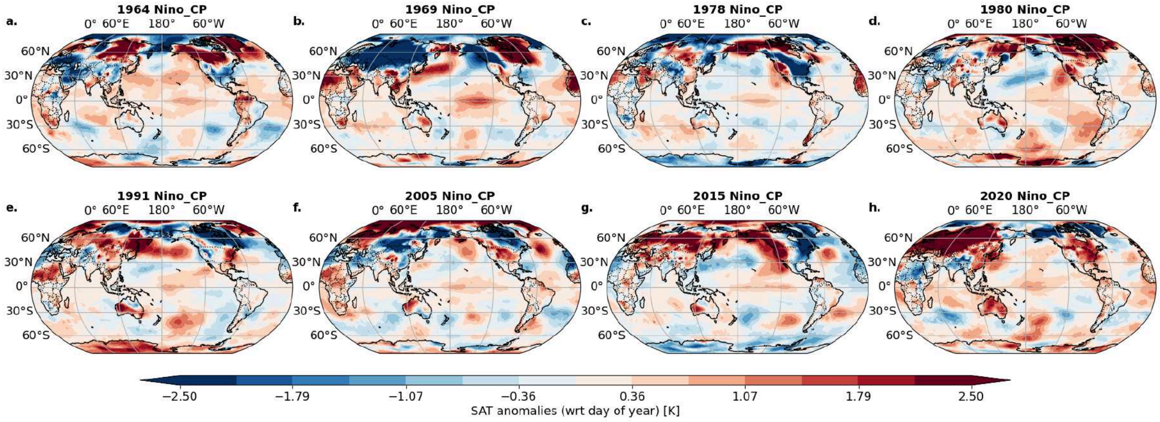


Figure S3. Seasonal DJF averages for all years identified as CP El Niño. We compute seasonal average anomalies for DJF time series separately for each year between 1959 and 2020. The title of each subplot denotes the year of the end of the season in which the event occurred. In total 8 events have been identified as CP El Niño.

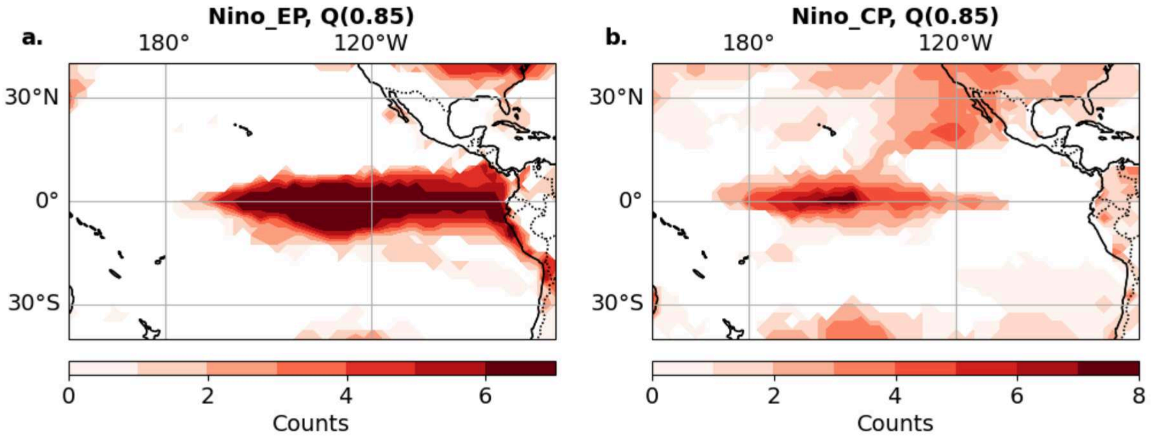


Figure S4. Counts of anomaly center during El Niño. We compute composite anomalies for the DJF time series for all years between 1959 and 2020. The year is indicated in which the event mainly occurred. In total 7 (8) events have been identified as EP (CP) El Niños. The colorbars indicate the number of counts at a spatial location. The contour was chosen according to the $Q(0.9)$ quantile of the spatial map values for each year separately.

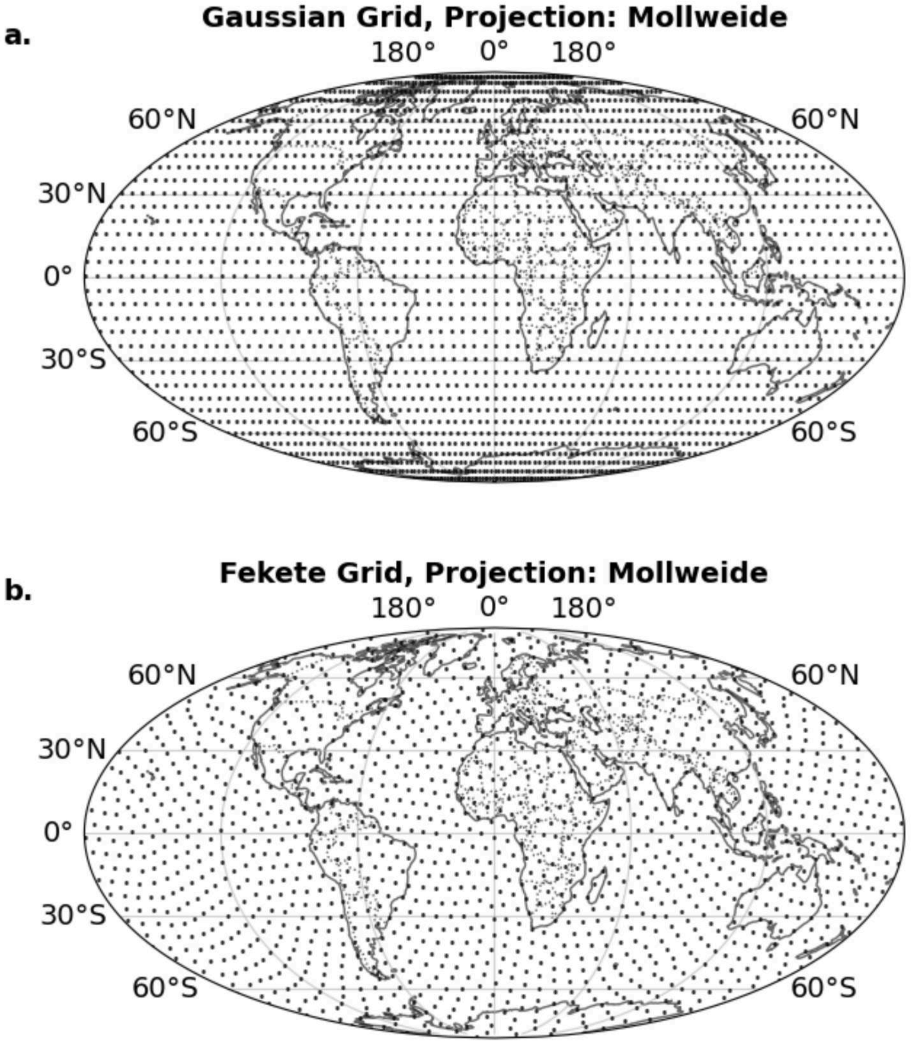


Figure S5. Gaussian Grid and Fekete Grid. We show the effect of the Fekete algorithm. In **a** displayed is the rectangular Gaussian grid. Note that at the poles the points are much closer to each other than at the equator. In **b** displayed is the Fekete grid with approximately uniformly distributed grid points. Neighboring points of both grids have at the equator the same spatial distance of 5°. Note that only for visual reasons, here we employ a grid resolution of 5° instead of the 2.5° resolution used in the analysis of this work. The Mollweide projection was used to minimize the distortion bias of a chosen projection.

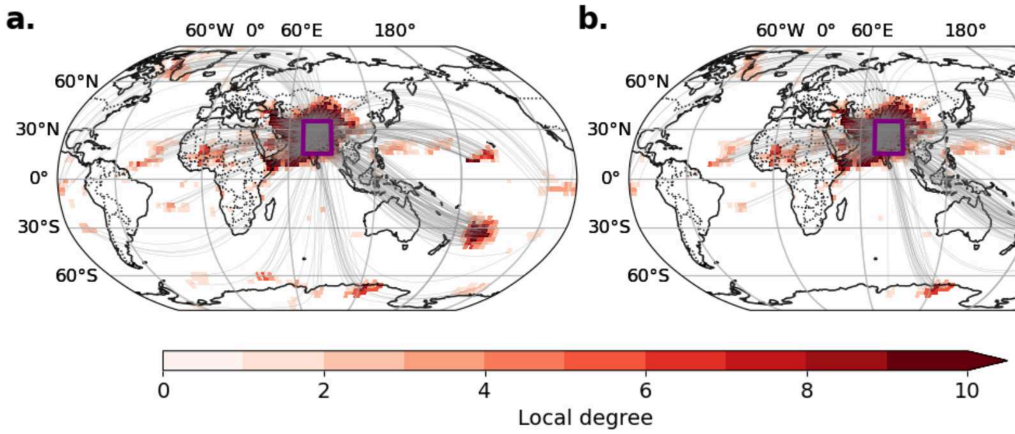


Figure S6. Link bundle correction to climate networks. We show a network for CP conditions. The network is further corrected by a Gaussian KDE to avoid spurious links that occur out of random coincidence. We analyze an example for links of the location of the Indian Subcontinent (70° E-90° E, 15° N-35° N, marked by the violet rectangle). **a** shows the uncorrected network, **b** shows the corrected network. For visual reasons, only every 3rd link is plotted.

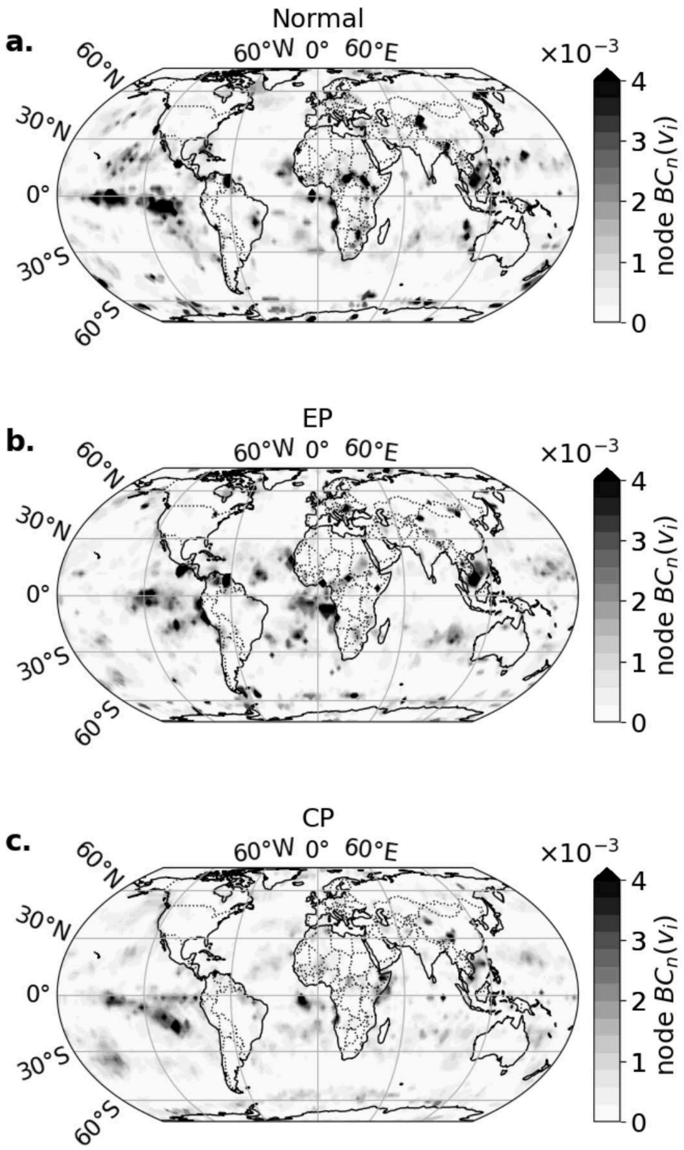


Figure S7. Node Betweenness Centrality for normal, EP and CP conditions. The betweenness centrality is computed using Eq. 2. The climate networks obtained for normal (**a**), on Eastern Pacific El Niño (**b**) and on Central Pacific El Niño (**c**) conditions.

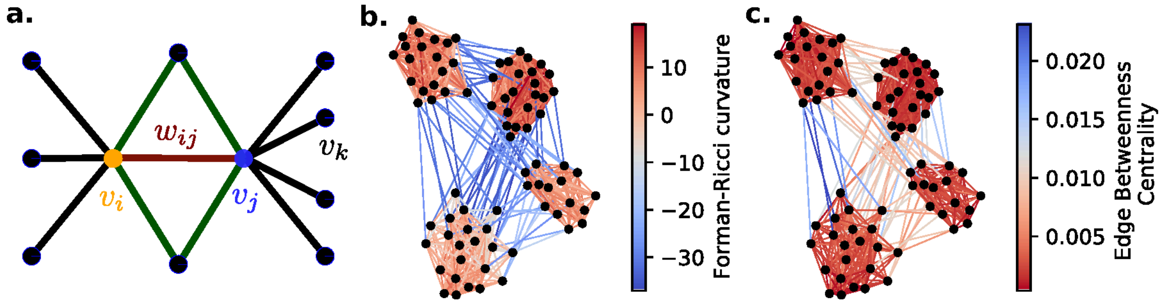


Figure S8. Curvature of complex networks. Forman-curvature of an edge with weight w_{ij} connecting nodes v_i and v_j is obtained by Eq. 4 (**a**). In the synthetic network constructed with a Stochastic Block Model, Forman-curvature (**b**) provides values to distinguish within-community links (red) from between-community links (blue) in contrast to betweenness centrality (**c**) with only a few links with high scores. The network with four communities is visualized using the spring layout from NetworkX (Hagberg et al., 2008).

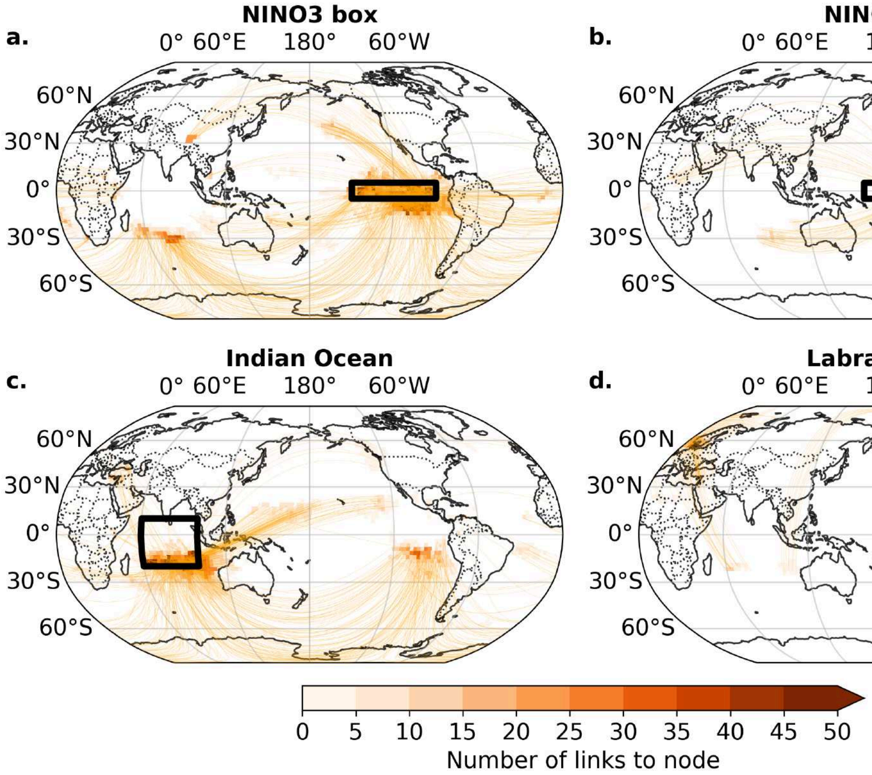


Figure S9. Teleconnections from eastern and central Pacific Ocean, IO, and the Labrador Sea. F^-_{ij}, i.e. teleconnections, for normal years (orange links) connected to the four selected regions (black rectangles): Niño 3 box (**a**), Niño 4 box (**b**), Indian Ocean (**c**), and Labrador Sea (**d**). For visual clarity, only a third of all links are shown.

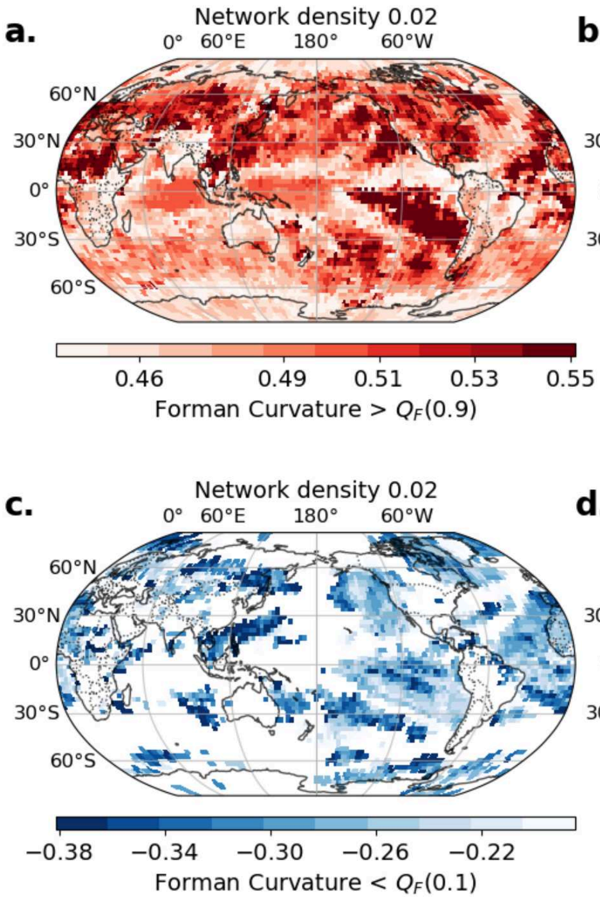


Figure S10. Sensitivity to network density for CP conditions. We analyze the sensitivity to different network densities. The first row shows results for the positive node curvature f_i^+ for **a** the CP network with a density of 0.02 and **b** a density of 0.05. The second row shows results for the negative node curvature f_i^-, for the network with density of 0.02 (**c**) and density of 0.05 (**d**).

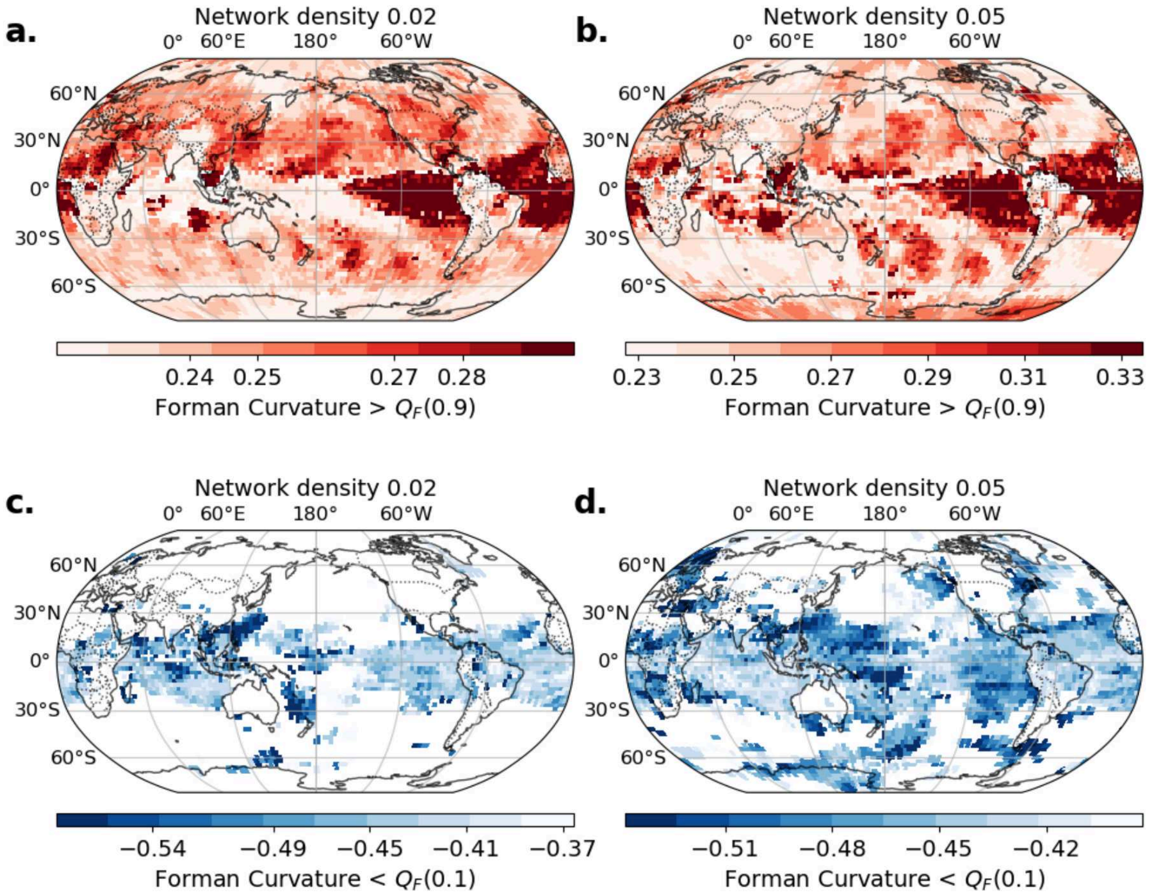


Figure S11. Sensitivity to network density for EP conditions. We analyze the sensitivity to different network densities. The first row shows results for the positive node curvature f_i^+ for **a** the EP network with a density of 0.02 and **b** a density of 0.05. The second row shows results for the negative node curvature f_{ij}^-, for the network with density of 0.02 (**c**) and density of 0.05 (**d**).

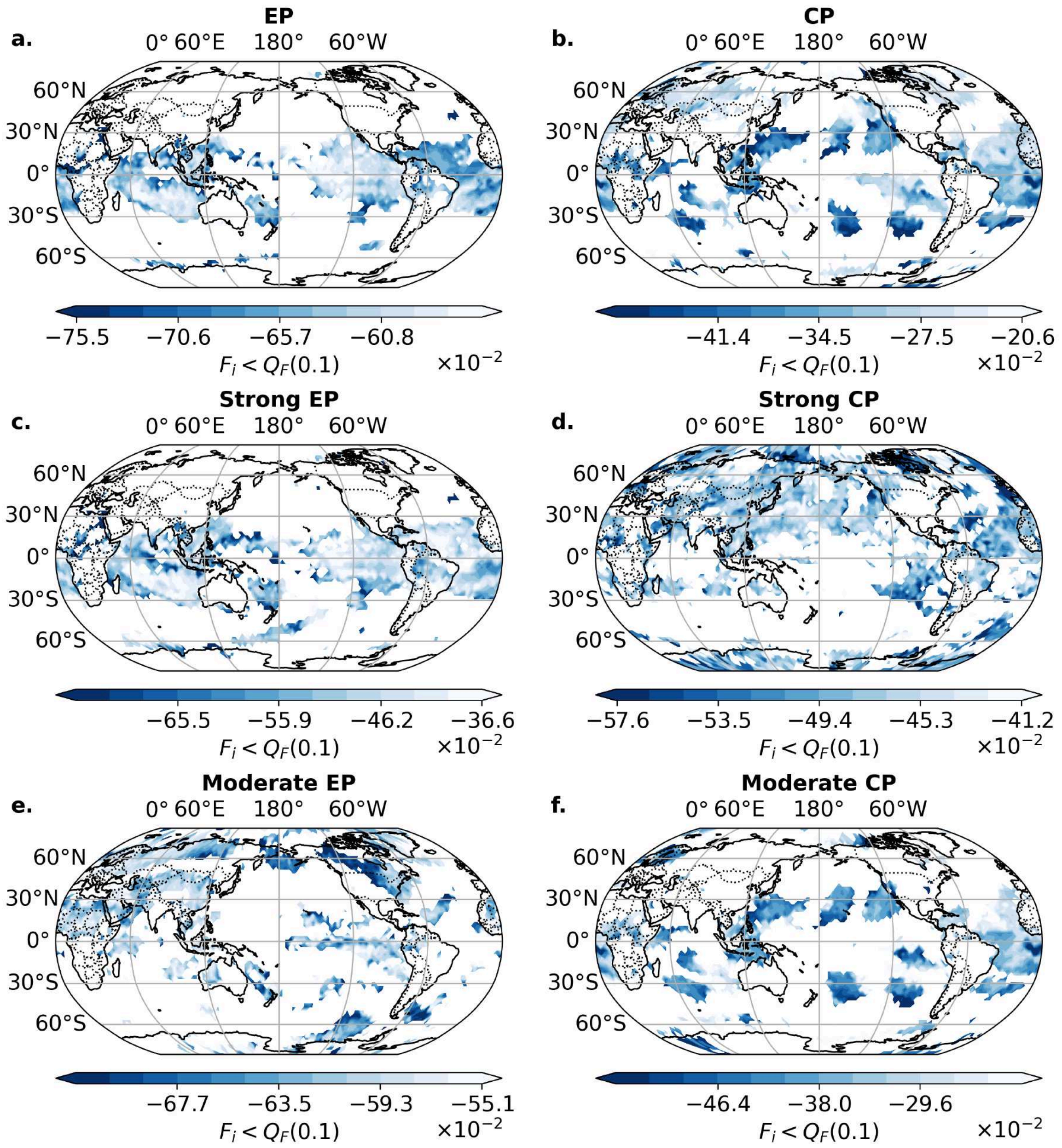


Figure S12. Sensitivity strong and moderate El Niños. We analyze the sensitivity to differently strong El Niños for negative node curvature f_i^-. Displayed are **a** all events that make the EP network**b** all events that make the CP network. The second row shows results for strong ElNiños defined as events where N3/N4 is larger than 1 (**c**,**d**). The third row shows results for moderate ElNiños defined as events where N3/N4 is between 0.5 and 1 (**e**,**f**).

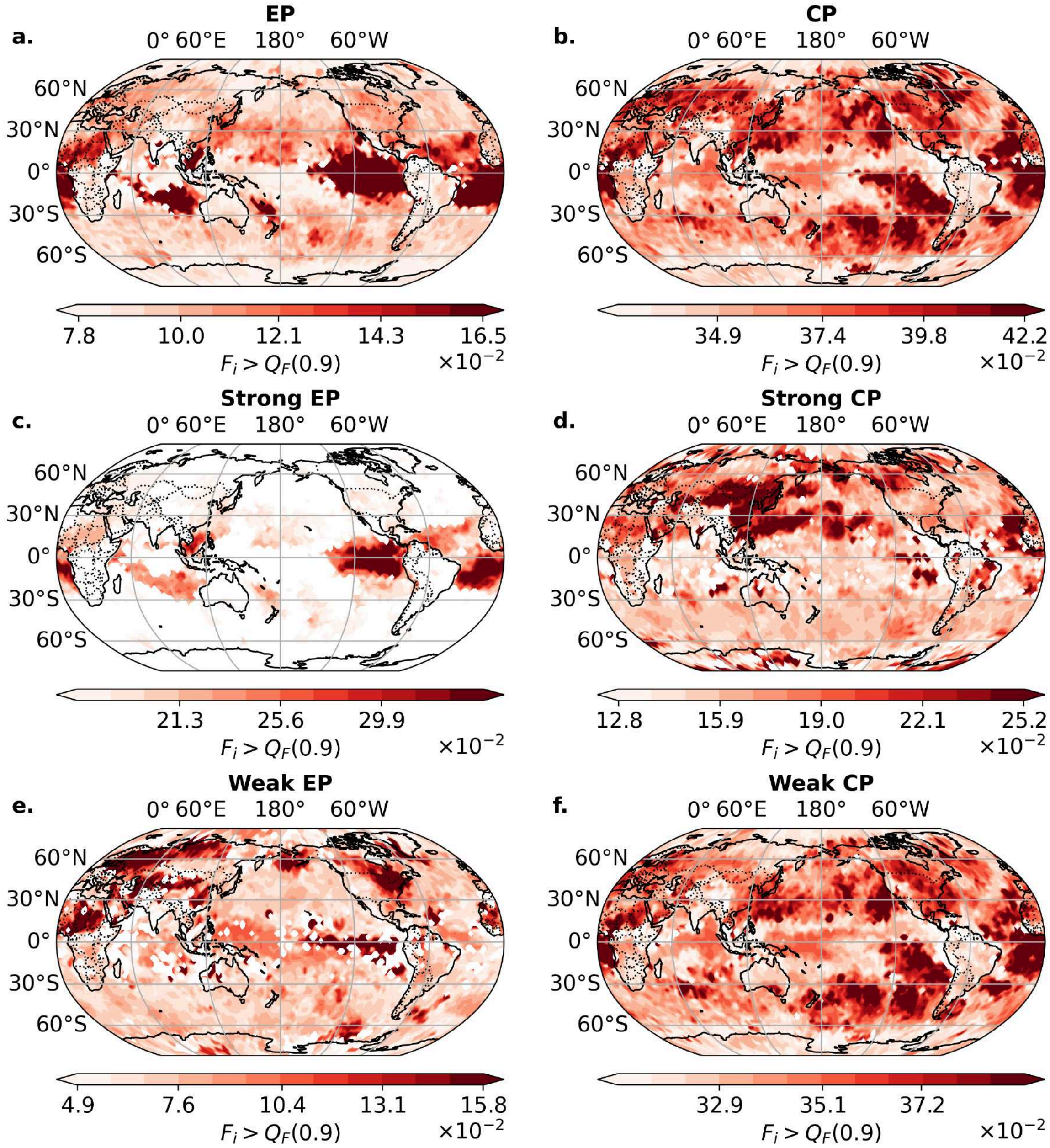


Figure S13. Sensitivity strong and moderate El Niños. We analyze the sensitivity to differently strong El Niños for positive node curvature f_i^+. Displayed are **a** all events that make the EP network**b** all events that make the CP network. The second row shows results for strong ElNiños defined as events where N3/N4 is larger than 1 (**c**,**d**). The third row shows results for moderate ElNiños defined as events where N3/N4 is between 0.5 and 1 (**e**,**f**).

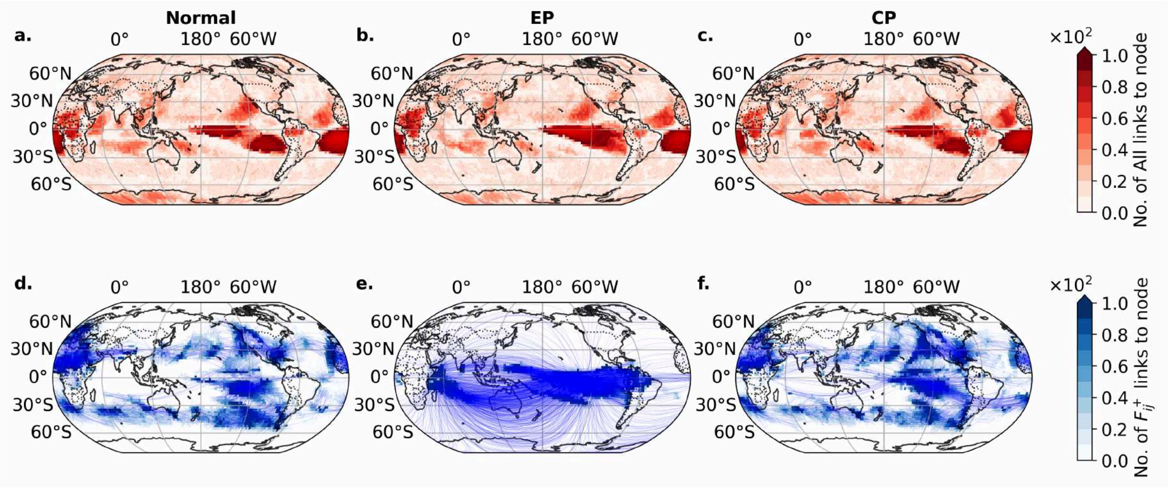


Figure S14. Same as Fig. 2 main text but for daily SAT data of the pre-industrial control run of the UKESM1-0-LL model.

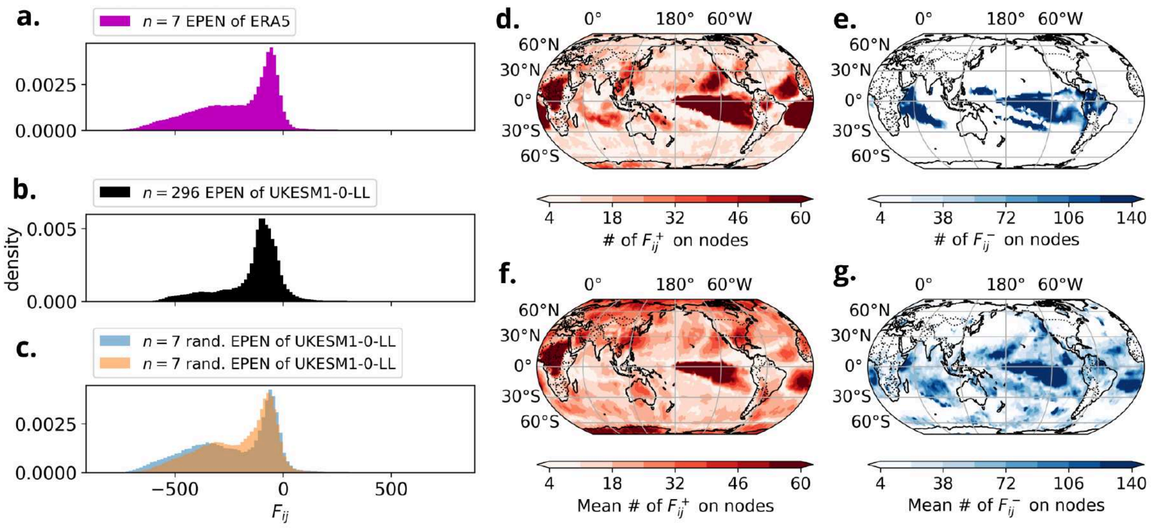


Figure S15. Effect of the finite amount of EP El Niño events. The distribution of curvature obtained from all 296 EP El Niño events in the pre-industrial control run of the UKESM1-0-LL model **(b)** is compared to the distributions obtained from 7 randomly sampled EP events from the pre-industrial control run (c) and the distribution obtained from EP events in ERA5 (a). The number of links with positive and negative curvature on each node is shown for the network constructed from 296 EP El Niño events of the pre-industrial control run **(d, e)**. The average number of positively/negatively curved links on each node over the networks constructed from 7 randomly sampled events are displayed in **(f/g)**,

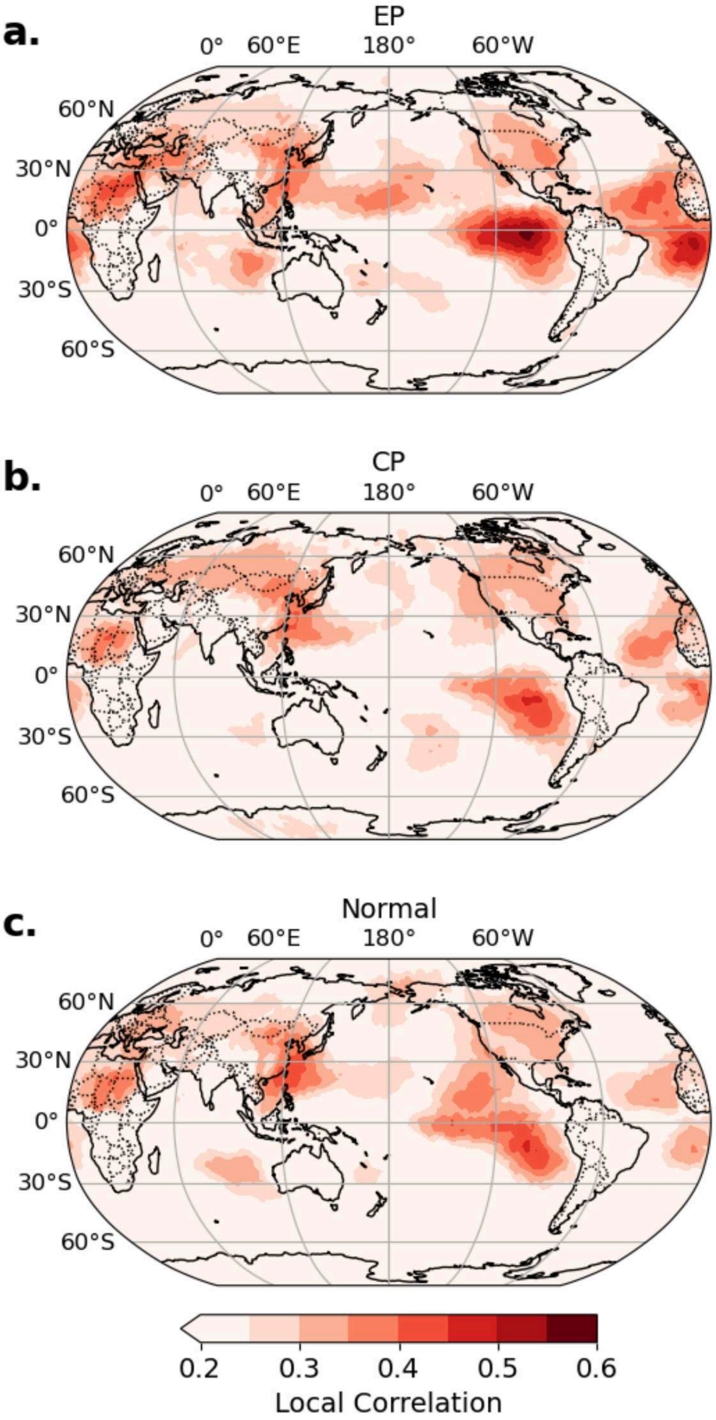


Figure S16. Local Correlation Analysis of ENSO conditions. We analyze the local correlation of nodes for the different ENSO conditions. The local correlation (equation 7) takes the $Nb = 200$ next neighbor points and uses Spearman's rank correlation. Displayed are local correlations for (**a**) EP (**b**) CP and (**c**) normal year DJF conditions.

Supplementary Material for Chapter 5

B.

Supporting Material to Propagation pathways of Indo-Pacific rainfall extremes are modulated by Pacific sea surface temperatures

Felix M. Strnad^{1,*}**, Jakob Schloer**^{1}**, Ruth Geen**^{2}**, Niklas Boers**^{3,4,5}**, and Bedartha Goswami**^{1,*}

^{1}Cluster of Excellence Machine Learning: New Perspectives for Science, University of Tübingen, Tübingen, Germany
^{2}School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK
^{3}School of Engineering & Design, Earth System Modelling, Technical University Munich, Munich, Germany
^{4}Potsdam Institute for Climate Impact Research, Potsdam, Germany
^{5}Department of Mathematics and Global Systems Institute, University of Exeter, Exeter, UK
*Corr. authors: felix.strnad@uni-tuebingen.de

ABSTRACT

Contents of this file: Text Supplementary Note 1 to Supplementary Note 14; Figures S1 to S41

Introduction

In this Supplementary Material to our article, we describe in detail the definition of Extreme Rainfall Events (EREs) in the Indo-Asia Pacific domain (Text Supplementary Note 1, Fig. S1). The schematic of the Event Synchronization approach is discussed in Sec. Supplementary Note 2 and Fig. S2. We show one schematic to exemplify the climate network approach (Fig. S3). A detailed analysis of the regions of synchronous EREs, detected by a community detection approach is given in Supplementary Note 3. Here, we discuss as well the synchronous rainfall index estimation per community (Fig. S4,S5) as well as its lead-lag behavior (Fig. S6). We also analyzed the synchronous rainfall indices with respect to their periodicity (Fig. S7). Also a Hovmöller diagram for initiation time points in the equatorial Indian Ocean (EIO) are shown in Fig. S8. Intra-annual distribution of occurrence of synchronous EREs is explored in section Supplementary Note 4 (Fig. S9, S10). In Sec. Supplementary Note 5, we show one example for the canonical BSISO propagation (Fig. S11, S12), as well as the spatially resolved propagation of the BSISO for the three discovered BSISO propagation modes (Fig. S13, Fig. S14, Fig. S15). The corresponding propagation of EREs is shown for the three modes in Fig. S16. The robustness of our results with respect to the chosen BSISO index is shown in section Supplementary Note 6, Fig. S17. We demonstrate in section Sec. Supplementary Note 7 that the organization of EREs through the BSISO can also be shown by using a simple Linear Regression Model (Fig. S18, S19). In Sec. Supplementary Note 8, we prove the robustness of our results with respect to the underlying dataset

(Fig. S20, S21) and the chosen community detection algorithm (Fig. S22, S23).) We provide further complex network analyses in Sec. Supplementary Note 9. Different network measures are analyzed globally (Fig. S24-S26) as well as community-specific (Fig. S27-S30). We provide a further discussion on the MJO similarity in Sec. Supplementary Note 10 (Fig. S31). More detailed information to the connection the El Niño Southern Oscillation (ENSO) is given in Sec. Supplementary Note 11 (Fig. S32, S33). We also show classical conditions for different ENSO types in boreal summer (Fig. S34). We also visualize the effect of the ENSO condition on the vertical velocities over the Maritime Continent for the 3 propagation modes (Fig. S35) To provide the reader with background to the current discussion in the literature, we outline the basic concepts of the moisture mode theory (Sec. Supplementary Note 12, Fig. S36, S37) and the vertical shear mechanism (Sec. Supplementary Note 13, Fig. S38). In the main text, we focused mainly on the BSISO-dominated communities. A detailed discussion on the North India-China region and its connection to current literature is provided in Sec. Supplementary Note 14 (Fig. S39, Fig. S40, Fig. S41).

Supplementary Note 1 Extreme rainfalls in the Indo-Asia Pacific domain

To construct the climate network of EREs in the Indo-Asia Pacific region, we only take into account those grid locations where the 95th percentile value of rainfall is greater than 10 mm/day or if we find more than 10 such 'event days' over the whole time period. Regions around the Arabian Peninsula and the Gobi and Taklamakan deserts were thus excluded from the analysis (compare Figure S1). We find that number of extreme events varies significantly between the tropics and the subtropics, with the tropics experiencing a far higher number of events (Figure S1). Our null model (sec. Methods) which depends on the number of events in each time series, helps to reduce the bias introduced in the event synchronization measure due to different event rates[1]. The upper bound on the dynamical delay is set to $\tau_{\max} = 10$ days for all grid location pairs.

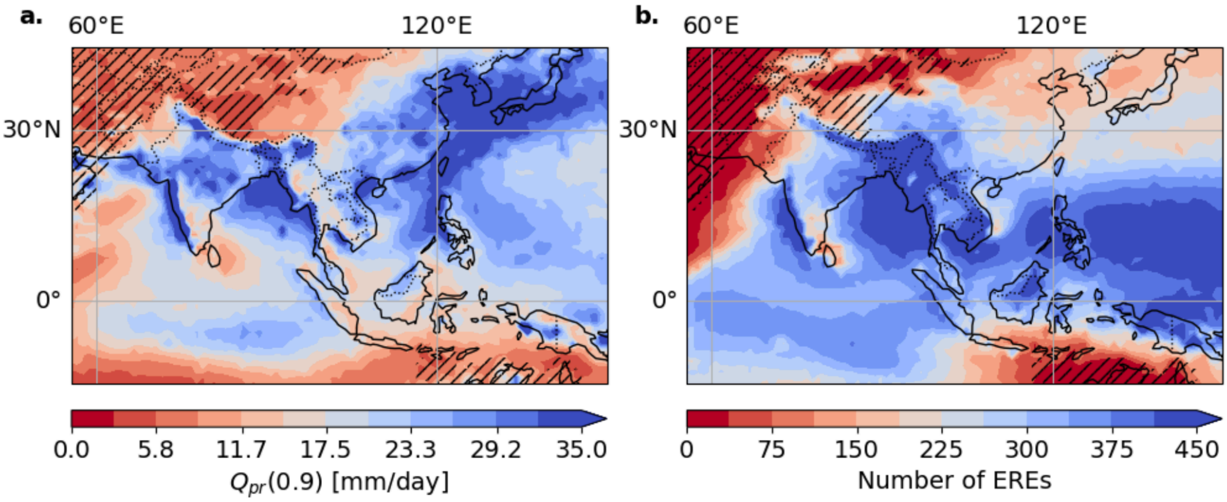


Figure S1. Quantiles and total numbers of extreme-rainfall events. a yields the number of counted EREs per spatial location. **b** shows the respective values of the quantile local function $Q_{0.9}$ for each location. Data is used from the MSWEP dataset[2] over the period from 1979–2021. An Extreme Rainfall Event (ERE) is defined as a day with a rainfall sum more than the 90th percentile of all wet days (i.e. days with rainfall more than 1 mm/day. Locations with less than 10 events in total are excluded from the analysis and marked as hatched areas.

Supplementary Note 2 Event Synchronization and network construction

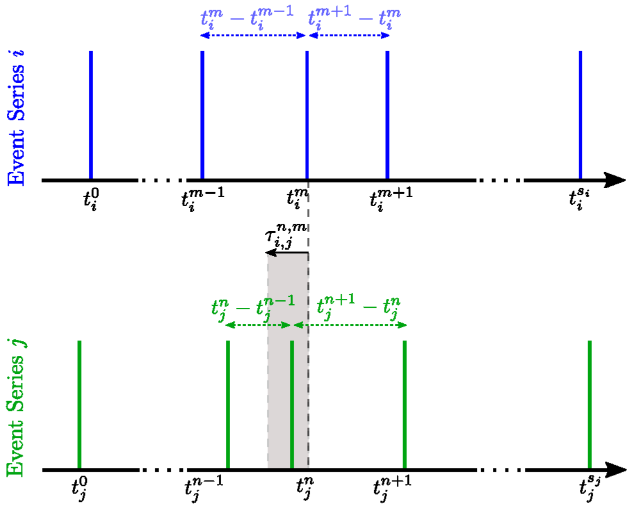


Figure S2. Scheme of time-delayed Event Synchronization. The point in time t_i^m is identified as synchronous to t_j^n since the time difference between t_i^m and t_j^n is within the allowed the range given by $\tau_{ij}^{m,n}$ according to the definition expressed in equation (see Methods). Not shown here is the maximum delay $\tau_{\max}$.

Fig. S2 describes the ES scheme. Note that if two or more events in one series occur at subsequent time steps the dynamical delay $\tau_{i,j}^{m,n}$ will result in a value of 1/2, leading to a case where it is likely that two sequentially occurring events are not recognized as synchronous. Therefore, in our analysis, blocks of consecutive events are counted as one event, placed on the point in time of the first event.

Supplementary Note 2.1 Event Synchronization based network construction

The construction of the network in our study is based on the event synchronization technique, which enables the estimation of statistically significant synchrony between events across different time series. This technique allows us to quantify the point-wise similarity of extreme event time series.

One key element is the adjacency matrix A, which holds significant importance in describing the network mathematically. The adjacency matrix is a square matrix of size $N \times N$, where N represents the number of spatial locations, equivalent to the number of time series under consideration. In this matrix, if $A_{i,j} = 1$, it indicates that events at location i are statistically significantly followed by events at location j. To aid in the interpretation of the adjacency matrix and provide a more intuitive understanding, we provide a schematic of the network construction in Fig. S3.

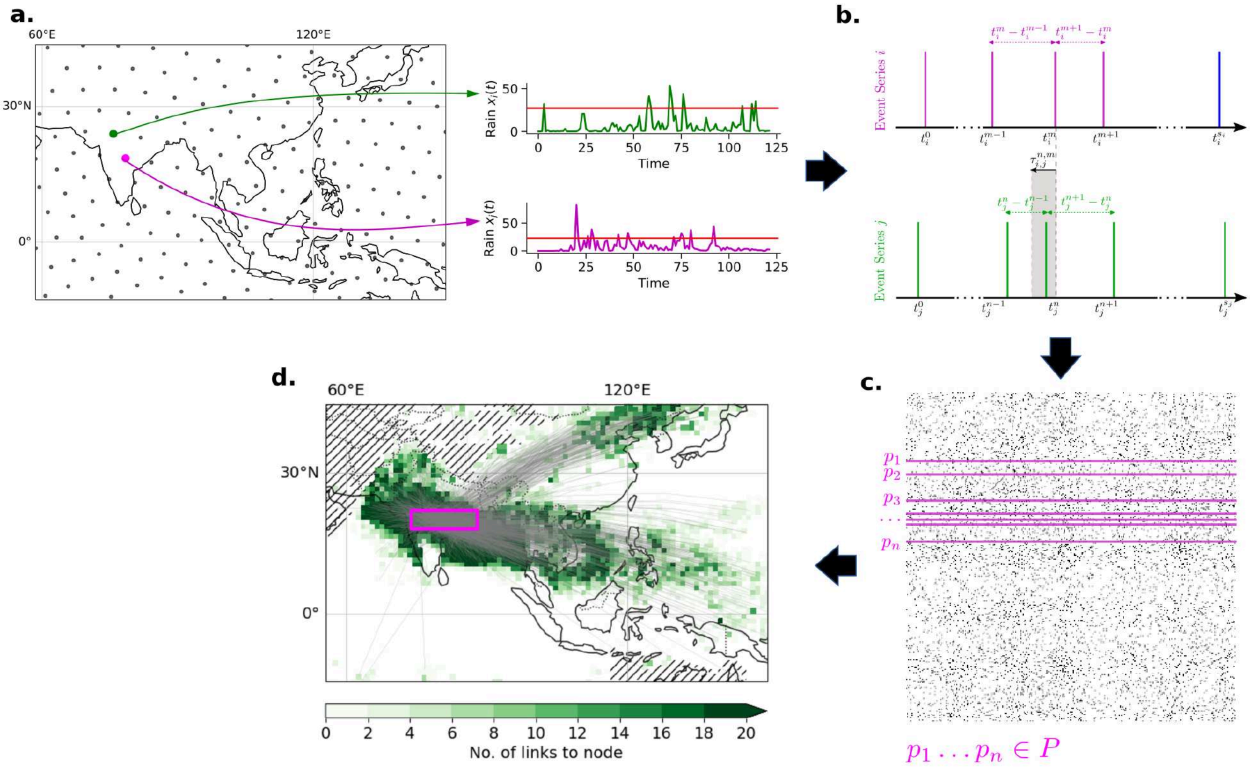


Figure S3. Scheme of network construction and interpretation using event synchronization. a Time series of precipitation data. These translate to event time series. **b** The point in time t_i^m is identified as synchronous to t_j^n since the time difference between t_i^m and t_j^n is within the allowed range given by $\tau_{ij}^{m,n}$ according to the definition (see Methods). **c** The network is constructed by point-to-point comparison of all time series to each other and mathematically expressed by the $N \times N$ adjacency matrix A, where N denotes the number of locations. **d** The highlighted rows of the adjacency matrix A translate to a set of points P (pink rectangle) that represent spatial locations that are statistically significantly synchronous to further locations (the black dots in the respective row) to network links.

Supplementary Note 3 BSISO drives the organization of synchronous EREs

We compute the membership likelihood for the 6 identified monsoon regions in Fig. 2 a. These are computed as it is described in Sec. . The membership likelihoods for the different climate network communities are visualized in Fig. S4. These plots show low spatial variability of the communities and therefore underline the robustness of the community detection algorithm.

The BSISO is directly reflected in the shape of the communities. The BSISO is known to emerge around the center of the Equatorial Indian Ocean characterized by intense rainfalls[3,4] consistent with our observation of an increased likelihood of synchronous EREs in the EIO community for BSISO phases 1 and 2 (Fig. 3 a)[5]. The BSISO-driven rainfalls are known to intensify significantly over the Bay of Bengal[6] consistent with our observations (Fig. 3 c). The increased probability of experiencing extreme precipitation during active BSISO by a factor of two to three in the MC community region is consistent with observations in other studies[7]. The strongest enhancement in SA occurs during BSISO phases 4 to 6 (Fig. 3 d) establishing a monsoon trough developing here over the Bay of Bengal and leading to winds blowing southwesterly across the Indian Ocean carrying a lot of moisture. Many studies suggest that the establishment of a such convergence zone along the monsoon trough is enhanced during active BSISO expressed by strong low-pressure systems[6,8–12]. In WP, EREs occur during BSISO phases 6 and 7 (Fig. 3). The intraseasonal variation of the convection anomalies over the tropical eastern Indian Ocean until the western Pacific Oceans, so-called Pacific Japan (JP)-mode, has been attributed to the summertime MJO[13].

For each of these regions we derive a community-specific synchronous rainfall index (SRI) (see Methods, Fig. 1). This index can also be used to determine the most synchronous days per community (Fig. S5). Using the synchronous rainfall index (Fig. S5), it is also possible to derive the power spectrum of each community-specific synchronous rainfall index. We compare it to an Autoregressive Model with lag 1 autocorrelation (AR(1)). We find that the characteristic 30-60 day period of the BSISO is also dominant for all communities except the North-India China community indicating that this community is not dominated by the BSISO (Fig. S7).

We can also show the BSISO propagation using Hovmöller diagrams for the most synchronous days in the EIO community. In Fig. S8, we observe an eastward propagation via the Maritime Continent and a northward propagation. However, the propagation pattern gets blurry around 5-10 days after the initiation suggesting a variability in the propagation pathways.

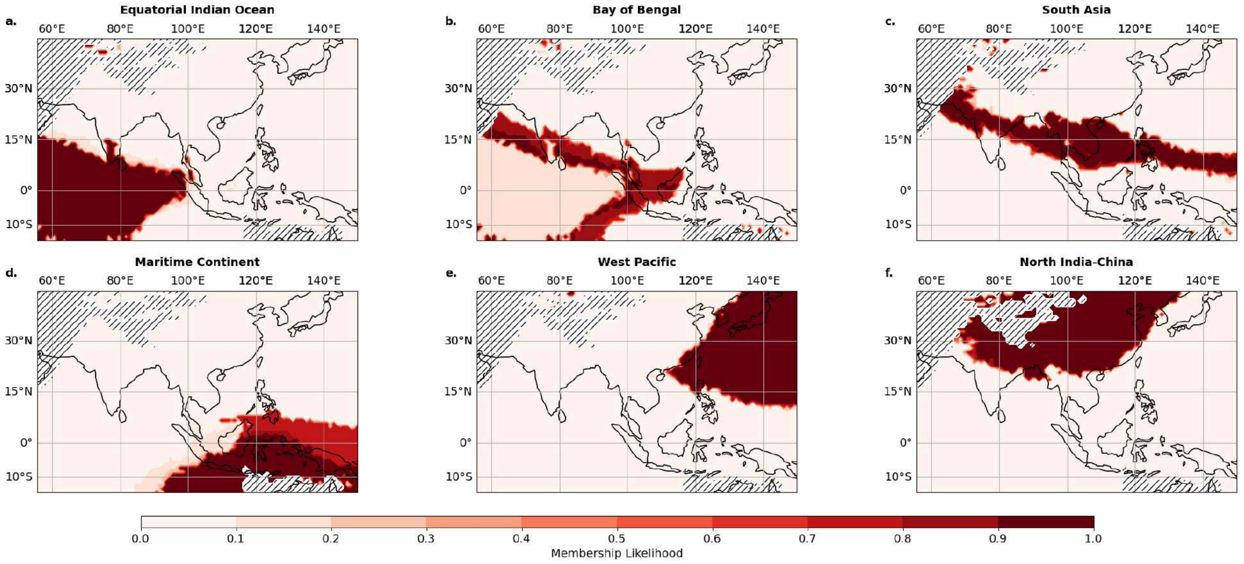


Figure S4. Membership Likelihoods for different communities. Using the heuristic outlined in sec. we find 6 stable communities. The colorbar shows the membership likelihood of a respective community. 100 independent runs of the community detection algorithm have been used for this analysis.

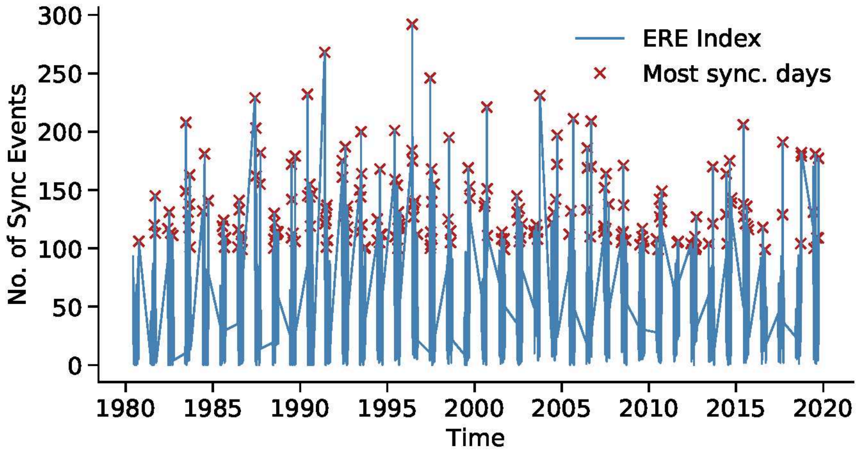


Figure S5. Definition of synchronous extreme rainfall event index. The synchronous extreme rainfall event index (ERE index) (see Methods) is shown here as exemplary for the community in the Equatorial Indian Ocean (Fig. 2). We derive the most synchronous days by taking the peaks of all values above the 90th percentile denoted by red crosses.

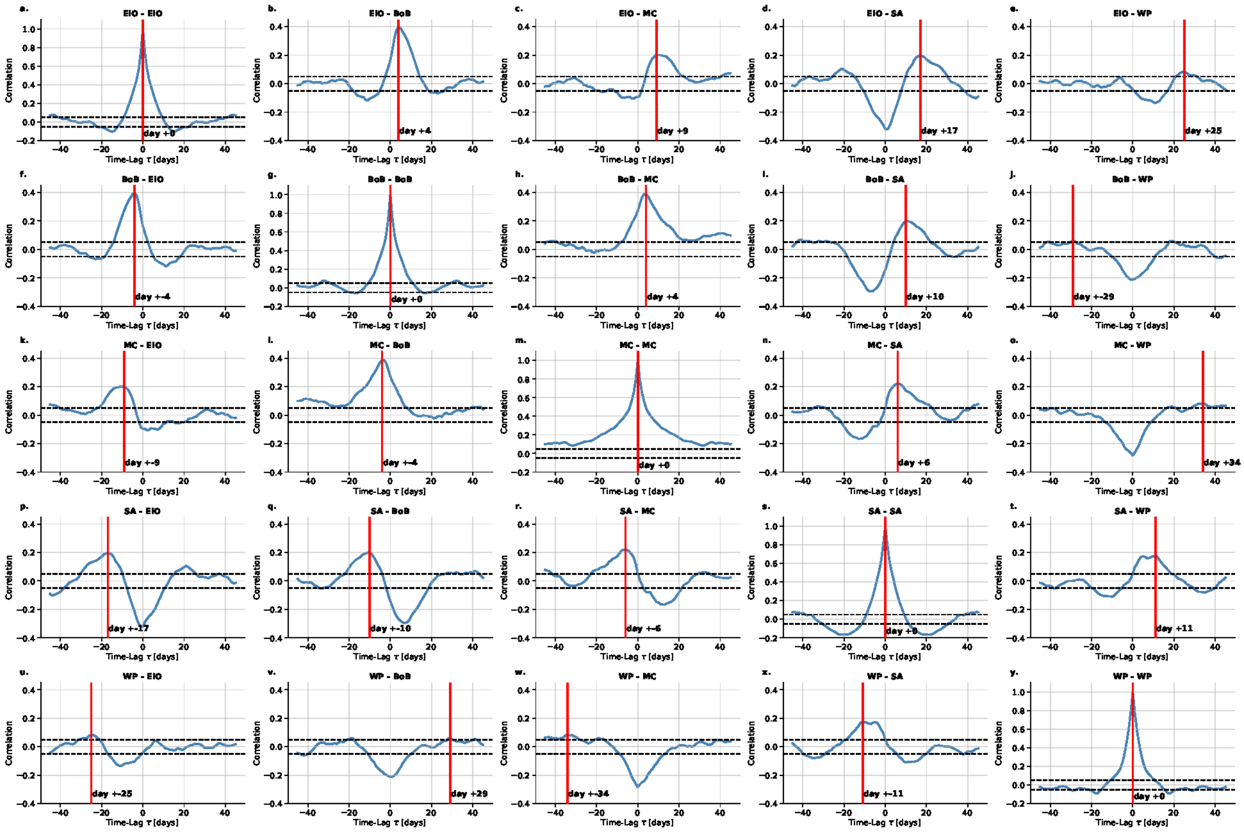


Figure S6. Lead-Lag analysis for different communities. Using the synchronous rainfall index (see Methods) we derive lead-lag correlation analysis based on Spearman's rank order correlation. The dashed line denotes the 0.95 confidence interval. Vertical solid red lines denoting the day of maximum correlation.

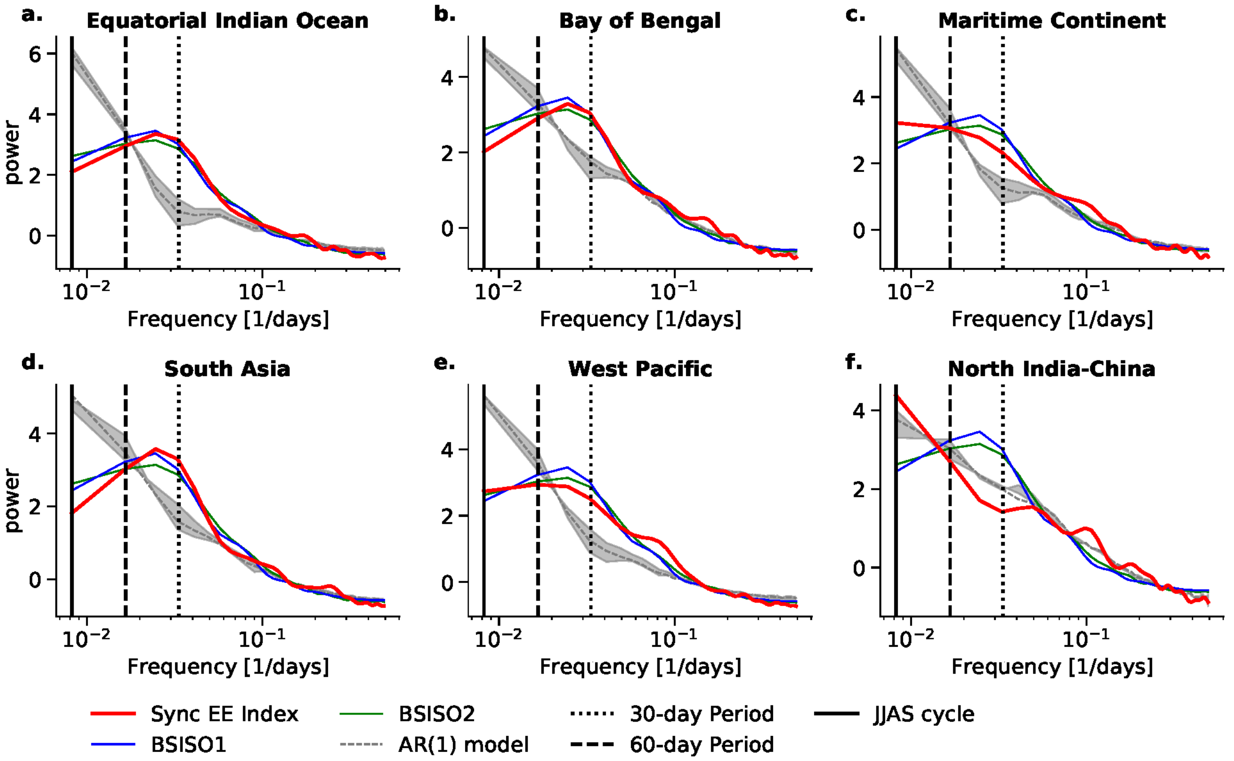


Figure S7. Power spectrum analysis of community-specific synchronous ERE index. Using the synchronous rainfall index (Fig. S5), we derive a power spectrum analysis for each community of its community-specific index (red line). It is compared to samples of an Auto-Regressive model with lag 1 autocorrelation (AR(1)). The dashed line denotes the mean and the grey band the 0.95 confidence interval of the AR(1) model. Vertical lines denoting the JJAS frequency and the dominant 30-60 day period of the BSISO. For comparison also the power spectrum of the BSISO1 (blue line) and BSISO2 (green line) index[14] are shown.

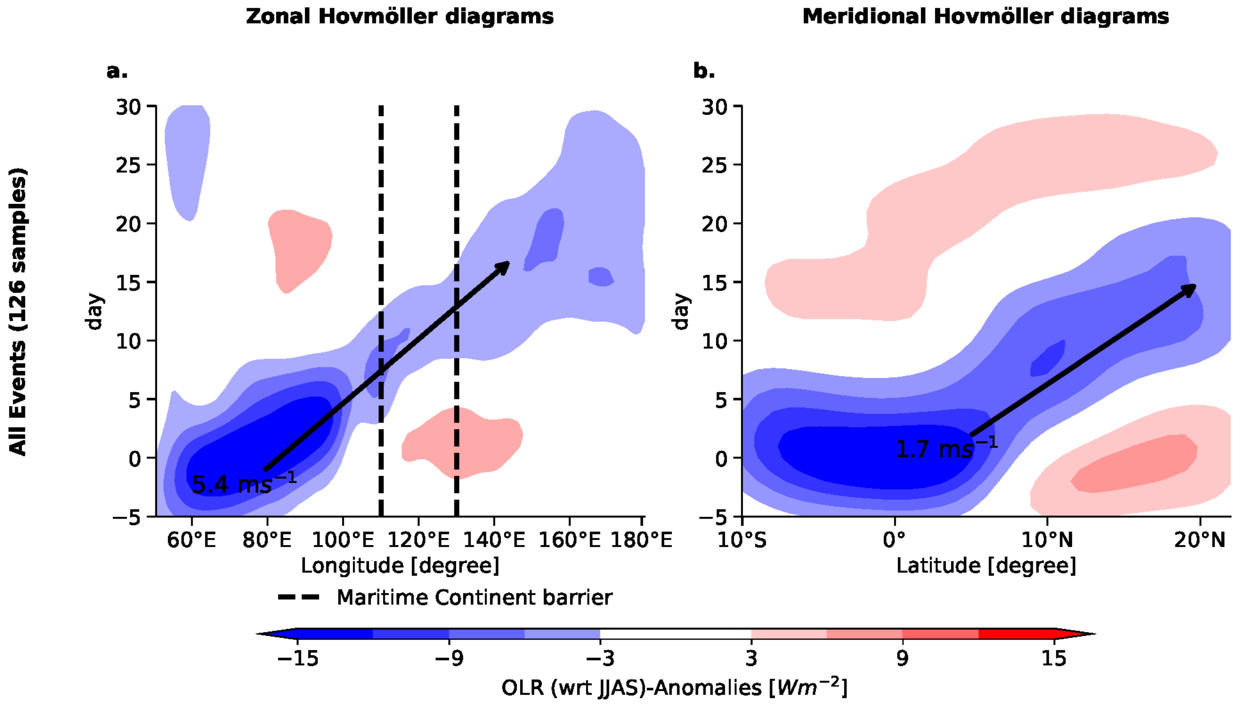


Figure S8. Hovmöller diagrams after initiation. The first column (**a**) shows the composited Hovmöller diagrams in the zonal direction, and the second column (**b**) in the meridional direction. All anomalies are computed with respect to the JJAS seasonality. Day 0 describes the days of maximum synchronization within the EIO community (Fig. 2 a). The dashed lines mark the area of the Maritime Continent barrier roughly estimated to be from 110° E-130° E and the arrows denote the estimated speed of the convective system.

Supplementary Note 4 Intra-annual distribution of occurrence of synchronous EREs

We count the occurrence of EREs for a day of maximum synchronization for the 6 identified monsoon regions in Fig. 2 a. Using the synchronous rainfall index (Fig. S5) per community it is possible to determine the average number of EREs per month (week) for a specific community, shown in Fig. S9 for the monthly distribution and Fig. S10 for the weekly distribution

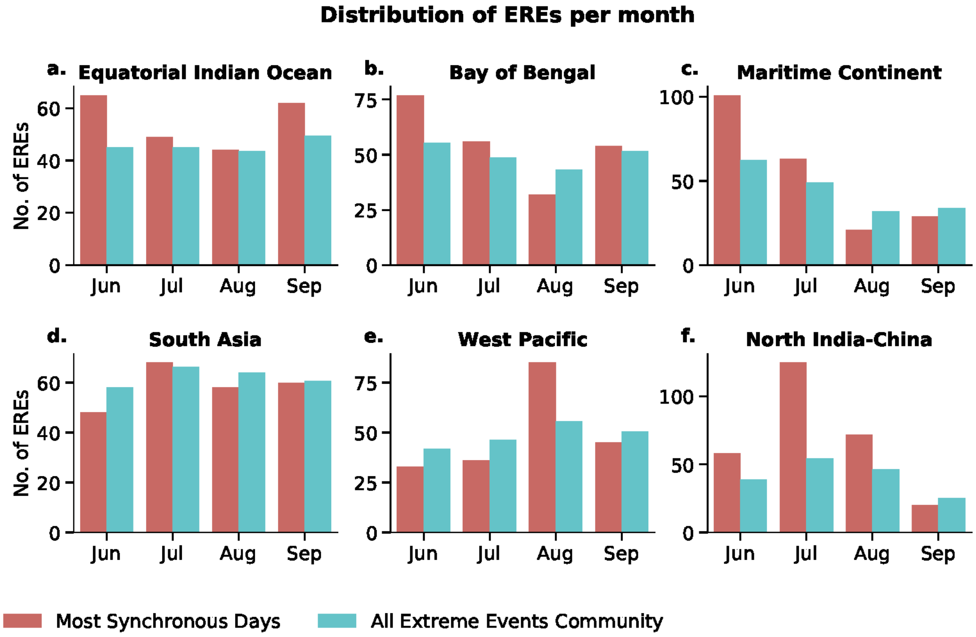


Figure S9. Histogram of the occurrence of EREs per month for different communities. Using the synchronous rainfall index (see Methods) we count the average number of EREs per day of maximum synchronization (red) and for all EREs that are in the community (green) normalized by the number of days.

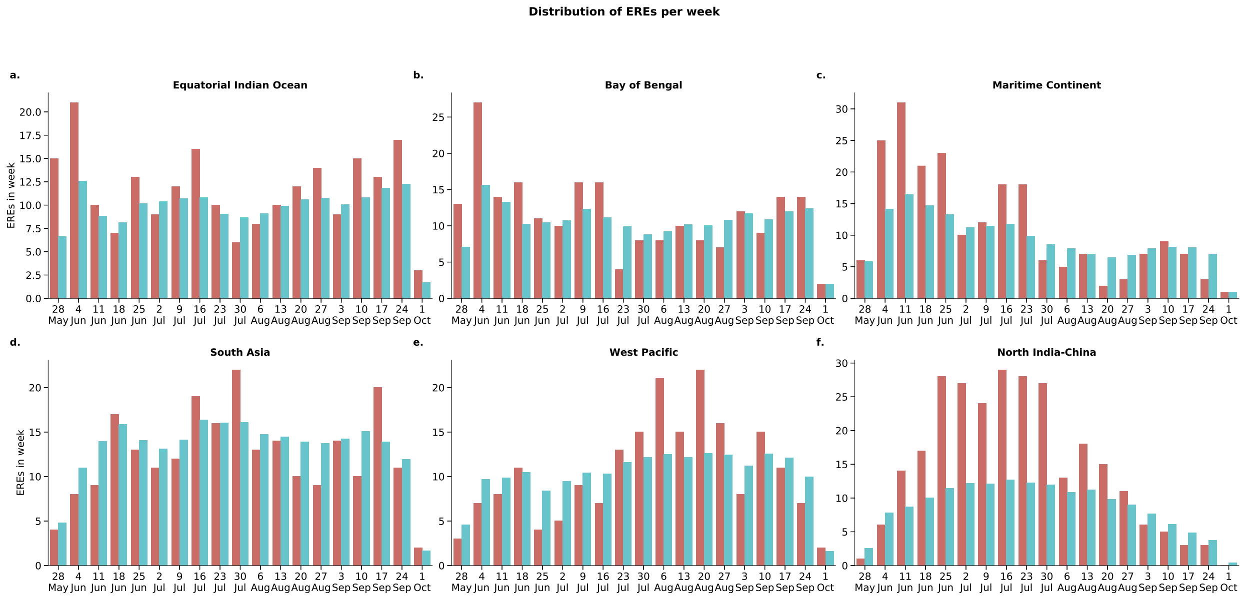


Figure S10. Histogram of the occurrence of EREs per week for different communities. Using the synchronous rainfall index (see Methods) we count the average number of EREs per day of maximum synchronization (red) and for all EREs that are in the community (green).

Supplementary Note 5 Propagation pattern of EREs associated with BSISO

Supplementary Note 5.1 Single events

In this subsection, for illustrative purposes, we investigate the propagation of individual Boreal Summer Intraseasonal Oscillation (BSISO) events across the Indian Ocean, the Maritime Continent, South Asia, and India, extending towards the Western Pacific specifically for the year 1990 to relate the propagation of EREs to the precipitation anomalies. By analyzing this selected year, we aim to provide a comprehensive understanding of the spatiotemporal dynamics and trajectories of BSISO-related events, shedding light on their propagation patterns.

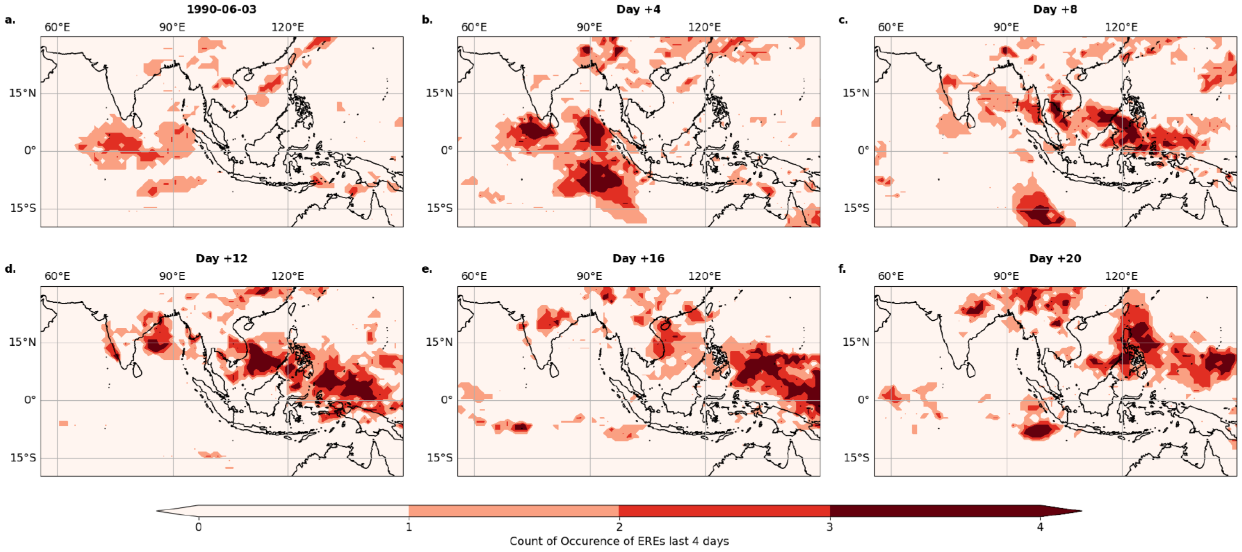


Figure S11. Propagation of Boreal Summer Intraseasonal Oscillation (BSISO) Events in terms of EREs for 1990. EREs are estimated following the methodology described in SI sec. Supplementary Note 1. The spatial evolution of BSISO-related ERE events for the year 1990 is shown. Each plot shows the accumulated sum of the occurrence of EREs in the last 4 days.

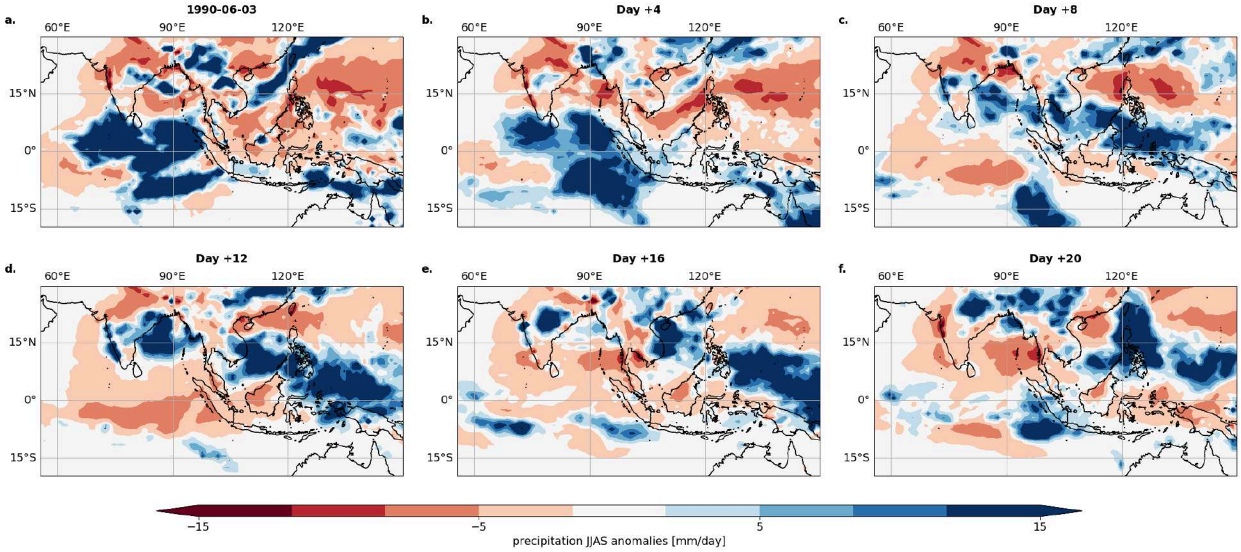


Figure S12. **Propagation of Boreal Summer Intraseasonal Oscillation (BSISO) Events in terms of precipitation anomalies for 1990.** The spatial evolution of a BSISO-related event of the year 1990 is shown here in terms of precipitation anomalies with respect to the JJAS climatology. The points in time are the same as in Fig. S11

Supplementary Note 5.2 BSISO propagation modes

The propagation of the BSISO is subdivided into 3 different modes. The propagation of the convective system is shown in a condensed way in Fig. 9. Here, we show spatial composites of all three propagation modes in their spatial progression until 24 days in advance. The Canonical propagation is shown in Fig. S13, the Eastward Blocked mode in Fig. S14, and the Stationary case in Fig. S15.

These propagation pathways also translate to propagation of EREs along the previously identified communities (Fig. 2 a). The Canonical propagation is shown in Fig. S16.

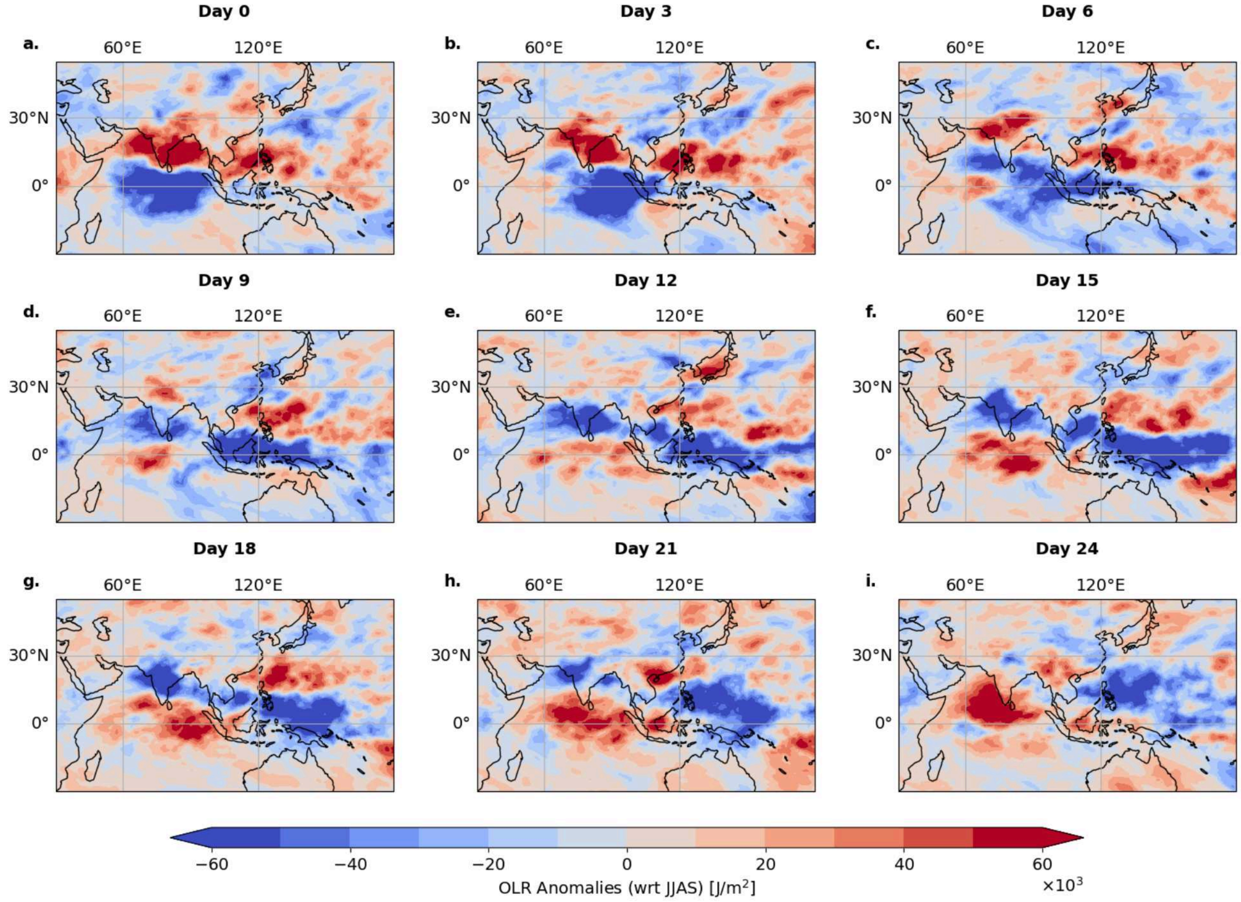


Figure S13. Propagation of anomaly precipitation for Canonical BSISO propagation mode. The rainfall anomalies are computed according to the JJAS climatology. Day 0 denotes the days of maximum synchronization in the EIO community (Fig. 2 a) classified as "Canonical" BSISO propagation mode (Fig. 4 2nd row). The following plots **b**-**i** are the composites of the respective days after day 0 in steps of 3 days.

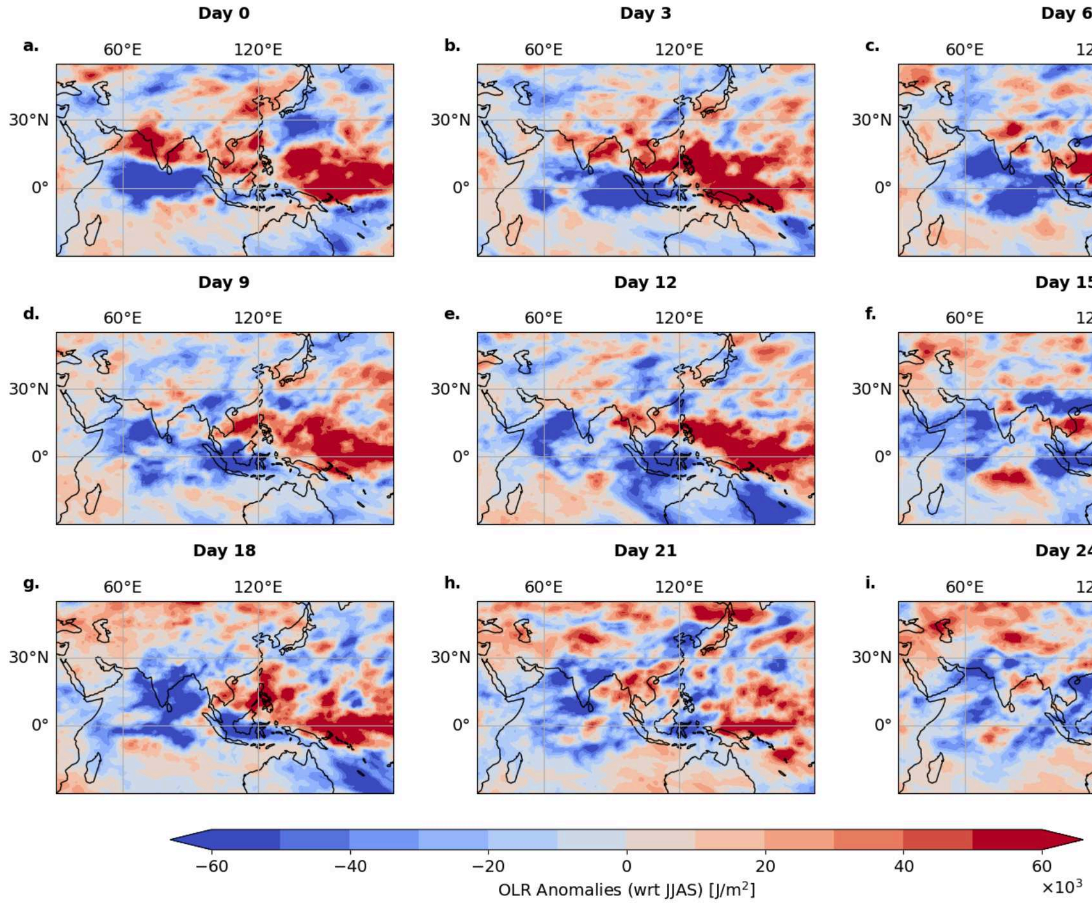


Figure S14. Propagation of anomaly precipitation for Eastward Blocked BSISO propagation mode. The rainfall anomalies are computed according to the JJAS climatology. Day 0 denotes the days of maximum synchronization in the EIO community (Fig. 2 a) classified as "Eastward Blocked" BSISO propagation mode (Fig. 4 2nd row). The following plots **b**-**i** are the composites of the respective days after day 0 in steps of 3 days.

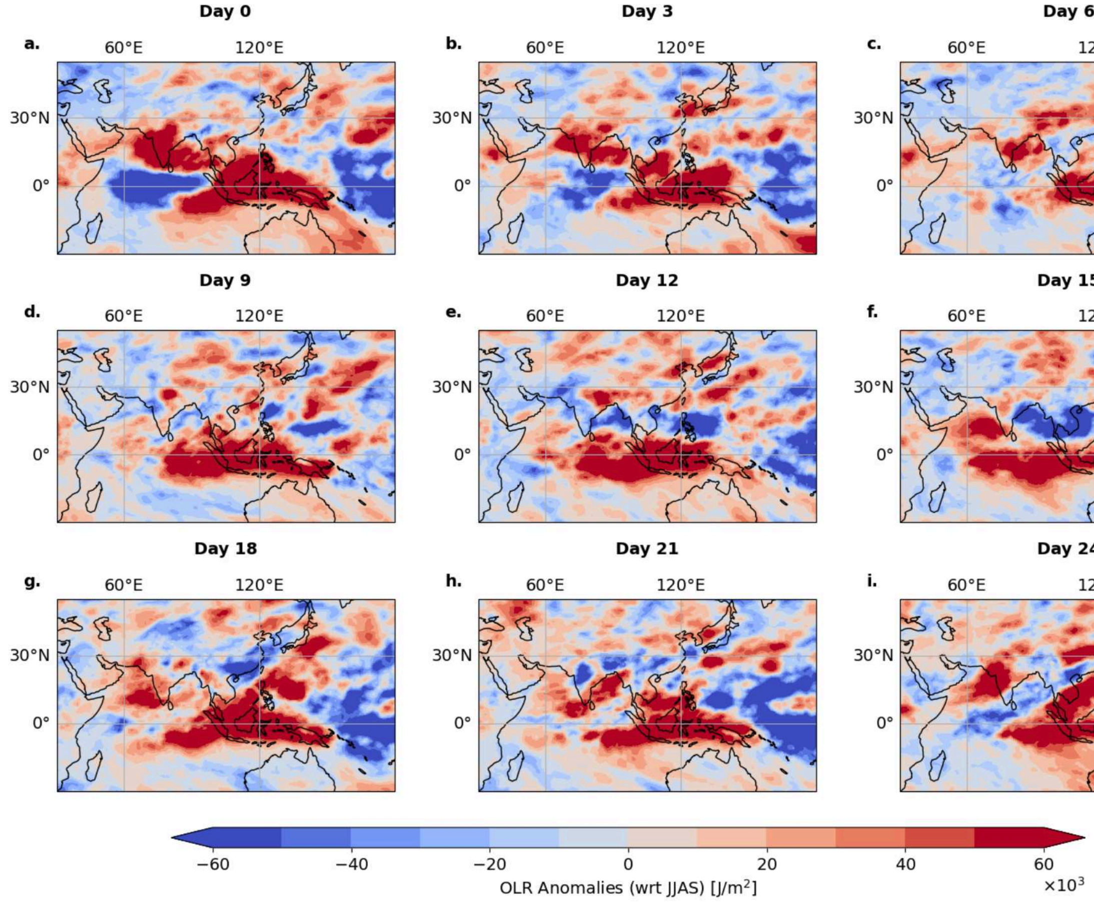


Figure S15. Propagation of anomaly precipitation for Stationary BSISO propagation mode. The rainfall anomalies are computed according to the JJAS climatology. Day 0 denotes the days of maximum synchronization in the EIO community (Fig. 2 a) classified as "Quasi-stationary" BSISO propagation mode (Fig. 4 3rd row). Rainfall anomalies occur solely in the region of the EIO community.

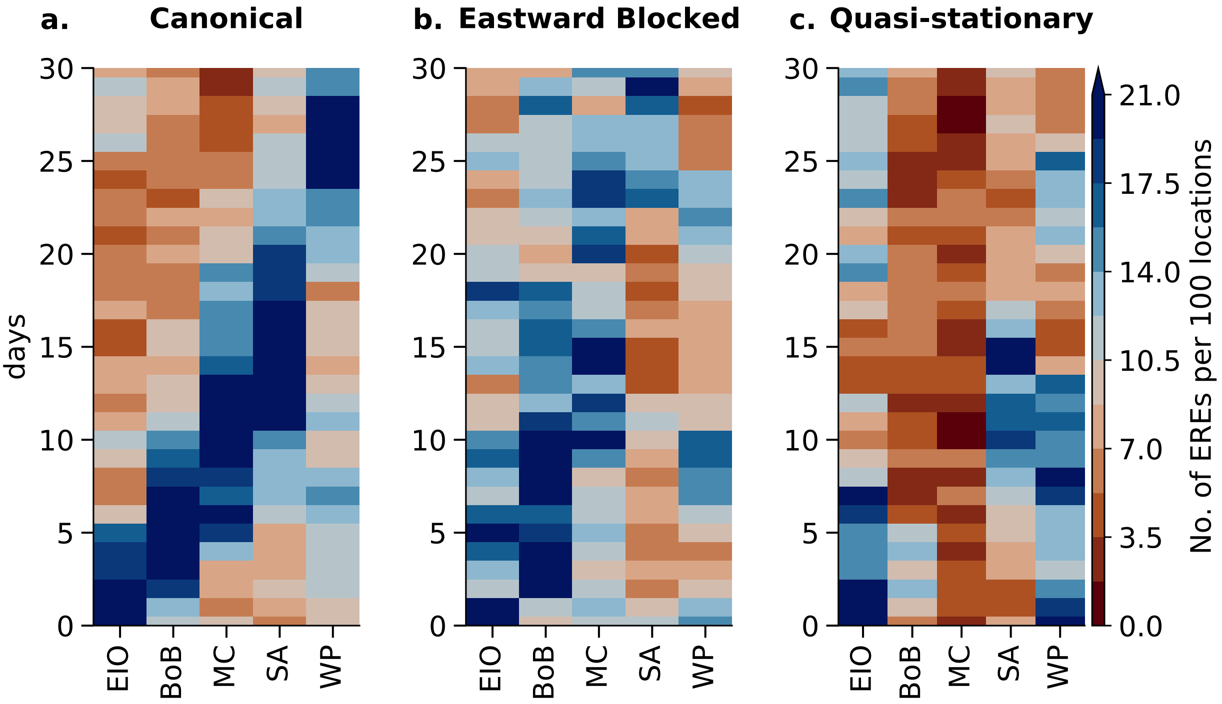


Figure S16. Propagation of EREs along communities. The rainfall extremes are computed according to section Supplementary Note 1. Day 0 denotes the days of maximum synchronization in the EIO community (Fig. 2 a). **a** shows BSISO propagation events classified as “Canonical”, **b** events classified as “Eastward blocked” and **c** events classified as “Quasi-stationary”.

Supplementary Note 6 Influence of the choice of BSISO indices

There are multiple indices that describe the characteristics of the BSISO propagation. While the index defined by[14] which is used in this study is better suited for tracking BSISO convection[15], the RMM index[16] and the BSISO index by[17] is better describing the circulation associated with BSISO when convection is reduced[15]. Here, we show that our qualitative results from Fig. 3 remain unaffected by the choice of the specific BSISO index. Fig. S17 shows the distribution over phases using the index described by[17]

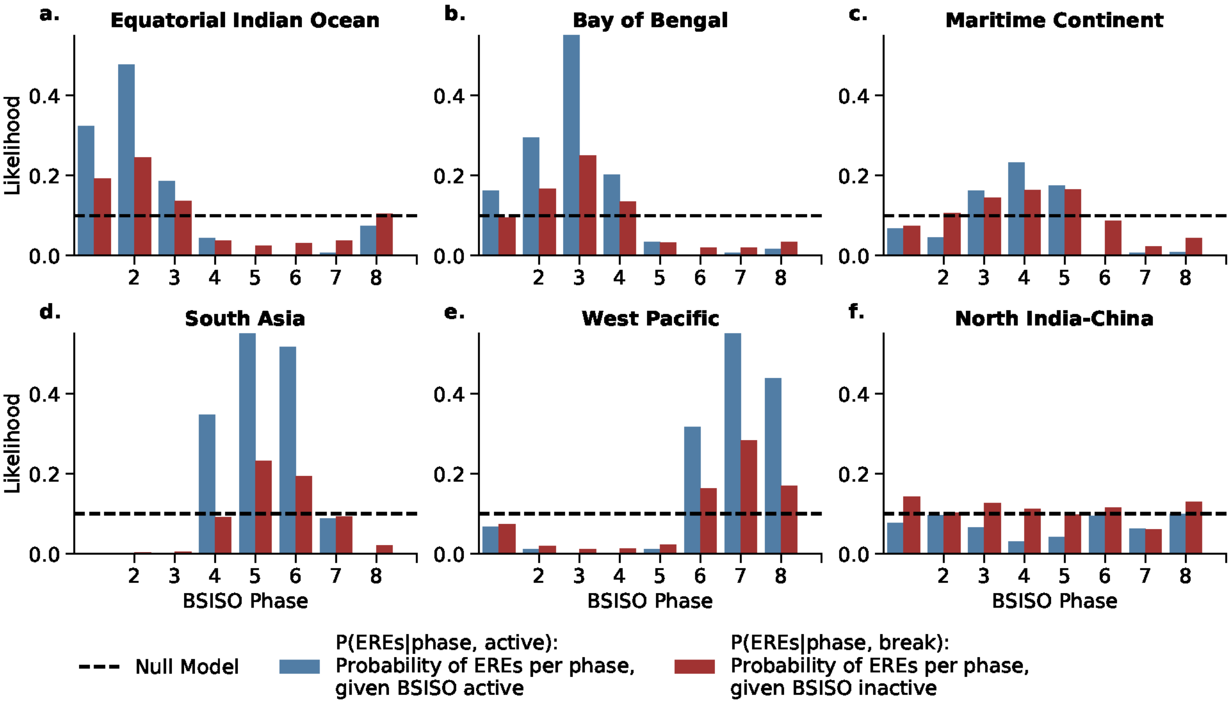


Figure S17. Likelihood of synchronous events for active/inactive BSISO phases. The likelihood of the occurrence of synchronous events ($s = 1$) is analyzed for active (blue) and inactive (red) BSISO phases (as defined by[17]) in the regions of Fig. 2 a: **a** equatorial Indian Ocean, **b** Bay of Bengal, **c** Maritime Continent, **d** South Asia, **e** Western Pacific, and **f** North India-China. The dashed line illustrates the likelihood of synchronous events estimated from a null model of randomly distributed synchronous events (i.e. by construction 10 %).

Supplementary Note 7 Correlation and linear model for BSISO index

The community-specific time series are statistically significantly correlated to the two components of the BSISO index using both the definition by[14] (Fig. S18) or the by[17] (Fig. S19). The time series of the communities EIO, BoB, MC, SA and WP are statistically significantly correlated with either or both the BSISO1 and BSISO2 index. The NIC ERE index is not significantly correlated.

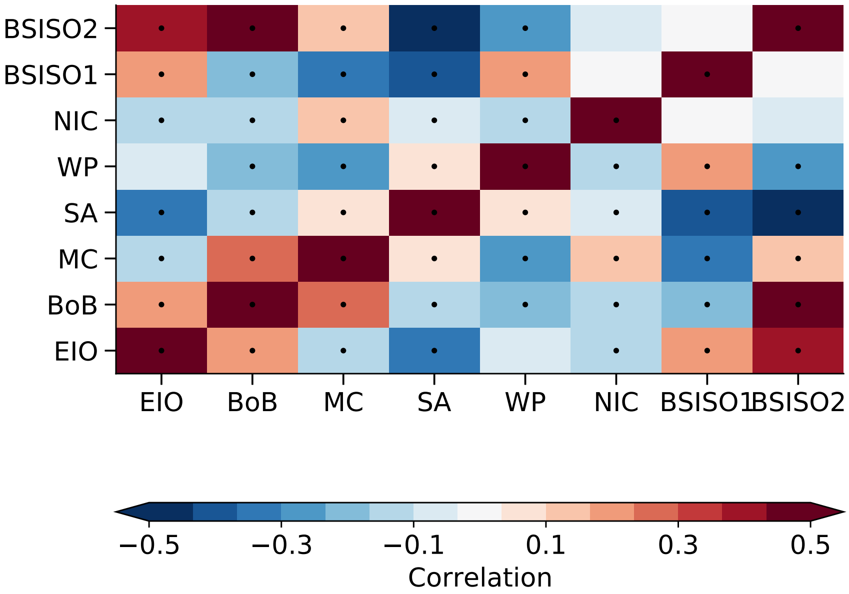


Figure S18. Correlation matrix of community-specific ERE index and BSISO index. The correlation of the community-specific ERE index and the two components of the BSISO index (as defined by Kikuchi et al.[14]) for the regions of Fig. 2 a: Equatorial Indian Ocean (EIO), Bay of Bengal (BoB), Maritime Continent (MC), South Asia (SA), Western Pacific (WP), and North India-China (NIC). Dots indicate significant correlations at the 99% confidence level using Student's t-test.

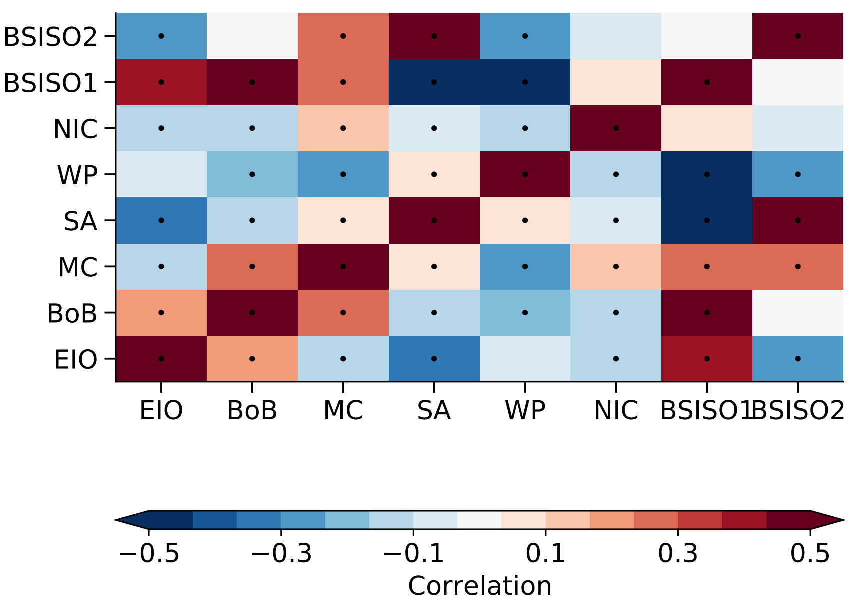


Figure S19. **Correlation matrix of community-specific ERE index and BSISO index**. The correlation of the community-specific ERE index and the two components of the BSISO index (as defined by Lee et al.[17]) for the regions of Fig. 2 a: Equatorial Indian Ocean (EIO), Bay of Bengal (BoB), Maritime Continent (MC), South Asia (SA), Western Pacific (WP), and North India-China (NIC). Dots indicate significant correlations at the 99% confidence level using Student's t-test.

The consistency between the community-specific ERE indices and the BSISO index can also be shown by a linear model. The model is estimated in the following way: A target time series $Y(t)$ is modelled by N input time series $X_i(t)$, where $i = 1,\dots,N$. We use a simple multi-dimensional linear regression incorporating time lags τ_i, expressed as:

$$Y(t) = \sum_{i=0}^{N} a_i X_i(t-\tau_i) + b + \varepsilon(t) \ , \tag{1}$$

where $a_0 \dots a_N, b$ are the fitting parameters and $\varepsilon(t)$ is the error term. The fitting parameters are estimated using the least square fit. The goodness of the fit (also denoted as "explained variance") is expressed by the square of the correlation between the observed Y values and the predicted $\hat{Y}$ values as $r^2 = \sum(\hat{Y}(t) - \overline{Y}(t)/Y(t) - \bar{Y}(t))$, where $\overline{Y}$ denotes the time average.

In order to quantify the amount of ERE variability that is explained by BSISO, we use the simple linear regression model between specific synchronous ERE indices (see Material&Methods Sec.) and the PC_1 and PC_2 indices of the BSISO index Kikuchi et al.[14]. We obtain the following explained variances: EIO $r = 67$ %, BoB $r = 77$ %, SA $r = 78$ %, MC $r = 54$ %, WP $r = 70$ % and NIC $r = 8$ %. Conversely, BSISO variability is explainable from the community synchronization indices with explained variances of $r = 83$ % ($r = 68$ %) for PC_1 (PC_2) indices. These results, obtained by using a simple linear regression (and applying a low-pass filter with 3 days cutoff on the time series to neglect small daily variations), affirm the close relationship of ERE synchronization to the BSISO.

Supplementary Note 8 Robustness tests

In this section, we show that our results are robust to variations of the choices presented in the main text on either the underlying dataset (sec. Supplementary Note 8.1) or the employed community detection algorithm (sec. Supplementary Note 8.2)

Supplementary Note 8.1 Sensitivity to dataset

To assess the robustness of our findings, we conducted a comparative analysis using the Tropical Rainfall Measuring Mission (TRMM) dataset[18], which provides global coverage with a spatial resolution of 0.25×0.25. The TRMM dataset was processed for the same spatial region and interpolated to match the resolution of the Fekete grid employed in the MSWEP dataset (see Section). By comparing the results presented in Figure 2 a obtained with the MSWEP dataset with the communities identified using the TRMM dataset (Fig. S20), we observe qualitatively similar spatial patterns. However, the membership likelihoods exhibit slightly larger variability in the TRMM dataset (Figure S21) compared to the MSWEP dataset (Figure S4). This discrepancy can be attributed to the shorter time length of the TRMM dataset, which reduces the statistical robustness of the climate network links. Notably, the Bay of Bengal (BoB) community exhibits the largest difference between the MSWEP and TRMM datasets.

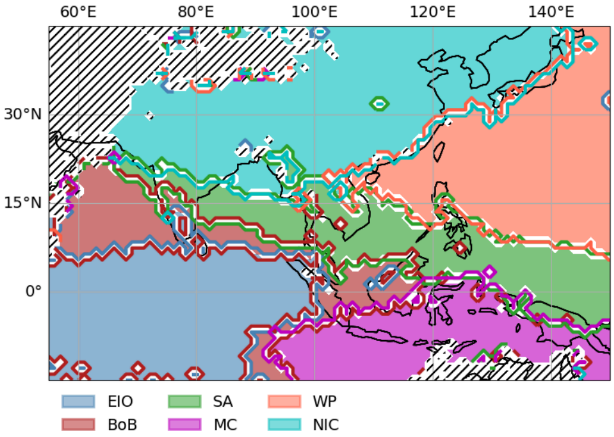


Figure S20. Membership Likelihoods for different communities using the TRMM dataset. Using the heuristic outlined in sec. we find 6 stable communities of a climate network that is constructed by using rainfall data from the TRMM dataset. These are labeled according to their spatial mean position, i.e. equatorial Indian Ocean (EIO), Bay of Bengal (BoB), Maritime Continent (MC), South Asia (SA), Western Pacific (WP), and North India-China (NIC). The color always shows the community for a respective location with its highest membership likelihood that comprises 100 independent runs of the community detection algorithm that has been used for this analysis. Hatched areas indicate regions with too little precipitation, which are excluded from the analysis and white regions indicate locations that cannot be assigned to a single community.

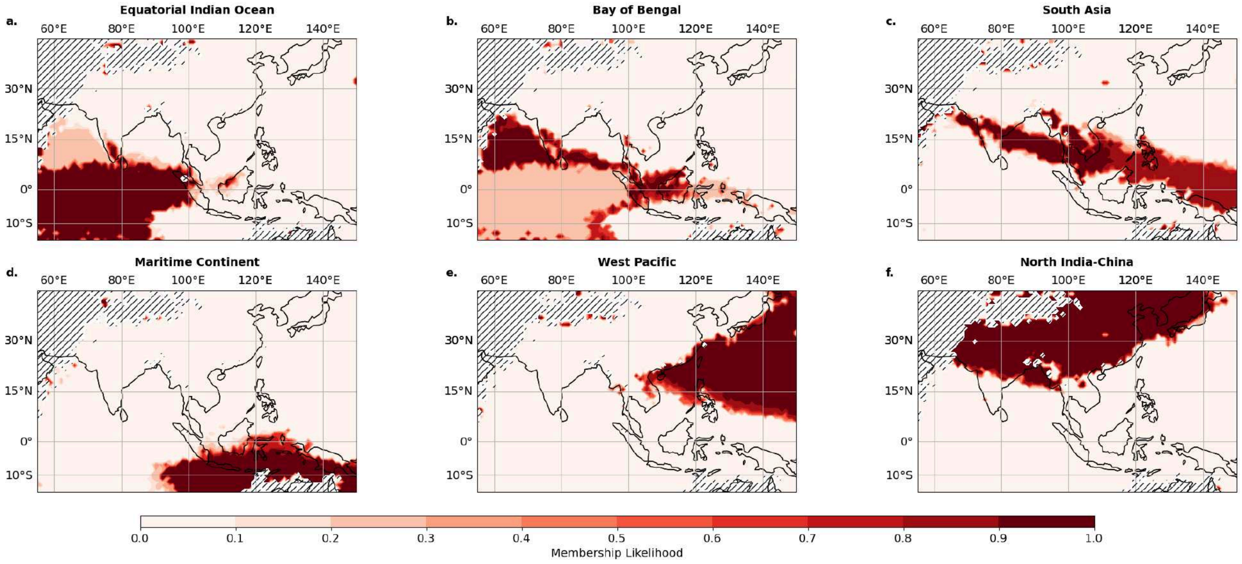


Figure S21. Membership Likelihoods for different communities using the TRMM dataset. Using the heuristic outlined in sec. we find 6 stable communities. The colorbar shows the membership likelihood of a respective community. 100 independent runs of the community detection algorithm have been used for this analysis.

Supplementary Note 8.2 Sensitivity to the choice of community detection algorithm

Here we employ the same climate network as used in the main text to a further community detection algorithm that uses the Parallel Louvain method[19] in the implementation of the NetworkIt package[20]. The observed communities exhibit a qualitative structure that is consistent with the results obtained by the SBM approach (Fig. 2 a). Nevertheless, we opted for the implementation of the Stochastic Block Model (SBM) due to its ability to yield more stable results in terms of the spatial configurations of the identified communities.

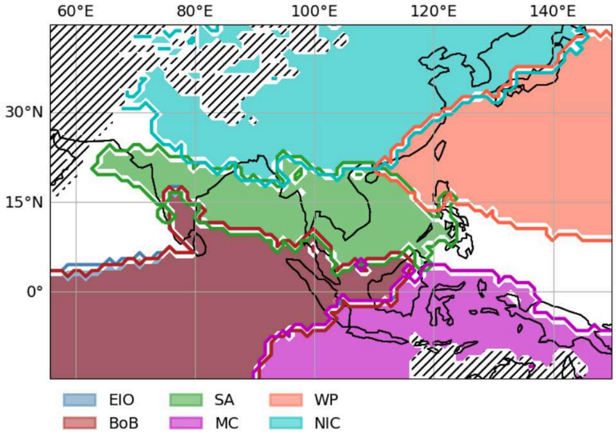


Figure S22. Different communities using PLM method implementation. Using the heuristic outlined in sec. we find 6 stable communities of a climate network that is constructed by using rainfall data from the TRMM dataset. These are labeled according to their spatial mean position, i.e. equatorial Indian Ocean (EIO), Bay of Bengal (BoB), Maritime Continent (MC), South Asia (SA), Western Pacific (WP), and North India-China (NIC). The color always shows the community for a respective location with its highest membership likelihood that comprises 100 independent runs of the community detection algorithm that has been used for this analysis. Hatched areas indicate regions with too little precipitation, which are excluded from the analysis.

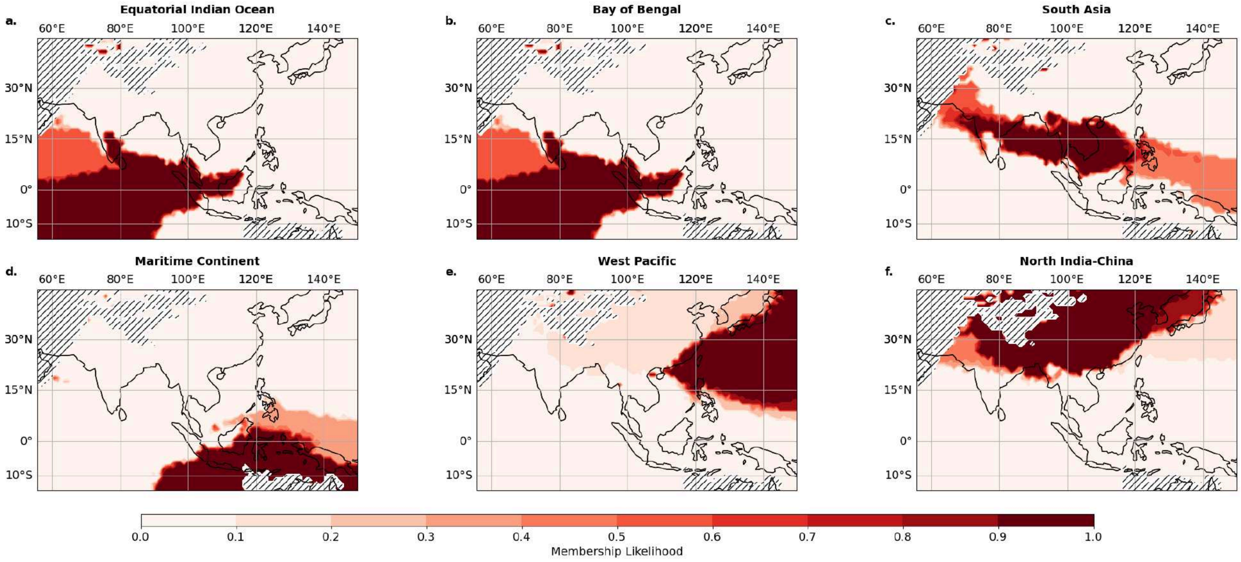


Figure S23. Membership Likelihoods for different communities using the MSWEP dataset with PLM community detection algorithm. Using the heuristic outlined in sec. we find 6 stable communities. The colorbar shows the membership likelihood of a respective community. 100 independent runs of the community detection algorithm have been used for this analysis.

Supplementary Note 9 Network Analysis

The dynamics of different mesoscale systems and the boreal summer atmospheric circulation is reflected by the network structure, and recent studies have shown that different network metrics help to uncover these systems which has been applied in various contexts for analyzing monsoon-related rainfall dynamics (eg.[21–25]). The event synchronization algorithm applied might be very context-specific and differs from study to study. For example, both the spatial range and the time range differ when analyzing tropical cyclones[26] compared with long range teleconnections[21,25]. Similarly, the term "community" is understood differently, for example, sometimes it is used to describe certain pronounced patterns in network measures[27]. Still, different network measures can help to better understand the network communities.

We have applied an analysis for node degree (Fig. S24), clustering coefficient (Fig. S25), betweenness centrality (Fig. S26), and network curvature. All measures are boundary corrected[28].

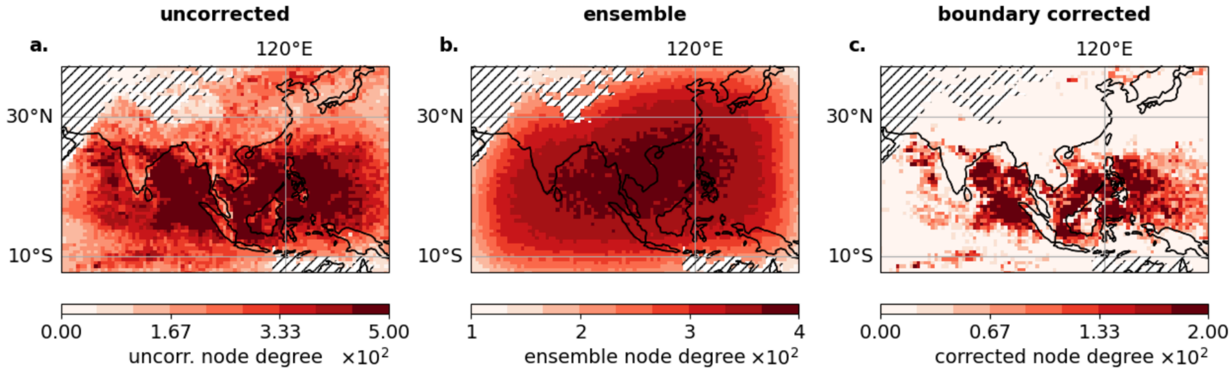


Figure S24. Node degree analysis. We investigate the node degree for the same network on which the community detection algorithm was applied (Fig. 2 a). We also visualize the effect of the boundary correction[28].

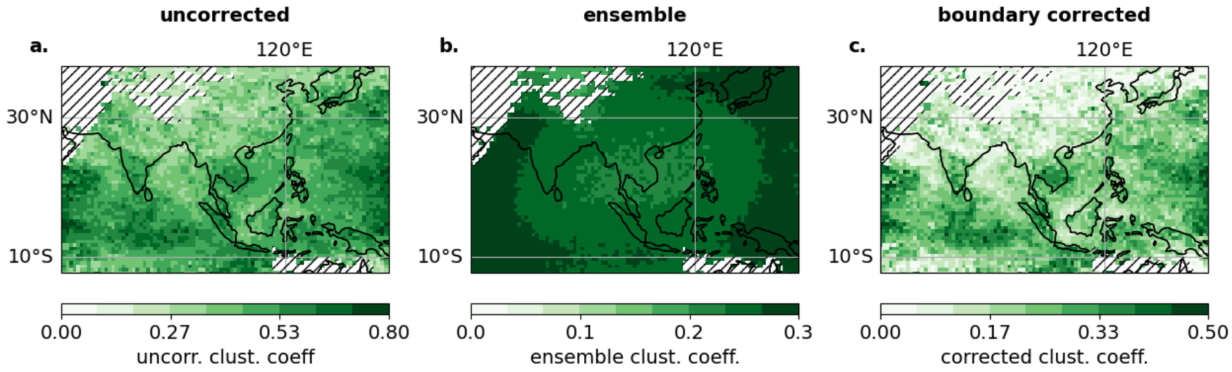


Figure S25. Clustering coefficient analysis. We investigate the clustering coefficient for the same network on which the community detection algorithm was applied (Fig. 2 a). We also visualize the effect of the boundary correction[28].

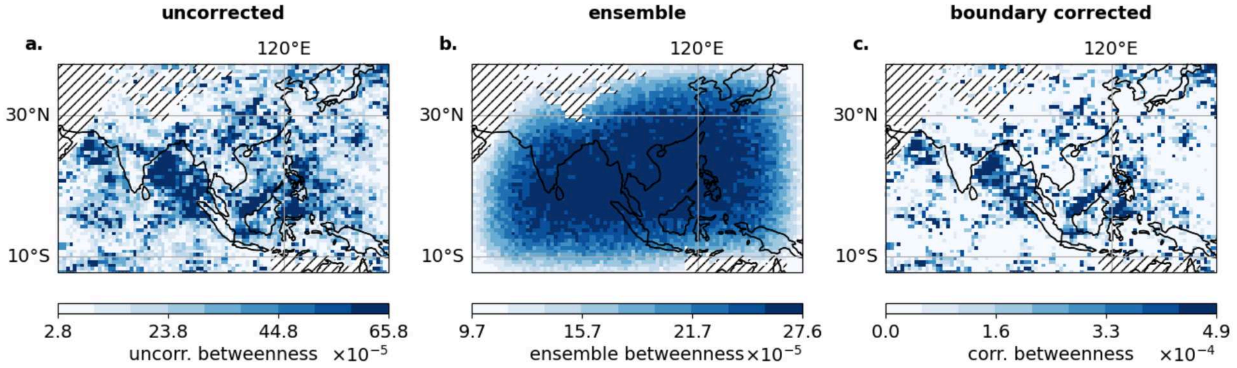


Figure S26. **Betweenness centrality analysis.** We investigate the betweenness centrality for the same network on which the community detection algorithm was applied (Fig. 2 a). We also visualize the effect of the boundary correction, using the method described in[28].

We further analyzed the community-specific network measures. Neither node degree (Fig. S27), nor betweenness centrality (Fig. S28) nor curvature (Fig. S29) provide substantial differences between the communities (Fig. 2 a). However, the local clustering coefficient (Fig. S30) between the NIC community and the other 5 communities shows a shifted distribution indicating that the density of the NIC is different to the other 5 communities. This might be related to the fact, that the community is mainly over land whereas the other communities cover substantial amounts over the ocean.

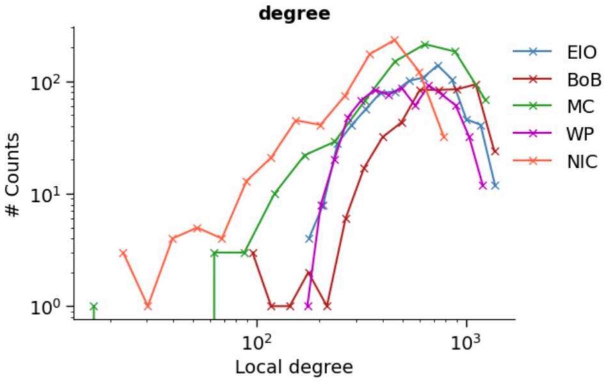


Figure S27. **histogram for node degree.** We investigate the node degree for the nodes of the network that are in the specific communities shown in Fig. 2 a.

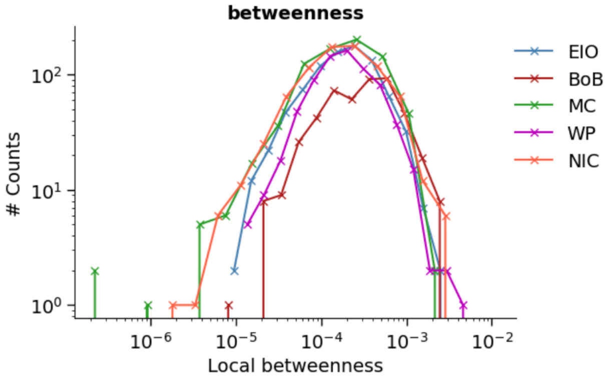


Figure S28. histogram for betweenness centrality We investigate the betweenness centrality for the nodes of the network that are in the specific communities shown in Fig. 2 a.

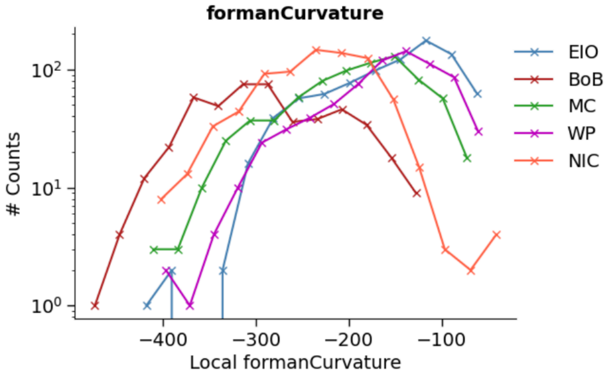


Figure S29. histogram for Forman Curvature We investigate the Forman Curvature for the nodes of the network that are in the specific communities shown in Fig. 2 a.

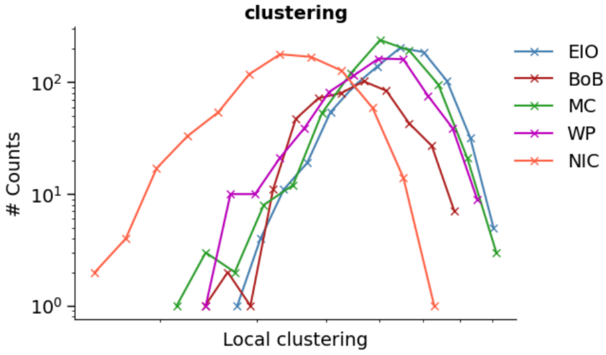


Figure S30. histogram for clusering coefficient. We investigate the local clustering coefficient for the nodes of the network that are in the specific communities shown in Fig. 2 a.

Supplementary Note 10 Comparison with Madden Julian Oscillation

The Madden Julian Oscillation (MJO)[29] is a closely related phenomenon. Here, we show that the MJO alone is not sufficient to explain the organization of EREs during JJAS. To do so, we apply the same conditional independence test as for the BSISO using the RMM index as it was introduced by[16]. The qualitative shape of the distribution remains, however, the lower likelihoods express that BSISO1 and BSISO2 indices are more suited for analysis of EREs during boreal summer (Fig. S31).

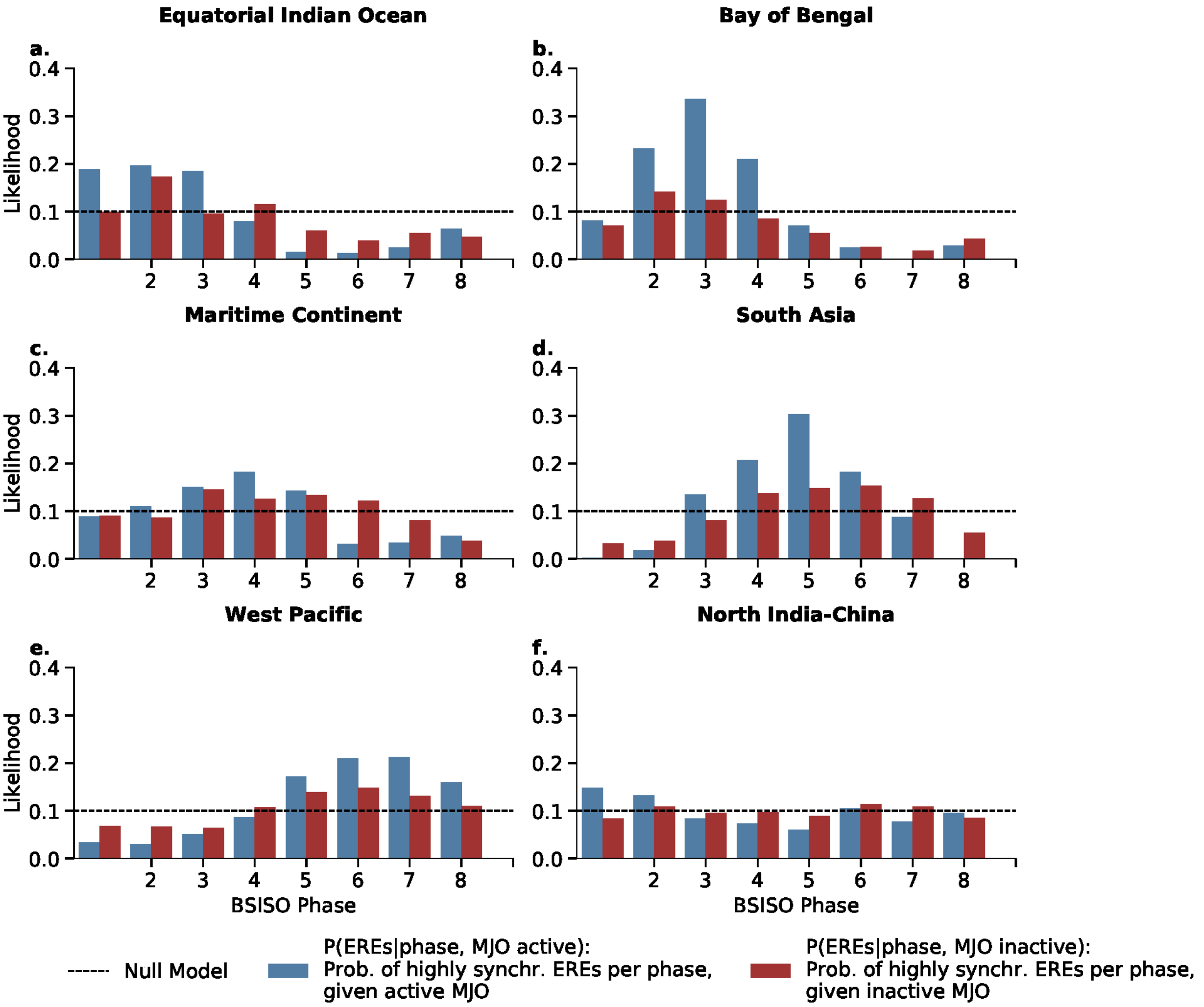


Figure S31. Likelihood of synchronous events for active MJO phases. The occurrence of synchronous events ($s = 1$) is analyzed for active (break) MJO phases displayed as blue (red) bars. The dashed line depicts the likelihood of synchronous events for a threshold of 10 % from a null model of a random distribution of synchronous events. **a-f** Likelihood of synchronous events $s = 1$ for specific phases. The histograms are for the communities received from the analysis in Figure 2 **d**, i.e. **a** Equatorial Indian Ocean, **b** Bay of Bengal, **c** Maritime continent, **d** South Asia, **e** Western Pacific and **f** North India-China.

Supplementary Note 11 Comparison to classical ENSO conditions

The likelihood for a specific propagation mode is substantially increased given the respective ENSO background state (Fig. S32). The conditional dependence test is designed to assess the likelihood of a specific BSISO propagation mode subject to the ENSO state (based on the NINO3.4 index definition[30]). Still, not all samples match exactly with ENSO (Fig. S33). ENSO induces changes in the local zonal as well as the local global meridional overturning circulation. These are visualized in Fig. S34. Also the updrift of vertical winds over the Maritime Continent experiences a modulation by ENSO. This modulation has in turn also influence on the modulation of the BSISO propagation pathways. The differences in the background vertical velocities for the three modes is therefore plotted in Fig.

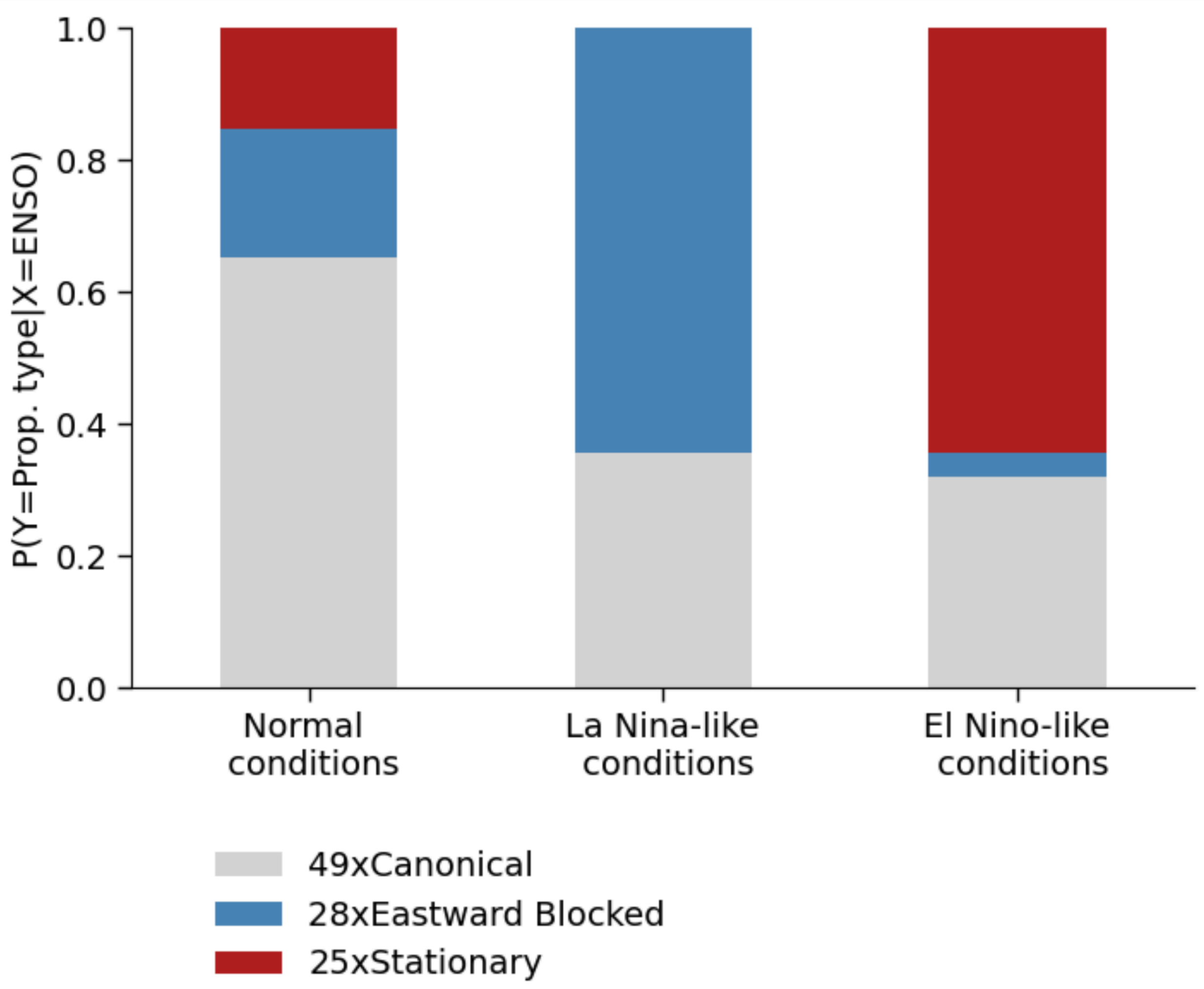


Figure S32. propagation modes at NINO3.4 index. The occurrence of synchronous events is displayed according to the identified propagation mode. The solid line depicts the NINO3.4 index. The definition of an El Niño (La Niña) year was adapted from[31], represented by the dashed line of 0.5 as the respective threshold.

We compare the single BSISO events to the respective ENSO state in the Pacific Ocean. Following[30,32], El Niño and La Niña are defined using the NINO3.4 index. We use June to September daily SST anomalies and select El Niño-like conditions (La Niña-like conditions) based on the average JJAS SST anomalies of the NINO3.4 index region. The respective plot is shown in Fig. S33.

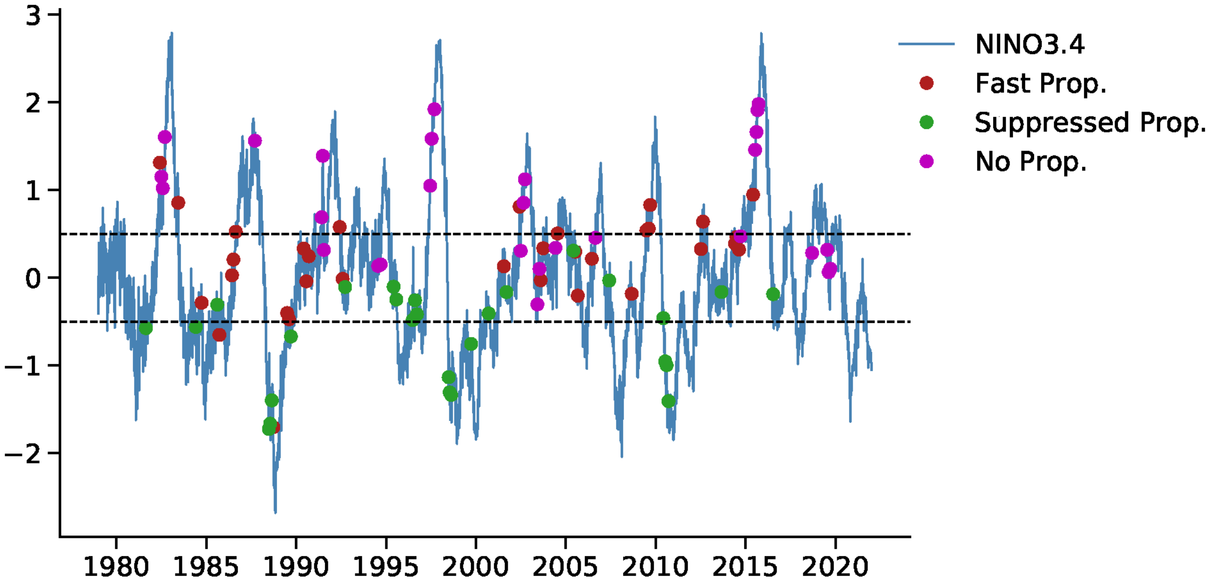


Figure S33. propagation modes at NINO3.4 index. The occurrence of synchronous events is displayed according to the identified propagation mode. The solid line depicts the NINO3.4 index. The definition of an El Niño (La Niña) year was adapted from[32], represented by the dashed line of 0.5 as the respective threshold.

We also analyze the classical overturning circulation conditions for Normal, El Niño-like and La Niña-like conditions in the Pacific Ocean. The pressure level dependent plots are shown in Fig. S34. We further show the spatial structure averaged for 400-600 hPa. The definition of an El Niño, La Niña like condition are based on the NINO3.4 index[32] derived from ERA5 data. The plot is analogous to Fig. 6.

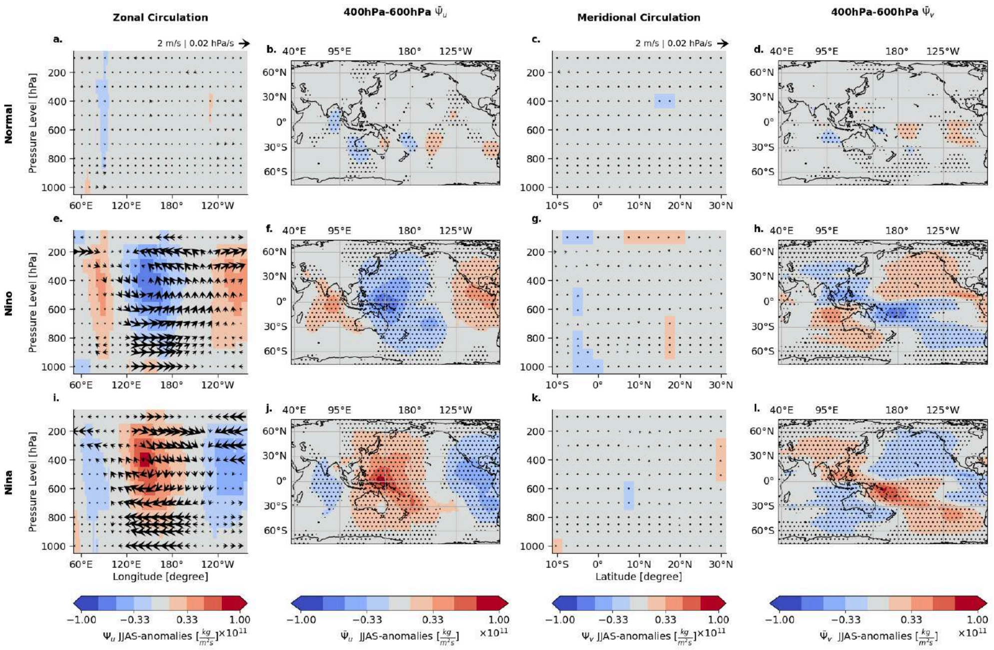


Figure S34. **Vertical cuts of overturning circulation for different ENSO types.** For the ENSO conditions normal (1st row), El Niño (2nd row) and La Niña (3rd row) the background overturning circulation anomalies are plotted. The ENSO state is based on the NINO3.4 definition[32] and the composited SSTs cover per event the full JJAS period. The first column (**a,c,e**) shows the composited vertical cuts in the zonal direction with the latitudes averaged between 10° S-10° N, the second column (**b,d,f**) in the meridional direction in the Indian Ocean with the latitudes averaged between 60° E-120° E. The color shading denotes the mass stream function in the zonal (meridional) direction. The wind fields are computed in the vertical direction from the u (v) components in m/s and in the horizontal direction in hPa/s.

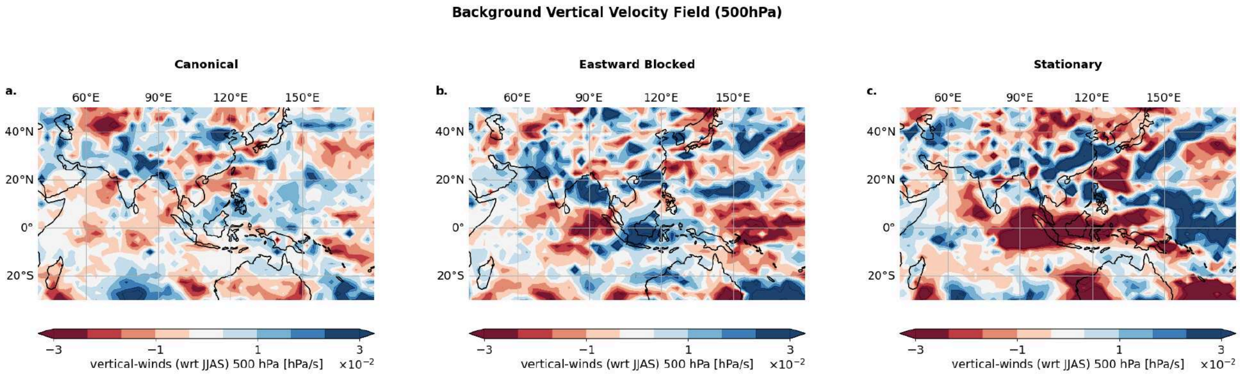


Figure S35. **Background Vertical winds at 500 hPa.**. The comparison of vertical velocity at 500 hPa is shown for the Canonical (1st column), Eastward Blocked (2nd column), and Stationary mode (3rd column) for the background state, averaged between 15-20 days before the BSISO initiation days. The horizontal direction is measured in hPa/s.

Supplementary Note 12 Characteristics of Moisture mode theory

One recently established mechanism for the canonical north-eastward propagation of the BSISO is the moisture mode theory[33, 34]. The BSISO's propagation mechanism is mainly determined by air-sea interaction which is through the feedback processes from updraft air masses associated with BSISO convection. This interaction can be demonstrated by composited plots of downward shortwave radiation (Fig. S36a,d,g), latent heat fluxes (Fig. S36b,e,h), and moisture (Fig. S36c,f,i).

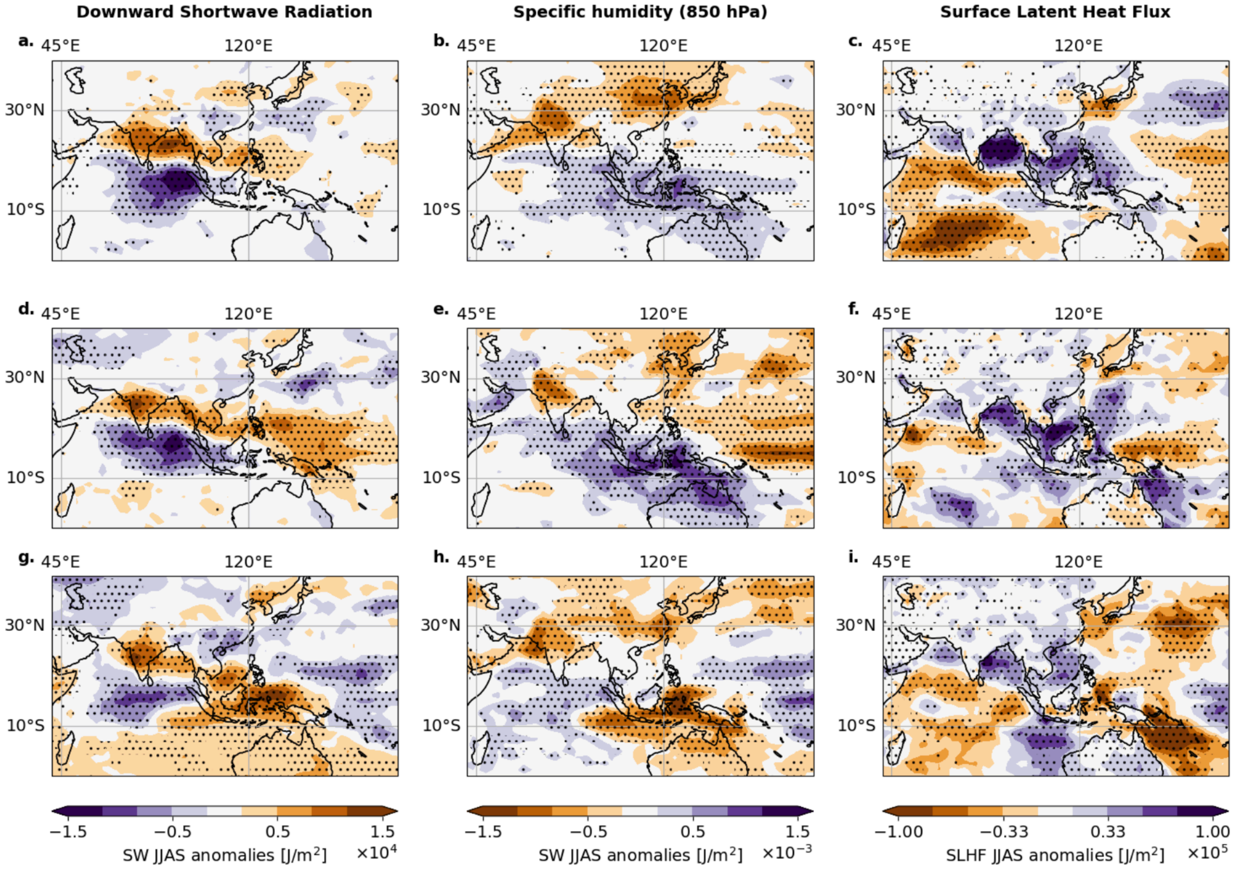


Figure S36. Radiation fluxes analysis. The comparison of Downward Shortwave Radiation (SW), specific humidity (Q) and surface latent heat flux (SLHF) is shown for the Canonical (1st row), Eastward Blocked (2nd row), and Stationary case (3rd row) for the BSISO initiation days. The first column (**a,d,g**) shows the column integrated SW and the second column (**b,e,h**) specific humidity composites. The third column (**c,f,g**) shows the composited surface latent heat flux anomalies. In all subplots, stipples denote anomalies that are significant at a 95 % confidence level using Student's t-test.

Theoretical frameworks[33–35] therefore often use moist static energy (MSE) as the characteristic variable demonstrating the BSISO propagation. Fig. S37a,b,c depict the spatial distribution of rainfall anomalies for the days of maximum synchronization, outgoing longwave radiation and moist static energy (MSE) anomalies at lag 0 day. The positive (negative) anomalies of MJO rainfall align with positive (negative) MSE anomalies, both demonstrating a southeast-northwestward tilt. This agreement between rainfall and MSE anomalies during the BSISO propagation is in accordance with the moisture mode theory and validates the utilization of MSE analysis to investigate the northward propagation mechanism of the BSISO.

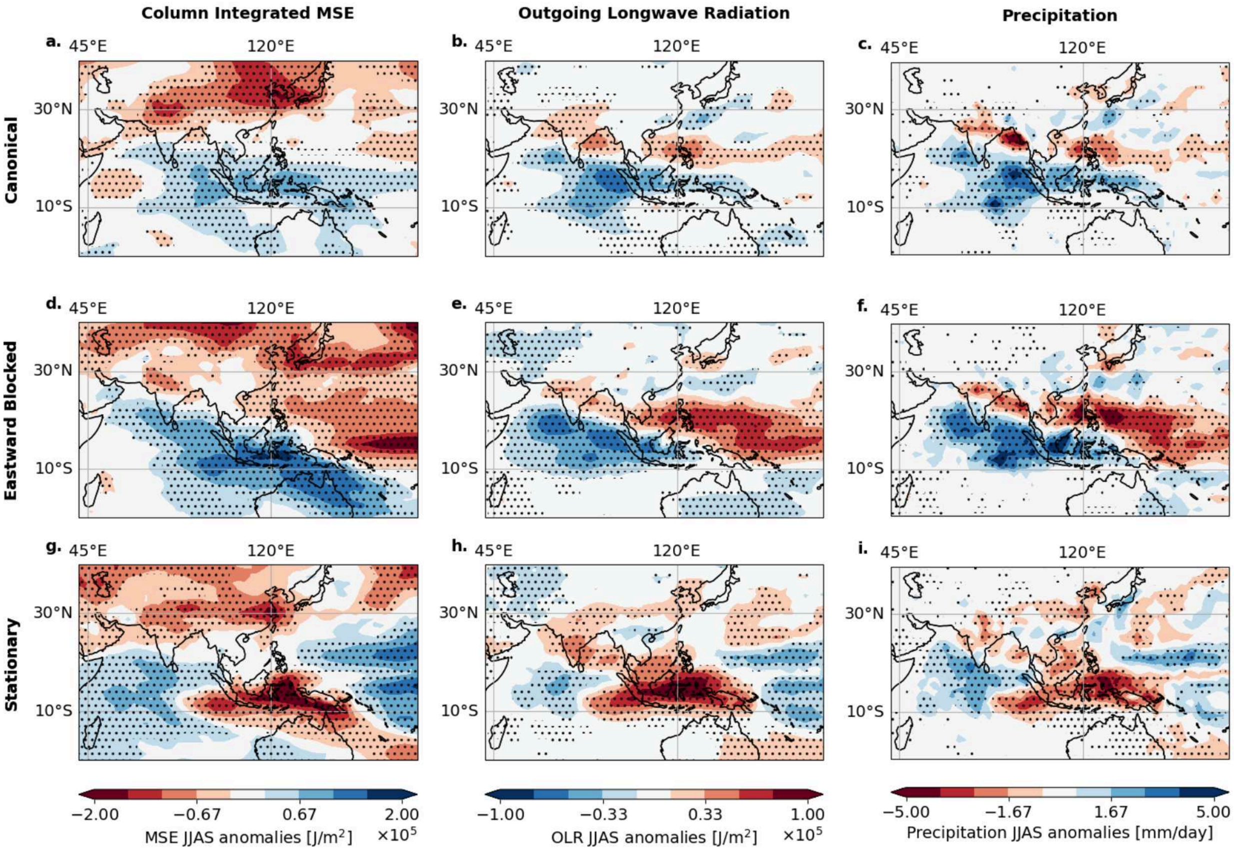


Figure S37. Moisture budget analysis. The comparison of MSE, Outgoing Longwave Radiation (OLR) and precipitation is shown for the Canonical (1st row), Eastward Blocked (2nd row), and Stationary case (3rd row) for the BSISO initiation days. The first column (**a,d,g**) shows the column integrated MSE and the second column (**b,e,h**) OLR composites. The third column (**c,f,g**) shows the composited rainfalls. In all subplots, stipples denote anomalies that are significant at a 95 % confidence level using Student's t-test.

Supplementary Note 13 Vertical Shear Mechanism

One proposed mechanism for the northward propagation of the BSISO is the vertical shear mechanism. It suggests that the easterly negative vertical shear of the zonal mean flow (i.e. low-level westerlies and upper-level easterlies) generates a northward component to the moist Rossby waves[36–38]. As the upward motion w induced by the convective activity in the Indian Ocean decreases northward, the interaction of the BSISO with the easterly vertical shear of the mean flow $\bar{U}_z$ generates cyclonic vorticity ζ_B and boundary layer convergence to the north of the anomalous convection band that promotes the northward movement. This process is described by a two-dimensional barotropic vorticity equation[36]:

$$\frac{\partial \zeta_B}{\partial t} \propto \bar{U}_z \frac{\partial w}{\partial y} \, . \tag{2}$$

Our observed characteristic northward propagation is also in agreement with the vertical shear mechanism[36–38]. The three modes reveal different northward propagation characteristics (Fig. 6 b,d,f) which relate well to the vertical shear mechanism (see Methods). The vertical shear is strongest in the northern Indian Ocean for all three propagation modes (Fig. S38 a,d,g) and hence the necessary condition for northward propagation is fulfilled. The Canonical mode reveals a strong anomalous meridional gradient of the vertical velocity over the Bay of Bengal (Fig. S38 b) and north of the Maritime Continent. This also explains the anomalous relative vorticity pattern over South India and north of the Maritime Continent (Fig. S38 c). We find a similar pattern for the Eastward Blocked mode in the meridional gradients (Fig. S38 e) and thus also in the relative vorticity (Fig. S38 f) with some intensification over the Maritime Continent, which is likely due to the shifted Walker circulation (Fig. 6 f). The Stationary propagation mode is not eastward moving and thus not emitting Rossby Waves upon arriving at the Maritime Continent. Therefore, the region in the northern Indian Ocean does not show significant anomalies in the meridional vertical velocity gradient (Fig. S38 h) and consequently also no significant relative vorticity (Fig. S38 i).

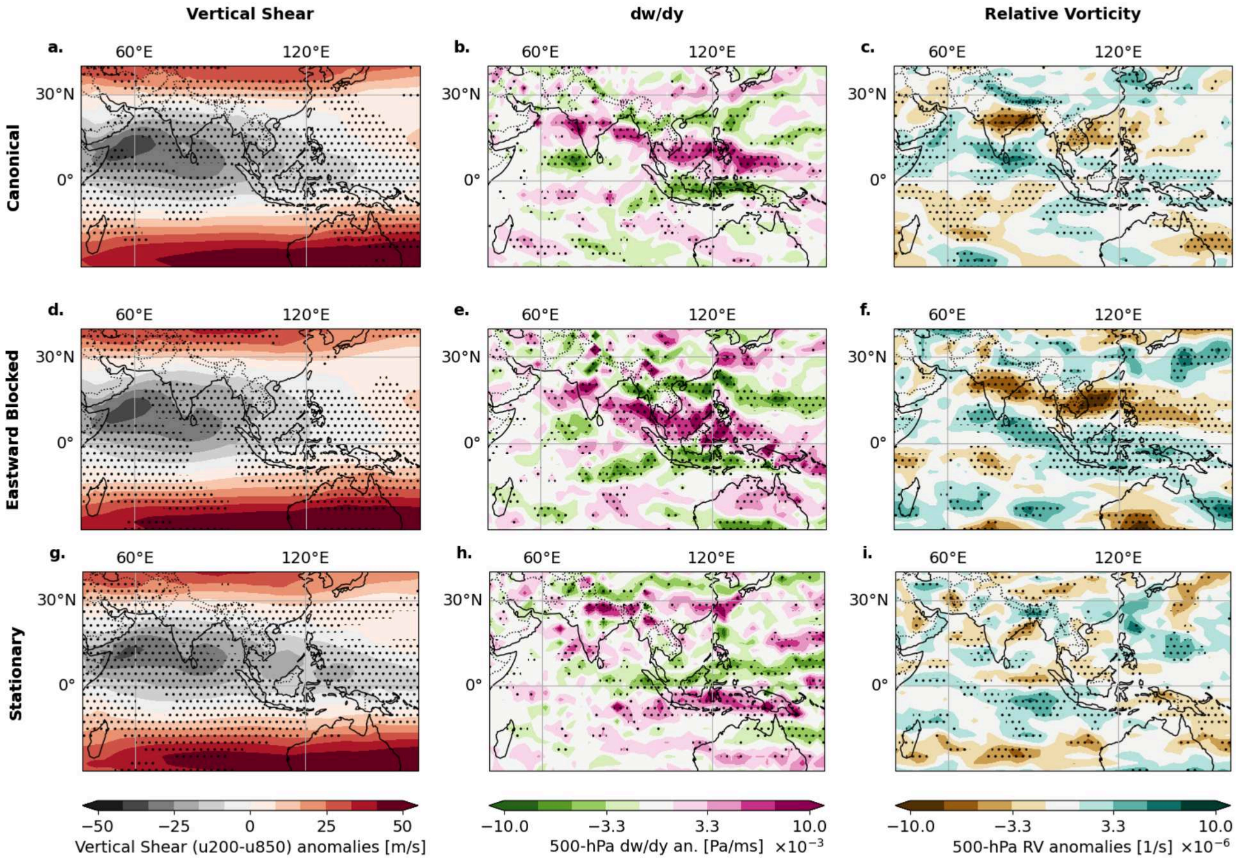


Figure S38. Northward propagation for different propagation modes. The evolution of the northward propagation as a response to eastward propagating Kelvin waves is shown for the Canonical (1st row), Eastward Blocked (2nd row), and Stationary case (3rd row). The first column (**a,d,g**) shows the asymmetrical mean vertical shear $\bar{U}_z$, calculated as $u200 - u850$ and the second column (**b,e,h**) the meridional gradients of the vertical velocity $\frac{\partial w}{\partial y}$ (**c,f,i**). The third column (**c,f,g**) shows the relative vorticity ζ as an indicator of the strength of the northward propagation. All plots are averaged for days 5 to 10 days after initiation. In all subplots, stipples denote anomalies that are significant at a 95 % confidence level using Student's t-test.

Supplementary Note 14 Connection of the rainfall extremes during Indian and East Asian monsoon

One region that emerges from the community detection approach is the NIC community which is of all six regions the only one solely over land (Fig 2 a). The community connects the northeastern part of India, the Tibetan Plateau, the Himalayan Mountains, and most parts of China of which the Himalayan foothills and the Ganges Delta in North East India are among those regions that experience the highest rainfall accumulation during SASM (Fig. S1) during core monsoon season in July.

In early work, the enhanced upper-level atmospheric wave train, known as circumglobal teleconnection[39] (or due to its near-equivalence over Eurasia[40] also known as the "Silk Road pattern"[41]) was suspected to connect the Northern Indian region with the China region[39], and[21] suspect this mechanism to be responsible for the synchronization of EREs between Northern India and Northern China[21]. Composite anomalies of the days of maximum synchronization within the NIC community reveal a large-scale wave train pattern originating from the mid-latitude Atlantic ocean. It is enhanced across Eurasia and connects the Northern India region with the Yellow River basin (Fig. S39) corroborating results from[25].

We also identify a further mechanism connecting parts of Northern India with the Yellow River basin in Northern China. Using composites of vertically integrated moisture vapor flux (IVF) uncovers a continuous path of anomalously high moisture transport established in a moisture corridor (Fig. S40 a-e), starting from the Ganges Delta and gated by the Tibetan Plateau towards the Yellow River Basin and Northern China possibly driven by the Silk Road pattern. The connection manifests itself in the northward displacement of the western North Pacific subtropical high (WNPSH) during the Asian summer monsoon[42] corroborating studies on the dominant route of stage 4 of the East Asian Atmospheric Rivers[42].

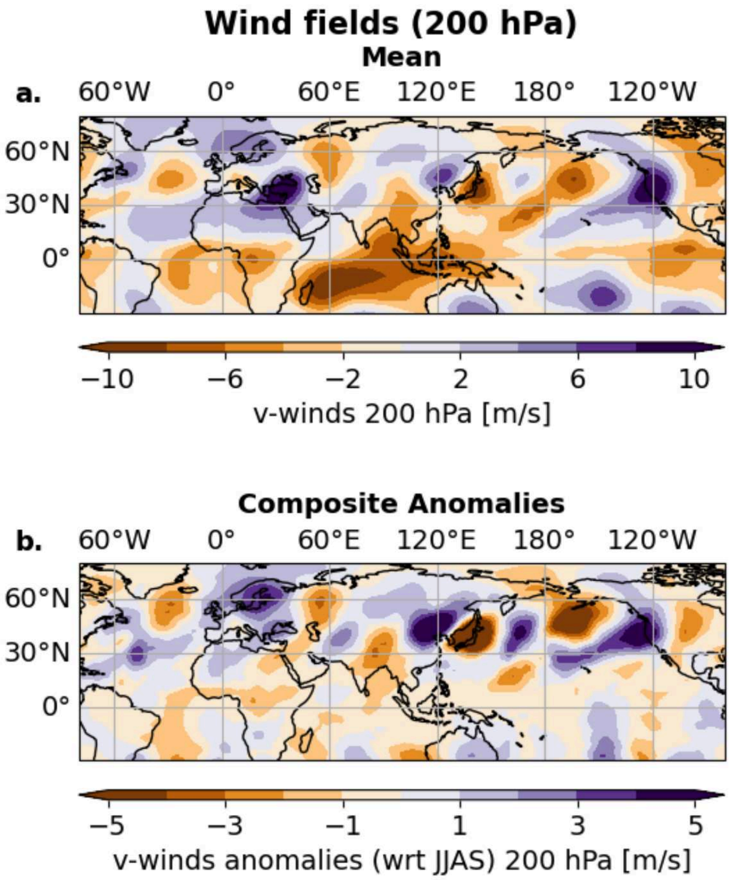


Figure S39. Atmospheric conditions for the synchronous day pattern. The v-winds components at 200 hPa is plotted for the mean (**a**) the composited anomalies **b**). The days were chosen as described in section using the days of maximum synchronization in the NIC region. The composite anomalies are computed with respect to the JJAS climatology. The Silk Road pattern connects clearly the North of India with the China area.

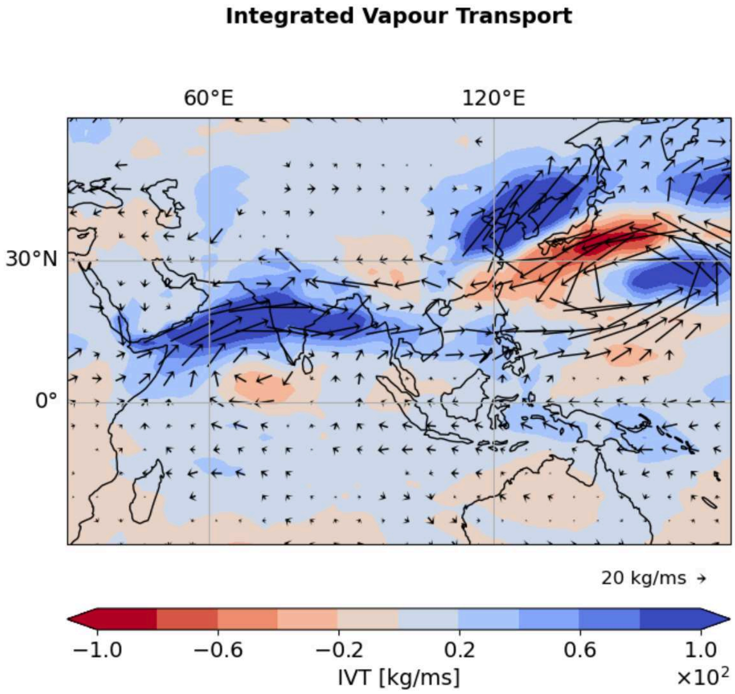


Figure S40. **Vertically integrated water vapor flux (IVT) for days of maximum synchronization for NIC region.** Same as Figure S39 but for the vertically integrated water vapor flux (IVF). Day 0 denotes the days of maximum synchronization. The composite anomalies are computed with respect to the JJAS climatology. Only IVT arrows that are significant at 95 % level following the Student's t-test are plotted.

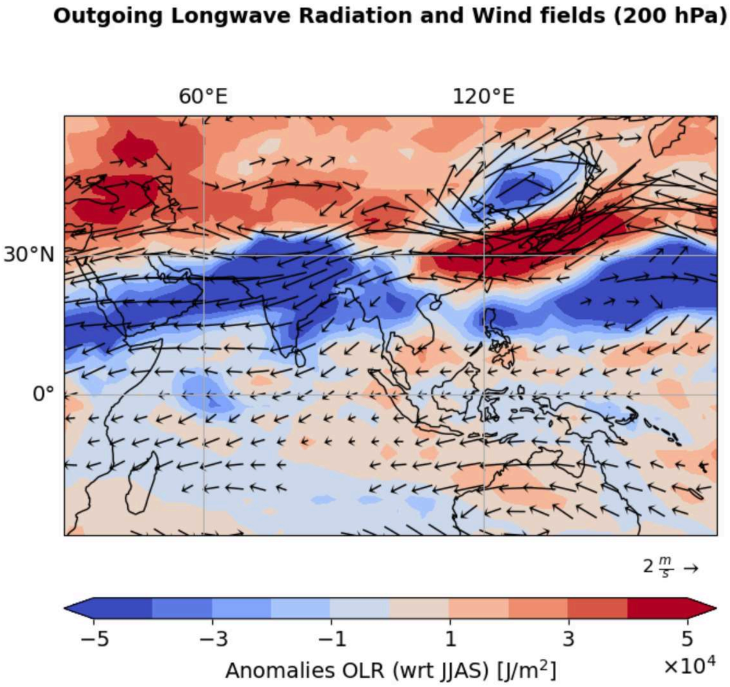


Figure S41. Outgoing Longwave Radiation for days of maximum synchronization for NIC region Same as Figure S39 but for Outgoing Longwave Radiation (IVF) overlapped by wind fields at from anomalous upper-level wind fields at 200 hPa. Composites are computed for the days of maximum synchronization. The composite anomalies are computed with respect to the JJAS climatology. Only wind arrows that are significant at 95 % level following the Student's t-test are plotted.

Supplementary References

1. Rheinwalt, A. *et al.* Non-linear time series analysis of precipitation events using regional climate networks for Germany. *Clim. Dyn.* **46**, 1065–1074, DOI: 10.1007/s00382-015-2632-z (2016).
2. Beck, H. E. *et al.* MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bull. Am. Meteorol. Soc.* **100**, 473–500, DOI: 10.1175/BAMS-D-17-0138.1 (2019).
3. Lau, W. K. M., Waliser, D. E., Majda, A. J. & Stechmann, S. N. Multiscale theories for the MJO. In *Intraseasonal Variability in the Atmosphere-Ocean Climate System*, 549–568, DOI: 10.1007/978-3-642-13914-7_17 (Springer, Berlin, Germany, 2011).
4. Kim, H., Vitart, F. & Waliser, D. E. Prediction of the Madden–Julian Oscillation: A Review. *J. Clim.* **31**, 9425–9443, DOI: 10.1175/JCLI-D-18-0210.1 (2018).
5. Kikuchi, K. The Boreal Summer Intraseasonal Oscillation (BSISO): A Review. *J. Meteorol. Soc. Jpn. Ser. II* 2021–045, DOI: 10.2151/jmsj.2021-045 (2021).
6. 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, DOI: 10.1007/s00703-016-0470-0 (2017).
7. Da Silva, N. A. & Matthews, A. J. Impact of the Madden–Julian Oscillation on extreme precipitation over the western Maritime Continent and Southeast Asia. *Q. J. R. Meteorol. Soc.* **147**, 3434–3453, DOI: 10.1002/qj.4136 (2021).
8. Anandh, P. C., Vissa, N. K. & Broderick, C. Role of MJO in modulating rainfall characteristics observed over India in all seasons utilizing TRMM. *Int. J. Clim.* **38**, 2352–2373, DOI: 10.1002/joc.5339 (2018).
9. Di Capua, G. *et al.* Tropical and mid-latitude teleconnections interacting with the Indian summer monsoon rainfall: a theory-guided causal effect network approach. *Earth Syst. Dyn.* **11**, 17–34, DOI: 10.5194/esd-11-17-2020 (2020).
10. Karmakar, N., Boos, W. R. & Misra, V. Influence of Intraseasonal Variability on the Development of Monsoon Depressions. *Geophys. Res. Lett.* **48**, e2020GL090425, DOI: 10.1029/2020GL090425 (2021).
11. Schreck, C. J. Global Survey of the MJO and Extreme Precipitation. *Geophys. Res. Lett.* **48**, e2021GL094691, DOI: 10.1029/2021GL094691 (2021).
12. Hunt, K. M. R. & Turner, A. G. Nonlinear intensification of monsoon low pressure systems by the BSISO. *Weather. Clim. Dyn. Discuss.* 1–28, DOI: 10.5194/wcd-2022-31 (2022).

13. Li, X., Gollan, G., Greatbatch, R. J. & Lu, R. Impact of the MJO on the interannual variation of the Pacific–Japan mode of the East Asian summer monsoon. *Clim. Dyn.* **52**, 3489–3501, DOI: 10.1007/s00382-018-4328-7 (2019).

14. Kikuchi, K., Wang, B. & Kajikawa, Y. Bimodal representation of the tropical intraseasonal oscillation. *Clim. Dyn.* **38**, 1989–2000, DOI: 10.1007/s00382-011-1159-1 (2012).

15. Wang, S., Ma, D., Sobel, A. H. & Tippett, M. K. Propagation Characteristics of BSISO Indices. *Geophys. Res. Lett.* **45**, 9934–9943, DOI: 10.1029/2018GL078321 (2018).

16. Wheeler, M. C. & Hendon, H. H. An All-Season Real-Time Multivariate MJO Index: Development of an Index for Monitoring and Prediction. *Mon. Weather. Rev.* **132**, 1917–1932, DOI: 10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2 (2004).

17. Lee, J.-Y. *et al.* Real-time multivariate indices for the boreal summer intraseasonal oscillation over the Asian summer monsoon region. *Clim. Dyn.* **40**, 493–509, DOI: 10.1007/s00382-012-1544-4 (2013).

18. Huffman, G. J. *et al.* The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* **8**, 38–55, DOI: 10.1175/JHM560.1 (2007).

19. Lancichinetti, A. & Fortunato, S. Community detection algorithms: A comparative analysis. *Phys. Rev. E* **80**, 056117, DOI: 10.1103/PhysRevE.80.056117 (2009).

20. Staudt, C. L., Sazonovs, A. & Meyerhenke, H. NetworKit: A Tool Suite for Large-scale Complex Network Analysis. *arXiv* DOI: 10.48550/arXiv.1403.3005 (2014). 1403.3005.

21. Boers, N. *et al.* Complex networks reveal global pattern of extreme-rainfall teleconnections. *Nature* **566**, 373–377, DOI: 10.1038/s41586-018-0872-x (2019).

22. Stolbova, V., Surovyatkina, E., Bookhagen, B. & Kurths, J. Tipping elements of the Indian monsoon: Prediction of onset and withdrawal. *Geophys. Res. Lett.* **43**, 3982–3990, DOI: 10.1002/2016GL068392 (2016).

23. Wolf, F., Ozturk, U., Cheung, K. & Donner, R. V. Spatiotemporal patterns of synchronous heavy rainfall events in East Asia during the Baiu season. *Earth Syst. Dyn.* **12**, 295–312, DOI: 10.5194/esd-12-295-2021 (2021).

24. Malik, N., Marwan, N. & Kurths, J. Spatial structures and directionalities in Monsoonal precipitation over South Asia. *Nonlinear Process. Geophys.* **17**, 371–381, DOI: 10.5194/npg-17-371-2010 (2010).

25. Gupta, S. *et al.* Interconnection between the Indian and the East Asian Summer Monsoon: spatial synchronization patterns of extreme rainfall events. *Int. J. Clim.* **n/a**, DOI: 10.1002/joc.7861 (2022).

26. Gupta, S., Boers, N., Pappenberger, F. & Kurths, J. Complex network approach for detecting tropical cyclones. *Clim. Dyn.* **57**, 3355–3364, DOI: 10.1007/s00382-021-05871-0 (2021).

27. Strnad, F. M., Schlör, J., Fröhlich, C. & Goswami, B. Teleconnection Patterns of Different El Niño Types Revealed by Climate Network Curvature. *Geophys. Res. Lett.* **49**, e2022GL098571, DOI: 10.1029/2022GL098571 (2022).

28. Rheinwalt, A., Marwan, N., Kurths, J., Werner, P. & Gerstengarbe, F.-W. Boundary effects in network measures of spatially embedded networks. *Eur. Lett.* **100**, 28002, DOI: 10.1209/0295-5075/100/28002 (2012).

29. Madden, R. A. & Julian, P. R. Detection of a 40–50 Day Oscillation in the Zonal Wind in the Tropical Pacific. *J. Atmos. Sci.* **28**, 702–708, DOI: 10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2 (1971).

30. Trenberth, K. E. The Definition of El Niño. *Bull. Am. Meteorol. Soc.* **78**, 2771–2778, DOI: 10.1175/1520-0477(1997)078<2771:TDOENO>2.0.CO;2 (1997).

31. Capotondi, A., Wittenberg, A. T., Kug, J.-S., Takahashi, K. & McPhaden, M. J. ENSO Diversity. In *El Niño Southern Oscillation in a Changing Climate*, 65–86, DOI: 10.1002/9781119548164.ch4 (American Geophysical Union (AGU), 2020).

32. Trenberth, K. E. & Stepaniak, D. P. Indices of El Niño Evolution. *J. Clim.* **14**, 1697–1701, DOI: 10.1175/1520-0442(2001)014<1697:LIOENO>2.0.CO;2 (2001).

33. Sobel, A. & Maloney, E. An Idealized Semi-Empirical Framework for Modeling the Madden–Julian Oscillation. *J. Atmos. Sci.* **69**, 1691–1705, DOI: 10.1175/JAS-D-11-0118.1 (2012).

34. Wang, S. & Sobel, A. H. A Unified Moisture Mode Theory for the Madden–Julian Oscillation and the Boreal Summer Intraseasonal Oscillation. *J. Clim.* **35**, 1267–1291, DOI: 10.1175/JCLI-D-21-0361.1 (2022).

35. Adames, Á. F. & Kim, D. The MJO as a Dispersive, Convectively Coupled Moisture Wave: Theory and Observations. *J. Atmos. Sci.* **73**, 913–941, DOI: 10.1175/JAS-D-15-0170.1 (2016).

36. Wang, B. & Xie, X. A Model for the Boreal Summer Intraseasonal Oscillation. *J. Atmos. Sci.* **54**, 72–86, DOI: 10.1175/1520-0469(1997)054<0072:AMFTBS>2.0.CO;2 (1997).

37. Wang, B., Webster, P. J. & Teng, H. Antecedents and self-induction of active-break south Asian monsoon unraveled by satellites. *Geophys. Res. Lett.* **32**, DOI: 10.1029/2004GL020996 (2005).

38. Hoskins, B. & Wang, B. Large-scale atmospheric dynamics. In *The Asian Monsoon*, 357–415, DOI: 10.1007/3-540-37722-0_9 (Springer, Berlin, Germany, 2006).

39. Ding, Q. & Wang, B. Circumglobal Teleconnection in the Northern Hemisphere Summer. *J. Clim.* **18**, 3483–3505, DOI: 10.1175/JCLI3473.1 (2005).

40. Zhou, F., Zhang, R. & Han, J. Relationship between the Circumglobal Teleconnection and Silk Road Pattern over Eurasian continent. *Sci. Bull.* **64**, 374–376, DOI: 10.1016/j.scib.2019.02.014 (2019).

41. Enomoto, T., Hoskins, B. J. & Matsuda, Y. The formation mechanism of the Bonin high in August. *Q. J. R. Meteorol. Soc.* **129**, 157–178, DOI: 10.1256/qj.01.211 (2003).

42. Pan, M. & Lu, M. East Asia Atmospheric River catalog: Annual Cycle, Transition Mechanism, and Precipitation. *Geophys. Res. Lett.* **47**, e2020GL089477, DOI: 10.1029/2020GL089477 (2020).

Supplementary Material for Chapter 6 C.

Supplementary Information

In this Supplementary Information, we provide additional information on the data and methods used in this study. We also provide additional figures that support the main text. The Supplementary Information is organized as follows: In Sec. SI 1, we provide additional information on the climate network used in this study. In Sec. SI 2, we provide additional information on the basic pattern of the synchronization between North India and the Sahel Zone. In Sec. SI 3 we provide additional information on the yearly fluctuations of the synchronization between North India and the Sahel. In Sec. SI 4, we give additional information on the Tropical Easterly Jet (TEJ) and its characteristics. In Sec. SI 5, we outline additional information on the Circumglobal Teleconnection (CGT) and its characteristics. SI 6 provides additional information on the Pacific Decadal Oscillation (PDO) and its characteristics. The characteristics of the Somali Jet are shown in Sec. SI 7. In Sec. SI 8, we provide additional information on the Boreal Summer Intraseasonal Oscillation (BSISO) and its characteristics. In section SI 9, we provide additional information on the propagation pathways of the synchronization between North India and the Sahel Zone as they are estimated by a Lagrangian trajectory model.

SI 1 Climate Network of Synchronous Rainfall Events

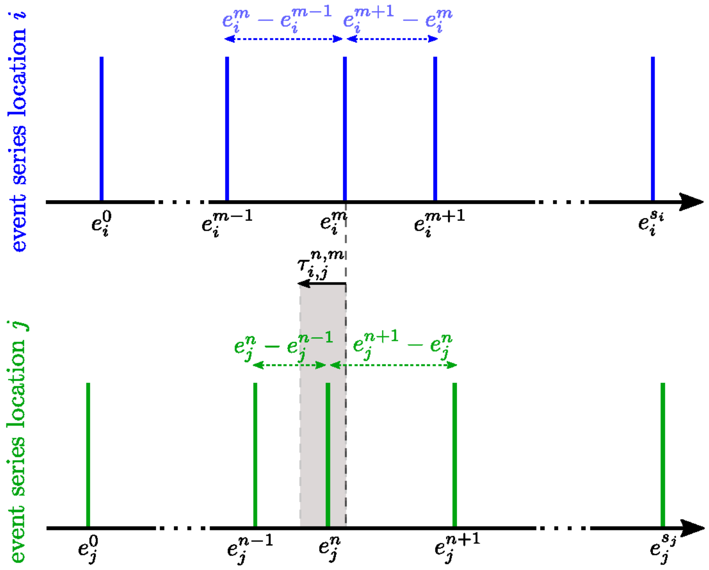


Figure S1: **Scheme of time-delayed Event Synchronization.** The blue (green) bars denote events that happen at the locations in i (j) at the time points e_i^m (e_j^n). m (n) denotes the event number at location i (j). The arrows indicate the time differences between neighboring events at the respective locations, i.e. $d_{i,i}^{m,m-1}$, $d_{i,i}^{m,m+1}$, $d_{j,j}^{n,n-1}$ and $d_{j,j}^{n,n+1}$. In this exemplary sketch, the point in time e_i^m is identified as synchronous to e_j^n since the time difference between e_i^m and e_j^n is within the allowed range given by $\tau_{ij}^{m,n}$ according to the definition expressed in equation 5. Not shown here is the maximum delay $\tau_{\max}$.

Event Synchronization To quantify the level of synchronization between pairs of time series, we employ the Event Synchronization algorithm Quian Quiroga et al. (2002). This algorithm counts the number of events that occur at the same time between pairs of event sequences $\{e_i^m\}_{m=1}^{s_i}$ and $\{e_j^n\}_{n=1}^{s_j}$, where s_i (s_j) represents the total number of events at location i (j), and e_i^m (e_j^n) denotes the point in time of the mth (nth) event in that location. Each event in each time series is compared to all events in the other time series, and the number of synchronous events are counted. To account for a certain delay between events, we introduce a dynamical delay $\tau_{ij}^{m,n}$. The time difference d between an event e_i^m in location i and an event e_j^n in location j is denoted as $d_{ij}^{m,n} = e_i^m - e_j^n$. The set $D_{i,j}^{m,n}$ is defined as the set that contains the time differences to the neighboring events of e_i^m and e_j^n respectively:

$$D_{i,j}^{m,n} = \left\{ d_{i,i}^{m,m-1}, d_{i,i}^{m,m+1}, d_{j,j}^{n,n-1}, d_{j,j}^{n,n+1}, 2\tau_{\max} \right\}, \tag{4}$$

The dynamical delay is now calculated as half of the minimum time difference between subsequent events in both time series around event e_i^m and e_j^n and is limited to a maximum value of τ_{max}:

$$\tau_{ij}^{m,n} = \frac{1}{2} \min_{\forall d \in D_{ij}^{m,n}} d \, . \tag{5}$$

In Fig. S1, we show an exemplary scheme of two events in a two event series that are considered as synchronous. The parameter τ_{max} represents the maximum allowable time delay between two events, capturing both short and long-spatial range interactions. In this study, we set τ_{max} to 10 days. The event synchronization strength $R_{i,j}$ between locations i and j is calculated as the sum of synchronous time points between all pairs of event sequences $\{e_i^m\}_{m=1}^{s_i}$ and $\{e_j^n\}_{n=1}^{s_j}$:

$$R_{i,j} = \sum_{m=1}^{s_i} \sum_{n=1}^{s_j} S_{i,j}^{m,n} \qquad \text{where} \qquad S_{i,j}^{m,n} = \begin{cases} 1 & 0 < d_{ij}^{m,n} < \tau_{i,j}^{m,n}\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{6}$$

We count blocks of consecutive events as a single event, positioned at the time of the first event. This avoids a hypothetical possible dynamical delay value of $1/2$, which would lead to a case where two sequentially occurring events are not considered synchronous.

Constructing the network. To determine the number of counted synchronous events between location i and j is statistically significant, we perform a null-model test. The null hypothesis assumes that the observed $R_{i,j}$ value arises from a pair of randomly generated event sequences $\{e_i'^m\}_{m=1}^{s_i}$ and $\{e_j'^n\}_{n=1}^{s_j}$ with the same number of events s_i, s_j as the observed sequences. We construct surrogate event sequences by randomly and uniformly distributing events over the length of the time series. We consider event series e_i to be significantly synchronous with event series e_j if their corresponding $R_{i,j}$ value exceeds the 95th percentile of the distribution $R'_{i,j}$ values obtained from 2000 pairs of surrogate event sequences e'_i, e'_j. When a significant $R_{i,j}$ value is found, we establish an edge from node n_i to n_j. Mathematically this network is represented by the adjacency matrix which is an $N \times N$ matrix. Each entry of the matrix can be either 0 or 1. One defines that $A_{i,j} = 1$ indicates that events at location i are statistically significantly followed by events at location j. If $A_{i,j} = 0$, there is no significant synchronization between the two locations.

Spatial link correction To avoid considering singular pairs of time series as statistically significant by coincidence, we perform a spatial link correction. We assume that synchronous time series are caused by physical mechanisms and exhibit spatially coherent patterns (Boers et al., 2019). For each spatial location, we randomly rewire its network links 2000 times. We then use a Gaussian kernel density estimator (KDE) with the bandwidth selected according to Scott's Rule of Thumb to compute the spatial link distribution of each random sample. A link is considered significant if its regional link distribution, obtained by a Gaussian KDE, is above the 99.9th percentile.

SI 2 Community Detection in the tropics and subtropics

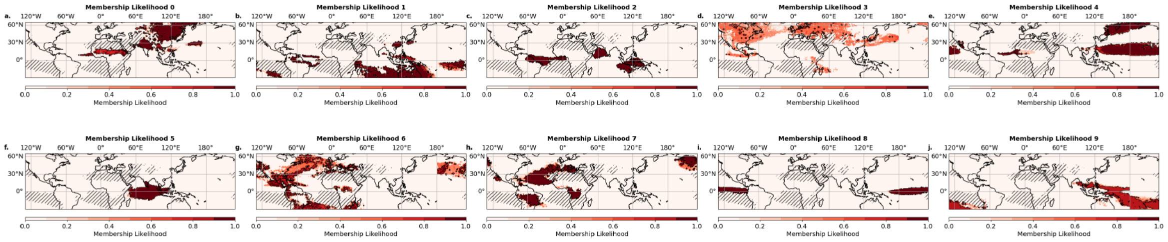


Figure S2: **Membership Likelihoods for different communities.** Using the heuristic outlined in sec. 3 we find 10 communities. The color bar shows the membership likelihood of a respective community. 100 independent runs of the community detection algorithm have been used for this analysis. We find that some communities show a quite stable occurrence, some others are more fluctuating. However, for this work, we are mainly interested in the North India - Sahel zone community which is one of the most stable communities of the algorithm.

SI 3 Yearly Fluctuations of the Synchronization

This section investigates the yearly fluctuations of the synchronization between North India and the Sahel Zone. The different years are clustered using a Gaussian Mixture model. We find that the synchronization is not constant but varies between different years (Fig. S3 a).

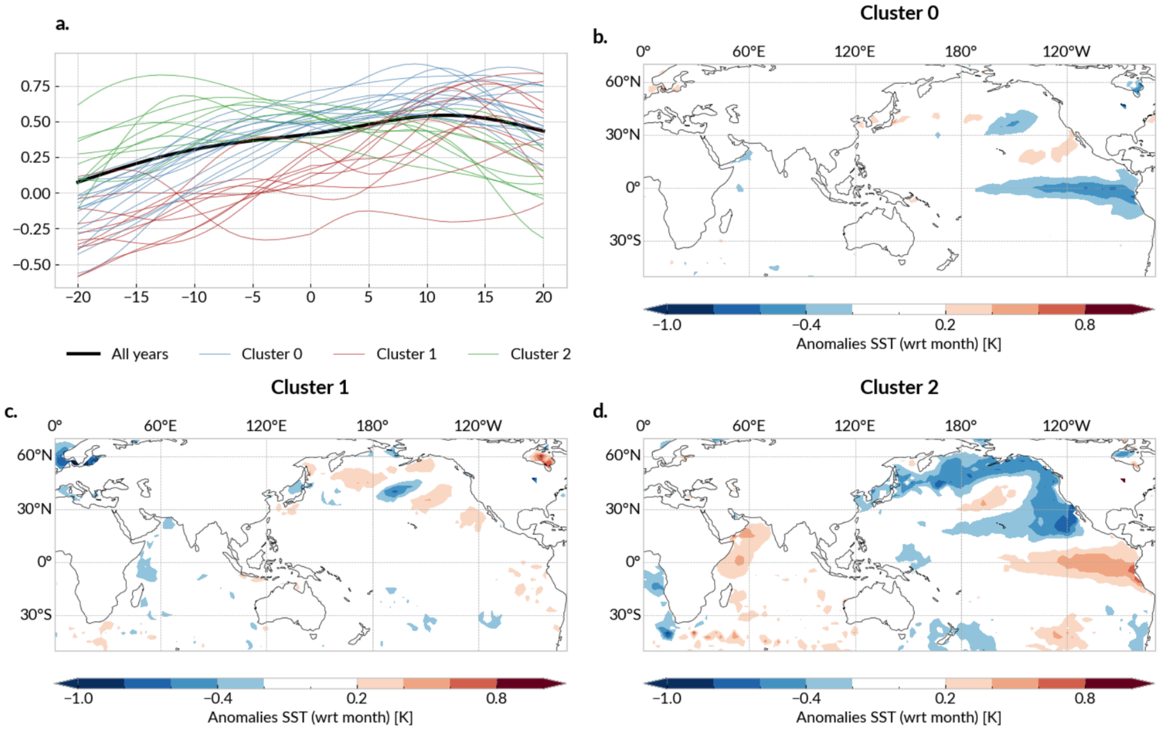


Figure S3: **Yearly fluctuations of the synchronization. a** The yearly lead-lag correlations of extreme rainfall events (EREs) between North India (NI) and the Sahel Zone (SZ) are displayed. The lead-lag correlation is computed using the counts of EREs per year in NI (SZ). The individual yearly lead-lag correlations are clustered using a Gaussian Mixture model. **b,c,d** The corresponding SST anomalies for the JJAS season for the clustered years are shown. Only statistically significant values are shown.

The corresponding SST anomalies for the JJAS season for the clustered years are shown in Fig. S3 b,c,d. We can see that the SST anomalies in the Pacific Ocean are different for the different clusters. This already suggests that the ENSO background state is influencing the synchronization between North India and the Sahel Zone. However, the pattern is not very pronounced which indicates that the synchronization is not solely driven by the ENSO background state.

SI 4 Tropical Easterly Jet Index

This section presents the characteristics of the Tropical Easterly Jet (TEJ) according to the definition by Huang et al. (2019). The TEJ is a prominent feature of the Asian summer monsoon circulation, driven by the meridional thermal contrast between the Asian landmass and the Indian Ocean. It is further intensified by the elevated heating over the Tibetan plateau. The variability of the TEJ substantially influences climate patterns in the tropical and subtropical domain over Asia and the Sahel zone and its surrounding regions.

Pointwise spatial correlation analysis of wind fields We uncover that multiple potential drivers come together when the probability of observing the delayed synchronization between North India and the Sahel Zone is notably increased (Fig.7). We find a substantially higher frequency of synchronizations during an active Boreal Summer Intraseasonal Oscillation (BSISO) in phases 5-8, conditioned on a La Niña background state and an intensified Tropical Easterly

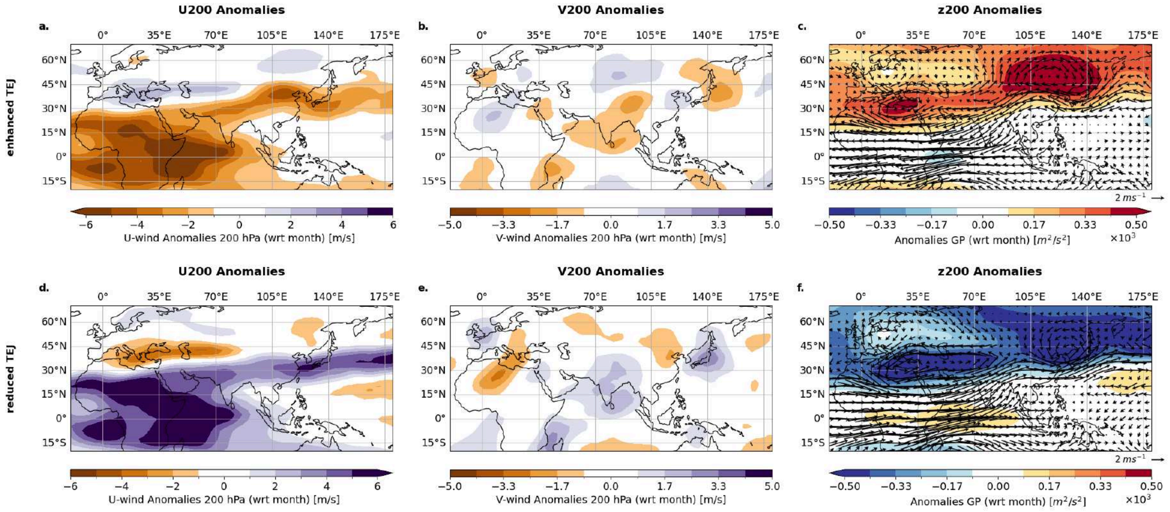


Figure S4: **Illustration of enhanced/reduced TEJ phases.** Composites of different phases of the TEJ are shown. The first row shows the enhanced state, and the second row the reduced state. The first column (**a**, **d**) displays the U200 anomalies (with respect to month), the second column (**b**, **e**) showcases the V200 anomalies, and third row (**c**, **f**) exhibits the Geopotential anomalies at 200 hPa overlapped by the wind field anomalies at 200 hPa, as identified by the TEJI metric described in Sec. 2.1. The magenta rectangle in pannel **a** visualizes the box that is used to define the TEJI.

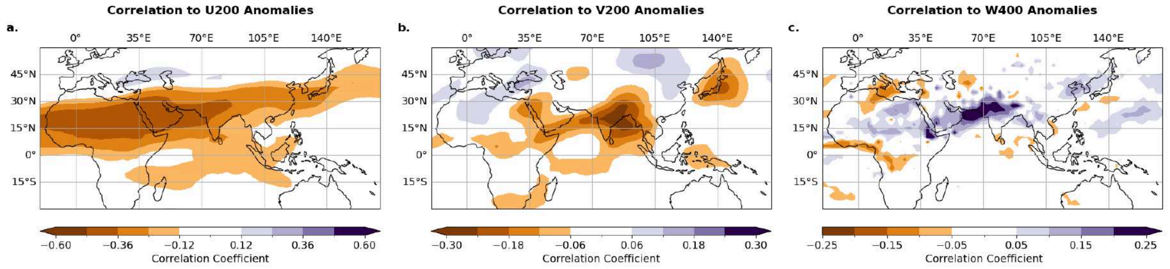


Figure S5: **Correlation of Synchronous Rainfall Index with wind fields. a** The pointwise Spearman correlation of the zonal u-wind anomalies at 200 hPa with the lagged synchronous rainfall index is shown. **b** Same as for **a** but for the meridional v-wind anomalies at 200 hPa. **c** Same as for **a/b** but for the vertical velocity ω at 500 hPa. Colored areas imply statistically significant correlations at the 99 % confidence interval using a two-sided Student's t-test.

Jet (TEJ) in a westerly direction (Fig. 7 c). The pronounced role of the TEJ in establishing the synchronization is underlined by the fact that in its absence there are hardly any synchronizations observed regardless of the respective ENSO state and the BSISO phases (Fig. 7 b,d,f). Similarly, El Niño-like conditions substantially decrease the likelihood of observing synchronizations (Fig. 7 a,b). Further, synchronizations during inactive BSISO are distributed around the null model (grey bars in Fig 7) and are not statistically significant. We estimate the relation between BSISO phases (as defined by Kikuchi et al. (2012)) and concurrent EREs (Fig.S2 a) utilizing a conditional dependence test constructed as the conditional probability of synchronous rainfall events contingent upon (i) phase, (ii) active (inactive) BSISO, (iii) state of the El Niño Southern Oscillation (ENSO), and (iv) state of the TEJ. This statistical analysis suggests the interpretation that La Niña-like conditions establish the foundational state in which the BSISO is predisposed to propagate poleward over Northern India, as reported by (Strnad et al., 2023). This convective cloud over Northern India (NI) is then prone to be advected westwards through the summer circulation over NI in July and August enforced by an enhanced TEJ. These three factors are now analyzed in detail.

SI 5 Circumglobal Teleconnection (CGT)

This section shows the characteristics of the Circumglobal Teleconnection Pattern (CGT) as defined by Ding and Wang (2005). The CGT is a wave pattern that is characterized by a zonal wavenumber of 5. It is a wave pattern that is known to be associated with the Asian Summer Monsoon and the North Atlantic Oscillation (NAO). The CGT is known to be associated with the Silk Road Pattern (SRP) (Enomoto et al., 2003) and the East Asian Wave train pattern (Ding and Wang, 2007). This characteristic wave pattern over Eurasia is shown in Fig. S6 which resembles the wave pattern identified by Ding and Wang (2007).

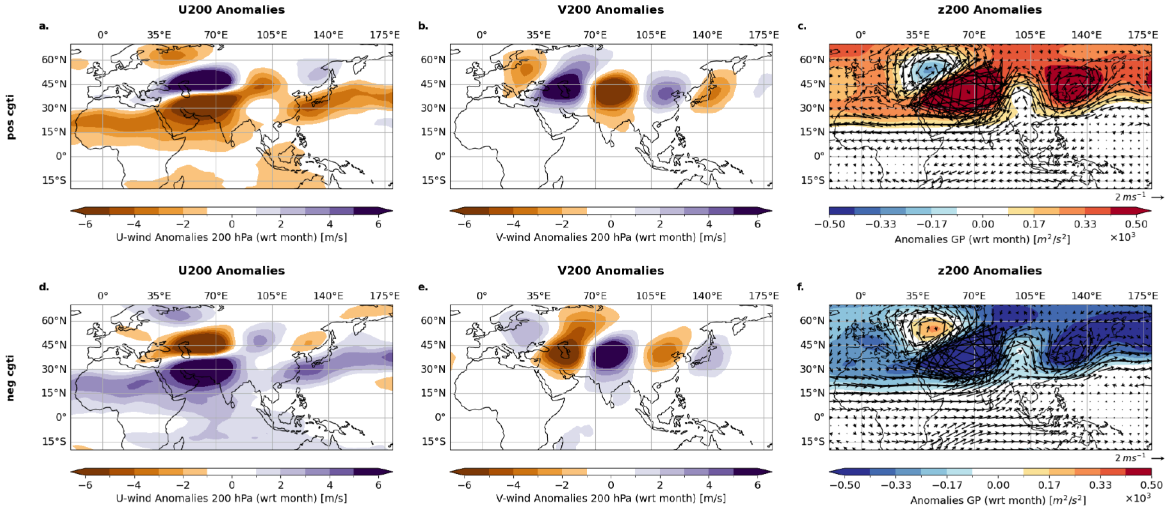


Figure S6: **Illustration of positive/negative CGT phases**. Composites of different phases of the CGT are shown. The first row shows the enhanced state, and the second row the reduced state. The first column (**a**, **d**) displays the U200 anomalies (with respect to month), the second column (**b**, **e**) showcases the V200 anomalies, and third row (**c**, **f**) exhibits the Geopotential anomalies at 200 hPa overlapped by the wind field anomalies at 200 hPa, as identified by the CGTI metric described in Sec. 2.1. The magenta rectangle in pannel **a** visualizes the box that is used to define the CGTI.

SI 6 Connection to Pacific Decadal Oscillation (PDO)

We observe that the Pacific Decadal Oscillation (PDO) has a signature in the SST plots of the Strong Synchronization (Fig. 5 a). The PDO is a long-term climate pattern that is characterized by a positive and negative phase (Fig. S7). The PDO is known to be connected with ENSO and influences the climate in the Indo-Pacific domain and the North American continent. We observe that starting from around 1986, the phases of the PDO roughly align with the phases of the synchronization (Fig. S8)

SI 7 The Somali Jet

The South Asian monsoon onset is marked by a rapid increase in rainfall over southwest India and is preceded by a reversal of winds in the lower troposphere. This occurs analogously to the formation of a low-level zonally oriented jet over the western Indian Ocean. The jet initially flows across the equator and then turns eastward near the East African coast into a zonal flow (Findlater, 1969). This jet is commonly known as the Somali jet. Its low-level southwesterly flow plays a crucial role in transporting moisture for the Indian and Asian Summer Monsoon. It precedes the onset of the monsoon over India and exhibits similar characteristics of rapid intensification and gradual retreat during the monsoon season (Halpern and Woiceshyn, 2001; Masiwal et al., 2023). This flow can be visualized using Integrated Vapor Transport (IVT). We observe that for the MSDs the flow is enhanced over the Arabian Sea and the Bay of Bengal (Fig. S9 a,b) bringing additional moisture towards the center and North of India. In dry years, consistently the flow is reduced (Fig. S9 c).

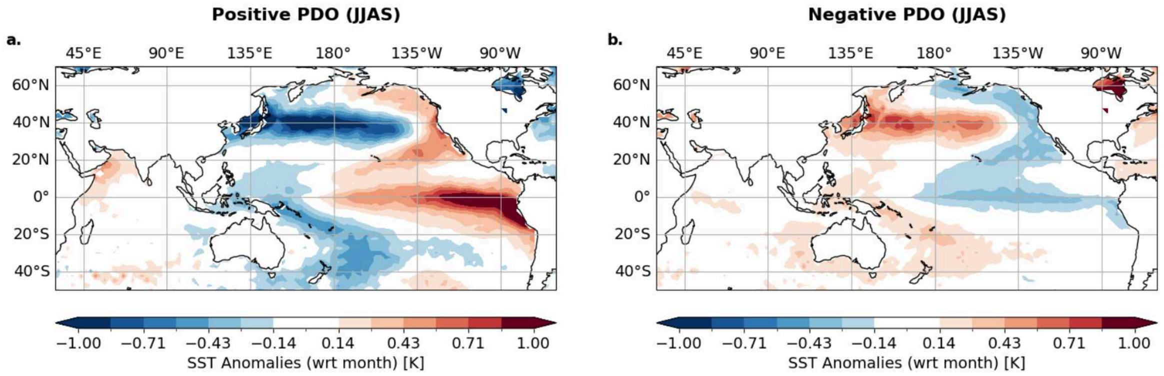


Figure S7: **Illustration of positive/negative PDO phases**. Composites of different phases of the PDO in JJAS are shown. Panel **a** shows the SST anomalies that are associated with the positive PDO phase, while panel **b** shows the SST anomalies that are associated with the negative PDO phase.

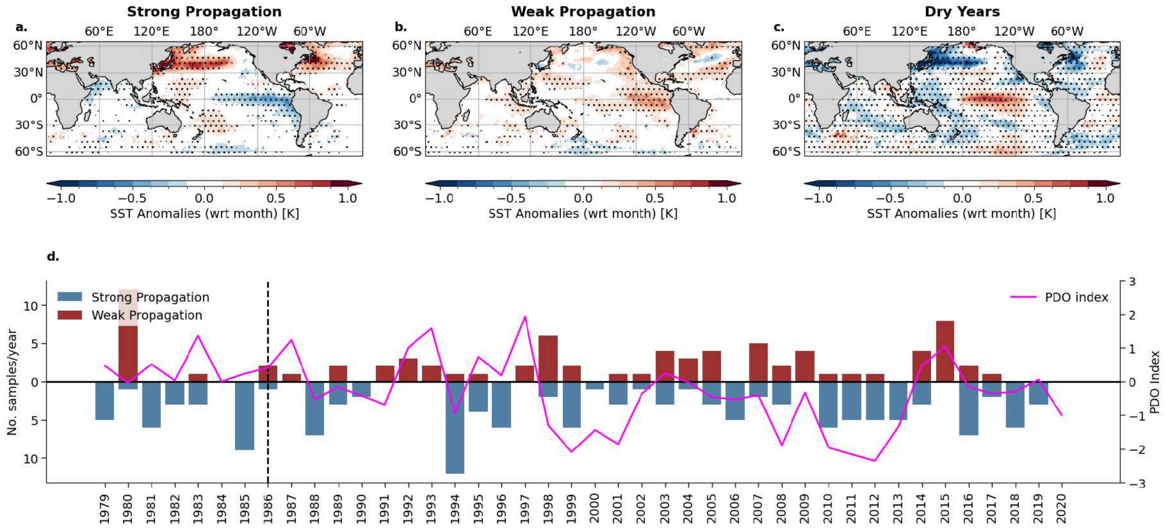


Figure S8: **Background SST states for latent cluster. a–c** shows the sea surface temperature (SST) background state conditions for the Strong Synchronization (**a**) and Weak Synchronization (**b**). For comparison, conditions for years with few/no synchronizations are also shown (**c**). Dotted areas indicate statistical significance at 95 % confidence level using Student's t-test. Panel **d** shows the distribution of days classified as Strong and Weak Synchronization alongside the PDO index (magenta line). The dotted vertical line indicates the year 1986 after which the clustered propagation pathways match the PDO phases quite well.

SI 8 Convection through BSISO

The Boreal Summer Intraseasonal Oscillation (BSISO) is a prominent feature of the Asian Summer Monsoon. It is characterized by a northward propagation of convection over the Indian Ocean and the Maritime Continent. The BSISO periodically occurs on time scales of around 30-60 days and is associated with the active and break events of the Asian Summer Monsoon (Kikuchi, 2021). The BSISO is characterized by eight phases, which are shown in Fig. S11. It has been shown that ENSO modulates substantially the propagation pathway of the BSISO (Strnad et al., 2023). In La Niña years, the BSISO is more likely to propagate over the Indian subcontinent and the North Indian Ocean. Further, in La Niña years the TEJ is enhanced. Therefore, also the convection centers and anomalies change and are more confined to the North and Northwest of India. This is shown in Fig. S13.

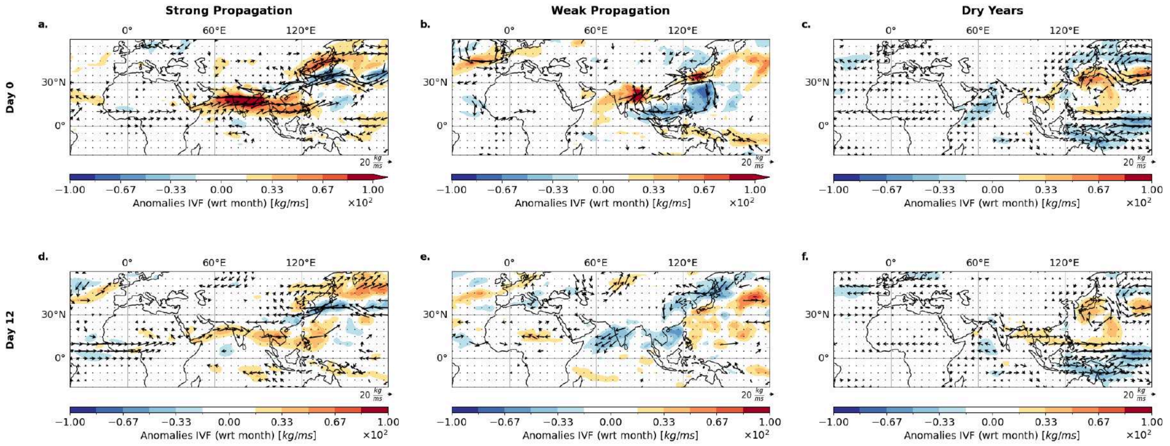


Figure S9: **Vertically integrated water vapor flux (IVT) for days of maximum synchronization for clustered regions.** Vertically integrated water vapor flux (IVF) in eastward and northward directions is visualized for the different clusters (**a,b**) and for the July-August background state for years with only a few synchronizations, called dry years. Only statistically significant anomalies at the 95 % confidence level are shown. Day 0 denotes the days of maximum synchronization. The composite anomalies are computed with respect to the monthly climatology. Only IVT arrows that are significant at 95 % level following the Student's t-test are plotted.

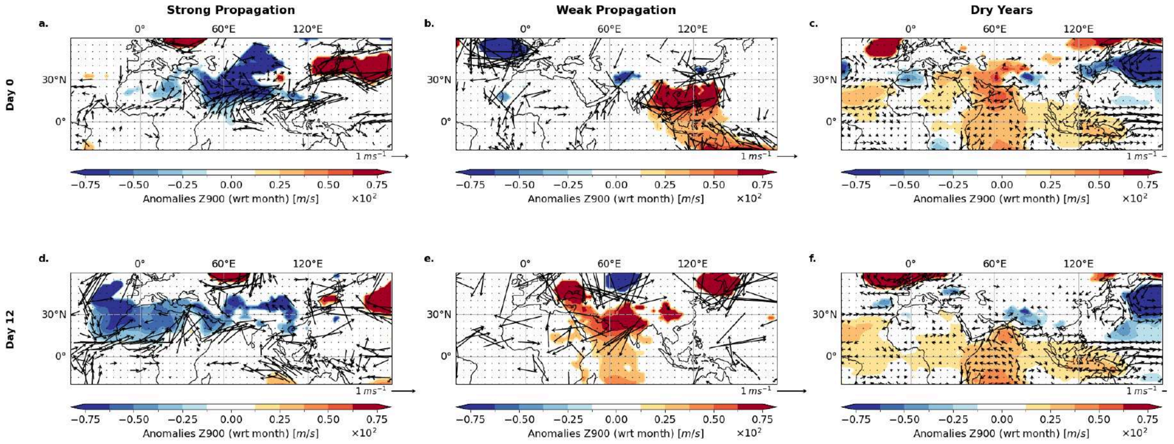


Figure S10: **Geopotential and horizontal wind fields at 900 hPa.** Wind fields and geopotential at 900 hPa are visualized for the different clusters (**a,b**) and for the July-August background state for years with only a few synchronizations, called dry years. Only statistically significant anomalies at the 95 % confidence level are shown. Day 0 denotes the days of maximum synchronization. The composite anomalies are computed with respect to the monthly climatology. Only wind arrows that are significant at 95 % level following the Student's t-test are plotted.

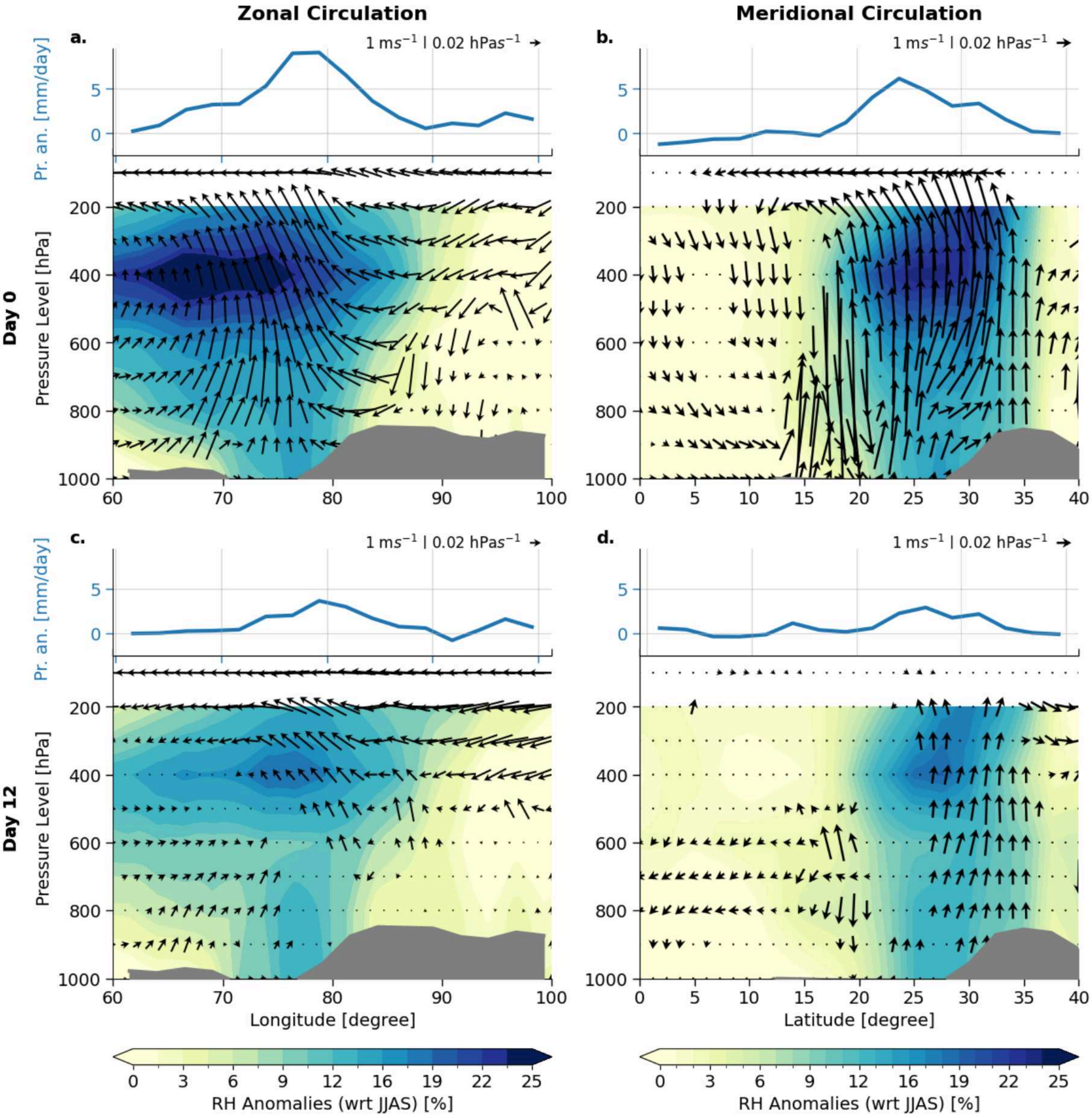


Figure S11: **Lagged precipitation occurrence in North India.** The first column visualizes the zonal circulation averaged between 20° N – 30° N latitude, while the second column the meridional circulation, averaged between 75° E – 85° E longitude. The first row shows the conditions on Day 0 and the second on Day 12, which is the day of the strongest correlation (see Fig. S2 b). Each subplot consists of two panels. The top panel shows the meridionally (zonally) averaged precipitation anomalies (with respect to the month). The down panel shows the vertical circulation. Here, the colored contours show statistically significant composites of V-(U-) wind anomalies at the left (right) column. Grey contours visualize the orography. The wind fields in the zonal (meridional) circulation plots are estimated using the meridionally (zonally) averaged u (v) anomalies, measured in m/s, and the vertical velocity ω in the horizontal direction, measured in hPa/s. Only statistically significant arrows at $95\,\%$ confidence level are shown.

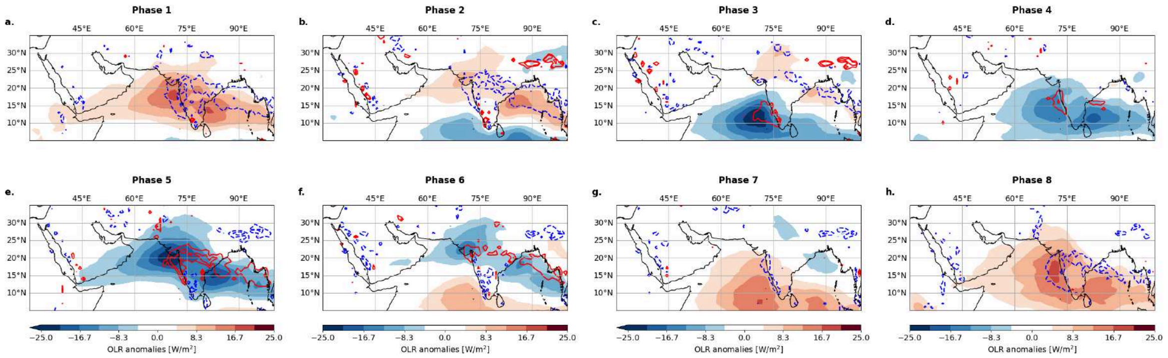


Figure S12: **Convection centers for BSISO phases.** For phases 1-8 of the BSISO the convection is shown in terms of OLR anomalies (color shading) and vertical velocity ω in contours where solid red (dashed blue) contours indicate anomalously upward (downward) motion.

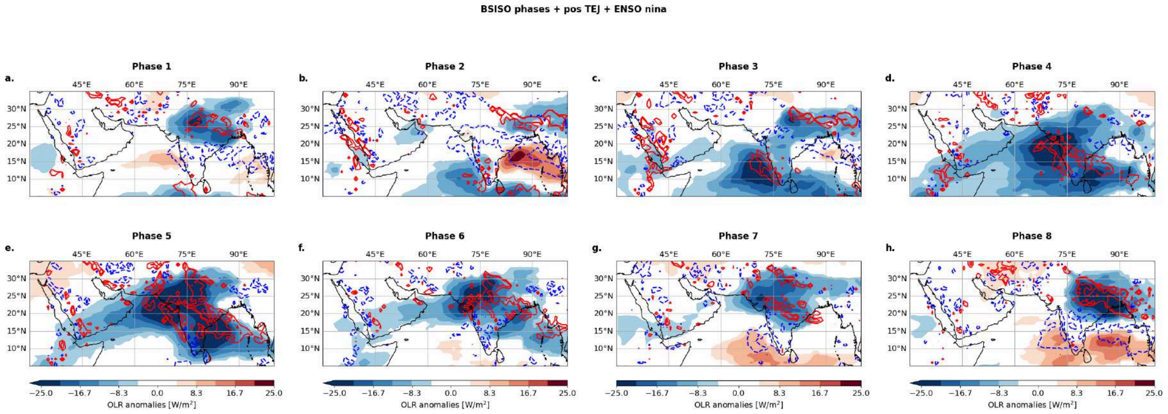


Figure S13: **Convection centers for BSISO phases.** For time points in La Niña years with enhanced TEJ the phases 1-8 of the BSISO are shown in terms of OLR anomalies (color shading) and vertical velocity ω in contours where solid red (dashed blue) contours indicate anomalously upward (downward) motion.

SI 9 Lagrangian Trajectories between North India and Sahel Zone

To visualize possible trajectories we use a Lagrangian trajectory model. The Lagrangian trajectory model is based on the LAGRANTO approach (Sprenger and Wernli, 2015) and is implemented in the geoutils package (Strnad, 2023). The trajectories are calculated for the most synchronous days in the wet years. The time point for the start of the trajectories is set as the days of maximum synchronization occur and we set the starting pressure level at 500 hPa. The resulting trajectories are clustered using a Gaussian Mixture model algorithm. We identify 3 different groups of trajectories. The most likely group is shown in Fig. S14 a. Air is rising at NI and then is deflected South- and Westward, merging with the Tropical Easterly Jet (TEJ) around the Sahel Zone. The second group is shown in Fig. S14 b. Air is rising at North India which then propagates towards SZ but is then transported via the subtropical jet stream further to North China. We conclude that the propagation of extremes seems to follow the summer mean flow in the upper atmosphere at around 200 hPa and the TEJ plays a role in modulating the westward propagation.

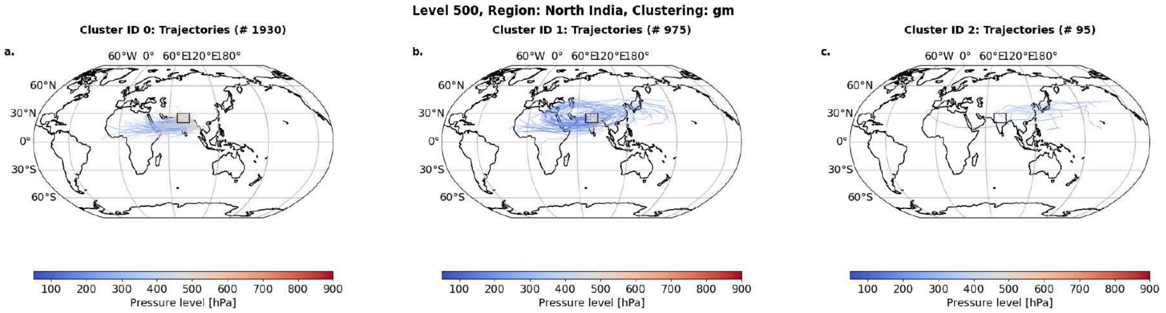


Figure S14: **Lagrangian Trajectory analysis for trajectories starting in North India.** For the most synchronous days in the wet years, we apply a Lagrangian Trajectory analysis in an implementation following the LAGRANTO approach Sprenger and Wernli (2015) for spatial location in the North India box (here marked by the rectangle). For visual reasons in every group, only every 10th trajectory is plotted. The time point for the start of the trajectories is set as the days of maximum synchronization occur and we set the starting pressure level at 500 hPa. The resulting trajectories (in total there are around 100 most synchronous days and 20 starting locations, i.e. $\approx$ 3000 trajectories) are clustered using a Gaussian Mixture model algorithm. We identify 3 different groups of trajectories. **a** shows the most likely group. Air is rising at NI and then is deflected South- and Westward, merging with the Tropical Easterly Jet (TEJ) around the Sahel Zone. **b** shows similarly the rising air at North India which then propagates towards SZ but is then transported via the subtropical jet stream further to North China. We conclude that the propagation of extremes seems to follow the summer mean flow in the upper atmosphere at around 200 hPa and the TEJ plays a role in the westward propagation.

Bibliography

Here are the references in citation order.

F. M. Strnad et al. (2022). 'Teleconnection Patterns of Different El Niño Types Revealed by Climate Network Curvature'. In: *Geophysical Research Letters* 49.17, e2022GL098571. DOI: `10.1029/2022GL098571`.

F. M. Strnad et al. (2023). 'Propagation pathways of Indo-Pacific rainfall extremes are modulated by Pacific sea surface temperatures'. In: *Nature Communications* 14.5708, pages 1–16. DOI: `10.1038/s41467-023-41400-9`.

F. M. Strnad et al. (2024). 'Intraseasonal synchronization of extreme rainfalls between North India and the Sahel'. In: *in review at Quarterly Journal of the Royal Meteorological Society*. DOI: `10.48550/arXiv.2405.08492`. eprint: `2405.08492`.

Hess, P. et al. (2022). 'Physically constrained generative adversarial networks for improving precipitation fields from Earth system models'. In: *Nature Machine Intelligence* 4, pages 828–839. DOI: `10.1038/s42256-022-00540-1`.

Schlör, J. et al. (2024). 'Contribution of El Niño Southern Oscillation (ENSO) Diversity to Low-Frequency Changes in ENSO Variance'. In: *Geophysical Research Letters* 51.14, e2024GL109179. DOI: `10.1029/2024GL109179`.

Lembo, V. et al. (2024). 'Dynamics, statistics and predictability of Rossby waves, heatwaves and spatially compounded extreme events'. In: *Bull. Am. Meteorol. Soc.* -1.aop. DOI: `10.1175/BAMS-D-24-0145.1`.

Trenberth, K. E. (2022). 'Teleconnections and Patterns of Variability'. In: *The Changing Flow of Energy Through the Climate System*. Cambridge, England, UK: Cambridge University Press, pages 162–179. DOI: `10.1017/9781108979030.014`.

Di Capua, G. et al. (2020). 'Dominant patterns of interaction between the tropics and mid-latitudes in boreal summer: causal relationships and the role of timescales'. In: *Weather Clim. Dyn.* 1.2, pages 519–539. DOI: `10.5194/wcd-1-519-2020`.

Trenberth, K. E. (1997). 'The Definition of El Niño'. In: *Bull. Am. Meteorol. Soc.* 78.12, pages 2771–2778. DOI: `10.1175/1520-0477(1997)078<2771:TDOENO>2.0.CO;2`.

Zhang, F. et al. (2019). 'What Is the Predictability Limit of Midlatitude Weather?' In: *J. Atmos. Sci.* 76.4, pages 1077–1091. DOI: `10.1175/JAS-D-18-0269.1`.

Vitart, F. and A. W. Robertson (2018). 'The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events'. In: *npj Clim. Atmos. Sci.* 1.3, pages 1–7. DOI: `10.1038/s41612-018-0013-0`.

Boers, N. et al. (2019). 'Complex networks reveal global pattern of extreme-rainfall teleconnections'. In: *Nature* 566.7744, pages 373–377. DOI: `10.1038/s41586-018-0872-x`.

Kotz, M., A. Levermann, and L. Wenz (2022). 'The effect of rainfall changes on economic production'. In: *Nature* 601, pages 223–227. DOI: `10.1038/s41586-021-04283-8`.

Gadgil, S. and S. Gadgil (2006). 'The Indian monsoon, GDP and agriculture'. In: *Economic & Political Weekly* 41.November 25, pages 4887–4895.

Gallé, J. and A. Katzenberger (2024). 'Indian Agriculture Under Climate Change: The Competing Effect of Temperature and Rainfall Anomalies'. In: *EconDisCliCha.*, pages 1–53. DOI: `10.1007/s41885-024-00154-4`.

Gupta, S. et al. (2022). 'Interconnection between the Indian and the East Asian Summer Monsoon: spatial synchronization patterns of extreme rainfall events'. In: *Int. J. Climatol.* n/a.n/a. DOI: `10.1002/joc.7861`.

Beverley, J. D. et al. (2021). 'Dynamical mechanisms linking Indian monsoon precipitation and the circumglobal teleconnection'. In: *Clim. Dyn.* 57.9, pages 2615–2636. DOI: `10.1007/s00382-021-05825-6`.

Madden, R. A. and P. R. Julian (1971). 'Detection of a 40–50 Day Oscillation in the Zonal Wind in the Tropical Pacific'. In: *J. Atmos. Sci.* 28.5, pages 702–708. DOI: `10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2`.

Hoskins, B. (2013). 'The potential for skill across the range of the seamless weather-climate prediction problem: a stimulus for our science'. In: *Q. J. R. Meteorolog. Soc.* 139.672, pages 573–584. DOI: `10.1002/qj.1991`.

Stott, P. A. et al. (2016). 'Attribution of extreme weather and climate-related events'. In: *WIREs Clim. Change* 7.1, pages 23–41. DOI: `10.1002/wcc.380`.

Gulev, S. et al. (2021). 'Changing State of the Climate System'. In: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 287–422. DOI: 10.1017/9781009157896.004.

Stan, C. et al. (2017). 'Review of Tropical-Extratropical Teleconnections on Intraseasonal Time Scales'. In: *Rev. Geophys.* 55.4, pages 902–937. DOI: 10.1002/2016RG000538.

Amaral, I. (2022). 'Complex Networks'. In: *Encyclopedia of Big Data*. Cham, Switzerland: Springer, pages 198–201. DOI: 10.1007/978-3-319-32010-6_40.

Strogatz, S. H. (2001). 'Exploring complex networks'. In: *Nature* 410, pages 268–276. DOI: 10.1038/35065725.

Watts, D. J. (2004). 'The "New" Science of Networks'. In: *Annual Review of Sociology* Volume 30, 2004, pages 243–270. DOI: 10.1146/annurev.soc.30.020404.104342.

Dijkstra, H. A. et al. (2019). *Networks in Climate*. Cambridge, England, UK: Cambridge University Press.

Boers, N., J. Kurths, and N. Marwan (2021). 'Complex systems approaches for Earth system data analysis'. In: *J. Phys.: Complexity* 2.1, page 011001. DOI: 10.1088/2632-072x/abd8db.

Tsonis, A. A. and K. L. Swanson (2008). 'Topology and Predictability of El Ni\~no and La Ni\~na Networks'. In: *Phys. Rev. Lett.* 100.22, page 228502. DOI: 10.1103/PhysRevLett.100.228502.

Donges, J. F. et al. (2009a). 'Complex networks in climate dynamics'. In: *Eur. Phys. J. Spec. Top.* 174.1, pages 157–179. DOI: 10.1140/epjst/e2009-01098-2.

Malik, N., N. Marwan, and J. Kurths (2010). 'Spatial structures and directionalities in Monsoonal precipitation over South Asia'. In: *Nonlinear Processes Geophys.* 17.5, pages 371–381. DOI: 10.5194/npg-17-371-2010.

Ludescher, J. et al. (2014). 'Very early warning of next El Niño'. In: *Proc. Natl. Acad. Sci. U.S.A.* 111.6, pages 2064–2066. DOI: 10.1073/pnas.1323058111.

Boers, N. et al. (2014). 'Prediction of extreme floods in the eastern Central Andes based on a complex networks approach'. In: *Nat. Commun.* 5.5199, pages 1–7. DOI: 10.1038/ncomms6199.

Stolbova, V. et al. (2016). 'Tipping elements of the Indian monsoon: Prediction of onset and withdrawal'. In: *Geophys. Res. Lett.* 43.8, pages 3982–3990. DOI: 10.1002/2016GL068392.

Gupta, S. et al. (2021). 'Complex network approach for detecting tropical cyclones'. In: *Clim. Dyn.* 57.11, pages 3355–3364. DOI: 10.1007/s00382-021-05871-0.

Liu, T. et al. (2023). 'Teleconnections among tipping elements in the Earth system'. In: *Nat. Clim. Change* 13, pages 67–74. DOI: 10.1038/s41558-022-01558-4.

Kumar, K. K. et al. (2006). 'Unraveling the Mystery of Indian Monsoon Failure During El Niño'. In: *Science* 314.5796, pages 115–119. DOI: 10.1126/science.1131152.

Strangeways, I. (2006). *Precipitation: Theory, Measurement and Distribution*. Cambridge, England, UK: Cambridge University Press.

Trenberth, K. E. et al. (2003). 'The Changing Character of Precipitation'. In: *Bull. Am. Meteorol. Soc.* 84.9, pages 1205–1218. DOI: 10.1175/BAMS-84-9-1205.

Houze, R. A. (2014). *Cloud dynamics*. Academic press.

– (1997). 'Stratiform Precipitation in Regions of Convection: A Meteorological Paradox?' In: *Bull. Am. Meteorol. Soc.* 78.10, pages 2179–2196. DOI: 10.1175/1520-0477(1997)078<2179:SPIROC>2.0.CO;2.

Achatz, U. (2022). *Atmospheric Dynamics*. Berlin, Germany: Springer.

Serinaldi, F. and C. G. Kilsby (2014). 'Rainfall extremes: Toward reconciliation after the battle of distributions'. In: *Water Resour. Res.* 50.1, pages 336–352. DOI: 10.1002/2013WR014211.

Lehmann, J., D. Coumou, and K. Frieler (2015). 'Increased record-breaking precipitation events under global warming'. In: *Clim. Change* 132.4, pages 501–515. DOI: 10.1007/s10584-015-1434-y.

Katzenberger, A. et al. (2022). 'Intensification of Very Wet Monsoon Seasons in India Under Global Warming'. In: *Geophys. Res. Lett.* 49.15, e2022GL098856. DOI: 10.1029/2022GL098856.

Katzenberger, A. et al. (2021). 'Robust increase of Indian monsoon rainfall and its variability under future warming in CMIP6 models'. In: *Earth Syst. Dyn.* 12.2, pages 367–386. DOI: 10.5194/esd-12-367-2021.

Wu, Z. et al. (2007). 'On the trend, detrending, and variability of nonlinear and nonstationary time series'. In: *Proceedings of the National Academy of Sciences* 104.38, pages 14889–14894.

Dai, A. and C. E. Bloecker (2019). 'Impacts of internal variability on temperature and precipitation trends in large ensemble simulations by two climate models'. In: *Clim. Dyn.* 52.1, pages 289–306. DOI: 10.1007/s00382-018-4132-4.

Gao, S. (2020). ‘Dynamical downscaling of surface air temperature and precipitation using RegCM4 and WRF over China’. In: *Clim. Dyn.* 55.5, pages 1283–1302. DOI: 10.1007/s00382-020-05326-y.

Cheng, Y. et al. (2024). ‘Assessment of dynamical downscaling performance over cordex east Asia using MPAS-A global variable resolution model: climatology, seasonal cycle, and extreme events’. In: *Clim. Dyn.*, pages 1–23. DOI: 10.1007/s00382-024-07265-4.

Kendon, E. J. et al. (2014). ‘Heavier summer downpours with climate change revealed by weather forecast resolution model’. In: *Nat. Clim. Change* 4, pages 570–576. DOI: 10.1038/nclimate2258.

O’Gorman, P. A. (2015). ‘Precipitation Extremes Under Climate Change’. In: *Curr. Clim. Change Rep.* 1.2, pages 49–59. DOI: 10.1007/s40641-015-0009-3.

Schwendike, J. et al. (2014). ‘Local partitioning of the overturning circulation in the tropics and the connection to the Hadley and Walker circulations’. In: *J. Geophys. Res. Atmos.* 119.3, pages 1322–1339. DOI: 10.1002/2013JD020742.

Schwendike, J. et al. (2021). ‘On the Relationship Between the Madden-Julian Oscillation and the Hadley and Walker Circulations’. In: *J. Geophys. Res. Atmos.* 126.4, e2019JD032117. DOI: 10.1029/2019JD032117.

Sobel, A. H. et al. (2008). ‘The role of surface heat fluxes in tropical intraseasonal oscillations’. In: *Nat. Geosci.* 1, pages 653–657. DOI: 10.1038/ngeo312.

Sobel, A. and E. Maloney (2012). ‘An Idealized Semi-Empirical Framework for Modeling the Madden–Julian Oscillation’. In: *J. Atmos. Sci.* 69.5, pages 1691–1705. DOI: 10.1175/JAS-D-11-0118.1.

Wang, S. and A. H. Sobel (2022). ‘A Unified Moisture Mode Theory for the Madden–Julian Oscillation and the Boreal Summer Intraseasonal Oscillation’. In: *J. Clim.* 35.4, pages 1267–1291. DOI: 10.1175/JCLI-D-21-0361.1.

Kim, D., J.-S. Kug, and A. H. Sobel (2014). ‘Propagating versus Nonpropagating Madden–Julian Oscillation Events’. In: *J. Clim.* 27.1, pages 111–125. DOI: 10.1175/JCLI-D-13-00084.1.

Wallace, J. M. and P. V. Hobbs (2006). *Atmospheric Science*. Elsevier, Academic Press.

Hoskins, B. J., A. J. Simmons, and D. G. Andrews (1977). ‘Energy dispersion in a barotropic atmosphere’. In: *Q. J. R. Meteorolog. Soc.* 103.438, pages 553–567. DOI: 10.1002/qj.49710343802.

Hoskins, B. J. and D. J. Karoly (1981). ‘The Steady Linear Response of a Spherical Atmosphere to Thermal and Orographic Forcing’. In: *J. Atmos. Sci.* 38.6, pages 1179–1196. DOI: 10.1175/1520-0469(1981)038<1179:TSLROA>2.0.CO;2.

Vallis, G. K. (2006). *Atmospheric and Oceanic Fluid Dynamics*. Cambridge, U.K.: Cambridge University Press, page 745.

Wolter, K. and M. S. Timlin (2011). ‘El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI.ext)’. In: *Int. J. Climatol.* 31.7, pages 1074–1087. DOI: 10.1002/joc.2336.

Bjerknes, J. (1969). ‘ATMOSPHERIC TELECONNECTIONS FROM THE EQUATORIAL PACIFIC’. In: *Mon. Weather Rev.* 97.3, pages 163–172. DOI: 10.1175/1520-0493(1969)097<0163:ATFTEP>2.3.CO;2.

Capotondi, A. et al. (2020). ‘ENSO Diversity’. In: *El Niño Southern Oscillation in a Changing Climate*. American Geophysical Union (AGU), pages 65–86. DOI: 10.1002/9781119548164.ch4.

Kripalani, R. and A. Kulkarni (1997). ‘Climatic impact of El Nino/La Nina on the Indian monsoon: A new perspective’. In: *Weather* 52.2, pages 39–46.

Hand, E. (2015). ‘The storm king’. In: *Science* 350.6256, pages 22–25. DOI: 10.1126/science.350.6256.22.

Zhang, C. (2005). ‘Madden-Julian Oscillation’. In: *Rev. Geophys.* 43.2. DOI: 10.1029/2004RG000158.

Zhang, C. et al. (2020). ‘Four Theories of the Madden-Julian Oscillation’. In: *Rev. Geophys.* 58.3, e2019RG000685. DOI: 10.1029/2019RG000685.

Rostami, M., B. Zhao, and S. Petri (2022). ‘On the genesis and dynamics of Madden–Julian oscillation-like structure formed by equatorial adjustment of localized heating’. In: *Q. J. R. Meteorolog. Soc.* 148.749, pages 3788–3813. DOI: 10.1002/qj.4388.

Wheeler, M. C. and H. H. Hendon (2004). ‘An All-Season Real-Time Multivariate MJO Index: Development of an Index for Monitoring and Prediction’. In: *Mon. Weather Rev.* 132.8, pages 1917–1932. DOI: 10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.

Wang, B. and X. Xie (1997). ‘A Model for the Boreal Summer Intraseasonal Oscillation’. In: *J. Atmos. Sci.* 54.1, pages 72–86. DOI: 10.1175/1520-0469(1997)054<0072:AMFTBS>2.0.CO;2.

Kikuchi, K. (2021). 'The Boreal Summer Intraseasonal Oscillation (BSISO): A Review'. In: *Journal of the Meteorological Society of Japan. Ser. II*, pages 2021–045. DOI: `10.2151/jmsj.2021-045`.

Kikuchi, K., B. Wang, and Y. Kajikawa (2012). 'Bimodal representation of the tropical intraseasonal oscillation'. In: *Clim. Dyn.* 38.9, pages 1989–2000. DOI: `10.1007/s00382-011-1159-1`.

Wang, B. and Q. Ding (2008). 'Global monsoon: Dominant mode of annual variation in the tropics'. In: *Dyn. Atmos. Oceans* 44.3, pages 165–183. DOI: `10.1016/j.dynatmoce.2007.05.002`.

Ramage, C. (1971). *Monsoon Meteorology*. Experimental Botany V. 10. Academic Press.

Trenberth, K. E., D. P. Stepaniak, and J. M. Caron (2000). 'The Global Monsoon as Seen through the Divergent Atmospheric Circulation'. In: *J. Clim.* 13.22, pages 3969–3993. DOI: `10.1175/1520-0442(2000)013<3969:TGMAST>2.0.CO;2`.

Bordoni, S. and T. Schneider (2008). 'Monsoons as eddy-mediated regime transitions of the tropical overturning circulation'. In: *Nat. Geosci.* 1, pages 515–519. DOI: `10.1038/ngeo248`.

Geen, R. et al. (2020). 'Monsoons, ITCZs, and the Concept of the Global Monsoon'. In: *Rev. Geophys.* 58.4, e2020RG000700. DOI: `10.1029/2020RG000700`.

Hoskins, B. J., G.-Y. Yang, and R. M. Fonseca (2020). 'The detailed dynamics of the June–August Hadley Cell'. In: *Q. J. R. Meteorolog. Soc.* 146.727, pages 557–575. DOI: `10.1002/qj.3702`.

Walker, J. M., S. Bordoni, and T. Schneider (2015). 'Interannual Variability in the Large-Scale Dynamics of the South Asian Summer Monsoon'. In: *J. Clim.* 28.9, pages 3731–3750. DOI: `10.1175/JCLI-D-14-00612.1`.

Schneider, T., T. Bischoff, and G. H. Haug (2014). 'Migrations and dynamics of the intertropical convergence zone'. In: *Nature* 513, pages 45–53. DOI: `10.1038/nature13636`.

Meijer, N. et al. (2024). 'Proto-monsoon rainfall and greening in Central Asia due to extreme early Eocene warmth'. In: *Nat. Geosci.* 17, pages 158–164. DOI: `10.1038/s41561-023-01371-4`.

Jeevanjee, N. et al. (2017). 'A perspective on climate model hierarchies'. In: *J. Adv. Model. Earth Syst.* 9.4, pages 1760–1771. DOI: `10.1002/2017MS001038`.

Eyring, V. et al. (2016). 'Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization'. In: *Geosci. Model Dev.* 9.5, pages 1937–1958. DOI: `10.5194/gmd-9-1937-2016`.

Beck, H. E. et al. (2019b). 'MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment'. In: *Bull. Am. Meteorol. Soc.* 100.3, pages 473–500. DOI: `10.1175/BAMS-D-17-0138.1`.

Beck, H. E. et al. (2019a). 'Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS'. In: *Hydrol. Earth Syst. Sci.* 23.1, pages 207–224. DOI: `10.5194/hess-23-207-2019`.

Hovmöller, E. (1949). 'The Trough-and-Ridge diagram'. In: *Tellus* 1.2, pages 62–66. DOI: `10.1111/j.2153-3490.1949.tb01260.x`.

Lorenz, E. N. (1956). *Empirical orthogonal functions and statistical weather prediction*. Volume 1. Massachusetts Institute of Technology, Department of Meteorology Cambridge.

Ding, Q. and B. Wang (2005). 'Circumglobal Teleconnection in the Northern Hemisphere Summer'. In: *J. Clim.* 18.17, pages 3483–3505. DOI: `10.1175/JCLI3473.1`.

Beverley, J. D. et al. (2019). 'The northern hemisphere circumglobal teleconnection in a seasonal forecast model and its relationship to European summer forecast skill'. In: *Clim. Dyn.* 52.5, pages 3759–3771. DOI: `10.1007/s00382-018-4371-4`.

Quian Quiroga, R., T. Kreuz, and P. Grassberger (2002). 'Event synchronization: A simple and fast method to measure synchronicity and time delay patterns'. In: *Phys. Rev. E* 66.4, page 041904. DOI: `10.1103/PhysRevE.66.041904`.

Rheinwalt, A. et al. (2015). 'Teleconnections in Climate Networks: A Network-of-Networks Approach to Investigate the Influence of Sea Surface Temperature Variability on Monsoon Systems'. In: *Machine Learning and Data Mining Approaches to Climate Science*. Cham, Switzerland: Springer, pages 23–33. DOI: `10.1007/978-3-319-17220-0_3`.

Donges, J. F. et al. (2009b). 'The backbone of the climate network'. In: *EPL (Europhysics Letters)* 87.4, page 48007. DOI: `10.1209/0295-5075/87/48007`.

Haas, M., B. Goswami, and U. von Luxburg (2023). 'Pitfalls of Climate Network Construction—A Statistical Perspective'. In: *J. Clim.* 36.10, pages 3321–3342. DOI: `10.1175/JCLI-D-22-0549.1`.

Benjamini, Y. and Y. Hochberg (1995). 'Controlling the false discovery rate: a practical and powerful approach to multiple testing'. In: *J. Roy. Statist. Soc. Ser. B* 57.1, pages 289–300.

Holm, S. (1979). *A Simple Sequentially Rejective Multiple Test Procedure on JSTOR*. [Online; accessed 17. May 2024].

Hochberg, Y. and Y. Benjamini (1990). 'More powerful procedures for multiple significance testing'. In: *Stat. Med.* 9.7, pages 811–818. DOI: `10.1002/sim.4780090710`.

Wolf, F. et al. (2021). 'Spatiotemporal patterns of synchronous heavy rainfall events in East Asia during the Baiu season'. In: *Earth Syst. Dyn.* 12.1, pages 295–312. DOI: `10.5194/esd-12-295-2021`.

Boers, N. et al. (2013). 'Complex networks identify spatial patterns of extreme rainfall events of the South American Monsoon System'. In: *Geophys. Res. Lett.* 40.16, pages 4386–4392. DOI: `10.1002/grl.50681`.

Donges, J. F. et al. (2015). 'How complex climate networks complement eigen techniques for the statistical analysis of climatological data'. In: *Clim. Dyn.* 45.9, pages 2407–2424. DOI: `10.1007/s00382-015-2479-3`.

Stolbova, V. et al. (2014). 'Topology and seasonal evolution of the network of extreme precipitation over the Indian subcontinent and Sri Lanka'. In: *Nonlinear Processes Geophys.* 21.4, pages 901–917. DOI: `10.5194/npg-21-901-2014`.

Peixoto, T. P. (2023). *Descriptive vs. Inferential Community Detection in Networks: Pitfalls, Myths and Half-Truths*. Cambridge, England, UK: Cambridge University Press.

Fortunato, S. and D. Hric (2016). 'Community detection in networks: A user guide'. In: *Phys. Rep.* 659, pages 1–44. DOI: `10.1016/j.physrep.2016.09.002`.

Peixoto, T. P. (2019). 'Bayesian Stochastic Blockmodeling'. In: *Advances in Network Clustering and Blockmodeling*. Chichester, England, UK: John Wiley & Sons, Ltd, pages 289–332. DOI: `10.1002/9781119483298.ch11`.

Kretschmer, M. et al. (2021). 'Quantifying Causal Pathways of Teleconnections'. In: *Bull. Am. Meteorol. Soc.* 102.12, E2247–E2263. DOI: `10.1175/BAMS-D-20-0117.1`.

Hannart, A. et al. (2016). 'Causal Counterfactual Theory for the Attribution of Weather and Climate-Related Events'. In: *Bull. Am. Meteorol. Soc.* 97.1, pages 99–110. DOI: `10.1175/BAMS-D-14-00034.1`.

Mengel, M. et al. (2021). 'ATTRICI v1.1 – counterfactual climate for impact attribution'. In: *Geosci. Model Dev.* 14.8, pages 5269–5284. DOI: `10.5194/gmd-14-5269-2021`.

Schölkopf, B. (2022). 'Causality for Machine Learning'. In: *Probabilistic and Causal Inference: The Works of Judea Pearl*. Volume 36. New York, NY, USA: Association for Computing Machinery, pages 765–804. DOI: `10.1145/3501714.3501755`.

Kaddour, J. et al. (2022). 'Causal Machine Learning: A Survey and Open Problems'. In: *arXiv*. DOI: `10.48550/arXiv.2206.15475`. eprint: `2206.15475`.

Runge, J. et al. (2019). 'Detecting and quantifying causal associations in large nonlinear time series datasets'. In: *Sci. Adv.* 5.11. DOI: `10.1126/sciadv.aau4996`.

Leamer, E. E. (1985). 'Vector autoregressions for causal inference?' In: *Carnegie-Rochester Conference Series on Public Policy* 22, pages 255–304. DOI: `10.1016/0167-2231(85)90035-1`.

Wiedermann, M. et al. (2016). 'A climate network-based index to discriminate different types of El Niño and La Niña'. In: *Geophys. Res. Lett.* 43.13, pages 7176–7185. DOI: `10.1002/2016GL069119`.

Fan, J. et al. (2017). 'Network analysis reveals strongly localized impacts of El Niño'. In: *Proceedings of the National Academy of Sciences* 114.29, pages 7543–7548.

Bendito, E. et al. (2007). 'Estimation of Fekete points'. In: *J. Comput. Phys.* 225.2, pages 2354–2376. DOI: `10.1016/j.jcp.2007.03.017`.

Lin, H. (2019). 'Long-lead ENSO control of the boreal summer intraseasonal oscillation in the East Asian-western North Pacific region'. In: *npj Clim. Atmos. Sci.* 2.31, pages 1–6. DOI: `10.1038/s41612-019-0088-2`.

Ahn, M.-S. et al. (2020). 'Role of Maritime Continent Land Convection on the Mean State and MJO Propagation'. In: *J. Clim.* 33.5, pages 1659–1675. DOI: `10.1175/JCLI-D-19-0342.1`.

Liu, F. et al. (2016). 'Modulation of Boreal Summer Intraseasonal Oscillations over the Western North Pacific by ENSO'. In: *J. Clim.* 29.20, pages 7189–7201. DOI: `10.1175/JCLI-D-15-0831.1`.

Lee, D. Y. et al. (2022). 'Dominant Processes for Dependence of Boreal Summer Intraseasonal Oscillation on El Niño Phases'. In: *Geophys. Res. Lett.* 49.14, e2022GL098669. DOI: `10.1029/2022GL098669`.

Cardil, A. et al. (2023). 'Climate teleconnections modulate global burned area'. In: *Nat. Commun.* 14.427, pages 1–10. DOI: `10.1038/s41467-023-36052-8`.

Dalelane, C., K. Winderlich, and A. Walter (2023). 'Evaluation of global teleconnections in CMIP6 climate projections using complex networks'. In: *Earth Syst. Dyn.* 14.1, pages 17–37. DOI: `10.5194/esd-14-17-2023`.

Chen, G. (2021). ‘Diversity of the Global Teleconnections Associated with the Madden–Julian Oscillation’. In: *J. Clim.* 34.1, pages 397–414. DOI: 10.1175/JCLI-D-20-0357.1.

Steptoe, H., S. E. O. Jones, and H. Fox (2018). ‘Correlations Between Extreme Atmospheric Hazards and Global Teleconnections: Implications for Multihazard Resilience’. In: *Rev. Geophys.* 56.1, pages 50–78. DOI: 10.1002/2017RG000567.

Trenberth, K. E. et al. (1998). ‘Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures’. In: *J. Geophys. Res. Oceans* 103.C7, pages 14291–14324. DOI: 10.1029/97JC01444.

Vitart, F. and A. W. Robertson (2019). ‘Introduction: Why Sub-seasonal to Seasonal Prediction (S2S)?’ In: *Sub-Seasonal to Seasonal Prediction*. Walthm, MA, USA: Elsevier, pages 3–15. DOI: 10.1016/B978-0-12-811714-9.00001-2.

Mariotti, A. et al. (2020). ‘Windows of Opportunity for Skillful Forecasts Subseasonal to Seasonal and Beyond’. In: *Bull. Am. Meteorol. Soc.* 101.5, E608–E625. DOI: 10.1175/BAMS-D-18-0326.1.

Woolnough, S. J. (2019). ‘The Madden-Julian Oscillation’. In: *Sub-Seasonal to Seasonal Prediction*. Walthm, MA, USA: Elsevier, pages 93–117. DOI: 10.1016/B978-0-12-811714-9.00005-X.

Lam, R. et al. (2023). ‘Learning skillful medium-range global weather forecasting’. In: *Science* 382.6677, pages 1416–1421. DOI: 10.1126/science.adi2336.

Bi, K. et al. (2023). ‘Accurate medium-range global weather forecasting with 3D neural networks’. In: *Nature* 619, pages 533–538. DOI: 10.1038/s41586-023-06185-3.

Lessig, C. et al. (2023). ‘AtmoRep: A stochastic model of atmosphere dynamics using large scale representation learning’. In: *arXiv*. DOI: 10.48550/arXiv.2308.13280. eprint: 2308.13280.

Rezaei, A. et al. (2023). ‘Changes in global teleconnection patterns under global warming and stratospheric aerosol intervention scenarios’. In: *Atmos. Chem. Phys.* 23.10, pages 5835–5850. DOI: 10.5194/acp-23-5835-2023.

Geng, X., J.-S. Kug, and Y. Kosaka (2024). ‘Future changes in the wintertime ENSO-NAO teleconnection under greenhouse warming’. In: *npj Clim. Atmos. Sci.* 7.81, pages 1–12. DOI: 10.1038/s41612-024-00627-z.

Ummenhofer, C. C. and G. A. Meehl (2017). ‘Extreme weather and climate events with ecological relevance: a review’. In: *Phil. Trans. R. Soc. B* 372.1723. DOI: 10.1098/rstb.2016.0135.

Mo, C. et al. (2019). ‘Frequency analysis of precipitation extremes under climate change’. In: *Int. J. Climatol.* 39.3, pages 1373–1387. DOI: 10.1002/joc.5887.

Rodell, M. and B. Li (2023). ‘Changing intensity of hydroclimatic extreme events revealed by GRACE and GRACE-FO’. In: *Nat. Water* 1, pages 241–248. DOI: 10.1038/s44221-023-00040-5.

Haszpra, T., M. Herein, and T. Bódai (2020). ‘Investigating ENSO and its teleconnections under climate change in an ensemble view – a new perspective’. In: *Earth Syst. Dyn.* 11.1, pages 267–280. DOI: 10.5194/esd-11-267-2020.

Hess, P. et al. (2023). ‘Deep Learning for Bias-Correcting CMIP6-Class Earth System Models’. In: *Earth's Future* 11.10, e2023EF004002. DOI: 10.1029/2023EF004002.

Hess, P. et al. (2024). ‘Fast, Scale-Adaptive, and Uncertainty-Aware Downscaling of Earth System Model Fields with Generative Foundation Models’. In: *arXiv*. DOI: 10.48550/arXiv.2403.02774. eprint: 2403.02774.

Cai, W. et al. (2015). ‘ENSO and greenhouse warming’. In: *Nat. Clim. Change* 5, pages 849–859. DOI: 10.1038/nclimate2743.

Alizadeh, O. (2024). ‘A review of ENSO teleconnections at present and under future global warming’. In: *WIREs Clim. Change* 15.1, e861. DOI: 10.1002/wcc.861.

Turner, A. G. and H. Annamalai (2012). ‘Climate change and the South Asian summer monsoon’. In: *Nat. Clim. Change* 2, pages 587–595. DOI: 10.1038/nclimate1495.

Papalexiou, S. M. and A. Montanari (2019). ‘Global and Regional Increase of Precipitation Extremes Under Global Warming’. In: *Water Resour. Res.* 55.6, pages 4901–4914. DOI: 10.1029/2018WR024067.

Maloney, E. D., Á. F. Adames, and H. X. Bui (2019). ‘Madden–Julian oscillation changes under anthropogenic warming’. In: *Nat. Clim. Change* 9, pages 26–33. DOI: 10.1038/s41558-018-0331-6.

Cui, J. and T. Li (2019). ‘Changes of MJO propagation characteristics under global warming’. In: *Clim. Dyn.* 53.9, pages 5311–5327. DOI: 10.1007/s00382-019-04864-4.

Du, D. et al. (2024). ‘Increase in MJO predictability under global warming’. In: *Nat. Clim. Change* 14, pages 68–74. DOI: 10.1038/s41558-023-01885-0.

Huang, S. et al. (2019). 'Interdecadal change in the relationship between the tropical easterly jet and tropical sea surface temperature anomalies in boreal summer'. In: *Clim. Dyn.* 53.3, pages 2119–2131. DOI: `10.1007/s00382-019-04801-5`.

Sharma, V. et al. (2024). 'Evidence of strengthening of tropical easterly jet after 1998 climate shift using ERA-5 datasets'. In: *Theor. Appl. Climatol.*, pages 1–16. DOI: `10.1007/s00704-024-04975-x`.